PM3386

S/UNI-2XGE

ETHERNET OVER SONET USING THE S/UNI-2XGE

APPLICATION NOTE

PRELIMINARY

ISSUE 1

REVISION HISTORY

Issue No.	Issue Date	Details of Change
1	Sept 18, 2000	Document Created

CONTENTS

LIST OF FIGURES

1 REFERENCES

1. PMC-Sierra Inc., PM3386 *Dual Gigabit Ethernet Controller Standard Product Data Sheet.* Issue 4, June 2000.

2. PMC-Sierra Inc., PM5381 *Saturn User Network Interface (S/UNI-2488) Telecom Standard Product Data Sheet.* Issue 1, February 2000.

3. PMC-Sierra, Inc., *Saturn Compatible Packet over SONET Interface Specification – POS-PHY Level 3.* Issue 4, June 2000.

4. IETF Network Working Group – RFC-1619, “Point to Point Protocol (PPP) over SONET/SDH Specification”, May 1994.

5. IETF Network Working Group – RFC-1661, “The Point to Point Protocol (PPP)”, July, 1994.

6. IETF Network Working Group – RFC-1662, “PPP in HDLC like framing”, July 1994.

7. IETF Network Working Group – RFC-2615, “Point to Point Protocol (PPP) over SONET/SDH”, June 1999.

8. Doshi, B.T. et al. *A Simple Data Link Protocol for High-Speed Packet Networks.* Bell Labs Technical Journal, Jan-March, 1999. pp.85-104.

2 OVERVIEW

As data rates expand into the 10 Gigabit range, Ethernet is beginning to gain popularity as a key technology for Wide Area Network providers. Scalable from 1Mbps to 10Gbps, Ethernet offers an attractive solution that is easily integrated into LAN, MAN and WAN. To facilitate this trend into the WAN, network providers are beginning to implement Ethernet over existing SONET/SDH optical networks. This application note outlines a number of possible Ethernet over SONET (EOS) implementations using the POS-PHY Level 3 system interface available on the S/UNI-2xGE Dual Gigabit Ethernet Controller and PMC's OC-48/PL3 products such as the S/UNI-2488.

The PM3386 S/UNI-2xGE standard product is a dual channel SERDES and GMAC with embedded FIFOs that provides a high density and low power Gigabit Ethernet solution for direct connection to electrical optical modules. The device also provides a standard GMII interface to Gigabit Ethernet transceivers for 1000BaseT interfaces. On the system side, the PM3386 provides a 104MHz 32bit POS-PHY Level 3 interface for connection to higher layer devices.

The PM5381 S/UNI-2488 standard product supports a single SATURN User Network Interface with SONET/SDH processing, ATM and Packet mapping functions at the STS-48c (STM-16-16c) 2488.32 Mbit/s rate. The PM5381 features integral CRU and CSU for clock/data recovery and generation. The device is intended for use in equipment implementing Asynchronous Transfer Mode (ATM) User-Network Interface (UNI), ATM Network-Network Interfaces (NNI), and Packet Over SONET/SDH (POS) interfaces. The POS interface can be used to support several packet based protocols, including Point-to-Point Protocol (PPP). The S/UNI-2488 may find application at either end of switch-to-switch links or switch-to-terminal links, both in public network (WAN) and private network (LAN) situations.

The PL3 system side architecture common to a number of PMC-Sierra devices is an ideal common interface for simple integration into a multiservice architecture that could include: OC-48, OC-12, OC-3, High Density DS3 or Gigabit Ethernet traffic.

