Implementation Agreement for 800ZR Coherent Interfaces

OIF-800ZR-01.0

October 8, 2024

Implementation Agreement created and approved

OIF

www.oiforum.com

The OIF is an international non-profit organization with over 150 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 EA, IEEE 802.3, INCITS T11, Infiniband, IPEC, ITU SG-15, PCI-SIG, SNIA SFF.

For additional information contact:
OIF
39221 Paseo Padre Pkwy, Suite J
Fremont, CA 94538 USA

+1.510.392.4903 info@oiforum.com

www.oiforum.com

Working Group: **Physical and Link Layer (PLL) Working Group**

TITLE: **800ZR Coherent Implementation Agreement**

SOURCE:

TECHNICAL EDITOR

Kishore Kota
Marvell Technology, Inc
15485 Sand Canyon Ave
Irvine, CA 92618
Phone: +1 949-864-5513
Email: kkota@marvell.com

WORKING GROUP CHAIR

David R. Stauffer, Ph D.
Kandou Bus, SA
EPFL Innovation Park, Building I
1015 Lausanne, Switzerland
Phone: +1.802.316.0808
Email: david@kandou.com

OPTICAL VICE CHAIR

Karl Gass
Phone: +1.505.301.1511
Email: iamthedonutking@mac.com

ABSTRACT: This Implementation Agreement specifies requirements for 800ZR coherent interfaces intended for single-span amplified DWDM links for datacenter interconnect (DCI).

Table of Contents

List of Figures

List of Tables

1 Introduction

This implementation agreement (IA) defines a single-wavelength 800G coherent line interface and frame format for single-span, amplified, 80-120km, point-to-point, DWDM noise limited links (e.g. data center interconnect applications).

The 800G coherent line interfaces defined in this IA are designed to support Ethernet clients (minimum 100GE) up to 800G aggregate bandwidth. The scope of this IA includes definitions of Ethernet client mappings (including muxing of lower-rate services into the 800G frame), definition of the 800G frames, forward-error correction (FEC)/modulation and optical specifications for these interfaces. The IA aims to enable interoperable, cost-effective, low power 800ZR implementations in small form-factor pluggable modules with port densities equivalent to client optics. No restriction on the physical form-factor of the modules is implied by this IA.

Figure 1 is a reference diagram for the 800ZR interface. The IA defines a single-channel line interface which operates on a black link which is defined only in terms of the channel characteristics in this implementation agreement.

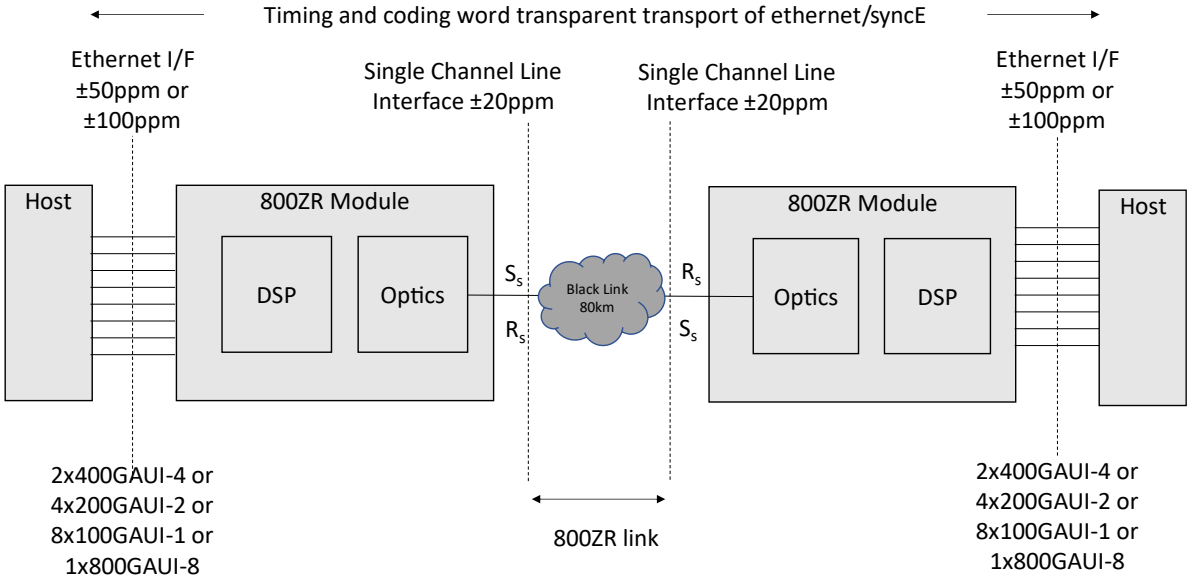


Figure 1: 800ZR Reference Diagram

2 800ZR Use cases

There are two use cases for 800ZR amplified DWDM noise-limited point-to-point links (no optical add/drop multiplexors) shown in Figure 2 and Figure 3. For amplified links the reach is dependent on the OSNR achieved at the receiver (noise-limited). The 800ZR targeted reach is 80-120km.

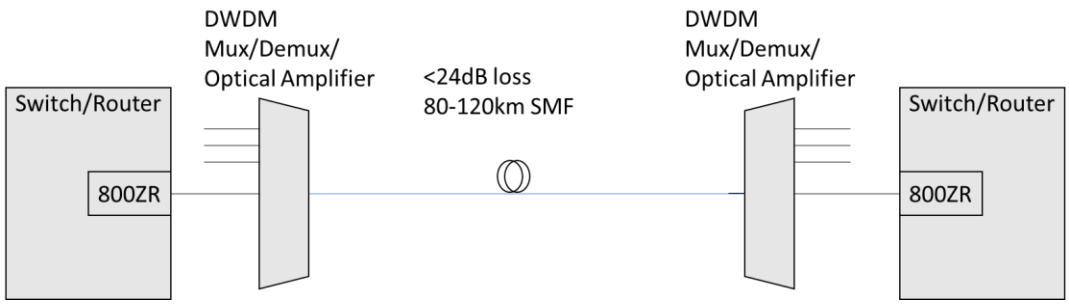


Figure 2: Switch/Router line card with 800ZR DWDM interfaces

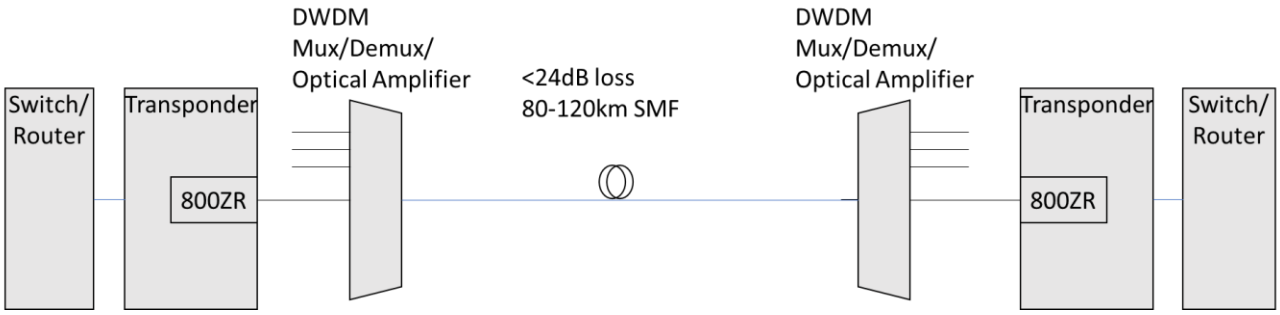


Figure 3: Transceiver line card with 800ZR interfaces

3 800ZR Client Interfaces

An 800ZR interface supports one or more Ethernet clients (minimum 100GE) with up to 800G aggregate bandwidth. The clients shown in Table 1 are supported by the muxponder specifications defined in Section 4. Support for multiple clients of different types is beyond the scope of this IA.

Client Type	Chip-to-Module interface	Number of clients
100GBASE-R	100GAUI-1	8
200GBASE-R	200GAUI-2	4
400GBASE-R	400GAUI-4	2
800G-ETC-R	800G-ETC-R	1
800GBASE-R	800GAUI-8	1

Table 1: Client types and interfaces

NOTE – 800G-ETC-R is functionally compatible to 800GBASE-R from [IEEE P802.3df]. The IA will generically use 800GE to refer to both variants, unless specifically called out.

4 Client Mapping to 800ZR frame

The 800ZR frame is the container used for client mapping purposes prior to the 800ZR FEC and DSP framing functions. The 800ZR frame and its DSP frame structure use the FlexO-8e and FlexO-8e-DO-16QAM frame structures defined in [ITU-T G.709.1] and [ITU-T G.709.6], with some of the FlexO overhead fields not used or set to fixed values because they are not necessary in the 800ZR application. Like the FlexO-8e frame structure over FlexO-8e-DO-16QAM interface, the 800ZR frame structure is formed through 128-bit interleaving of eight 100G FlexO instances, referred to in this document as 100G ZR instances. The 100G ZR instances carry payload, overhead and stuff bits to enable mapping and de-mapping of the clients listed in Table 1. The frame structure for each 100G ZR instance is described in Section 4.3.

The 800ZR coherent interface uses the Generic Mapping Procedure (GMP) to map Ethernet clients to the 800ZR frame. The complete process to map clients to the 800ZR frame is shown in Figure 4.

NOTE: Throughout this document transmit direction refers to client to line and receive direction refers to line to client.

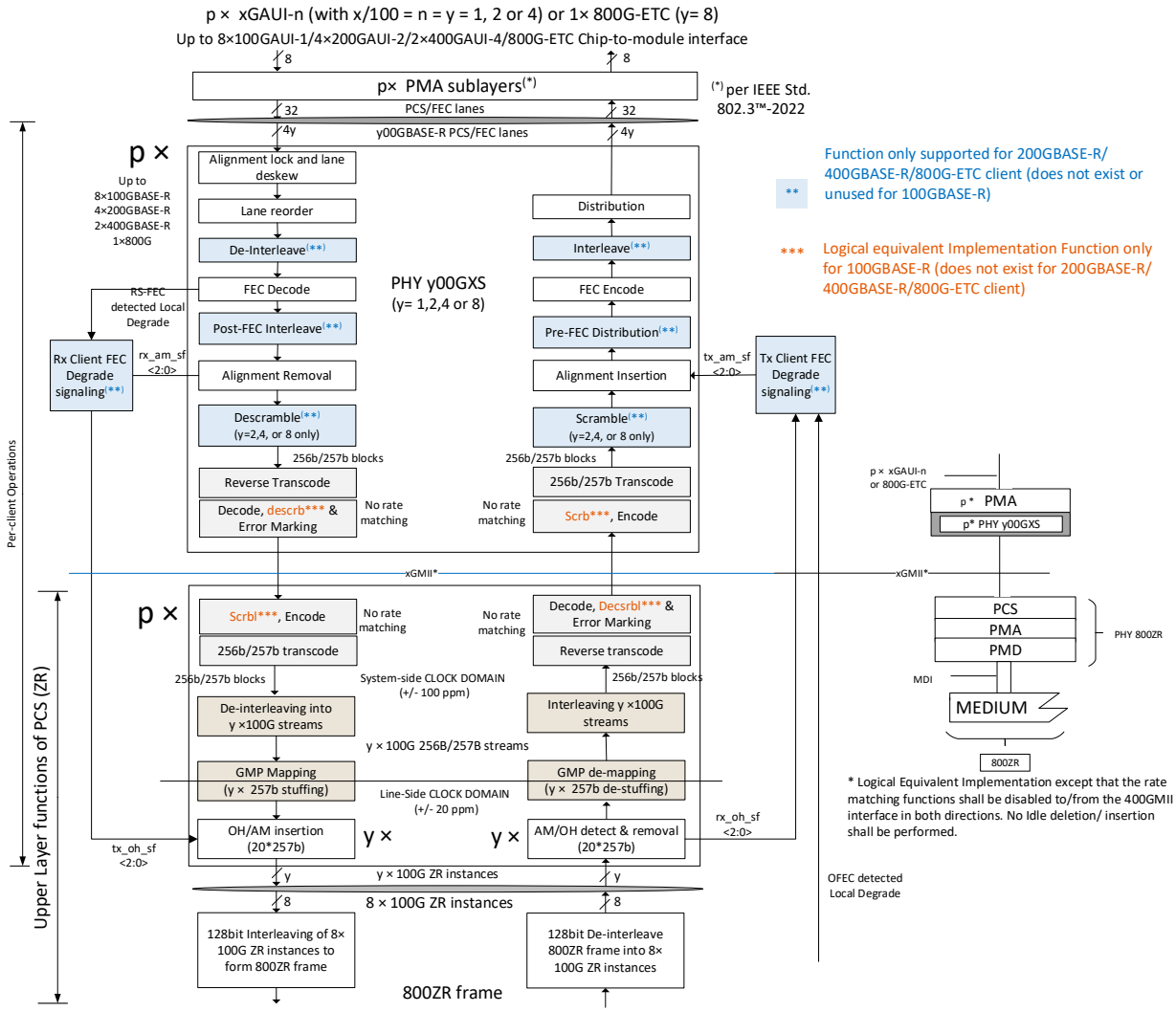


Figure 4: 100GE/200GE/400GE/800GE Clients to 800ZR data path

4.1 Client Adaptation Processes

The client mapping process for each 200GE/400GE/800GE client into 100G ZR instances includes the following steps for transmit towards 800ZR line:

1. PMA Adaptation
2. Alignment lock and lane de-skew
3. Lane re-order and de-interleave
4. FEC Decode and generation of “Local Degrade” (LD) signal
5. Post-FEC interleave
6. Alignment marker removal and extraction of rx_am_sf<2:0> fields
7. Descramble
8. De-interleave of 257b blocks of the client into aligned 100G streams
9. GMP procedure to map to multi-frame aligned 100G ZR instances

10. Overhead and Alignment Mechanism (AM) insertion into each 100G ZR instance

The client mapping process for each 100GE client into a 100G ZR instance includes the following steps for transmit towards 800ZR line:

1. PMA Adaptation
2. Alignment lock and lane de-skew
3. Lane re-order and de-interleave
4. FEC Decode
5. Alignment marker removal
6. GMP procedure to map to 100G ZR instance
7. Overhead and Alignment Mechanism (AM) insertion into 100G ZR instance

The 800ZR frame is always formed from 128-bit interleaving of eight aligned 100G ZR instances carrying overhead and payload. If some of the clients are not provisioned, the corresponding 100G ZR instances will contain overhead to indicate that the frame does not carry any client payload and the content of the payload area is undefined in this IA.

4.1.1 PMA Adaptation

This implementation agreement supports the client interfaces listed in Table 1. All these clients use 100Gb/s per lane chip-to-module interfaces defined in Annex 120G of [IEEE 802.3ck] and Annex 120G of [IEEE 802.3df].

The 100GAUI-1 signal for a 100GBASE-R client is processed in the transmit direction using the PMA procedures described in Clause 83 of [IEEE 802.3], as amended by [IEEE 802.3ck], to extract four FEC lanes (100GBASE-R) for each client. The inverse process is used in the receive direction.

The 200GAUI-2 signal for a 200GBASE-R client is processed in the transmit direction using the PMA procedures described in Clause 120 of [IEEE 802.3], as amended by [IEEE 802.3ck], to extract eight PCS lanes (200GBASE-R) for each client. The inverse process is used in the receive direction.

The 400GAUI-4 signal for each 400GBASE-R client is processed in the transmit direction using the PMA procedures described in Clause 120 of [IEEE 802.3], as amended by [IEEE 802.3ck], to extract sixteen PCS lanes (400GBASE-R) for each client. The inverse process is used in the receive direction.

The 800GAUI-8 signal for an 800GBASE-R client is processed in the transmit direction using the PMA procedures described in Clause 173 of [IEEE 802.3df], to extract thirty-two PCS lanes (800GBASE-R). The inverse process is used in the receive direction.

The 800G signal for an 800G-ETC-R client is processed in the transmit direction using the PMA procedures described in [ETC-800G], to extract thirty-two PCS lanes (800G-ETC-R). The inverse process is used in the receive direction.

4.1.2 Alignment lock, lane de-skew, lane reorder and de-interleaving

The PCS/FEC lanes of each client are individually processed in the transmit direction to achieve alignment lock on each individual lane. Following alignment lock, the lanes are de-skewed to align all lanes of each client signal and the lanes are re-ordered and de-interleaved. The inverse processes are used in the receive direction.

The lanes of 100GBASE-R client are processed using procedures described in Clause 91 of [IEEE 802.3] with amendments described in [IEEE 802.3ck].

The lanes of 200GBASE-R and 400GBASE-R clients are processed using the procedures described in Clause 119 of [IEEE 802.3] with amendments described in [IEEE 802.3ck].

The lanes of an 800GBASE-R client are processed using the procedures described in Clause 172 of [IEEE 802.3df].

The lanes of an 800G-ETC-R client are processed using the procedures described in [ETC-800G] specification.

4.1.3 FEC Decode/Encode

In the transmit direction, the FEC code-words of each client signal are decoded to correct for any accumulated errors. A local degrade signal is generated for each 200GBASE-R, 400GBASE-R, 800GBASE-R, or 800G-ETC-R client signal for insertion into the 800ZR frame. The inverse process of FEC encoding is implemented in the receive direction. Further details of processing the local degrade signal are described in Section 4.7.3.

For 100GBASE-R clients, the Reed-Solomon Forward Error Correction (FEC) sublayer is described in Clause 91 of [IEEE 802.3] for RS(544,514) with amendments described in [IEEE 802.3ck].

For 200GBASE-R and 400GBASE-R clients, the Reed-Solomon Forward Error Correction (FEC) is described in Clause 119 of [IEEE802.3] with amendments described in [IEEE 802.3ck].

