Multi-link Gearbox Implementation Agreement

IA # OIF-MLG-03.0

April 2016

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 COBO, Ethernet Alliance, IEEE 802.1, IEEE 802.3, IETF, INCITS T11, ITU-T SG15, MEF, ETSI NFV, ONF, SFF Committee, ATIS-COAST, ATIS-TMOC, TMF, the CFP-MSA Group and the XFP MSA Group.

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

Working Group: **Physical and Link Layer**

TITLE: **Multi-link Gearbox Implementation Agreement 3.0**

SOURCE:

TECHNICAL EDITOR

Scott A. Irwin, Ph.D.
MoSys, Inc.
1606 Golden Aspen Dr, Ste 102
Ames, IA 50010 USA
Phone: +1 515 663 8820 x201
Email: sairwin@mosys.com

WORKING GROUP CHAIR

David R. Stauffer, Ph.D.
Kandou Bus SA
EPFL Innovation Park Bldg. I
1015 Lausanne Switzerland
Phone: +1 802 316 0808
Email: david@kandou.com

ABSTRACT: The MLG (Multi-Link Gearbox) 3.0 Implementation Agreement defines three MLG configurations: A 4x25G lane configuration is comprised of 20 MLG lanes (similar to MLG 1.0/2.0), where 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 (similar to MLG 2.0), where each group of eight MLG lanes can carry either four 10GBASE-R signals or a single 40GBASE-R signal. A 2x20G/1x40G lane configuration is comprised of 4 MLG lanes (similar to 40GBASE-R) to carry up to four 10GBASE-R signals. MLG 3.0 adds support for transport of MLG lanes across 4x25G RS-FEC encoded links (per IEEE 802.3 Clause 91). In addition, MLG 3.0 defines an optional inband communication mechanism to support monitoring and control of remote devices.

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.

© 2015 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 6
4 Document Revision History 7
5 Introduction 8
5.1 Requirements 8
5.2 Sample Applications 9
6 MLG40 17
6.1 MLG40 General Mechanism 17
6.2 MLG40 Detailed Block Diagrams 18
7 MLG100 26
7.1 MLG100 General Mechanism 26
7.2 MLG100 Detailed Block Diagrams (without RS-FEC and without RM) 30
7.3 MLG100 Detailed Block Diagrams (without RS-FEC and with RM) 41
7.4 MLG100 Detailed Block Diagrams (with RS-FEC and optional RM) 56
8 MLG200 71
8.1 MLG200 General Mechanism 71
8.2 MLG200 Detailed Block Diagrams (without RS-FEC and RM) 76
8.3 MLG200 Detailed Block Diagrams (without RS-FEC and with RM) 87
8.4 MLG200 Detailed Block Diagrams (with RS-FEC and optional RM) 102
9 MLG Management 114
9.1 MLG Mux Management 114
9.2 MLG Demux Management 115
9.3 Generic MLG Management 116
10 References 117
10.1 Normative references 117
10.2 Informative references 117
11 Appendix A (Informative): Suggested MPO Receptacle Physical Lane Assignments for various MLG applications 118
12 Appendix B (Informative): Alternate Approach for mapping of 40GBASE-R over 4x20G lanes 124
13 Appendix C: List of companies belonging to OIF when document was approved 127

2 List of Figures

FIGURE 1: MLG100 10GBASE-R VIRTUAL LINK 9
FIGURE 2: MLG100 40GBASE-R AND 10GBASE-R VIRTUAL LINK 10
FIGURE 3: MLG200 40GBASE-R VIRTUAL LINK 10
FIGURE 4: MLG40 10GBASE-R VIRTUAL LINK (2X20G) 11
FIGURE 5: MLG40 10GBASE-R VIRTUAL LINK (1X40G) 11
FIGURE 6: MLG100 10GBASE-R PORT EXPANDER 12
FIGURE 7: MLG100 40GBASE-R AND 10GBASE-R PORT EXPANDER 13
FIGURE 8: MLG200 40GBASE-R PORT EXPANDER 14
FIGURE 9: MLG40 10GBASE-R PORT EXPANDER (2X20G) 15
FIGURE 10: MLG40 10GBASE-R PORT EXPANDER (1X40G) 15
FIGURE 11: APPLICATION OF RS-FEC WITH MLG LINK 16
FIGURE 12: MLG MONITORING AND CONTROL APPLICATION 16
FIGURE 13: MLG40 10GBASE-R MUX BLOCK DIAGRAM 18
FIGURE 14: MLG40 SCRAMBLER DIAGRAM AND 66B BLOCK FLOW 20
FIGURE 15: MLG40 10GBASE-R DEMUX BLOCK DIAGRAM 22
FIGURE 16: MLG40 DESCRAMBLER DIAGRAM AND 66B BLOCK FLOW 24
FIGURE 17: MLG100 RS-FEC AM MAPPING 28
FIGURE 18: MLG100 RS-FEC BIP MAPPING 28
FIGURE 19: MLG100 RM DATA CHANNEL 29
FIGURE 20: MLG100 MUX 10GBASE-R MAPPING BLOCK DIAGRAM 30
FIGURE 21: MLG100 MUX 40GBASE-R MAPPING BLOCK DIAGRAM 33
FIGURE 22: MLG100 DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM 36
FIGURE 23: MLG100 DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM 39
FIGURE 24: MLG100 MUX 10GBASE-R MAPPING BLOCK DIAGRAM 42
FIGURE 25: MLG100 SCRAMBLER DIAGRAM AND 66B BLOCK FLOW 45
FIGURE 26: MLG100 MUX 40GBASE-R MAPPING BLOCK DIAGRAM 47
FIGURE 27: MLG100 DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM 50
FIGURE 28: MLG100 DESCRAMBLER DIAGRAM AND 66B BLOCK FLOW 52
FIGURE 29: MLG100 DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM 54
FIGURE 30: MLG100 MUX 10GBASE-R MAPPING BLOCK DIAGRAM 57
FIGURE 31: MLG100 MUX 40GBASE-R MAPPING BLOCK DIAGRAM 61
FIGURE 32: MLG100 DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM 65
FIGURE 33: MLG100 DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM 68
FIGURE 34: MLG200 RS-FEC AM MAPPING 73
FIGURE 35: MLG200 RS-FEC BIP MAPPING 74
FIGURE 36: MLG200 RM DATA CHANNEL 75
FIGURE 37: MLG200 MUX 10GBASE-R MAPPING BLOCK DIAGRAM 76
FIGURE 38: MLG200 MUX 40GBASE-R MAPPING BLOCK DIAGRAM 79
FIGURE 39: MLG200 DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM 82
FIGURE 40: MLG200 DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM 85
FIGURE 41: MLG200 MUX 10GBASE-R MAPPING BLOCK DIAGRAM 88
FIGURE 42: MLG200 SCRAMBLER DIAGRAM AND 66B BLOCK FLOW 91
FIGURE 43: MLG200 MUX 40GBASE-R MAPPING BLOCK DIAGRAM 93
FIGURE 44: MLG200 DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM 96
FIGURE 45: MLG200 DESCRAMBLER DIAGRAM AND 66B BLOCK FLOW 98
FIGURE 46: MLG200 DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM 100
FIGURE 47: MLG200 MUX 10GBASE-R MAPPING BLOCK DIAGRAM 103
FIGURE 48: MLG200 MUX 40GBASE-R MAPPING BLOCK DIAGRAM 106
FIGURE 49: MLG200 DEMUX 10GBASE-R DEMAPPING BLOCK DIAGRAM 109

FIGURE 50: MLG200 DEMUX 40GBASE-R DEMAPPING BLOCK DIAGRAM 112
FIGURE 51: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG 10GBASE-SR OR 10GBASE-LR OPTICAL PORT EXPANDER 118
FIGURE 52: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 10GBASE-SR AND 40GBASE-SR4 PORT EXPANDER 119
FIGURE 53: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 40GBASE-SR4 PORT EXPANDER 120
FIGURE 54: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 10GBASE-LR AND 40GBASE-LR4 PORT EXPANDER 121
FIGURE 55: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X25G MLG OPTICAL 40GBASE-LR4 PORT EXPANDER 122
FIGURE 56: MPO CONNECTOR PHYSICAL LANE ASSIGNMENTS FOR 8X25G OPTICAL 40GBASE-R PORT EXPANDER 123
FIGURE 57: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 8X25G OPTICAL 40GBASE-SR4 PORT EXPANDER 124
FIGURE 58: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X20G OPTICAL 40GBASE-SR4 PORT EXPANDER 125
FIGURE 59: MPO RECEPTACLE PHYSICAL LANE ASSIGNMENTS FOR 4X20G OPTICAL 40GBASE-LR4 PORT EXPANDER 126

3 List of Tables

TABLE 1: MLG40 LANE ALIGNMENT MARKER VALUES 17
TABLE 2: MLG100 LANE ALIGNMENT MARKER VALUES 27
TABLE 3: MLG200 LANE ALIGNMENT MARKER VALUES 72

4 Document Revision History

Working Group: Physical and Link Layer

SOURCE:

Editor's Name
Scott A. Irwin, Ph.D.
MoSys, Inc.
1606 Golden Aspen Dr, Ste 102
Ames, IA 50010 USA
Phone: +1 515 663 8820 x201
Email: sairwin@mosys.com

Working Group Chair
David R. Stauffer, Ph.D.
Kandou Bus SA
PFL Innovation Park Bldg. I
1015 Lausanne Switzerland
Phone: +1 802 316 0808
Email: david@kandou.com

DATE: **Date Approved**

Issue No.	Issue Date	Details of Change
oif2014.148.03	January 2015	Baseline Text
oif2015.148.04	April 2015	Consolidated MLG Management section. General corrections throughout.
oif2015.148.05	July 2015	Updates per 3Q2015 comment resolution. RS-FEC descriptions in MLG100 and MLG200. Misc corrections.
oif2015.148.06	Oct 2015	Updates per 4Q2015 comment resolution. Defined 5b and 8b pad of RS-FEC AM mapping. Misc grammatical and punctuation corrections.

5 Introduction

The MLG (Multi-link Gearbox) 3.0 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). MLG 3.0 supports 100G links with or without 4x25G RS-FEC (per IEEE Std 802.3bj-2014 Clause 91). MLG 3.0 also defines an in-band coding that allows independent 10GBASE-R signals to transit physical 20G and 40G lanes. This enables a variety of applications to reuse 40GBASE-R and 100GBASE-R technology for the transport of individual 10G and 40G links.

In addition, MLG 3.0 defines an optional inband communication protocol to support monitoring and control of remote devices.

5.1 Requirements

Three applications are defined:

- The MLG 3.0 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 3.0 4x25G gearbox is compatible with, and can be interconnected with, an MLG 1.0 gearbox. The term “MLG100” will be used in reference to this application.
- The MLG 3.0 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 term “MLG200” will be used in reference to this application.
- The MLG 3.0 2x20G/1x40G mux application maps 10GBASE-R signals over 4 (10G) MLG lanes, with 2 MLG lanes bit-multiplexed over each physical 20G lane, or 4 MLG lanes bit-multiplexed over each physical 40G lane. The term “MLG40” will be used in reference to this application.

The MLG100 (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 ±100ppm). The MLG100 demux decodes the format produced by the MLG100 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 MLG100 mux can be interconnected with an MLG 1.0 demux, and an MLG 1.0 mux can be interconnected with an MLG100 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 MLG200 (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 MLG200 demux decodes the format produced by the MLG200 mux to produce from zero to twenty 10GBASE-R signals and from zero to five 40GBASE-R signals.

The MLG40 (2x20G/1x40G) mux encodes from zero to four 10GBASE-R signals compliant with IEEE Std 802.3™-2012 Clauses 49 and 51 into a format consistent with the IEEE Std 802.3-2012 Clause 83 PMA service interface (e.g., two physical lanes of 20.625 Gb/s ±100ppm or one physical lane of 41.25 Gb/s ±100ppm). The MLG40 demux decodes the format produced by the MLG40 mux to produce from zero to four 10GBASE-R signals.

The MLG mux to 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 an MLG mux to demux link. The bit sequence may be modified by the insertion and/or deletion of idles and rescrambling, 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 mux to 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 mux to 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 MLG 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:

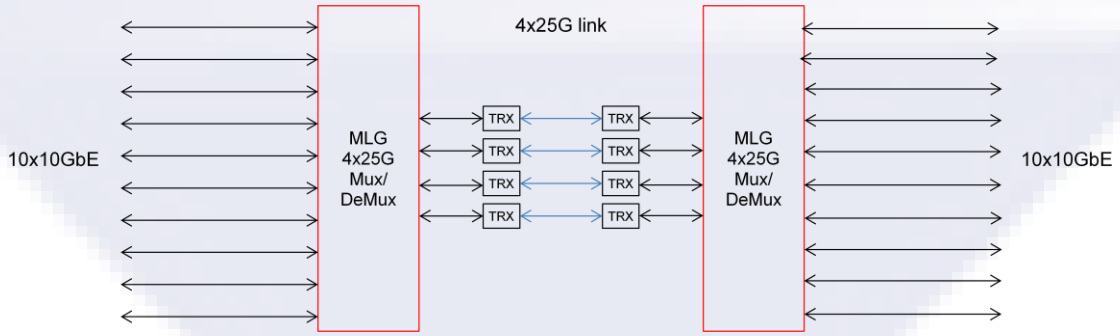


Figure 1: MLG100 10GBASE-R Virtual Link

Up to two (asynchronous) 40GBASE-R signals may be transported over a virtual link using a MLG100 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:

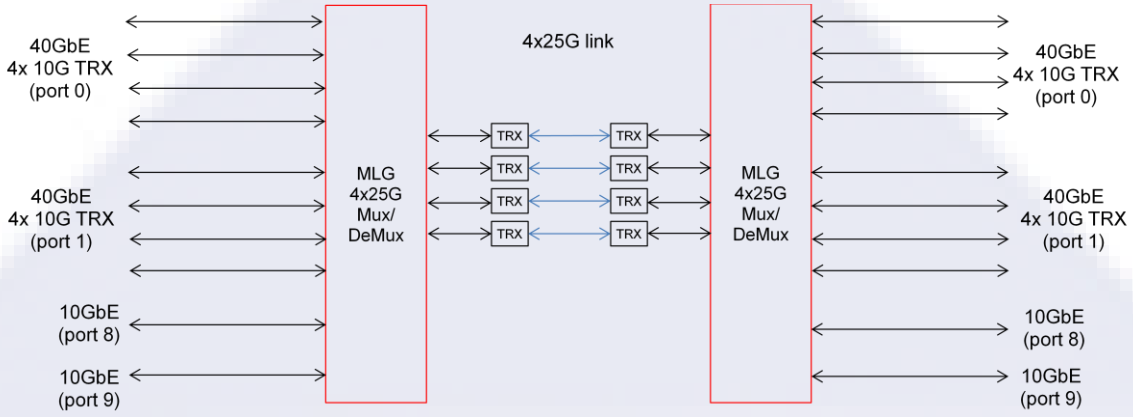


Figure 2: MLG100 40GBASE-R and 10GBASE-R Virtual Link

For the most efficient implementation of an “all 40G” application, a MLG200 gearbox can be used to provide a virtual link for up to five 40GBASE-R signals:

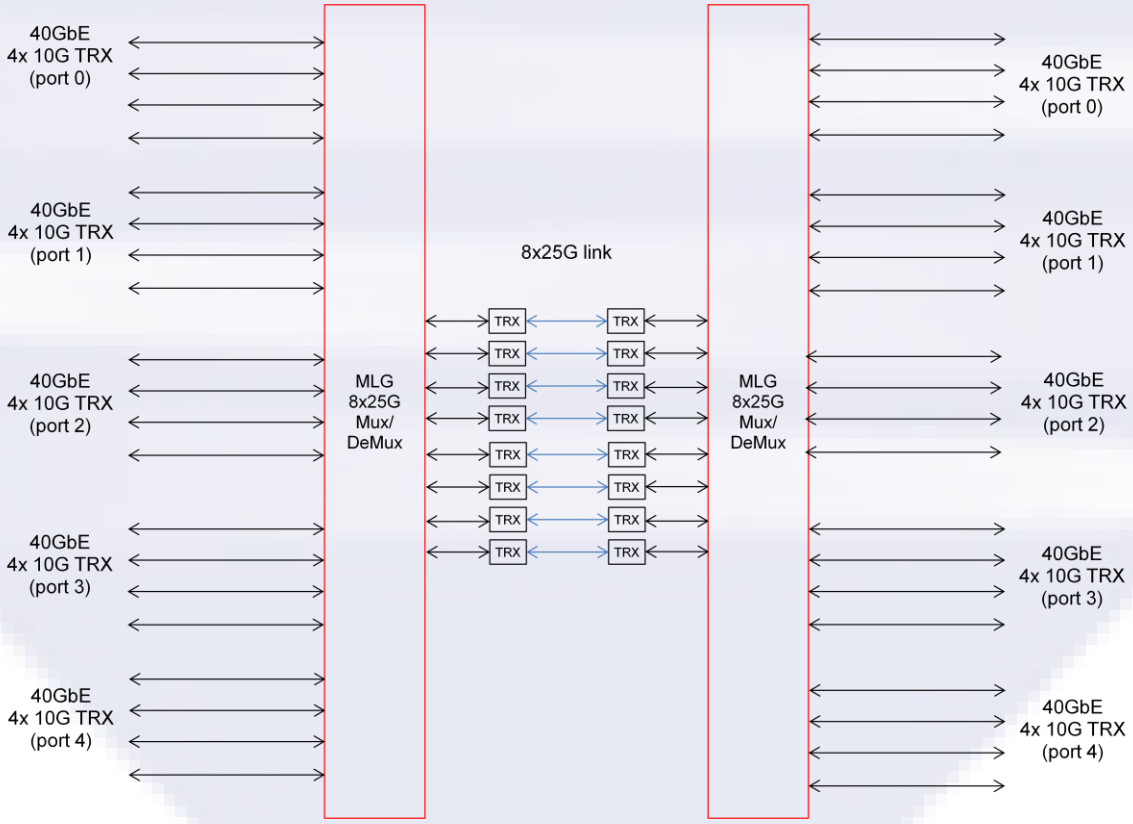


Figure 3: MLG200 40GBASE-R Virtual Link

Up to four (asynchronous) 10GBASE-R signals may also be transported over a virtual link using 2x20G or 1x40G physical lanes, similar to a 40GBASE-R link (per IEEE 802.3 Clause 82 and 83).

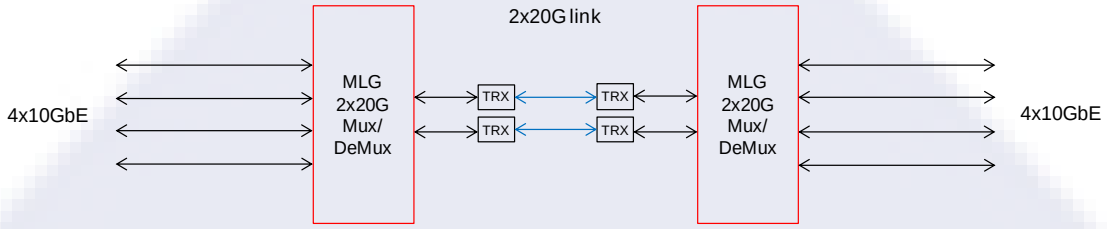


Figure 4: MLG40 10GBASE-R Virtual Link (2x20G)

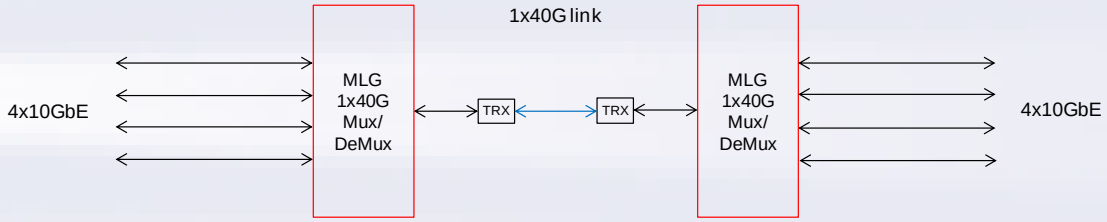


Figure 5: MLG40 10GBASE-R Virtual Link (1x40G)

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.

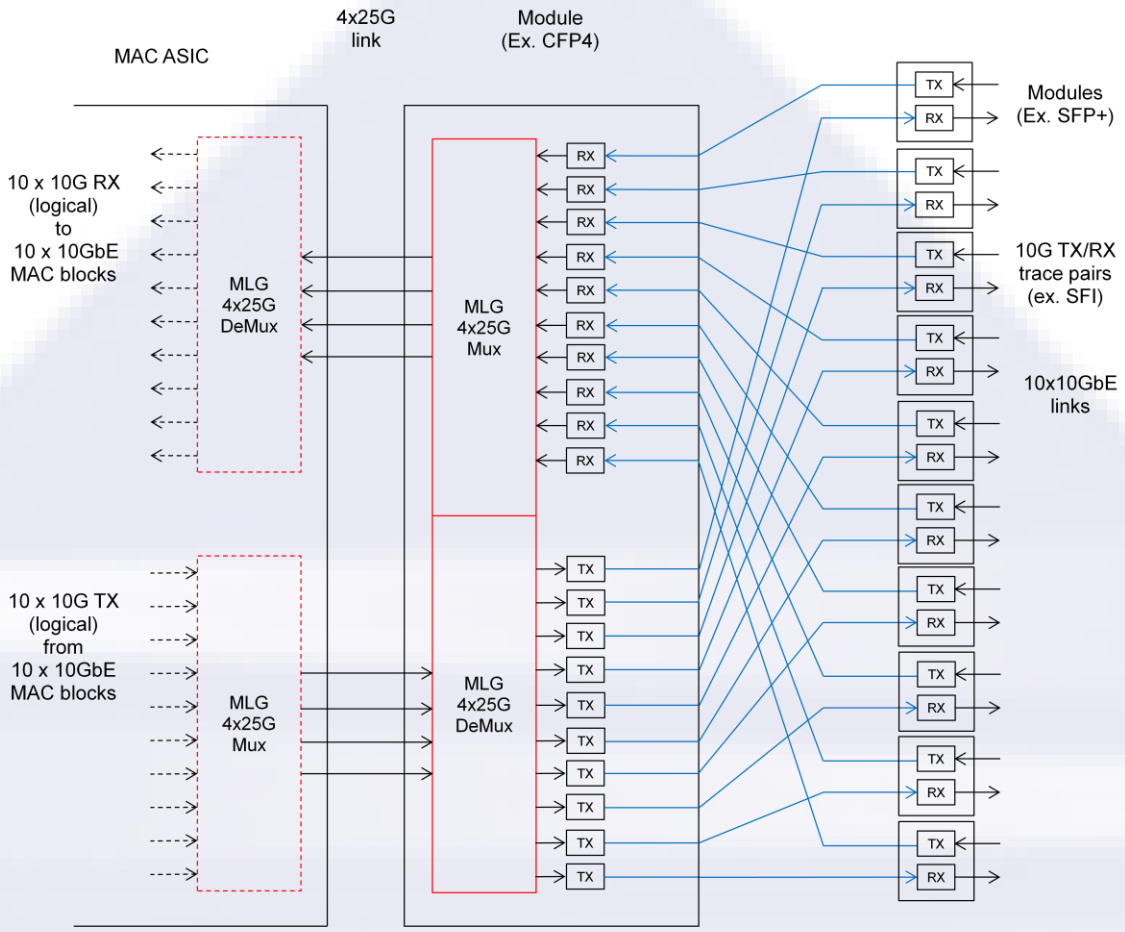


Figure 6: MLG100 10GBASE-R Port Expander

The MLG100 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.

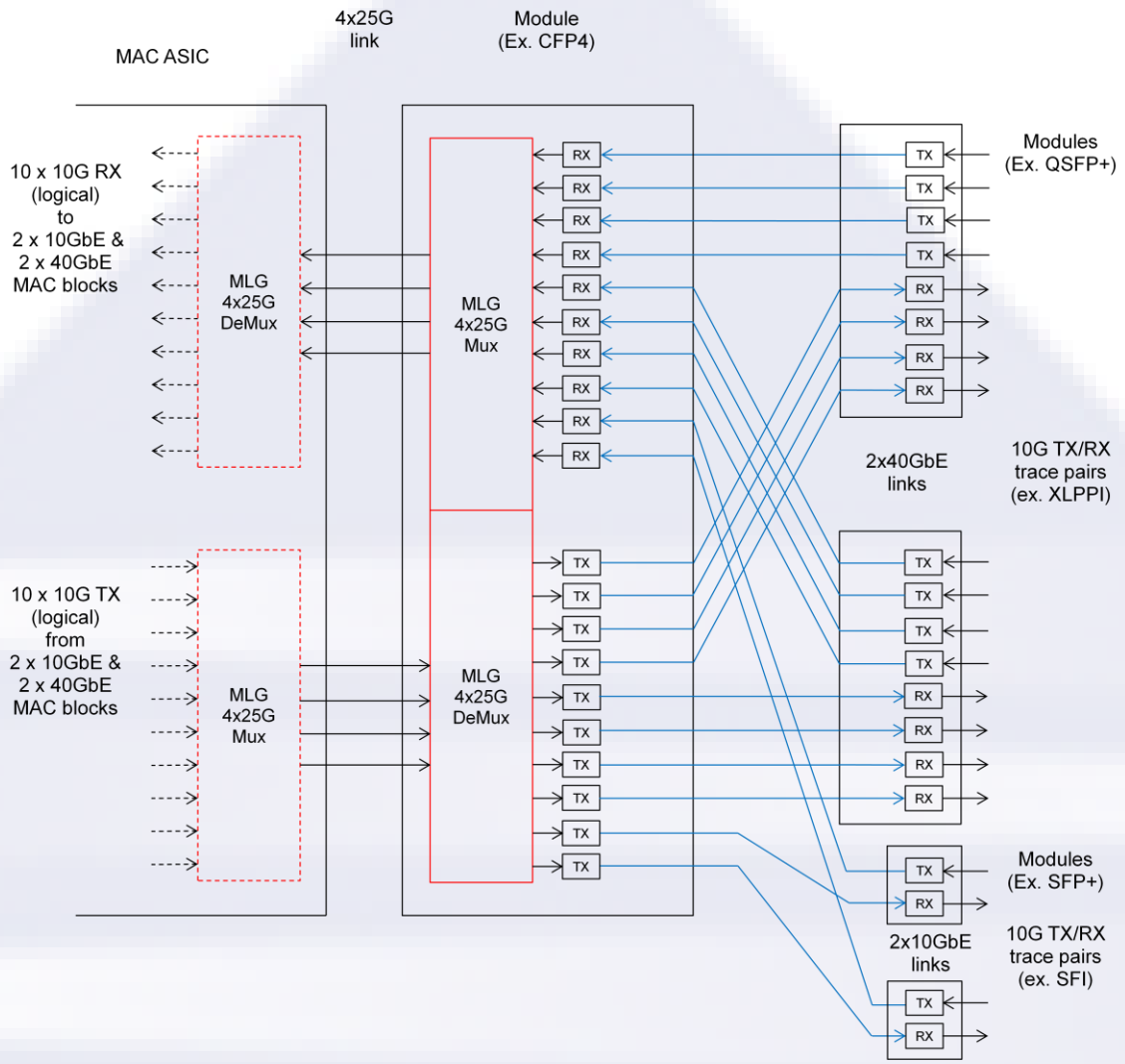


Figure 7: MLG100 40GBASE-R and 10GBASE-R Port Expander

Higher capacity and 40G-only applications with 100% utilization can be provided using the MLG200 gearbox configuration:

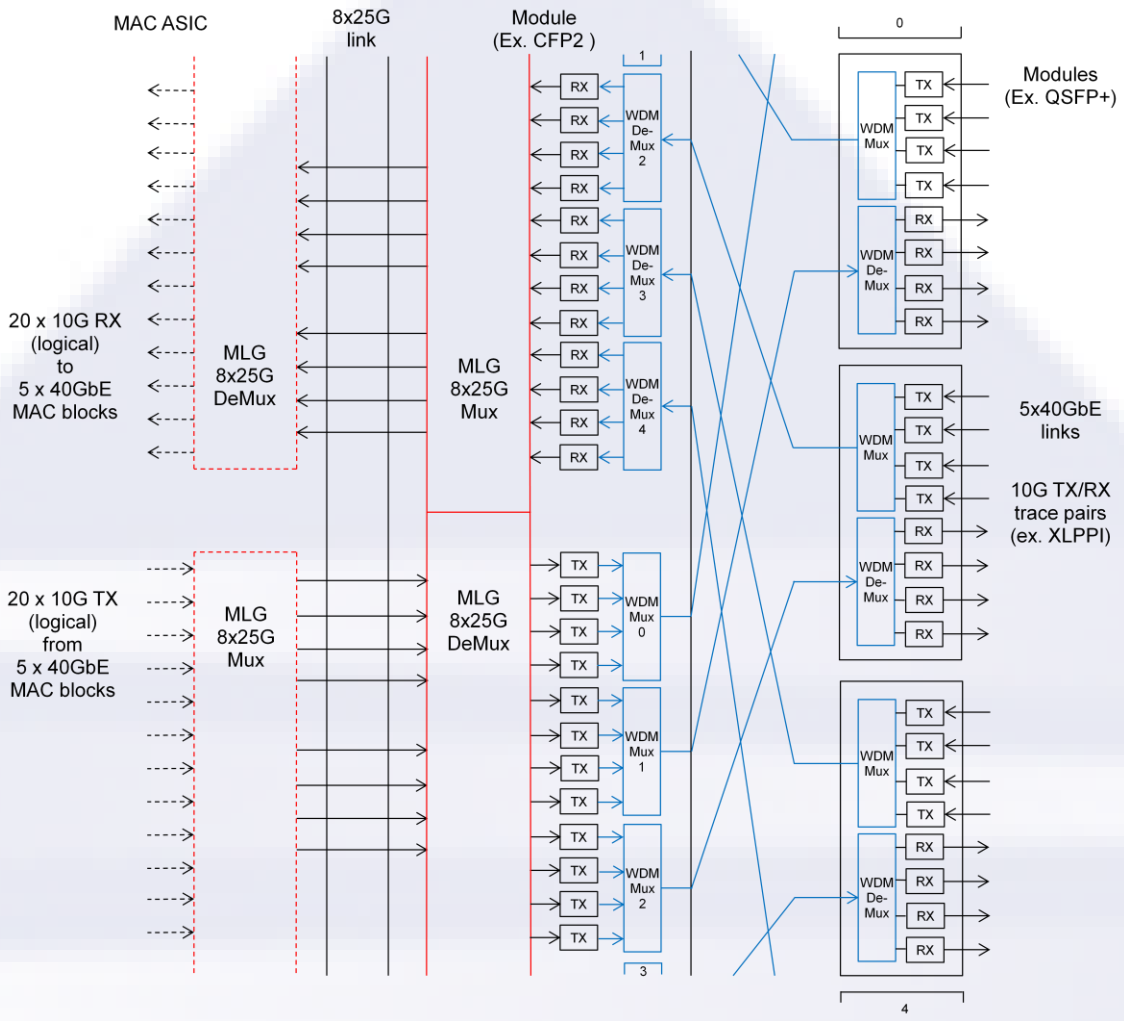


Figure 8: MLG200 40GBASE-R Port Expander

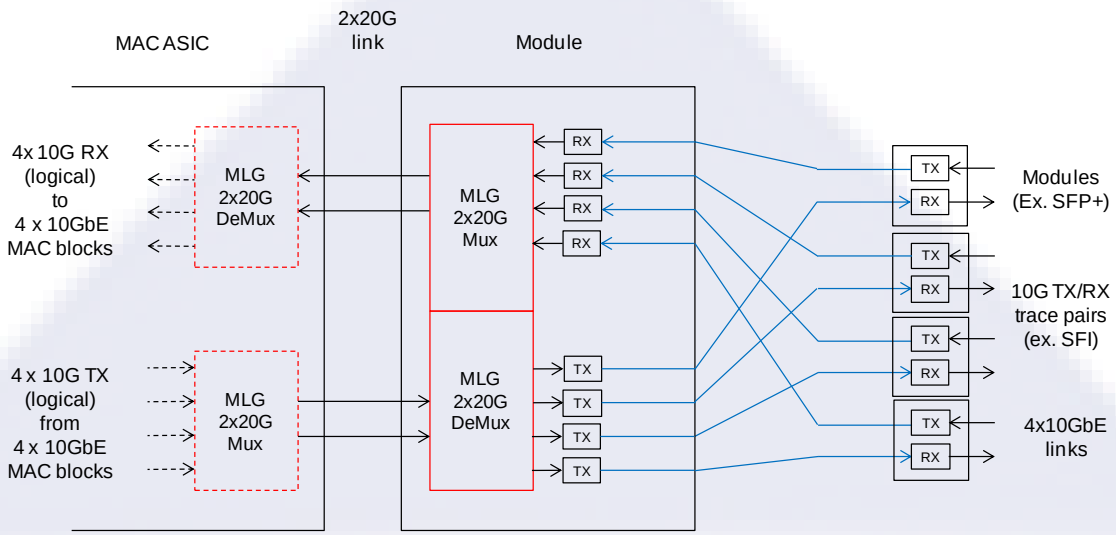


Figure 9: MLG40 10GBASE-R Port Expander (2x20G)