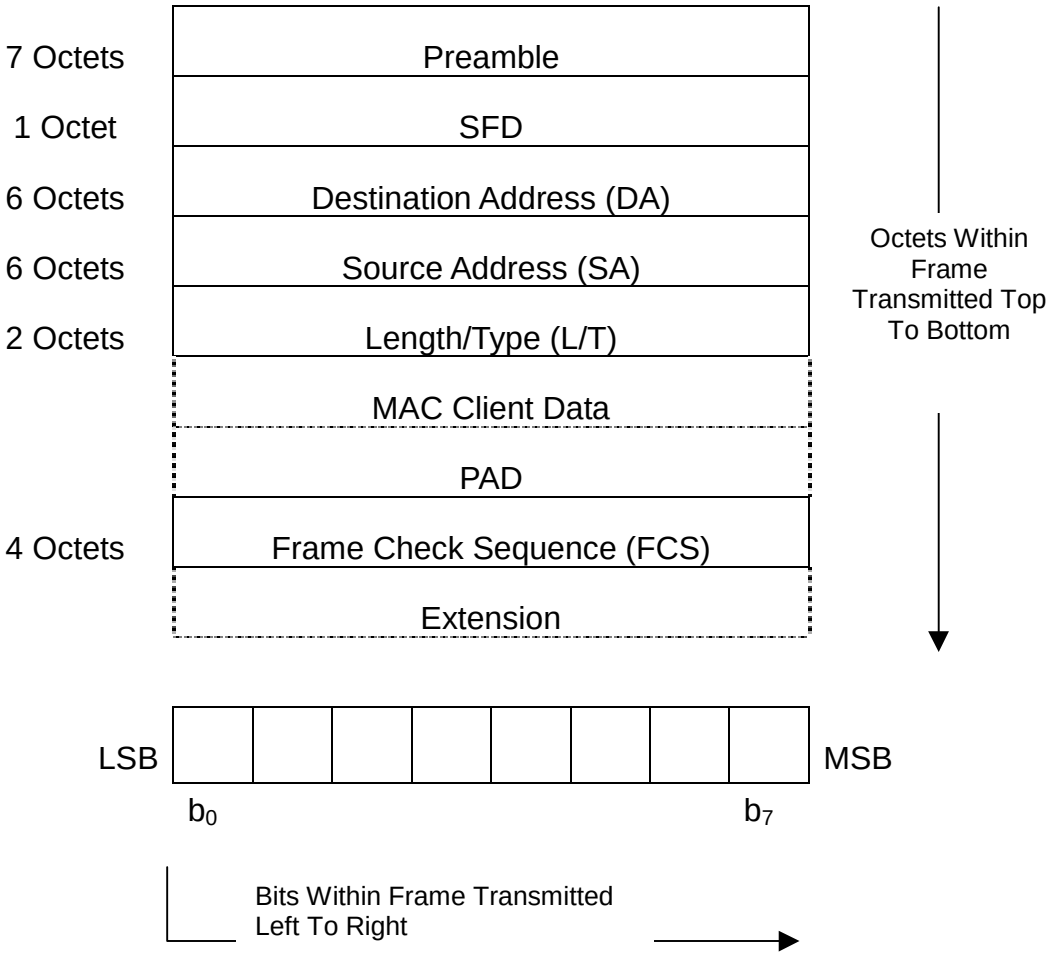
3 ETHERNET OVER SONET WITHIN THE PL3 ARCHITECTURE

PMC offers a fully integrated Dual Gigabit Ethernet solution with it's S/UNI-2xGE device. This device features direct line side interfaces to 1000BASESX and 1000-BASE-LX Gigabit Ethernet optics and interfaces to higher layer devices via a 104MHz POS-PHY Level 3 interface that is common to a number of PMC devices. This common architecture lends itself naturally to multiservice implementations including Ethernet over SONET.

3.1 The S/UNI-2xGE

The S/UNI-2xGE is a dual channel Gigabit Ethernet device that performs data recovery and MAC processing on incoming frames before forwarding them to higher layer devices via the PL3 interface. Figure 1 below outlines the format of a valid MAC frame and Table 1 shows how this frame is mapped into the Big Endian POS-PHY Level 3 compliant packet.

Figure 1 - MAC Frame Format



The POS-PHY interface on the PM3386 supports Big Endian data transfer via packets formatted as shown below in Table 1. Both ingress and egress directions use the same data mapping. Note that the Preamble and Start of Frame Delimiter (SFD) fields are removed from the MAC frame for transmission over the PL3 bus. These frames can be dropped as the Ethernet frame is fully defined for the MAC layer using the octets from the Destination Address to the FCS. Additionally, the Preamble and SFD are ignored during FCS calculation.