For an 800GBASE-R client, the Reed-Solomon Forward Error Correction (FEC) is described in Clause 172 of [IEEE 802.3df].

For 800G-ETC-R client the processing is described in the [ETC-800G] specification.

4.1.4 Alignment Marker removal/insertion and Interleaving/Deinterleaving

In the transmit direction, the decoded FEC code-words are re-interleaved to form a stream of 257b blocks. For 100GBASE-R, 200GBASE-R, or 400GBASE-R clients, the order and bit ordering of these blocks shall match the ordering in the signal prior to FEC decoding. For 800G-ETC-R or 800GBASE-R clients, the order and bit ordering of the 257b blocks within each 400G flow shall match the ordering prior to FEC decoding and the 800G stream of 257b blocks shall be formed by round-robin interleaving of the two 400G streams of 257b blocks.

The 257b blocks containing alignment markers (AM) and pad bits are removed from the client stream, without inserting any idle blocks for rate matching. For 100GBASE-R clients, the AM fields contain BIP counters which are discarded along with the AM. For all other clients, the client am_sf<2:0> bits are extracted prior to AM removal for transmission in the overhead fields of the 800ZR frame. The 257b blocks of 200GBASE-R/400GBASE-R clients are descrambled following the procedures in Clause 119.2.5.6 of [IEEE 802.3]. The 257b blocks of 800GBASE-R clients are descrambled following the procedures in Clause 172.2.5.6 of [IEEE 802.3df]. The 257b blocks of 800G-ETC-R clients are descrambled following the procedures in Clause 3.2.4.2.6 of [ETC-800G]. For 800GBASE-R and 800G-ETC-R clients, the am_sf<2:0> bits extracted from the two 400G PCS streams are processed as described in Section 4.7.3 prior to transmission in the overhead fields of the 800ZR frame. For 100GBASE-R clients the 257b blocks are not descrambled.

The inverse process is implemented in the receive direction. For 200GBASE-R/400GBASE-R clients, the 257b blocks are scrambled following the procedures in Clause 119.2.4.3 of [IEEE 802.3]. 800GBASE-R clients are scrambled following the procedures in Clause 172.2.4.5 of [IEEE 802.3df]. 800G-ETC-R clients

are scrambled following the procedures in Clause 3.2.4.1.4 of [ETC-800G]. For 100GBASE-R clients, no scrambling is performed (the blocks are intrinsically scrambled).

For 200GBASE-R, 400GBASE-R, 800GBASE-R and ETC-800G-R clients, the am_sf<2:0> bits extracted from the received 800ZR frame overhead fields are processed as described in Section 4.7.3 and inserted in the appropriate locations in the AM fields. For 100GBASE-R clients the BIP counters are recomputed and inserted.

4.2 Client Mapping to 100G ZR instances

The 800ZR frame is formed by interleaving eight 100G ZR instances as described in section 4.4.

In the transmit direction, any failed client signal is replaced with a replacement signal as described in section 4.7. Any clients which are not provisioned are indicated through the MSI overhead field of the corresponding 100G ZR instances with the payload undefined in this IA.

For 100GBASE-R clients, the 257b stream is mapped directly into the payload area of the corresponding 100G ZR instance.

For 200GBASE-R clients, the 257b stream is distributed on a 257b block basis, in a round-robin fashion into two 100G 257b streams which are mapped into the payload areas of the corresponding two 100G ZR instances.

For 400GBASE-R clients, the 257b stream is distributed on a 257b block basis, in a round-robin fashion into four 100G 257b streams which are mapped into the payload areas of the corresponding four 100G ZR instances.

For an 800G-ETC-R or 800GBASE-R client, the 257b stream (formed as described in Section 4.1.4) is distributed on a 257b block basis, in a round-robin fashion into eight 100G 257b streams which are mapped into the payload areas of the corresponding eight 100G ZR instances.

4.2.1 Profiles for 800ZR client assignment to 100G ZR Instances

As described above, the 800ZR frame comprises eight 100G ZR instances into which clients can be mapped. An 800ZR implementation shall support the assignment of clients to 100G ZR Instances defined in Table 2. For clients which span multiple 100G ZR instances, the data shall be mapped to contiguous 100G ZR instances with ascending IID order. All 100G ZR instances are phase and frequency aligned. Section 4.3.3.3 describes the IID overhead field in the 100G ZR instance.

Client Type	Mapping to 100G ZR instances
100GBASE-R	Client #1 ⇒ 100G ZR Instance IID=1 Client #2 ⇒ 100G ZR Instance IID=2 Client #3 ⇒ 100G ZR Instance IID=3 Client #4 ⇒ 100G ZR Instance IID=4 Client #5 ⇒ 100G ZR Instance IID=5 Client #6 ⇒ 100G ZR Instance IID=6 Client #7 ⇒ 100G ZR Instance IID=7

	Client #8 ⇒ 100G ZR Instance IID=8
200GBASE-R	Client #1 ⇒ 100G ZR Instances IID=1, 2 Client #2 ⇒ 100G ZR Instances IID=3, 4 Client #3 ⇒ 100G ZR Instances IID=5, 6 Client #4 ⇒ 100G ZR Instances IID=7, 8
400GBASE-R	Client #1 ⇒ 100G ZR Instances IID=1,2,3,4 Client #2 ⇒ 100G ZR Instances IID=5,6,7,8
800G-ETC-R 800GBASE-R	100G ZR Instances IID=1,2,…,8

Table 2: Recommended profiles for client mapping to 100G ZR instances

4.3 Frame structure of 100G ZR Instances

The 100G ZR frame is a profile of the FlexO frame structure described in clause 8 of [ITU-T G.709.1] and is shown in Figure 5. The 100G ZR frame may be viewed as consisting of 128 rows of 5140 bits each (aligning with Figure 8-1 of [ITU-T G.709.1]). The rows are numbered 1-128, and the bit columns are numbered 1-5140.

Each 100G frame consists of an overhead area in bit columns 1-1280 in row 1, and a payload area that occupies the rest of the frame. During normal operation, the payload area is used to map Ethernet clients and consists of a 5-bit PAD field in row 1 bit columns 1281-1285 that is transmitted as all zeros and ignored at the receiver, followed by 656,635 bits (2555 × 257 bits) for client data mapping. The 100G client stream consisting of 257b blocks is mapped into this payload area using GMP as described in Section 4.5.

The frame overhead consists of three fields:

- Bits 1-480 are used as an alignment mechanism (AM).
- Bits 481-960 are unused in 800ZR (this field is designated EOH and used for other purposes in [ITU-T G.709.1].
- Bits 961-1280 (BOH) are used to carry overhead information. For 800ZR, the overhead is a subset of what is defined in [ITU-T G.709.1], as described in the remainder of this section.

All relevant overhead in this IA is transmitted in the transmission order specified in [ITU-T G.709.1].

The 100G ZR frame structure may equivalently be viewed as 5140 blocks of 128 bits each as shown in the lower part of Figure 5 (aligning with Figure 8-1 of [ITU-T G.709.6]). This representation is convenient when discussing the interleaving of the 100G frame structure into the 800ZR frame.

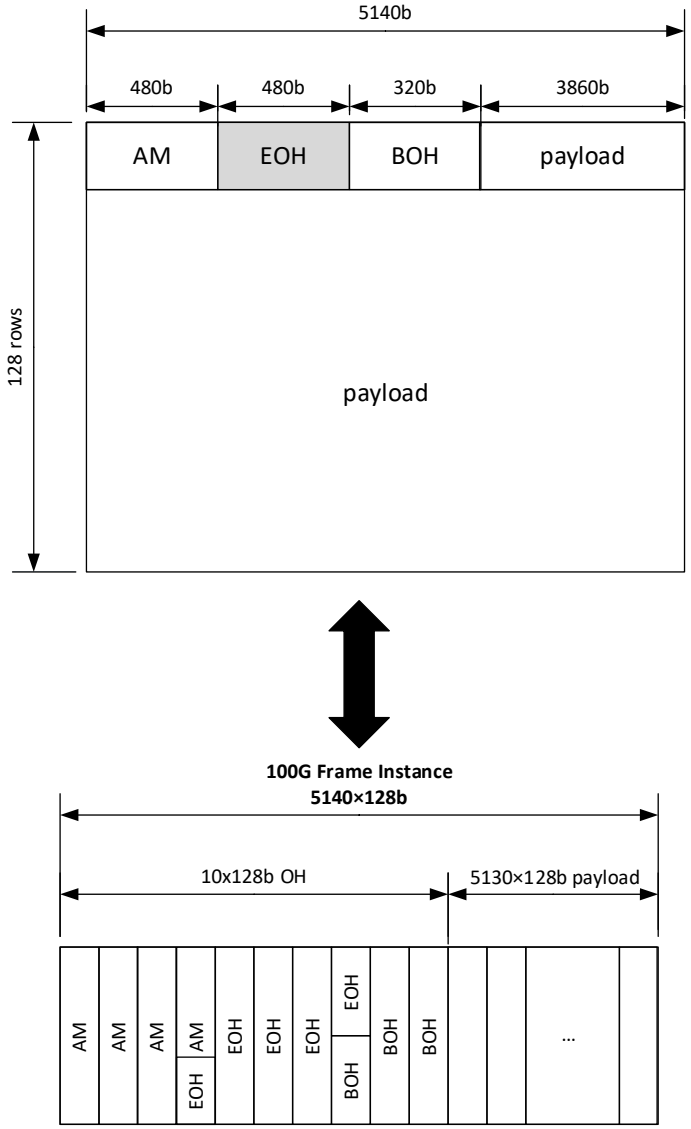


Figure 5: 100G Frame Structure

4.3.1 Alignment Mechanism (AM)

The role of the AM is to find the 800ZR frame boundary. The AM field is in bit columns 1-480 of the first row of each frame and is protected by the line-side FEC. The contents of the AM field are specified in clause 9.1 of [ITU-T G.709.6].

Figure 6 illustrates the structure of the AM field within each 100G ZR instance. The 480-bit field consists of 16 octets of FA1 in bit columns 1 to 128, 16 octets of FA2 in bit columns 129 to 256, and 224 reserved bits in bit columns 257 to 480. The FA1 and FA2 octets are shown in Table 3. Each of the FA1 and FA2 octets shall be transmitted MSB first. The reserved bits are transmitted as all-zeros and ignored upon receipt. Frame alignment can be done across a subset of the AM field.

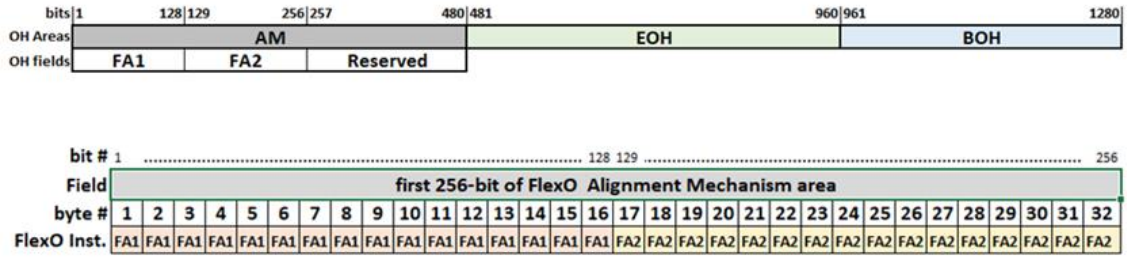


Figure 6: Alignment Mechanism definition

AM Field	Value
FA1	0x09
FA2	0xD7

Table 3: Alignment Mechanism Encodings

4.3.2 Extended Overhead (EOH) Field

The 480-bit extended overhead (EOH) field is in bit columns 481 to 960 within each 100G ZR instance. This field is not used in this implementation agreement (but used in related standards for other overhead purposes; see Clause 9.3 of [ITU-T G.709.1]). It is transmitted as all-zeros and ignored at the receiver.

4.3.3 Basic Overhead (BOH) Field

The 320-bit basic overhead (BOH) field is in bit columns 961 to 1280 within each 100G ZR instance. This overhead field is used to carry information needed to identify and de-map the client information carried within the payload area of the frame as described in clause 9.2 of [ITU-T G.709.1].

For convenience, Figure 7 shows details of the BOH. The BOH is distributed over an eight-frame multiframe. The BOH fields defined in [ITU-T G.709.1] and used in this IA are indicated with white background; those defined in [ITU-T G.709.1] but not used in the IA are shown with yellow background. Overhead fields that are not used are set to zero and ignored at the receiver.

The acceptance of overhead fields at the receiver is described in [ITU-T G.798].

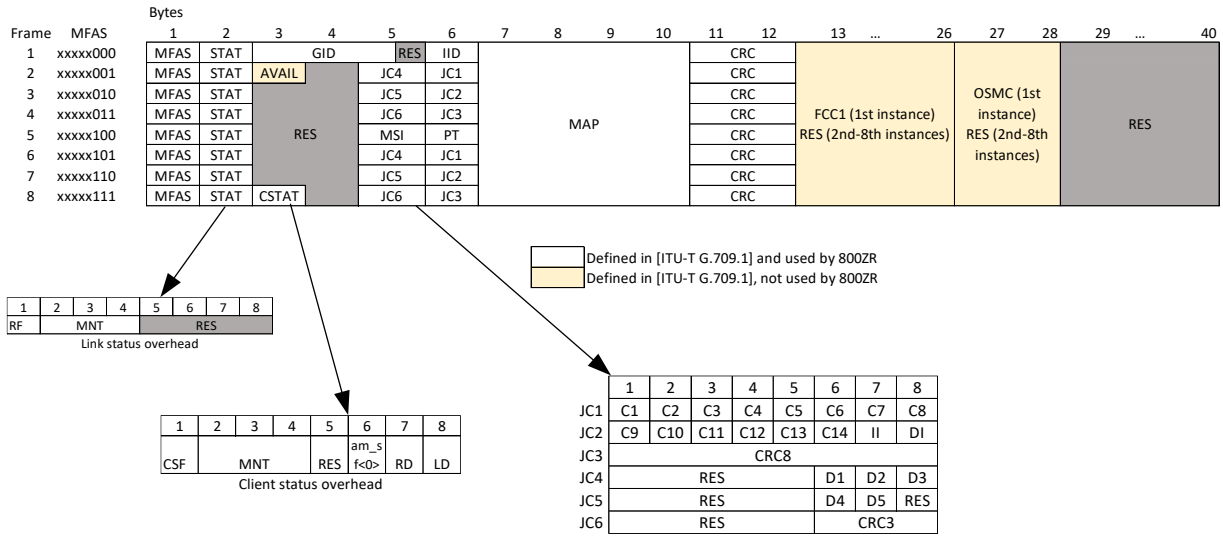


Figure 7: Overhead (BOH) field within each 100G ZR instance

4.3.3.1 Multi-frame Alignment Signal (MFAS)

The multi-frame alignment signal (MFAS) is the first byte in the 320-bit BOH of each 100G ZR instance. It is present and incremented in every 100G ZR instance. It counts from 0x00 to 0xFF and provides a multi-frame sequence following [ITU-T G.709.1] Clause 9.2.1 definition. Note that the MFAS sequences of the eight 100G ZR instances must be synchronized.

4.3.3.2 Group ID (GID)

The GID field is a 20-bit value that is present in each 100G ZR instance and located in BOH bytes 3, 4 and 5 of frame 1. It is defined in Clause 9.2.2.1 of [ITU-T G.709.1] for the purpose of managing groups that consist of multiple PHYs. For 800ZR, there is only one PHY in the group, and the GID is set to 0x1 and replicated in each 100G ZR instance. The bits of the GID shall be transmitted MSB first.

4.3.3.3 IID

The IID is an eight-bit value that is present in BOH byte 6 of frame 1 in each 100G ZR instance. It is defined in Clause 9.2.2.2 of [ITU-T G.709.1] for the purpose of identifying the individual instances that compose a FlexO group. For an 800ZR interface, there are always eight 100G ZR instances, so the IID values are fixed to 1-8 with value 0x01 carried by the first instance and value 0x08 carried by the eighth instance. The IID is transmitted MSB first.

4.3.3.4 MAP

The 256-bit MAP field is spread across BOH bytes 7-10 of all eight frames of the multiframe and is replicated in every 100G ZR instance. This field is defined in Clause 9.2.2.3 of [ITU-T G.709.1] and for convenience is shown in Figure 8. This field is used to indicate the IID of each member in a group. For 800ZR, the IIDs are fixed to 1-8, and as such, this field is set to the value of MAP[1:8]=0xFF and 0 in all other bits.

	MAP bytes																													
Frame	7								8								9								10							
1	RES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
2	32	33																														
3																																
4																																
5																																
6																																
7																																
8																						245	246	247	248	249	250	251	252	253	254	RES

Figure 8: 800ZR MAP Field

4.3.3.5 CRC

The CRC field is calculated and populated as specified in Clause 9.2.5 of [ITU-T G.709.1]. It is a CRC-16 carried in BOH bytes 11 and 12 and protects the integrity of the BOH fields in bytes 2 to 10 (and therefore excludes the MFAS, OSMC and FCC1 fields) of the same 100G ZR instance.

4.3.3.6 Payload Type (PT)

The payload type (PT) field shall be transmitted in BOH byte 6 of frame 5 of the multi-frame and replicated in every 100G ZR instance. The payload type values are assigned by ITU-T SG15 and documented in clause 9.2.4.2 of [ITU-T G.709.1]. Table 4: lists the values of this field used by 800ZR. The PT field is transmitted in order with the most-significant bit transmitted first.