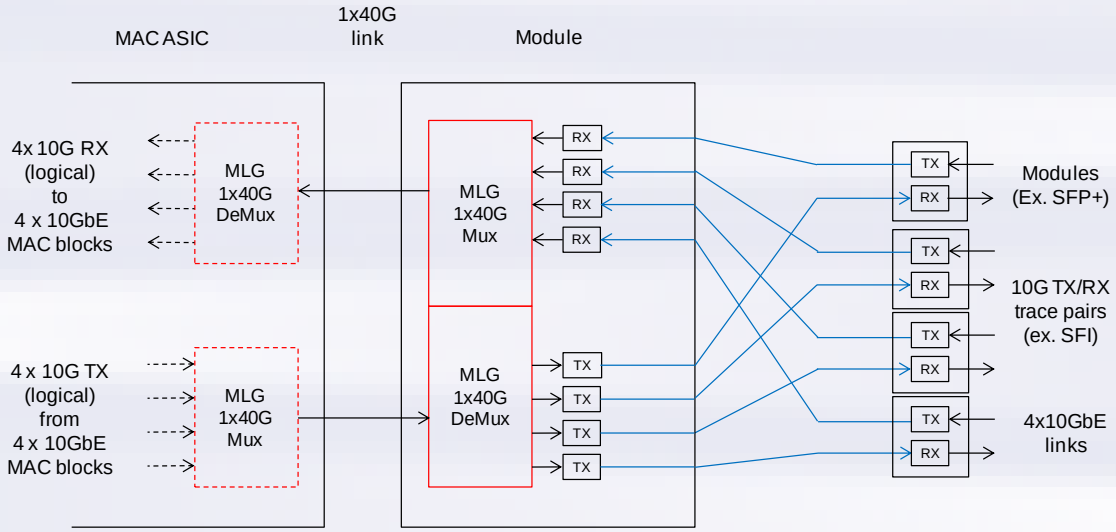


Figure 10: MLG40 10GBASE-R Port Expander (1x40G)

5.2.3 4x25G RS-FEC Links

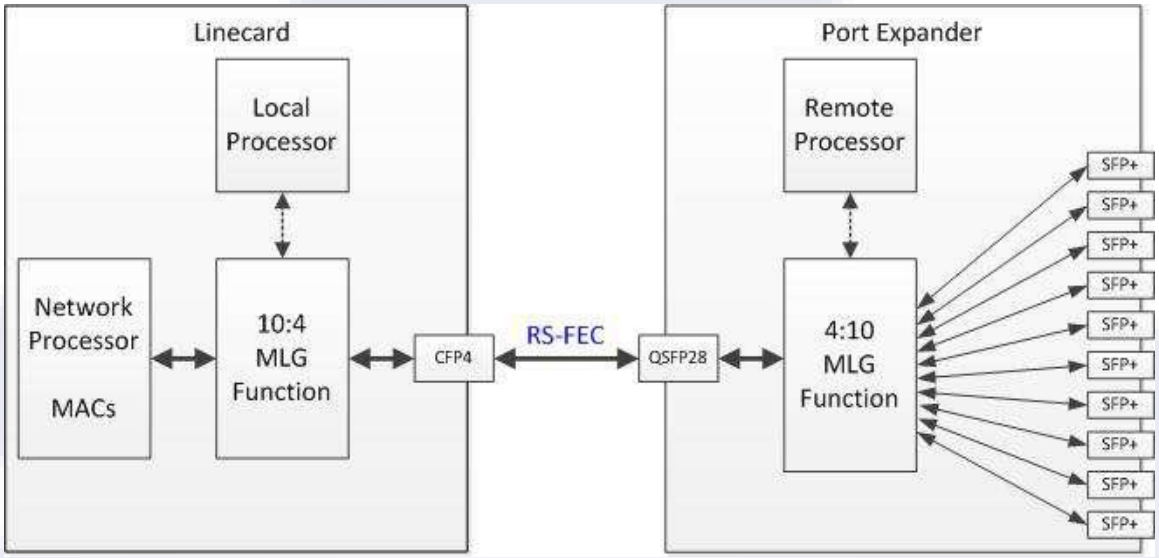


Figure 11: Application of RS-FEC with MLG Link

5.2.4 Management Channel

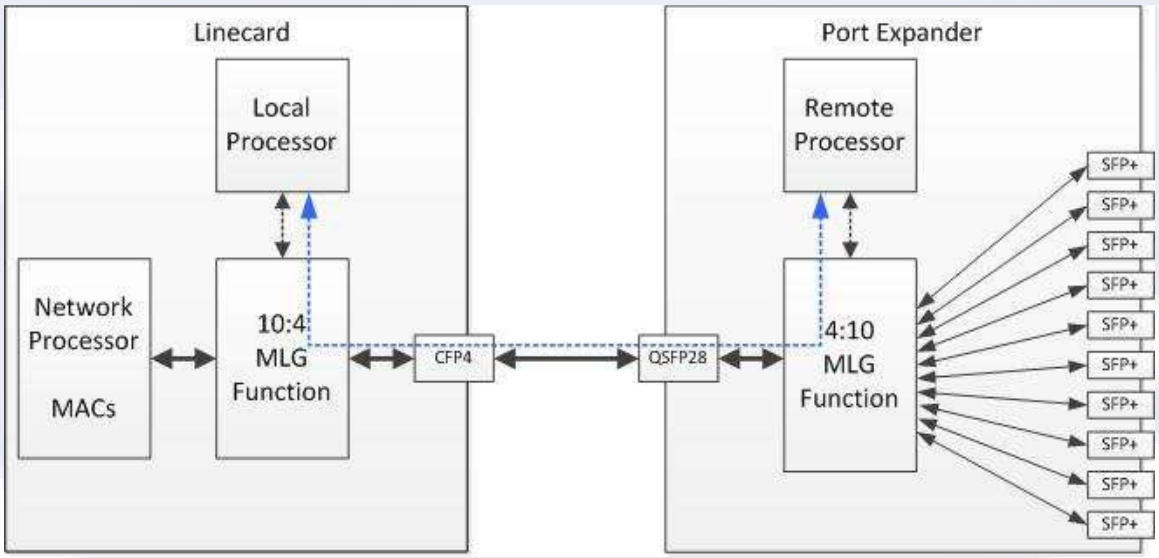


Figure 12: MLG Monitoring and Control Application

6 MLG40

6.1 MLG40 General Mechanism

The MLG40 mechanism reuses the 40GBASE-R PMA, which can combine the information from the 40GBASE-R PCS into a variety of different physical lane widths. The 40GBASE-R PCS is divided into 4 PCS lanes, distributing the 66-bit blocks of the 41.25 Gb/s aggregate at the PCS TX round-robin to four lanes of 10.3125 Gb/s. Sufficient idles are deleted from the 41.25 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 MLG40 to transport 10GBASE-R signals. Sufficient idles are added or removed from each of the incoming 10GBASE-R signals to map them all to a common clock domain and to make room for the insertion of the 66-bit MLG40 lane alignment markers described below. Each 10GBASE-R signal is mapped to a MLG40 lane running at 10.3125 Gb/s. A 66-bit MLG40 lane alignment (and identification) marker is simultaneously inserted on each of these MLG40 lanes after every 16383 66-bit blocks. The MLG40 lanes are identified as lane 0 to 3, each corresponding to one of the mapped 10GBASE-R signals being carried.

At the MLG40 demux, the MLG40 lanes are identified, deskewed, reordered, and the MLG40 lane alignment markers removed. Each of the MLG40 lanes is then mapped back to a 10GBASE-R signal. Idles can then be added or removed from each 10GBASE-R signal to map it onto a new 10G output clock domain (if required).

6.1.1 MLG40 Lane Markers

The MLG40 lane marker values are chosen so as not to replicate the values used in the 40GBASE-R PCS. This prevents bringing up a link which accidentally interconnects an MLG40 formatted signal with a 40GBASE-R PCS.

Table 1 provides the lane alignment markers used for the MLG40 applications with up to four 10GBASE-R signals.

Table 1: MLG40 Lane Alignment Marker Values

MLG40 lane number	Encoding[a] {M0, M1, M2, BIP_3, M4, M5, M6, BIP_7}
0	0x80, 0xB4, 0xAF, BIP_3, 0x7F, 0x4B, 0x50, BIP_7
1	0x29, 0x85, 0x1D, BIP_3, 0xD6, 0x7A, 0xE2, BIP_7
2	0x11, 0x2A, 0xD8, BIP_3, 0xEE, 0xD5, 0x27, BIP_7
3	0xBF, 0x7E, 0x4D, BIP_3, 0x40, 0x81, 0xB2, BIP_7

[a] Each octet is transmitted LSB to MSB.

6.2 MLG40 Detailed Block Diagrams

6.2.1 10GBASE-R Mux

The process for multiplexing 10GBASE-R signals onto MLG40 lanes is shown in Figure 13.

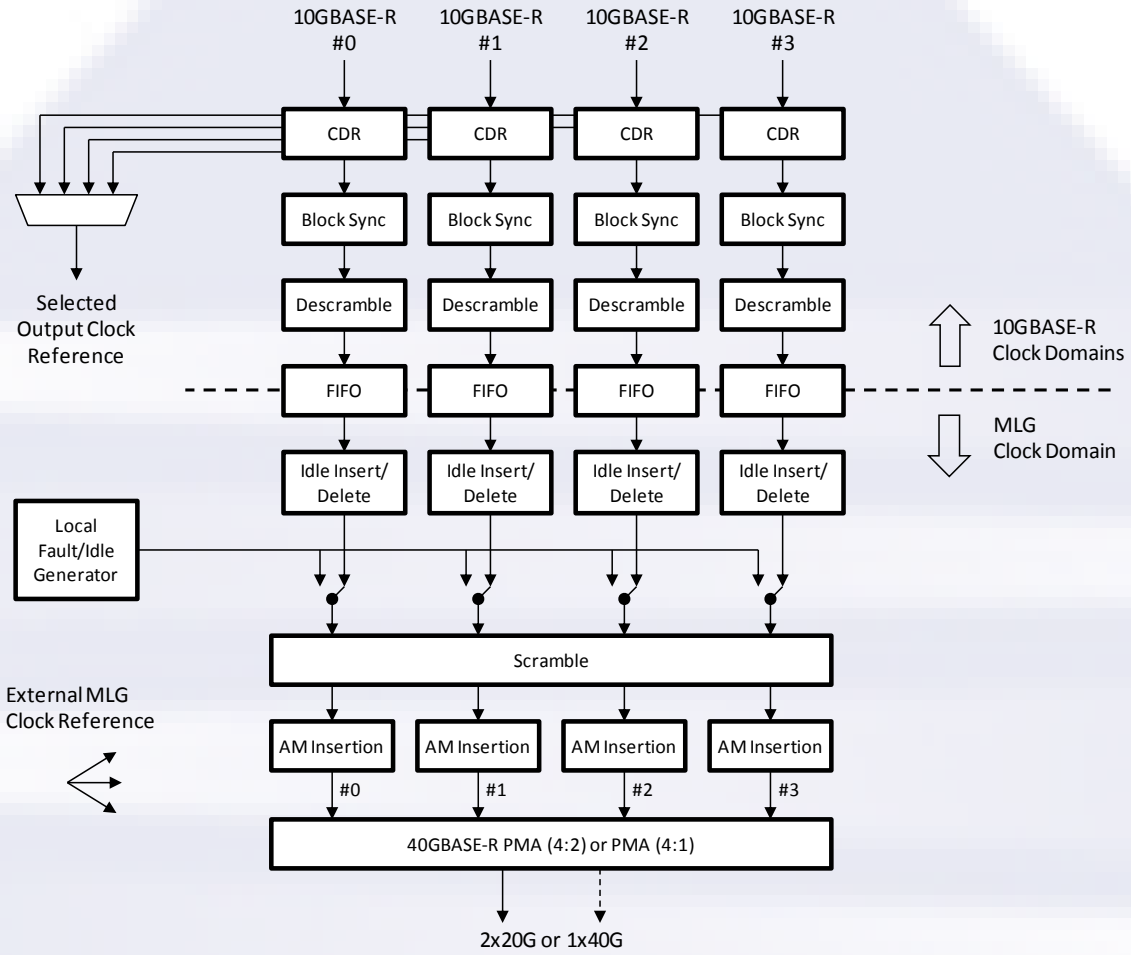


Figure 13: MLG40 10GBASE-R Mux Block Diagram

6.2.1.1 Clock and Data Recovery (CDR)

Each of the input 10GBASE-R signals to the MLG40 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.

6.2.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

6.2.1.3 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.

6.2.1.4 Idle Insert/Delete

Idles are inserted or deleted as necessary in order that the bitstream, after MLG40 lane marker insertion, will match a common MLG40 clock reference. Idle insertion and deletion shall comply with the rules defined in IEEE Std 802.3-2012 Clause 49.2.4.7.

An external MLG40 clock reference is provided as an input to the MLG40 mux block. The external MLG40 clock reference may be sourced from either a local free-running oscillator (± 100ppm), or from a system level BITS/SyncE timing architecture. The exact implementation of the external MLG40 clock reference is a system architecture decision, and outside the scope of this IA.

6.2.1.5 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 6.2.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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: 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.

6.2.1.6 Scrambled Idle Test Pattern Generation

The MLG40 mux may optionally generate for each 10G lane a scrambled idle test pattern. This test pattern can traverse a mux, and can be checked by the far-end MLG40 demux, or looped back at the MLG40 demux or Ethernet RX and checked at the near end MLG40 demux. An MLG40 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 6.2.1.7). 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.

6.2.1.7 Scramble

The four 10GBASE-R signals are 66-bit block interleaved, from lane 0 up to lane 3, to create the 40G MLG40 block stream for scrambling, similar to 40GBASE-R. The non-sync header bits of the MLG40 block stream are scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6. Once scrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG40 PCS lanes. This lane interleave, scramble, and block distribution function must maintain the mapping of 10GBASE-R signals 0 to 3 to MLG40 PCS lanes 0 to 3, respectively. For clarity, this process is illustrated in Figure 14.

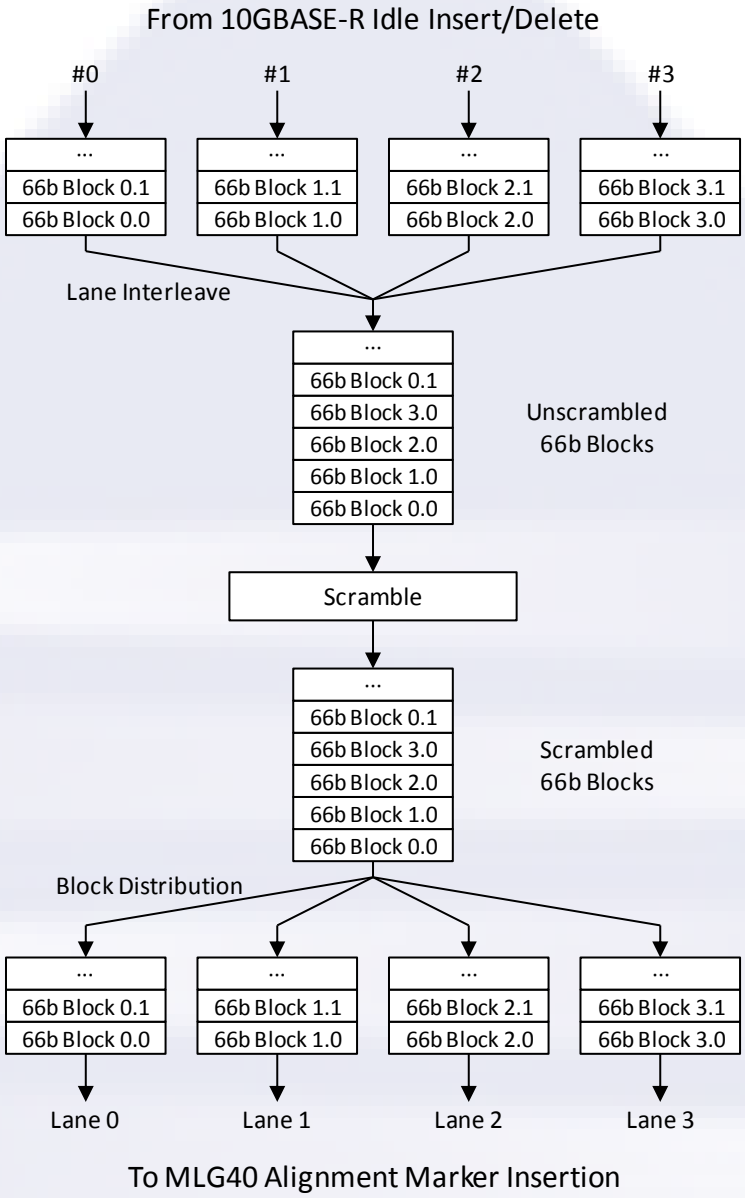


Figure 14: MLG40 Scrambler Diagram and 66b Block Flow

6.2.1.8 Alignment Marker Insertion

In order to support deskew and reordering of the MLG40 lanes into the constituent 10GBASE-R signals at the MLG40 demux, alignment markers are added periodically to each MLG40 lane. Each alignment marker is a special 66-bit block. The alignment markers shall be inserted after every 16383 data blocks on each MLG40 lane in the same way (at the same time on all four MLG40 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. These values are distinct from the PCS lane alignment markers used for 40GBASE-R, so there is no possibility of mistaking four 10GBASE-R signals for a 40GBASE-R link or vice-versa. The BIP_3 and BIP_7 fields of each MLG40 lane alignment marker are calculated over all of the previous bits on the

given MLG40 lane, from and including the previous MLG40 lane alignment marker but not the current MLG40 lane alignment marker per the procedures described for PCS lanes in Clause 82.2.8 of IEEE Std 802.3-2012.

6.2.1.9 PMA (4:2) or PMA (4:1)

The processes described in Clauses 6.2.1.1 through 6.2.1.8 will produce four MLG40 lanes, each with a bit rate of 10.3125 Gb/s ±100ppm (all MLG40 lanes locked to the common clock source). These MLG40 lanes can now be combined onto physical lanes as if they were 40GBASE-R PCS lanes using a PMA (4:2) or PMA (4:1). A PMA (4:2) will produce a 2-lane by 20.625 Gb/s interface, where each of these physical lanes carries two of the four 10G MLG40 lanes, bit multiplexed. A PMA (4:1) will produce a single lane 41.25 Gb/s interface, which carries all four of the 10G MLG40 lanes, bit multiplexed. Note that an implementation may support generation of any of the optional test patterns specified for a 40GBASE-R PMA across the MLG40 link: see IEEE Std 802.3-2012 Clause 83.5.10. If supported, the corresponding MDIO control variables must also be supported.

6.2.2 10GBASE-R Demux

The process for demultiplexing 10GBASE-R signals from MLG40 lanes is shown in Figure 15.

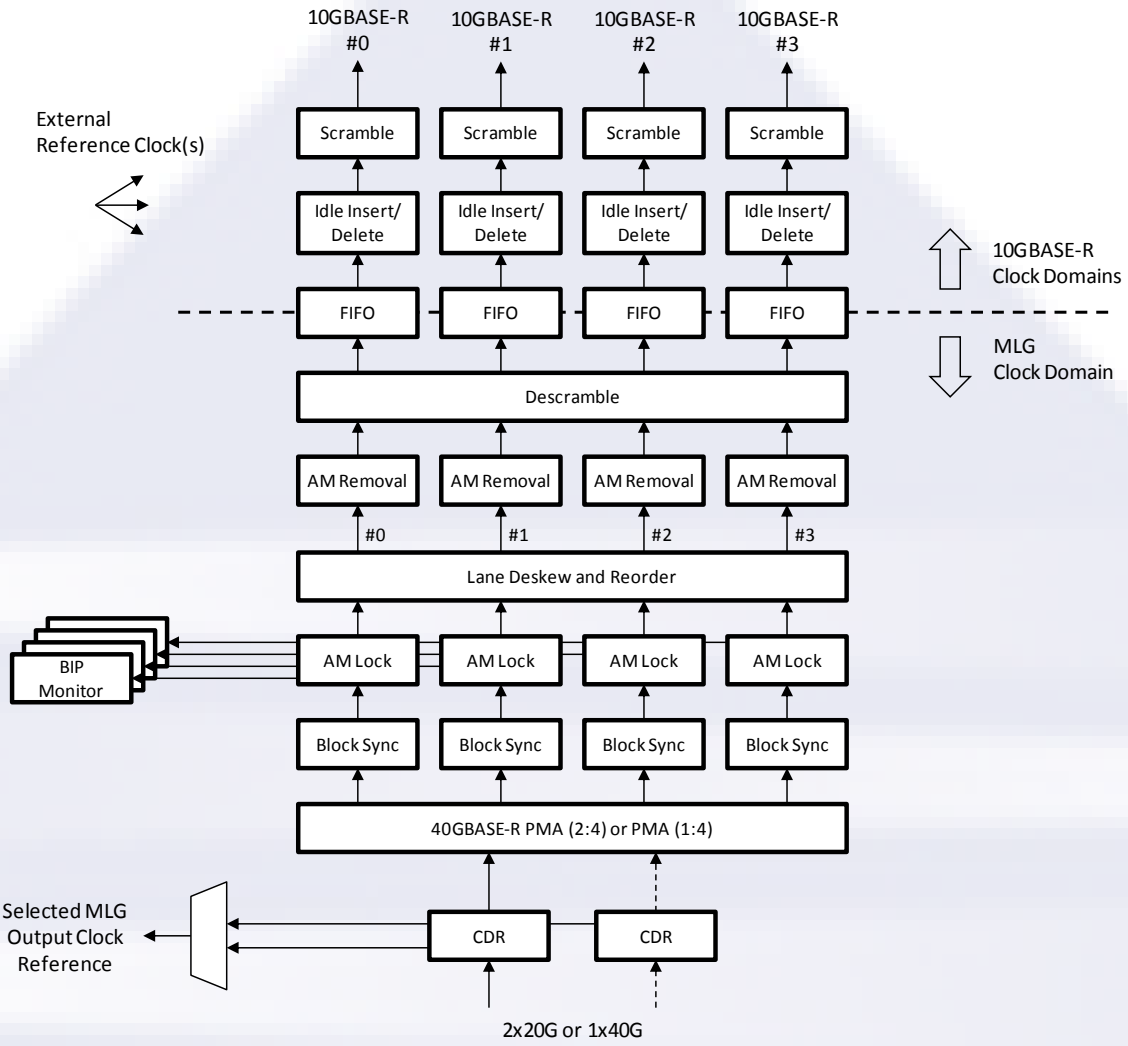


Figure 15: MLG40 10GBASE-R Demux Block Diagram

6.2.2.1 PMA (2:4) or PMA (1:4)

The input to the MLG40 demux will come from a physical interface that is similar to that of 40GBASE-R application. For the most common MLG40 configuration, this is expected to be a 2-lane by 20.625 Gb/s interface or single lane of 41.25 Gb/s, where each of these physical lanes carries 2 or 4 of the four MLG40 lanes, bit multiplexed. The PMA (2:4) or PMA (1:4) will demultiplex the MLG40 lanes as if they were PCS lanes into four individual bit streams of 10.3125 Gb/s ±100ppm. Note that an implementation may support detection of any of the optional test patterns specified for a 40GBASE-R PMA across the MLG40 link: see IEEE Std 802.3-2012 Clause 83.5.10. If supported, the corresponding MDIO control and status variables must also be supported.

6.2.2.2 MLG40 Lane Block Sync

66-bit block sync is obtained on each of the MLG40 lanes in the same manner as 40GBASE-R PCS lanes using the state diagram in Figure 82-10 of IEEE Std 802.3-2012.

6.2.2.3 MLG40 Lane Alignment Lock

Alignment marker lock is obtained on each MLG40 lane in the same manner as 40GBASE-R PCS lanes using the state diagram in Figure 82-11 of IEEE Std 802.3-2012, using the expected MLG40 lane alignment marker values from Table 1.

6.2.2.4 BIP Monitor

The expected BIP value is compared to the received BIP value in each received alignment marker and BIP error counters for each MLG40 lane are incremented as described for PCS lanes in Clause 82.2.14 of IEEE Std 802.3-2012.

6.2.2.5 MLG40 Lane Reorder

Once alignment marker lock has been obtained on the MLG40 lanes, the MLG40 lanes are reordered to the appropriate 10GBASE-R lane per the AM mapping of Table 1.

6.2.2.6 MLG40 Lane Alignment Marker Removal

The alignment markers in each MLG40 lane are removed in the same way as for 40GBASE-R PCS lane alignment markers described in Clause 82.2.14 of IEEE Std 802.3-2012.

6.2.2.7 Descramble

The four aligned MLG40 PCS lanes are 66-bit block interleaved, from lane 0 up to lane 3, to create the 40G MLG40 block stream for descrambling, similar to 40GBASE-R. The non-sync header bits of the MLG40 block stream are 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. Once descrambled, the 66-bit blocks are distributed round-robin to the appropriate 10GBASE-R signal. This lane interleave, descramble, and block distribution function must maintain the mapping of MLG40 PCS lanes 0 to 3 to 10GBASE-R signals 0 to 3, respectively. For clarity, this process is illustrated in Figure 16.

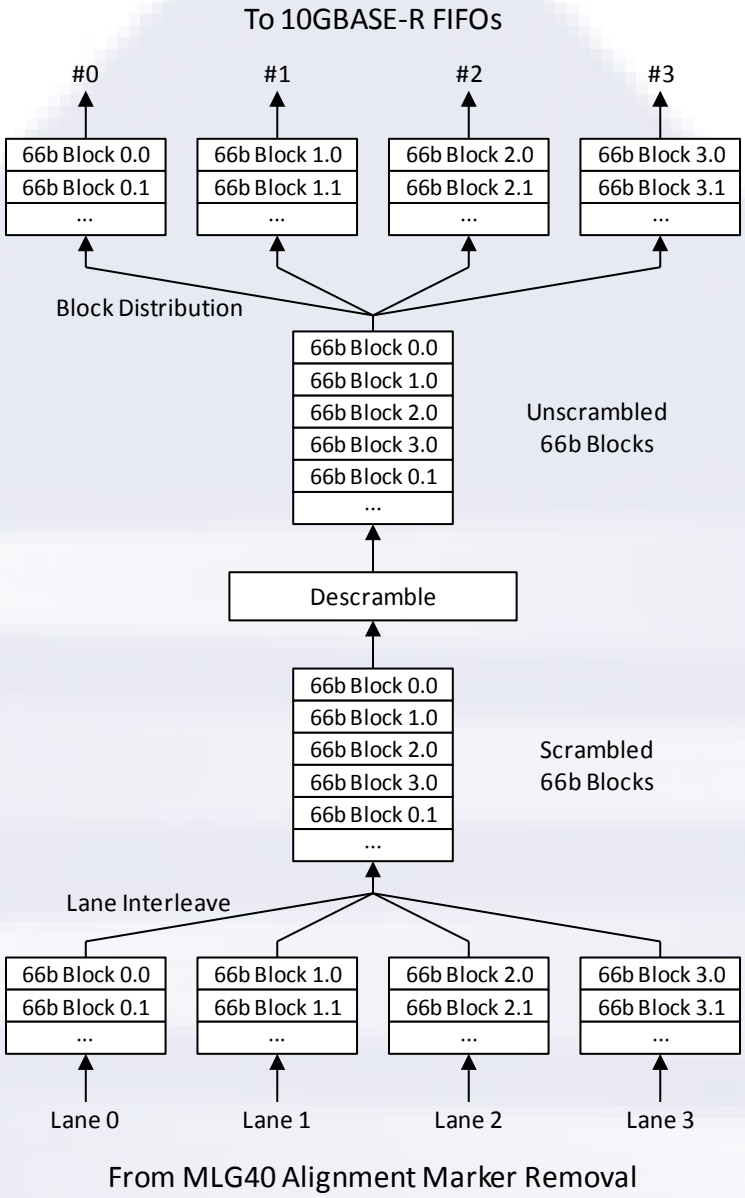


Figure 16: MLG40 Descrambler Diagram and 66b Block Flow

6.2.2.8 Scrambled Idle Test Pattern Checker

The MLG40 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.

6.2.2.9 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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 input signals to be carried across a MLG40 link (and used to clock one or more of the 10GBASE-R output signals).

6.2.2.10 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 MLG100

7.1 MLG100 General Mechanism

The MLG100 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 MLG100 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 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 7.1.1.

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 rescrambling the non-lane marker blocks.

7.1.1 MLG100 Lane Markers

The MLG100 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 2 provides the MLG lane alignment markers used for the MLG100 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 are identical to MLG 2.0, and, when all MLG lanes of a MLG100 application are provisioned to carry 10GBASE-R, are identical to those for MLG 1.0, so these implementations can be interconnected with appropriate provisioning.

If Remote Management (RM) is enabled, the fields M0, M1, M2, BIP3 are swapped with the fields M4, M5, M6, BIP7 in the AM for MLG lane 0.0 (0.0-10 or 0.0-40). This is a similar format as used for Rapid Alignment Markers, as shown in Figure 82-9b of IEEE 802.3bj-2014.

Table 2: MLG100 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[b]	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[b]	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	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

[a] Each octet is transmitted LSB to MSB.

[b] The upper and lower half of the AM Encoding is reversed on MLG lane 0.0 if RM channel is in data stream.

7.1.2 MLG100 RS-FEC

Some 4x25G physical links require the signals to be encoded with the Reed-Solomon Forward Error Correction (RS-FEC) sublayer, as specified in IEEE Std 802.3bj-2014

Clause 91. If RS-FEC is required, the 20 MLG lanes of an MLG100 application are mapped to the 100G RS-FEC sublayer of IEEE Std 802.3bj-2014, Clause 91. A 100G scrambler is used to scramble the MLG data stream to support the 256b/257b inverse transcoding within the RS-FEC sublayer.

The MLG AM mapping to RS-FEC lanes is shown in Figure 17. As done in Clause 91, the Clause 82 AM0 and AM16 are replicated across the all four RS-FEC lanes, and AM4 through AM7 are used to denote RS-FEC lanes 0 through 3. The MLG AMs are then mapped into the remaining section, and can be used to distinguish the specific 10GE/40GE client and Remote Management configurations.

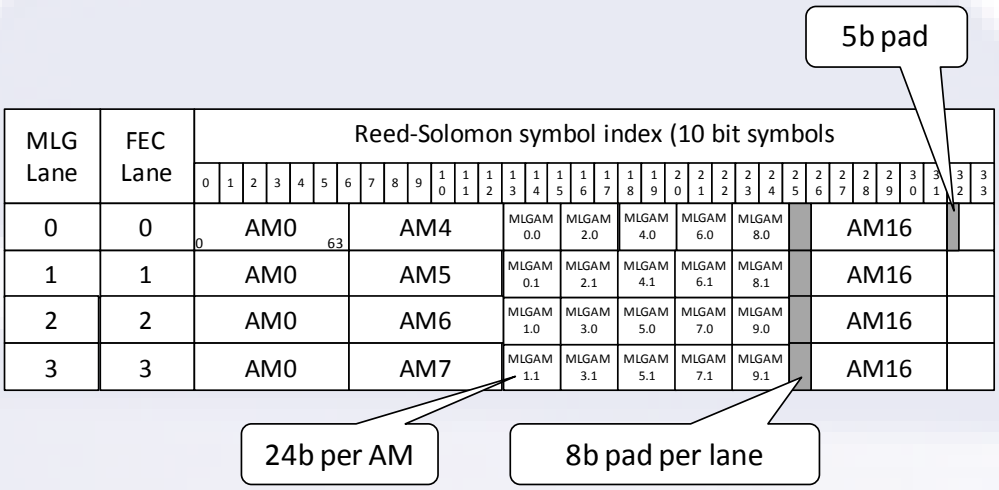


Figure 17: MLG100 RS-FEC AM mapping

The 5b pad of Figure 17 is added at the end of the alignment marker mapping as described in Section 91.5.2.6 of IEEE Std 802.3bj-2014. The 8b pad is added per lane in the same manner, except with the alternating 8b pattern “01100101” and “10011010”.

