



**Multi-link Gearbox
Implementation Agreement**

IA # OIF-MLG-02.0

April 2013

Implementation Agreement created and approved
by the Optical Internetworking Forum
www.oiforum.com



The OIF is an international non profit organization with over 90 member companies, including the world's leading carriers and vendors. Being an industry group uniting representatives of the data and optical worlds, OIF's purpose is to accelerate the deployment of interoperable, cost-effective and robust optical internetworks and their associated technologies. Optical internetworks are data networks composed of routers and data switches interconnected by optical networking elements.

With the goal of promoting worldwide compatibility of optical internetworking products, the OIF actively supports and extends the work of national and international standards bodies. Working relationships or formal liaisons have been established with IEEE 802.1, IEEE 802.3, IETF, IP-MPLS Forum, IPv6 Forum, ITU-T SG13, ITU-T SG15, MEF, ATIS-OPTXS, ATIS-TMOC, TMF and the XFP MSA Group.

For additional information contact:
The Optical Internetworking Forum, 48377 Fremont Blvd.,
Suite 117, Fremont, CA 94538
510-492-4040 ☎ info@oiforum.com

www.oiforum.com

Working Group: Physical and Link Layer

TITLE: Multi-link Gearbox Implementation Agreement 2.0

SOURCE:	TECHNICAL EDITOR	WORKING GROUP CHAIR
	Stephen J. Trowbridge, Ph. D. Alcatel-Lucent 5280 Centennial Trail Boulder, CO 80303 USA Phone: +1 720 945 6885 Email: steve.trowbridge@alcatel-lucent.com	David R. Stauffer, Ph.D. IBM Corporation 1000 River Road, MC 862J Essex Jct., VT 05452, USA Phone: +1 802 769-6914 Email: dstauffe@us.ibm.com

ABSTRACT: The MLG (Multi-Link Gearbox) 2.0 Implementation Agreement defines two MLG configurations: A 4x25G lane configuration is comprised of 20 MLG lanes (similar to MLG 1.0), adding that two groups of eight MLG lanes can be configured to carry either four 10GBASE-R signals or a single 40GBASE-R signal, while the remaining 4 MLG lanes can carry two 10GBASE-R signals. An 8x25G lane configuration is comprised of 40 MLG lanes, where each group of eight MLG lanes can carry either four 10GBASE-R signals or a single 40GBASE-R signal.

Notice: This Technical Document has been created by the Optical Internetworking Forum (OIF). This document is offered to the OIF Membership solely as a basis for agreement and is not a binding proposal on the companies listed as resources above. The OIF reserves the rights to at any time to add, amend, or withdraw statements contained herein. Nothing in this document is in any way binding on the OIF or any of its members.

The user's attention is called to the possibility that implementation of the OIF implementation agreement contained herein may require the use of inventions covered by the patent rights held by third parties. By publication of this OIF implementation agreement, the OIF makes no representation or warranty whatsoever, whether expressed or implied, that implementation of the specification will not infringe any third party rights, nor does the OIF make any representation or warranty whatsoever, whether expressed or implied, with respect to any claim that has been or may be asserted by any third party, the validity of any patent rights related to any such claim, or the extent to which a license to use any such rights may or may not be available or the terms hereof.

© 2013 Optical Internetworking Forum

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction other than the following, (1) the above copyright notice and this paragraph must be included on all such copies and derivative works, and (2) this document itself may not be modified in any way, such as by removing the copyright notice or references to the OIF, except as needed for the purpose of developing OIF Implementation Agreements.

By downloading, copying, or using this document in any manner, the user consents to the terms and conditions of this notice. Unless the terms and conditions of this notice are breached by the user, the limited permissions granted above are perpetual and will not be revoked by the OIF or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE OIF DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY, TITLE OR FITNESS FOR A PARTICULAR PURPOSE.

1 Table of Contents

1	Table of Contents.....	4
2	List of Figures.....	5
3	List of Tables.....	5
4	Document Revision History.....	6
5	Introduction.....	7
5.1	Requirements.....	7
5.2	Sample Applications.....	8
6	General Mechanism.....	12
6.1	MLG Lane Markers.....	13
7	Detailed Block Diagrams.....	15
7.1	MLG mux.....	16
7.2	MLG demux.....	22
7.3	Generic MLG Management.....	30
8	References.....	31
8.1	Normative references.....	31
9	Appendix A (Informative): Suggested MPO Receptacle Physical Lane Assignments for various MLG applications.....	31
10	Appendix B (Informative): Alternate Approach for mapping of 40GBASE-R over 4x20G lanes.....	37
11	Appendix C: List of companies belonging to OIF when document was approved 40	

2 List of Figures

FIGURE 1: 10GBASE-R VIRTUAL LINK	8
FIGURE 2: 40GBASE-R AND 10GBASE-R VIRTUAL LINK.....	9
FIGURE 3: 40GBASE-R VIRTUAL LINK USING 8X25G MLG GEARBOX	9
FIGURE 4: 10GBASE-R PORT EXPANDER	10
FIGURE 5: 40GBASE-R AND 10GBASE-R PORT EXPANDER.....	11
FIGURE 6: 40GBASE-R PORT EXPANDER	12
FIGURE 7: MLG MUX 10GBASE-R MAPPING BLOCK DIAGRAM	16
FIGURE 8: MLG MUX 40GBASE-R MAPPING BLOCK DIAGRAM	19
FIGURE 9: MLG DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM	23
FIGURE 10: MLG DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM	26
FIGURE 11: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG 10GBASE-SR OR 10GBASE-LR OPTICAL PORT EXPANDER.....	31
FIGURE 12: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 10GBASE-SR AND 40GBASE-SR4 PORT EXPANDER	32
FIGURE 13: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 40GBASE-SR4 PORT EXPANDER.....	33
FIGURE 14: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 10GBASE-LR AND 40GBASE-LR4 PORT EXPANDER.....	34
FIGURE 15: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 40GBASE-LR4 PORT EXPANDER.....	35
FIGURE 16: MPO CONNECTOR PHYSICAL LANE ASSIGNMENTS FOR 8X25G OPTICAL 40GBASE-R PORT EXPANDER	36
FIGURE 17: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 8X25G OPTICAL 40GBASE- SR4 PORT EXPANDER.....	37
FIGURE 18: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X20G OPTICAL 40GBASE- SR4 PORT EXPANDER.....	38
FIGURE 19: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X20G OPTICAL 40GBASE- LR4 PORT EXPANDER	39

3 List of Tables

TABLE 1: MLG 2.0 4X25G LANE ALIGNMENT MARKER VALUES.....	14
TABLE 2: MLG 2.0 8X25G LANE ALIGNMENT MARKER VALUES.....	15

4 Document Revision History

Working Group: Physical and Link Layer

SOURCE:	Editor's Name Stephen J. Trowbridge, Ph. D. Alcatel-Lucent 5280 Centennial Trail Boulder, CO 80303 USA Phone: +1 720 945 6885 Email: steve.trowbridge@alcatel-lucent.com	Working Group Chair David R. Stauffer, Ph.D. IBM Corporation 1000 River Road, MC 862J Essex Jct., VT 05452, USA Phone: +1 802 769-6914 Email: dstauffe@us.ibm.com

DATE: April 2013

Issue No.	Issue Date	Details of Change
OIF2012.325.00	October 2012	Initial Text Proposal
OIF2012.325.01	January 2013	Text after resolution of Straw Ballot #172 comments
OIF2012.325.02	April 2013	Text basis for publication after Principal Ballot #82

5 Introduction

The MLG (Multi-link Gearbox) 2.0 implementation agreement defines an in-band coding that allows independent 10GBASE-R and 40GBASE-R signals to transit 4x25G and 8x25G gearboxes implementing a 100GBASE-R PMA function (or an expanded 8x25G lane variant). This enables a variety of applications to reuse 100GBASE-R technology for the transport of individual 10G and 40G links.

5.1 Requirements

Two applications are defined:

- The MLG 2.0 4x25G gearbox application maps 10GBASE-R and/or 40GBASE-R signals over 20 MLG lanes, with five MLG lanes bit-multiplexed over each physical 25G lane. Two groups of eight MLG lanes can each be configured to carry four 10GBASE-R signals or a single 40GBASE-R signal, while the remaining four MLG lanes can carry two 10GBASE-R signals. When both of the 8 MLG lane groups are configured to carry four 10GBASE-R signals, the MLG 2.0 4x25G gearbox is compatible with, and can be interconnected with, an MLG 1.0 gearbox.
- The MLG 2.0 8x25G gearbox application maps 10GBASE-R and/or 40GBASE-R signals over 40 MLG lanes, with five MLG lanes bit-multiplexed over each physical 25G lane. Five groups of eight MLG lanes can each be configured to carry four 10GBASE-R signals or a single 40GBASE-R signal. While a particular device might be implemented so that it can be configured to provide two 4x25G gearbox applications or a single 8x25G gearbox application, the two applications cannot be interconnected.

The MLG 2.0 4x25G mux encodes from zero to ten 10GBASE-R signals compliant with IEEE Std 802.3™-2012 clauses 49 and 51, and from zero to two 40GBASE-R signals compliant with IEEE Std 802.3-2012 clauses 82 and 83 into a format consistent with the IEEE Std 802.3-2012 clause 83 PMA service interface (e.g., four physical lanes of 25.78125 Gb/s \pm 100ppm). The MLG 2.0 4x25G demux decodes the format produced by the MLG 2.0 4x25G mux to produce from zero to ten 10GBASE-R signals and from zero to two 40GBASE-R signals. When configured to carry only 10GBASE-R signals, the MLG 2.0 4x25G mux can be interconnected with an MLG 1.0 demux, and an MLG 1.0 mux can be interconnected with an MLG 2.0 4x25G demux. Note – for applications requiring only support of two 40GBASE-R signals over four physical lanes (without 10GBASE-R support), the approach described in Appendix B may also be considered.

The MLG 2.0 8x25G mux encodes from zero to twenty 10GBASE-R signals compliant with IEEE Std 802.3-2012 clauses 49 and 51 and from zero to five 40GBASE-R signals compliant with IEEE Std 802.3-2012 clauses 82 and 83 into a format consistent with an expanded interface similar to a double-width 100GBASE-R PMA (as if it were comprised of double the number of PCS lanes striped over double the number of physical lanes within the same skew and skew variation limits). The MLG 2.0 8x25G demux decodes the format produced by the MLG 2.0 8x25G mux to produce from zero to twenty 10GBASE-R signals and from zero to five 40GBASE-R signals.

The MLG 2.0 mux to MLG 2.0 demux link preserves the compliance to IEEE Std 802.3-2012 clauses 49 and 51 of each 10GBASE-R signal and the compliance to IEEE Std 802.3-2012 clauses 82 and 83 of each 40GBASE-R signal. Note that the exact bit sequence of a given 10GBASE-R signal or 40GBASE-R signal is not preserved across

Multi-link Gearbox 2.0 Implementation Agreement

an MLG 2.0 mux to MLG 2.0 demux link. The bit sequence may be modified by the insertion and/or deletion of idles and rescrumbling, or in the case of 40GBASE-R, restriping the 66-bit blocks over different PCS lanes after idle insertion or deletion. Note also that non-64B/66B formatted 10G signals such as pseudo-random or PRBS cannot transit an MLG link.

The skew and skew variation limits for IEEE 802.3-2012 clause 80 from SP1 to SP6 for a 100GBASE-R PCS lane are met by an MLG 2.0 mux to MLG 2.0 demux link.

Operating 10G lanes and 40G lanes are not affected by turning on or off any other 10G or 40G lane, nor by the failure of any other 10G or 40G lane.

An MLG 2.0 mux to MLG 2.0 demux link can preserve the long-term average clock frequency of a single selected 10GBASE-R or a selected 40GBASE-R signal, or all signals if they are originally from a common clock source, but how this is achieved is outside the scope of this IA.

5.2 Sample Applications

5.2.1 Virtual Link

The 10GBASE-R Virtual Link function uses the MLG to transport up to ten (asynchronous) 10GBASE-R signals across a 4x25G MLD gearbox using a PMD defined for 100GBASE-R (for example, a 100GBASE-LR4 or 100GBASE-ER4 PMD per IEEE Std 802.3-2012 clause 88). Note that up to twenty 10GBASE-R signals could be carried using an 8x25G MLG gearbox:

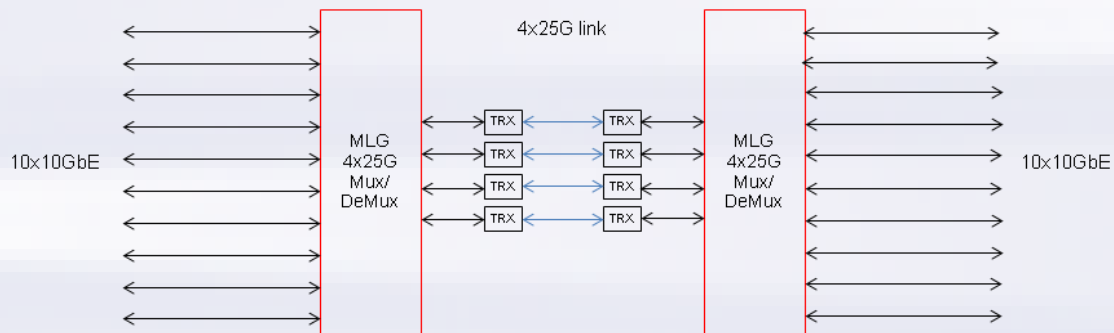


Figure 1: 10GBASE-R Virtual Link

Up to two (asynchronous) 40GBASE-R signals may be transported over a virtual link using a 4x25G MLG gearbox. Since this does not use all of the capacity of the 4x25G link, the remaining capacity can carry up to two 10GBASE-R signals:

Multi-link Gearbox 2.0 Implementation Agreement

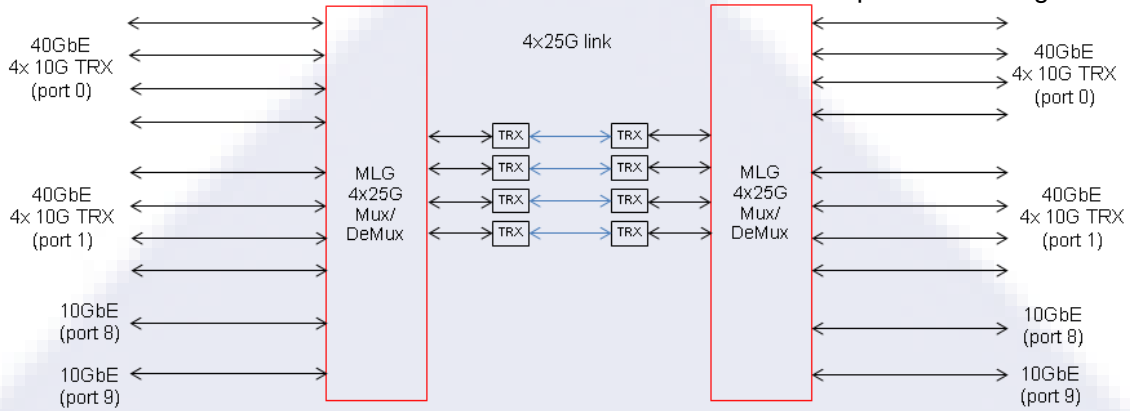


Figure 2: 40GBASE-R and 10GBASE-R Virtual Link

For the most efficient implementation of an “all 40G” application, an 8x25G MLG gearbox can be used to provide a virtual link for up to five 40GBASE-R signals:

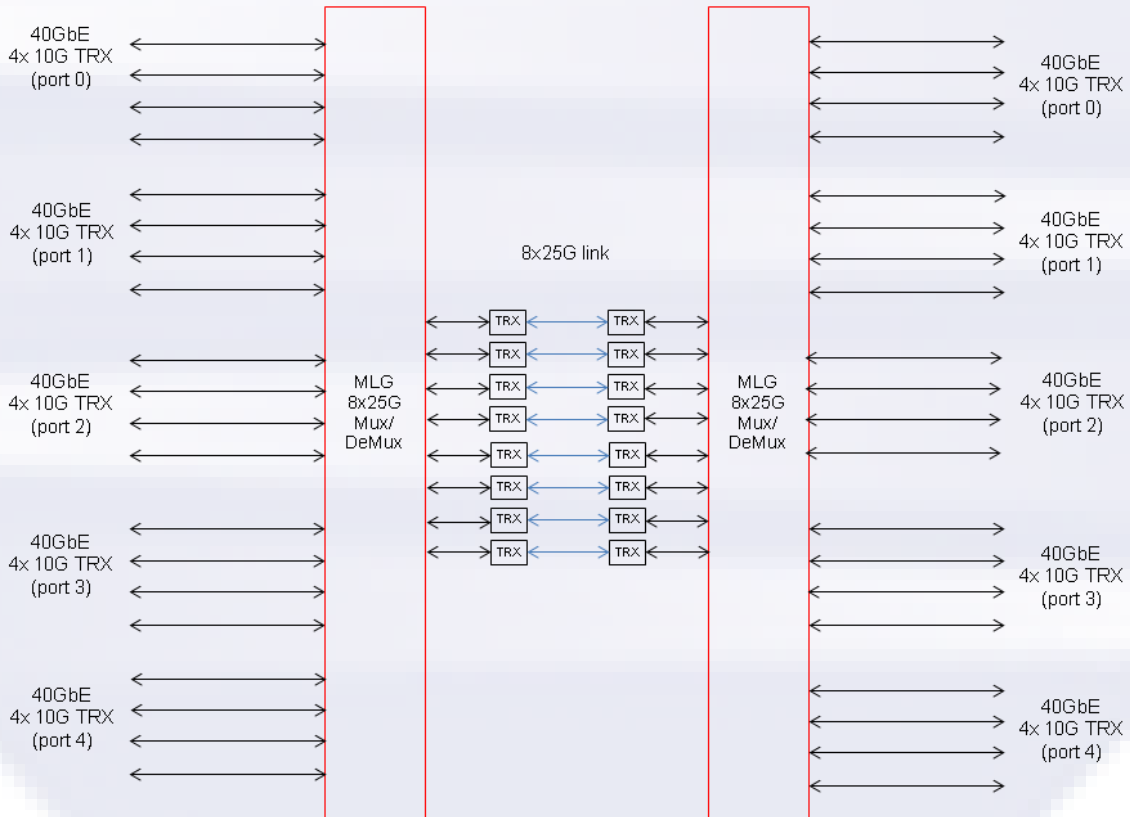


Figure 3: 40GBASE-R Virtual Link using 8x25G MLG gearbox

5.2.2 Port Expander

The 10GBASE-R port expander enables high density 10G I/O using module interfaces, form factors, and higher speed ASIC interfaces designed for 100GBASE-R. This could be applied to electrical or optical interfaces.

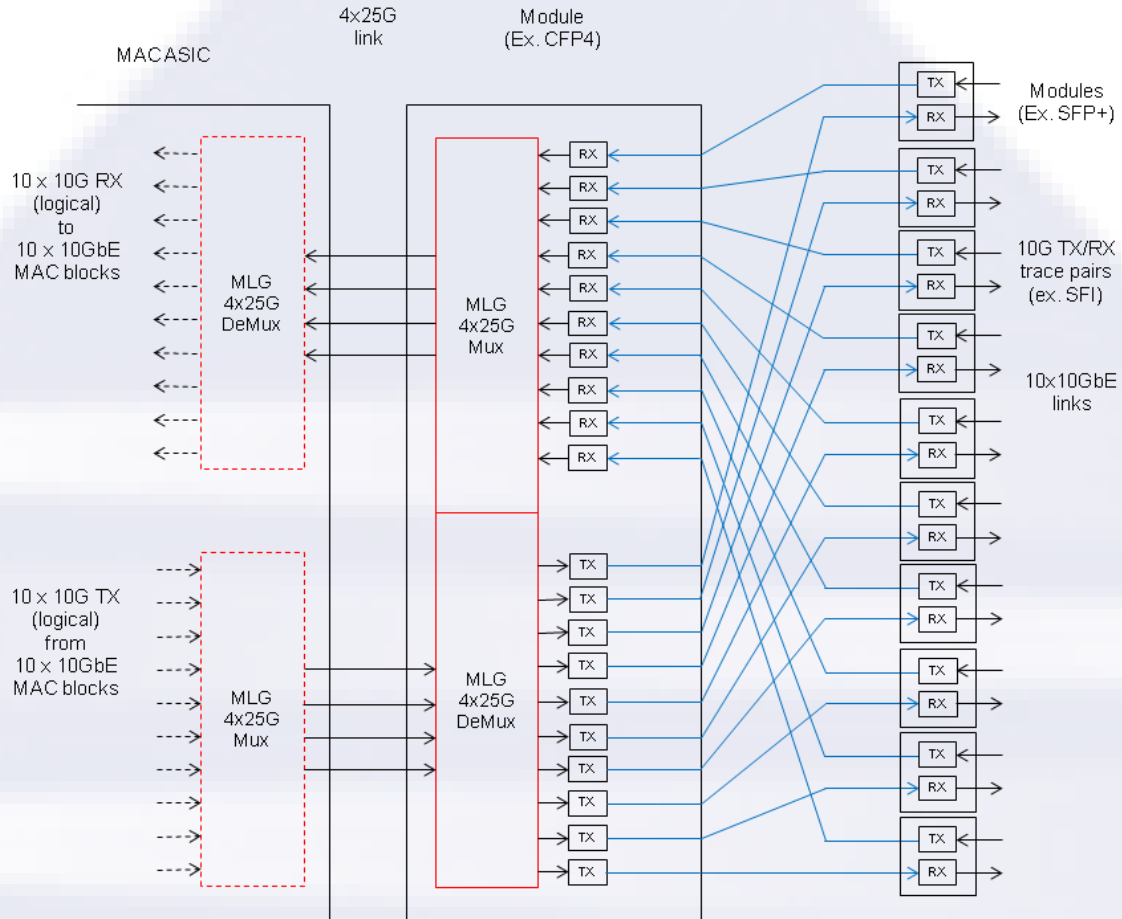


Figure 4: 10GBASE-R Port Expander

The 4x25G MLG gearbox may also be used to support 40GBASE-R. Since the full 100G link cannot be 100% utilized for 40GBASE-R, capacity is available to support two 10GBASE-R in addition to two 40GBASE-R.

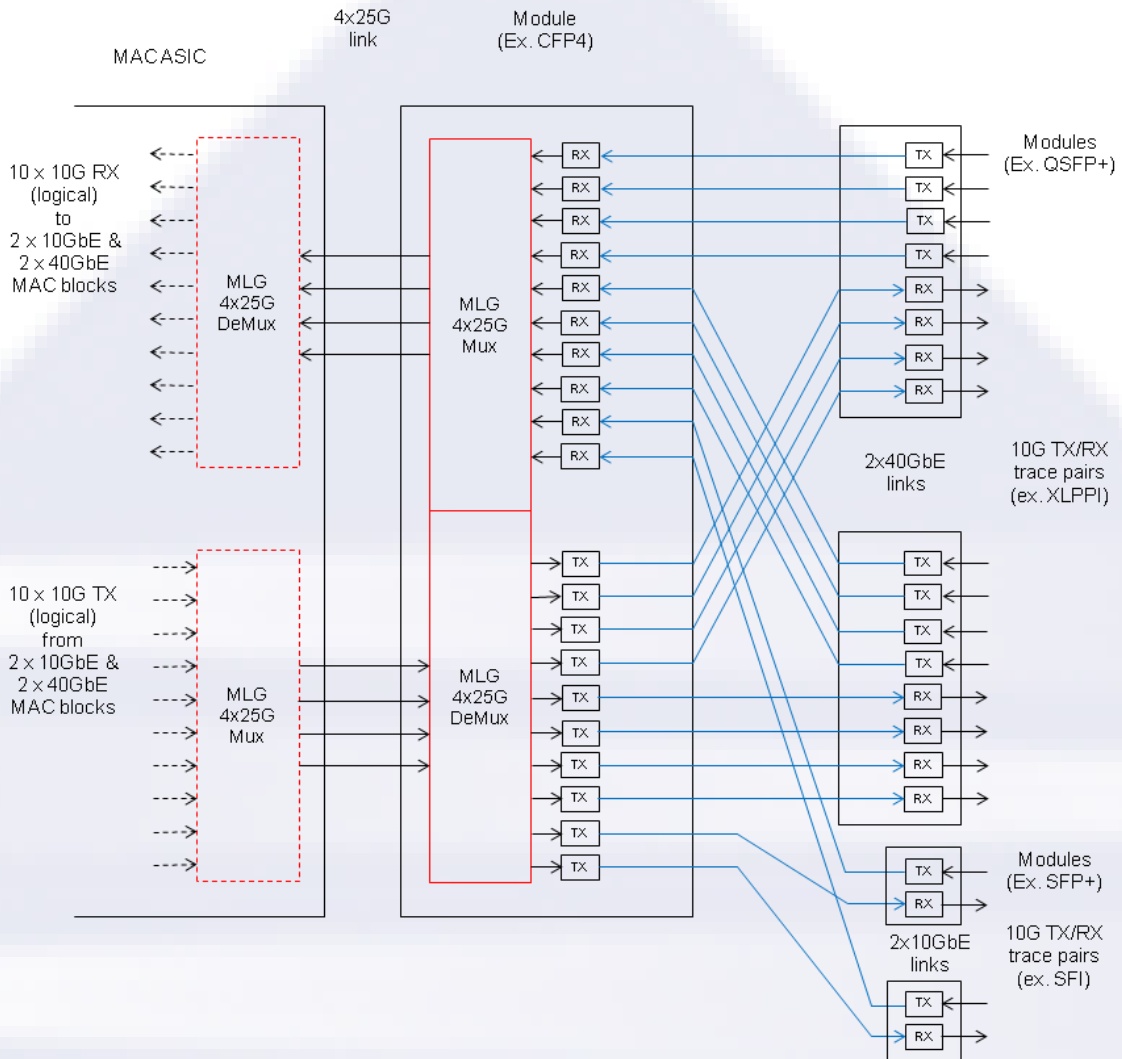


Figure 5: 40GBASE-R and 10GBASE-R Port Expander

Higher capacity and 40G-only applications with 100% utilization can be provided using the 8x25G gearbox configuration:

Multi-link Gearbox 2.0 Implementation Agreement

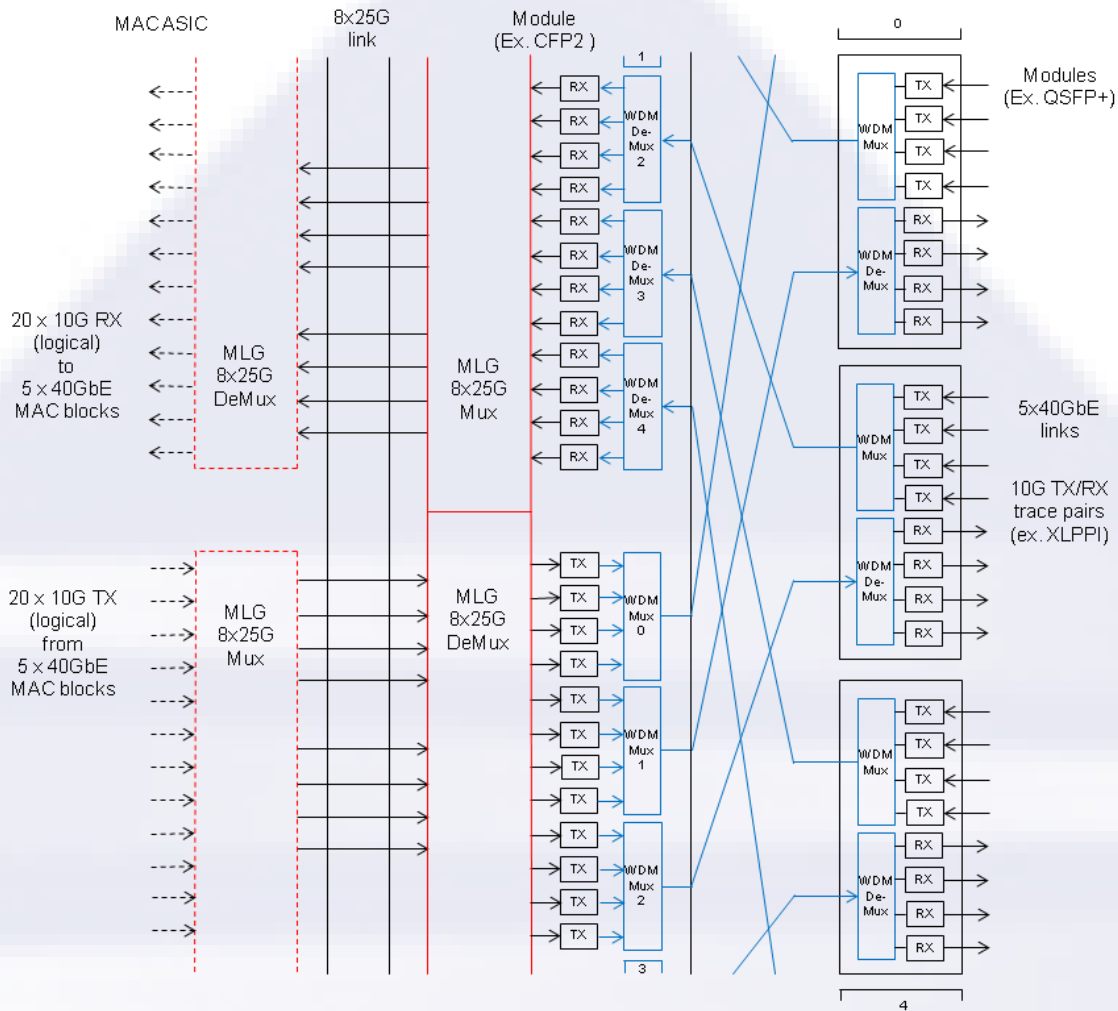


Figure 6: 40GBASE-R Port Expander

6 General Mechanism

The MLG 2.0 4x25G mechanism reuses the 100GBASE-R PMA, which can combine the information from the 100GBASE-R PCS into a variety of different physical lane widths. The 100GBASE-R PCS is divided into 20 PCS lanes, distributing the 66-bit blocks of the 103.125 Gb/s aggregate at the PCS Tx round-robin to twenty lanes of 5.15625 Gb/s. Sufficient idles are deleted from the 103.125 Gb/s aggregate bit stream (prior to scrambling and block distribution) to allow for insertion of a 66-bit PCS lane alignment marker after every 16383 66-bit blocks on each PCS lane. The PCS lane alignment markers allow for identification and deskew of the PCS lanes at the Rx end of the link. At the Rx, the lanes are identified, reordered, and deskewed and the PCS lane alignment markers are removed to reassemble the original aggregate sequence of 66-bit blocks. This PCS mechanism is described in IEEE Std 802.3-2012 clause 82.

A similar approach is employed by the MLG to transport 10GBASE-R and 40GBASE-R signals. The incoming PCS lane markers are removed from the 40GBASE-R PCS lanes and sufficient idles are added or removed from each of the 10GBASE-R or 40GBASE-R signals to map them all to a common clock domain and to make room for the insertion of

the 66-bit MLG lane alignment markers described below. Each 10GBASE-R signal is then demultiplexed into two MLG lanes (running at 5.15625 Gb/s) by alternating 66-bit blocks on each of the two MLG lanes, and each 40GBASE-R signal is demultiplexed into eight MLG lanes (running at 5.15625 Gb/s), distributing 66-bit blocks round-robin to each of the eight MLG lanes. A 66-bit MLG lane alignment (and identification) marker is simultaneously inserted on each of these MLG lanes after every 16383 66-bit blocks. The MLG lanes are identified as lane x.y, where x=0 to 9 (for the 4x25G gearbox application) or x=0 to 19 (for the 8x25G gearbox application), and y=0 or 1. For mapping of 10GBASE-R, x indicates which of the 10GBASE-R signals is being carried, and y=0 or 1 identifies the two MLG lanes that comprise a particular 10GBASE-R signal. Groups of MLG lanes which can carry four 10GBASE-R signals can be configured to carry a single 40GBASE-R signal using the same MLG lane identifiers, but different MLG lane alignment markers as described in section 6.1.