Payload Type	Value
Multiplex of 100GBASE-R, 200GBASE-R, 400GBASE-R, 800GBASE-R, or 800G-ETC-R clients	0x40
Framed PRBS Test Signal (Ref. 4.9**)**	0x41

Table 4: Payload Type (PT) Values

4.3.3.7 Client Multiplex Structure (MSI)

The client multiplex structure identifier (MSI) field is specified in clause 10.2.3.1 of [ITU-T G.709.1] as FlexO MSI and is transmitted in frame 5 byte 5 of the BOH. It indicates the assignment of the 100G ZR instances to clients. The MSI consists of a one-bit OCC field to indicate whether a client occupies the instance and a 7-bit tributary port identifier (TPID) field to indicate the client ID when the instance is occupied. All instances occupied by the same client use the same TPID value. Table 5 indicates the TPID values used for the each of the client profiles defined in Table 2, where client #1 is assigned a TPID value of 0 and so on.

Client Types	TPID values
8x100GE	0,1,2,3,4,5,6,7
4x200GE	0,0,2,2,4,4,6,6
2x400GE	0,0,0,0,4,4,4,4

1x800GE	0,0,0,0,0,0,0,0

Table 5: Client multiplex profiles for 800ZR

4.3.3.8 Link status monitoring and signaling (STAT)

The link status (STAT) overhead byte is specified in clause 9.2.3 of [ITU-T G.709.1] to be present in BOH byte 2 of every 100G ZR instance. This byte is used to signal the status of the entire 800ZR link and replicated in every 100G ZR instance. It includes the 1-bit RF, 3-bit MNT field.

The 3-bit MNT signal is used to indicate whether normal traffic or a maintenance signal is present on the link. It shall be set to 0x0 during normal operation. The MNT field is transmitted MSB first and shall take values from Table 6.

Signal	Value	Condition
MNT-LCK	0x5	Maintenance lock
Normal signal	0x0	Normal operation

Table 6: MNT field values

4.3.3.9 Client status monitoring and signaling (CSTAT)

The client status (CSTAT) overhead bytes are specified in clause 10.2.3.3 of [ITU-T G.709.1] to be present in frame 8 BOH byte 3 of every 100G ZR instance. This field is used to signal the status of each client carried in the 800ZR link. The CSTAT byte for each client is replicated in every 100G ZR instance used to carry the client. The overhead includes the 1-bit CSF, 3-bit MNT field, am_sf<0>, LD and RD bits.

The 3-bit MNT signal is used as a maintenance signal indication on a per-client basis. It shall be set to 0x0 during normal operation. The MNT field is transmitted MSB first and shall take values from Table 7.

Signal	Value	Condition
MNT-LCK	0x5	Maintenance lock
Normal	0x0	Normal Operation

Table 7: CSTAT MNT field values

4.4 Generation of 800ZR frame from 100G ZR instances

The 800ZR frame is generated through 128 bit interleaving of eight 100G ZR instances as described in clause 8.4 of [ITU-T G.709.1] and clause 8 of [ITU-T G.709.6]. Figure 9 illustrates this process. In accordance with [ITU-T G.709.1], the eight 100G ZR instances are frame and multi-frame aligned prior to the interleaving to form the 800ZR frame. The resulting 800ZR frame consists of 41,120 blocks of 128 bits. The 800ZR frame overhead occupies the first 80 blocks of the frame followed by 41,040 blocks of payload information.

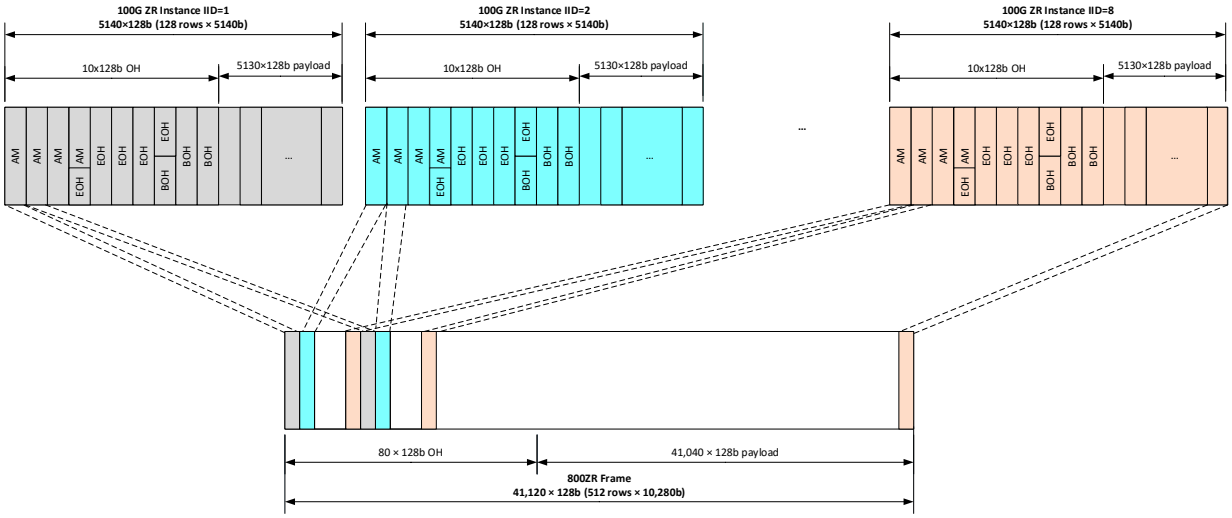


Figure 9: Generation of 800ZR frame from eight 100G ZR instances

4.5 GMP Mapping Processes

The mapping of clients into the 100G ZR instances is as specified in clause 10.2 and Annex B of [ITU-T G.709.1].

4.5.1 Multi-frame structure

For each of the eight 100G ZR instances used for client mapping, a multi-frame structure is used for GMP mapping. The multi-frame structure of all eight 100G ZR instances is synchronized at source. The multi-frame structure is shown in Figure 10 and is constructed from 4 consecutive frames identified by the MFAS bits 7 and 8 set to 00, 01, 10 and 11 respectively. A 5-bit PAD is inserted in each 100G ZR instance to maintain 257b alignment of the client signal. This PAD area is transmitted as zeros and ignored at the receiver. The entire remaining payload area of these four frames consisting of 10,220 × 257 bits is used to map a 100G 257b stream of a y00GE client (y=1,2,4 or 8). The nominal C_m for the mapping is 10,216. Each individual multi-frame may carry more or fewer than 10,216 client payload blocks. The unused blocks contain a fixed stuff and shall be transmitted as zeros and ignored at the receiver. The entire GMP process for client mapping is detailed in Section 4.5.2.

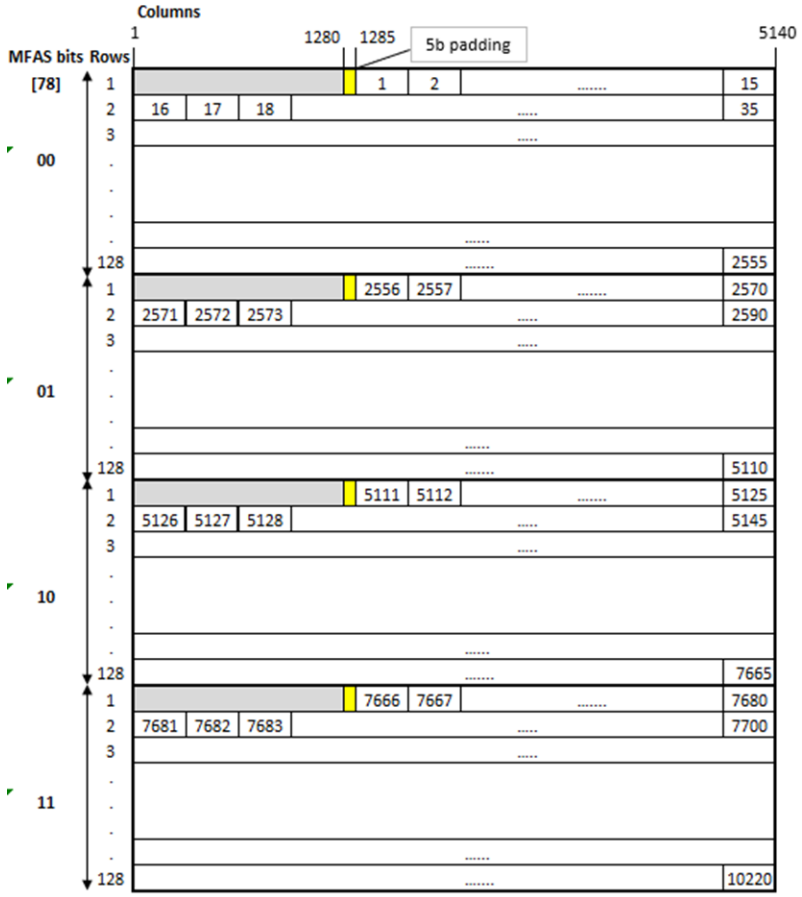


Figure 10: 100G ZR instance payload 4-frame multi-frame structure with 257-bit payload blocks for the GMP mapping of Ethernet client signal

4.5.2 Details of client mapping process

The y 100G streams of a y00GE client (y=1,2,4 or 8) as described in Section 4.2 are asynchronously mapped into y phase-locked 100G multi-frame structure instances described in 4.3 and 4.5.1. The 800ZR timing is de-correlated from the client clock to support multiplexing of asynchronous client signals with independent clock frequency offsets. The client signals are treated as a stream of 257b blocks. Data and timing transparency shall be supported using information fields which are inserted by the GMP process for use upon de-mapping.

The GMP Justification Control bytes (JC1-6) are carried in the client-specific overhead (bytes 5 and 6) of rows 2, 3 and 4 of the four-frame multi-frame used for GMP mapping (i.e., they appear in rows 2, 3, 4, 6, 7 and 8 of the eight-frame multi-frame shown in Figure 7). For clients that are carried by multiple 100G ZR instances, the mappers are locked and the JC bytes in all the associated 100G ZR instances have the same values. The 100G ZR instance used by the receiver for the extraction of JC bytes is vendor specific. The JC bytes signal the GMP parameters C_m and ΣC_{nD} from the mapper to the de-mapper. The parameters received in the overhead of multi-frame n indicate the data and stuff locations in multi-frame n+1.

Refer to [ITU-T G.709] Annex D and [ITU-T G.709.1] Annex B for the general principles of the Generic Mapping Procedure (GMP). For each y00GE client (with y=1,2,4 or 8):

- m = GMP data/stuff granularity = y × 257 bit.
- n = y × 257/32 = y × 8.03125-bit unit and represents the timing granularity of the GMP mapping present in C_m and ΣC_{nD} parameters.
- $P_{m,server}$ = maximum number of m-bit data entities in 4-frame multi-frame server payload of the y multi-frame aligned 100G ZR instances = 10,220
- C_m = number of client m-bit data entities in 4-frame multi-frame server payload of the y multi-frame aligned 100G ZR instances. It is encoded with 14 bits and carried in JC1 and JC2 control OH bytes.
- C_n = number of equivalent client n-bit data entities in 4-frame multi-frame server payload of the y multi-frame aligned 100G ZR instances. This value provides additional 'n'-bit timing information.
- ΣC_{nD} – accumulated value of the remainder of C_n and C_m. It is encoded with 5-bits and carried in JC4 and JC5 control OH bytes.
- C_n and C_m being integer values, then: $C_n(t) = 32\times C_m(t) + (\Sigma C_{nD}(t) - \Sigma C_{nD}(t-1))$

The support for n-bit timing information (ΣC_{nD}) in the JC4/JC5/JC6 OH is required.

Each of the y 100G streams of a client signal is mapped to the y 100G ZR instances payload structure as y 100G aligned 257b block streams. The payload area for this mapping consists of the payload of the aligned 4-frame multi-frames of y 100G ZR instances in ascending IID order. Groups of y × 257 consecutive bits of the client are mapped into y groups of 257b blocks of the y aligned 4-frame 100G multi-frame payload areas under control of the GMP data/stuff control mechanism. Each group of y × 257b in the y aligned 4-frame 100G multi-frame payload areas may carry either y × 257b client bits or y × 257b stuff bits. The stuff bits shall be transmitted as zeros and shall be ignored on receive.

The server input bit rate for 100GE clients after RS(544,514) FEC decode and AM removal is 100,384,497 kbits/s. The server input nominal bit rate for 200GE/400GE/800GE clients after RS(544,514) FEC decode and AM removal and distribution to 100G streams is 100,385,723 kbit/s.

The de-mapping process uses [ITU-T G.709.1] Annex B.1.2 to decode $C_m(t)$ from JC1/JC2/JC3. Errors in CRC8, computed over JC1/JC2, are handled using the method described in Annex B.1.2.1 of [ITU-T G.709.1]. The de-mapping process leverages [ITU-T G.709.1] Annex B.1.3 to decode $C_{nD}(t)$ from JC4/JC5/JC6. Errors in the JC4/JC5/JC6 fields are detected as described in Section 4.5.6. If the CRC3 protecting JC4/JC5/JC6 is good, the new values of JC4/JC5 are accepted, otherwise, these values are discarded, and previous values are retained.

The derivation of the GMP parameters is shown in Table 31 in Annex A.

4.5.3 Stuffing Locations

GMP is a positional mapping with non-fixed stuff locations. The stuff locations within the payload are determined using a delta-sigma algorithm based on the Cm(t) value. In the case of 800ZR, the payload into which GMP maps client data is the four-frame multi-frame described in Section 4.5.1.

Table 8 shows the location of the "stuff" GMP blocks for a few specific C_m values.

C_m	GMP stuff locations
10220	N/A
10219	1
10218	1, 5111
10217	1, 3407, 6814
10216	1, 2556, 5111, 7666
10215	1, 2045, 4089, 6133, 8177
10214	1, 1704, 3407, 5111, 6814, 8517

Table 8: GMP stuff locations

4.5.4 GMP overhead Encoding

GMP overhead (JC Bytes OH) is carried once per 4-frame multi-frame. GMP overhead shall carry the encoded 14-bit $C_m(t)$ (i.e., m-bit block count value) in C1-14 bits of JC1 & JC2 (C1=MSB, …, C14=LSB) and the encoded 5-bit $\Sigma C_{nD}(t)$ (cumulative value of $C_{nD}(t)$) in D1-D5 bits of JC4 and JC5 (D1=MSB, … D5=LSB) GMP parameters.

$C_m(t)$ shall be protected with a CRC8 (carried in JC3 OH byte) and $\Sigma C_{nD}(t)$ shall be protected with a CRC3 (carried in the three LSBs of JC6 OH byte).

The JC3 OH CRC8 and JC6 OH CRC3 calculations shall follow the descriptions in ITU-T G.709.1 Annex B.1.2.1 and B.1.3.1 respectively.

4.5.5 GMP OH – CRC8 calculation

The CRC8 located in JC3 is calculated over the JC1 and JC2 bits. The CRC8 uses the generator polynomial:

$$g(x) = x^8 + x^3 + x^2 + 1$$

- The JC1 and JC2 octets are taken in order, most significant bit first, to form a 16-bit pattern representing the coefficient of a polynomial M(x) of degree 15.
- M(x) is multiplied by x^8 and divided (modulo 2) by G(x), producing a remainder R(x) of degree 7 or less.
- The coefficients of R(x) are considered an 8-bit sequence, where x^7 is the most significant bit.
- The 8-bit sequence is the CRC8 where the MSB of the CRC8 is the coefficient of x^7 and the LSB is the coefficient of x^0.

The de-mapper process performs the same steps as the mapper, except, the M(x) polynomial of the first step includes the CRC bits of JC3, resulting in M(x) having degree 23. In the absence of bit errors, the remainder shall be 0000 0000.

4.5.6 GMP OH – CRC3 calculation

The CRC3 located in JC6 uses the generator polynomial:

$$g(x) = x^3 + x^2 + 1$$

- The three least significant bits of the JC4 and JC5 octets (JC4 D1-3 and JC5 D4-5 + RES) are taken in order, most significant bit first, to form a 6-bit pattern representing the coefficients of a polynomial M(x) of degree 5.
- M(x) is multiplied by x^3 and divided (modulo 2) by G(x), producing a remainder R(x) of degree 2 or less.
- The coefficients of R(x) are considered a 3-bit sequence, where x^2 is the most significant bit.
- This 3-bit sequence is the CRC3 where the MSB of the CRC3 is the coefficient of x^2 and the LSB is the coefficient of x^0.

The de-mapper process performs the same steps as the mapper, except, the M(x) polynomial in the first step includes the CRC bits of JC6, resulting in M(x) having degree 8. In the absence of bit errors, the remainder shall be 000.

4.6 Datapath Management and Control

The Datapath (see Figure 4) of an 800ZR application consists of all involved resources between and including the Host Interface and Media Interface. An 800ZR application can be delineated by System-side resources (Host Path) and Line-side resources (Network Path), each having their own clock domain and management state machines for managing the turn-up of the Host Path and the Network Path (ref. [CMIS] Section 7.6.2).

4.6.1 Datapath Initialization Behaviour

In the event either the Host Path or the Network Path is not initialized, an 800ZR application can (optionally) source Idle Control characters as 256b/257b encoded blocks. The Host Path state is per client port #p as identified by the TPID.

Host Path State	Network Path State	Host Output	Media Output
Uninitialized	Uninitialized	Electrical Squelch[1]	TX Disabled
Uninitialized	Initialized	Electrical Squelch	MSI(p) OCC bit set to 0
Initialized	Uninitialized	GMP PCS generator inserts a continuous stream of Idle control characters /I/ per [IEEE 802.3] Clause 91, 119.2.3.5 or [IEEE P802.3df] Clause 172 to the Rx datapath.	TX Disabled
Initialized	Initialized	Normal Data[2]	Normal Data

[1] See definition of squelch in [CMIS]

[2]If the expected payload type (PT) is 0x41 on the RX media to host direction, the host output is provisioned to either local fault (LF) or electrical squelch

Table 9: Datapath Initialization

4.7 Error Detection and Signaling

For errors detected at the Media Interface, Table 10 specifies the data path layer replacement signal sent toward the host interface on all affected clients. If a detected defect is associated with client port #p, then only that client output signal is replaced. If a detected defect relates to the overall media interface receiver, then all client port output signals are replaced. The recommended behavior of defect assertion and de-assertion is described in Section 4.7.1. The insertion of payload local fault (LF) insertion or electrical squelch is subject to a user configurable "Consequent Action Hold-off" timer. The recommended behavior of LF insertion or electrical squelch is described in Section 4.7.2.