Each MLG AM includes a BIP field. The mapping of these fields into RS-FEC lanes is shown in Figure 18.

0	AM0	BIP 0.0		BIP 2.0	AM4	BIP 4.0		BIP 6.0		AM16	BIP 8.0		!BIP 8.0
1	AM0	BIP 0.1		BIP 2.1	AM5	BIP 4.1		BIP 6.1		AM16	BIP 8.1		!BIP 8.1
2	AM0	BIP 1.0		BIP 3.0	AM6	BIP 5.0		BIP 7.0		AM16	BIP 9.0		!BIP 9.0
3	AM0	BIP 1.1		BIP 3.1	AM7	BIP 5.1		BIP 7.1		AM16	BIP 9.1		!BIP 9.1

Figure 18: MLG100 RS-FEC BIP mapping

Alignment lock, deskew, and lane reordering are done within the four RS-FEC lanes (0-3), as defined in Sections 91.5.3.1 and 91.5.3.2 of IEEE Std 802.3bj-2014.

7.1.3 MLG100 Remote Management Channel

The optional Remote Management (RM) channel is designed to carry maintenance information, which may include status, control, software images, or other information deemed appropriate by the two end points. The format of data on the RM channel is governed by RFC 1662, which describes PPP in HDLC-like framing. This RFC describes the octet synchronous HDLC to be used by the RM channel. The flag, address, control and escape octets are as described in this RFC. The RM channel uses the 32-bit FCS as described in this RFC. The 16-bit FCS option is not used. Data transparency is

maintained as described in the RFC. The format of the data carried over this PPP communication channel is not further described in this IA. The two endpoints must agree to the data format for successful communications.

RM data is 64b/66b encoded as data blocks, per IEEE Std 802.3-2012 Clause 82, Figure 82-5. The unscrambled 66b RM data blocks are inserted into the unscrambled MLG data stream in a similar manner as AM blocks, except they are scrambled with MLG data. They are simultaneously striped across all 20 MLG lanes, as shown in Figure 19. The RM blocks are spaced at a distance of 16384 from each other, and at a distance of 8192 from MLG AM blocks. The resulting RM channel has a raw bandwidth of ~6.1 Mb/s.

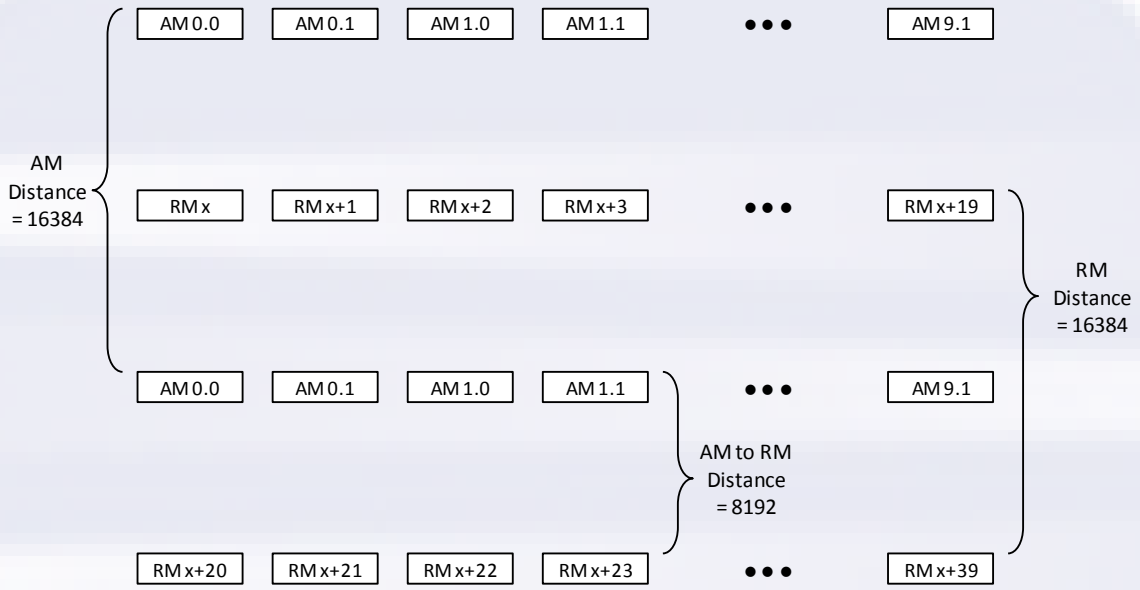


Figure 19: MLG100 RM Data Channel

7.2 MLG100 Detailed Block Diagrams (without RS-FEC and without RM)

The MLG100 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 composed 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 7.1.1, while the other 4 MLG lanes can carry two 10GBASE-Rs.

The MLG demux receives twenty MLG lanes for the MLG100 application. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

7.2.1 MLG100 mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 20.

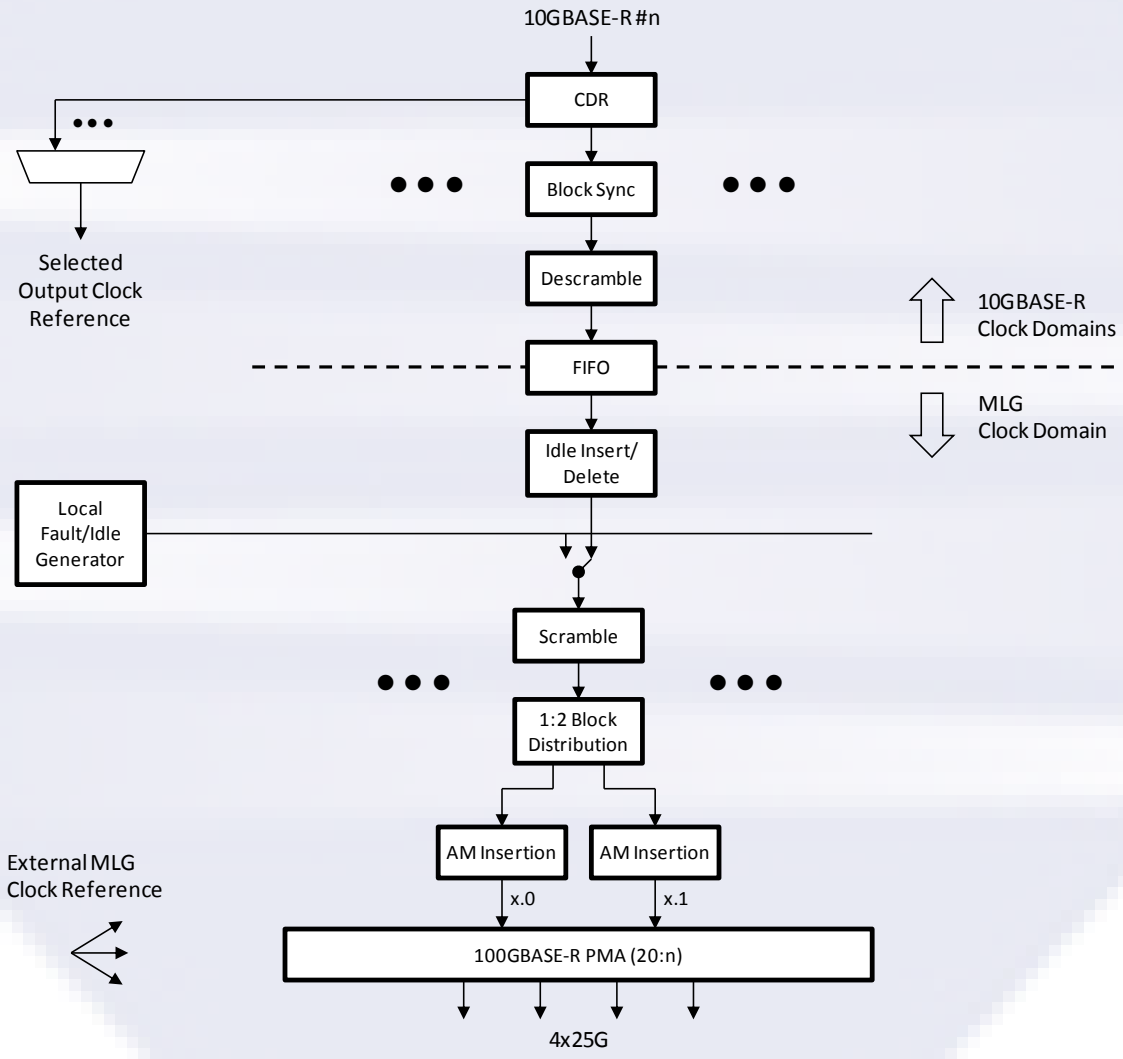


Figure 20: MLG100 mux 10GBASE-R mapping Block Diagram

7.2.1.1 Clock and Data Recovery (CDR)

Each of the 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.

7.2.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

7.2.1.3 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.

7.2.1.4 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 49.2.4.7.

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 (± 100ppm), 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.

7.2.1.5 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 7.2.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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.

7.2.1.6 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 7.2.1.7). 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.

7.2.1.7 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.

7.2.1.8 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 twenty PCS lanes.

7.2.1.9 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 2 for a MLG100 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 BIP_3 and BIP_7 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.

7.2.1.10 100GBASE-R PMA(20:n)

The processes described in Clauses 7.2.1.1 through 7.2.1.9 will produce twenty MLG lanes, each with a bit rate of 5.15625 Gb/s ±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. The most typical value of “n” is expected to be 4 for the MLG100 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. 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.

7.2.2 MLG100 mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 21. These may be eight of the twenty MLG lanes of a MLG100.

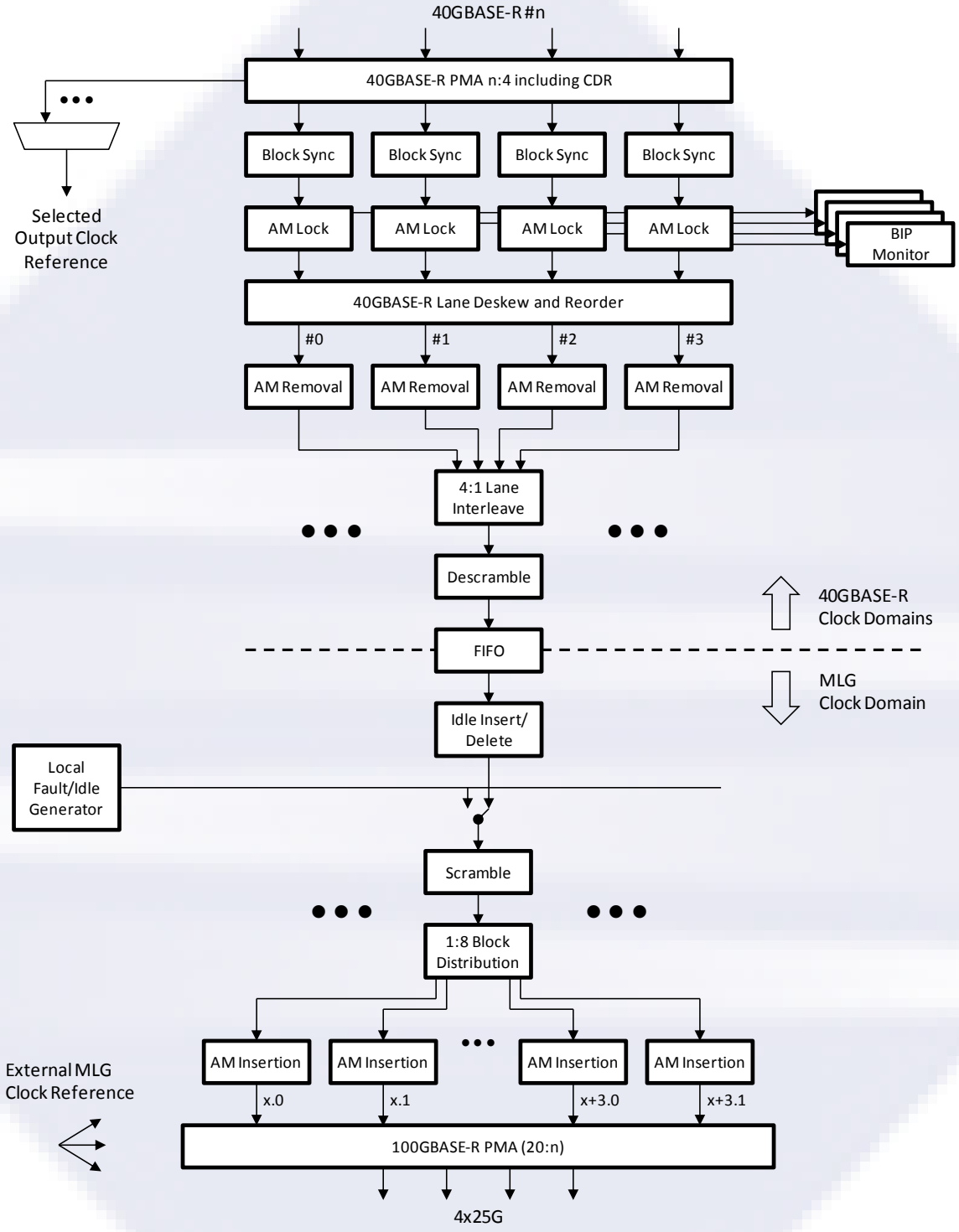


Figure 21: MLG100 mux 40GBASE-R mapping Block Diagram

7.2.2.1 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.

7.2.2.2 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.

7.2.2.3 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.

7.2.2.4 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.

7.2.2.5 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.

7.2.2.6 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.

7.2.2.7 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 (± 100ppm), 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.

7.2.2.8 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 7.2.2.1), 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 D_1, D_2, D_3, and zeros in Z_4, Z_5, Z_6, Z_7. 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.

7.2.2.9 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. A 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 7.2.2.10). 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.

7.2.2.10 Scramble

After idle insertion/deletion and Local Fault insertion, the resulting stream is re-scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6.

7.2.2.11 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 twenty PCS lanes.

7.2.2.12 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 2 for a MLG100 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 BIP_3 and BIP_7 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.

7.2.2.13 100GBASE-R PMA(20:n)

See Section 7.2.1.10. 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.

7.2.3 MLG100 demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 22.

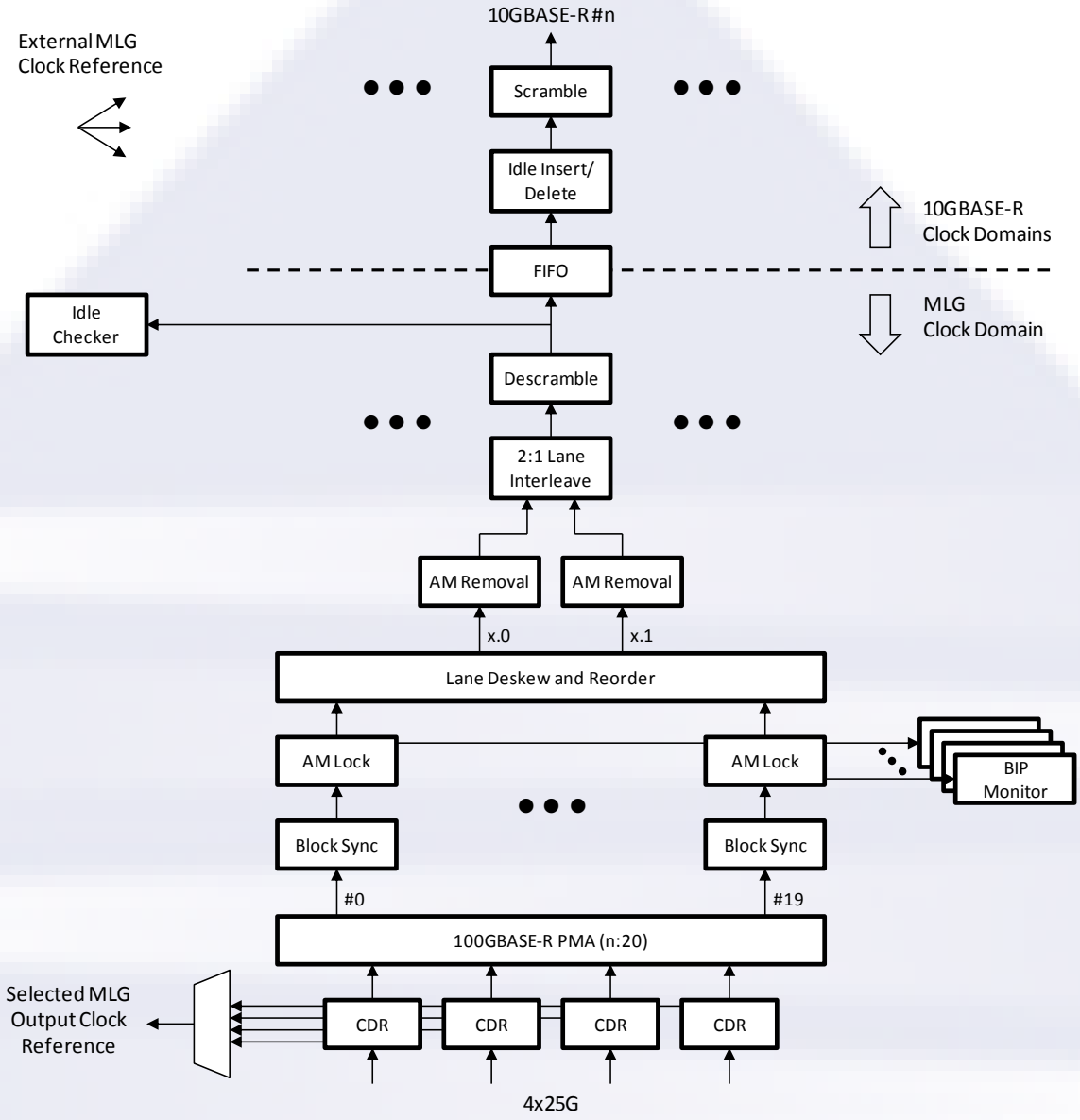


Figure 22: MLG100 demux 10GBASE-R demapping Block Diagram

7.2.3.1 100GBASE-R PMA(n:20)

The input to the MLG demux will come from a physical interface that is similar to that of 100GBASE-R for the MLG100 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 are possible. The PMA (n:20) will demultiplex the MLG lanes as if they were PCS lanes into twenty individual bit streams of 5.15625 Gb/s ±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.

7.2.3.2 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.

7.2.3.3 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 2 for the 4x25G MLG100 application.

7.2.3.4 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.

7.2.3.5 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.

7.2.3.6 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.

7.2.3.7 MLG Lane Interleave

The pairs of MLG lanes carrying each 10GBASE-R signal are 66-bit block interleaved to reconstitute the original 10GBASE-R signals.

7.2.3.8 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.

7.2.3.9 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.

7.2.3.10 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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).

7.2.3.11 Scramble

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

7.2.4 MLG100 demux for 40GBASE-R

The process for demapping a 40GBASE-R from eight MLG lanes is shown in Figure 23.

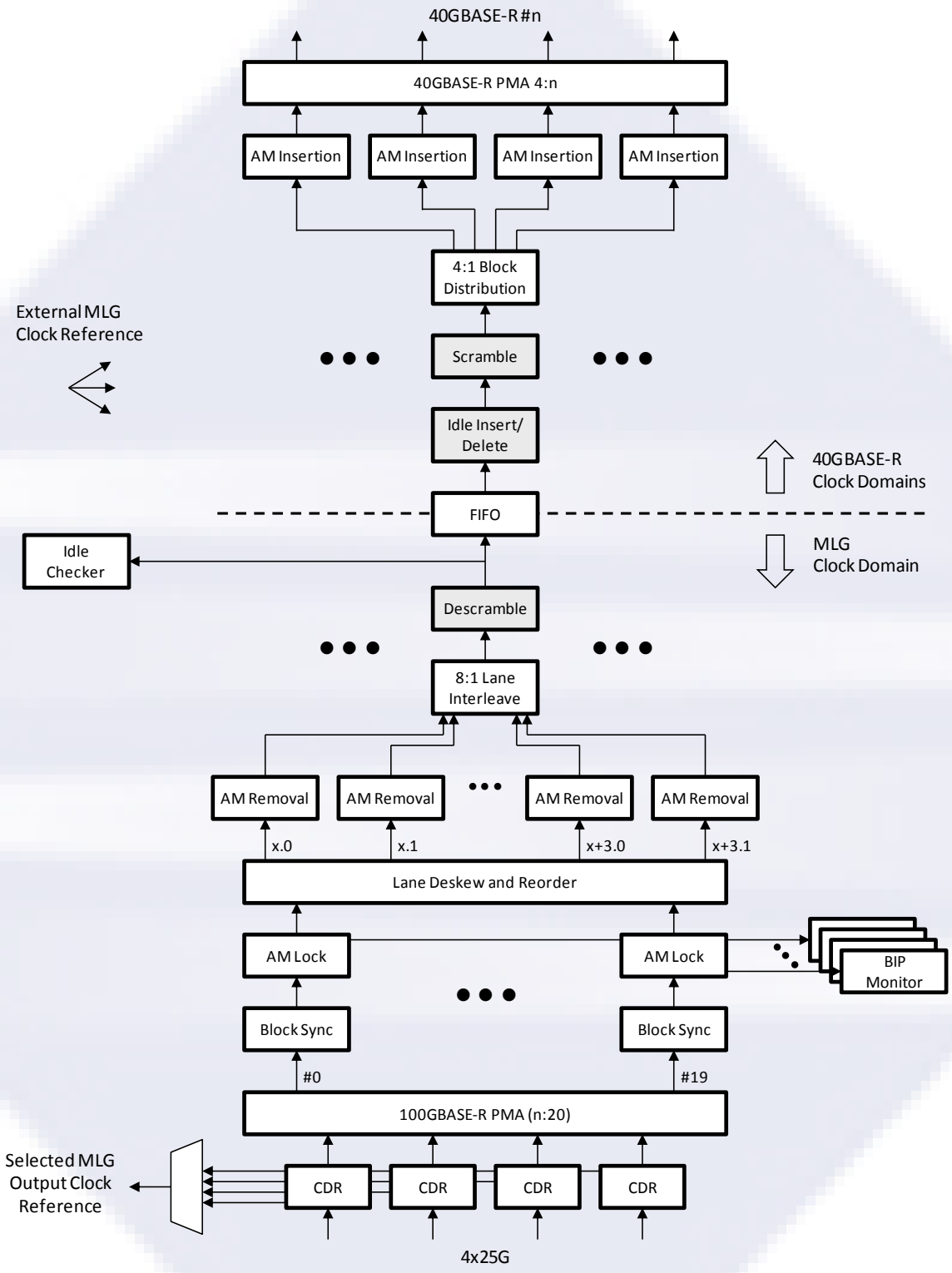


Figure 23: MLG100 demux 40GBASE-R demapping Block Diagram

7.2.4.1 100GBASE-R PMA(n:20)

See Section 7.2.3.1. 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.

7.2.4.2 MLG Lane block sync

See Section 7.2.3.2. 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.

7.2.4.3 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 2 for the MLG100 application.

7.2.4.4 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.

7.2.4.5 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.

7.2.4.6 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.

7.2.4.7 MLG Lane Interleave

The eight 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.

7.2.4.8 Descramble

Note that there are three shaded boxes in Figure 23: 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 7.2.4.11 if a different clock domain is not needed for the 40GBASE-R.

When required, the non-sync header bits of the 40GBASE-R signal are 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 of IEEE Std 802.3-2012.

7.2.4.9 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 may 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 (±100ppm), 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).

7.2.4.10 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.

7.2.4.11 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.

7.2.4.12 Alignment Insertion

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

7.2.4.13 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.

7.3 MLG100 Detailed Block Diagrams (without RS-FEC and with RM)

The MLG100 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 7.1.1, while the other 4 MLG lanes can carry two 10GBASE-Rs.

The MLG demux receives twenty MLG lanes for the MLG100 application. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

7.3.1 MLG100 mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 24.

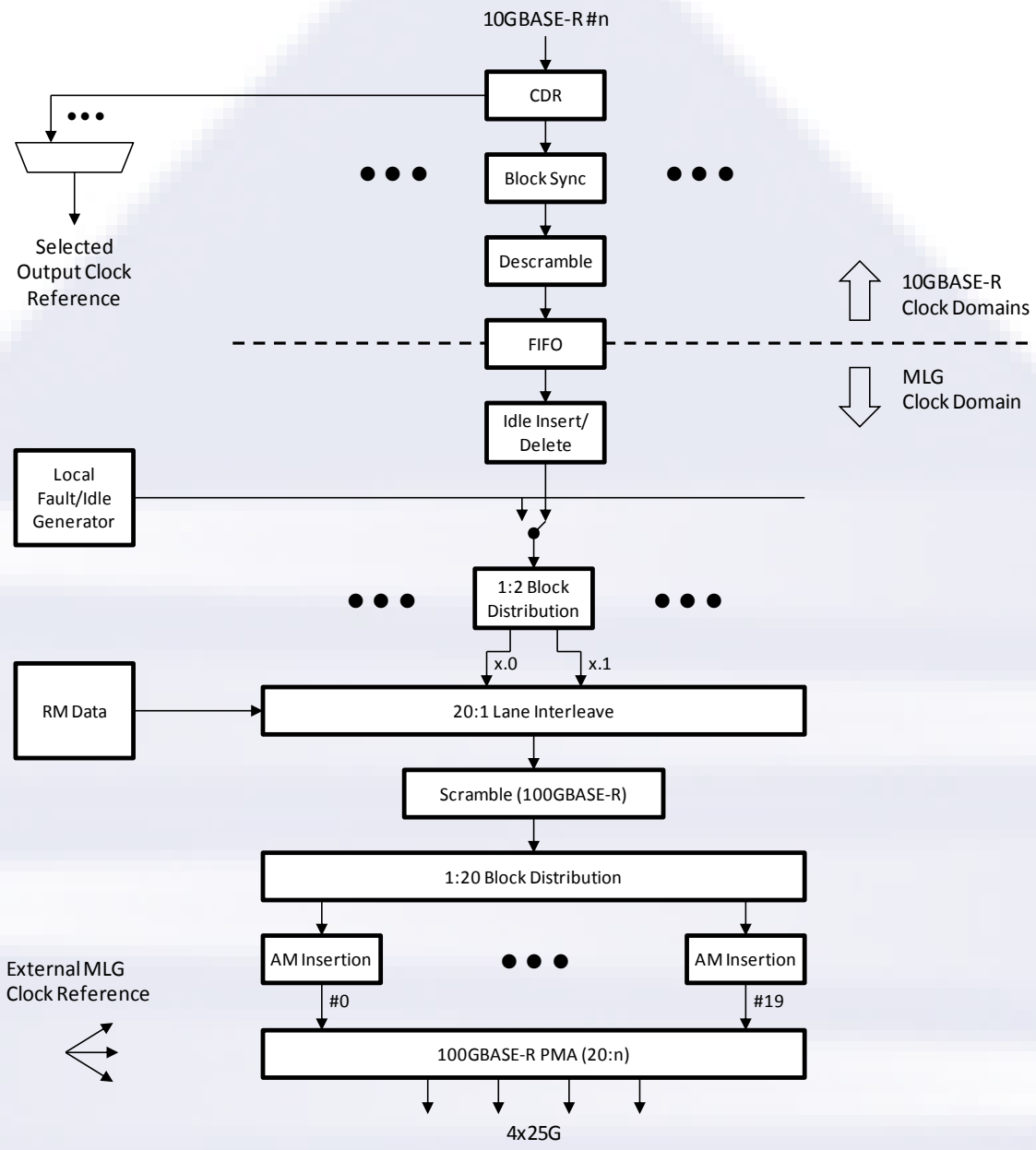


Figure 24: MLG100 mux 10GBASE-R mapping Block Diagram

7.3.1.1 Clock and Data Recovery (CDR)

Each of the 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.

7.3.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

7.3.1.3 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.

7.3.1.4 Idle Insert/Delete

Idles are inserted or deleted as necessary in order that the bitstream, after MLG lane marker and RM channel data 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 49.2.4.7.

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 (± 100ppm), 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.

7.3.1.5 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 7.3.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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.

7.3.1.6 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 7.3.1.9). 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.

7.3.1.7 2:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

7.3.1.8 RM data insertion and 20:1 Lane Interleave

After 10GBASE-R and 40GBASE-R block distribution to MLG lanes, the RM channel data blocks are inserted periodically into each MLG lane, as shown in Figure 25. The 66b blocks shall be inserted 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 RM blocks shall be inserted after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

The twenty unscrambled MLG100 lanes, including RM data, are 66-bit block interleaved, from lane 0.0 up to lane 9.1, to create a 100G block stream, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

7.3.1.9 Scramble

The non-sync header bits of the MLG100 block stream are scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6.

7.3.1.10 1:20 Block Distribution

Once scrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG100 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, scrambling, and 1:20 block distribution mechanism must maintain the mapping of MLG lanes 0.0 to 9.1 to PCS lanes 0 to 19. For clarity, this process is illustrated in Figure 25.

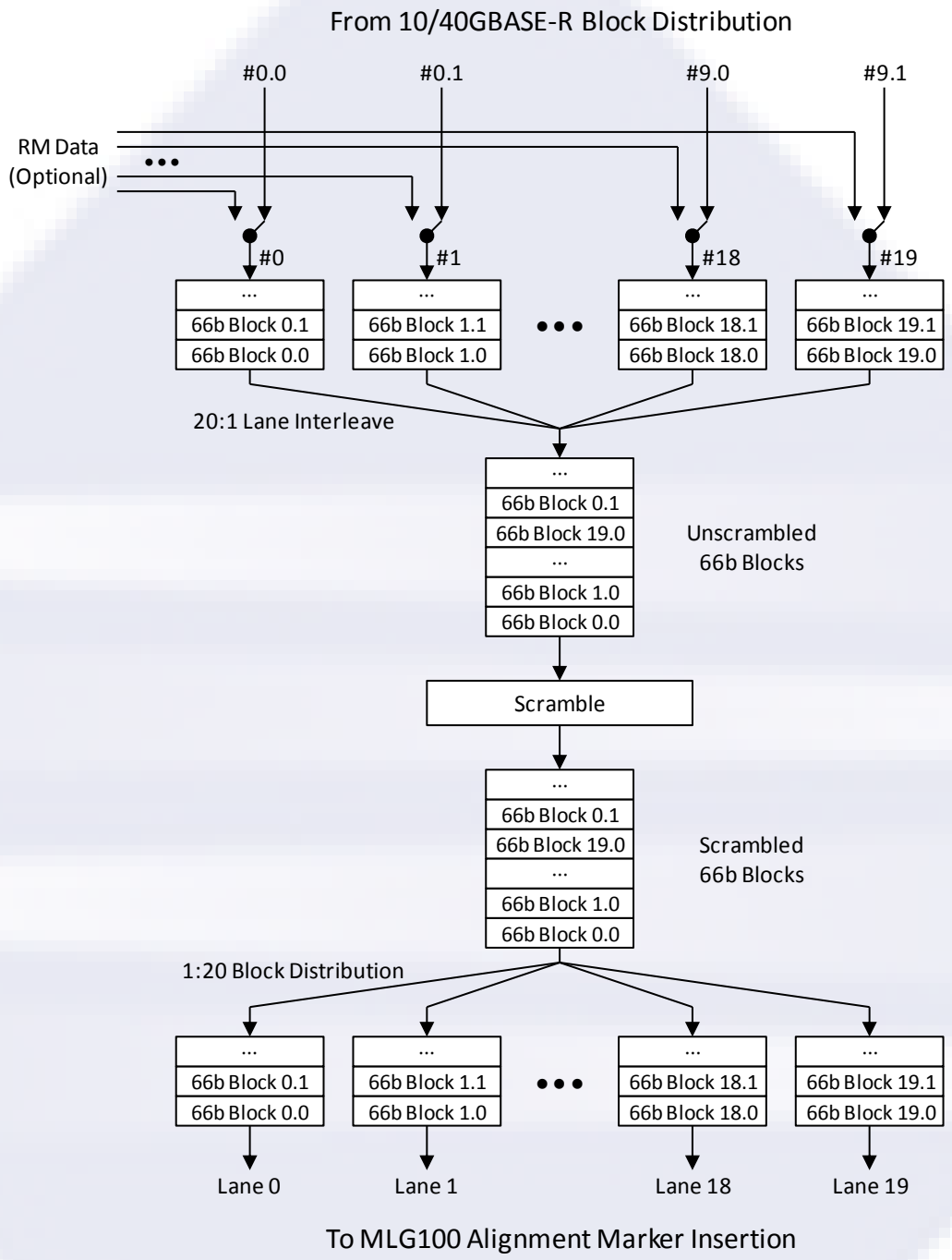


Figure 25: MLG100 Scrambler Diagram and 66b Block Flow

7.3.1.11 Alignment Marker Insertion

In order to support deskew and reordering of the MLG lanes into the constituent 10GBASE-R or 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 2 for a MLG100 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 BIP_3 and BIP_7 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.