At the MLG demux, the MLG lanes are identified, deskewed, reordered, and the MLG lane alignment markers removed. For demapping of 10GBASE-R, pairs of MLG lanes comprising each 10GBASE-R signal are reinterleaved on a 66-bit block basis. Idles can then be added or removed from each 10GBASE-R signal to map it onto a new 10G output clock domain (if required). For demapping of 40GBASE-R, the eight MLG lanes comprising the 40GBASE-R are reinterleaved on a 66-bit block basis. Idles can be added or removed from the 40GBASE-R to map it to a new 40G output clock domain (if required). The 66-bit blocks are then distributed round-robin to four 40GBASE-R PCS lanes and the 40G PCS lane alignment markers for those lanes are inserted.

Note that in the case that mapping to a new 40G output clock domain is not required, in the demapping of 40GBASE-R by the MLG demux, the number of MLG lane alignment markers removed is the same as the number of PCS lane alignment markers inserted. So while the 40GBASE-R MLG lanes will still need to be recovered, deskewed and reordered to stay within 40GBASE-R skew and skew variation limits, this process can be done without descrambling and rescrumbling the non-lane marker blocks.

6.1 MLG Lane Markers

The MLG 2.0 lane marker values are chosen so as not to replicate the values used in the 100GBASE-R PCS or 40GBASE-R PCS. This prevents bringing up a link which accidentally interconnects an MLG formatted signal with a 100GBASE-R PCS. In addition, when 40GBASE-R is mapped into those MLG lanes half of the lane markers have different values than they do when 10G is mapped into those MLG lanes. This prevents mistakenly demapping 66-bit blocks of a 40GBASE-R as four 10GBASE-Rs or demapping 66-bit blocks of four 10GBASE-Rs as a 40GBASE-R in the case of mis-configuration between the MLG mux and the MLG demux.

Table 1 provides the MLG lane alignment markers used for the MLG 2.0 4x25G application. MLG lanes 0.0, 0.1, 1.0, 1.1, 2.0, 2.1, 3.0, and 3.1 may be provisioned to carry four 10GBASE-Rs or a single 40GBASE-R. The MLG lane marker values x.0 have different values depending on whether 10GBASE-R or 40GBASE-R is mapped into those MLG lanes as indicated. Similarly MLG lanes 4.0, 4.1, 5.0, 5.1, 6.0, 6.1, 7.0 and 7.1 may be provisioned to carry four 10GBASE-Rs or a single 40GBASE-R. In the 4x25G application, MLG lanes 8.0, 8.1, 9.0, and 9.1 are only capable of carrying two 10GBASE-Rs. Note that the MLG lane marker values used when all MLG lanes of an MLG 2.0 4x25G application are provisioned to carry 10GBASE-R are identical to those for MLG 1.0, so these implementations can be interconnected with appropriate provisioning.

Table 1: MLG 2.0 4x25G Lane Alignment Marker Values

MLG lane number	Encoding ^a {M0, M1, M2, BIP ₃ , M4, M5, M6, BIP ₇ }	MLG lane number	Encoding ^a {M0, M1, M2, BIP ₃ , M4, M5, M6, BIP ₇ }
0.0-10	0x80, 0xB4, 0xAF, BIP ₃ , 0x7F, 0x4B, 0x50, BIP ₇	0.1	0x29, 0x85, 0x1D, BIP ₃ , 0xD6, 0x7A, 0xE2, BIP ₇
0.0-40	0x89, 0x40, 0x9F, BIP ₃ , 0x76, 0xBF, 0x60, BIP ₇		
1.0-10	0x11, 0x2A, 0xD8, BIP ₃ , 0xEE, 0xD5, 0x27, BIP ₇	1.1	0xBF, 0x7E, 0x4D, BIP ₃ , 0x40, 0x81, 0xB2, BIP ₇
1.0-40	0xAA, 0x39, 0xE3, BIP ₃ , 0x55, 0xC6, 0x1C, BIP ₇		
2.0-10	0x7C, 0x3F, 0x1C, BIP ₃ , 0x83, 0x00, 0xE3, BIP ₇	2.1	0xEE, 0x8B, 0xBA, BIP ₃ , 0x11, 0x74, 0x45, BIP ₇
2.0-40	0x14, 0x6B, 0xD7, BIP ₃ , 0xEB, 0x94, 0x28, BIP ₇		
3.0-10	0xD1, 0x87, 0x25, BIP ₃ , 0x2E, 0x78, 0xDA, BIP ₇	3.1	0xD0, 0x02, 0x39, BIP ₃ , 0x2F, 0xFD, 0xC6, BIP ₇
3.0-40	0xE1, 0xDB, 0x6C, BIP ₃ , 0x1E, 0x24, 0x93, BIP ₇		
4.0-10	0x6D, 0xFE, 0x11, BIP ₃ , 0x92, 0x01, 0xEE, BIP ₇	4.1	0xA1, 0xD2, 0xAB, BIP ₃ , 0x5E, 0x2D, 0x54, BIP ₇
4.0-40	0x39, 0xB8, 0x5C, BIP ₃ , 0xC6, 0x47, 0xA3, BIP ₇		
5.0-10	0x0E, 0xC6, 0x3C, BIP ₃ , 0xF1, 0x39, 0xC3, BIP ₇	5.1	0x98, 0x78, 0x07, BIP ₃ , 0x67, 0x87, 0xF8, BIP ₇
5.0-40	0x4A, 0x59, 0x12, BIP ₃ , 0xB5, 0xA6, 0xED, BIP ₇		
6.0-10	0x1B, 0xBF, 0xA0, BIP ₃ , 0xE4, 0x40, 0x5F, BIP ₇	6.1	0x31, 0x90, 0xC3, BIP ₃ , 0xCE, 0x6F, 0x3C, BIP ₇
6.0-40	0x55, 0xD3, 0xC6, BIP ₃ , 0xAA, 0x2C, 0x39, BIP ₇		
7.0-10	0x0D, 0x9A, 0x46, BIP ₃ , 0xF2, 0x65, 0xB9, BIP ₇	7.1	0x9F, 0x08, 0xB6, BIP ₃ , 0x60, 0xF7, 0x49, BIP ₇
7.0-40	0xB6, 0xA2, 0xCF, BIP ₃ , 0x49, 0x5D, 0x30, BIP ₇		
8.0	0xBB, 0x55, 0x9D, BIP ₃ , 0x44, 0xAA, 0x62, BIP ₇	8.1	0xA8, 0x05, 0xFC, BIP ₃ , 0x57, 0xFA, 0x03, BIP ₇
9.0	0x04, 0xA1, 0x94, BIP ₃ , 0xFB, 0x5E, 0x6B, BIP ₇	9.1	0x07, 0x72, 0xDB, BIP ₃ , 0xF8, 0x8D, 0x24, BIP ₇

The MLG 2.0 lane alignment markers for the 8x25G application are given in Table 2. MLG lanes can be provisioned so that each of five groups of eight MLG lanes (MLG lanes 0.x-3.x, 4.x-7.x, 8.x-11.x, 12.x-15.x, and 16.x-19.x) can be provisioned to carry either four 10GBASE-Rs or a single 40GBASE-R. As above for the 4x25G application, the 8x25G application uses different MLG lane alignment markers in half of the MLG lane positions when a group of MLG lanes is carrying 40GBASE-R rather than 10GBASE-R to prevent demapping at the wrong rate in the case of mis-provisioning between the MLG mux and the MLG demux.

Table 2: MLG 2.0 8x25G Lane Alignment Marker Values

MLG lane number	Encoding ^a {M0, M1, M2, BIP ₃ , M4, M5, M6, BIP ₇ }	MLG lane number	Encoding ^a {M0, M1, M2, BIP ₃ , M4, M5, M6, BIP ₇ }
0.0-10	0x80, 0xB4, 0xAF, BIP ₃ , 0x7F, 0x4B, 0x50, BIP ₇	0.1	0x29, 0x85, 0x1D, BIP ₃ , 0xD6, 0x7A, 0xE2, BIP ₇
0.0-40	0x89, 0x40, 0x9F, BIP ₃ , 0x76, 0xBF, 0x60, BIP ₇		
1.0-10	0x11, 0x2A, 0xD8, BIP ₃ , 0xEE, 0xD5, 0x27, BIP ₇	1.1	0xBF, 0x7E, 0x4D, BIP ₃ , 0x40, 0x81, 0xB2, BIP ₇
1.0-40	0xAA, 0x39, 0xE3, BIP ₃ , 0x55, 0xC6, 0x1C, BIP ₇		
2.0-10	0x7C, 0x3F, 0x1C, BIP ₃ , 0x83, 0xC0, 0xE3, BIP ₇	2.1	0xEE, 0x8B, 0xBA, BIP ₃ , 0x11, 0x74, 0x45, BIP ₇
2.0-40	0x14, 0x6B, 0xD7, BIP ₃ , 0xEB, 0x94, 0x28, BIP ₇		
3.0-10	0xD1, 0x87, 0x25, BIP ₃ , 0x2E, 0x78, 0xDA, BIP ₇	3.1	0xD0, 0x02, 0x39, BIP ₃ , 0x2F, 0xFD, 0xC6, BIP ₇
3.0-40	0xE1, 0xDB, 0x6C, BIP ₃ , 0x1E, 0x24, 0x93, BIP ₇		
4.0-10	0x6D, 0xFE, 0x11, BIP ₃ , 0x92, 0x01, 0xEE, BIP ₇	4.1	0xA1, 0xD2, 0xAB, BIP ₃ , 0x5E, 0x2D, 0x54, BIP ₇
4.0-40	0x39, 0xB8, 0x5C, BIP ₃ , 0xC6, 0x47, 0xA3, BIP ₇		
5.0-10	0x0E, 0xC6, 0x3C, BIP ₃ , 0xF1, 0x39, 0xC3, BIP ₇	5.1	0x98, 0x78, 0x07, BIP ₃ , 0x67, 0x87, 0xF8, BIP ₇
5.0-40	0x4A, 0x59, 0x12, BIP ₃ , 0xB5, 0xA6, 0xED, BIP ₇		
6.0-10	0x1B, 0xBF, 0xA0, BIP ₃ , 0xE4, 0x40, 0x5F, BIP ₇	6.1	0x31, 0x90, 0xC3, BIP ₃ , 0xCE, 0x6F, 0x3C, BIP ₇
6.0-40	0x55, 0xD3, 0xC6, BIP ₃ , 0xAA, 0x2C, 0x39, BIP ₇		
7.0-10	0x0D, 0x9A, 0x46, BIP ₃ , 0xF2, 0x65, 0xB9, BIP ₇	7.1	0x9F, 0x08, 0xB6, BIP ₃ , 0x60, 0xF7, 0x49, BIP ₇
7.0-40	0xB6, 0xA2, 0xCF, BIP ₃ , 0x49, 0x5D, 0x30, BIP ₇		
8.0-10	0xBB, 0x55, 0x9D, BIP ₃ , 0x44, 0xAA, 0x62, BIP ₇	8.1	0xA8, 0x05, 0xFC, BIP ₃ , 0x57, 0xFA, 0x03, BIP ₇
8.0-40	0x20, 0x37, 0x16, BIP ₃ , 0xDF, 0xC8, 0xE9, BIP ₇		
9.0-10	0x04, 0xA1, 0x94, BIP ₃ , 0xFB, 0x5E, 0x6B, BIP ₇	9.1	0x07, 0x72, 0xDB, BIP ₃ , 0xF8, 0x8D, 0x24, BIP ₇
9.0-40	0xCA, 0xBC, 0x1A, BIP ₃ , 0x35, 0x43, 0xE5, BIP ₇		
10.0-10	0x37, 0xA0, 0x0D, BIP ₃ , 0xC8, 0x5F, 0xF2, BIP ₇	10.1	0xE6, 0xEF, 0x8C, BIP ₃ , 0x19, 0x10, 0x73, BIP ₇
10.0-40	0x7F, 0xC0, 0xCC, BIP ₃ , 0x80, 0x3F, 0x33, BIP ₇		
11.0-10	0x6A, 0xD9, 0x73, BIP ₃ , 0x95, 0x26, 0x8C, BIP ₇	11.1	0x41, 0xDE, 0x0A, BIP ₃ , 0xBE, 0x21, 0xF5, BIP ₇
11.0-40	0x32, 0xFA, 0xE2, BIP ₃ , 0xCD, 0x05, 0x1D, BIP ₇		
12.0-10	0xDF, 0x13, 0x45, BIP ₃ , 0x20, 0xEC, 0xBA, BIP ₇	12.1	0x8D, 0x0D, 0xBC, BIP ₃ , 0x72, 0xF2, 0x43, BIP ₇
12.0-40	0xD6, 0x9B, 0x18, BIP ₃ , 0x29, 0x64, 0xE7, BIP ₇		
13.0-10	0x43, 0xBB, 0x33, BIP ₃ , 0xBC, 0x44, 0xCC, BIP ₇	13.1	0xD2, 0xE9, 0x97, BIP ₃ , 0x2D, 0x16, 0x68, BIP ₇
13.0-40	0x4E, 0xBA, 0x03, BIP ₃ , 0xB1, 0x45, 0xFC, BIP ₇		
14.0-10	0x2D, 0x17, 0x04, BIP ₃ , 0xD2, 0xE8, 0xFB, BIP ₇	14.1	0x3A, 0x83, 0x31, BIP ₃ , 0xC5, 0x7C, 0xCE, BIP ₇
14.0-40	0xC6, 0x57, 0x48, BIP ₃ , 0x39, 0xA8, 0xB7, BIP ₇		
15.0-10	0xED, 0xCD, 0x5E, BIP ₃ , 0x12, 0x32, 0xA1, BIP ₇	15.1	0x65, 0xB2, 0x32, BIP ₃ , 0x9A, 0x4D, 0xCD, BIP ₇
15.0-40	0xEF, 0x22, 0x4A, BIP ₃ , 0x10, 0xDD, 0xB5, BIP ₇		
16.0-10	0x2B, 0x1A, 0x05, BIP ₃ , 0xD4, 0xE5, 0xFA, BIP ₇	16.1	0x22, 0x0C, 0xDC, BIP ₃ , 0xDD, 0xF3, 0x23, BIP ₇
16.0-40	0x4B, 0x5C, 0x41, BIP ₃ , 0xB4, 0xA3, 0xBE, BIP ₇		
17.0-10	0x47, 0xAE, 0x79, BIP ₃ , 0xB8, 0x51, 0x86, BIP ₇	17.1	0xB8, 0xA3, 0xD6, BIP ₃ , 0x47, 0x5C, 0x29, BIP ₇
17.0-40	0x69, 0x19, 0xAE, BIP ₃ , 0x96, 0xE6, 0x51, BIP ₇		
18.0-10	0xC3, 0xF7, 0x82, BIP ₃ , 0x3C, 0x08, 0x7D, BIP ₇	18.1	0xCD, 0x0A, 0x6F, BIP ₃ , 0x32, 0xF5, 0x90, BIP ₇
18.0-40	0x86, 0x8D, 0xC5, BIP ₃ , 0x79, 0x72, 0x3A, BIP ₇		
19.0-10	0x21, 0xCE, 0xB9, BIP ₃ , 0xDE, 0x31, 0x46, BIP ₇	19.1	0x8E, 0x81, 0x3B, BIP ₃ , 0x71, 0x7E, 0xC4, BIP ₇
19.0-40	0x09, 0xF2, 0xA6, BIP ₃ , 0xF6, 0x0D, 0x59, BIP ₇		

7 Detailed Block Diagrams

The detailed processing to implement the functionality described in clause 6 is provided in this clause.

7.1 MLG mux

The 4x25G MLG mux multiplexes from zero to two 40GBASE-R signals and from zero to ten 10GBASE-R signals, with a maximum combined bandwidth of 103.125 Gb/s, into a format consistent with the IEEE Std 802.3-2012 clause 83 PMA. The signal is comprised of 20 MLG lanes of 5.1625 Gb/s. Two groups of 8 MLG lanes can be configured to carry four 10GBASE-Rs or a single 40GBASE-R as described in section 6, while the other 4 MLG lanes can carry two 10GBASE-Rs.

The 8x25G MLG mux multiplexes from zero to five 40GBASE-R signals and from zero to twenty 10GBASE-R signals, with a maximum combined bandwidth of 206.25 Gb/s, into a format consistent with a double-width IEEE Std 802.3-2012 clause 83 PMA. This signal is comprised of 40 MLG lanes of 5.1625 Gb/s. Each of five groups of 8 MLG lanes can be configured to carry four 10GBASE-Rs or a single 40GBASE-R.