Defects	Description	Consequent Action
LOL	Loss of DSP frame lock	Payload local fault (LF) insertion or electrical squelch on client port #p (if defect is per port) or all client ports (if defect is for entire media interface).
LOF/LOM	Loss of ZR frame/multi-frame	
LOS (Optional)	Loss of optical signal	
FED (Optional)[1]	FEC Excess Degrade Detected (ref. 4.7.4)	
CSF (CSTAT[1]) (p) (optional)	Client Signal Fail on client port #p	
MSIM(p)	MSI mismatch on client port #p	
PLM	Payload Type Mismatch	
STAT MNT LCK	STAT Maintenance Lock	
CSTAT MNT LCK (p)	CSTAT Maintenance Lock on client port #p	
Uncorrectable FEC errors	CRC32 detected as errored (Ref. 5.3)	Error Mark /E/ [2] all client port outputs
	FEC Uncorrected Block Error	
LD (CSTAT[8]) or FDD	LD detected, or FEC Degrade Detected (ref 4.7.4)	Signal LD (AM_sf<1>) on client port #p if LD(p) or on all client ports if FDD.

Table 10: Error detection and consequent actions – Media to Host

[1] FED defect and consequent action are optional. Depending on the FED threshold setting, FED can be used as a pre-emptive condition to force a protection switch to avoid excess data loss due to a degraded link.

[2] Error marking is optional.

For errors detected on the host interfaces (per Ethernet client port #p), Table 11 specifies the data path layer replacement signal sent toward the media interface for the Ethernet clients with detected defects. Note that other Ethernet clients without any detected defect remain unaffected.

Defects on client port #p	Description	Consequent Action on client port #p
LOL/LOA(p)	Loss of FEC Alignment Lock on client port #p (ref. 802.3 Clause 91, 119 or 172)	Payload local fault (LF) insertion using a clock that is within the client bit rate tolerance. Signal CSF in 800ZR CSTAT[1] (p).
LOS(p)	Loss of signal on client port #p	
FED(p) (Optional)[1]	FEC Excess Degrade Detected on client port #p (ref. 4.7.4)	
FEC UCB(p)	FEC Uncorrected Block Error on client port #p	Error Mark \E
LD(p) (AM_sf<1>) or FDD(p)	FEC Degrade Detected on client port #p (ref [IEEE 802.3] Clause 119.2.5.3)	Signal LD in 800ZR CSTAT[8] (p). Corresponds with FEC_degraded_SER alarm in [IEEE 802.3] Clause 119.2.5.3 for 200GBASE-R and 400GBASE-R clients and [IEEE 802.3df] Clause 172.2.5.3 for 800GBASE-R

Table 11: Error detection and consequent actions – Host to Media

[1] FED defect and consequent action are optional. Depending on the FED threshold setting, FED can be used as a pre-emptive condition to force a protection switch to avoid excess data loss due to a degraded link.

The replacement signals used for each client are specified in Table 12 and shall be generated with a clock that is within the client bit rate tolerance. The replacement signals for the y00GBASE-R clients are aligned with Table 10-3 of [ITU-T G.709.1].

Client Type	Replacement Signal
100GBASE-R	Scrambled Local Fault sequence ordered-sets encoded as per [IEEE 802.3] Clause 82
200GBASE-R and 400GBASE-R	Local Fault sequence ordered-sets encoded as per [IEEE 802.3] Clause 119
800GBASE-R	Local Fault sequence ordered-sets encoded as per [IEEE 802.3df] Clause 172
800G-ETC-R	Local Fault sequence ordered sets per [ETC-800G]

Table 12: Client replacement signal

4.7.1 Media-side defects

The detection of media-side defects is described in the following sections for each defect. This behavior is intended to be consistent with the description of similar defects in [ITU-T G.798].

4.7.1.1 Loss of frame/multi-frame (dLOF/dLOM)

The loss of frame defect (dLOF) is based on searching for the AM pattern contained in the 800ZR frame structure. The process has two states, out-of-frame (OOF) and in-frame (IF). dLOF is generated when OOF state persists for 3ms. After assertion, the defect is cleared only if in the IF state continuously for 3ms.

The multiframe alignment defect is based on searching for the MFAS byte incremental sequence contained in the BOH of each 100G ZR instance (see Section 4.3.3.1). The process has two states, out-of-multiframe (OOM) and in-multiframe (IM). dLOM is generated when OOM state persists for 3ms for any 100G ZR instance. After assertion, the defect is cleared immediately if in the IM state in all the 100G ZR instances.

LOF and LOM processes are based on [ITU G.798] clauses 8.2.2, 8.2.8, 6.2.5.1 and 6.2.5.2.

4.7.1.2 Loss of Lock Defect (dLOL)

The loss of lock defect (dLOL) is based on pilot symbols and DSP frame/super-frame alignment. The 800ZR DSP frame structure and pilot sequence are described in Sections 5.10-5.10.5. The dLOL is asserted if DSP frame alignment is lost for 3ms. After assertion, the defect is cleared only if the DSP frame alignment is detected continuously for 3ms.

4.7.1.3 Loss of Signal Defect (dLOS)

The loss of signal defect (dLOS) implementation is vendor-specific and defined by ***RxLOSType***.

4.7.1.4 Client Signal Fail Defect (dCSF)

The client signal fail defect (dCSF) is based on the CSF overhead bit described in Section 4.3.3.9 and detected in the media to host direction separately for each demapped client. A new CSF value is accepted for a 100G ZR instance if a new value of the CSF field is received in this 100G ZR instance overhead frame with good CRC (section 4.3.3.5). The dCSF defect of a client is asserted if the accepted CSF value of any of the 100G ZR instances carrying this client is “1”. The defect is cleared if the accepted CSF value of all 100G ZR instances carrying this client is “0”.

4.7.1.5 MSI Mismatch (dMSIM)

The MSI mismatch defect (dMSIM) is based on the client multiplex structure derived from the MSI overhead described in Section 4.3.3.7 and detected in the media to host direction separately for each demapped client.

The multiplex structure identifier (MSI) of the eight 100G ZR instances consists of one byte per 100G ZR instance. A new multiplex structure identifier for a 100G ZR instance is accepted if a new value of the MSI byte field of that 100G ZR instance is received in an overhead frame with good CRC (section 4.3.3.5).

The configuration of the expected MSI is implementation dependent.

The dMSIM defect for each client port is based on comparing the accepted and expected MSI for all 100G ZR instances. The detection process described below reuses dMSIM specifications in clause 6.2.9.2 of [ITU-T G.798].

For each case where the accepted and expected MSI values pertaining to the same 100G ZR instance are not equal,

a) if the 100G ZR instance is occupied in the expected MSI value, the dMSIM defect of the Ethernet client indicated by the TPID subfield of the expected MSI value is set,

b) if the 100G ZR instance is occupied in the accepted MSI value, the dMSIM defect of the Ethernet client indicated by the TPID subfield of the accepted MSI value is set.

dMSIM shall be detected within 100 ms of changes to accepted or expected MSI.

4.7.1.6 Payload Mismatch Defect (dPLM)

The payload mismatch detect (dPLM) is based on the payload type overhead field described in Section 4.3.3.6 and detected for the entire 800ZR payload structure.

A new payload type is accepted per 100G ZR instance if a new value of the PT byte field of the 100G ZR instance overhead is received in an overhead frame with good CRC (section 4.3.3.5). dPLM shall be declared if the accepted payload type is not equal to the expected payload type(s), in any of the eight 100G ZR instances. dPLM shall be cleared if the accepted payload type is equal to the expected payload type(s) in all of the eight 100G ZR instances.

Note that the above detection process is based on dPLM specifications in clause 6.2.4.2 of [ITU-T G.798].

4.7.1.7 STAT Maintenance Lock Defect (dLCK)

The maintenance lock defect (dLCK) is based on the STAT MNT overhead bits described in Section 4.3.3.8 and detected in the media to host direction. A new MNT value is accepted for a 100G ZR instance if a new value of the MNT field is received in this 100G ZR instance overhead frame with good CRC (section 4.3.3.5). The dLCK defect is asserted if the accepted MNT value of any of 100G ZR instances is “101”. The defect is cleared if the MNT value of all 100G ZR instances are not equal to “101”.

4.7.1.8 Client CSTAT Maintenance Lock Defect

The client maintenance lock defect is based on the CSTAT MNT overhead bits described in Section 4.3.3.9 and detected in the media to host direction separately for each demapped client. A new CSTAT MNT value is accepted for a 100G ZR instance if a new value of the CSTAT MNT field is received in this 100G ZR instance overhead frame with good CRC (section 4.3.3.5). The defect of a client is asserted if the accepted CSTAT MNT value of any of the 100G ZR instances carrying this client is “101”. The defect is cleared if the accepted CSTAT MNT value of all 100G ZR instances carrying this client are not equal to “101”.

4.7.1.9 Remote Fault Defect (dRF)

The remote fault (dRF) is based on the STAT RF overhead bit described in Section 4.3.3.8 and detected in the media to host direction. A new RF value is accepted for a 100G ZR instance if a new value of the RF field is received in this 100G ZR instance overhead frame with good CRC (section 4.3.3.5). The dRF defect is asserted if the accepted RF value of any of 100G ZR instances is “1”. The defect is cleared if the RF value of all 100G ZR instances are “0”.

4.7.2 Consequent Action Hold-off

The insertion of the local fault replacement signal or electrical squelch in response to the media-side defects described in Table 10 is subject to a hold-off timer illustrated in Figure 11. A C-CMIS user-configurable parameter “consequent action holdoff time” is used to delay the insertion of local faults (LF) or electrical squelch after detection of a defect. Idles are inserted during this interval with the clock

frequency held at the last estimate prior to defect detection. If the failure and the corresponding defect is cleared prior to the expiry of the hold-off timer, the traffic and clock resume as normal without ever inserting local faults or electrical squelch. If the failure persists longer than the hold-off timer, local faults are inserted and sent to the host with a nominal clock within the client bit rate tolerance derived from a local timing reference or electrical squelch is asserted. Normal traffic and clock timing are restored only after the failure and the corresponding defect are cleared. Since the usage of hold-off is user dependent, a "holdoff time" value of zero should be supported by implementations to ensure the default behavior of no holdoff.

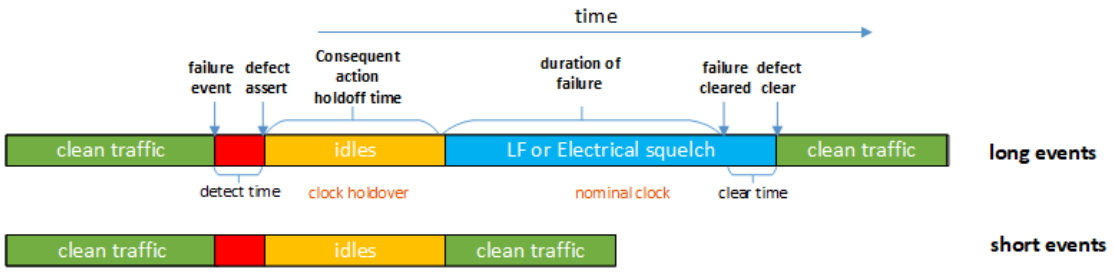


Figure 11: Local Fault (LF) or electrical squelch hold-off timing

4.7.3 Link Degrade Indication

800ZR modules support the detection and signaling of end-to-end link degradations for use by switch/routers with soft reroute capabilities as described in Annex C.1 of [ITU-T G.709.1]. Figure 12 illustrates the bidirectional signaling between an 800ZR module and two Routers (A and B). Pre-FEC BER monitors are used to detect and insert link degrade at both the 800ZR optical link and the Ethernet client PCS interfaces.

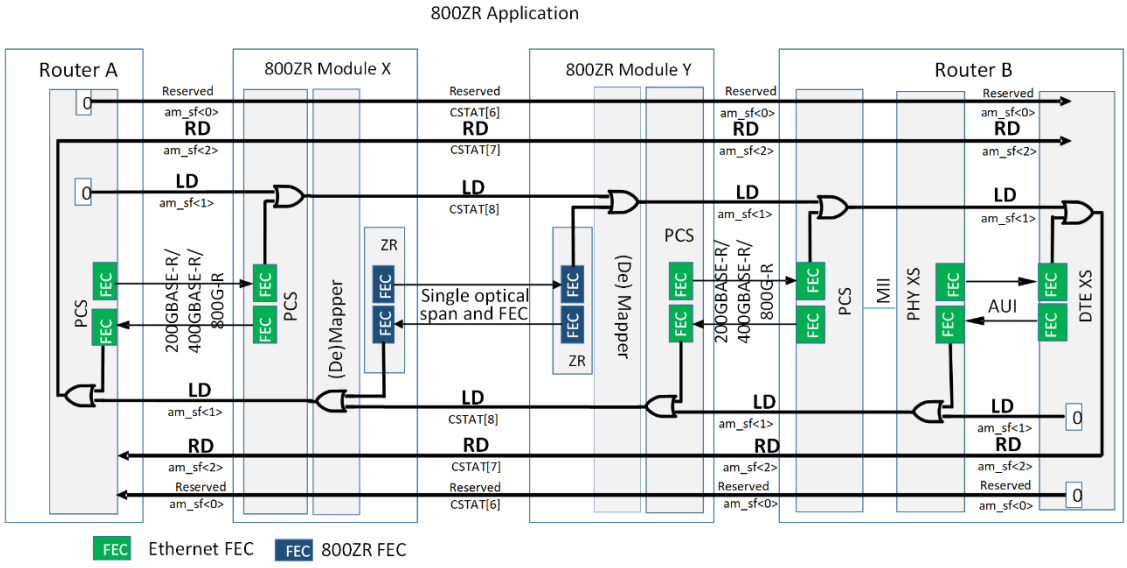


Figure 12: Local/Remote Degrade interworking between Switch/Router and 800ZR transceiver

For 200GE and 400GE client types, [IEEE 802.3] defines three bits in the AM field (am_sf<2:0>) to carry Link Degrade Indication (LDI). Bit am_sf<2> is defined as a Remote Degrade (RD) signal, bit am_sf<1> is defined as a Local Degrade (LD) signal and bit am_sf<0> is reserved. For 800GE clients, the am_sf<2:0> bits are present in both 400G streams at the client interface. For these clients, the bits present in the two 400G streams shall be ORed to form a single set of am_sf<2:0> for further processing and insertion into the CSTAT overhead field in the host-to-media direction. Conversely, in the media-to-host direction, the single set of am_sf<2:0> bits extracted and processed from the 800ZR interface shall be replicated in both streams for 800GE clients.

The 800ZR transceivers X and Y forward the information in the Reserved (am_sf<0>) and RD (am_sf<2>) bits between the transceivers as illustrated in Figure 12. The information in am_sf<0> shall be carried in CSTAT overhead bit 6 of the 100ZR frame instances carrying the client. The status information in am_sf<2> shall be carried in CSTAT overhead bit 7 of the 100ZR frame instances carrying the client. The status information in the LD (am_sf<1>) bit shall be carried after some additional processing in the CSTAT overhead, bit 8 of the 100ZR frame instances carrying the client.

In the host-to-media datapath, the additional processing consists of ORing the ingress LD status in the am_sf<1> bit of the 200GBASE-R/400GBASE-R or 800GE client signals with the local host interface RS(544,514) FEC degrade status and signaling LD in CSTAT<8> to the media interface. In the media-to-host datapath, the CSTAT<8> bit from the media interface is ORed with the 800ZR FEC degrade status and signaled on the am_sf<1> bit to the local host.

4.7.4 Link Degrade Warning and Alarming

FEC Detected Degrade (FDD) and FEC Excessive Degrade (FED) is an optional link monitoring feature, indicating a link degrade condition to the local host and remote transmitter. It can be used, for example, to pre-emptively move traffic away from a degraded link (e.g., traffic re-route). This feature requires capturing the pre-FEC BER from the FEC decoder block over a Performance Monitor (PM) interval. Statistics are gathered by HW and reported by SW. FED and FDD are determined by comparing the HW BER reported statistics against [user configurable] thresholds.

Link Degrade (LD) signaling shall be based on the FEC decoder statistics (number of corrected errored bits, and uncorrectable blocks). Fault detection calculation and threshold settings may be implementation dependent (e.g., based on FEC decoder pre-FEC BER detection capabilities).

The following Performance Monitoring (PM) parameters are defined for determining a Link Degrade (LD) condition over a PM interval. The PM interval and the collection of the statistics to determine LD is defined by the Management Interface Spec specific to the module which this IA is implemented.

FEC decoder bit counters:

- *pFECbitcount* = total number of bits counted over PM interval (64-bit value)
- *pFECcorrbit* = total number of FEC corrected bits over PM interval (64-bit value).

Pre-FEC BER bit counters:

- *pFECBER* = FEC BER over PM interval = (*pFECcorrbit* / *pFECbitcount*)

Pre-FEC threshold settings:

- *FEC_excessive_BER_activate_threshold* (programmable)

- *FEC_excessive_BER_deactivate_threshold* (programmable)
- *FEC_degraded_BER_activate_threshold* (programmable)
- *FEC_degraded_BER_deactivate threshold* (programmable)

FEC degrade settings:

- *FECdetectdegraded* = FEC degraded status condition over PM interval.
- *FECdetectexcessdegraded* = FEC excessively degraded status condition over PM interval.