7.3.1.12 100GBASE-R PMA(20:n)

The processes described in Clauses 7.2.1.1 through 7.2.1.9 will produce twenty MLG lanes, each with a bit rate of 5.15625 Gb/s ±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. The most typical value of “n” is expected to be 4 for the MLG100 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. 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.

7.3.2 MLG100 mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 21. These may be eight of the twenty MLG lanes of a MLG100.

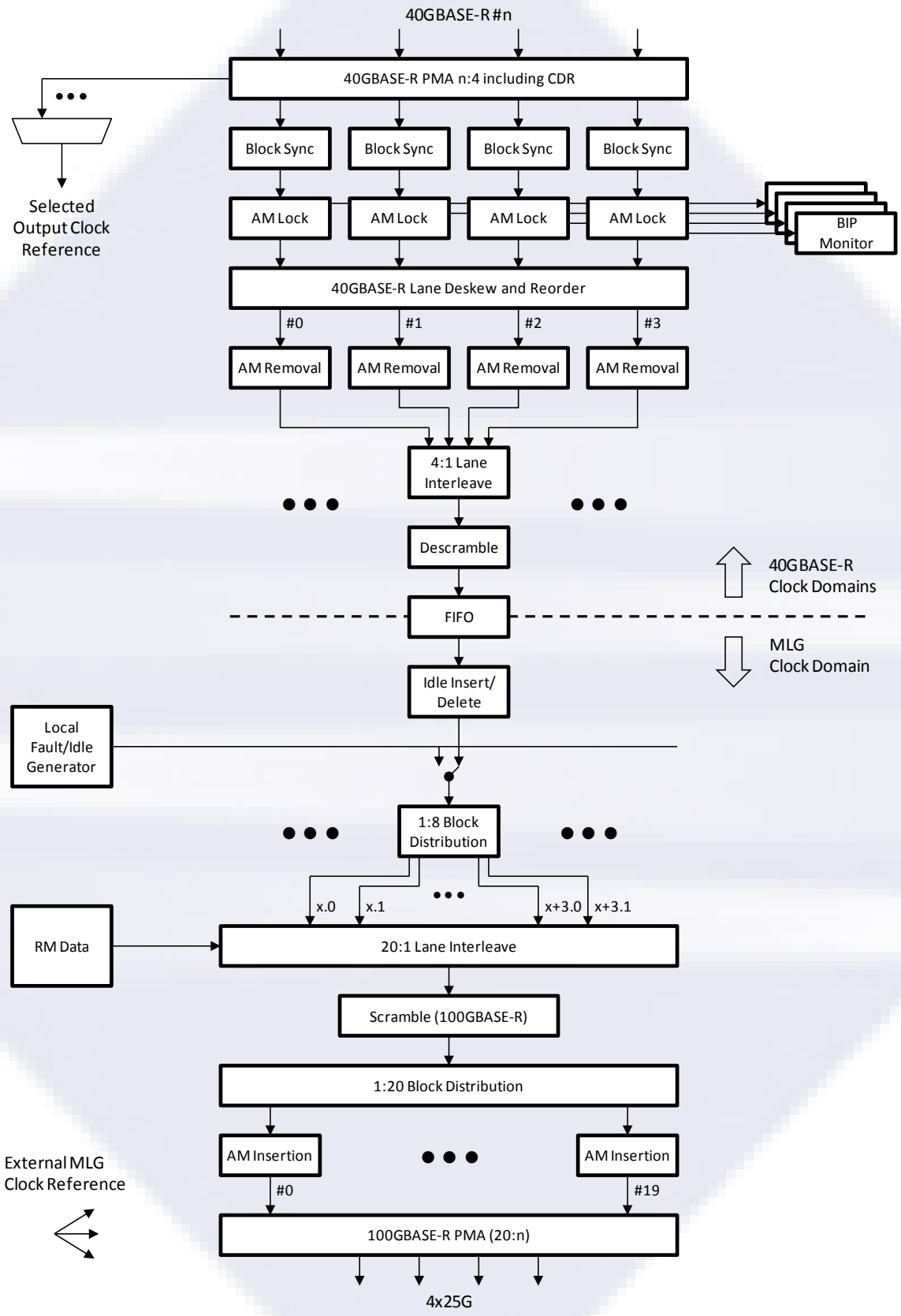


Figure 26: MLG100 mux 40GBASE-R mapping Block Diagram

7.3.2.1 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.

7.3.2.2 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.

7.3.2.3 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.

7.3.2.4 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.

7.3.2.5 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.

7.3.2.6 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.

7.3.2.7 Idle insert/delete

Idles are inserted or deleted as necessary in order that the bitstream, after PCS lane marker and RM channel data 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 (± 100ppm), 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.

7.3.2.8 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 7.3.2.1), 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 D_1, D_2, D_3, and zeros in Z_4, Z_5, Z_6, Z_7. 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.

7.3.2.9 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 7.3.2.12). 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.

7.3.2.10 8:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

7.3.2.11 RM data insertion and 20:1 Lane Interleave

See Section 7.3.1.8. 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.

7.3.2.12 Scramble

See Section 7.3.1.9. 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.

7.3.2.13 1:20 Block Distribution

See Section 7.3.1.10. 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.

7.3.2.14 Alignment Marker Insertion

See Section 7.3.1.11. 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.

7.3.2.15 100GBASE-R PMA(20:n)

See Section 7.3.1.12. 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.

7.3.3 MLG100 demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 27.

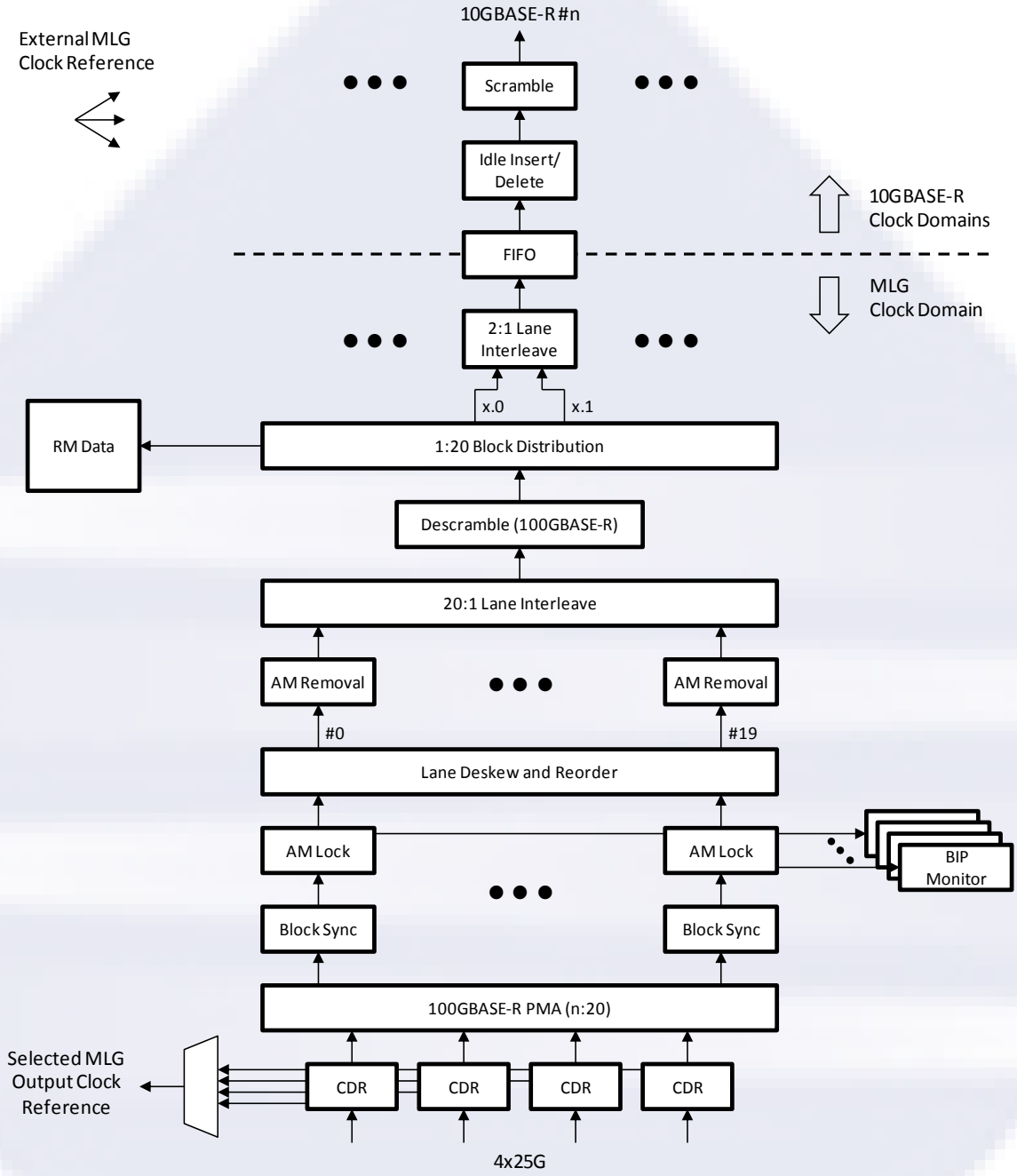


Figure 27: MLG100 demux 10GBASE-R demapping Block Diagram

7.3.3.1 100GBASE-R PMA (n:20)

The input to the MLG demux will come from a physical interface that is similar to that of 100GBASE-R for the MLG100 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 are possible. The PMA (n:20) will demultiplex the MLG lanes as if they were PCS lanes into twenty individual bit streams of 5.15625 Gb/s ±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.

7.3.3.2 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.

7.3.3.3 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 2 for the 4x25G MLG100 application.

7.3.3.4 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.

7.3.3.5 MLG Lane reorder

The twenty MLG lanes are deskewed and reordered according to the procedures described in IEEE Std 802.3-2012 Clauses 82.2.12 and 82.2.13. 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.

7.3.3.6 20:1 Lane Interleave

After AM removal, the twenty scrambled MLG100 lanes, including RM data, are 66-bit block interleaved, from lane 0 up to lane 19, to create a 100G block stream, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

7.3.3.7 Descramble

The non-sync header bits of the MLG100 block stream are 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.

7.3.3.8 1:20 Block Distribution RM data removal

Once descrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG100 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, descrambling, and 1:20 block distribution mechanism must maintain the mapping of PCS lanes 0 to 19 to MLG lanes 0.0 to 9.1. For clarity, this process is illustrated in Figure 28.

After block distribution to the twenty MLG lanes, the RM channel data blocks are removed periodically from each MLG lane, as shown in Figure 28. The 66b blocks shall be removed on each MLG lane in the same way (at the same time on all twenty MLG

lanes) after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

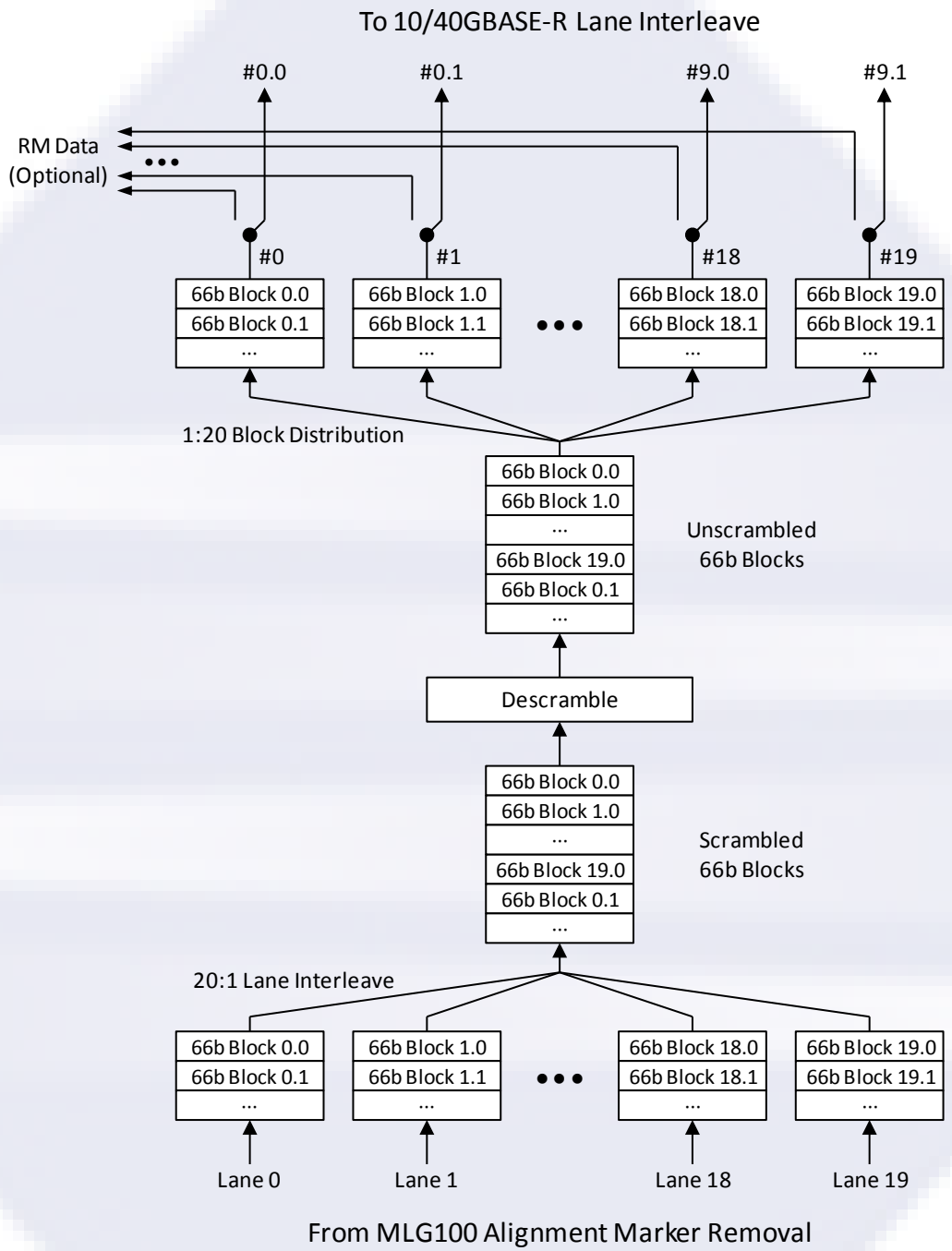


Figure 28: MLG100 Descrambler Diagram and 66b Block Flow

7.3.3.9 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.

7.3.3.10 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.

7.3.3.11 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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).

7.3.3.12 Scramble

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

7.3.4 MLG100 demux for 40GBASE-R

The process for demapping a 40GBASE-R from eight MLG lanes is shown in Figure 29.

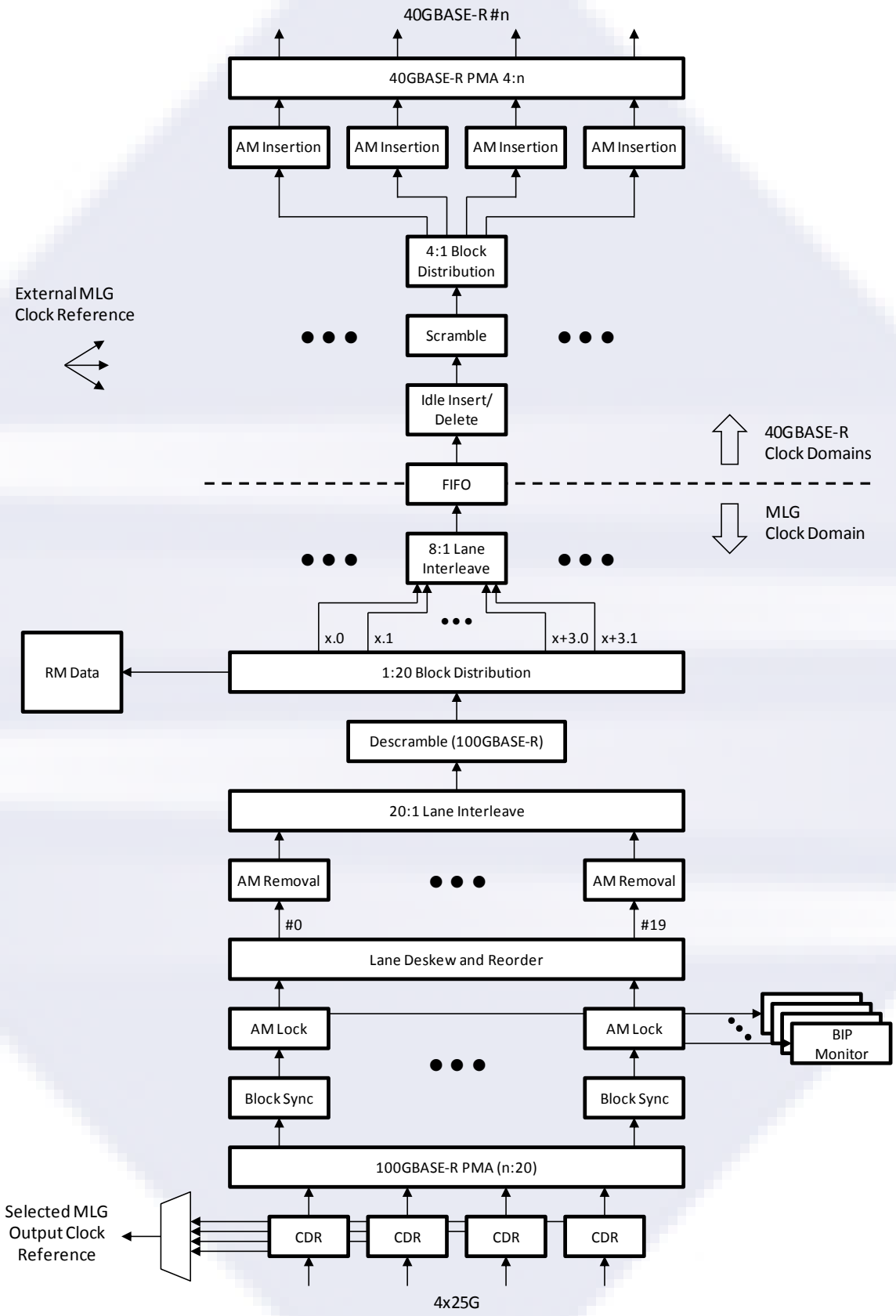


Figure 29: MLG100 demux 40GBASE-R demapping Block Diagram

7.3.4.1 100GBASE-R PMA (n:20)

See Section 7.3.3.1. 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.

7.3.4.2 MLG Lane block sync

See Section 7.3.3.2. 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.

7.3.4.3 MLG Lane Alignment lock

See Section 7.3.3.3. 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.

7.3.4.4 BIP monitor

See Section 7.3.3.4. 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.

7.3.4.5 MLG Lane reorder

See Section 7.3.3.5. 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.

7.3.4.6 MLG Lane Alignment marker removal

See Section 0. 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.

7.3.4.7 20:1 Lane Interleave

See Section 7.3.3.6. 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.

7.3.4.8 Descramble

See Section 7.3.3.7. 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.

7.3.4.9 1:20 Block Distribution RM data removal

See Section 7.3.3.8. 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.

7.3.4.10 MLG Lane Interleave

The eight 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.

7.3.4.11 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 may 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 (±100ppm), 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).

7.3.4.12 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.

7.3.4.13 4:1 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.

7.3.4.14 Alignment Insertion

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

7.3.4.15 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.

7.4 MLG100 Detailed Block Diagrams (with RS-FEC and optional RM)

The MLG100 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 7.1.1, while the other 4 MLG lanes can carry two 10GBASE-Rs.

The MLG demux receives twenty MLG lanes for the MLG100 application. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

7.4.1 MLG100 mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 30.

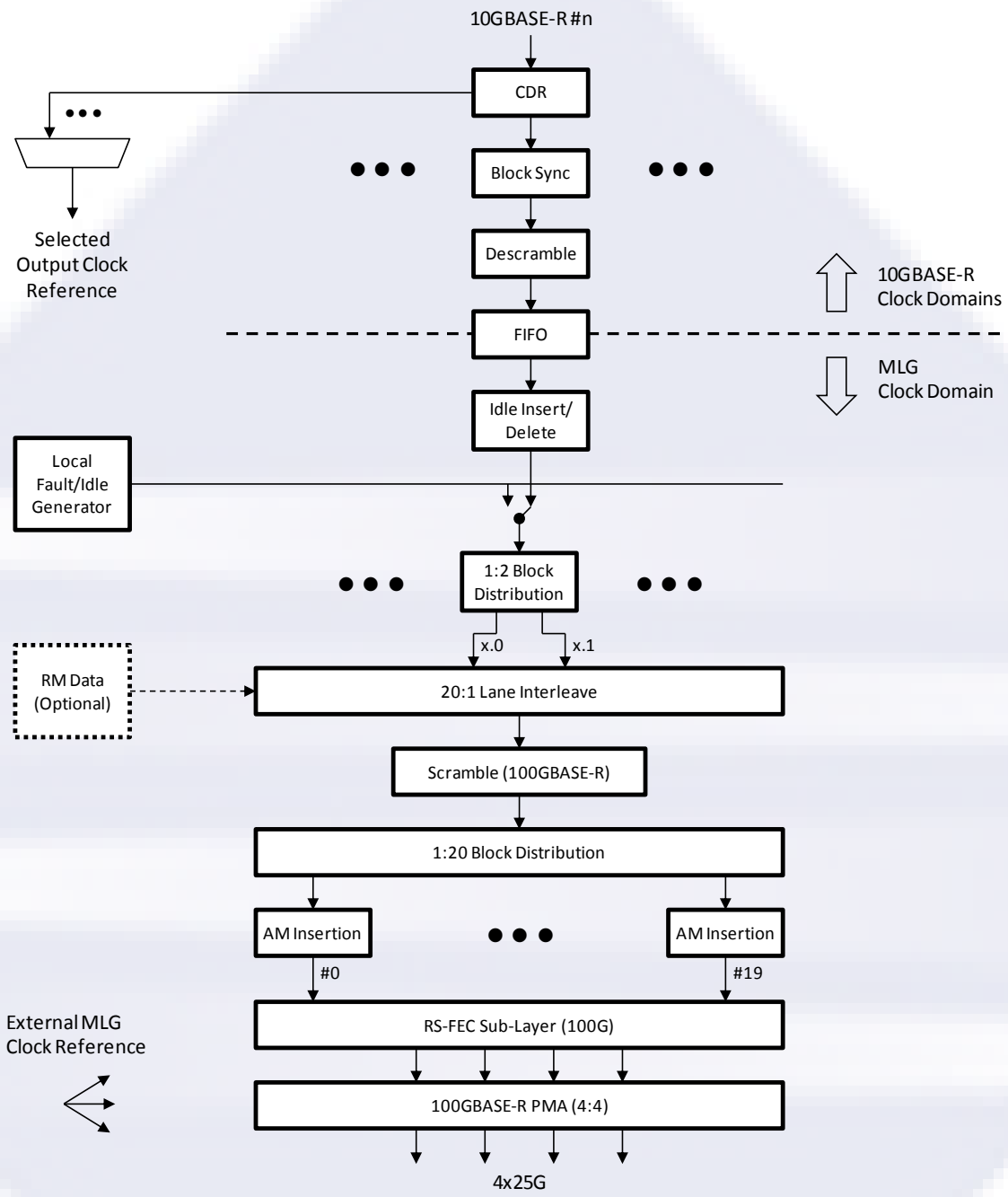


Figure 30: MLG100 mux 10GBASE-R mapping Block Diagram

7.4.1.1 Clock and Data Recovery (CDR)

Each of the 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.

7.4.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

7.4.1.3 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.

7.4.1.4 Idle Insert/Delete

Idles are inserted or deleted as necessary in order that the bitstream, after MLG lane marker and optional RM channel data 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 49.2.4.7.

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 (± 100ppm), 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.

7.4.1.5 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 7.4.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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.

7.4.1.6 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 7.4.1.9). 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.

7.4.1.7 2:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

7.4.1.8 RM data insertion and 20:1 Lane Interleave

After 10GBASE-R and 40GBASE-R block distribution to MLG lanes, if RM channel is enabled, the RM channel data blocks are inserted periodically into each MLG lane, as shown in Figure 25. The 66b blocks shall be inserted 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 RM blocks shall be inserted after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

The twenty unscrambled MLG100 lanes, including RM data, if enabled, are 66-bit block interleaved, from lane 0.0 up to lane 9.1, to create a 100G block stream, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

7.4.1.9 Scramble

The non-sync header bits of the MLG100 block stream are scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6.

7.4.1.10 1:20 Block Distribution

Once scrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG100 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, scrambling, and 1:20 block distribution mechanism must maintain the mapping of MLG lanes 0.0 to 9.1 to PCS lanes 0 to 19. For clarity, this process is illustrated in Figure 25.

7.4.1.11 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 2 for a MLG100 gearbox. If RM channel is present, the upper and lower half of AM 0.0 are reversed, as described in 7.1.1. 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 BIP_3 and BIP_7 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.

7.4.1.12 RS-FEC Sublayer (Transmit)

See Section 7.1.2 for description of the general mechanism. A MLG100 mux supporting RS-FEC encoded physical interfaces implements the transmit function within the RS-FEC Sublayer of IEEE Std 802.3bj-2014, except for the following:

- The AM mapping and insertion mechanism described in Clause 91.5.2.6 is replaced with the function described in Section 7.1.2.

7.4.1.13 100GBASE-R PMA (4:4)

The four RS-FEC lanes produced by the RS-FEC sublayer are connected to physical medium via a PMA (4:4) to produce a 4-lane by 25.78125 Gb/s interface. Note that an implementation may support generation of any of the optional test patterns specified for a 100GBASE-R PMA across the MLG100 link: see IEEE Std 802.3-2012 Clause 83.5.10. If supported, the corresponding MDIO control variables must also be supported.

7.4.2 MLG100 mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 31. These may be eight of the twenty MLG lanes of a MLG100.

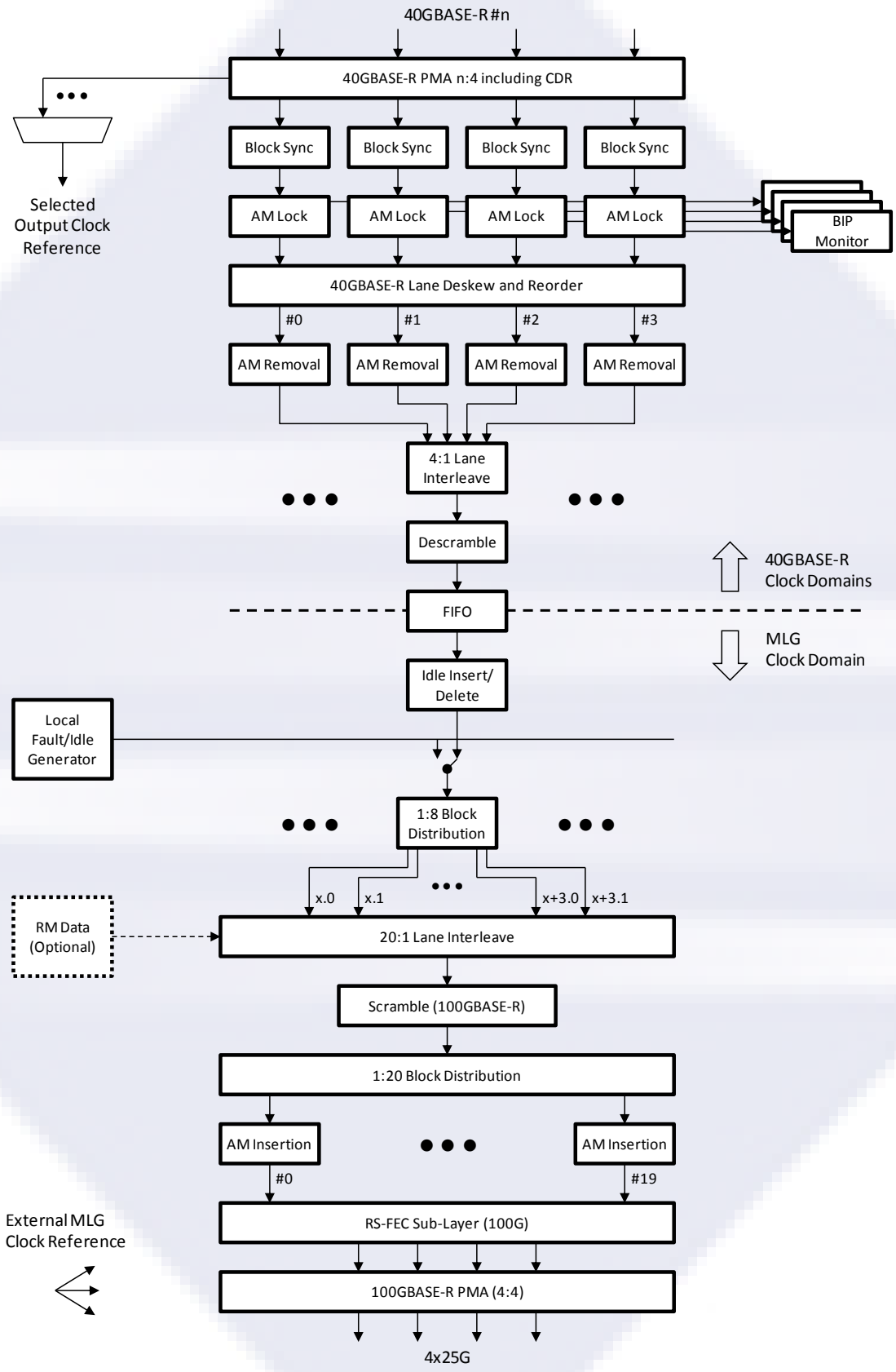


Figure 31: MLG100 mux 40GBASE-R mapping Block Diagram

7.4.2.1 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.

7.4.2.2 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.

7.4.2.3 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.

7.4.2.4 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.

7.4.2.5 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.

7.4.2.6 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.

7.4.2.7 Idle insert/delete

Idles are inserted or deleted as necessary in order that the bitstream, after PCS lane marker and RM channel data 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 (± 100ppm), 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.

7.4.2.8 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 7.2.2.1), 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 D_1, D_2, D_3, and zeros in Z_4, Z_5, Z_6, Z_7. 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.

7.4.2.9 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 7.2.1.7). 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.

7.4.2.10 8:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

7.4.2.11 RM data insertion and 20:1 Lane Interleave

After 10GBASE-R and 40GBASE-R block distribution to MLG lanes, if RM channel is enabled, the RM channel data blocks are inserted periodically into each MLG lane, as shown in Figure 25. The 66b blocks shall be inserted 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 RM blocks shall be inserted after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

The twenty unscrambled MLG100 lanes, including RM data, if enabled, are 66-bit block interleaved, from lane 0.0 up to lane 9.1, to create a 100G block stream, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

7.4.2.12 Scramble

The non-sync header bits of the MLG100 block stream are scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6.

7.4.2.13 1:20 Block Distribution

Once scrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG100 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, scrambling, and 1:20 block distribution mechanism must maintain the mapping of MLG lanes 0.0 to 9.1 to PCS lanes 0 to 19. For clarity, this process is illustrated in Figure 25.

7.4.2.14 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 2 for a MLG100 gearbox. If RM channel is present, the upper and lower half of AM 0.0 are reversed, as described in 7.1.1. 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 BIP_3 and BIP_7 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.

7.4.2.15 RS-FEC Sublayer

See Section 7.4.1.12. 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.

7.4.2.16 100GBASE-R PMA (4:4)

See Section 7.4.1.13. 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.