Table 1. PM3386 POS-PHY Level 3 Packet Format

Bits 31:24	Bits 23:16	Bits 15:8	Bits 7:0
DA[7:0]	DA[15:8]	DA[23:16]	DA[31:24]
DA[39:32]	DA[47:40]	SA[7:0]	SA[15:8]
SA[23:16]	SA[31:24]	SA[39:32]	SA[47:40]
L/T[7:0]	L/T[15:8]	Data[7:0]	Data[15:8]
Data[23:16]	…	…	…
…	…	…	…
FCS[24:31]	FCS[16:23]	FCS[8:15]	FCS[0:7]

Traffic from the two GE channels is identified via an in-band address that is output at the start of each packet. For additional details regarding the function of the PL3 interface, refer to the POS-PHY Level 3 Specification (PMC-1980495).

The flow control features of the S/UNI-2xGE which include side band control inputs and outputs, or flow control via internal watermark levels in the receive FIFOs can be used in conjunction with features of the S/UNI-2488 (outlined below in Section 3.2) to implement EOS.

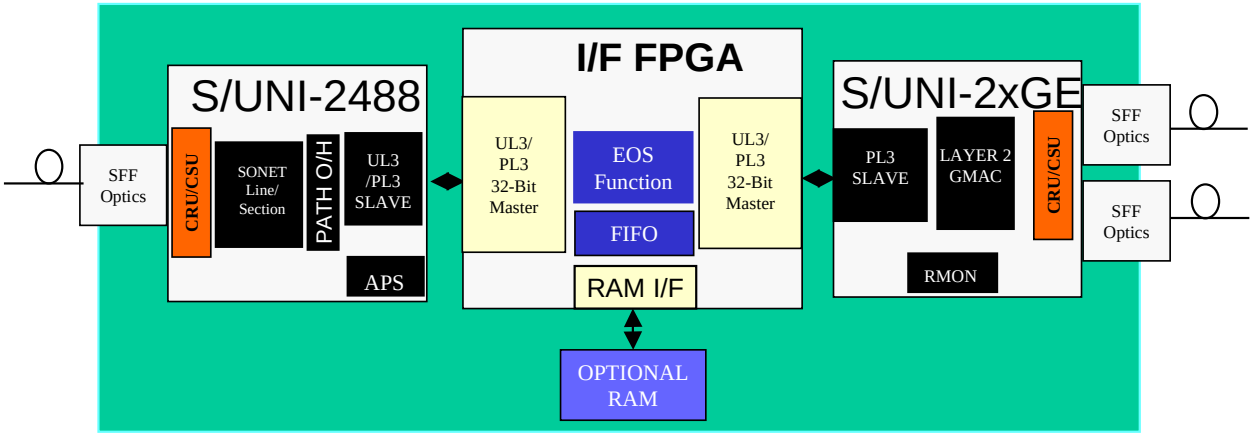
3.2 The S/UNI-2488

The S/UNI-2488 is capable of mapping POS frames into the SONET/SDH STS-48 (STM-16-16c) SPE and terminates applicable section, line and path overhead. The system side POS-PHY interface is capable of processing several protocols including PPP/HDLC and can operate in a transparent mode to support alternate packet transmission protocols such as InterWAN Packet Transfer (IPT) or Simple Data Link Protocol (SDL).

4 PHYSICAL IMPLEMENTATION

Figure 2 below outlines a basic hardware architecture that could be used to implement EOS.

Figure 2 -EOS Physical Architecture



As shown above, the FPGA used to implement the EOS function requires two PL3 interfaces which would use approximately 180 I/Os, and must be capable of internal operation at speeds up to 104MHz. An optional RAM interface for packet buffering, along with additional I/Os required for a micro interface and other generic logic functions would drive the device to over 200 I/Os. The Altera APEX or ACEX devices and the Xilinx Virtex or XC4000 families feature devices with well over 200 I/Os at various gate counts. In addition, both companies are currently developing POS-PHY Level 3 macrocells that can be employed to reduce development time and simplify the design process.

5 ETHERNET OVER SONET MAPPING

The following sections discuss the issues associated with implementing EOS using the S/UNI-2xGE and S/UNI-2488.