Each of the above registers shall have a corresponding enable, status, and latch bit settings. *FECdetectdegraded* and *FECdetectexcessdegraded* shall also be a maskable interrupt.

PM interval:

- *PM_Interval* = (programmable); default = 1 second.

The FEC decoder counts and reports the number of bits detected in error over the PM interval per FEC block (min., max., avg.).

- When the (avg) number of bit errors exceed the threshold set in *FEC_degraded_BER_activate_threshold, FECdetectdegraded* is set and latched.
- When the (avg) number of bit errors fall below the threshold *FEC_degraded_BER_deactivate_threshold*, *FECdetectdegraded* is cleared.
- When the (avg) number of bit errors exceed the threshold set in *FEC_excessive_BER_activate_threshold, FECexcessdegraded* is set and latched.
- When the (avg) number of bit errors fall below the threshold *FEC_excessive_BER_deactivate_threshold*, *FECexcessdegraded* is cleared.

4.8 Maintenance Signals

Maintenance signals are used to replace certain portions of the 800ZR frame under specific conditions. Currently, maintenance signals are defined for maintenance lock (MNT LCK) conditions signaled in the STAT and CSTAT overhead defined in Sections 4.3.3.8 and 4.3.3.9.

4.8.1 STAT LCK maintenance signal

The STAT LCK pattern is specified in clause 8.5 of [ITU-T G.709.1].

During maintenance lock condition signaled in the MNT field of the STAT overhead (see section 4.3.3.8), a repeating “0101 0101” pattern replaces all the frame fields except the AM, MFAS and BOH CRC of the eight 100G ZR instances. Implementations may populate other overhead fields. This condition is signaled during “Modem Tx Loopback” as shown in Table 23.

4.8.2 CSTAT LCK maintenance signal

During maintenance lock condition signaled in the MNT field of the CSTAT overhead (see section 4.3.3.9) of a specific client, the client payload is replaced with an Ethernet local-fault (LF) signal. The LF signal must be generated at the client bit rate and within the client bit rate tolerance, possibly from a local clock.

4.9 Framed PRBS Test Signal

A framed 800ZR PRBS is used for validating ZR protocol framing, OFEC/DSP framing, symbol mapping, and FAW/TS/PS insertion. The required PRBS31 is per [IEEE 802.3] with initial state being all 1’s.

- Generation/checking is towards/from the media interface.
- The PRBS polynomial pattern is inserted per 100G ZR instance, replicated in all instances.
- The PRBS pattern will be identified by a unique payload type identifier.
- The PRBS in the Tx data path is inserted in the ZR payload area of a ZR frame structure with OH (see Figure 5 showing the payload area of each 100G ZR instance).
- The PRBS in the Rx data path is monitored from the ZR payload area. The PRBS checker shall recover and verify the PRBS31 sequence.

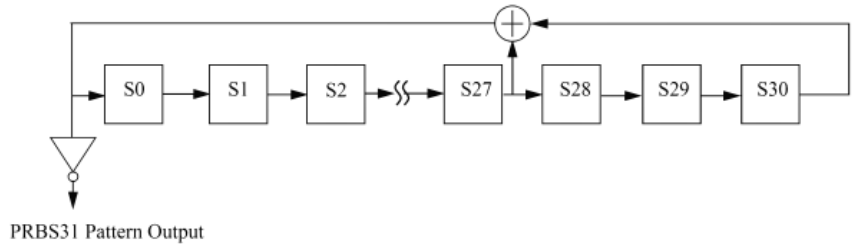


Figure 13: Framed PRBS generator

5 800ZR frame adaptation to 800ZR Coherent Interface

The 800ZR coherent interface uses OFEC for forward error correction and dual-polarization 16QAM (DP-16QAM) modulation. The complete process to adapt the 800ZR frame to the 800ZR coherent line interface is shown in Figure 14. The frame and adaptation processes described in this clause are the same as what is described in clauses 11 and 12 of [ITU-T G.709.6] for the case of x=8 and k=8. The 800ZR frame is viewed as 512 rows of 10,280 bits for FEC adaptation. The process includes the following steps:

1. Computation and insertion of CRC to aid detection of uncorrected codewords
2. Insertion of padding to align the payload blocks with the required FEC block lengths
3. Scrambling of the bits to remove any patterns present in the payload
4. Encoding the scrambled bits to add the OFEC parity bits
5. Interleaving to improve the burst error performance of the code
6. Mapping of the encoded bits to 16QAM symbols
7. Distribution of the resulting symbols to the X and Y polarizations to form DP-16QAM symbols
8. Insertion of FAW/TS/RES symbols to facilitate frame alignment and training
9. Insertion of pilot symbols to aid carrier phase recovery.

These steps are described in detail in the following sections.

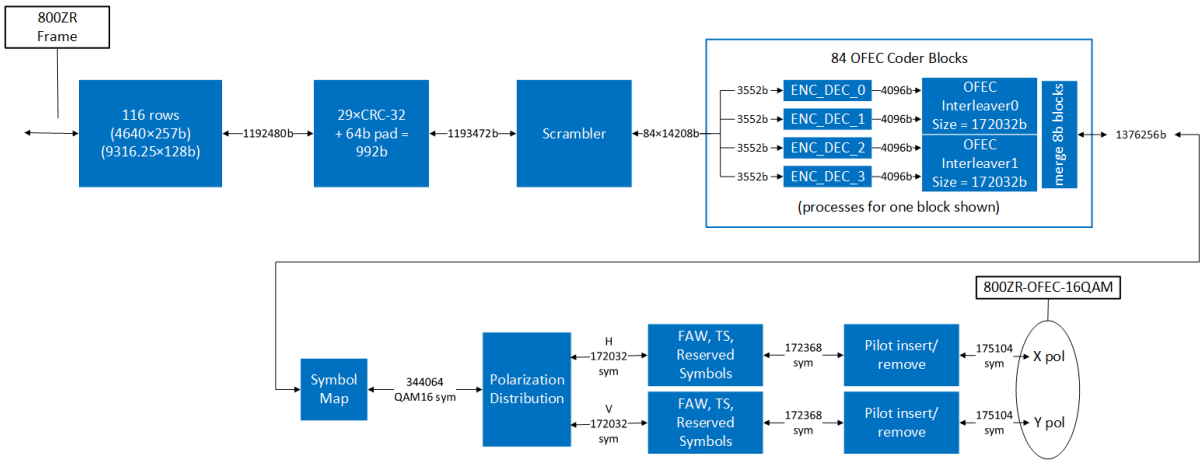


Figure 14: Adaptation of 800ZR frame to the 800ZR coherent line interface

5.1 Adaptation of 800ZR Frame payload to OFEC Block

The 800ZR coherent signal may be considered to comprise of a continuous series of DSP super-frames which are formed from the interleaved output of four independent OFEC encoders which encode a continuous series of 84 OFEC Coder Blocks numbered B_1 to B_{84}. Each OFEC Coder block B_i consists of 3,552 bits from each of the 4 encoders, for a total of 14,208 bits. The relationship of the 800ZR frame to the OFEC Coder blocks is shown in Figure 15.

- 116 rows × 10,280 bits of information (1,192,480bits) are adapted to 84 OFEC coder blocks of 14208 bits each.
- Each set of 84 OFEC coder blocks of [84×14,208] bits = 1,193,472bits carries 1,192,480 bits from the 800ZR frame + CRC [29×32b=928bits] + Pad [64bits].

The starting row of the 116 rows which comprise the OFEC coder block is arbitrary and does not have any specific relationship to the row containing the 800ZR alignment mechanism fields. As a result, the 116 rows of the 800ZR frame that are used to form the set of 84 OFEC Coder blocks “floats” with respect to the 800ZR frame structure.

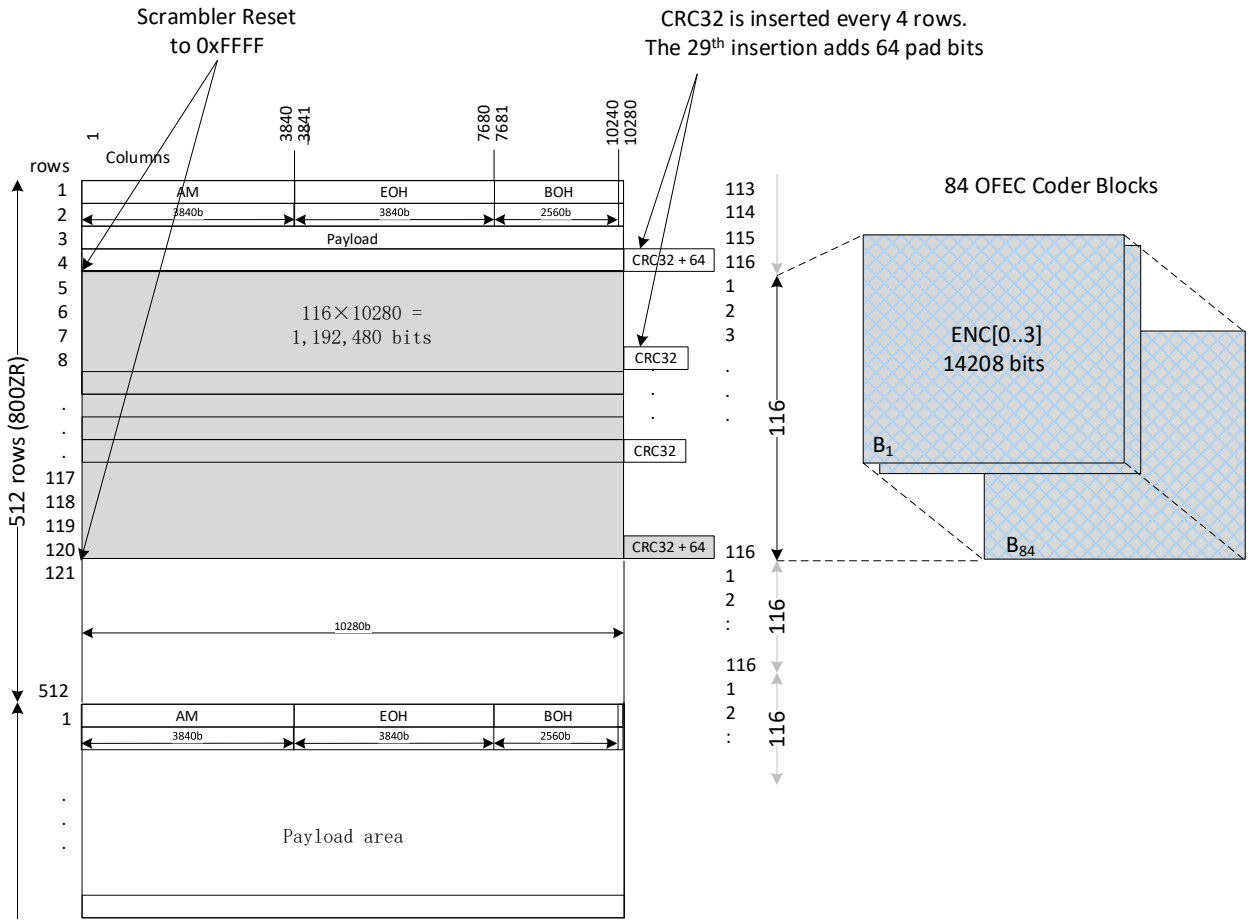


Figure 15: 800ZR frame to OFEC block relationship

5.2 800ZR CRC + Pad Insertion

Insertion of the CRC and pad bits is specified in clause 12.4.1 of [ITU-T G.709.6].

A 32-bit CRC is calculated over the [4×10,280 = 41,120] bits corresponding to 4 rows of the 800ZR frame as shown in Figure 15 and shall be inserted at the end of those 4 rows. The CRC is calculated with the generator polynomial from IEEE 802.3.

$$G(x) = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$$

Mathematically, the CRC value corresponding to the 41,120 input bits is defined by the following procedures:

- The first 32-bits of the input block are complemented.
- The 41,120 bits of the protected fields are the coefficients of a polynomial $M(x)$ of degree 41,119. (The first bit of the 41,120 input bits corresponds to the $x^{41,119}$ term and the last bit of the 41,120 input bits corresponds to the x^0 term).
- $M(x)$ is multiplied by x^{32} and divided (modulo 2) by $G(x)$, producing a remainder $R(x)$ of degree ≤ 31.

- The coefficients of $R(x)$ are a 32-bit sequence.
- The bit sequence is complemented, and the result is the CRC.

The 32-bits of the CRC value are placed with the x^{31} term as the left-most bit of the CRC32 field and the x^0 term as the right-most bit of the CRC32 field. (The bits of the CRC are thus transmitted in the order: $x^{31}, x^{30}, \cdots, x^1, x^0$).

At the end of the 116 rows of the 800ZR frame which map to 84 OFEC coder blocks, a 64-bit all-zero pad is inserted to align to a multiple of 3,552 bits required for the OFEC encoder.

5.3 Error Marking

The CRC field described in Section 5.2 may optionally be used in a receiver for error marking as described in 12.4.1.1.2 of [ITU-T G.709.6]. This error marking may be used to meet requirements of mean-time to false packet acceptance (MTTFPA) requirements of Ethernet clients. It should be noted that the relationship between the rows of the 800ZR frame used for CRC computation and the individual clients is complex due to the multiplexing and interleaving structure. As a result, individual implementations could differ in how and where the CRC errors are mapped to error blocks within the client streams.

5.4 Bit Transmission Order at input of scrambler

The transmission of the 84 OFEC Coder block payload shall start with row1, column 1 of the first row of the 116-row block, followed by row 1, column 2 until the row 4, column 10,280 has been transmitted. This shall be followed by the 32 bits of the CRC in the order described in Section 5.2.

The transmission shall continue in a similar fashion for every 4 rows until the CRC has been transmitted for row 116. This shall be followed by 64 bits of an all-zero PAD.

The entire sequence of steps shall be repeated for the next set of 116 rows of the 800ZR frame.

5.5 Frame Synchronous Scrambling

The scrambler/descrambler is located before the OFEC encoder on transmit and after the OFEC decoder on receive. The input to the scrambler is the 1,193,472 bits of the 84 OFEC Coder Block. Figure 16 shows a pictorial depiction of the scrambler function.

The operation of the scrambler shall be functionally equivalent to that of a frame-synchronous scrambler with a sequence length of 65535 and the generating polynomial shall be:

$$x^{16} + x^{12} + x^3 + x + 1$$

The scrambler/descrambler shall reset to 0xFFFF on the first bit of the 1,193,472 input sequence (which corresponds to row 1, column 1 of the group of 84 OFEC coder blocks) and advance during each bit of the input. All bits of the input sequence are scrambled. At the receiver the descrambler is synchronized (initialized) at the start of each block.

The scrambler is identical to the one specified in clause 12.4.3 of [ITU-T G.709.6].

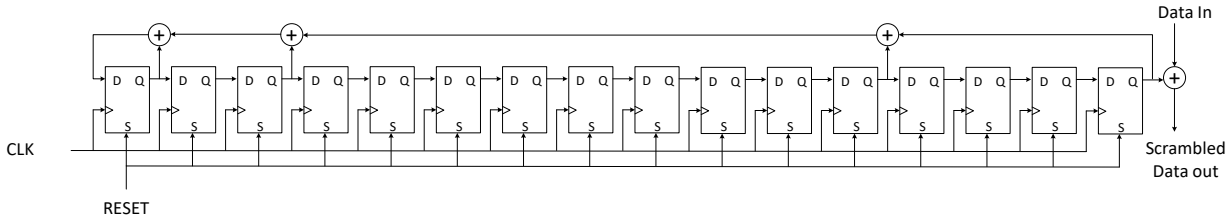


Figure 16: Pictorial depiction of the scrambler function

5.6 Encoding of OFEC Coder Blocks

The 1,193,472 bits at the output of the scrambler shall be processed as a set of 84 OFEC Coder blocks B_1 to B_{84} of 14,208 bits each. The OFEC Coder blocks shall be formed from 14,208 contiguous bits at the scrambler output in the time sequence B_1, B_2, …, B_{84}.

The 14,208 bits of each OFEC Coder block shall be distributed at a granularity of 1-bit to each of four parallel OFEC encoders in a round-robin fashion. The OFEC encoders are labelled ENC0, ENC1, ENC2, ENC3 in Figure 14 and shall receive the bits in time order to the encoder in the order ENC0, ENC1, ENC2 and ENC3. The OFEC encoders shall process the 3,552-bit input block to generate a 4,096-bit output block as described in Section 5.7.

The output of the OFEC Encoders is interleaved to improve the burst tolerance of the code. The processing for 800ZR involves the use of two parallel interleavers. The first interleaver shall receive the encoded output of ENC0 and ENC1. The second interleaver shall receive the encoded output of ENC2 and ENC3. The interleaving process implemented within each interleaver is described in Section 5.8.

The structure of the OFEC Coder Block is identical to the description in clause 12.4.4 of [ITU-T G.709.6].

5.7 OFEC Encoder

Each of the four OFEC encoders ENC0, ENC1, ENC2 and ENC3 shown in Figure 14 operates in parallel to produce an OFEC codeword. A codeword in OFEC is a semi-infinite set of bits organized in a matrix with semi-infinite number of rows and N columns (N=128).

It has the property that each bit is part of two “constituent component codewords,” in which each constituent component codeword is a binary vector x of length 2N satisfying the parity check constraint $xH = 0$, where H is a $(2N, 2N - k)$ binary parity check matrix, with $2N > k > N$. Here $k = 239$, and each constituent component codeword has $(2N - k) = 17$ parity bits. The fraction of bits that are parity bits is 17/111=15.3%.

Specifically in OFEC, H is a parity check matrix of an extended BCH(256,239) code. This BCH code has a minimum Hamming distance of 6. OFEC uses a textbook BCH encoding with a parity check matrix H that is specified below.