7.4.3 MLG100 demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 27.

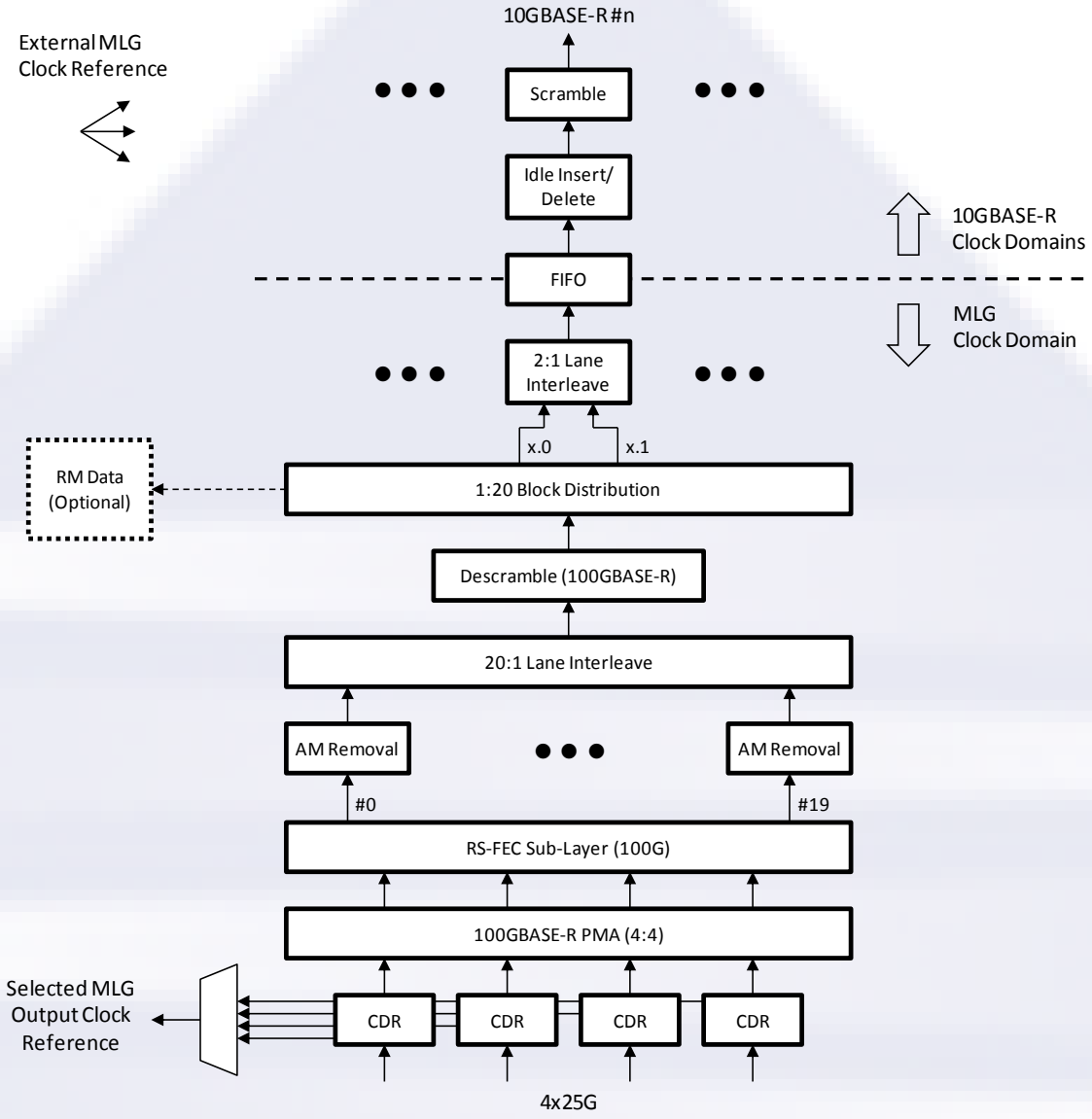


Figure 32: MLG100 demux 10GBASE-R demapping Block Diagram

7.4.3.1 100GBASE-R PMA (4:4)

The input to the MLG demux will come from a physical interface that is similar to that of 4x25G RS-FEC for the MLG100 application. 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.

7.4.3.2 RS-FEC Sublayer (Receive)

See Section 7.1.2 for description of the general mechanism. A MLG100 demux supporting RS-FEC encoded physical interfaces implements the receive function within the RS-FEC Sublayer of IEEE Std 802.3bj-2014, except for the following:

- Since the MLG100 demux payload may be from a 10GE client, the block type field restoration of step f2 of Clause 91.5.3.5 shall be cross-referenced using Figure 49-7 of IEEE Std 802.3-2012, which is a superset of Figure 82-5.
- The AM mapping and insertion mechanism described in Clause 91.5.3.7 is replaced with the mechanism described in Section 7.1.2.

7.4.3.3 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.

7.4.3.4 20:1 Lane Interleave

After AM removal, the twenty scrambled MLG100 lanes are 66-bit block interleaved, from lane 0 up to lane 19, to create a 100G block stream, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

7.4.3.5 Descramble

The non-sync header bits of the MLG100 block stream are 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.

7.4.3.6 1:20 Block Distribution RM data removal

Once descrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG100 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, descrambling, and 1:20 block distribution mechanism must maintain the mapping of PCS lanes 0 to 19 to MLG lanes 0.0 to 9.1. For clarity, this process is illustrated in Figure 28.

After block distribution to the twenty MLG lanes, if RM channel is enabled, the RM channel data blocks are removed periodically from each MLG lane, as shown in Figure 28. The 66b blocks shall be removed on each MLG lane in the same way (at the same time on all twenty MLG lanes) after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

7.4.3.7 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.

7.4.3.8 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.

7.4.3.9 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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).

7.4.3.10 Scramble

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

7.4.4 MLG100 demux for 40GBASE-R

The process for demapping a 40GBASE-R from eight MLG lanes is shown in Figure 33.

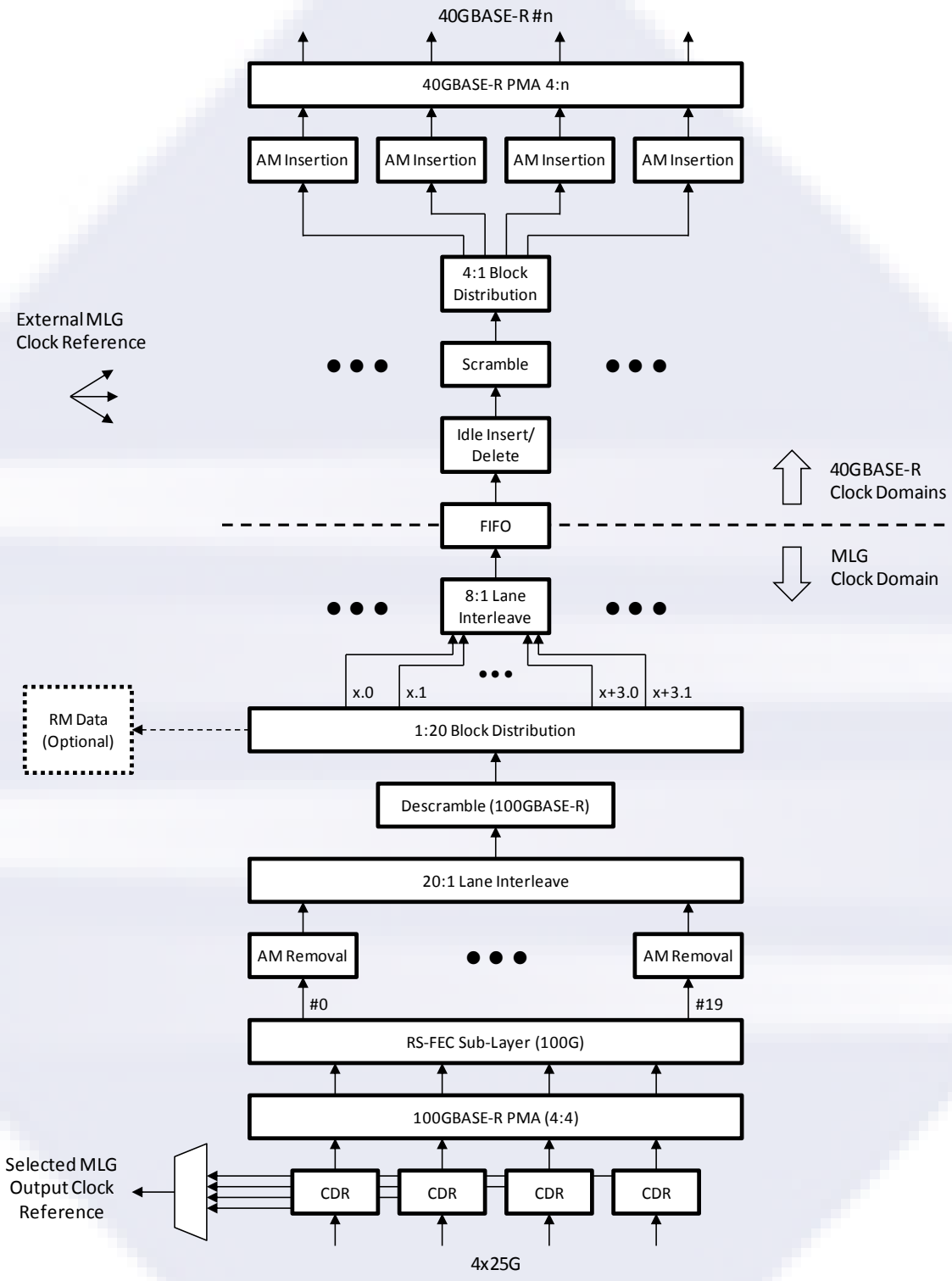


Figure 33: MLG100 demux 40GBASE-R demapping Block Diagram

7.4.4.1 100GBASE-R PMA (4:4)

See Section 7.4.3.1. 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.

7.4.4.2 RS-FEC Sublayer

See Section 7.4.3.2. 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.

7.4.4.3 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.

7.4.4.4 20:1 Lane Interleave

After AM removal, the twenty scrambled MLG100 lanes are 66-bit block interleaved, from lane 0 up to lane 19, to create a 100G block stream, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

7.4.4.5 Descramble

The non-sync header bits of the MLG100 block stream are 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.

7.4.4.6 1:20 Block Distribution RM data removal

Once descrambled, the 66-bit blocks are distributed round-robin to the appropriate MLG100 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, descrambling, and 1:20 block distribution mechanism must maintain the mapping of PCS lanes 0 to 19 to MLG lanes 0.0 to 9.1. For clarity, this process is illustrated in Figure 28.

After block distribution to the twenty MLG lanes, if RM channel is enabled, the RM channel data blocks are removed periodically from each MLG lane, as shown in Figure 28. The 66b blocks shall be removed on each MLG lane in the same way (at the same time on all twenty MLG lanes) after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

7.4.4.7 MLG Lane Interleave

The eight 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.

7.4.4.8 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 may 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 (±100ppm), 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).

7.4.4.9 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.

7.4.4.10 4:1 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.

7.4.4.11 Alignment Insertion

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

7.4.4.12 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.

8 MLG200

8.1 MLG200 General Mechanism

MLG200 employs the same approach as MLG100 (see Section 7.1) for the transport of 10GBASE-R and 40GBASE-R signals, but with double the number of MLG lanes (40 instead of 20). Each 10GBASE-R signal is 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 19 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 8.1.1.

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 rescrambling the non-lane marker blocks.

8.1.1 MLG200 Lane Markers

The MLG200 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 3 provides the MLG alignment markers used for the MLG200 application. 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 carry either 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.

If Remote Management (RM) is enabled, the AM values for M0, M1, M2, BIP3 and M4, M5, M6, BIP7 are swapped on MLG lane 0.0 (0.0-10 or 0.0-40). This is a similar format as used for Rapid Alignment Markers, as shown in Figure 82-9b of IEEE 802.3bj-2014.

Table 3: MLG200 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[b]	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[b]	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		

[a] Each octet is transmitted LSB to MSB.

[b] The upper and lower half of the AM Encoding is reversed on MLG lane 0.0 if RM channel is in data stream.

8.1.2 MLG200 RS-FEC

If RS-FEC is required by the physical link, the 40 MLG lanes of an MLG200 application are mapped to two instances of the 100G RS-FEC sublayer of IEEE Std 802.3bj-2014, Clause 91. Each uses a similar approach as defined in Section 7.1.2, but with one carrying MLG lanes 0.0 through 9.1 and the other MLG lanes 10.0 through 19.1. Separate 100G scramblers are used to scramble the MLG data for each instance to support the 256b/257b inverse transcoding within the RS-FEC sublayer.

The MLG AM mapping to RS-FEC lanes is shown in Figure 34. As done in Clause 91, the Clause 82 AM0 and AM16 are replicated across the four lanes of each RS-FEC instance and AM4 through AM7 are used to denote RS-FEC lanes 0 through 3 within each. The MLG AMs are then mapped into the remaining sections, and can be used to distinguish between MLG lanes 0-3 and 4-7, and the specific 10GE/40GE client and Remote Management configurations.

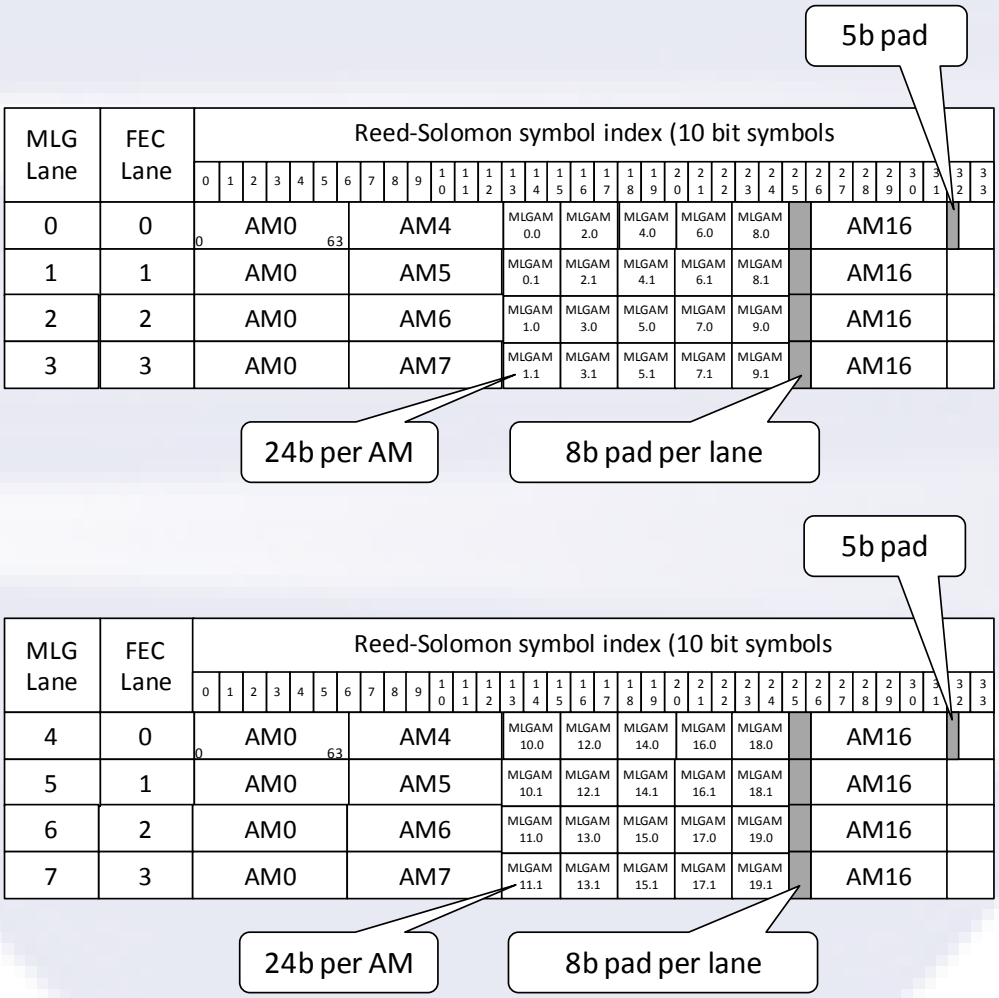


Figure 34: MLG200 RS-FEC AM mapping

Each 5b pad of Figure 34 is added at the end of the alignment marker mapping as described in Section 91.5.2.6 of IEEE Std 802.3bj-2014. The 8b pad is added per lane in the same manner, except with the alternating 8b pattern “01100101” and “10011010”.

Each MLG AM includes a BIP field. The mapping of these fields into RS-FEC lanes is shown in Figure 35.

0	AM0	BIP 0.0		BIP 2.0	AM4	BIP 4.0		BIP 6.0		AM16	BIP 8.0		!BIP 8.0
1	AM0	BIP 0.1		BIP 2.1	AM5	BIP 4.1		BIP 6.1		AM16	BIP 8.1		!BIP 8.1
2	AM0	BIP 1.0		BIP 3.0	AM6	BIP 5.0		BIP 7.0		AM16	BIP 9.0		!BIP 9.0
3	AM0	BIP 1.1		BIP 3.1	AM7	BIP 5.1		BIP 7.1		AM16	BIP 9.1		!BIP 9.1

4	AM0	BIP 10.0		BIP 12.0	AM4	BIP 14.0		BIP 16.0		AM16	BIP 18.0		!BIP 18.0
5	AM0	BIP 10.1		BIP 12.1	AM5	BIP 14.1		BIP 16.1		AM16	BIP 18.1		!BIP 18.1
6	AM0	BIP 11.0		BIP 13.0	AM6	BIP 15.0		BIP 17.0		AM16	BIP 19.0		!BIP 19.0
7	AM0	BIP 11.1		BIP 13.1	AM7	BIP 15.1		BIP 17.1		AM16	BIP 19.1		!BIP 19.1

Figure 35: MLG200 RS-FEC BIP mapping

Alignment lock, deskew, and lane reordering are done within each 4-lane sub-group (0-3 and 4-7) as defined in Sections 91.5.3.1 and 91.5.3.2 of IEEE Std 802.3bj-2014. To maintain link coherency of a 40GE client mapped to MLG lanes 8.0 through 11.1, or, if enabled, the RM channels, the output of the two RS-FEC receive sublayers must be deskewed and reordered between the upper and lower links.

8.1.3 MLG200 Remote Management Channels

The optional Remote Management (RM) channels are designed to carry maintenance information, which may include status, control, software images, or other information deemed appropriate by the two end points. In MLG200 applications, there are two independent RM data channels (RM 0 and RM 1), each formatted as described in Section 7.1.3. If RM is enabled, both RM channels are inserted into the data stream as described below. Separate or aggregated usage of the channels is outside the scope of this IA, and the two endpoints must agree to the data format for successful communications.

Data of each RM channel is 64b/66b encoded as data blocks, per IEEE Std 802.3-2012 Clause 82, Figure 82-5. The unscrambled 66b data blocks of the two RM channels are inserted into the unscrambled MLG data stream in a similar manner as AM blocks, except they are scrambled with MLG data. They are simultaneously striped across all 40 MLG lanes, as shown in Figure 36. The RM blocks are spaced at a distance of 16384 from each other, and at a distance of 8192 from MLG AM blocks. The resulting raw bandwidth of each RM channel is ~6.1 Mb/s.

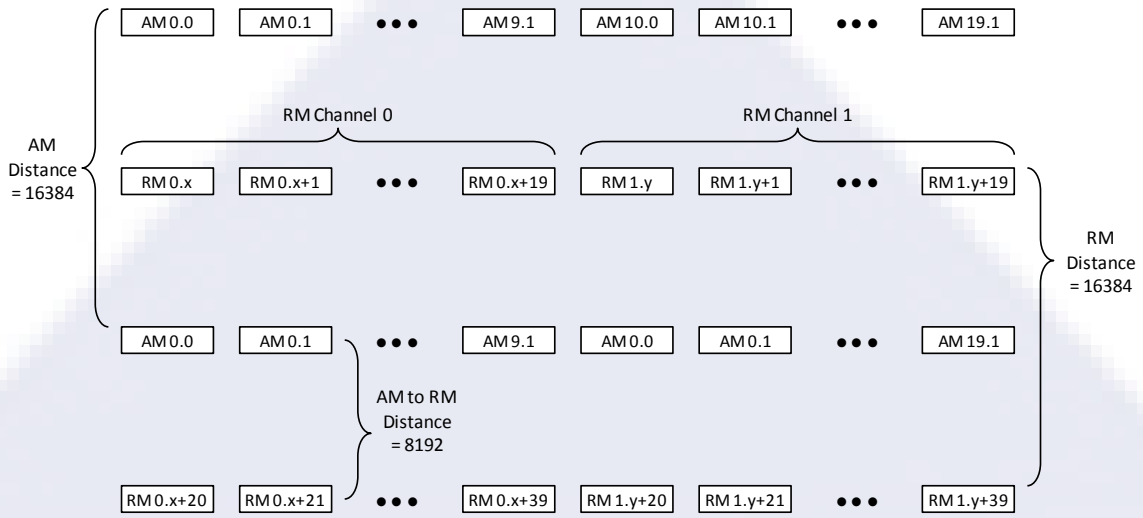


Figure 36: MLG200 RM Data Channel

8.2 MLG200 Detailed Block Diagrams (without RS-FEC and RM)

The MLG200 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.

The MLG200 Demux receives forty MLG lanes. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

8.2.1 MLG200 mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 37.

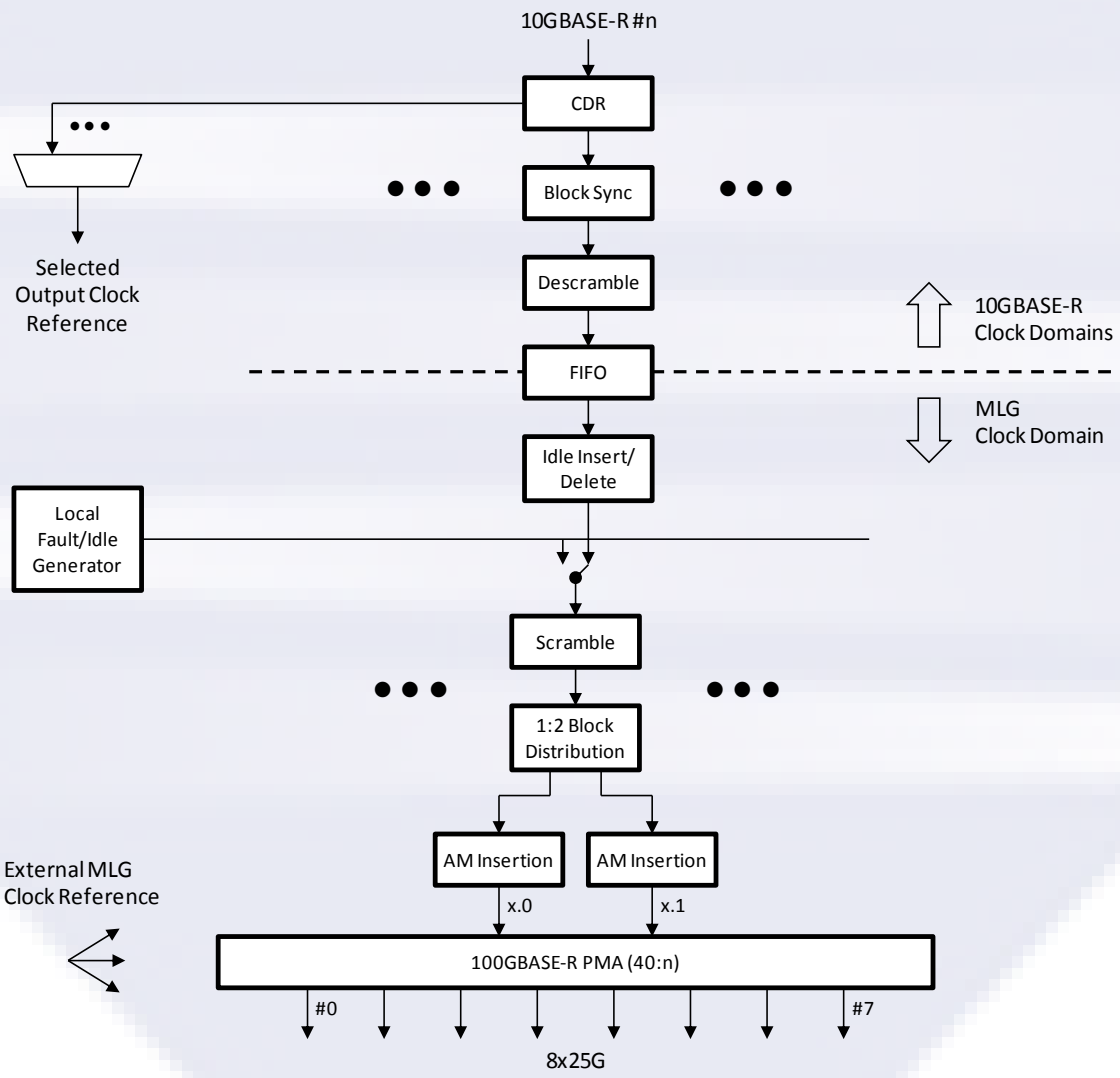


Figure 37: MLG200 mux 10GBASE-R mapping Block Diagram

8.2.1.1 Clock and Data Recovery (CDR)

Each of the 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.

8.2.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

8.2.1.3 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.

8.2.1.4 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 49.2.4.7.

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 (± 100ppm), 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.

8.2.1.5 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 7.2.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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.

8.2.1.6 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 8.2.1.7). 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.

8.2.1.7 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.

8.2.1.8 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 twenty PCS lanes.

8.2.1.9 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 forty 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 3 for an MLG200 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 BIP_3 and BIP_7 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.

8.2.1.10 100GBASE-R PMA (40:n)

The processes described in sections 8.2.1.1 through 8.2.1.9 will produce forty MLG lanes, each with a bit rate of 5.15625 Gb/s ±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 a double width 100GBASE-R PMA (40:n). The most typical value of “n” is expected to be 8 for the MLG200 application. A PMA (40:8) 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.

8.2.2 MLG200 mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 38. These may be eight of the forty MLG lanes of a MLG200.

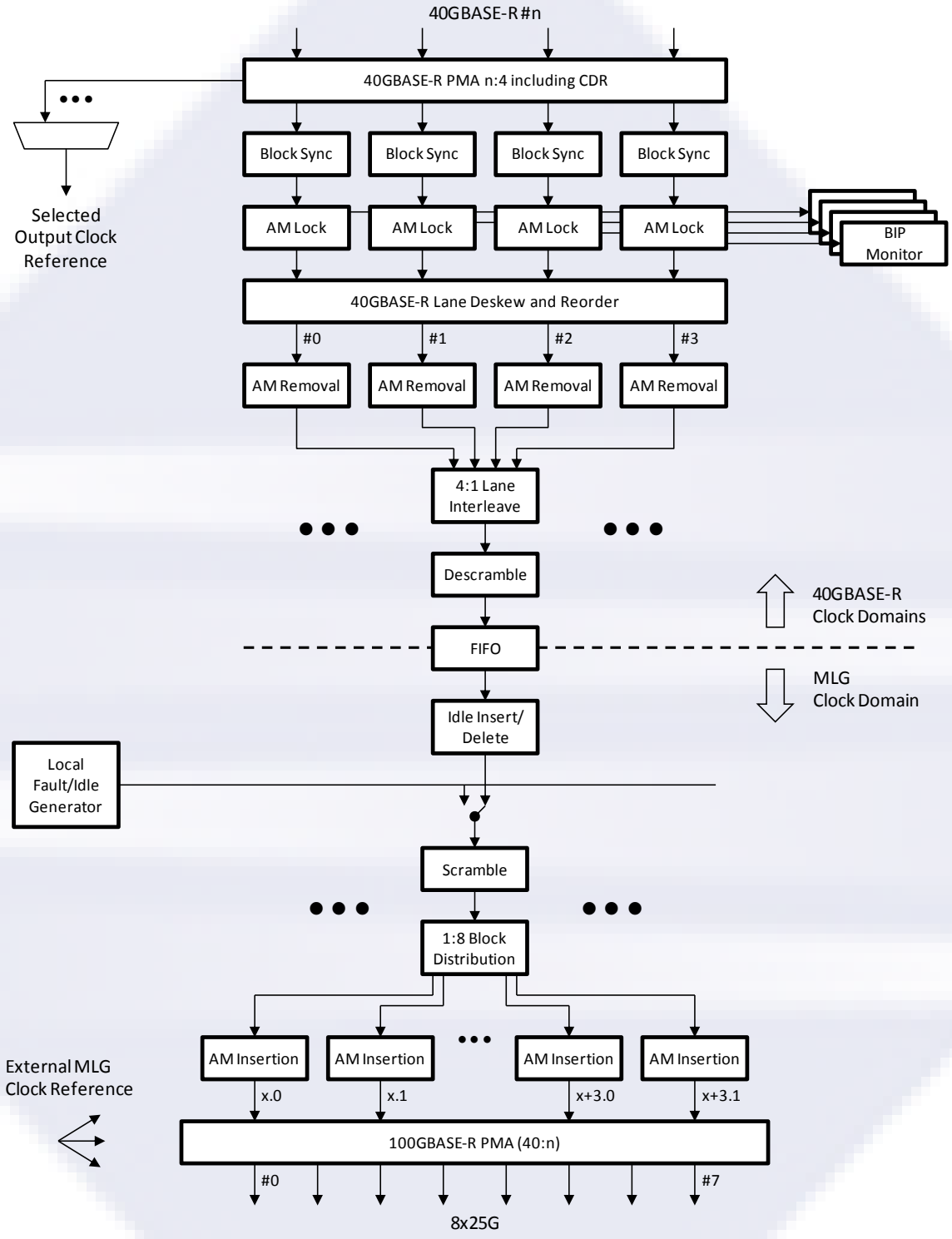


Figure 38: MLG200 mux 40GBASE-R mapping Block Diagram

8.2.2.1 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.

8.2.2.2 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.

8.2.2.3 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.

8.2.2.4 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.

8.2.2.5 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.

8.2.2.6 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.

8.2.2.7 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 (± 100ppm), 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.

8.2.2.8 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 8.2.2.1), 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 D_1, D_2, D_3, and zeros in Z_4, Z_5, Z_6, Z_7. 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.

8.2.2.9 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. A 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 8.2.2.10). 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.

8.2.2.10 Scramble

After idle insertion/deletion and Local Fault insertion, the resulting stream is re-scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6.

8.2.2.11 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 twenty PCS lanes.

8.2.2.12 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 forty 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 3 for a MLG200 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 BIP_3 and BIP_7 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.

8.2.2.13 100GBASE-R PMA (40:n)

See Section 8.2.1.10. 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.

8.2.3 MLG200 demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 39.

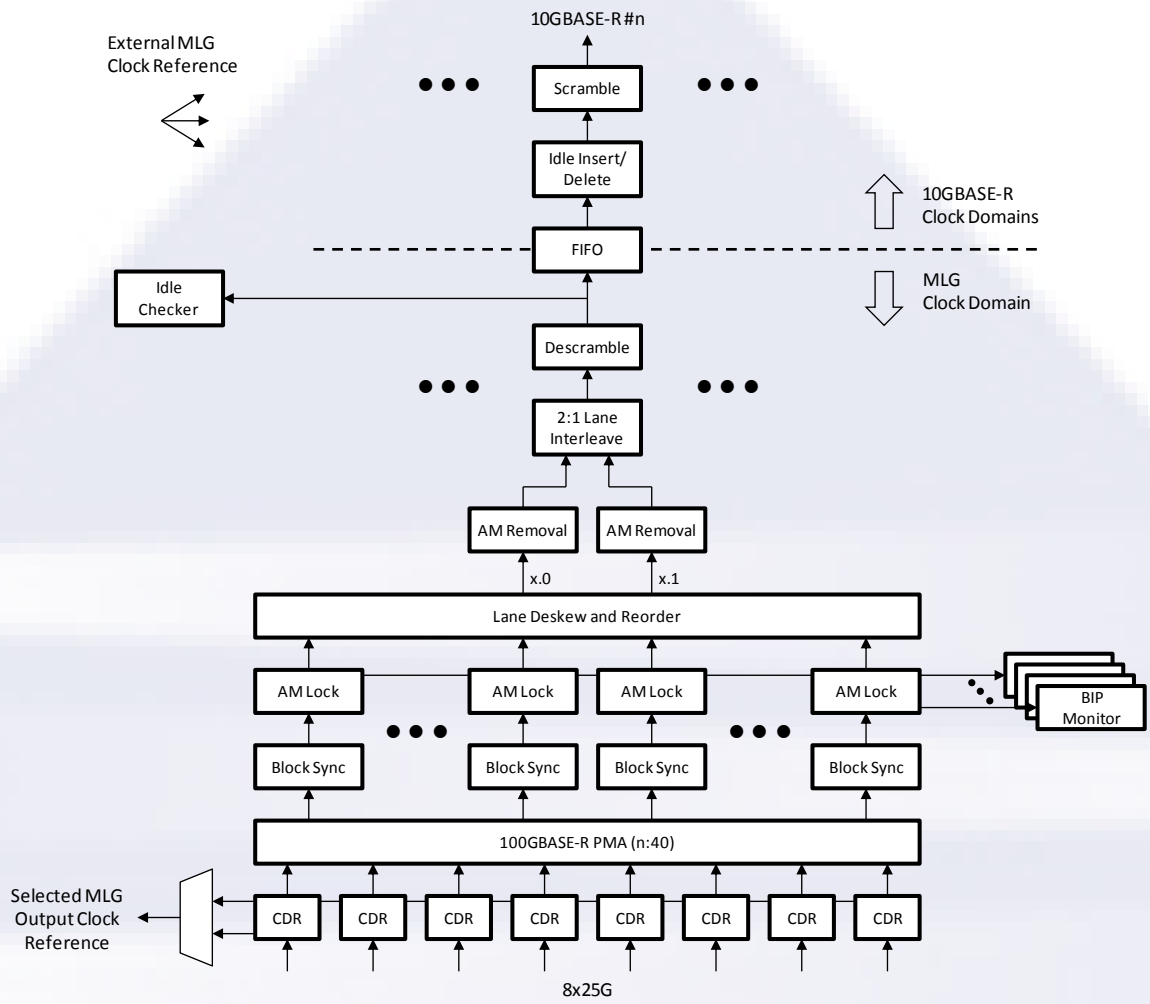


Figure 39: MLG200 demux 10GBASE-R demapping Block Diagram

8.2.3.1 100GBASE-R PMA (n:40)

The input to the MLG demux will come from a physical interface that is similar to that of a double-width 100GBASE-R for the MLG200 application. For the most common MLG configuration, this is expected to be an 8-lane by 25.78125 Gb/s interface, where each of these physical lanes carries five of the forty MLG lanes, bit multiplexed, but other physical lane counts which are divisors of 40 are possible. The PMA (n:40) will demultiplex the MLG lanes as if they were PCS lanes intoforty individual bit streams of 5.15625 Gb/s ±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.