5.1 Packet Mapping

The basic functionality required to implement EOS is to properly delineate the incoming or outgoing frames so that they can be processed by either the S/UNI-2488 or the S/UNI-2xGE. Using the PPP over SONET capabilities of the S/UNI-2488 reduces the packet processing requirements of the EOS functions implemented in the interface FPGA, but this is by no means the greatest or only solution. There are alternative packet framing schemes that may be applied to the EOS hardware architecture outlined in this application note. These schemes include InterWAN Packet Transport (IPT) and Simple Data Link (SDL) Protocol.

Specific discussion of the advantages or disadvantages of the various packet framing protocols are beyond the scope of this application note. By operating in transparent mode, the S/UNI-2488 is capable of functioning with any packet mapping function. Section 6 below outlines how to configure the S/UNI-2488 properly for transparent or PPP operation.

Additional details regarding SDL can be found in a paper written by Doshi et.al. entitled *A Simple Data Link Protocol for High-Speed Packet Networks,* Bell Labs Technical Journal, Jan-March, 1999.

5.1.1 Design Considerations

This section discusses the design considerations specific to the hardware architecture outlined in Figure 2 above.

5.1.1.1 Channel Delineation

In transmission from Gigabit Ethernet to SONET, traffic from two Ethernet channels must be mapped into a single SONET channel, and then demapped at the far end. A method for identifying which Ethernet channel the received EOS packets should be routed to is required. This can be done by adding an extra byte to the outgoing (S/UNI-2xGE to S/UNI-2488) packets which contains a channel identifier that can be read at the far end.

Alternatively, channel delineation could be achieved by modifying an existing field within the PL3 packet, such as the Source Address field. This method is easier to

implement than tagging each packet with an additional byte, and will not have any implications on the bandwidth of the system. However, it should be noted that modifying an existing field of the frame will corrupt the FCS causing a receive error at the far end. If there are no concerns regarding modifying a field of the ingress Ethernet Frame, such as with assertion of flow control functionality discussed below in Section 5.1.1.2, this method would be simpler to implement within the FPGA.

5.1.1.2 Far End Flow Control

The appended channel control byte (or modified Source Address field) used to identify the physical channel allocated to the Ethernet traffic can also be used to provide flow control functionality across the SONET link. The S/UNI-2xGE provides side band flow control outputs that identify when PAUSE frames are being received, and inputs that allow the user to assert flow control.

Figure 3 -Far End Flow Control.