The constituent component codewords are ordered as explained below to allow high speed parallel encoding and decoding. To define what bits are part of a given constituent component codeword the following structure is used:

- The infinite matrix of bits is partitioned in square blocks of $B \times B$ bits ($B = 16$), arranged in rows and columns as shown in Figure 17. There are N/B blocks per row (N/B=8), and each square block is identified by a square block row number, R, and a square block column number,

C, where $C = 0,1,\cdots,\frac{N}{B} - 1$, appearing respectively on the left hand side and at the top of Figure 17.

- Each bit inside a square block is identified by its row number, r, where $r = 0,1,\cdots,B-1$, and column number, c, where $c = 0,1,\cdots,B-1$, where bit $(0,0)$ is at the upper left corner of a block. Overall, each bit in the infinite matrix is identified by a quadruple $\{R, C, r, c\}$.
- The number of guard-block rows needs to be even with a value $2G$. The value $G = 2$, or $2G = 4$ in Figure 17.

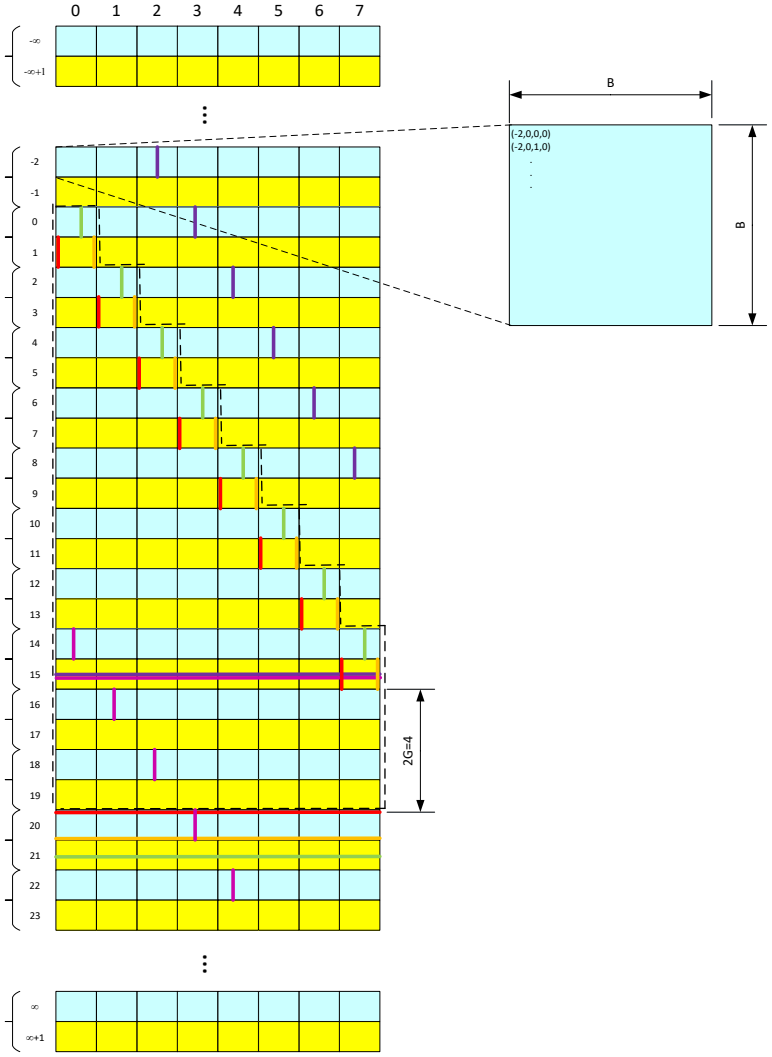


Figure 17: Structure of the OFEC codeword

A row of bits is identified by (R, r), with a square block row number R and a bit row number r within that block, where $r = 0,1,\cdots,B-1$. A constituent component codeword can be identified by the number of the row that contains all bits of the 2nd half of the codeword. The kth bit ($k = 0,1,\cdots,2N-1$) of a constituent component codeword (R, r) is the bit identified with the quadruple:

$$\text{If } k < N: \quad \left\{(R\hat{}1) - 2G - \frac{2N}{B} + 2\left\lfloor\frac{k}{B}\right\rfloor, \left\lfloor\frac{k}{B}\right\rfloor, (k\%B)\hat{}r, r\right\} \quad (1)$$

$$\text{If } k \geq N: \quad \left\{R, \left\lfloor\frac{k-N}{B}\right\rfloor, r, (k\%B)\hat{}r\right\} \quad (2)$$

Where,

$\lfloor . \rfloor$ denotes the floor operator

(a % b) denotes the value of a modulo b, and

(a ^ b) represents the number with a binary representation equal to the bitwise "exclusive or" of the binary representations of the numbers a and b

These formulas are illustrated in Figure 17. The union of line segments (both vertical and horizontal) of a given color shows the bits forming a constituent component codeword but the ordering in the segments is not the ordering in the codeword.

For example, consider the constituent component codeword (20,0). The position of its bits in the semi-infinite matrix are indicated by the red line segments. Bits 0 to 15 are in column 0 of block (1,0), bits 16 to 31 in column 0 of block (2,1), …, and bits 112 to 127 in column 0 of block (15,7). The bit indices go up as one descends in the columns.

Bits 128 to 255 are in row 0 of blocks (20,0) to (20,7), and their indices go up moving to the right within a row.

Bits 0 to 127 are referred to as the "front" of a constituent component codeword, and bits 128 to 255 as the "back".

Note that each bit in the OFEC encoder belongs to the front of a constituent component codeword and to the back of another one. Also, if the back of a constituent component codeword is an odd-numbered row of square blocks (yellow background), then its front is an even-numbered row of square blocks (blue background), and conversely.

The square blocks located below the "front bits" and above the "back bits" of a given constituent component codeword are so called guard blocks, relative to the constituent component codeword of interest.

Continuing the example, the bits of constituent component codeword (20,15), identified by the orange line segments, are in the same blocks as the segments of constituent component codeword (20,0). However, because "r" is 15 instead of 0 as in the previous example, the expressions "^r" in the formulas become significant, and the bits are taken in reverse order in each block. For example, bits 0 to 15 in the front of codeword (20,15) are bits 15 to 0 in column 15 of block (1,0).

The OFEC encoder is identical to the description in clause 12.4.5 of [ITU-T G.709.6].

Note: The OFEC code is a block-convolutional code, and its performance is characterized by its "error events". Without the "^r" permutation, there are about 625,000 possible error events of weight 36 that can start at every decoding of a constituent component codeword. For comparison, a Product Code based on the same constituent component codeword has more than 3.3e13 codewords of weight 36. The presence of "^r" permutation can be observed to eliminate error events of weight 36. Consequently, the minimum Hamming distance of the OFEC code is at least 42.

5.7.1 Encoding

Encoding is done sequentially, in order of increasing row index. At the time when a constituent component codeword (R, r) is being encoded, all constituent component codewords (R', r') with $R' < R - 2G$ would already be encoded.

To encode a constituent component codeword (R, r), form a vector x of length $2N$ where the front N bits are read from previously encoded bits in the infinite matrix according to formula (1) above. In the back, the first $k - N = 111$bits are fresh information bits. The last $2N - k = 17$ back bits are the parity bits which are calculated to satisfy the parity check equation $xH = 0$. After encoding, the N back bits are placed at their positions in the infinite matrix according to equation (2) above, and bits in their positions are output to the interleaver.

Considering Figure 17, we see that G is large enough to allow parallel encoding of $2B(G + 1) = 96$ constituent component codewords, assuming the pipeline delay is small. This number is considerably reduced when the pipeline delay increases, which is typically the case in the decoder.

One can also see that at most N/B(N/B+2G+1) =104 square blocks need to be kept in the encoder memory (excluding the current input). The square blocks that must be kept in memory to encode block rows 20 and 21 are surrounded by the dashed line in Figure 17.

A large G allows for longer pipeline delays in the encoding and decoding operations and allows for more parallel execution in the encoder and decoder, at the expense of increased memory.

5.7.2 Encoder interface

The encoder input consists of rectangular blocks of size $(2B) \times (2N - k) = 32 \times 111$ bits. The encoder input blocks are numbered $0,1,2,\cdots$. The input bits into the encoder are sequenced. The i[th] input bit is placed in the encoder input block $\lfloor i/\ (32 \times 111) \rfloor$ at the position indicated by the value $i\%(32 \times 111)$ in Figure 18. Note that an encoder input block is divided into 16×16-bit blocks, except along the right edge where the size is 16×15.

Bit $k = 0,1,2,\cdots$ of row p in encoder input block P is placed in position N+k of constituent component codeword $(2P + \lfloor p/B \rfloor, p\%B)$.

0 1 … 15 16 17 … 31 32… . . . . 240 .. . 255	512… . . . . . … 767	1024 … . . . . . … 1279	1536 … . . . . . … 1791	2048 … . . . . . … 2303	2560 … . . . . . … 2815	3072 3073 … 3086 3087 … 3101 3102 … 3599 . . . . 3297 … 3311
256… 271 272 . . . . 496 511	768 … . . . . . …. 1023	1280 … . . . . . … 1535	1792 .. . . … 2047	2304… . . . . . … 2559	2816.. . . . . . … 3071	3312 … 3326 3327 … . . . . 3537 … 3551

Figure 18: Sequencing of bits within an OFEC encoder input block

The encoder output consists of rectangular blocks of size $(2B) \times N = 32 \times 128$ bits. The encoder output blocks are numbered $0,1,2,\cdots$. Bit k=0,1,2,… of row p in rectangle P is the bit $\{2P + \left\lfloor \frac{p}{B} \right\rfloor, \left\lfloor \frac{k}{B} \right\rfloor, p\%B, k\%B\}$ of the semi-infinite array.

The bits within an output block of the encoder are sequenced as shown in Figure 19.

0 1 … 15 16 17 … 31 32… . . . . 240 .. . 255	512… . . . . . . … 767	1024 … . . . . . . … 1279	1536 … . . . . . . … 1791	2048 … . . . . . . … 2303	2560 … . . . . . . … 2815	3072 . . . . . . … 3327	3584 . . . . . . … 3839
256… 271 272 . . . . . 496 511	768 … . . . . . . …. 1023	1280 … . . . . . . … 1535	1792 .. . . … 2047	2304… . . . . . . … 2559	2816.. . . . . . . … 3071	3328 . . . . . . … 3583	3840 . . . . . . … 4095

Figure 19: Sequencing of bits within an OFEC encoder output block

5.7.3 Formal encoder definition

This section directly describes the encoder (ENC0, ENC1, ENC2 or ENC3) output bits as a function of the input bits, integrating the diverse elements that have been described in previous sections.

An OFEC encoder is an entity that produces a binary output y(i) from a binary input u(i), where i=0,1,2,…

The relationship between y and u is expressed through intermediate variables.

In particular, there is a four dimensional array V(R,C,r,c), where R is an integer; C=0,1,…,7; r=0,1,…,15; and c=0,1,…,15.

Associated with array V, there are constituent component codeword vectors $W_{R,rw}$ with elements $W_{R,r}(i)$, where R ≥ 0, r = 0, 1. 2….15, and i = 0, 1, …, 255.

$$\text{For } R \geq 0,\ W_{R,r}(k) = \begin{cases} V((R \wedge 1) - 20 + 2\,[k/16],\ [k/16],\ (k\ \%\ 16) \wedge r,\ r) \text{ for } k < 128 \\ V(R,\ [(k - 128)/16],\ r,\ (k\ \%\ 16) \wedge r) \text{ for } 128 \leq k < 256 \end{cases}$$

where,

- [.] denotes the floor operator,
- (a % b) denotes the value of a modulo b, and

- (a ^ b) represents the number with a binary representation equal to the bitwise “exclusive or” of the binary representations of the numbers a and b.

The bits in the $W_{R,r}$ satisfy the following equalities:

For R ≥ 0, r = 0, 1, …, 15 and k = 0, 1, …, 110

$$W_{R,r}(128 + k) = u([R/2] \times 32 \times 111 + ((R \% 2) \times 16 + r) \times (16 – [k/96]) + [k/16] \times 512 + k \% 16)$$

For R ≥ 20, $W_{R,r}$ H = 0, where H is a parity check matrix of an extended BCH(256, 239) code, using a textbook encoding; i.e., if x is a vector satisfying xH = 0, then

1. x has an even parity, and
2. if the first 255 bits of x are seen as the binary coefficients of a polynomial x(t) of degree 254 (with bit 0 of x being the coefficient of power 254), with t being the indeterminate, then this binary polynomial x(t) is divisible by the binary codeword generator polynomial $t^{16} + t^{14} + t^{13} + t^{11} + t^{10} + t^9 + t^8 + t^6 + t^5 + t + 1$.

The output y satisfies the relationship

For R ≥ 0; C = 0, 1, ..., 7; r = 0, 1, …, 15; and c = 0, 1, …, 15.

$$V (R, C, r, c) = y([R/2] \times 32 \times 128 + (R \% 2) \times 256 + C \times 16 \times 32 + r \times 16 + c)$$

It can be observed that 20 × 16 ×17 values are left undefined in $W_{R,r}$ and in V(R, C, r, c) for 0 ≤ R < 20, and thus also in the output y. This is by design; an implementation can choose any convenient values.

However, for test vectors, the output needs to be totally specified. To that end, the following additional constraints are added:

For 0 ≤ R < 20, $W_{R,r}$ H’ = 0, where H’ is a 256 × 17 binary matrix where the first 128 rows are all zero and the last 128 rows are equal to the last 128 rows of H.

5.8 OFEC Interleaver

This section describes details of each of the two parallel interleavers (Interleaver0 and Interleaver1) shown in Figure 14. After OFEC encoding, the output of two encoders is interleaved by a block interleaver. The interleaver block size is 172,032 bits (42 encoder output blocks, 21 blocks each from ENC0/ENC1 or ENC2/ENC3 at the input of each interleaver). The number of interleaver blocks per each codeblock for each interleaver is 1376256/172032=8.

The OFEC interleaver is identical to the description in clause 12.4.6 of [ITU-T G.709.6].

5.8.1 OFEC Interleaver architecture

The 172,032 bits in an interleaver block are organized as an (84,8) array of 16-bit×16-bit square blocks as shown in Figure 21. Note that the format is similar to the format used by the encoder and decoder. The interleaver then consists of the following two mechanisms:

1. An intra-block interleaver that reorders the bits in each 16×16 square block to ensure that the bits in each row and column of a square block at the encoder output are remapped almost uniformly in the square block for transmission on the line. That operation can be seen as happening at the input to the interleaver.
2. An inter-block interleaver that attempts to have nearby symbols on the line contain bits that are widely separated in the encoder output.

The interleaver is half rate and fed by two quarter-rate encoders ENC0/ENC1 or ENC2/ENC3. Successive rows of square blocks from ENC0 or ENC2 will be written in even block rows of the interleaver buffers (yellowish colors in Figure 21), whereas successive rows of square blocks from ENC1 or ENC3 will be written in odd block rows (pinkish colors). Consequently, the content of an interleaver buffer is the square block row by square block row interleaving of vertical segments of the semi-infinite matrices of encoders ENC0/ENC1 or ENC2/ENC3.

5.8.2 Intra-block interleaving

For the purposes of intra-block interleaving, the interleaver is considered to receive 16×16 square blocks of bits from the encoders, and each square block is considered separately.

The intra-block interleaving is specified in Figure 20, which indicates the row and column of the source bit for each destination bit in the square block. For example, bit (14,15) [base 0] encoder output block is place in row 1 of column 0 of the corresponding interleaver square block.

0,0	1,1	2,2	3,3	4,4	5,5	6,6	7,7	8,8	9,9	10,10	11,11	12,12	13,13	14,14	15,15
14,15	15,0	0,1	1,2	2,3	3,4	4,5	5,6	6,7	7,8	8,9	9,10	10,11	11,12	12,13	13,14
12,14	13,15	14,0	15,1	0,2	1,3	2,4	3,5	4,6	5,7	6,8	7,9	8,10	9,11	10,12	11,13
10,13	11,14	12,15	13,0	14,1	15,2	0,3	1,4	2,5	3,6	4,7	5,8	6,9	7,10	8,11	9,12
8,12	9,13	10,14	11,15	12,0	13,1	14,2	15,3	0,4	1,5	2,6	3,7	4,8	5,9	6,10	7,11
6,11	7,12	8,13	9,14	10,15	11,0	12,1	13,2	14,3	15,4	0,5	1,6	2,7	3,8	4,9	5,10
4,10	5,11	6,12	7,13	8,14	9,15	10,0	11,1	12,2	13,3	14,4	15,5	0,6	1,7	2,8	3,9
2,9	3,10	4,11	5,12	6,13	7,14	8,15	9,0	10,1	11,2	12,3	13,4	14,5	15,6	0,7	1,8
15,7	0,8	1,9	2,10	3,11	4,12	5,13	6,14	7,15	8,0	9,1	10,2	11,3	12,4	13,5	14,6
13,6	14,7	15,8	0,9	1,10	2,11	3,12	4,13	5,14	6,15	7,0	8,1	9,2	10,3	11,4	12,5
11,5	12,6	13,7	14,8	15,9	0,10	1,11	2,12	3,13	4,14	5,15	6,0	7,1	8,2	9,3	10,4
9,4	10,5	11,6	12,7	13,8	14,9	15,10	0,11	1,12	2,13	3,14	4,15	5,0	6,1	7,2	8,3
7,3	8,4	9,5	10,6	11,7	12,8	13,9	14,10	15,11	0,12	1,13	2,14	3,15	4,0	5,1	6,2
5,2	6,3	7,4	8,5	9,6	10,7	11,8	12,9	13,10	14,11	15,12	0,13	1,14	2,15	3,0	4,1
3,1	4,2	5,3	6,4	7,5	8,6	9,7	10,8	11,9	12,10	13,11	14,12	15,13	0,14	1,15	2,0
1,0	2,1	3,2	4,3	5,4	6,5	7,6	8,7	9,8	10,9	11,10	12,11	13,12	14,13	15,14	0,15

Figure 20: Source positions (row,col) for intra-block interleaving

Note: The left entries of the pairs in Figure 20 *form a Latin Square. The right entries almost form a Latin Square, but they are duplicated in the first and last rows.*

5.8.3 Inter-block interleaving

The intra-block permutation described in the previous section is applied to each square block in the buffer as it comes in from the encoder.