8.2.3.2 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.

8.2.3.3 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 2 for the MLG200 application.

8.2.3.4 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.

8.2.3.5 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.

8.2.3.6 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.

8.2.3.7 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.

8.2.3.8 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.

8.2.3.9 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.

8.2.3.10 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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).

8.2.3.11 Scramble

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

8.2.4 MLG200 demux for 40GBASE-R

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

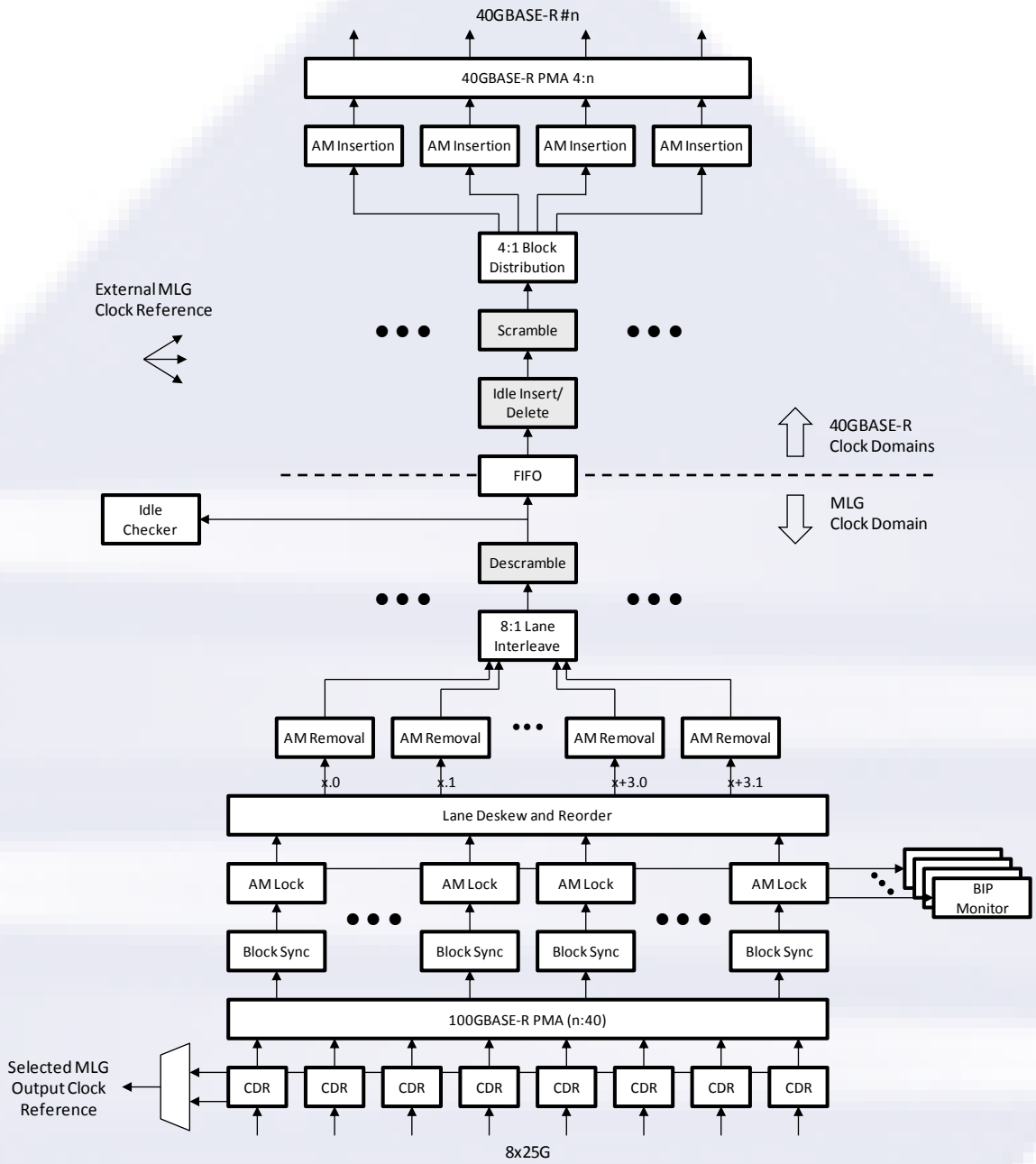


Figure 40: MLG200 demux 40GBASE-R demapping Block Diagram

8.2.4.1 100GBASE-R PMA (n:40)

See Section 8.2.3.1. 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.

8.2.4.2 MLG Lane block sync

See Section 8.2.3.2. 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.

8.2.4.3 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 3 for the MLG200 application.

8.2.4.4 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.

8.2.4.5 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.

8.2.4.6 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.

8.2.4.7 MLG Lane Interleave

The eight 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.

8.2.4.8 Descramble

Note that there are three shaded boxes in Figure 23: 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 8.2.4.11 if a different clock domain is not needed for the 40GBASE-R.

When required, the non-sync header bits of the 40GBASE-R signal are 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 of IEEE Std 802.3-2012.

8.2.4.9 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 may 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 (±100ppm), 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).

8.2.4.10 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.

8.2.4.11 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.

8.2.4.12 Alignment Insertion

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

8.2.4.13 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.

8.3 MLG200 Detailed Block Diagrams (without RS-FEC and with RM)

The MLG200 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.

The MLG200 Demux receives forty MLG lanes. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

8.3.1 MLG200 mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 41.

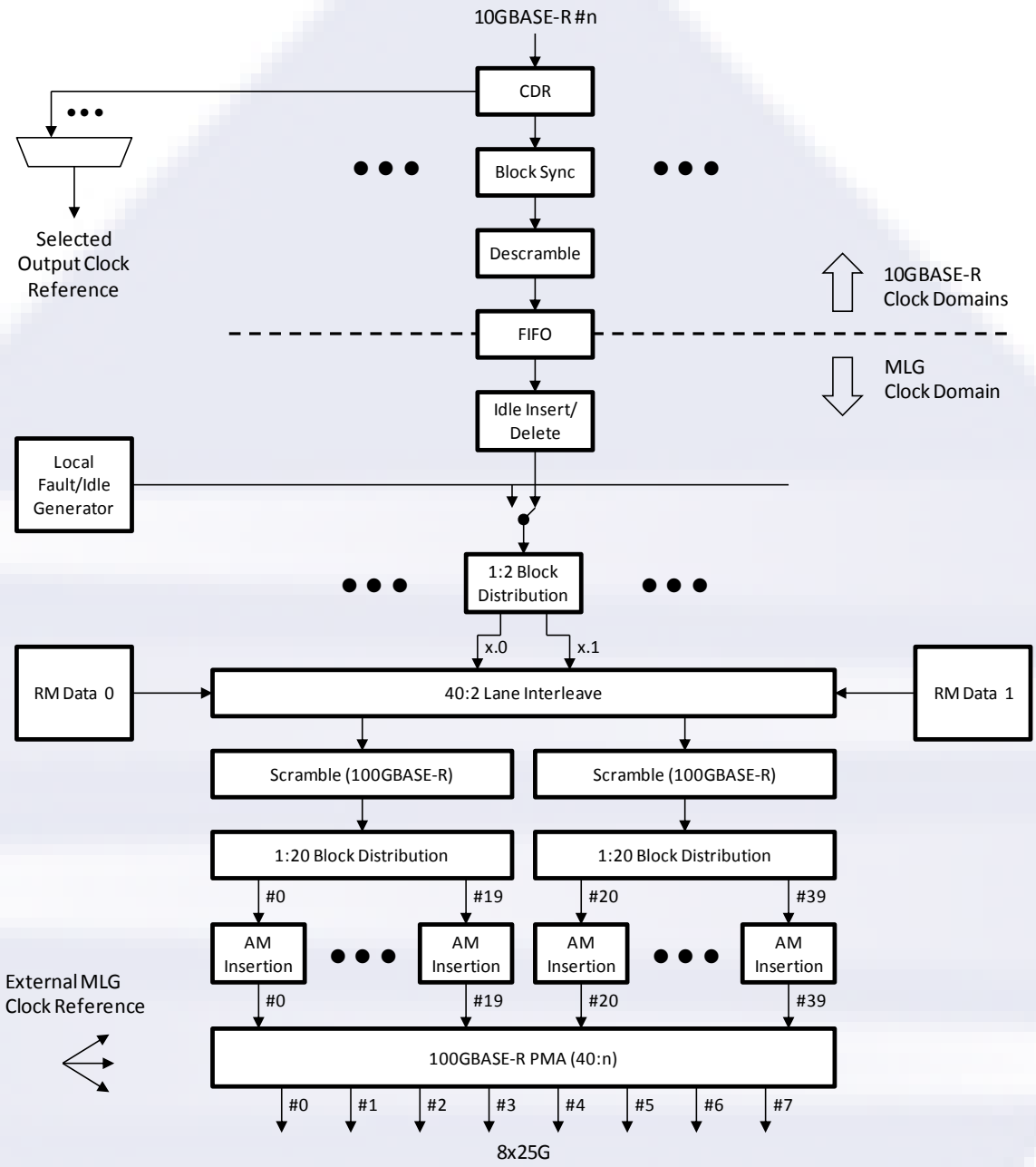


Figure 41: MLG200 mux 10GBASE-R mapping Block Diagram

8.3.1.1 Clock and Data Recovery (CDR)

Each of the 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.

8.3.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

8.3.1.3 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.

8.3.1.4 Idle Insert/Delete

Idles are inserted or deleted as necessary in order that the bitstream, after MLG lane marker and RM channel data 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 49.2.4.7.

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 (± 100ppm), 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.

8.3.1.5 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 8.3.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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.

8.3.1.6 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 8.3.1.9). 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.

8.3.1.7 2:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

8.3.1.8 RM data insertion and 20:1 Lane Interleave

After 10GBASE-R and 40GBASE-R block distribution to MLG lanes, the RM channel data blocks are inserted periodically into each MLG lane, as shown in Figure 42. The 66b blocks shall be inserted on each MLG lane in the same way (at the same time on all forty MLG lanes) as the insertion of PCS lane alignment markers described in Clause 82.2.7 of IEEE Std 802.3-2012. The RM blocks shall be inserted after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

The upper and lower set of twenty unscrambled MLG200 lanes, including RM data, are 66-bit block interleaved, from lane 0.0 up to lane 9.1 and 10.0 up to lane 19.1, to create two 100G block streams, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

8.3.1.9 Scramble

The non-sync header bits of each 100G block stream are scrambled according to IEEE Std 802.3-2012 Clause 82.2.5, which is identical to the 10GBASE-R scrambler described in Clause 49.2.6.

8.3.1.10 1:20 Block Distribution

Once scrambled, the 66-bit blocks from each 100G block stream are distributed round-robin to the appropriate MLG200 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 40:2 lane interleave, scrambling, and 2:40 block distribution mechanism must maintain the mapping of MLG lanes 0.0 to 19.1 to PCS lanes 0 to 39. For clarity, this process is illustrated in Figure 42.

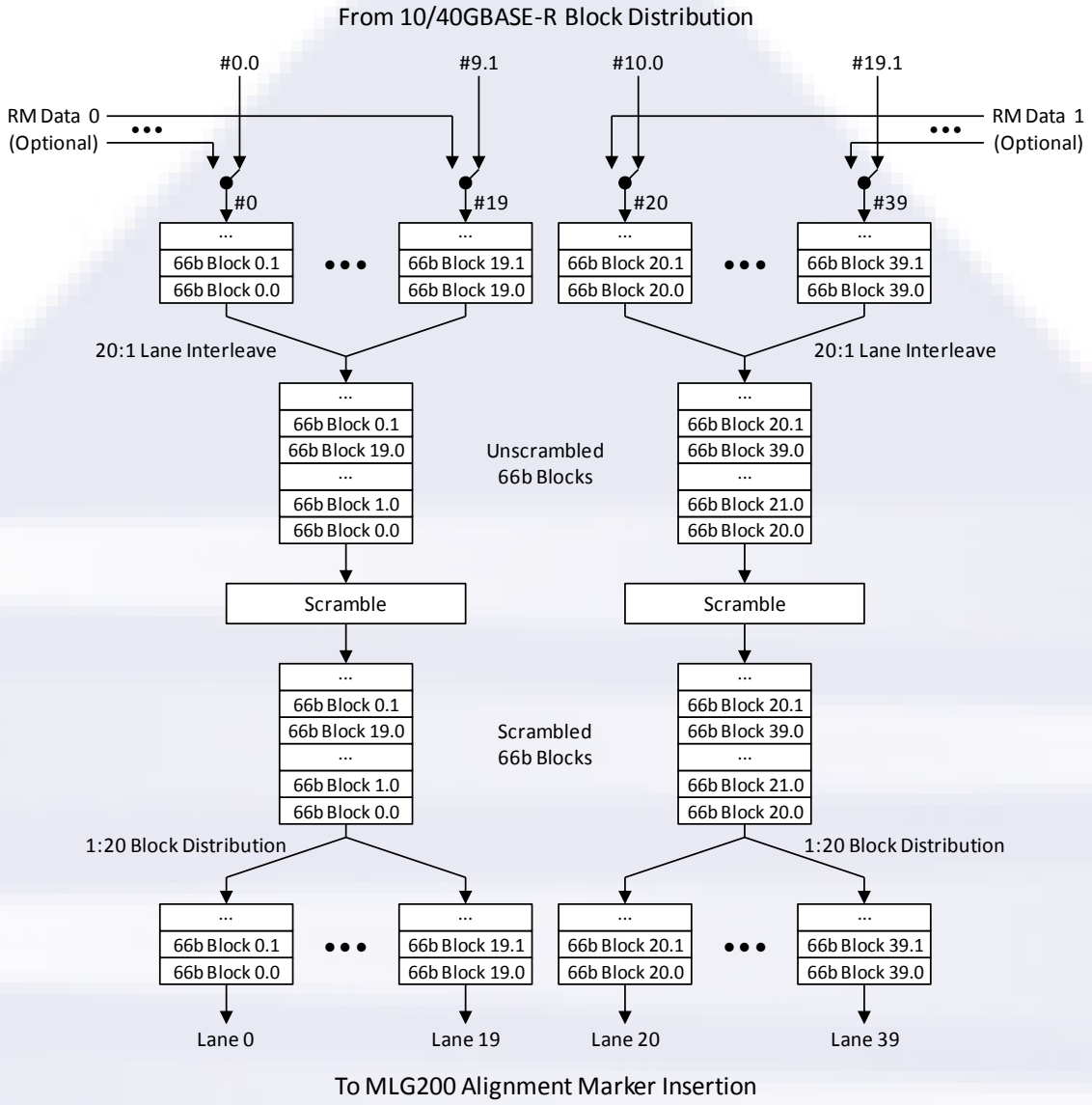


Figure 42: MLG200 Scrambler Diagram and 66b Block Flow

8.3.1.11 Alignment Marker Insertion

In order to support deskew and reordering of the MLG lanes into the constituent 10GBASE-R or 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 3 for a MLG200 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 BIP_3 and BIP_7 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.

8.3.1.12 100GBASE-R PMA (40:n)

The processes described in sections 8.3.1.1 through 8.3.1.11 will produce forty MLG lanes, each with a bit rate of 5.15625 Gb/s ±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 a double width 100GBASE-R PMA (40:n). The most typical value of “n” is expected to be 8 for the MLG200 application. A PMA (40:8) 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.

8.3.2 MLG200 mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 43. These may be eight of the twenty MLG lanes of a MLG100.

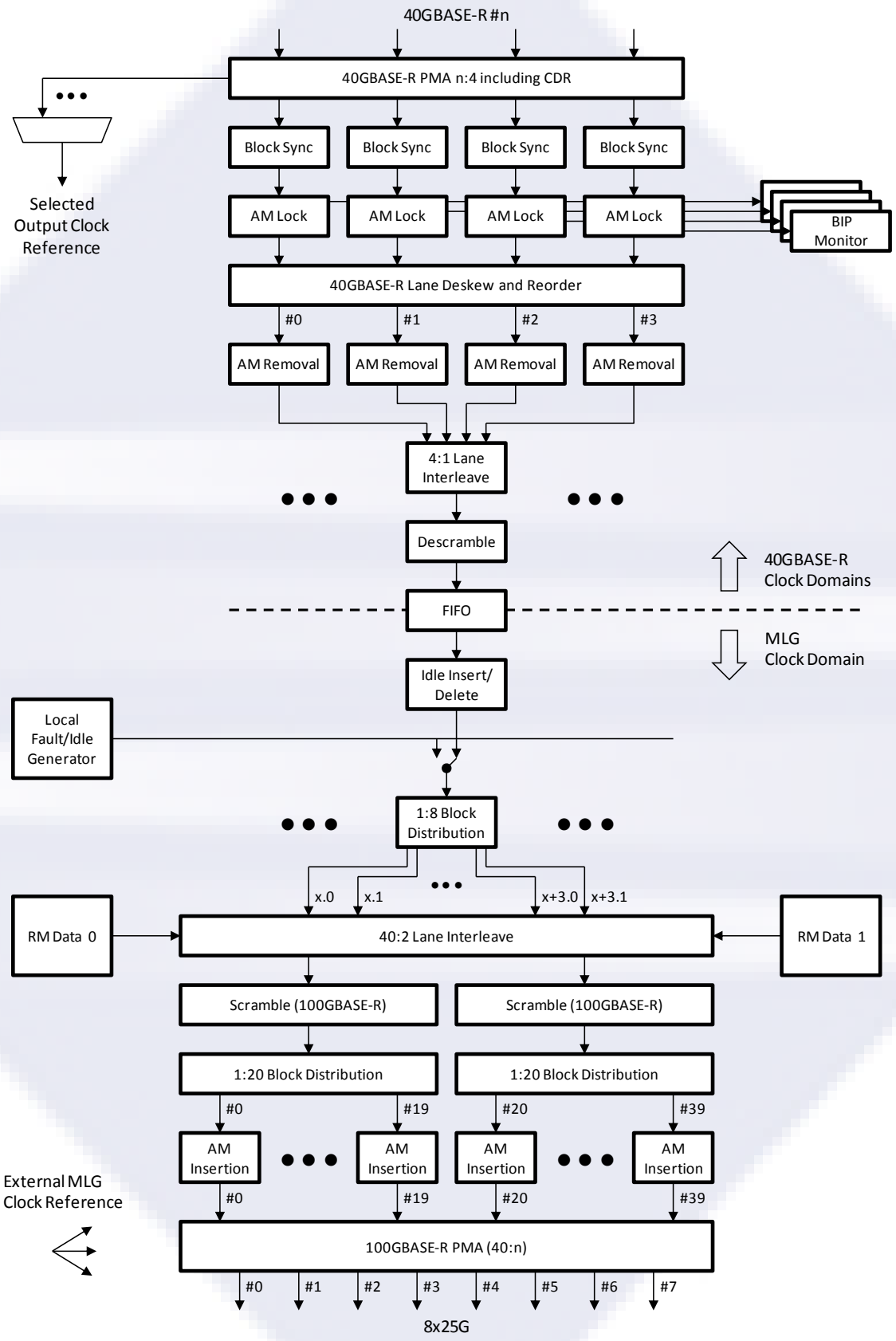


Figure 43: MLG200 mux 40GBASE-R mapping Block Diagram

8.3.2.1 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.

8.3.2.2 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.

8.3.2.3 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.

8.3.2.4 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.

8.3.2.5 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.

8.3.2.6 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.

8.3.2.7 Idle insert/delete

Idles are inserted or deleted as necessary in order that the bitstream, after PCS lane marker and RM channel data 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 (± 100ppm), 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.

8.3.2.8 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 8.3.2.1), 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 D_1, D_2, D_3, and zeros in Z_4, Z_5, Z_6, Z_7. 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.

8.3.2.9 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 8.3.2.12). 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.

8.3.2.10 8:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

8.3.2.11 RM data insertion and 20:1 Lane Interleave

See Section 8.3.1.8. 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.

8.3.2.12 Scramble

See Section 8.3.1.9. 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.

8.3.2.13 1:20 Block Distribution

See Section 8.3.1.10. 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.

8.3.2.14 Alignment Marker Insertion

See Section 8.3.1.11. 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.

8.3.2.15 100GBASE-R PMA (40:n)

See Section 8.3.1.12. 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.

8.3.3 MLG200 demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 44.

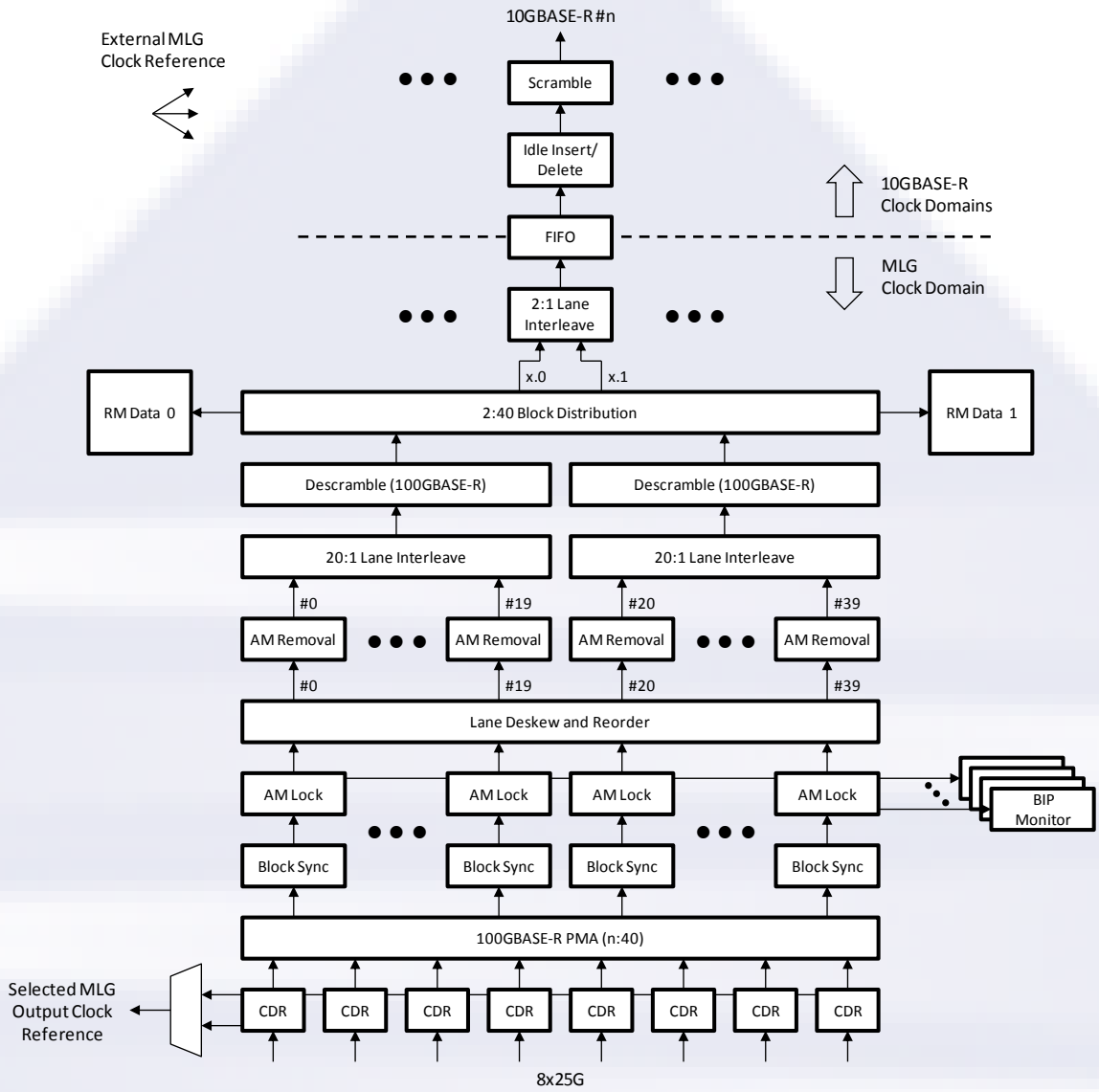


Figure 44: MLG200 demux 10GBASE-R demapping Block Diagram

8.3.3.1 100GBASE-R PMA(n:40)

The input to the MLG demux will come from a physical interface that is similar to that of a double-width 100GBASE-R for the MLG200 application. For the most common MLG configuration, this is expected to be an 8-lane by 25.78125 Gb/s interface, where each of these physical lanes carries five of the forty MLG lanes, bit multiplexed, but other physical lane counts which are divisors of 40 are possible. The PMA (n:40) will demultiplex the MLG lanes as if they were PCS lanes into forty individual bit streams of 5.15625 Gb/s ±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.

8.3.3.2 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.

8.3.3.3 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 2 for the 4x25G MLG100 application.

8.3.3.4 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.

8.3.3.5 MLG Lane reorder

The forty MLG lanes are deskewed and reordered according to the procedures described in IEEE Std 802.3-2012 Clauses 82.2.12 and 82.2.13.

8.3.3.6 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.

8.3.3.7 20:1 Lane Interleave

After AM removal, the upper and lower set of twenty scrambled MLG200 lanes, including RM data, are 66-bit block interleaved, from lane 0 up to lane 19 and lane 20 up to lane 39, to create two 100G block streams, similar to 100GBASE-R as described in IEEE Std 802.3-2012 Clause 82.2.14.

8.3.3.8 Descramble

The non-sync header bits of each 100G block stream are 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.

8.3.3.9 1:20 Block Distribution RM data removal

Once descrambled, the 66-bit blocks from each 100G block stream are distributed round-robin to the appropriate MLG200 lanes, as described in Clause 82.2.6 of IEEE Std 802.3-2012. The above 20:1 lane interleave, descrambling, and 1:20 block distribution mechanism must maintain the mapping of PCS lanes 0 to 19 and 20 to 39 to MLG lanes 0.0 to 9.1 and 10.1 to 19.1, respectively. For clarity, this process is illustrated in Figure 45.

After block distribution to the forty MLG lanes, the RM channel data blocks are removed periodically from each MLG lane, as shown in Figure 45. The 66b blocks shall be removed on each MLG lane in the same way (at the same time on all forty MLG lanes) after every 16383 data blocks and at an offset of 8192 from MLG alignment markers.

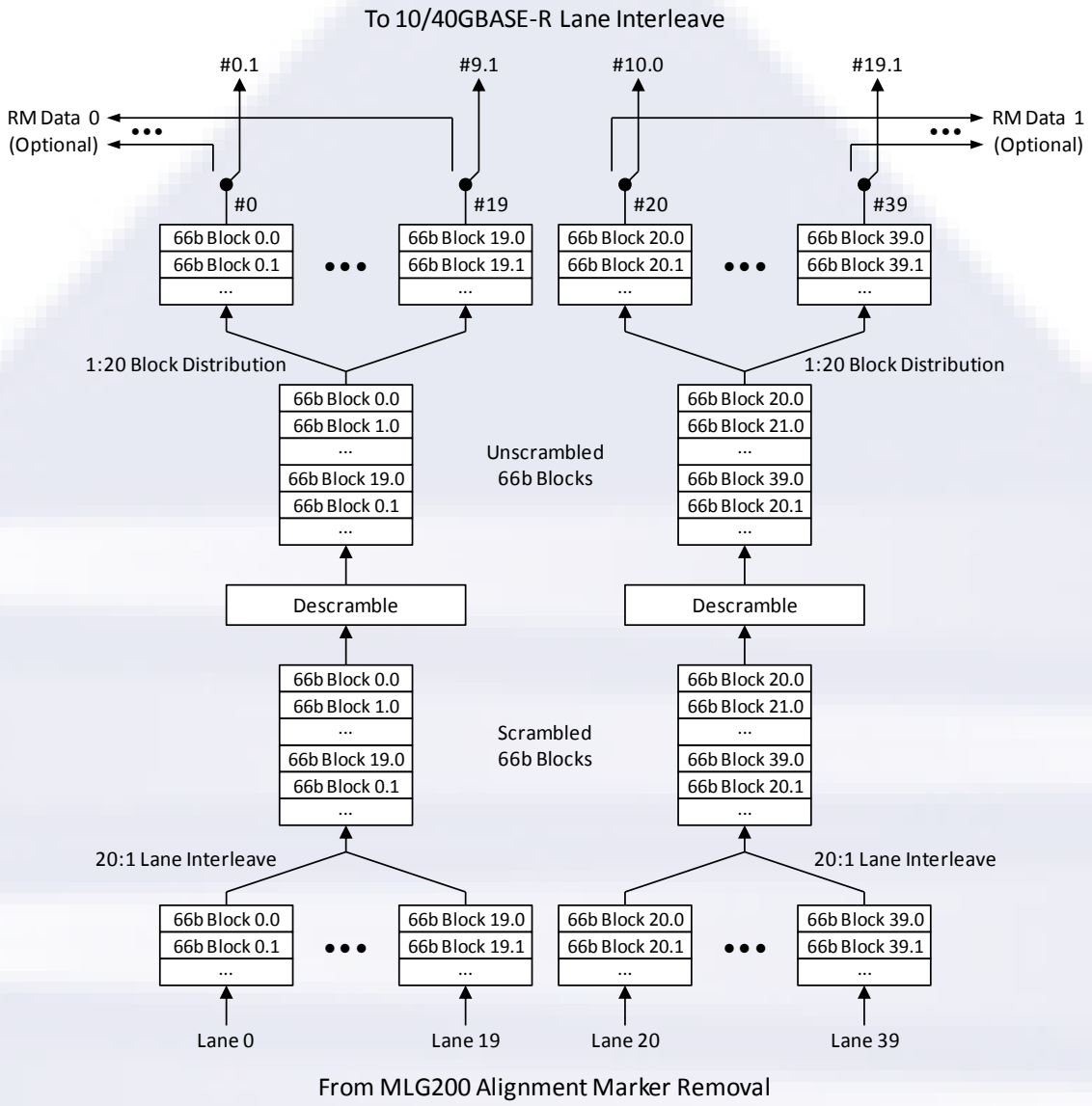


Figure 45: MLG200 Descrambler Diagram and 66b Block Flow

8.3.3.10 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.

8.3.3.11 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.

8.3.3.12 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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).

8.3.3.13 Scramble

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

8.3.4 MLG200 demux for 40GBASE-R

The process for demapping a 40GBASE-R from eight MLG lanes is shown in Figure 46.