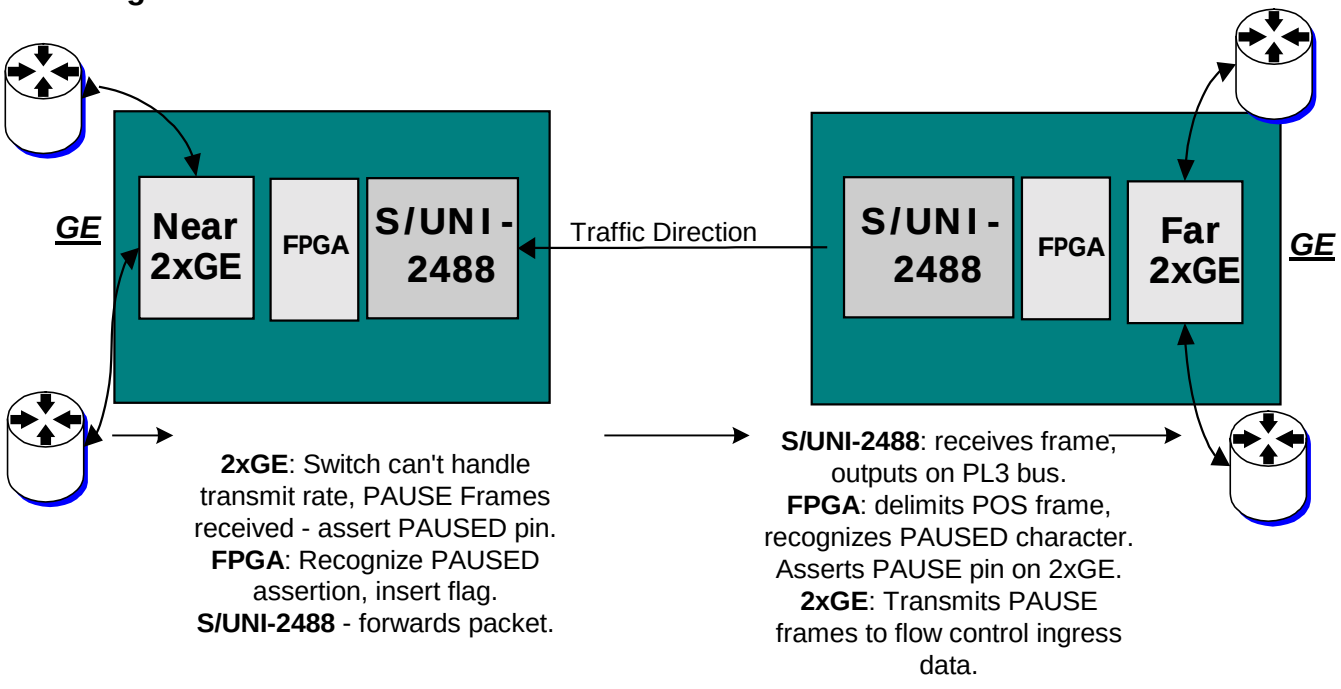


Figure 3 above outlines an implementation to assert flow control to the far side. Using this method the S/UNI-2xGE can filter all PAUSE frames, rather than send them out over the PL3 interface. Upon reception of a PAUSE frame on the near side, the PAUSED pin on the S/UNI-2xGE will be asserted. The FPGA, upon detection of the PAUSED pin, inserts a “PAUSED” character attached to the next outgoing packet being sent to the far side. Using the side band control pins

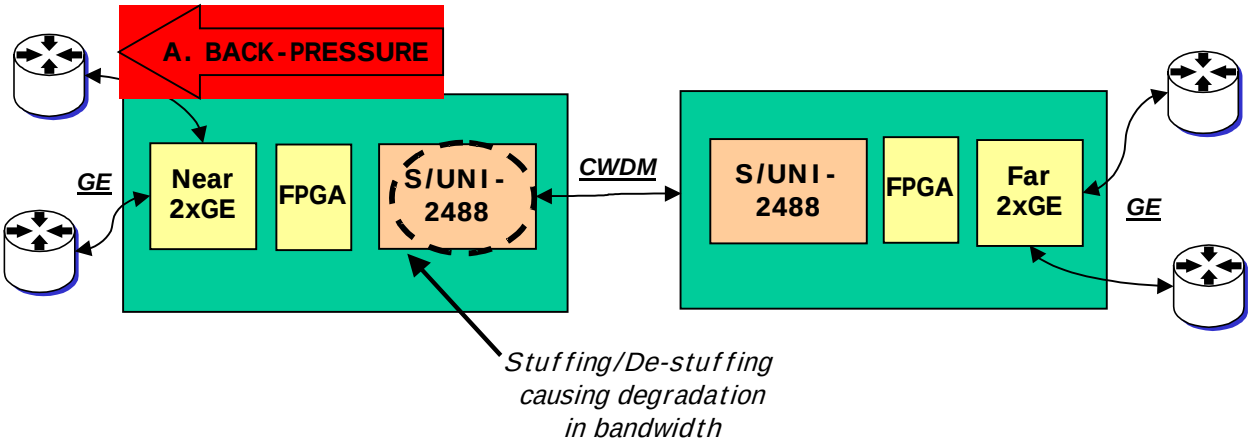
eliminates the need for the FPGA to detect entire PAUSE frames, the FPGA logic only needs to decode the “PAUSED” character in the channel control byte. If packet buffering is implemented in the EOS application to accommodate for delays over long haul distances (See Appendix B for more details on buffering functions) PAUSE frames that are passed through the network are queued along with the rest of the frames and will cause additional delay in asserting flow control. Using the side band pins eliminates delay or loss of PAUSE frames due to queuing.