7.1.1 MLG mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 7.

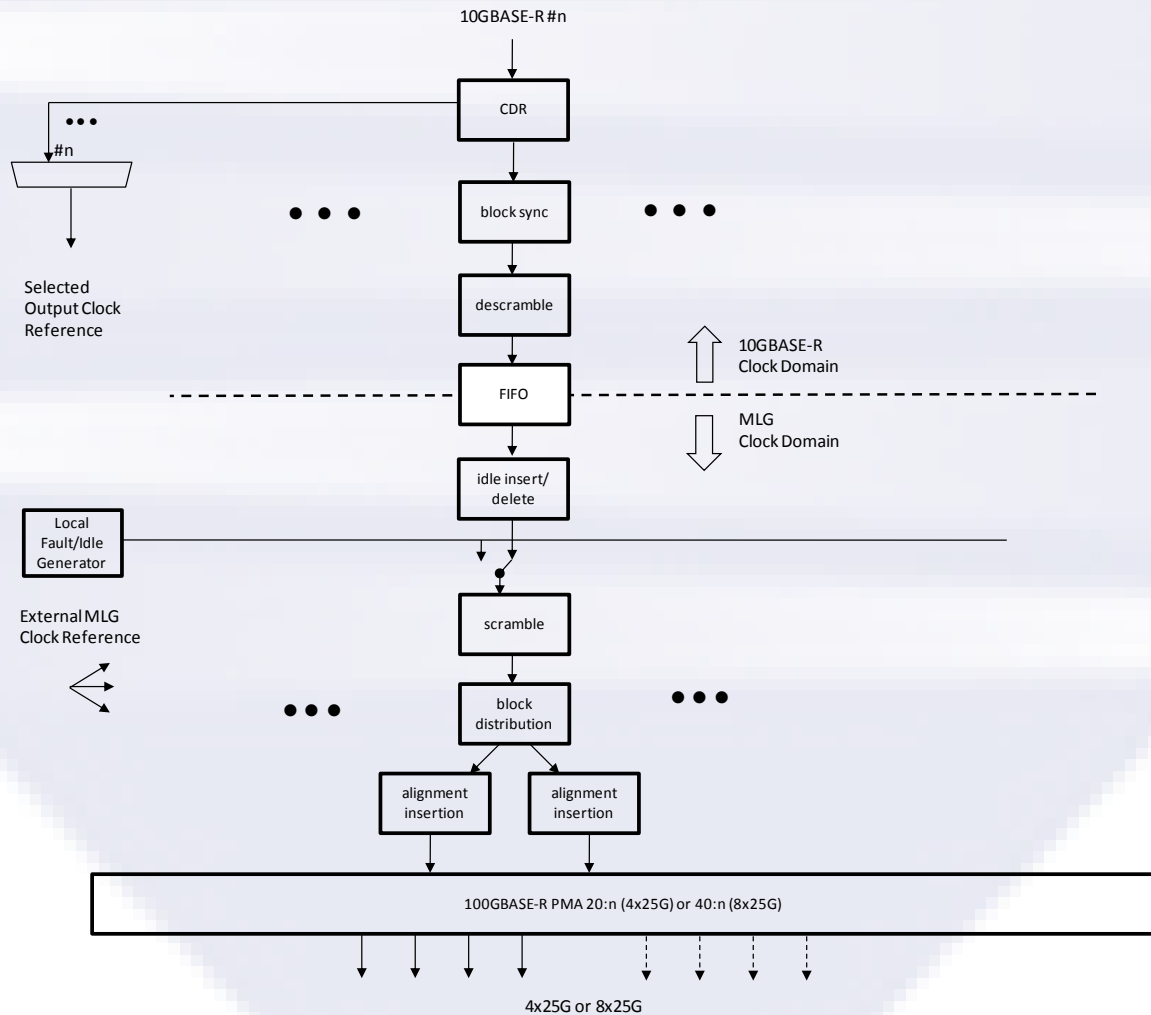


Figure 7: MLG mux 10GBASE-R mapping Block Diagram

7.1.1.1 Clock and Data Recovery (CDR)

Each of the ten input 10GBASE-R signals to the MLG mux will undergo clock and data recovery. This function also provides input to the signal detect function which is used to indicate failure of the input signal to the management interface and downstream functions. One of the recovered 10GBASE-R clocks can also be selected to be output as a 10G timing reference, and used (if required) to drive a network timing architecture such as SyncE.

7.1.1.2 Block Sync

Once clock and data has been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-12 of IEEE Std 802.3-2012. When block lock=false per this state diagram, this is considered a failure of the input signal equivalent to not being able to recover clock and data as part of the signal detect function.

7.1.1.3 Descramble

The non-sync header bits of the 10GBASE-R signal are descrambled using the process as described in clause 49.2.10 of IEEE Std 802.3-2012.

7.1.1.4 Idle Insert/Delete

Idles are inserted or deleted as necessary in order that the bitstream, after MLG lane marker insertion, will match a common MLG clock reference. Idle insertion and deletion shall comply with the rules defined in IEEE Std 802.3-2012 clause 48.2.4.2.3.

An external MLG clock reference is provided as an input to the MLG mux block. The external MLG clock reference may be sourced from either a local free-running oscillator (± 100 ppm), or from a system level BITS/SyncE timing architecture. The exact implementation of the external MLG clock reference is a system architecture decision, and outside the scope of this IA.

7.1.1.5 Local Fault (LF) Insertion

If the incoming 10GBASE-R signal is disabled or has failed (see the signal detect function as part of CDR in 7.1.1.1), the signal is replaced with an Ethernet Local Fault sequence ordered set. This is a 66-bit block with the contents: Sync Header=10; Control block type=0x55; O1=0x0; O4=0x0; and the local fault encoding indicated in Table 46-5 of IEEE Std 802.3-2012 in D₁, D₂, D₃ and D₅, D₆, D₇. There should be sufficient buffer to allow replacement of a failed incoming signal with Local Fault without interruption to the traffic in the other 10GBASE-R or 40GBASE-R signals: for example, if there is not a full 66-bit block available for transmission, the Local Fault control block will be transmitted instead, maintaining the 66-bit block alignment with the data it replaces.

7.1.1.6 Scrambled Idle Test Pattern Generation

The MLG mux may optionally generate for each 10G or 40G lane a scrambled idle test pattern. This test pattern can traverse a gearbox, and can be checked by the far-end MLG demux, or looped back at the MLG demux or Ethernet Rx and checked at the near end MLG demux. An MLG implementation that includes this capability shall perform it as described in this clause

When a scrambled idle pattern is enabled for a given 10G signal, it shall be generated at the input to the scrambler (see 7.1.1.7). The input to the scrambler is a control block (block type=0x1E) with all idles as defined in IEEE Std 802.3-2012 Figure 82-5. When switching between scrambled idle test pattern mode and normal operation, the 66-bit block alignment shall be maintained.

7.1.1.7 Scramble

After idle insertion/deletion and Local Fault insertion if necessary, the resulting stream is re-scrambled. This is done according to clause 49.2.6 of IEEE Std 802.3-2012.

7.1.1.8 Block Distribution

The 66-bit blocks of each 10GBASE-R signal are distributed to two MLG lanes, alternating 66-bit blocks on each MLG lane. This is a similar process to that described in clause 82.2.6 of IEEE Std 802.3-2012, but distributing to two MLG lanes rather than four or ten PCS lanes.

7.1.1.9 Alignment Marker Insertion

In order to support deskew and reordering of the MLG lanes into the constituent 10GBASE-R signals at the MLG demux, alignment markers are added periodically to each MLG lane. Each alignment marker is a special 66-bit block. The alignment markers shall be inserted after every 16383 data blocks on each MLG lane in the same way (at the same time on all twenty MLG lanes) as the insertion of PCS lane alignment markers described in clause 82.2.7 of IEEE Std 802.3-2012. The alignment marker values are given in Table 1 for a 4x25G gearbox or Table 2 for an 8x25G gearbox. These values are distinct from the PCS lane alignment markers used for 40GBASE-R and 100GBASE-R signals, so there is no possibility of mistaking ten 10GBASE-R signals for a 100GBASE-R or vice-versa. The BIP₃ and BIP₇ fields of each MLG lane alignment marker are calculated over all of the previous bits on the given MLG lane, from and including the previous MLG lane alignment marker but not the current MLG lane alignment marker per the procedures described for PCS lanes in clause 82.2.8 of IEEE Std 802.3-2012.

7.1.1.10 100GBASE-R PMA(20:n) or PMA(40:n)

The processes described in clauses 7.1.1.1 through 7.1.1.9 will produce twenty or forty MLG lanes, each with a bit rate of 5.15625 Gb/s \pm 100ppm (all MLG lanes locked to the common clock source). These MLG lanes can now be combined as if they were 100GBASE-R PCS lanes using the 100GBASE-R PMA 20:n into any physical lane configuration used to support 100GBASE-R, or using a PMA 40:n onto a double-width interface. The most typical value of “n” is expected to be 4 for the 4x25G MLG application, or 8 for the 8x25G MLG application. A PMA(20:4) will produce a 4-lane by 25.78125 Gb/s interface, where each of these physical lanes carries five of the twenty MLG lanes, bit multiplexed. A PMA (40:4) will produce an 8-lane by 25.78125 Gb/s interface, where each of these physical lanes carries five of the forty MLG lanes, bit multiplexed. Note that an implementation may support generation of any of the optional test patterns specified for a 100GBASE-R PMA across the MLG link: see IEEE Std 802.3-2012 clause 83.5.10. If supported, the corresponding MDIO control variables must also be supported.

7.1.2 MLG mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 8. These may be eight of the twenty MLG lanes of a 4x25G MLG, or eight of the forty MLG lanes of an 8x25G MLG.

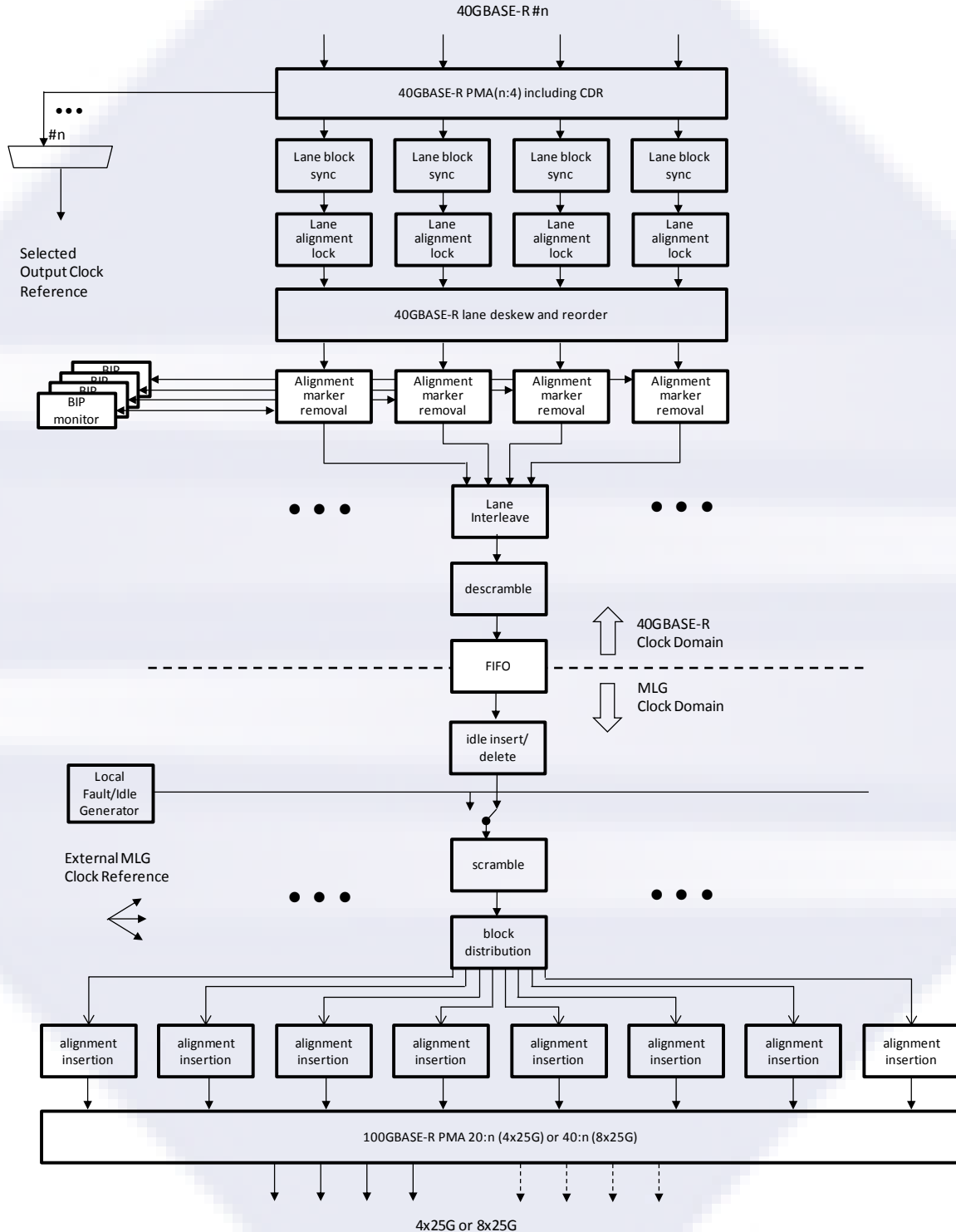


Figure 8: MLG mux 40GBASE-R mapping Block Diagram

7.1.2.1 40GBASE-R PMA (n:4) including CDR

This function is as per IEEE Std 802.3-2012 clause 83. For 40GBASE-FR, n is 1. For other 40GBASE-R PMDs, n is 4. This includes recovery of clock and data, which may be selected as the output clock reference based on provisioning.

7.1.2.2 40GBASE-R PCS Lane block sync

66-bit block lock is obtained on each PCS lane of the 40GBASE-R signal using the state diagram in IEEE Std 802.3-2012 Figure 82-10.

7.1.2.3 40GBASE-R PCS Lane alignment lock

Alignment marker lock is obtained on each PCS lane of the 40GBASE-R signal using the state diagram in IEEE Std 802.3-2012 Figure 82-11.

7.1.2.4 40GBASE-R lane deskew and reorder

The PCS lanes of the 40GBASE-R signal are deskewed and reordered according to the procedures described in IEEE Std 802.3-2012 clauses 82.2.12 and 82.2.13.

7.1.2.5 Alignment marker removal, BIP monitor, and PCS Lane Interleave

The PCS lane alignment markers are removed from the 40GBASE-R signal, and the PCS lanes are interleaved as described in IEEE Std 802.3-2012 clause 82.2.14. The expected BIP and received BIP values from each received alignment marker are compared and error counters are updated as described in 82.2.14.

7.1.2.6 Descramble

The 40GBASE-R signal is descrambled according to IEEE Std 802.3-2012 clause 82.2.15, which is identical to the 10GBASE-R descrambler described in clause 49.2.10.

7.1.2.7 Idle insert/delete

Idles are inserted or deleted as necessary in order that the bitstream, after PCS lane marker insertion, will match a common MLG clock reference. Idle insertion and deletion shall comply with the rules defined in IEEE Std 802.3-2012 clause 82.2.3.6.

An external MLG clock reference is provided as an input to the MLG mux block. The external MLG clock reference may be sourced from either a local free-running oscillator (± 100 ppm), or from a system level BITS/SyncE timing architecture. The exact implementation of the external MLG clock reference is a system architecture decision, and outside the scope of this IA.

7.1.2.8 Local Fault (LF) Insertion

If the incoming 40GBASE-R signal is disabled or has failed (see the signal detect function as part of CDR in the PMA described in 7.1.2.1), the signal will be replaced with an Ethernet Local Fault sequence ordered set. This is a 66-bit block with the contents: Sync Header=10; Control block type=0x4B; O1=0x0; and the local fault encoding indicated in Table 81-5 of IEEE Std 802.3-2012 in D₁, D₂, D₃, and zeros in Z₄, Z₅, Z₆, Z₇. There should be sufficient buffer to allow replacement of a failed incoming signal with Local Fault without interruption to the traffic in the other 10GBASE-R or 40GBASE-R signals: for example, if there is not a full 66-bit block available for transmission, the Local

Fault control block will be transmitted instead, maintaining the 66-bit block alignment with the data it replaces.