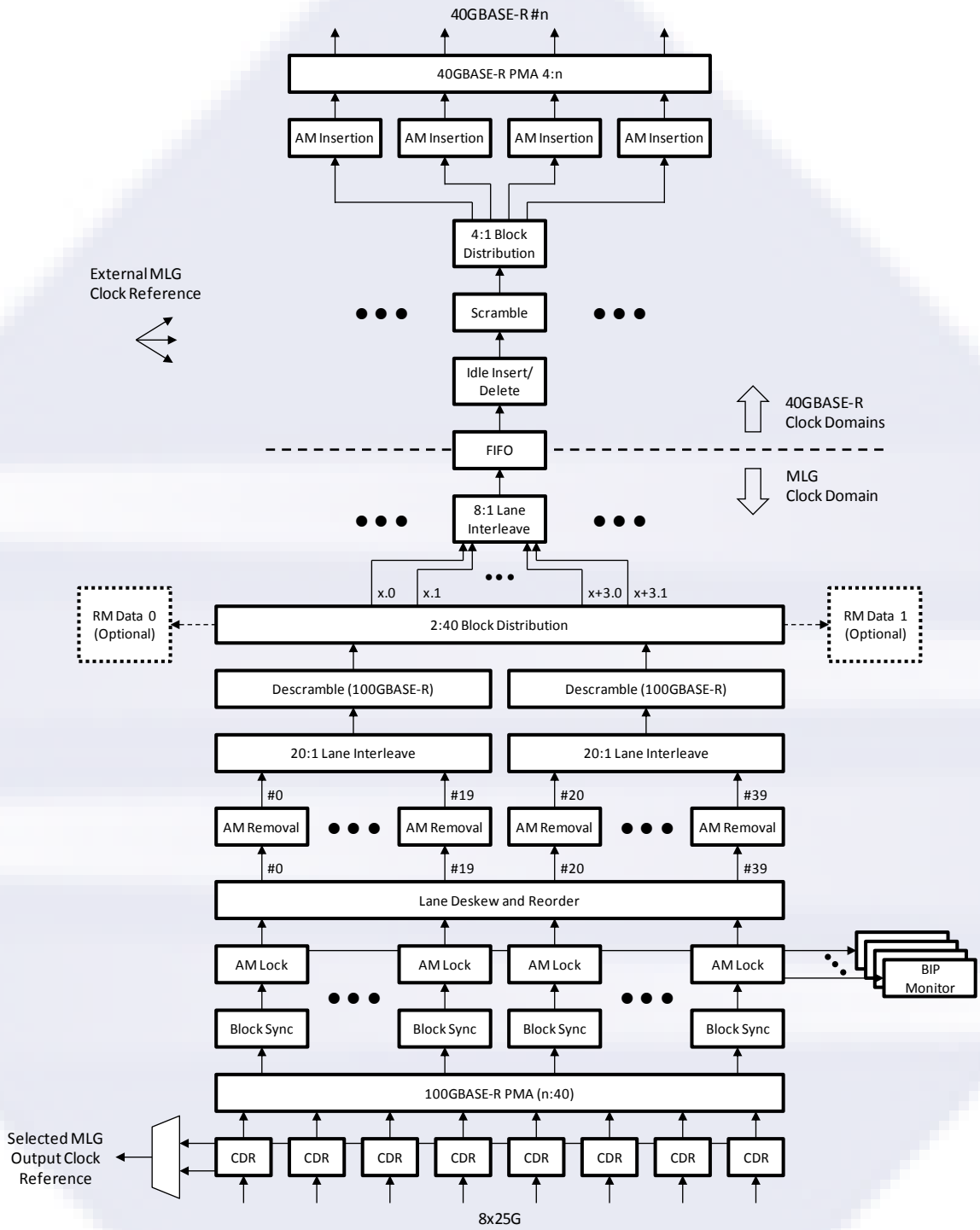


Figure 46: MLG200 demux 40GBASE-R demapping Block Diagram

8.3.4.1 100GBASE-R PMA(n:40)

See Section 8.3.3.1. 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.

8.3.4.2 MLG Lane block sync

See Section 8.3.3.2. 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.

8.3.4.3 MLG Lane Alignment lock

See Section 8.3.3.3. 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.

8.3.4.4 BIP monitor

See Section 8.3.3.4. 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.

8.3.4.5 MLG Lane reorder

See Section 8.3.3.5. 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.

8.3.4.6 MLG Lane Alignment marker removal

See Section 8.3.3.6. 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.

8.3.4.7 20:1 Lane Interleave

See Section 8.3.3.7. 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.

8.3.4.8 Descramble

See Section 8.3.3.8. 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.

8.3.4.9 1:20 Block Distribution RM data removal

See Section 8.3.3.9. 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.

8.3.4.10 MLG Lane Interleave

The eight 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.

8.3.4.11 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 may 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 (±100ppm), 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).

8.3.4.12 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.

8.3.4.13 4:1 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.

8.3.4.14 Alignment Insertion

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

8.3.4.15 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.

8.4 MLG200 Detailed Block Diagrams (with RS-FEC and optional RM)

The MLG200 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.

The MLG200 Demux receives twenty forty MLG lanes. Two MLG lanes may carry a 10GBASE-R, and eight MLG lanes may carry a 40GBASE-R.

8.4.1 MLG200 mux for 10GBASE-R

The process for mapping a 10GBASE-R into two MLG lanes is shown in Figure 47.

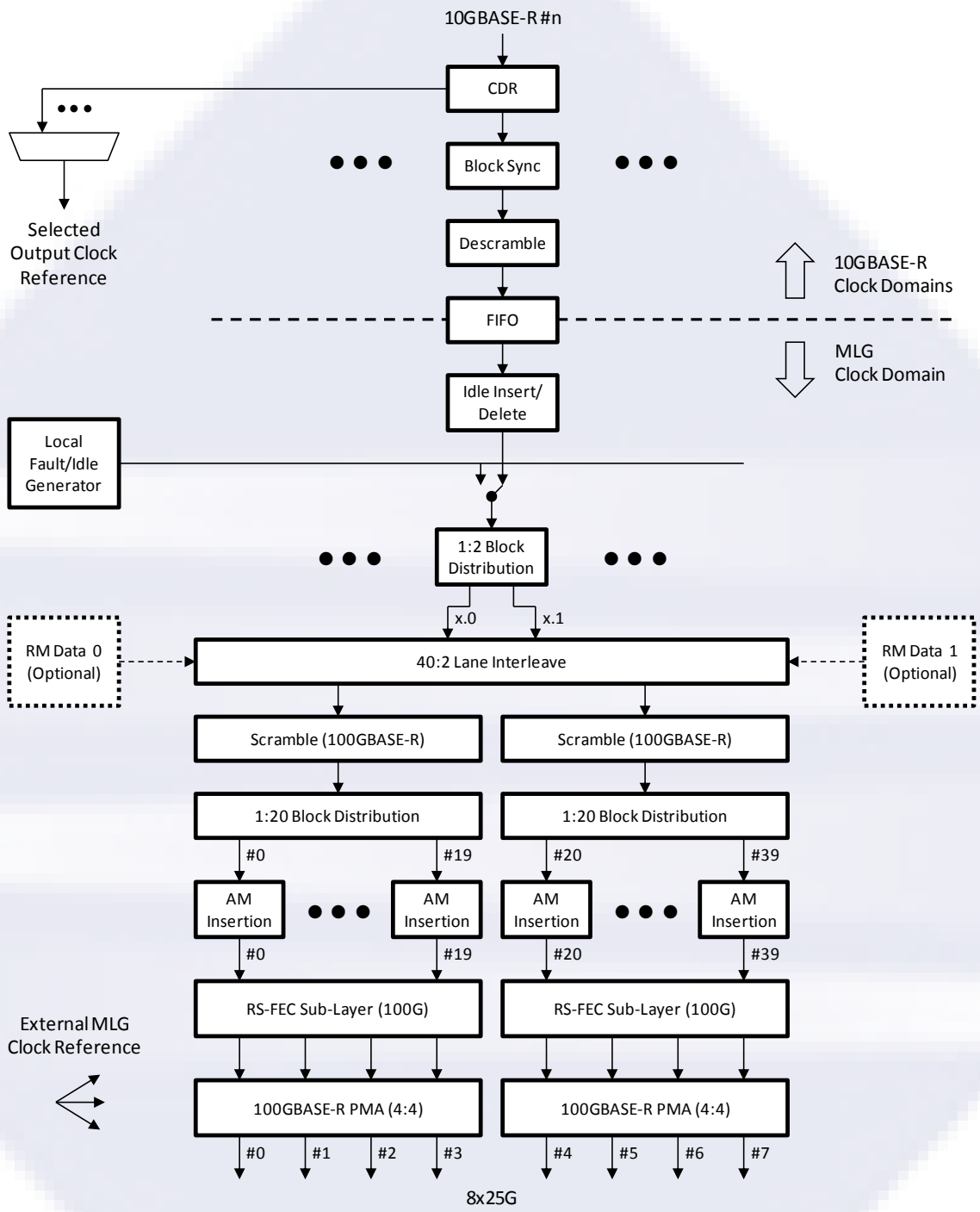


Figure 47: MLG200 mux 10GBASE-R mapping Block Diagram

8.4.1.1 Clock and Data Recovery (CDR)

Each of the 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.

8.4.1.2 Block Sync

Once clock and data have been recovered from a 10GBASE-R signal, 66-bit block synchronization is obtained. This is done per the state diagram in Figure 49-14 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.

8.4.1.3 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.

8.4.1.4 Idle Insert/Delete

Idles are inserted or deleted as necessary in order that the bitstream, after MLG lane marker and optional RM channel data 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 49.2.4.7.

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 (± 100ppm), 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.

8.4.1.5 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 8.4.1.1), 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 D_1, D_2, D_3 and D_5, D_6, D_7. 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.

8.4.1.6 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 8.4.1.9). 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.

8.4.1.7 2:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

8.4.1.8 RM data insertion and 20:1 Lane Interleave

See Section 8.3.1.8.

8.4.1.9 Scramble

See Section 8.3.1.9.

8.4.1.10 1:20 Block Distribution

See Section 8.3.1.10.

8.4.1.11 Alignment Marker Insertion

See Section 8.3.1.11.

8.4.1.12 RS-FEC Sublayer (Transmit)

See Section 8.1.2 for description of the general mechanism. A MLG200 mux supporting RS-FEC encoded physical interfaces implements the transmit function within the RS-FEC Sublayer of IEEE Std 802.3bj-2014 on each of two 4x25G physical interfaces (0-3 and 4-7), except for the following:

- AM mapping and insertion mechanism described in Clause 91.5.2.6 is replaced with the function described in Section 8.1.2.

8.4.1.13 100GBASE-R PMA(4:4)

The four RS-FEC lanes produced by each RS-FEC sublayer instance are connected to physical medium via a PMA 4:4 to produce a 8-lane by 25.78125 Gb/s interface. Note that an implementation may support generation of any of the optional test patterns specified for a 100GBASE-R PMA across the MLG100 link: see IEEE Std 802.3-2012 Clause 83.5.10. If supported, the corresponding MDIO control variables must also be supported.

8.4.2 MLG200 mux for 40GBASE-R

The mapping for 40GBASE-R into eight MLG lanes is illustrated in Figure 48. These may be eight of the forty MLG lanes of a MLG200.

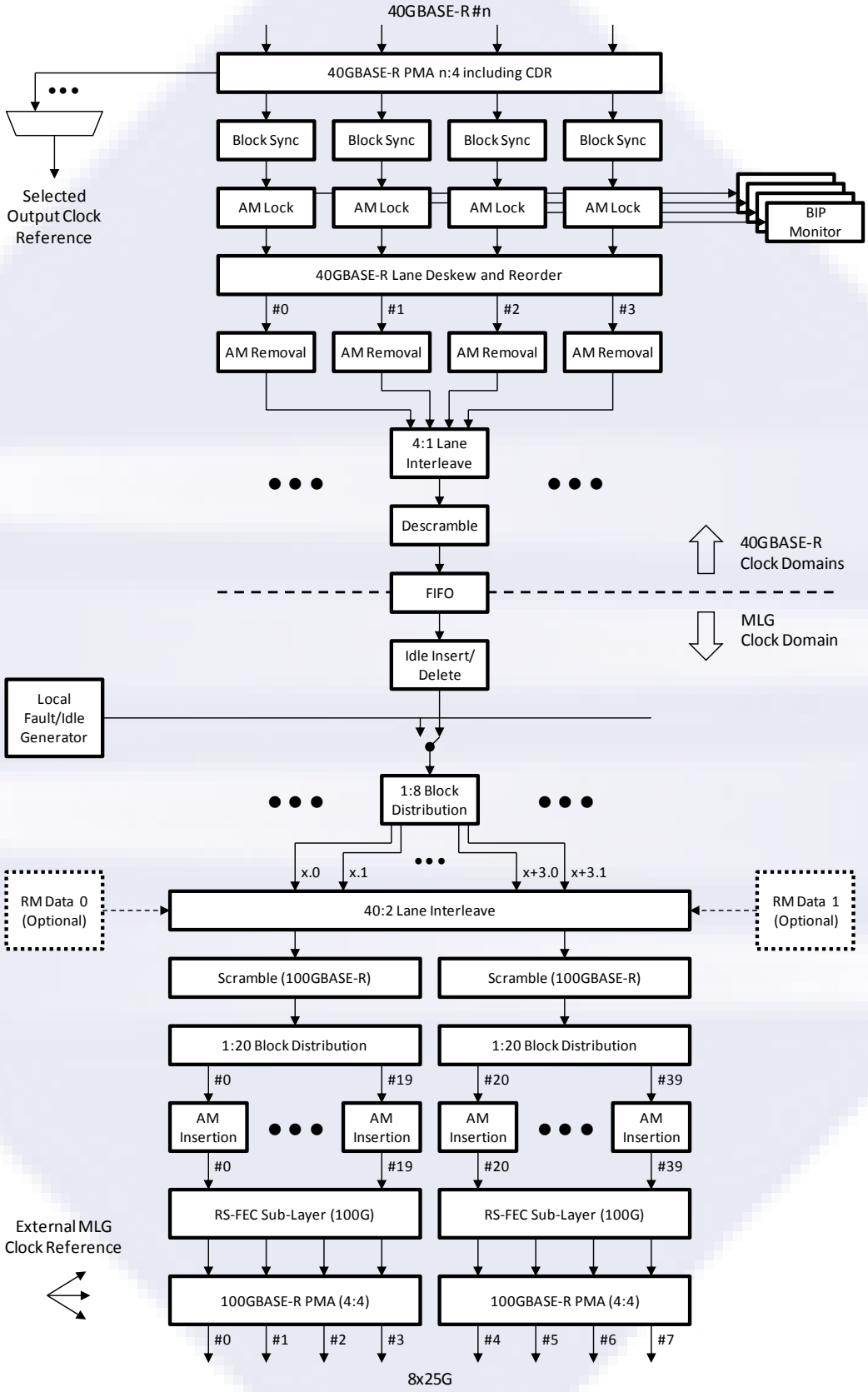


Figure 48: MLG200 mux 40GBASE-R mapping Block Diagram

8.4.2.1 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.

8.4.2.2 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.

8.4.2.3 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.

8.4.2.4 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.

8.4.2.5 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.

8.4.2.6 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.

8.4.2.7 Idle insert/delete

Idles are inserted or deleted as necessary in order that the bitstream, after PCS lane marker and RM channel data 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 (± 100ppm), 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.

8.4.2.8 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 8.4.2.1), 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 D_1, D_2, D_3, and zeros in Z_4, Z_5, Z_6, Z_7. 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.

8.4.2.9 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 8.4.2.12). 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.

8.4.2.10 8:1 Block Distribution

The unscrambled 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 twenty PCS lanes.

8.4.2.11 RM data insertion and 20:1 Lane Interleave

See Section 8.3.1.8.

8.4.2.12 Scramble

See Section 8.3.1.9.

8.4.2.13 1:20 Block Distribution

See Section 8.3.1.10.

8.4.2.14 Alignment Marker Insertion

See Section 8.3.1.11.

8.4.2.15 RS-FEC Sublayer

See Section 8.4.1.12. 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.

8.4.2.16 100GBASE-R PMA(4:4)

See Section 8.4.1.13. 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.

8.4.3 MLG200 demux for 10GBASE-R

The process for demapping a 10GBASE-R from two MLG lanes is shown in Figure 49.

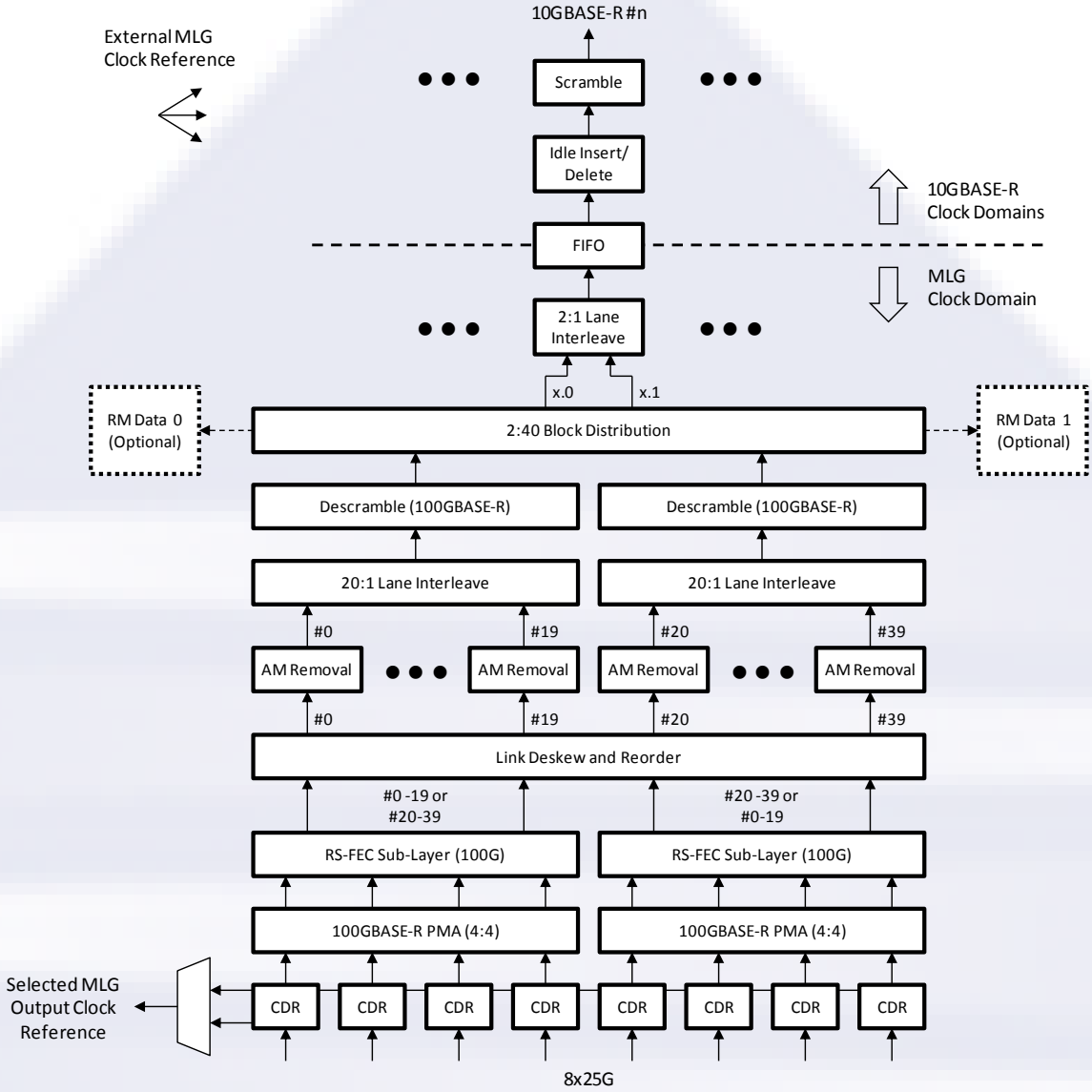


Figure 49: MLG200 demux 10GBASE-R demapping Block Diagram

8.4.3.1 100GBASE-R PMA (4:4)

The input to the MLG demux will come from a physical interface that is similar to that of double width (8x25G) RS-FEC for the MLG200 application. 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.

8.4.3.2 RS-FEC Sublayer (Receive)

See Section 8.1.2 for description of the general mechanism. A MLG200 demux supporting RS-FEC encoded physical interfaces implements the receive function within the RS-FEC Sublayer of IEEE Std 802.3bj-2014 on each of two 4x25G physical interfaces (0-3 and 4-7), except for the following:

- Since the MLG200 demux payload may be from a 10GE client, the block type field restoration of step f2 of Clause 91.5.3.5 shall be cross-referenced using Figure 49-7 of IEEE Std 802.3-2012, which is a superset of Figure 82-5.
- The AM mapping and insertion mechanism described in Clause 91.5.3.7 is replaced with the mechanism described in Section 8.1.2. Each of the two RS-FEC receive functions will receive either MLG lanes 0-19 or 20-39.

8.4.3.3 Link Deskew and Reorder

Each of the two RS-FEC sublayers will produce a 100G link consisting of either MLG lanes 0-19 or MLG lanes 20-39. In order to restore the appropriate 10GE and 40GE data to the proper clients, it may be necessary to swap (reorder) the two 100G MLG links. To support 40GE client payloads on MLG lanes 16-23, the two 100G links must also be deskewed, utilizing MLG lane alignment markers, in order to maintain 40GE data coherency across the 8 MLG lanes.

8.4.3.4 MLG Lane Alignment marker removal

See Section 8.3.3.6.

8.4.3.5 20:1 Lane Interleave

See Section 8.3.3.7.

8.4.3.6 Descramble

See Section 8.3.3.8.

8.4.3.7 1:20 Block Distribution RM data removal

See Section 8.3.3.9.

8.4.3.8 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.

8.4.3.9 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.

8.4.3.10 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 49.2.4.7.

Note, each 10GBASE-R signal may 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 (±100ppm), 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).

8.4.3.11 Scramble

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

8.4.4 MLG200 demux for 40GBASE-R

The process for demapping a 40GBASE-R from eight MLG lanes is shown in Figure 50.

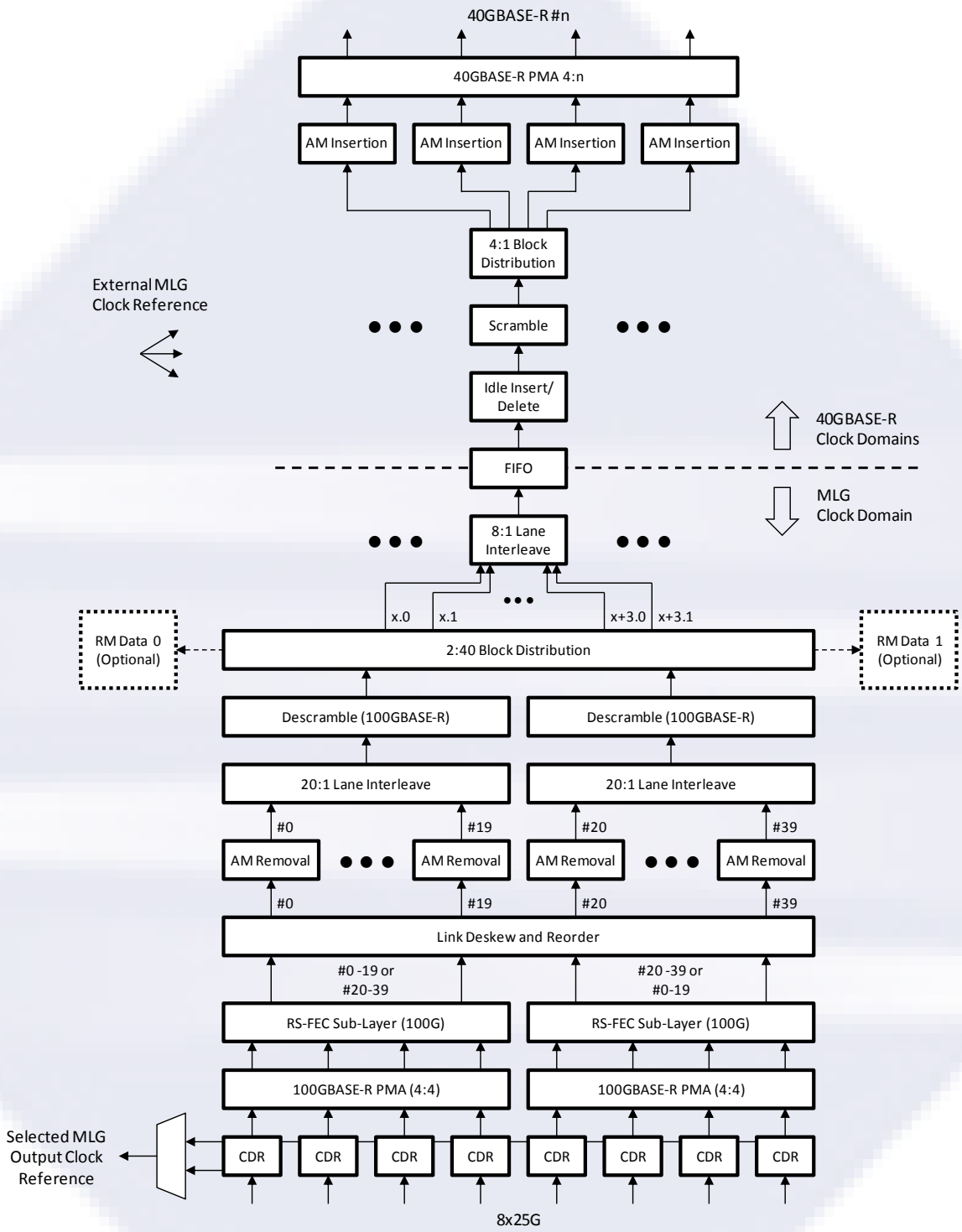


Figure 50: MLG200 demux 40GBASE-R demapping Block Diagram

8.4.4.1 100GBASE-R PMA(4:4)

See Section 8.4.3.1. 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.

8.4.4.2 RS-FEC Sublayer

See Section 8.4.3.2. 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.

8.4.4.3 MLG Lane Alignment marker removal

See Section 8.3.3.6.

8.4.4.4 20:1 Lane Interleave

See Section 8.3.3.7.

8.4.4.5 Descramble

See Section 8.3.3.8.

8.4.4.6 1:20 Block Distribution RM data removal

See Section 8.3.3.9.

8.4.4.7 MLG Lane Interleave

The eight 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.

8.4.4.8 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 may 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 (±100ppm), 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).

8.4.4.9 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.

8.4.4.10 4:1 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.

8.4.4.11 Alignment Insertion

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

8.4.4.12 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.

9 MLG Management

9.1 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 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 7.2.1.5 for 10GBASE-R or 7.2.2.8 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_4x10G_ability – indicates whether the MLG mux supports the MLG40 application
- MLG_mux_4x25G_ability – indicates whether the MLG mux supports the MLG100 application
- MLG_mux_8x25G_ability – indicates whether the MLG mux supports the MLG200 application
- MLG_mux_40G_ability – for a MLG100 and MLG200 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 only 10GBASE-R 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_4 through Signal_Detect_19 are not defined for the MLG40 mux application. Signal_Detect_10 through Signal_Detect_19 are not defined for the MLG100 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.

9.2 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. These are not valid in MLG40 application. Only MLG_demux_40G_select_0 and MLG-demux_40G_select_4 are valid for the MLG100 application. All are valid for the MLG200 application.
- 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_4 through MLG_demux_10G_Enable_19 are not valid for MLG40 application. MLG_demux_10G_Enable_10 through MLG_demux_10G_Enable_19 are not valid for MLG100 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_4 through MLG_demux_scrambled_idle_enable_19 are not valid for MLG40 application. MLG_demux_scrambled_idle_enable_10 through MLG_demux_scrambled_idle_enable_19 are not valid for MLG100 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 40GBASE-R or 100GBASE-R PMA(s). Prior to the MLG lane reorder block, MLG lanes are known simply by the position 0..3 (for MLG40 application), 0..19 (for MLG100 application), or 0..39 (for MLG200 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..3 for MLG40, 0.0, 0.1, … 9.1 for MLG200, or 0.0, 0.1, 1.0, …, 19.1 for MLG200). 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_4x10G_ability – indicates whether the MLG demux supports the MLG40 application
- MLG_demux_4x25G_ability – indicates whether the MLG demux supports the MLG100 application

- MLG_demux_8x25G_ability – indicates whether the MLG demux supports the MLG200 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:4), 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_4 through block_lock_39 are not valid for MLG40 application. block_lock_20 through block_lock_39 are not valid for MLG100 application.
- am_lock_0 through am_lock_39 – indicates whether alignment marker lock has been achieved on each of the MLG lanes. am_lock_4 through am_lock_39 are not valid for MLG40 application. am_lock_20 through am_lock_39 are not valid for MLG100 application.
- MLG_demux_lane_alignment_status – indicates whether all 4 (MLG40), 20 (MLG100), or 40 (MLG200) MLG lanes have achieved MLG lane alignment marker lock, that the 4, 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_4 through BIP_error_counter_39 are not valid for MLG40 application. BIP_error_counter_20 through BIP_error_counter_39 are not valid for MLG100 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..3 for MLG40 application, 0.0, 0.1, 1.0, …, 9.1 for MLG100 application, and 0.0, 0.1, 1.0, …, 19.1 for MLG200 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_4 through MLG_demux_scrambled_idle_19 are not valid for MLG40 application. MLG_demux_scrambled_idle_error_10 through MLG_demux_scrambled_idle_19 are not valid for MLG100 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.

9.3 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_4 through MLG_10G_loopback_enable_19 are not valid for MLG40 application. MLG_10G_loopback_enable_10 through MLG_10G_loopback_enable_19 are not valid for the MLG100 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.

10 References

10.1 Normative references

[1] IEEE, *IEEE Standard for Ethernet*, IEEE Std 802.3-2012, December 28, 2012.

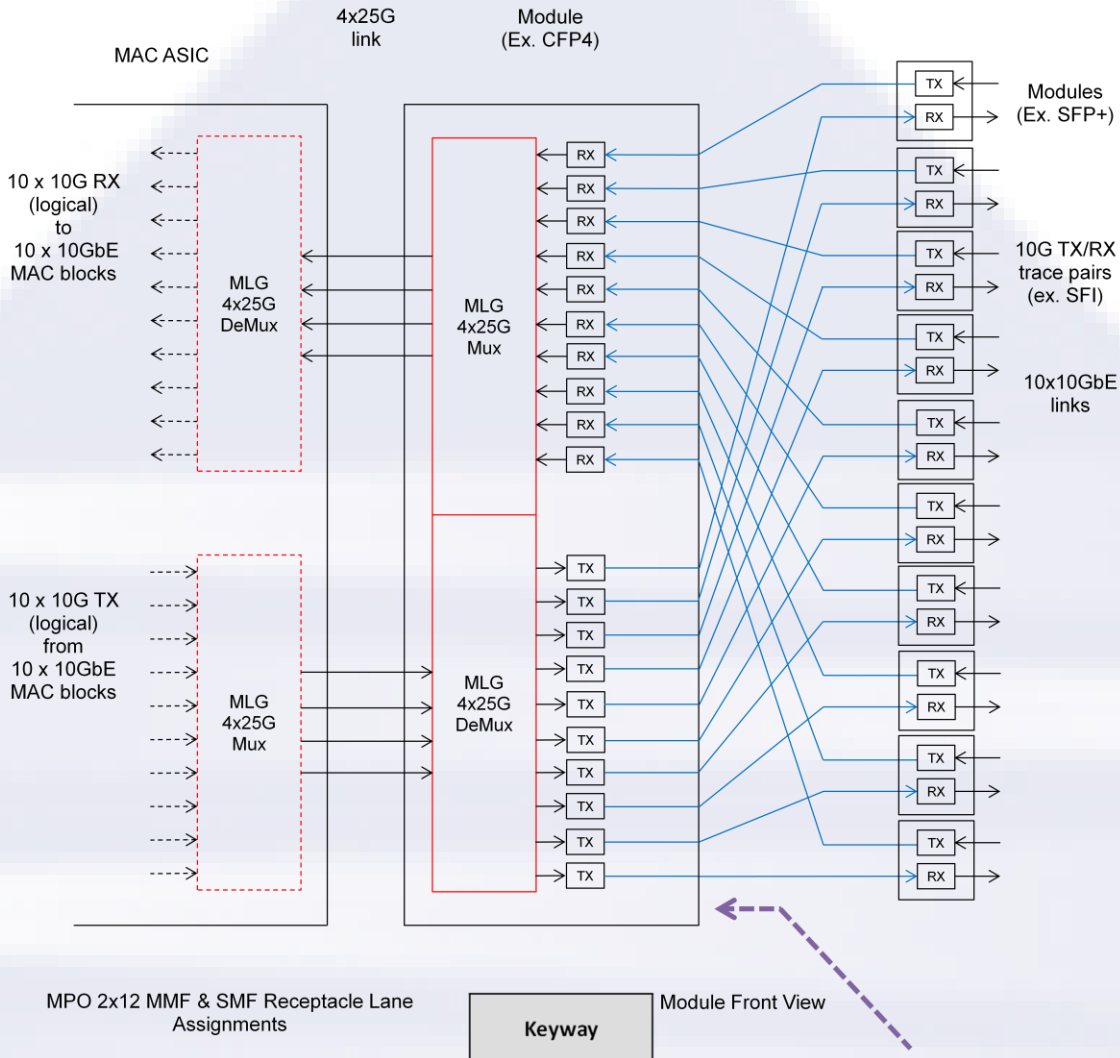
[2] IEEE, *Standard for Ethernet Amendment X: Physical Layer Specifications and Management Parameters for 100 Gb/s Operation Over Backplanes and Copper Cables*, IEEE Std 802.3bj-2014, June 12, 2014.