In addition to partitioning the interleaver buffer as a function of the encoder, ENC0/ENC1 or ENC2/ENC3, it is also partitioned to an upper half of 42 block rows (light color tones) and a lower half of 42 block rows (dark color tones). Overall, the buffer is then partitioned into 4 subsets, each containing 21×8 square blocks or 336×128 bits.

Subset number	Row Blocks
0	0, 2, …, 40

1	1, 3, …, 41
2	42, 44, …, 82
3	43, 45, …, 83

Table 13: Interleaver subsets

On output, groups of 8 bits are taken in turn from each subset, reading them out of a column of bits before proceeding to the next columns of bits.

As shown in Figure 21, the first 8 bits are read from the top of first column of subset 0, then the first 8 bits from the first column of subsets 1,2 and 3. Those 32 bits are then followed by taking the next 8 bits in the first column of each of the subsets 0,1,2 and 3. After 42 such cycles of 4×8 bits each, the first bit column of the interleaver buffer will be completely read out, and the output process continues by reading bit columns 1 to 127.

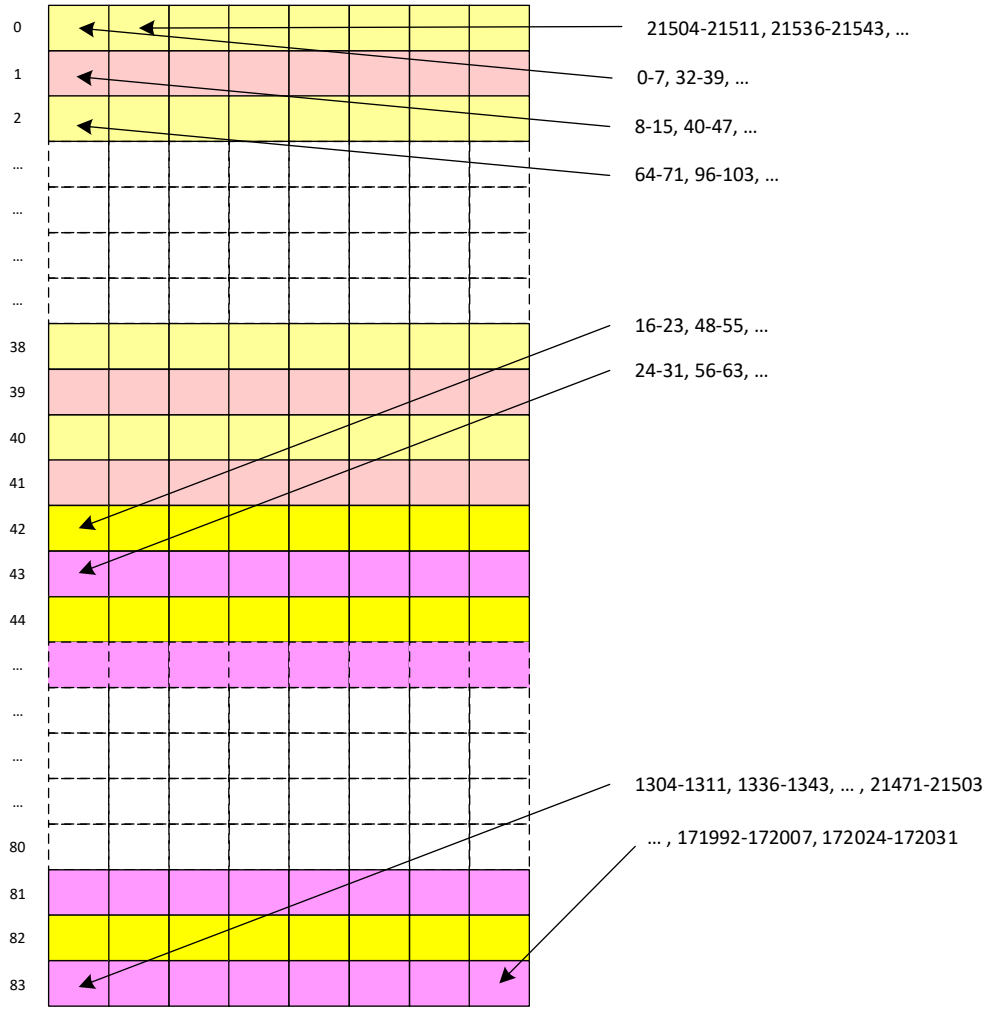


Figure 21: Inter-block interleaving

Note: Bits are read by columns, rather than rows because interleaver columns are much longer than rows, so bits in a column are spread over more constituent component codewords than bits in a row, which increases the tolerance to long bursts. The maximum correctable burst length, when used with a hard decoder, is a traditional measure of interleaver quality. In this case it can be shown to be 2,681 bits.

The bits read out of the interleavers are passed to the modulator where they are used in groups of 8 bits to map to DP-16QAM symbols.

5.9 Symbol Mapping and Polarization Distribution

The output bits of the two OFEC interleavers (interleaver0 and interleaver1) are processed in groups of 8 bits by a symbol mapper and polarization distribution i.e., 8 bits are mapped to a DP-16QAM symbol from interleaver0 followed by 8 bits from interleaver1 and so on. The 84 OFEC Coder Block output of 1,376,256 bits after interleaving shall be mapped to a group of 172,032 DP-16QAM symbols.

The input bits of the symbol mapper are denoted by c_k (k=0…1376255). The symbol mapper shall receive 8-bit blocks from each interleaver in time order in a round-robin fashion i.e., bit 0-7 from the first interleaver, followed by bit 0-7 from the second interleaver, followed by bit 8-15 from the first interleaver, bit 8-15 from the second interleaver and so on.

The symbol mapper shall map the input bits c_k (k=0…1376255) to DP-16QAM symbols denoted s_i (i=0…172031), where,

- (c_{8i}, c_{8i+2}) shall map to the in-phase (I) component of the X-polarization of s_i
- (c_{8i+4}, c_{8i+6}) shall map to the quadrature-phase (Q) component of the X-polarization of s_i
- (c_{8i+1}, c_{8i+3}) shall map to the in-phase (I) component of the Y-polarization of s_i
- (c_{8i+5}, c_{8i+7}) shall map to the quadrature-phase (Q) component of the Y-polarization of s_i

In each signaling dimension (XI/XQ/YI/YQ), the mapping from binary label to relative symbol amplitude shall be:

$$(0,0) \rightarrow -3, \qquad (0,1) \rightarrow -1, \qquad (1,1) \rightarrow +1, \qquad (1,0) \rightarrow +3$$

This mapping per polarization is further detailed in Table 14 below.

Symbol mapping and polarization distribution is identical to the description in clause 12.4.7.1 of [ITU-T G.709.6].

(c_{8i}, c_{8i+2}, c_{8i+4}, c_{8i+6}) or (c_{8i+1}, c_{8i+3}, c_{8i+5}, c_{8i+7})	**I**	**Q**
(0,0,0,0)	-3	-3
(0,0,0,1)	-3	-1
(0,0,1,0)	-3	+3
(0,0,1,1)	-3	+1
(0,1,0,0)	-1	-3
(0,1,0,1)	-1	-1
(0,1,1,0)	-1	+3

(0,1,1,1)	-1	+1
(1,0,0,0)	+3	-3
(1,0,0,1)	+3	-1
(1,0,1,0)	+3	+3
(1,0,1,1)	+3	+1
(1,1,0,0)	+1	-3
(1,1,0,1)	+1	-1
(1,1,1,0)	+1	+3
(1,1,1,1)	+1	+1

Table 14: DP-16QAM symbol amplitude map

5.10 800ZR DSP Framing

A DSP super-frame is formed from the group of 172,032 DP-16QAM symbols at the output of the symbol mapper and polarization distribution. The DSP super-frame is defined as a set of 175,104 DP-16QAM symbols. A DSP super-frame is further divided into 24 DSP sub-frames, each with 7,296 symbols. The details of the DSP super-frame are shown in Figure 22.

Pilot symbols shall be inserted every 64 symbols, starting with the first symbol of each DSP super-frame. The first 11 symbols of each DSP sub-frame can also be used for training (e.g., frame acquisition). The first symbol of the Training Sequence (TS) is a Pilot Symbol (PS).

- Every DSP subframe has the same structure based on a fixed TS with the first symbol processed as a pilot.
- The TS includes 11 QPSK symbols in each polarization. The TS is different between X and Y polarizations.
- The PS sequence includes (1+113) QPSK symbols based on PRBS. The first TS symbol is also the first symbol of the PS sequence.

The DSP frame is identical to the description in clause 11 of [ITU-T G.709.6].

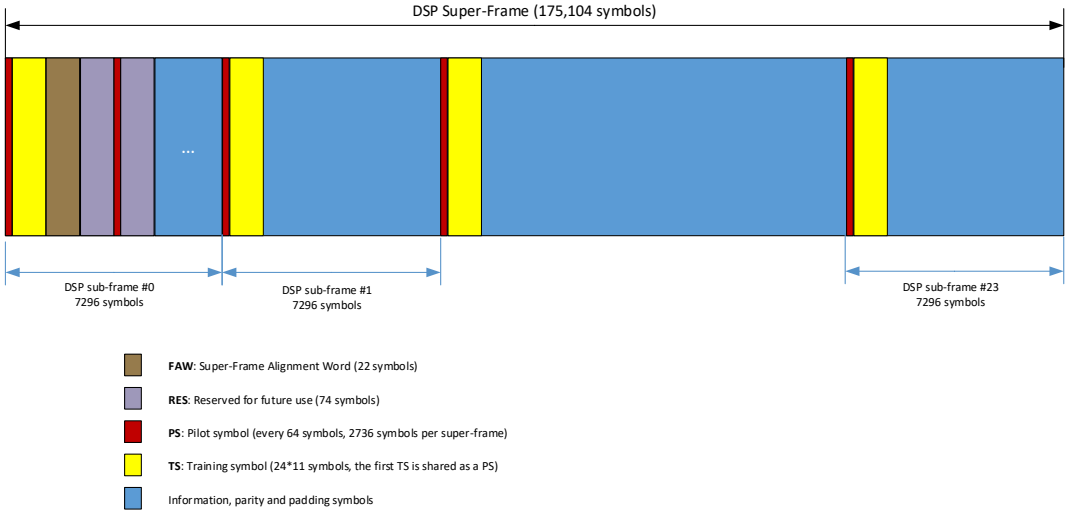


Figure 22: DSP frame format

5.10.1 First DSP sub-frame

The first DSP sub-frame of the super-frame shall include a 22 symbol Frame Alignment Word (FAW) used to align the 84 OFEC Code block frame. An additional 74 symbols are reserved for future use.

The first DSP sub-frame includes:

- 22 symbols used as the Frame Alignment Word (FAW). The FAW is different between X and Y polarizations.
- 74 symbols are reserved for future use. The symbols should be randomized to avoid strong tones. These symbols should be selected from 16QAM modulation.

5.10.2 FAW Sequence

The FAW sequence is the DSP frame MFAS specified in Annex A.1.3 of [ITU-T G.709.6]. It is referred to as the FAW in this document, and is represented as constellation points for convenience.

The FAW sequence at the beginning of each DSP super-frame shall match the sequence described in Table 15.

Index	FAW X	FAW Y	Index	FAW X	FAW Y
1	3-3j	3+3j	12	3-3j	-3+3j
2	3+3j	-3+3j	13	-3-3j	-3+3j
3	3+3j	-3-3j	14	-3-3j	3+3j
4	3+3j	-3+3j	15	-3+3j	-3-3j

5	3-3j	3-3j	16	3+3j	3+3j
6	3-3j	3+3j	17	-3-3j	-3-3j
7	-3-3j	3-3j	18	3-3j	-3+3j
8	3+3j	3-3j	19	-3+3j	3-3j
9	-3-3j	-3-3j	20	3+3j	-3-3j
10	-3+3j	3-3j	21	-3-3j	3-3j
11	-3+3j	3+3j	22	-3+3j	-3+3j

Table 15: FAW Sequence

5.10.3 Subsequent DSP sub-frames

Each subsequent DSP sub-frame after the first includes an 11 symbol TS, the first symbol of which is a PS.

5.10.4 Training Sequence

The training sequence is specified in Annex A.1.1 of [ITU-T G.709.6]. It is represented here using the constellation points for convenience.

The training sequence symbols TS shall match the symbols defined in Table 16 and transmitted in the same order. The first symbol of the TS is processed as a pilot symbol.

Index	Training X	Training Y
1*	-3+3j	-3-3j
2	3+3j	-3-3j
3	-3+3j	3-3j
4	3+3j	-3+3j
5	-3-3j	-3+3j
6	3+3j	3+3j
7	-3-3j	-3-3j
8	-3-3j	-3+3j
9	3+3j	3-3j
10	3-3j	3+3j
11	3-3j	3-3j

Table 16: Training symbol sequence

5.10.5 Pilot Sequence

The pilot sequence is specified in Annex A.1.2.1 of [ITU-T G.709.6]. It is represented here using the constellation points for convenience.

A pilot sequence formed from the outer symbols of the 16QAM constellation is inserted every 64 symbols in each DSP sub-frame. The pilot sequence is a fixed 114 symbol sequence formed from a PRBS10 sequence mapped to QPSK with different seeds for the X/Y polarizations as shown in Figure 23.

- Seeds are selected so that the pilot and training sequence combined are DC balanced
- Seeds are selected so that the first symbol in the training sequence is also the first symbol in the pilot sequence
- The seed is reset at the start of every DSP sub-frame

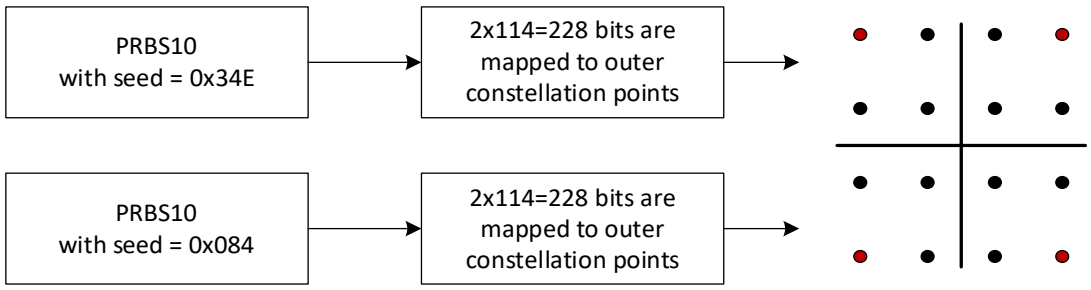


Figure 23: QPSK mapped pilot sequence

The generator polynomial and seed values for the pilot sequence are shown in Table 17.

Generator polynomial	Seed X	Seed Y
$x^{10} + x^7 + x^3 + x + 1$	0x34E	0x084

Table 17: PRBS10 polynomial and seed values for pilot sequence

Figure 24 shows the sequencing.

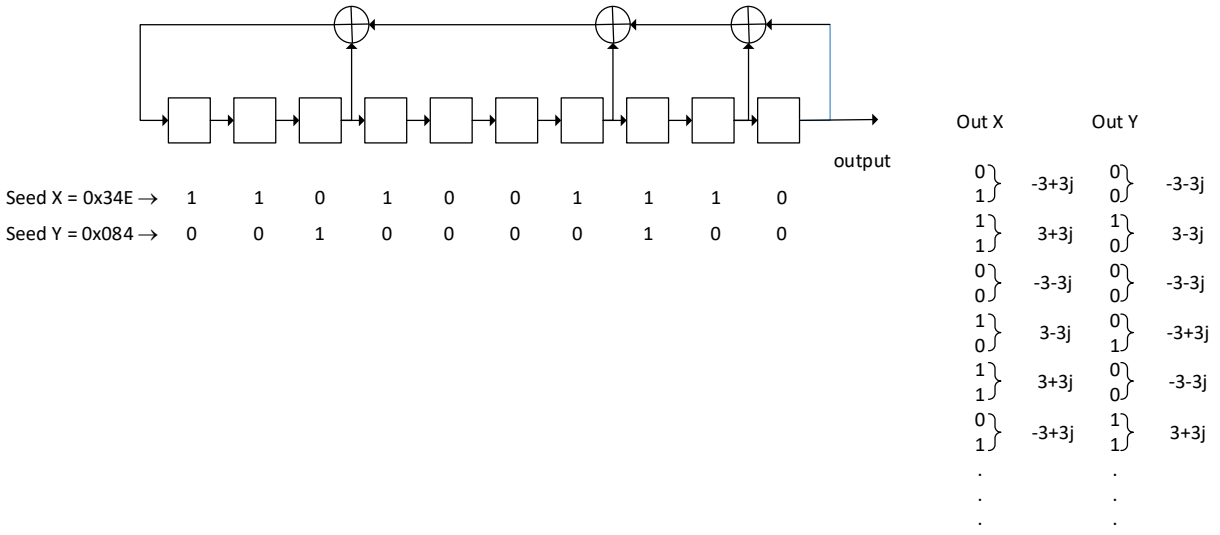


Figure 24: Pilot seed and sequencing

The complete pilot sequence is shown in Table 18.