7.1.2.9 Scrambled Idle Test Pattern Generation

The MLG mux may optionally generate for each 40GBASE-R a scrambled idle test pattern. This test pattern can traverse a gearbox, and can be checked by the far-end MLG demux, or looped back at the MLG demux or Ethernet Rx and checked at the near end MLG demux. An MLG implementation that includes this capability shall perform it as described in this clause.

When a scrambled idle pattern is enabled for a given 40G signal, it shall be generated at the input to the scrambler (see 7.1.1.7). The input to the scrambler is a control block (block type=0x1E) with all idles as defined in IEEE Std 802.3-2012 Figure 82-5. When switching between scrambled idle test pattern mode and normal operation, the 66-bit block alignment shall be maintained.

7.1.2.10 Block Distribution

The 66-bit blocks of each 40GBASE-R signal are distributed to eight MLG lanes, distributing the 66-bit blocks round-robin on each MLG lane. This is a similar process to that described in clause 82.2.6 of IEEE Std 802.3-2012, but distributing to eight MLG lanes rather than four or ten PCS lanes.

7.1.2.11 Alignment Marker Insertion

In order to support deskew and reordering of the MLG lanes into the constituent 40GBASE-R signals at the MLG demux, alignment markers are added periodically to each MLG lane. Each alignment marker is a special 66-bit block. The alignment markers shall be inserted after every 16383 data blocks on each MLG lane in the same way (at the same time on all twenty MLG lanes) as the insertion of PCS lane alignment markers described in clause 82.2.7 of IEEE Std 802.3-2012. The alignment marker values are given in Table 1 for a 4x25G gearbox or Table 2 for an 8x25G gearbox. These values are distinct from the PCS lane alignment markers used for 40GBASE-R and 100GBASE-R signals, so there is no possibility of mistaking ten 10GBASE-R signals as a 100GBASE-R or vice-versa. The BIP₃ and BIP₇ fields of each MLG lane alignment marker are calculated over all of the previous bits on the given MLG lane, from and including the previous MLG lane alignment marker but not the current lane alignment marker per the procedures described for PCS lanes in clause 82.2.8 of IEEE Std 802.3-2012.

7.1.2.12 100GBASE-R PMA(20:n) or PMA(40:n)

See 7.1.1.10. At this position in the stack, there is no difference based on whether 10GBASE-R or 40GBASE-R is carried on any given MLG lane.

7.1.3 MLG mux Management

Control Variables:

- MLG_mux_Enable – Enables or disables the MLG function output
- MLG_mux_40G_select_0, MLG_mux_40G_select_4, MLG_mux_40G_select_8, MLG_mux_40G_select_12, MLG_mux_40G_select_16 – Configures MLG lanes 0.x-3.x, 4.x-7.x, 8.x-11.x, 12.x-15.x, or 16.x-19.x, respectively to carry a single

Multi-link Gearbox 2.0 Implementation Agreement

40GBASE-R rather than four 10GBASE-Rs. Only MLG_mux_40G_select_0 and MLG_mux_40G_select_4 are valid for the 4x25G application, while all are valid for the 8x25G application.

- MLG_mux_10G_Enable_0 through MLG_mux_10G_Enable_19 – Enables or disables each of the 10G lanes. Any disabled 10G lane will have Local Fault inserted as described in 7.1.1.5 for 10GBASE-R or 7.1.2.8 for 40GBASE-R. MLG_mux_10G_Enable_10 through MLG_mux_10G_Enable_19 are not defined for the 4x25G application. When a group of MLG lanes is configured to carry 40GBASE-R, the first of the control variables (MLG_mux_10G_Enable_0, 4, 8, 12, 16) is used to enable the respective 40GBASE-R.
- MLG_mux_10G_Output_Timing_Reference – Selects which of the 10GBASE-R or 40GBASE-R signals provides the 10G timing output reference.
- MLG_mux_scrambled_idle_enable_0 through MLG_mux_scrambled_idle_enable_19 – if implemented, enables or disables the scrambled idle test pattern generated on a given 10G or 40G lane. For MLG lanes configured as 40G, the first 10G lane number is used, i.e., MLG_mux_scrambled_idle_enable_0, 4, 8, 12, 16.

Status Variables:

- MLG_mux_4x25G_ability – indicates whether the MLG mux supports the 4x25G application
- MLG_mux_8x25G_ability – indicates whether the MLG mux supports the 8x25G application
- MLG_mux_40G_ability – for a 4x25G MLG mux, indicates whether the mux is capable of carrying both 40GBASE-R and 10GBASE-R. If this value is FALSE, the MLG mux is only capable of MLG 1.0 operation
- Signal_Detect_0 through Signal_Detect_19 – Indicates whether a 10GBASE-R signal was successfully recovered through CDR and the 66-bit block lock process on each of the 10G input lanes. Signal_Detect_10 through Signal_Detect_19 are not defined for the 4x25G mux application. When a group of MLG lanes is configured to carry 40GBASE-R, the first of the status variables (Signal_Detect_0, 4, 8, 12, 16) is used to signal the status of the respective 40GBASE-R.
- MLG_mux_scrambled_idle_ability – Indicates whether this implementation of MLG has the ability to generate the 10G scrambled idle test pattern on each 10G lane.

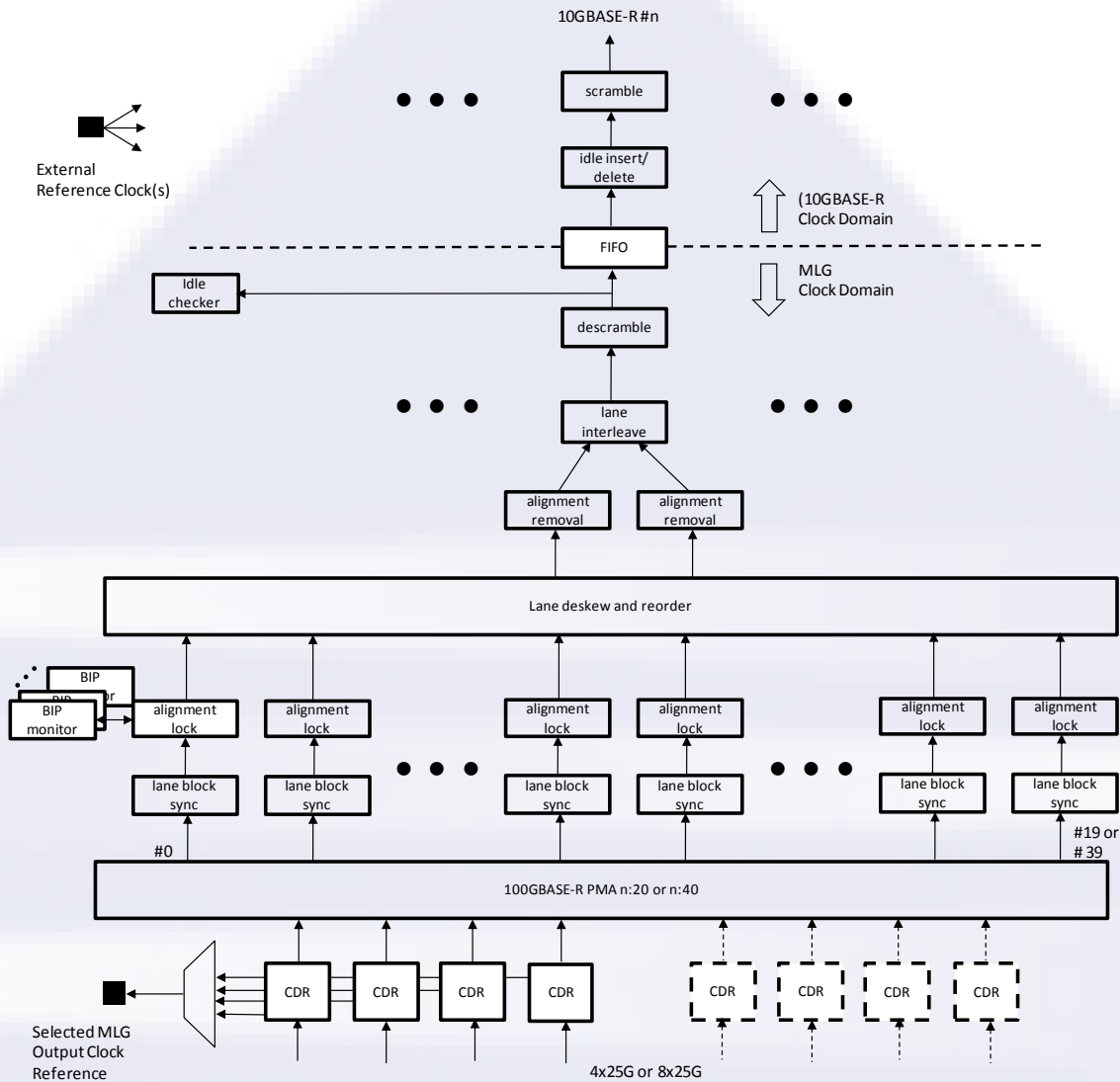
7.2 MLG demux

The MLG demux receives twenty (for the 4x25G application) or forty (for the 8x25G application) MLG lanes. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

7.2.1 MLG demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 9.

Multi-link Gearbox 2.0 Implementation Agreement


Figure 9: MLG demux 10GBASE-R demapping Block Diagram

7.2.1.1 100GBASE-R PMA(n:20 or n:40)

The input to the MLG demux will come from a physical interface that is similar to that of 100GBASE-R for the 4x25G application or a double-width 100GBASE-R for the 8x25G application. For the most common MLG configuration, this is expected to be a 4-lane by 25.78125 Gb/s interface or an 8-lane by 25.78125 Gb/s interface, where each of these physical lanes carries five of the twenty MLG lanes, bit multiplexed, but other physical lane counts which are divisors of 20 (or 40) are possible. The PMA(n:20) or PMA(n:40) will demultiplex the MLG lanes as if they were PCS lanes into twenty or forty individual bit streams of 5.15625 Gb/s ± 100 ppm. Note that an implementation may support detection of any of the optional test patterns specified for a 100GBASE-R PMA across the MLG link: see IEEE Std 802.3-2012 clause 83.5.10. If supported, the corresponding MDIO control and status variables must also be supported.

7.2.1.2 MLG Lane block sync

66-bit block sync is obtained on each of the MLG lanes in the same manner as 100GBASE-R PCS lanes using the state diagram in Figure 82-10 of IEEE Std 802.3-2012.

7.2.1.3 MLG Lane Alignment lock

Alignment marker lock is obtained on each MLG lane in the same manner as 100GBASE-R PCS lanes using the state diagram in Figure 82-11 of IEEE Std 802.3-2012, using the expected MLG lane alignment marker values from Table 1 for the 4x25G application or Table 2 for the 8x25G application.

7.2.1.4 BIP monitor

The expected BIP value is compared to the received BIP value in each received alignment marker and BIP error counters for each MLG lane are incremented as described for PCS lanes in clause 82.2.14 of IEEE Std 802.3-2012.

7.2.1.5 MLG Lane reorder

Once alignment marker lock has been obtained on the two MLG lanes that comprise a 10GBASE-R signal, those MLG lanes are deskewed, reordered and the 66-bit blocks interleaved to reconstitute the original 10GBASE-R signal.

7.2.1.6 MLG Lane Alignment marker removal

The alignment markers in each MLG lane are removed in the same way as for 100GBASE-R PCS lane alignment markers described in clause 82.2.14 of IEEE Std 802.3-2012.

7.2.1.7 MLG Lane Interleave

The pairs of MLG lanes comprising each 10GBASE-R signal being carried over the MLG link, are 66-bit block interleaved to reconstitute the original 10GBASE-R signals.

7.2.1.8 Descramble

The non-sync header bits of the 10GBASE-R signal are descrambled using the process as described in clause 49.2.10 of IEEE Std 802.3-2012.

7.2.1.9 Scrambled Idle Test Pattern Checker

The MLG demux may optionally implement a scrambled idle test pattern checker for each 10GBASE-R signal. If implemented, it shall be as described in this clause.

When `align_status` is true and the scrambled idle receive test-pattern mode is active, the scrambled idle test-pattern checker observes the sync header and the output from the descrambler. When the sync header and the output of the descrambler is the all idle pattern, a match is detected. When operating in scrambled idle test pattern, the test-pattern error counter counts blocks with a mismatch. Any mismatch indicates an error and shall increment the test-pattern error counter. Due to the error multiplication characteristics of the descrambler, the incoming bit error ratio can be estimated by dividing the 66-bit block error ratio by a factor of 124.

7.2.1.10 Idle Insert/Delete

Idles are inserted or deleted as necessary to map the rate of each 10GBASE-R signal to the required 10G output clock rate. Idle insertion and deletion shall comply with the rules defined in IEEE Std 802.3-2012 clause 48.2.4.2.3.

Note, each 10GBASE-R signal can in theory be mapped to an independent 10G output clock domain, driven by an external 10G clock reference. The external 10G reference clock(s) may be sourced from either a local, free-running oscillator (± 100 ppm), or from a system/network level BITS/SyncE timing architecture. The exact implementation of the external 10G reference clock(s) for a given application is a system architecture decision, and outside the scope of this IA.

However for reference it is noted that a BITS/SyncE timing architecture can be used to allow a timing reference from one of the 10GBASE-R or 40GBASE-R input signals to be carried across a MLG link (and used to clock one or more of the 10GBASE-R and 40GBASE-R output signals).

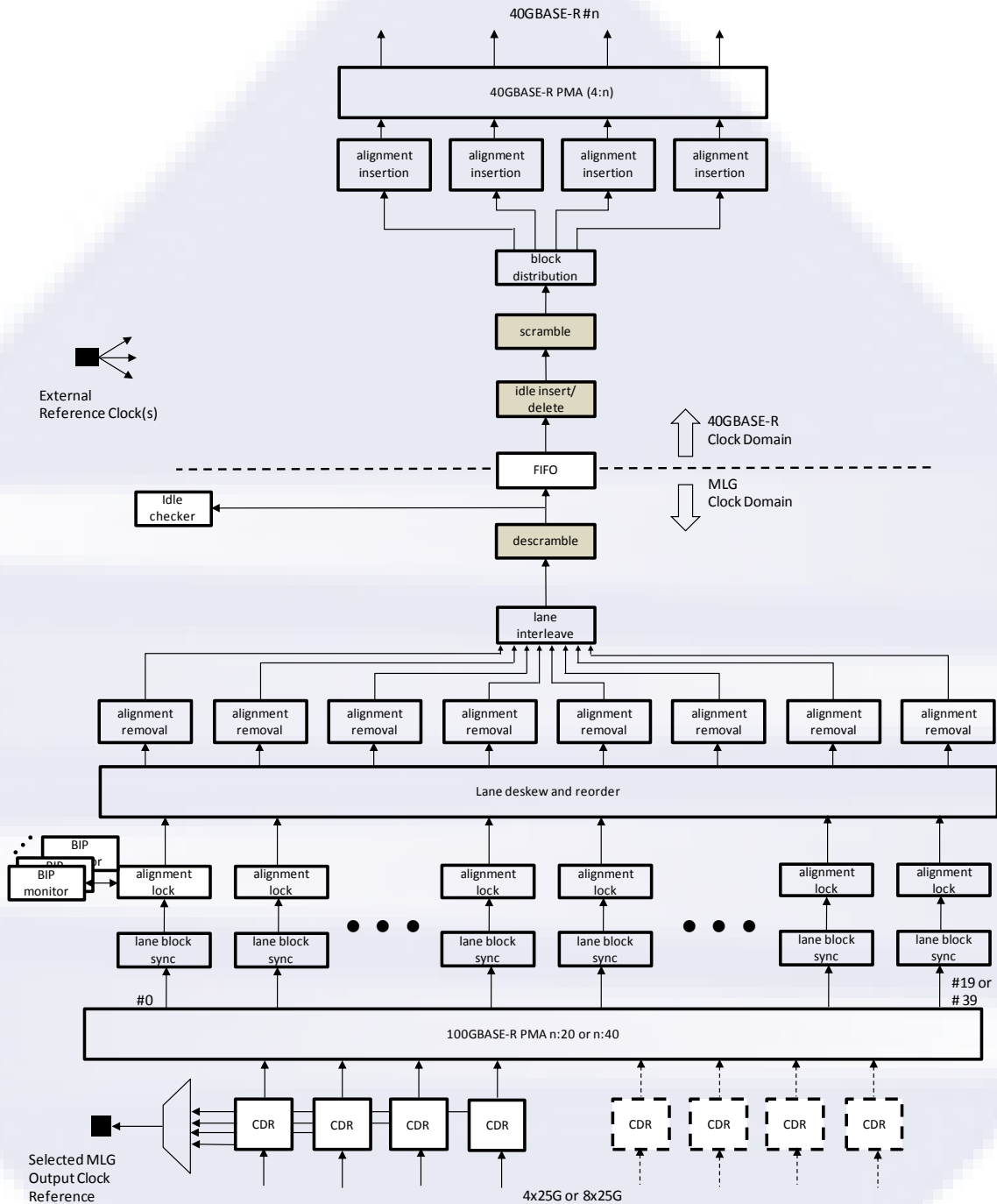
7.2.1.11 Scramble

After idle insertion/deletion, each 10GBASE-R signal is re-scrambled according to clause 49.2.6 of IEEE Std 802.3-2012.