[3] IETF, *PPP in HDLC-like Framing*, RFC 1662, July 1994.

10.2 Informative references

[4] IEEE, *Standard for Ethernet Amendment 3: Physical Layer Specifications and Management Parameters for 40 Gb/s and 100 Gb/s Operation Over Fiber Optic Cables*, IEEE Std 802.3bm-2015, February 16, 2015.

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



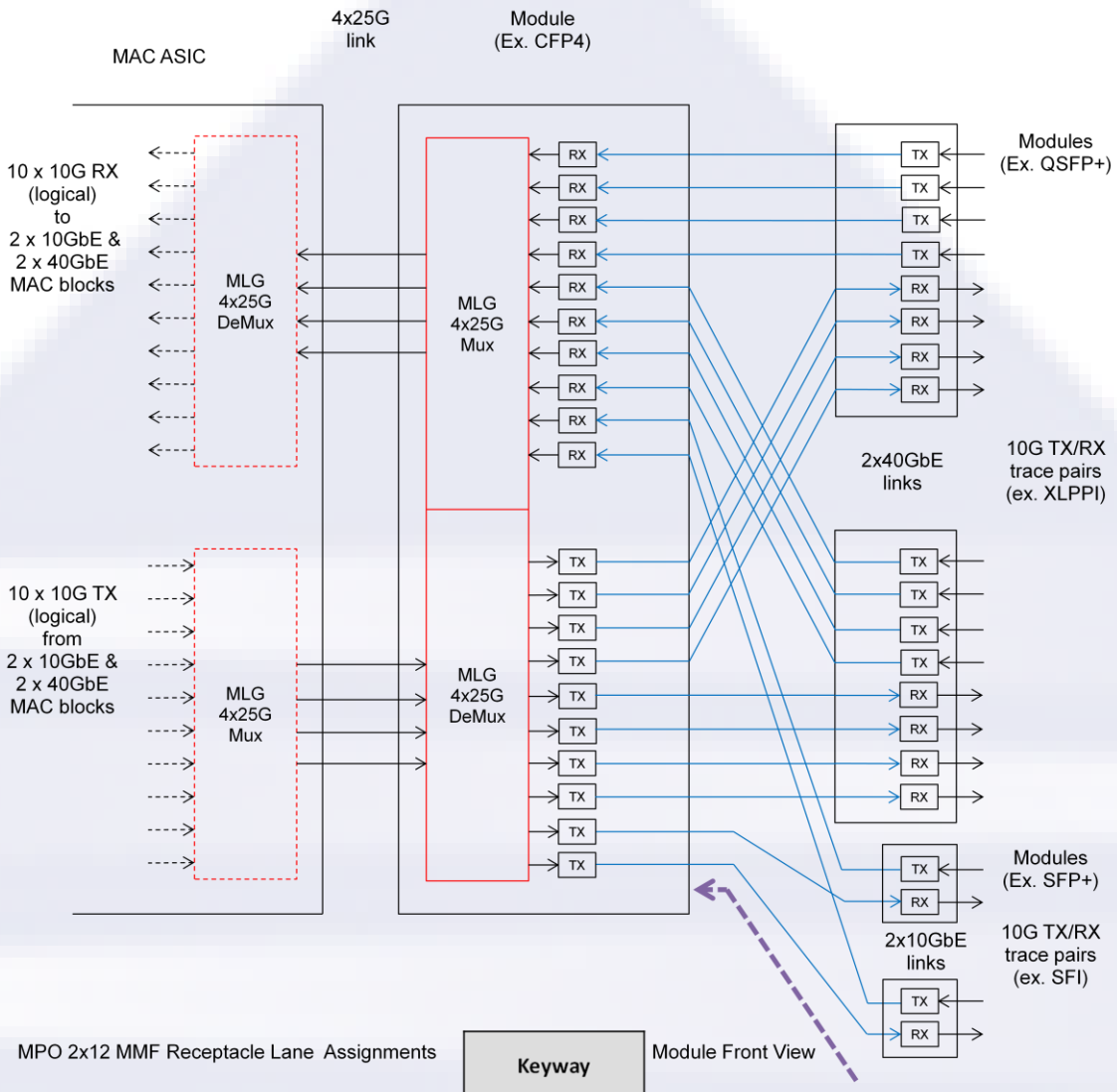
MPO 2x12 MMF & SMF Receptacle Lane Assignments

Keyway

Module Front View

	10GE_9 Rx	10GE_8 Rx	10GE_7 Rx	10GE_6 Rx	10GE_5 Rx	10GE_4 Rx	10GE_3 Rx	10GE_2 Rx	10GE_1 Rx	10GE_0 Rx	
	10GE_9 Tx	10GE_8 Tx	10GE_7 Tx	10GE_6 Tx	10GE_5 Tx	10GE_4 Tx	10GE_3 Tx	10GE_2 Tx	10GE_1 Tx	10GE_0 Tx	

Figure 51: MPO Receptacle Physical Lane Assignments for 4x25G MLG 10GBASE-SR or 10GBASE-LR Optical Port Expander



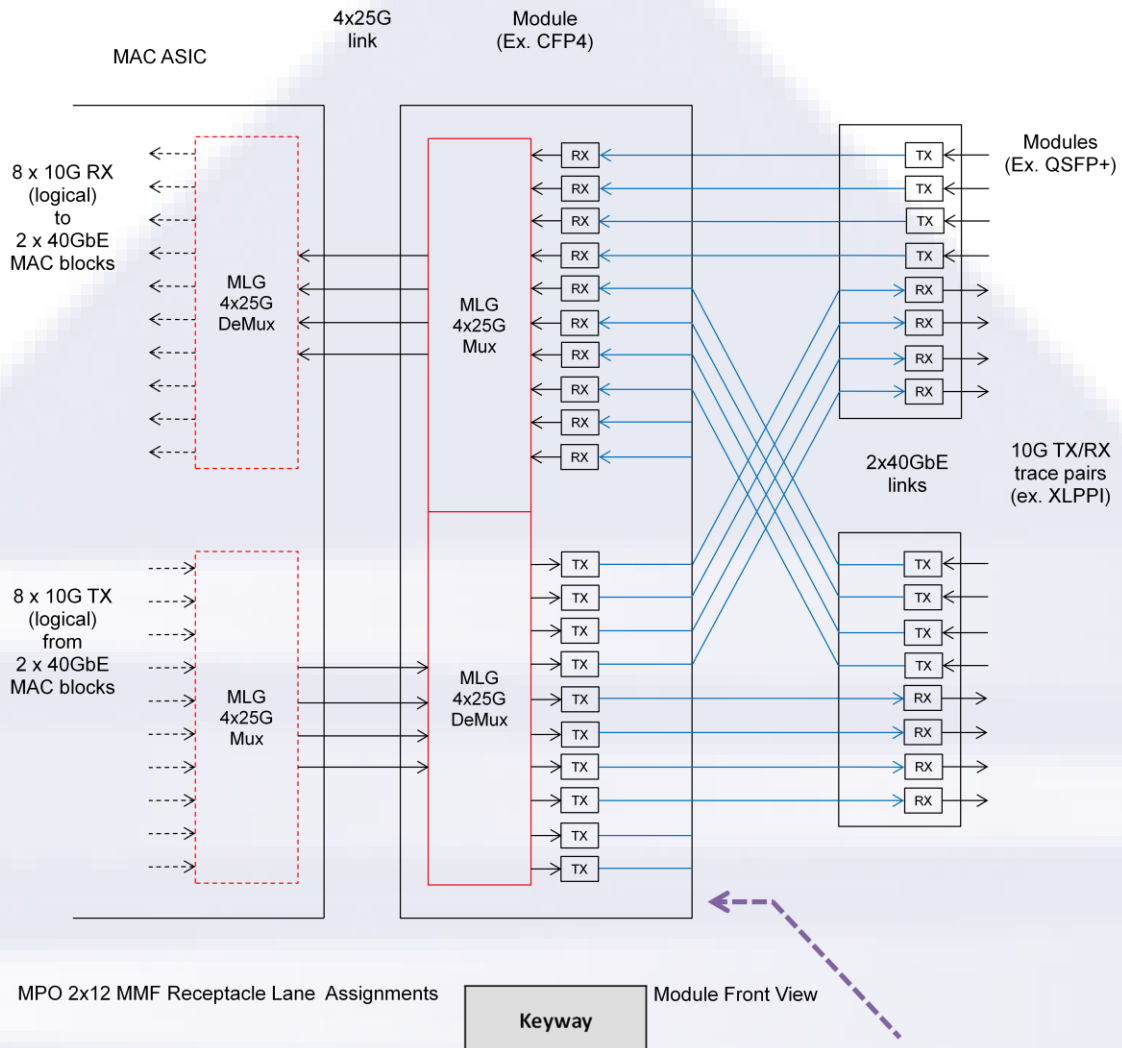
MPO 2x12 MMF Receptacle Lane Assignments

Keyway

Module Front View

	10GE_9 Rx	10GE_8 Rx	40GE_1 Rx	40GE_1 Rx	40GE_1 Rx	40GE_1 Rx	40GE_0 Rx	40GE_0 Rx	40GE_0 Rx	40GE_0 Rx	
	10GE_9 Tx	10GE_8 Tx	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 52: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 10GBASE-SR and 40GBASE-SR4 Port Expander



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

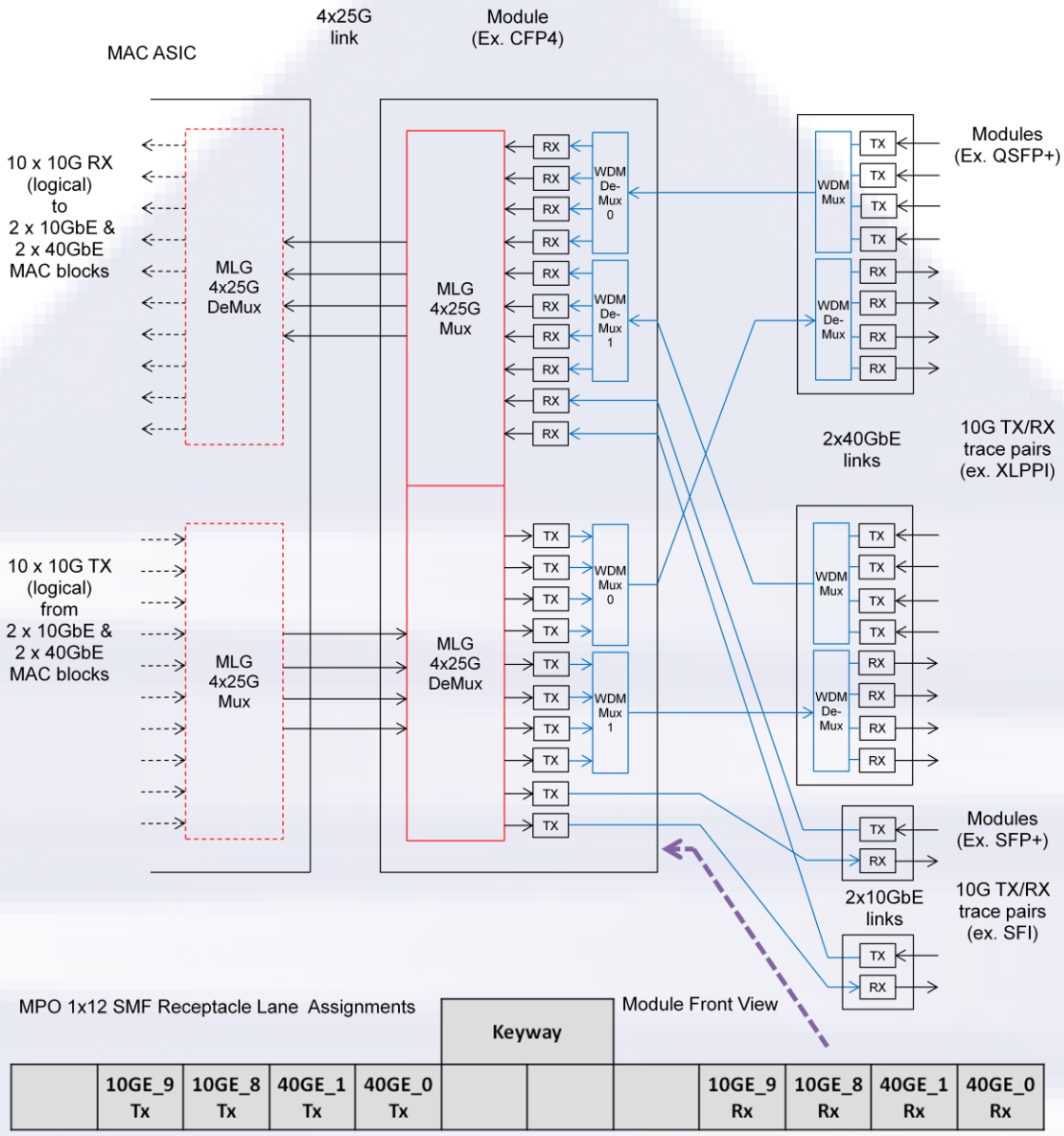


Figure 54: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 10GBASE-LR and 40GBASE-LR4 Port Expander

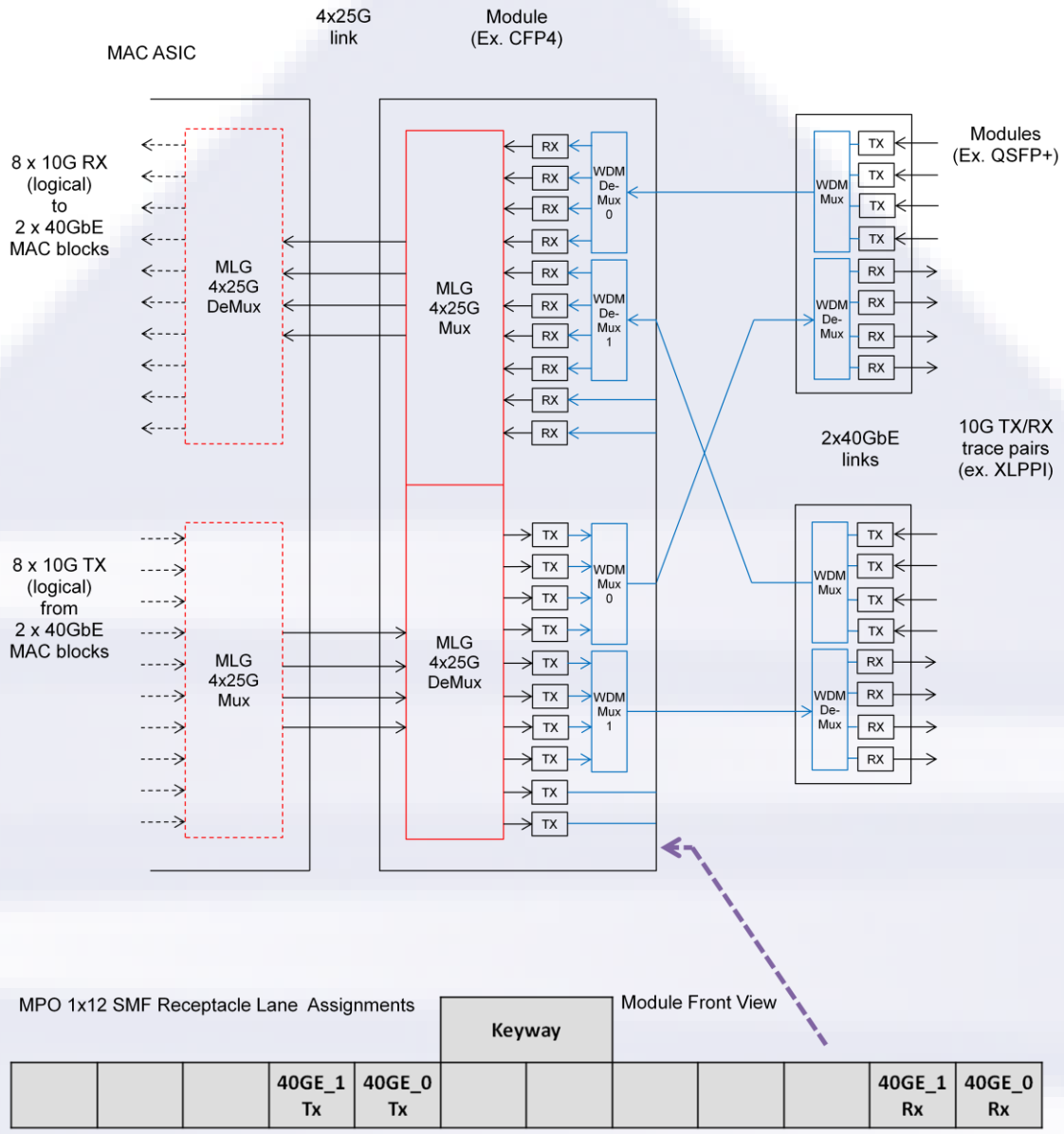


Figure 55: MPO Receptacle Physical Lane Assignments for 4x25G MLG Optical 40GBASE-LR4 Port Expander

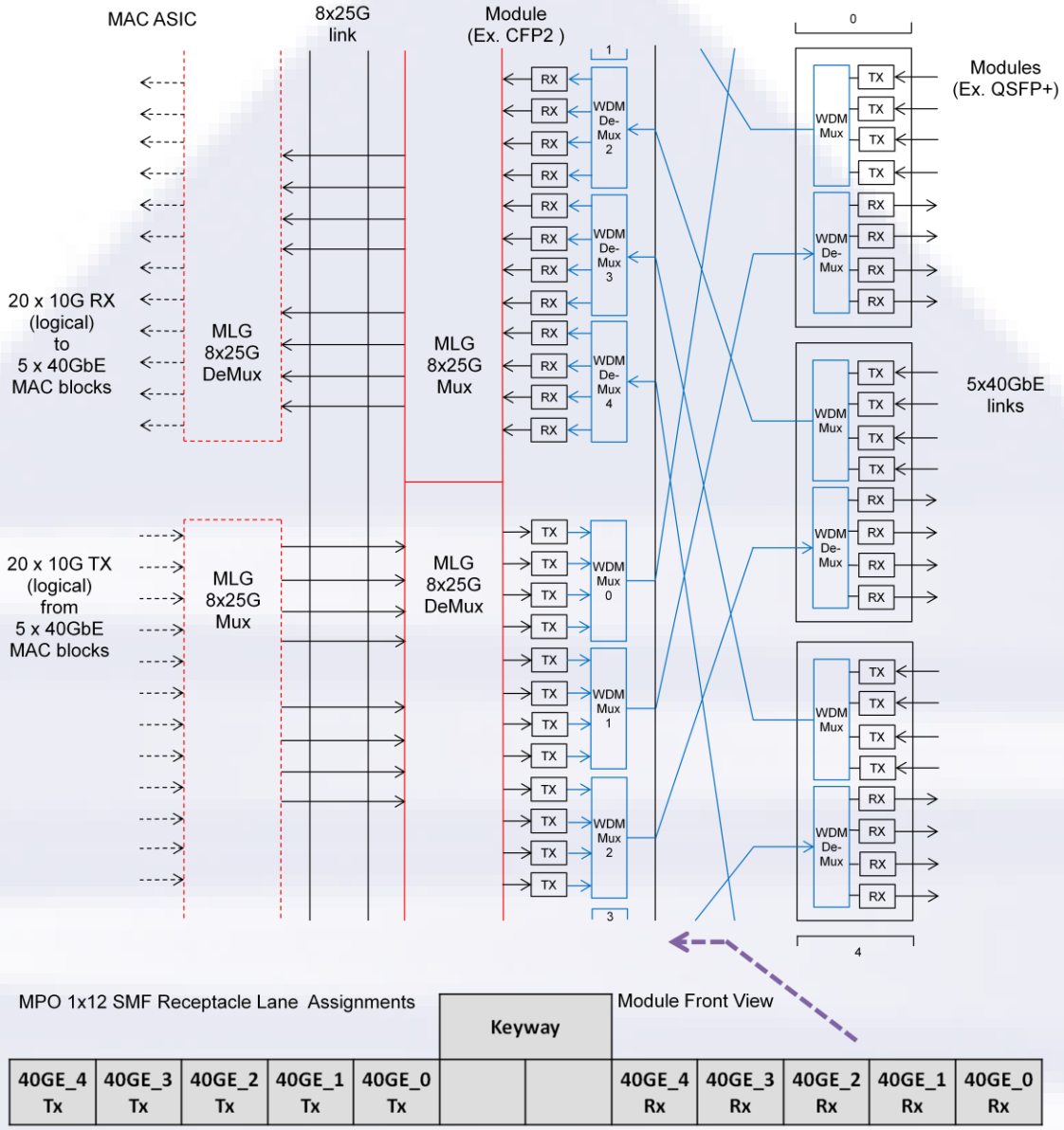


Figure 56: MPO Connector Physical Lane assignments for 8x25G Optical 40GBASE-R Port Expander

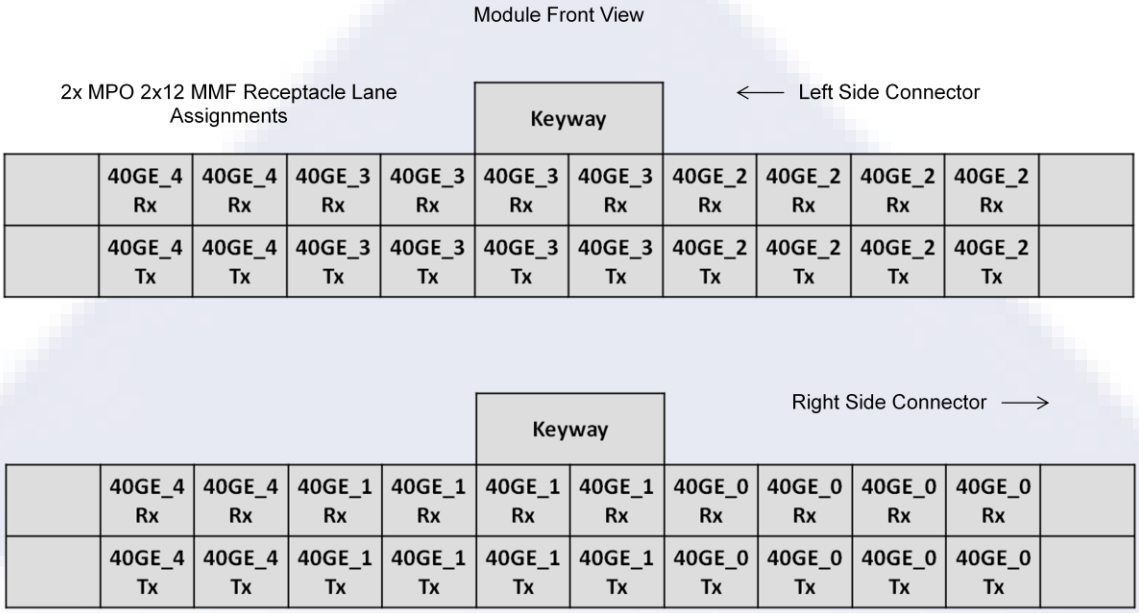


Figure 57: MPO Receptacle Physical Lane Assignments for 8x25G Optical 40GBASE-SR4 Port Expander

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

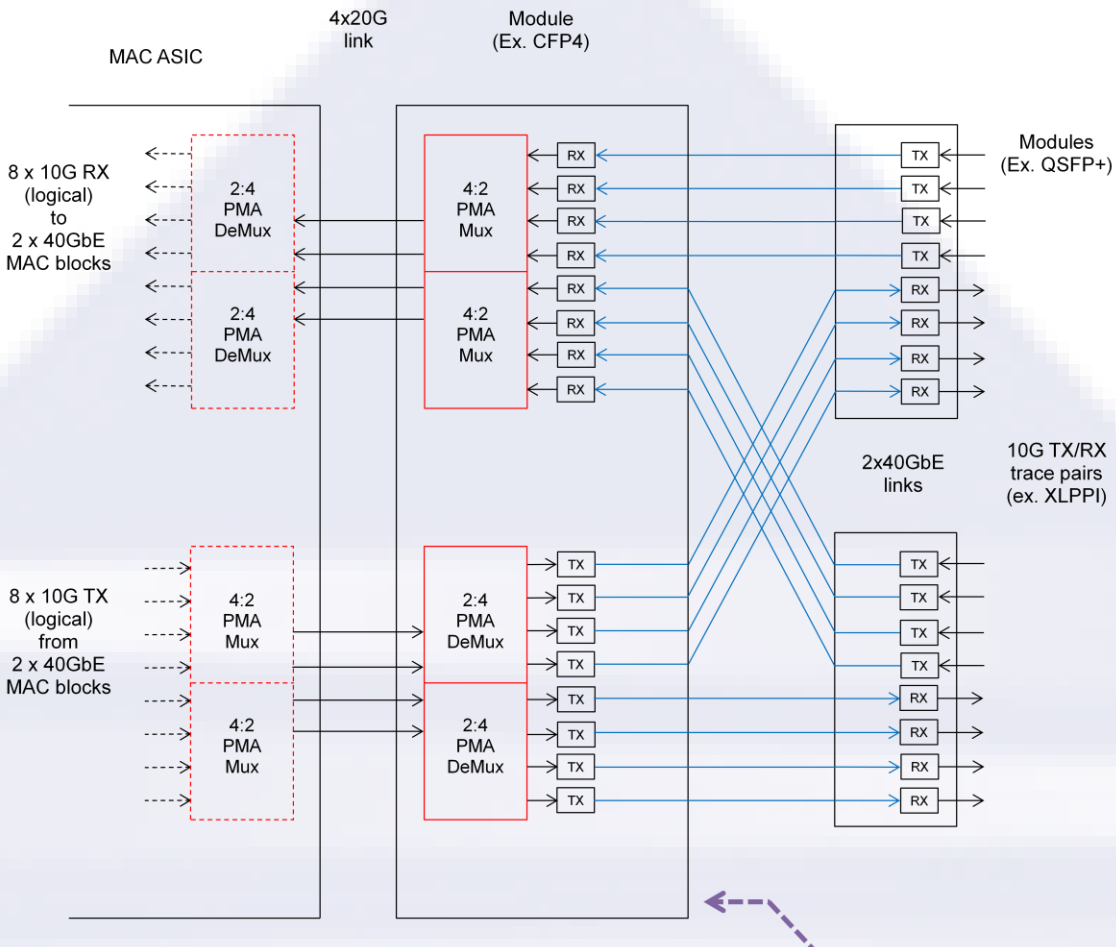
Applications which are restricted to the need to support two 40GBASE-R interfaces over a four-lane interface, do not need 10GBASE-R support, and hence would only result in 80% utilization of the 4x25G MLG specified in this document could consider using a simpler approach described in this Appendix.

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 58 and Figure 59 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 lane rate instead of a 25.78125 Gb/s lane rate is outside the scope of this implementation agreement.



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

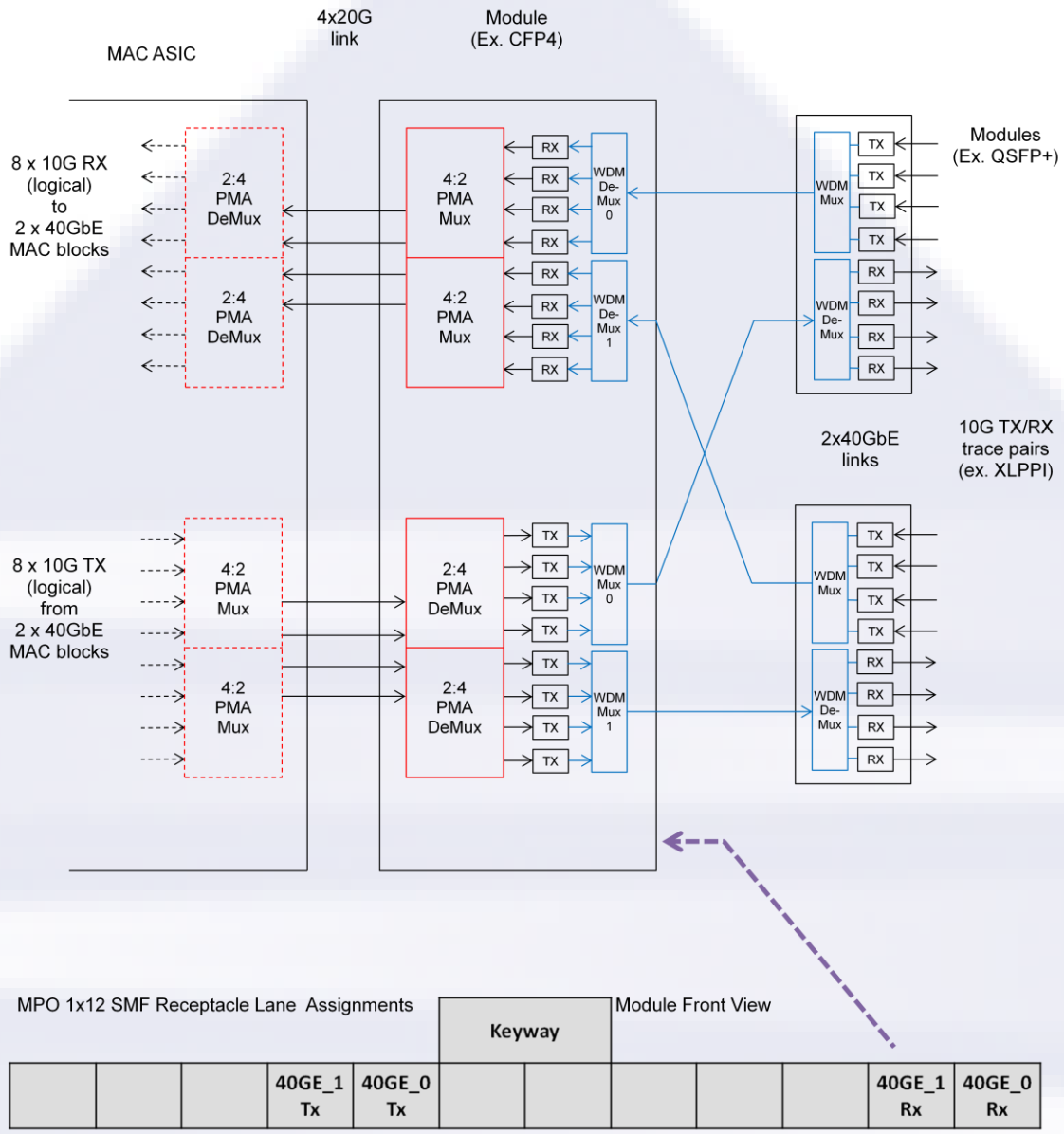


Figure 59: MPO Receptacle Physical Lane Assignments for 4x20G Optical 40GBASE-LR4 Port Expander

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

Acacia Communications
ADVA Optical Networking
Alcatel-Lucent
AMCC
Amphenol Corp.
Analog Devices
Anritsu
Broadcom Limited
Brocade
BTI Systems
China Telecom
Ciena Corporation
Cisco Systems
ClariPhy Communications
Compass Networks
Coriant
CPqD
Credo Semiconductor (HK) LTD
EMC Corporation
Ericsson
ETRI
FCI USA LLC
Fiberhome Technologies Group
Finisar Corporation
Fujikura
Fujitsu
Furukawa Electric Japan
GigOptix Inc.
Global Foundries
Google
Hitachi
Infinera
Inphi
Intel
Ixia
Juniper Networks
Kandou Bus
KDDI R&D Laboratories
Keysight Technologies, Inc.
Leaba
Lumentum
Luxtera
M/A-COM Technology Solutions
Marvell Technology
Mellanox Technologies
Microsemi Inc.
Microsoft Corporation
Mitsubishi Electric Corporation
Molex
MoSys, Inc.
NEC
NeoPhotonics
NTT Corporation
O-Net Communications (HK) Limited
Oclaro
Orange
PETRA
Picometrix
QLogic Corporation
Qorvo
Rockley Photonics
Samtec Inc.
Semtech
Socionext Inc.
Spirent Communications
Sumitomo Electric Industries
Sumitomo Osaka Cement
TE Connectivity
Tektronix
TELUS Communications, Inc.
TeraXion
Texas Instruments
Time Warner Cable
US Conec
Verizon
Viavi Solutions Deutschland GmbH
Xilinx
Yamaichi Electronics Ltd.
ZTE Corporation