In addition, there are 16K byte transmit FIFOs associated with each channel. Status pins (STPA, DTPA, PTPA) are available on the PL3 bus and indicate a user configurable fill level of the FIFOs. They allow the user to select a FIFO fill level at which to initiate flow control at the far end, and can also be used to optimize latency through the S/UNI-2xGE.

5.1.1.3 Near End Flow Control

A packet mapping implementation such as HDLC uses unique flag bytes to delineate each frame. To guarantee the flag pattern is unique, it must be removed from the data prior to transmission. Two escape bytes are inserted in place of the flag patterns within the data to avoid misdetection of the frame boundaries. This byte stuffing, and subsequent de-stuffing at the receive side, causes non-deterministic payload expansion, and can affect the ability of the S/UNI-2488 to maintain full bandwidth. Figure 4 below illustrates that the S/UNI-2xGE can assert backpressure onto the near Ethernet switch to compensate for variations in bandwidth.

Figure 4 - Near End Back Pressure



The S/UNI-2xGE has 64Kbyte FIFOs on the receive channels to accommodate for the latency in flow controlling the switch. High and Low watermarks within the

FIFOs can be set to control the point at which flow control is asserted and unasserted.

5.2 Summary

Implementing EOS functionality using the S/UNI-2xGE and the S/UNI-2488 requires an interface FPGA to interconnect the PL3 busses and provide some packet processing capability that can delineate the Ethernet frames for transmission over SONET. The architecture proposed in Figure 2 above can support any packet framing scheme that the user wants to implement by setting the S/UNI-2488 to operate in transparent mode. The S/UNI-2488 also supports PPP over SONET functionality which could be utilized in the EOS application. Modifying the Source Address field or appending an additional byte to the frame would identify which physical channel on the S/UNI-2xGE the traffic is associated with, and can be utilized to carry flow control information to the far side.

6 APPENDIX A: CONFIGURING THE S/UNI-2488

6.1 Transparent POS Mode

The following steps outline how to configure the S/UNI-2488 to operate in transparent POS mode:

1. On the receive side, register 0x740 RCFP Configuration controls the operating mode of the receive system side interface.

- Bit 0 – PROV – Enable the receive processor when set to 1. Defaults to 0.
- Bit 1 – DESCRMBL – Disable packet descrambling when set to 0. Defaults to 1.
- Bits 3 and 4 – CRC_SEL – Disable automatic CRC insertion with 00. Defaults to 11 which enables CRC_32 insertion.
- Bit 5 – CRCPASS – Set to logic 1, packets with FCS errors are not marked as such and are passed to the external FIFO interface as if no FCS error occurred. Defaults to 0.
- Bit 7 –DELINDIS – When set to logic 1 the DELINDIS bit is used to disable the HDLC flag alignment, byte destuffing and flag removal. The data stream is arbitrarily segmented into 62 byte long packets. FCS and descrambling operations still follow how they have been set in their respective configuration registers.
- Bit 10 – POS_SEL – Setting to logic 1 enables Packet processing.

2. On the transmit side the TCFP Configuration register (0x750) is used to control the operating mode of the transmit system side interface.

- Bit 0 – PROV – Enable the transmit processor when set to 1. Defaults to 0.
- Bit 1 –SCRMBL – Disable packet scrambling when set to 0. Defaults to 1.
- Bits 6 and 7 – CRC_SEL – Disable automatic CRC insertion with 00. Defaults to 11 which enables CRC_32 insertion.
- Bit 8 – POS_SEL – Setting to logic 1 enables Packet processing.

- Bit 10 –DELINDIS – When set to logic 1 the DELINDIS bit is used to disable the HDLC flag alignment, byte stuffing and flag insertion. FCS and descrambling operations still follow how they have been set in their respective configuration registers.

Refer to the S/UNI-2488 Data Sheet for more detailed information.

6.2 PPP/HDLC over SONET

The following steps outline how to set up the PPP over SONET functionality on the S/UNI-2488.