7.2.2 MLG demux for 40GBASE-R

The process for demapping a 40GBASE-R from eight MLG lanes is shown in Figure 10.

Multi-link Gearbox 2.0 Implementation Agreement


Figure 10: MLG demux 40GBASE-R demapping Block Diagram

7.2.2.1 100GBASE-R PMA(n:20 or n:40)

The input to the MLG mux will come from a physical interface that is similar to that of 100GBASE-R for the 4x25G application or a double-width 100GBASE-R for the 8x25G application. For the most common MLG configuration, this is expected to be a 4-lane by 25.78125 Gb/s interface or an 8-lane by 25.78125 Gb/s interface, where each of these physical lanes carries five of the twenty MLG lanes, bit multiplexed, but other physical

Multi-link Gearbox 2.0 Implementation Agreement

lane counts which are divisors of 20 (or 40) are possible. The PMA(n:20) or PMA(n:40) will demultiplex the MLG lanes as if they were PCS lanes into twenty or forty individual bit streams of 5.15625 Gb/s \pm 100ppm. Note that an implementation may support detection of any of the optional test patterns specified for a 100GBASE-R PMA across the MLG link: see IEEE Std 802.3-2012 clause 83.5.10. If supported, the corresponding MDIO control and status variables must also be supported.

7.2.2.2 MLG Lane block sync

66-bit block sync is obtained on each of the MLG lanes in the same manner as 100GBASE-R PCS lanes using the state diagram in Figure 82-10 of IEEE Std 802.3-2012.

7.2.2.3 MLG Lane Alignment lock

Alignment marker lock is obtained on each MLG lane in the same manner as 100GBASE-R PCS lanes using the state diagram in Figure 82-11 of IEEE Std 802.3-2012, using the expected MLG lane alignment marker values from Table 1 for the 4x25G application or Table 2 for the 8x25G application.

7.2.2.4 BIP monitor

The expected BIP value is compared to the received BIP value in each received alignment marker and BIP error counters for each MLG lane are incremented as described for PCS lanes in clause 82.2.14 of IEEE Std 802.3-2012.

7.2.2.5 MLG Lane reorder

Once alignment marker lock has been obtained on the eight MLG lanes that comprise a 40GBASE-R signal, those lanes are deskewed, reordered and the 66-bit blocks interleaved to reconstitute the original 40GBASE-R signal.

7.2.2.6 MLG Lane Alignment marker removal

The alignment markers in each MLG lane are removed in the same way as for 100GBASE-R PCS lane alignment markers described in clause 82.2.14 of IEEE Std 802.3-2012.

7.2.2.7 MLG Lane Interleave

The four MLG lanes comprising each 40GBASE-R signal being carried over the MLG link, are 66-bit block interleaved to reconstitute the original 40GBASE-R signals.

7.2.2.8 Descramble

Note that there are three shaded boxes in Figure 10: descramble, idle insert/delete, and scramble. For an implementation that does not require that the demapped 40GBASE-R be in a different clock domain from the MLG clock domain, these three steps can be omitted, as the ratio of bytes in the total bit stream from the MLG lane markers removed is exactly the same as that for the 40GBASE-R PCS lane markers that are inserted by the process. If the optional scrambled idle test pattern checker is implemented, the signal will also need to be descrambled to perform the check, but the scrambled signal can be passed directly through the FIFO to the block distribution described in 7.2.2.11 if a different clock domain is not needed for the 40GBASE-R.

When required, the non-sync header bits of the 10GBASE-R signal are descrambled using the process as described in clause 49.2.10 of IEEE Std 802.3-2012.

7.2.2.9 Idle insert/delete

Idles are inserted or deleted as needed to map the rate of each 40GBASE-R signal to the required 40G output clock rate. Idle insertion and deletion shall comply with the rules defined in IEEE Std 802.3-2012 clause 82.2.3.6.

Note, each 40GBASE-R signal can in theory be mapped to an independent 40G output clock domain, driven by an external 40G clock reference. The external 40G reference clock(s) may be sourced from either a local, free-running oscillator (± 100 ppm), or from a system/network level BITS/SyncE timing architecture. The exact implementation of the external 40G reference clock(s) for a given application is a system architecture decision, and outside the scope of this IA.

However, for reference it is noted that a BITS/SyncE timing architecture can be used to allow a timing reference from one of the 10GBASE-R or 40GBASE-R input signals to be carried across a MLG link (and used to clock one or more of the 10GBASE-R and 40GBASE-R output signals).

7.2.2.10 Scramble

If the 40GBASE-R signal has been descrambled, after idle insertion/deletion, each 40GBASE-R signal is re-scrambled according to clause 82.2.5, which reuses the clause 49.2.6 scrambler in IEEE Std 802.3-2012.

7.2.2.11 Block Distribution

The 66-bit blocks of the 40GBASE-R signal are distributed round-robin to four PCS lanes as described in clause 82.2.6 of IEEE Std 802.3-2012.

7.2.2.12 Alignment Insertion

The 40GBASE-R PCS alignment markers are inserted as described in clause 82.2.7 of IEEE Std 802.3-2012.

7.2.2.13 40GBASE-R PMA (4:n)

Once the PCS lane formatted 40GBASE-R signal is created, it can be carried over a standard 40GBASE-R PMA as described in IEEE Std 802.3-2012 clause 83. The physical lane count for the output signal depends on the actual PMD type: e.g., a PMA(4:1) is used for a 40GBASE-FR PMD and PMA(4:4) for other PMD types.

7.2.3 MLG demux Management

Control variables:

- MLG_demux_40G_select_0, MLG_demux_40G_select_4, MLG_demux_40G_select_8, MLG_demux_40G_select_12, MLG_demux_40G_select_16 – Configures MLG lanes 0.x-3.x, 4.x-7.x, 8.x-11.x, 12.x-15.x, or 16.x-19.x, respectively to demap a single 40GBASE-R from those MLG lanes rather than four 10GBASE-Rs. Only MLG_demux_40G_select_0 and MLG-demux_40G_select_4 are valid for the 4x25G application, while all are valid for the 8x25G application.

Multi-link Gearbox 2.0 Implementation Agreement

- MLG_demux_10G_Enable_0 through MLG_demux_10G_Enable_19 – Enables or disables each of the 10G output lanes. Any disabled 10G lane will have local fault inserted for the appropriate rate. MLG_demux_10G_Enable_10 through MLG_demux_10G_Enable_19 are not valid for the 4x25G application. When a group of MLG lanes is configured to carry 40GBASE_R, the first of the 10G control variables (MLG_demux_10G_Enable_0, 4, 8, 12, 16) is used to enable the respective 40GBASE-R.
- MLG_demux_scrambled_idle_enable_0 through MLG_demux_scrambled_idle_enable_19 – if implemented, enables or disables the scrambled idle test pattern checker for the indicated lane. MLG_demux_scrambled_idle_enable_10 through MLG_demux_scrambled_idle_enable_19 are not valid for the 4x25G application. For MLG lanes configured for 40GBASE-R, the first 10G lane number will be used to enable or disable the respective 40G scrambled idle checker, i.e., MLG_demux_scrambled_idle_enable_0, 4, 8, 12, 16.

Status variables:

Note that MLG lanes may be received by the MLG demux on different physical lanes than those on which they were originally transmitted by the MLG mux due to skew between MLG lanes and multiplexing by the intervening 100GBASE-R PMA(s). Prior to the MLG lane reorder block, MLG lanes are known simply by the position 0..19 (for the 4x25G application) or 0..39 (for the 8x25G application) on which they were received. After alignment lock on all MLG lanes and MLG lane reorder, MLG lanes are known by which half of which 10G signal they represent (0.0, 0.1, 1.0, ..., 19.1). Four-lane groups may also be used to carry 40GBASE-R, specifically 0.x-3.x, 4.x-7.x, 8.x-11.x, 12.x-15.x, 16.x-19.x.

- MLG_demux_4x25G_ability – indicates whether the MLG demux supports the 4x25G application
- MLG_demux_8x25G_ability – indicates whether the MLG demux supports the 8x25G application
- MLG_demux_40G_ability – for a 4x25G MLG demux, indicates whether the demux is capable of carrying both 40GBASE-R and 10GBASE-R. If this value is FALSE, the MLG demux is only capable of MLG 1.0 operation.
- MLG_demux_Link_Status – indicates whether the physical input lanes of the MLG demux are being received at the PMA(n:20) or PMA(n:40). Depending on n, individual lane status variables may be available. Since n is implementation dependent, so are the status registers.
- block_lock_0 through block_lock_39 – indicate whether 66-bit block lock has been achieved on each of the MLG lanes. block_lock_20 through block_lock_39 are not valid for the 4x25G application.
- am_lock_0 through am_lock_39 – indicates whether alignment marker lock has been achieved on each of the MLG lanes. am_lock_20 through am_lock_39 are not valid for the 4x25G application.
- MLG_demux_lane_alignment_status – indicates whether all 20 (or 40) MLG lanes have achieved MLG lane alignment marker lock, that the 20 (or 40) distinct MLG lane markers are received, and inter-MLG lane skew permits the 10GBASE-R signals or 40GBASE-R signals to be reassembled.

- BIP_error_counter_0 through BIP_error_counter_39 – contains the count of BIP errors counted on each MLG lane. BIP_error_counter_20 through BIP_error_counter_39 are not valid for the 4x25G application.
- lane_0_mapping through lane_39_mapping – indicates which (logical) MLG lane is received in each (physical) MLG lane position. Note that the MLG lanes that may be received in a MLG lane position are numbered 0.0, 0.1, 1.0, ..., 19.1. lane_20_mapping through lane_39_mapping are not valid for the 4x25G application.
- MLG_demux_scrambled_idle_ability – indicates whether this implementation implements the optional scrambled idle pattern checker in the MLG demux
- MLG_demux_scrambled_idle_error_0 through MLG_demux_scrambled_idle_error_19 – When the test pattern checker is enabled, counts the pattern mismatches on the indicated 10GBASE-R or 40GBASE-R signal. The counter is reset to zero at the point that the test pattern checker is enabled. MLG_demux_scrambled_idle_error_10 through MLG_demux_scrambled_idle_error_19 are not valid for the 4x25G application. For lanes configured for 40GBASE-R, the first 10G counter position is used, i.e., MLG_demux_scrambled_idle_error_counter_0, 4, 8, 12, 16.

7.3 Generic MLG Management

Control Variables:

- MLG_10G_loopback_enable_0 through MLG_10G_loopback_enable_19: When enabled, loops back the 10GBASE-R signal recovered on the given 10G output interface at the MLG demux to the corresponding 10G input interface at the MLG mux. MLG_10G_loopback_enable_10 through MLG_10G_loopback_enable_19 are not valid for the 4x25G application. For lanes configured for 40G, the first 10G lane position enables a 40G loopback, i.e., MLG_10G_loopback_enable_0, 4, 8, 12, 16.

8 References

8.1 Normative references

[1] IEEE Std 802.3-2012 Standard for Ethernet.

9 Appendix A (Informative): Suggested MPO Receptacle Physical Lane Assignments for various MLG applications

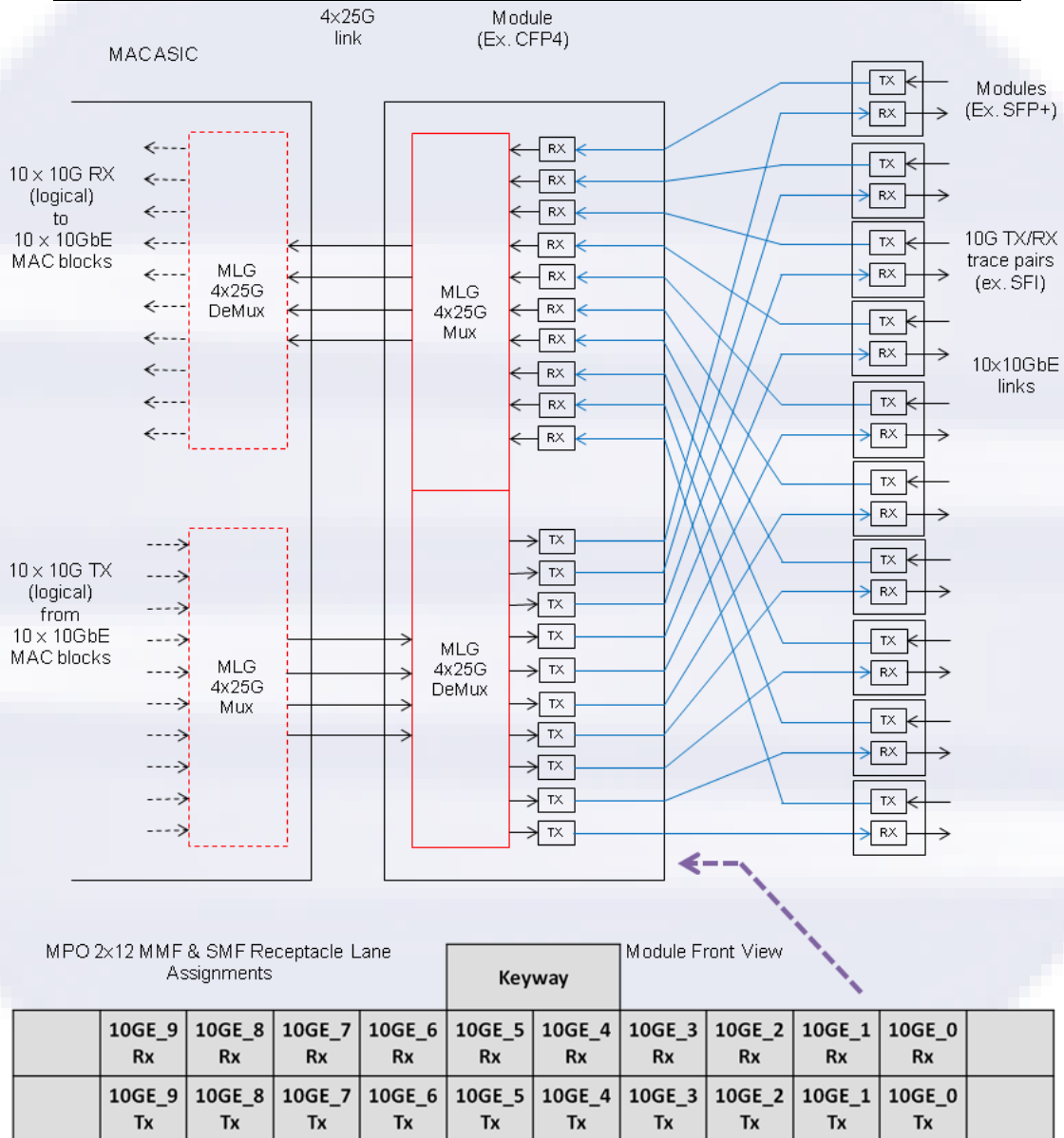


Figure 11: MPO Receptacle Physical Lane Assignments for 4x25G MLG 10GBASE-SR or 10GBASE-LR Optical Port Expander