1. Enable the transmitter and receiver by setting bit 0 in the TCFP and RCFP Configuration registers.

2. Enable POS mode by setting the POS_SEL bit in the Configuration registers.

3. Enable the desired CRC insertion, idle cell processing, and scrambling functions via the Configuration registers.

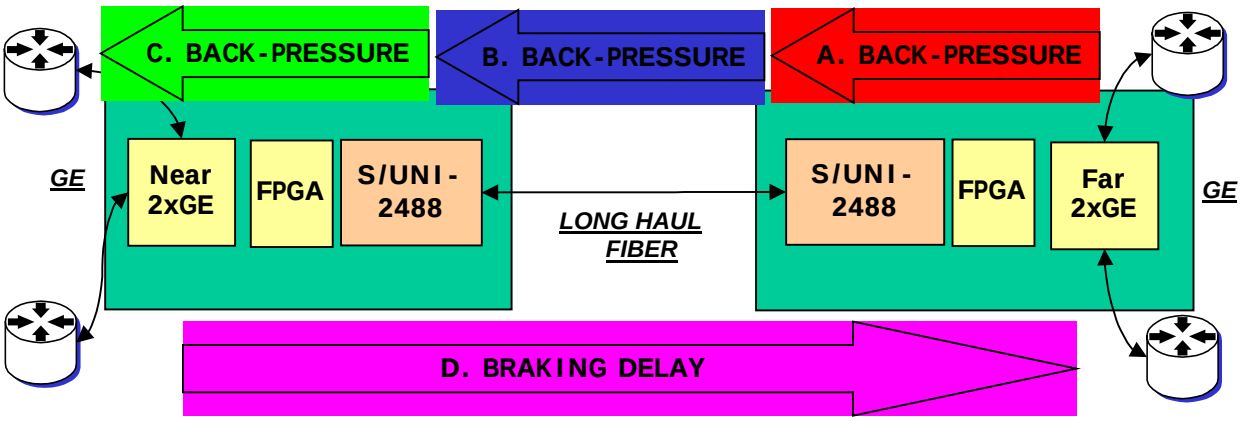
7 APPENDIX B: PACKET BUFFERING

The EOS architecture outlined in Figure 2 shows an optional RAM on the interface FPGA that may be required for packet buffering in certain EOS applications.

If the EOS function is implemented over a long fiber connection that has a considerable delay associated with it, there may be the need or desire to buffer data to accommodate flow control functionality in the presence of far end backpressure.

Figure 5 below shows a possible connection between two EOS ports, and shows the delay associated with a PAUSE request at the near end before it is asserted at the far end.

Figure 5 - Far End Back Pressure Delay



If the far side GE port is requesting flow control, the delay until that flow control function is observed by the far end point is two times the trip time from one GE network to the other, since the PAUSE request has to propagate from the far side to the near side, and then the entire near side to far side pipe has to be cleared before the far side stops seeing frames (the fiber acts like a really big FIFO).

The total delay associated with a single GE port is A+B+C+D or 2x(A+B+C) since A+B+C=D. Where:

	Bytes
Fiber Propagation Delay = (µs/km)	5.085

Fiber Propagation Delay = (bytes/km)	635.625
S/UNI-2488 Delay (worst case) - Based on SONET Overhead, FIFO depth.	1269
FPGA Delay TX – Based on small FIFO	1024
FPGA Delay RX – Based on small FIFO	32000
S/UNI-2xGE Delay – Based on short delay settings	600
Switch Delay- VERY arbitrary. Dependent on individual switches	6000

Note that the values shown in this Appendix are given as an example and are an approximation – not 100% exact values. Especially the Switch delay which is independent for each vendor, and could be very long.

Delay Parameter Calculations:

	Bytes
A = S/UNI-2488+S/UNI-2xGE+FPGA TX Delay =	2893
B = Fiber Delay x 50Km =	31781
C = S/UNI-2488+FPGA RX Delay + S/UNI-2xGE =	39869
D = A+B+C	74543
TOTAL DELAY (bytes)	149087

This shows that the total delay for the effect of a PAUSE frame request to propagate through a 50KM fiber link is approximately 150Kbytes.

The internal FIFO architecture on the S/UNI-2xGE (16K TX + 64K RX) would only be sufficient for short links of less than 1km. Longer links may require external buffering to accommodate far end backpressure.

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
(604) 415-4533
Web Site: http://www.pmc-sierra.com

PMC-2001398 (P1)