Multi-link Gearbox 2.0 Implementation Agreement

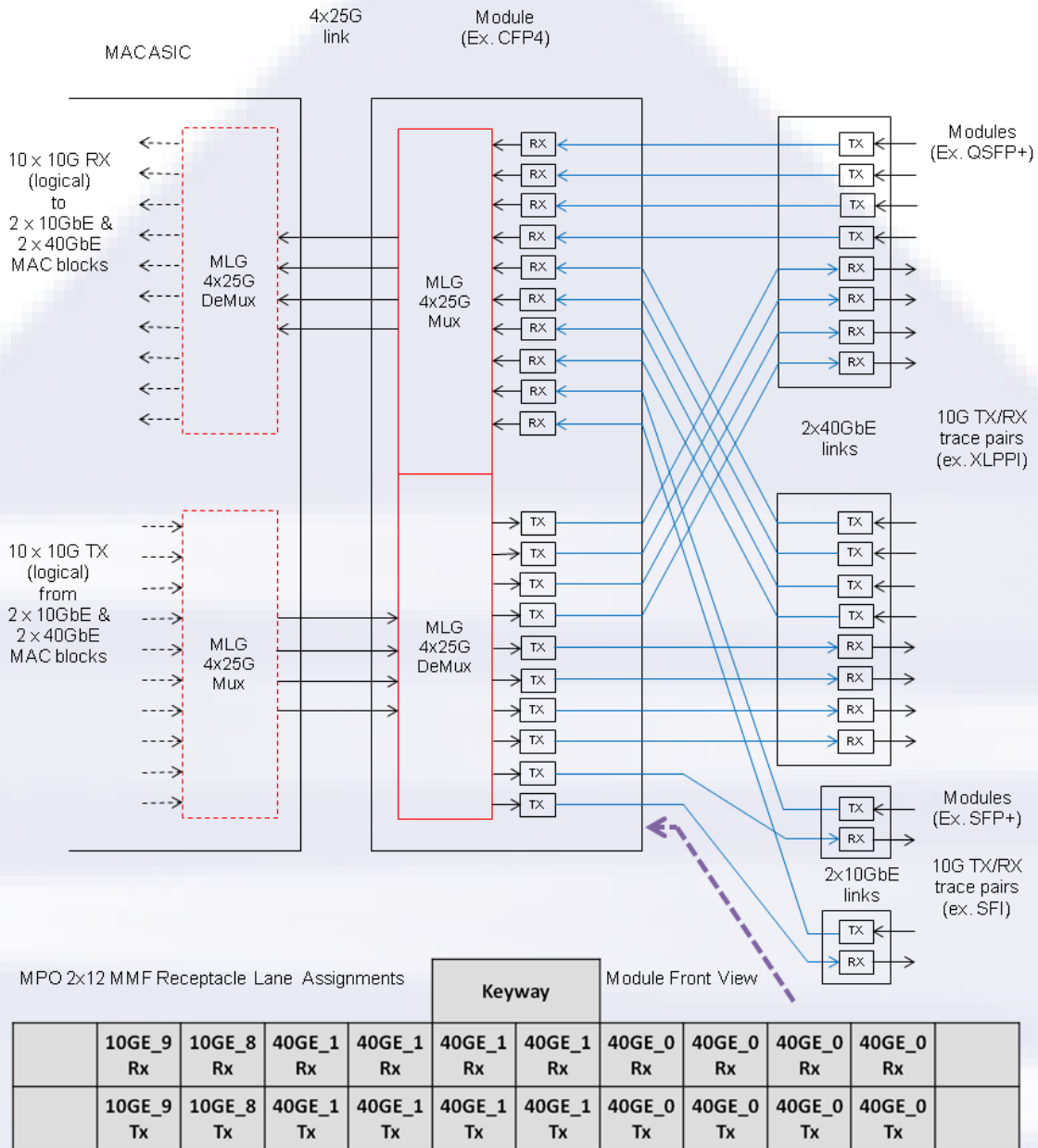
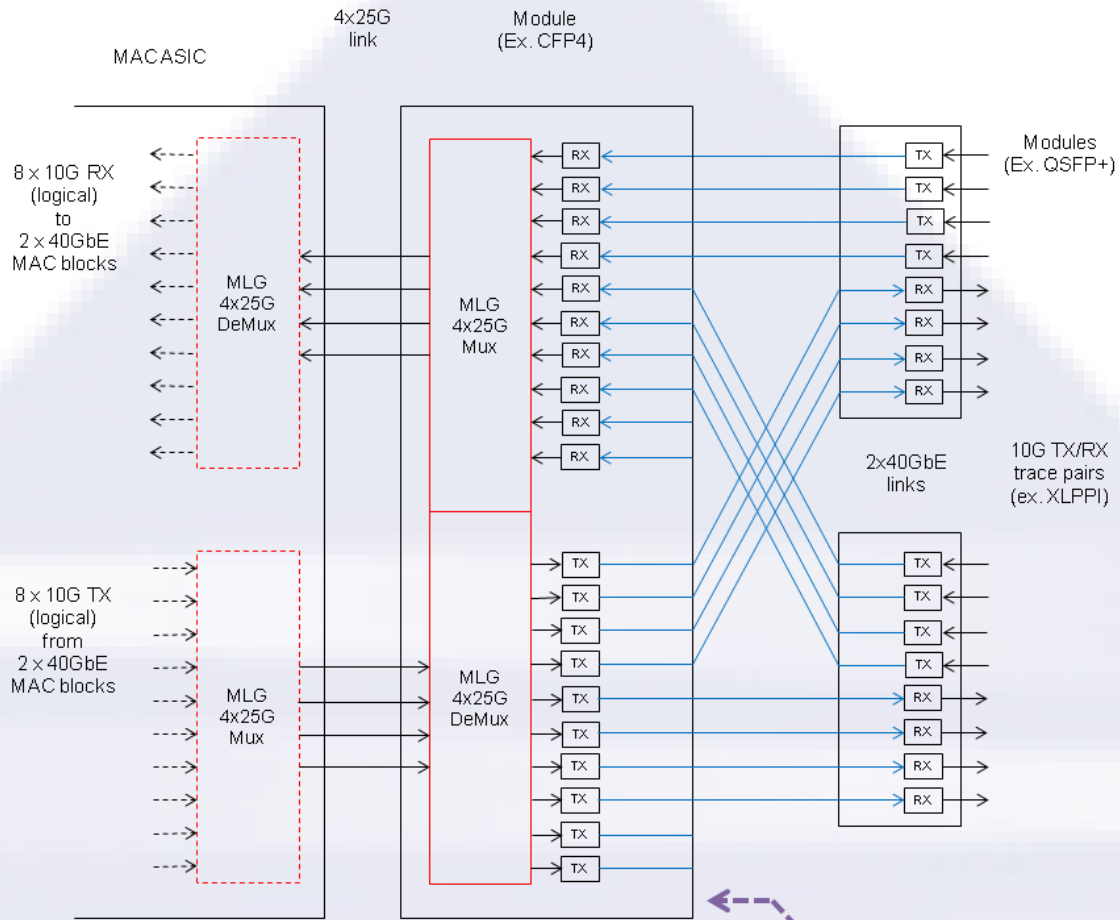


Figure 12: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 10GBASE-SR and 40GBASE-SR4 Port Expander

Multi-link Gearbox 2.0 Implementation Agreement



MPO 2x12 MMF Receptacle Lane Assignments

MPO 2x12 MMF Receptacle Lane Assignments											
Module Front View											
Keyway											
			40GE_1 Rx	40GE_1 Rx	40GE_1 Rx	40GE_1 Rx	40GE_0 Rx	40GE_0 Rx	40GE_0 Rx	40GE_0 Rx	
			40GE_1 Tx	40GE_1 Tx	40GE_1 Tx	40GE_1 Tx	40GE_0 Tx	40GE_0 Tx	40GE_0 Tx	40GE_0 Tx	

Figure 13: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 40GBASE-SR4 Port Expander

Multi-link Gearbox 2.0 Implementation Agreement

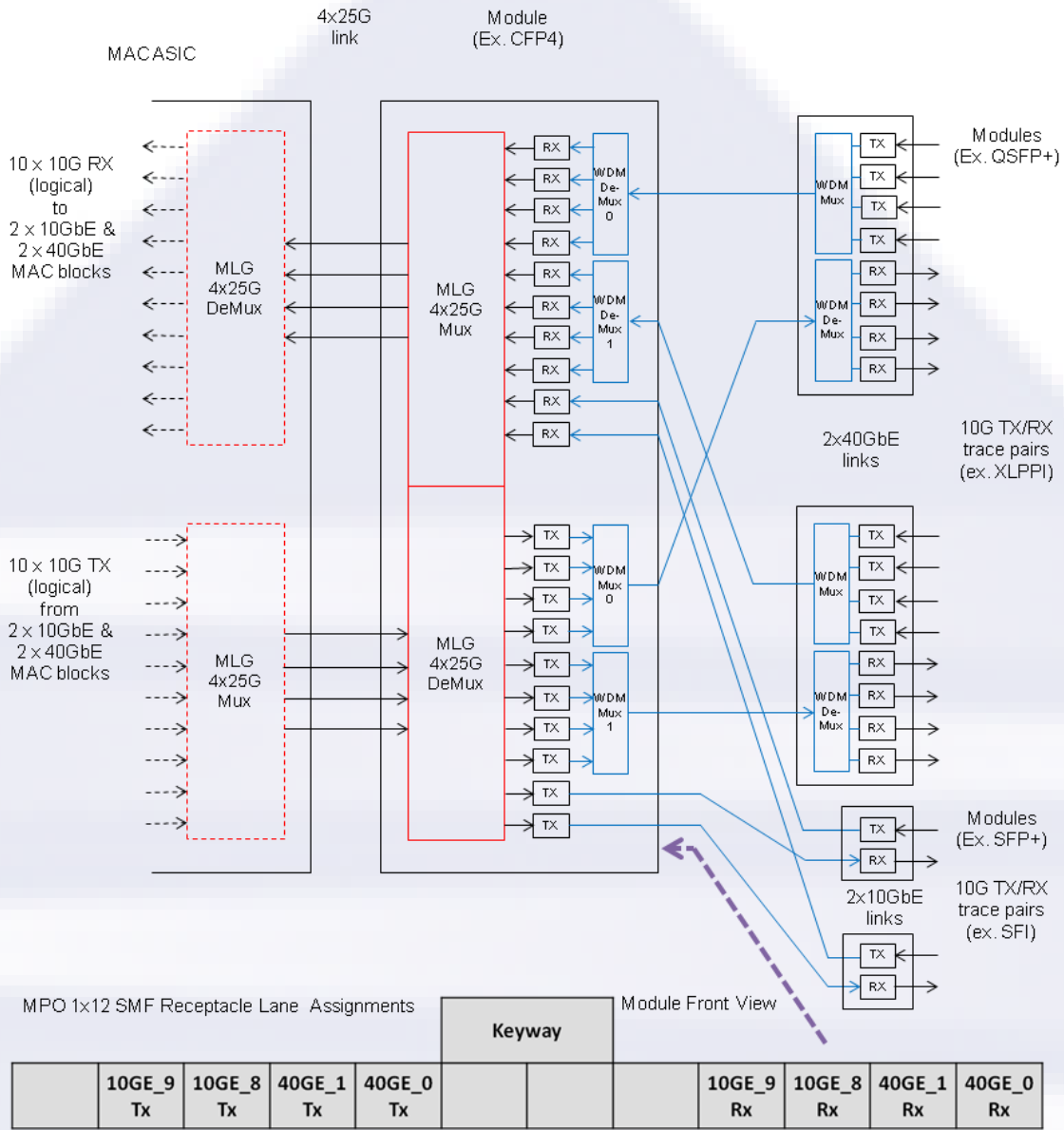


Figure 14: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 10GBASE-LR and 40GBASE-LR4 Port Expander

Multi-link Gearbox 2.0 Implementation Agreement

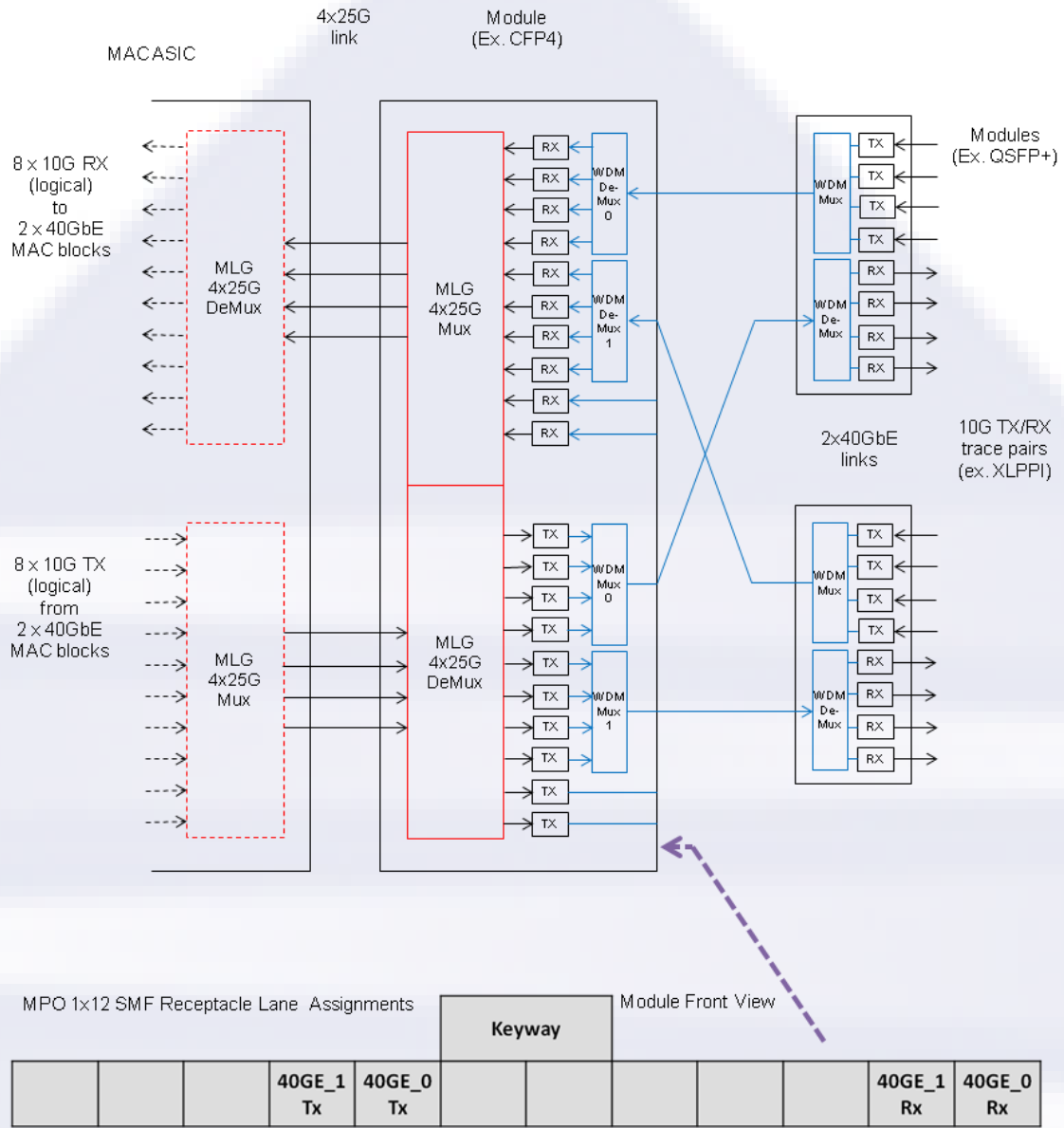


Figure 15: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 40GBASE-LR4 Port Expander

Multi-link Gearbox 2.0 Implementation Agreement

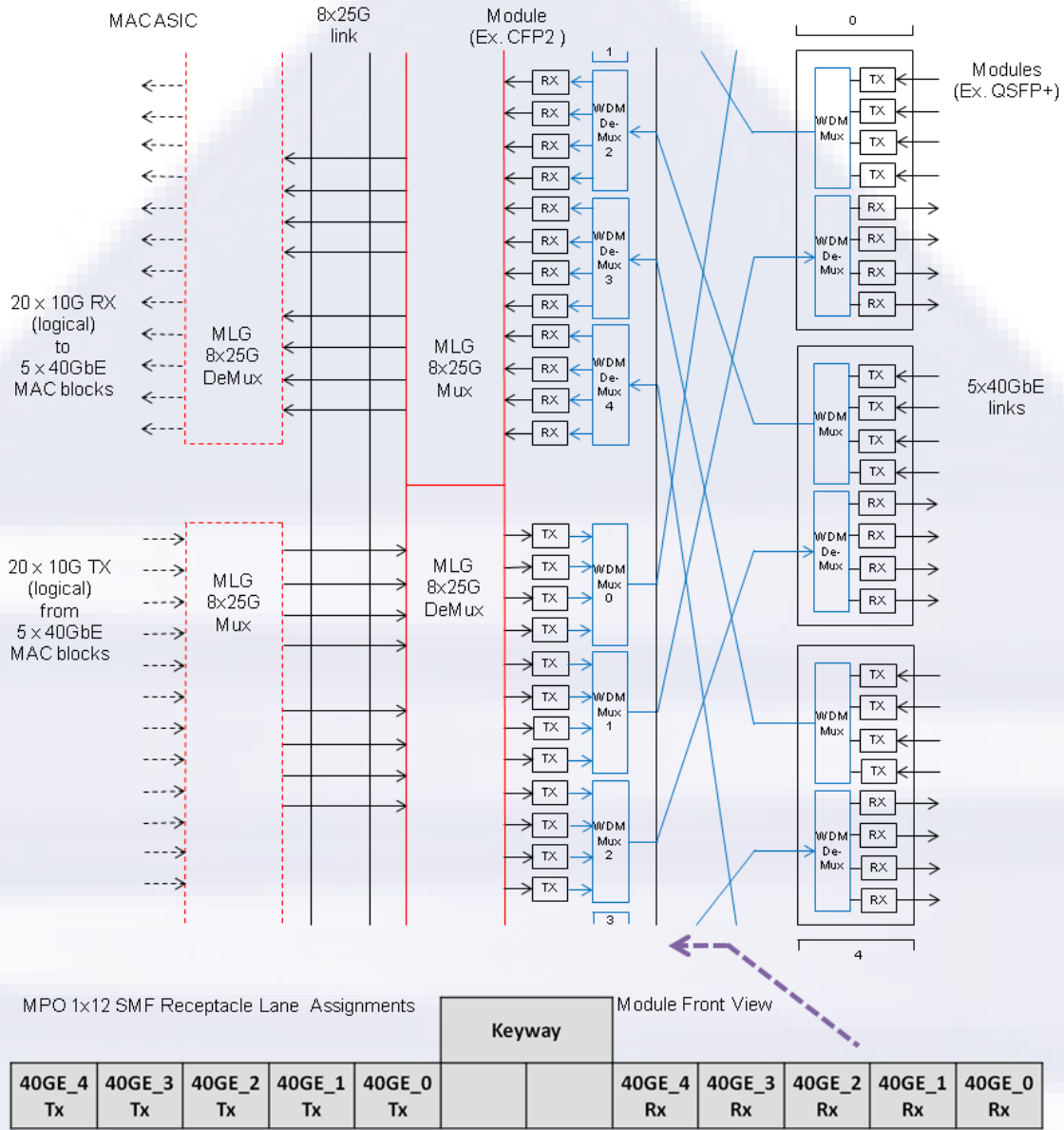


Figure 16: MPO Connector Physical Lane assignments for 8x25G Optical 40GBASE-R Port Expander

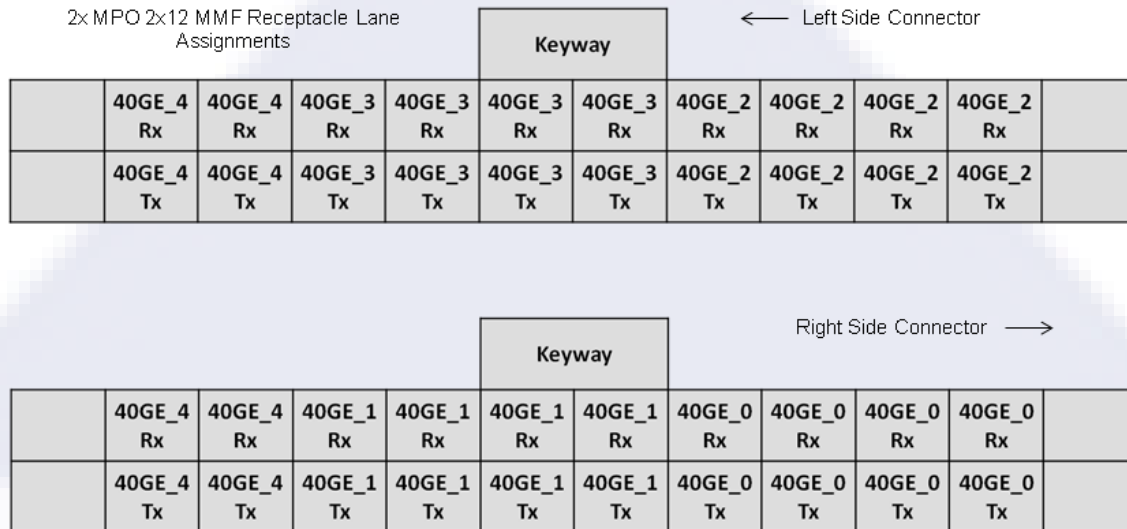


Figure 17: MPO Receptacle Physical Lane Assignments for 8x25G Optical 40GBASE-SR4 Port Expander

10 Appendix B (Informative): Alternate Approach for mapping of 40GBASE-R over 4x20G lanes

Applications which are restricted to the need to support two 40GBASE-R interfaces over a four-lane interface, do not need 10GBASE-R support, and hence would only result in 80% utilization of the 4x25G MLG specified in this document could consider using a simpler approach described in this Appendix.

The PMA specified in IEEE Std 802.3-2012 clause 83 is fully parameterized, and therefore can be used to describe the behavior of a 4:2 bit-mux and 2:4 bit-demux that could allow transport of 40GBASE-R over two physical lanes of 20.625 Gb/s. Note that no currently standardized 40GBASE-R PHY uses a PMA that would adapt to this lane rate.

Such physical lanes could be carried over electrical interfaces such as the CEI-28G-VSR interface currently under development in OIF, and might be transportable over a pair of lanes designed for the “CAUI-4” interface expected to be specified by the IEEE P802.3bm project which will be designed to operate at lane rate of 25.78125 Gb/s.

Examples of 40GBASE-SR4 and 40GBASE-LR4 port expanders using this approach are illustrated in Figure 18 and Figure 19 respectively.

It may also be possible to use certain non-gearbox 100GBASE-R modules expected to arise from the work of the IEEE P802.3bm project (with 4x25G electrical lanes directly driving 4x25G optical lanes) for something like a 2x40GBASE-R virtual link application. The engineering to verify the correct operation of such modules using a 20.625 Gb/s

Multi-link Gearbox 2.0 Implementation Agreement

lane rate instead of a 25.78125 Gb/s lane rate is outside the scope of this implementation agreement.

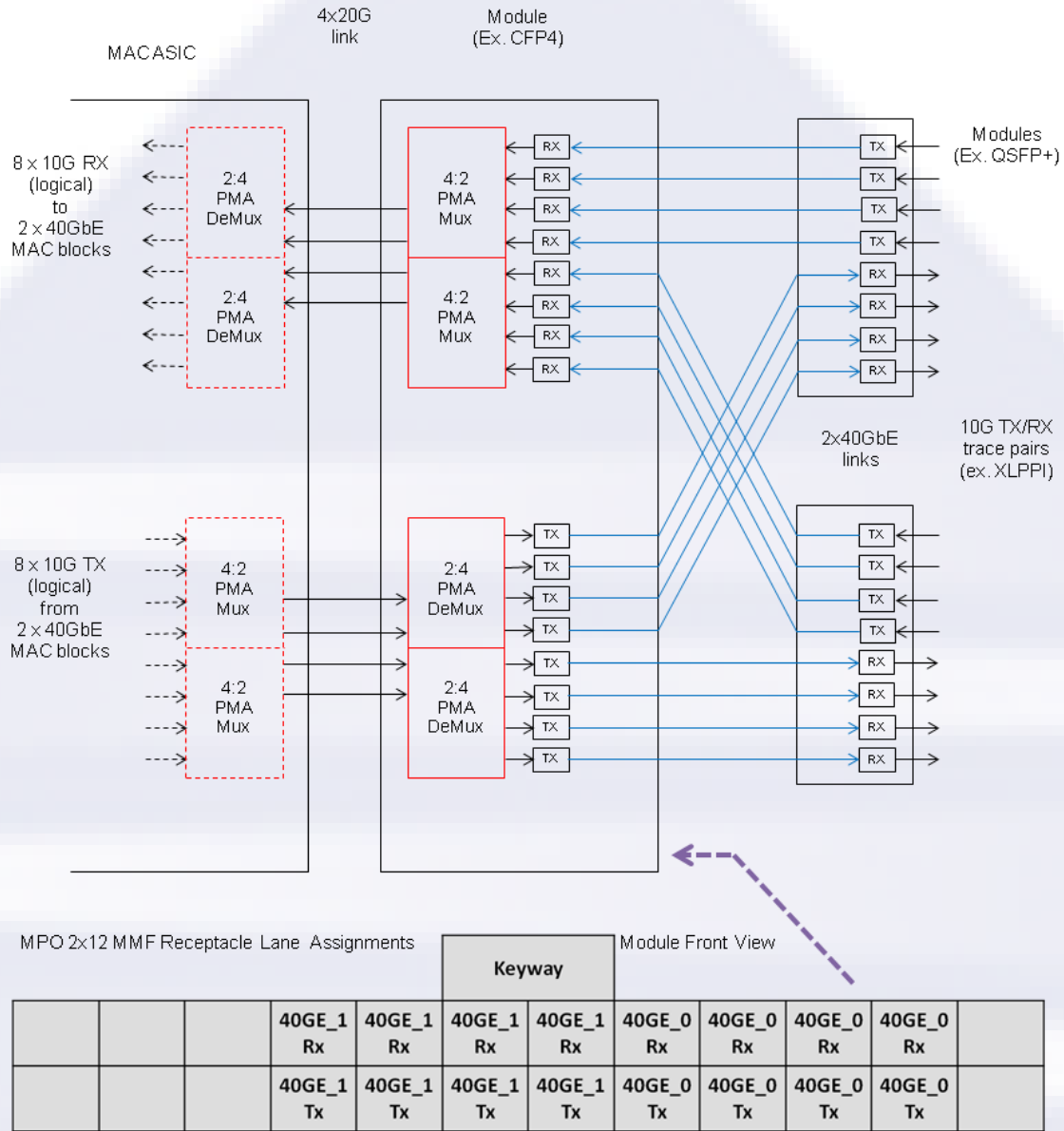


Figure 18: MPO Receptacle Physical Lane Assignments for 4x20G Optical 40GBASE-SR4 Port Expander

Multi-link Gearbox 2.0 Implementation Agreement

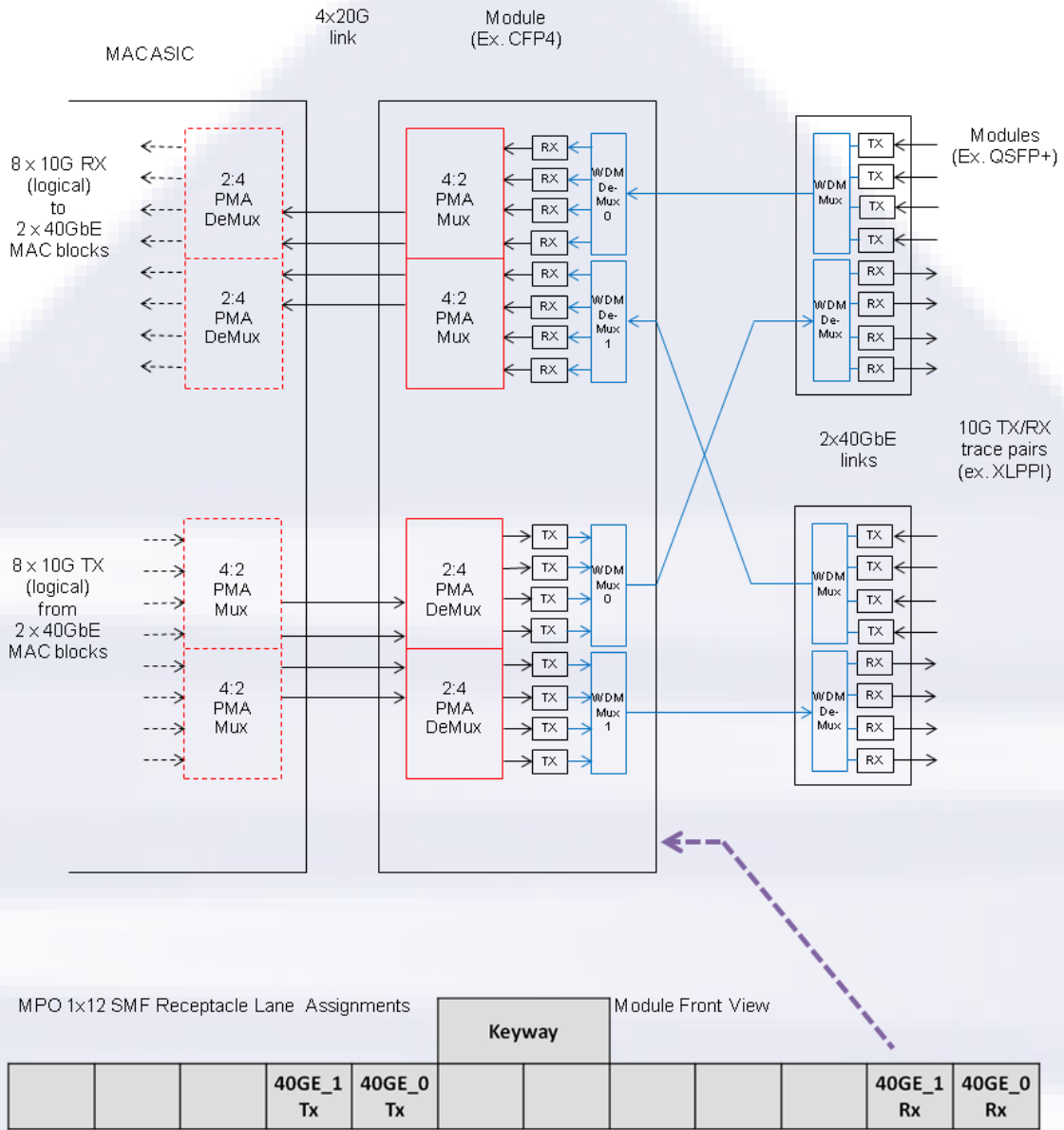


Figure 19: MPO Receptacle Physical Lane Assignments for 4x20G Optical 40GBASE-LR4 Port Expander

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

Acacia Communications	EXFO	MoSys, Inc.
ADVA Optical Networking	FCI USA LLC	MultiPhy Ltd
Agilent Technologies R&D and Marketing GmbH & Co.	Fiberhome Technologies Group	NEC
Alcatel-Lucent	Finisar Corporation	NeoPhotonics
Altera	France Telecom Group/Orange	Nokia Siemens Networks
AMCC	Fujitsu	NTT Corporation
Amphenol Corp.	Furukawa Electric Japan	Oclaro
Anritsu	GigOptix Inc.	Optoplex
Applied Communication Sciences	Hewlett Packard	PETRA
AT&T	Hitachi	Picomatrix
Avago Technologies Inc.	Hittite Microwave Corp	PMC Sierra
Broadcom	Huawei Technologies	QLogic Corporation
Brocade	IBM Corporation	Reflex Photonics
Centellax, Inc.	Infinera	Semtech
China Telecom	Inphi	SHF Communication Technologies
Ciena Corporation	Intel	Skorpios Technologies
Cisco Systems	IPtronics	Sumitomo Electric Industries
ClariPhy Communications	JDSU	Sumitomo Osaka Cement
Cogo Optronics	Juniper Networks	TE Connectivity
Comcast	Kandou	Tektronix
Cortina Systems	KDDI R&D Laboratories	Tellabs
CPqD	Kotura, Inc.	TeraXion
CyOptics	LeCroy	Texas Instruments
Dell, Inc.	LSI Corporation	Time Warner Cable
Department of Defense	Luxtera	TriQuint Semiconductor
Deutsche Telekom	M/A-COM Technology Solutions, Inc.	u2t Photonics AG
EigenLight.com	Marben Products	Verizon
Emcore	Metaswitch	Vitesse Semiconductor
Ericsson	Mindspeed	Xilinx
ETRI	Mitsubishi Electric Corporation	Xtera Communications
	Molex	Yamaichi Electronics Ltd.