Index	Pilot X	Pilot Y	Index	Pilot X	Pilot Y	Index	Pilot X	Pilot Y	Index	Pilot X	Pilot Y
1	-3+3j	-3-3j	30	3-3j	-3+3j	59	3-3j	-3+3j	88	-3+3j	-3-3j
2	3+3j	3-3j	31	3+3j	3+3j	60	3-3j	-3+3j	89	3+3j	3+3j
3	-3-3j	-3-3j	32	3-3j	3+3j	61	-3-3j	-3-3j	90	3-3j	3+3j
4	3-3j	-3+3j	33	-3+3j	-3-3j	62	3+3j	-3-3j	91	3-3j	3-3j
5	3+3j	-3-3j	34	-3-3j	-3+3j	63	3-3j	3+3j	92	-3-3j	3+3j
6	-3+3j	3+3j	35	3-3j	-3+3j	64	-3-3j	-3-3j	93	-3+3j	3+3j
7	3-3j	3+3j	36	-3+3j	-3-3j	65	3+3j	-3+3j	94	-3+3j	-3-3j
8	3+3j	-3+3j	37	-3-3j	3+3j	66	3+3j	-3+3j	95	-3+3j	3+3j
9	-3-3j	-3-3j	38	-3-3j	3+3j	67	3+3j	-3+3j	96	-3-3j	-3+3j
10	-3-3j	-3-3j	39	-3+3j	3-3j	68	-3-3j	3+3j	97	-3-3j	-3-3j
11	3+3j	3-3j	40	-3-3j	-3+3j	69	-3-3j	-3+3j	98	-3+3j	3-3j
12	-3-3j	3-3j	41	-3+3j	3-3j	70	3+3j	3+3j	99	-3+3j	3-3j
13	3+3j	3-3j	42	-3-3j	-3-3j	71	3+3j	3+3j	100	-3+3j	3-3j
14	-3-3j	-3-3j	43	-3-3j	-3-3j	72	3-3j	3+3j	101	3-3j	3+3j
15	-3+3j	-3-3j	44	3+3j	-3+3j	73	3-3j	-3+3j	102	-3+3j	-3+3j
16	-3-3j	3-3j	45	3+3j	3-3j	74	3+3j	3+3j	103	-3+3j	-3-3j

17	-3+3j	3-3j	46	-3-3j	-3+3j	75	3-3j	-3-3j	104	3-3j	-3-3j
18	-3+3j	3+3j	47	3+3j	-3-3j	76	3-3j	-3-3j	105	-3+3j	-3+3j
19	-3-3j	-3-3j	48	3+3j	-3+3j	77	3+3j	3-3j	106	-3-3j	3+3j
20	-3+3j	3-3j	49	-3+3j	3-3j	78	-3+3j	-3+3j	107	3-3j	-3-3j
21	3-3j	3+3j	50	3-3j	3+3j	79	3+3j	-3+3j	108	3+3j	3-3j
22	3-3j	-3-3j	51	-3+3j	3-3j	80	-3+3j	3+3j	109	-3+3j	3-3j
23	3-3j	3-3j	52	3-3j	-3-3j	81	3+3j	3-3j	110	-3-3j	-3+3j
24	3-3j	-3+3j	53	-3+3j	3+3j	82	-3+3j	3+3j	111	3-3j	-3-3j
25	3-3j	-3+3j	54	3-3j	-3+3j	83	3-3j	-3+3j	112	-3-3j	3-3j
26	3+3j	3-3j	55	3-3j	3-3j	84	-3+3j	3+3j	113	3+3j	3-3j
27	3+3j	3-3j	56	3+3j	-3-3j	85	-3-3j	3-3j	114	3+3j	-3+3j
28	-3+3j	-3+3j	57	-3-3j	3+3j	86	-3-3j	3-3j			
29	-3-3j	-3-3j	58	-3+3j	3+3j	87	3-3j	3+3j			

Table 18: 800ZR Pilot Sequence

5.11 Channel Mappings

X and Y indicate a pair of mutually orthogonal polarizations of any orientation and I and Q are mutually orthogonal phase channels in each polarization. The four data path channels are therefore labeled XI, XQ, YI, and YQ.

All coherent channel mappings provided in Table 19 are allowed for the TX signal. The Rx should work in all cases because it can unambiguously identify the polarization and phase of the signal based on the FAW.

The Tx mapping is specified in Table 19 by two designations: X:Y and I,Q where a ":" is used to separate X &Y, and a "," is used to separate I & Q.

Mapping	X:Y	I,Q	Notes
[0,x]	X:Y		Polarization cannot be interleaved
[1,x]	Y:X		
[x,0]		I,Q:I,Q	I/Q Same across Pol
[x,1]		Q,I:Q,I	
[x,2]		I,Q:Q,I	I/Q Flip across Pol
[x,3]		Q,I:I,Q	

Table 19: Channel Mappings

5.12 Frame Expansion Rate

The 800ZR optical signal is DP-16QAM with a nominal symbol rate of 118.203350603 Gbaud. Table 20 details the bit rates at various stages of the frame adaptation.

Offset	800ZR Frame (bps)	After PAD and CRC insertion (bps)	After OFEC encoding (bps)	After FAW,TS,RES insertion (bps)	After PS insertion (bps)	Symbol Rate (Baud)
-20ppm	804,899,008,669	805,633,039,811	929,018,280,143	930,832,768,971	945,607,892,288	118,200,986,536
0ppm	804,979,506,619	805,649,152,794	929,036,860,880	930,851,385,999	945,626,804,824	118,203,350,603
20ppm	805,060,004,570	805,665,265,777	929,055,441,617	930,870,003,027	945,645,717,360	118,205,714,670

Table 20: 800ZR bit rate expansion

6 Optical Specification

The 800ZR optical parameters are organized by Application codes (defined in Table 21) for the Tx, Rx, and the Optical Channel.

Ref	Application Description	Minimum Reach	Application Code – Name
6.0.010	120km or less, amplified, point-to-point, DWDM noise-limited links.	80km	***0x01*** – 800ZR-A, 150GHz DWDM, Tx Output Range A.
6.0.020	120km or less, amplified, point-to-point, DWDM noise-limited links.	80km	***0x02*** – 800ZR-B, 150GHz DWDM, Tx Output Range B.
6.0.030	120km or less, amplified, point-to-point, DWDM noise-limited links. An engineered link may be required to achieve full fill DWDM.	80km	***0x03*** – 800ZR-C, 150GHz DWDM, Tx Output Range C.

Table 21: 800ZR Applications codes

Note: All specifications are defined after calibration and compensation, at EOL over temperature and wavelength.

Transmitter specifications are relative to S_s, whereas Receiver specifications are relative to R_s as shown in Figure 1.

Bold italicized items found in tables indicate a reference to a Coherent Common Management Interface Specification [C-CMIS] or a Common Management Interface Specification [CMIS] defined function, state, or status condition.

6.1 800ZR, DWDM, Amplified - Application Code (***0x01, 0x02, 0x03***)

6.1.1 Optical Channel Specifications

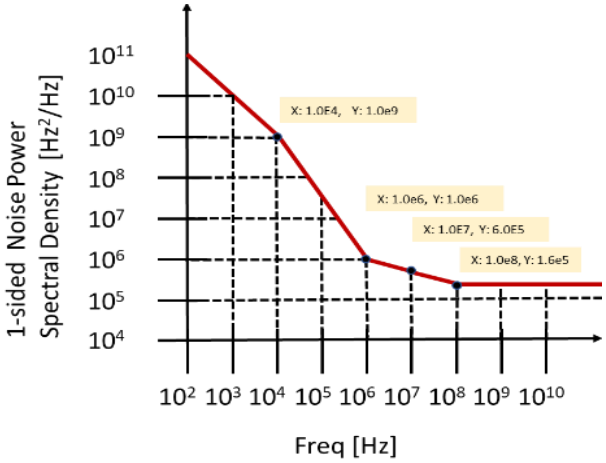
The mux/demux filters in the optical channel are not normatively defined as was done previously in the 400ZR IA.

Ref.	Parameter	Min	Max	Unit	Conditions/Comments
6.1.100	Channel Frequency	191.375	196.025	THz	
6.1.110	Channel spacing[†]	150		GHz	The same center frequencies as the even 400ZR channels are used
6.1.120	Post FEC BER		10^{-15}		For optical systems that conform to this IA
6.1.160	Chromatic Dispersion	0	2400	ps/nm	Frequency dependent change in phase velocity due to fiber based on G.652 fiber and a target reach of 80km
6.1.161	Optical Return Loss at S_s	24		dB	See definition in 6.4.2.
6.1.162	Discrete Reflectance between S_s and R_s		-27	dB	See definition in 6.4.3.
6.1.170	Maximum Differential Group Delay (DGD_{max})		28	ps	See definition in 6.4.4.
6.1.171	Polarization Dependent Loss (PDL)		2	dB	See definition in 6.4.5.
6.1.172	Polarization Rotation Speed		50	krad/s	See definition in 6.4.6.

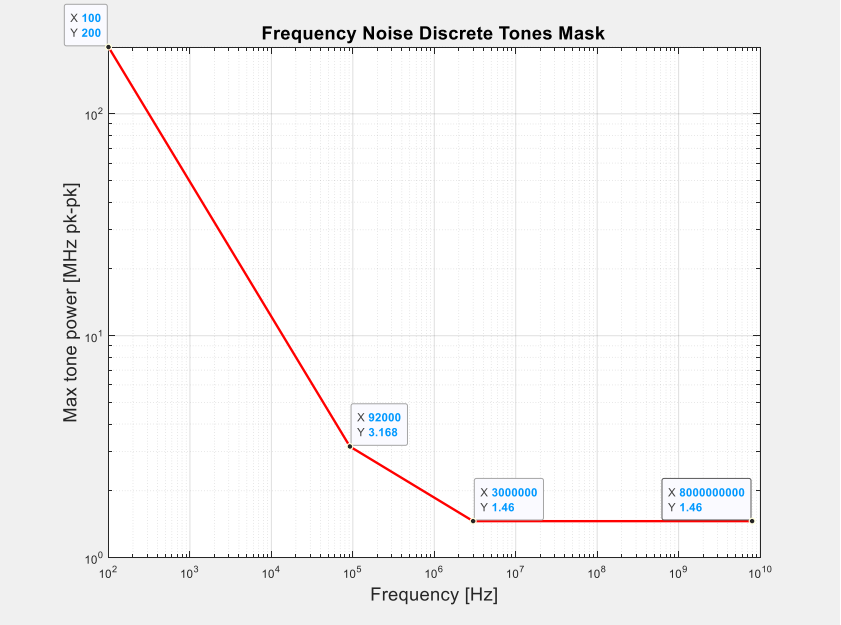
Table 22 800ZR Optical Channel Specifications

[†]The allowed channel frequencies (in THz) are defined by 193.1 + (*n*+3) × 0.025 where *n* is a positive or negative integer divisible by 6, including 0. For 800ZR modules, *n* = 114 to -72. The normative 32 × 150 GHz DWDM application channels are as defined in Section 8.1.

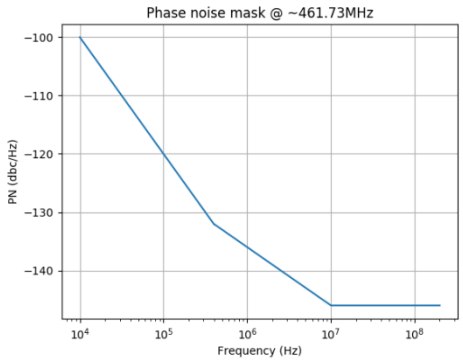
6.1.2 Transmitter Optical Specifications

<table>
<tr><th>Ref</th><th>Parameter</th><th>Unit</th><th>Min</th><th>Max</th><th>Conditions/Comments</th></tr>
<tr><td>6.1.200a</td><td>Channel Frequency</td><td>THz</td><td>191.375</td><td>196.025</td><td>Required channel spacings are defined in Section 8</td></tr>
<tr><td>6.1.200b</td><td>Laser frequency accuracy</td><td>GHz</td><td>-1.8</td><td>1.8</td><td>Offset from channel frequency set point.
The receiver LO has the same frequency accuracy.</td></tr>
<tr><td>6.1.201a</td><td>TX spectral Upper Mask</td><td>(GHz, dB)</td><td></td><td>(59.3, 0)
(73.1, -10)
(77.5, -15)
(79.9, -20)</td><td>See definition and mask in Section 6.4.1.</td></tr>
<tr><td>6.1.201b</td><td>TX spectral Lower Mask</td><td>(GHz, dB)</td><td>(59.3, -9)
(61.9, -20)
(61.9, -35)</td><td></td><td>See definition and mask in Section 6.4.1.</td></tr>
<tr><td>6.1.210a</td><td>Laser frequency noise</td><td>Hz^2/Hz</td><td colspan="2">Maximum laser frequency noise defined by mask,
 

<table>
<tr><th>Frequency [Hz]</th><th>1-sided frequency noise power spectral density (PSD) [Hz^2/Hz]</th></tr>
<tr><td>1.0e2</td><td>1.0E+11</td></tr>
<tr><td>1.0e4</td><td>1.0E+09</td></tr>
<tr><td>1.0e6</td><td>1.0E+06</td></tr>
<tr><td>1.0e7</td><td>6.0E+05</td></tr>
<tr><td>1.0e8</td><td>1.6E+05</td></tr>
<tr><td>1.0e9</td><td>1.6E+05</td></tr>
</table>
Mask does not apply to spurs. Measurement resolution bandwidth shall be within 10^{-1} to 10^{-6} of the frequency of interest.
High frequency component of the phase noise (100 MHz and above) is consistent with a 500 kHz laser linewidth. The receiver LO has the same linewidth.</td></tr>
</table>

<table>
<tr><th>Ref</th><th>Parameter</th><th>Unit</th><th>Min</th><th>Max</th><th>Conditions/Comments</th></tr>
<tr><td>6.1.210b</td><td>Laser frequency noise - discrete tone amplitude</td><td>MHz</td><td colspan="3">Maximum laser frequency noise discrete tone amplitude defined by mask:
 

<table>
<tr><th>Frequency [Hz]</th><th>Max tone power [MHz pk-pk]</th></tr>
<tr><td>1.0e2</td><td>200</td></tr>
<tr><td>9.2e4</td><td>3.1678</td></tr>
<tr><td>3.0e6</td><td>1.46</td></tr>
<tr><td>8.0e9</td><td>1.46</td></tr>
</table>
Tone power [MHz pk-pk] is calculated as:
$TonePower_{pkpk} = 2 * \sqrt{2} * \sqrt{\int Tone}$,
where $\int Tone$ = laser frequency noise PSD (Hz^2/Hz) at tone frequency multiplied by measurement resolution bandwidth (Hz) at tone frequency.
If there is more than one discrete tone present, the relative power of each tone normalized to the mask are summed squared, and the result must be less than 1, i.e. $\sum_i \left[{TonePower_i}/{MaxTonePower_i} \right]^2 \leq 1$</td></tr>
<tr><td rowspan="2">6.1.212</td><td rowspan="2">Laser RIN</td><td rowspan="2">dB/ Hz</td><td></td><td>-145</td><td>0.2 GHz≤ f ≤10 GHz Avg</td></tr>
<tr><td></td><td>-140</td><td>0.2 GHz≤ f ≤10 GHz Peak</td></tr>
</table>

<table>
<tr><th>Ref</th><th>Parameter</th><th>Unit</th><th>Min</th><th>Max</th><th>Conditions/Comments</th></tr>
<tr><td>6.1.213a</td><td>Tx clock phase noise (PN): Maximum PN mask§</td><td>dBc / Hz</td><td colspan="3">Maximum Phase Noise ℒ(f) defined by mask,
 

<table><tr><th>Frequency [Hz]</th><th>Tx clock phase noise [dBc/Hz]</th></tr><tr><td>1.0E+04</td><td>-100</td></tr><tr><td>4.0E+05</td><td>-132</td></tr><tr><td>1.0E+06</td><td>-136</td></tr><tr><td>>1.0E+07</td><td>-146</td></tr></table>$f_c = \frac{f_{baud}}{256} = 461.73MHz$
Mask does not apply to spurs, broadband phase noise only. Spurs are considered separately per 6.1.213c.</td></tr>
<tr><td>6.1.213b</td><td>Tx clock phase noise (PN); Maximum total integrated random jitter</td><td>UIrms</td><td></td><td>0.015</td><td>Weighted integrated random jitter:
$$\boldsymbol{\sigma_{rj} = \frac{1}{2\pi f_c}\sqrt{2\int_{f_1}^{f_2} 10^{\frac{\mathcal{L}(f)+W(f)}{10}}}}$$
Where,
• $f_c = f_{baud}/256$ (461.73 MHz)
• f_1 = 10 kHz
• $f_2 = f_c/2$
• $\mathcal{L}$ (f) is the phase noise (dBc/Hz) excluding spurs
• $W(f) = \begin{cases} 20\log_{10}\frac{f}{3\times10^6} & f < 3\ MHz \\ 0 & f \geq 3\ MHz \end{cases}$</td></tr>
</table>

Ref	Parameter			Unit	Min	Max	Conditions/Comments
6.1.213c	Tx clock phase noise (PN); Maximum total periodic jitter			UI_{pp}		0.03	Weighted total periodic jitter: $$PJ_{pp} = 2\sqrt{2\sum_{i=0}^{N}\sigma_{pj,i}^2}$$ $$\sigma_{pj,i} = \frac{1}{\sqrt{2}\pi f_c}10^{\frac{s_i+W(f_i)}{20}}$$ Where, • $f_c = f_{baud}/256$ (461.73 MHz) • s_i is the spur magnitude in dBc • f_i is the spur frequency in Hz • $W(f) = \begin{cases} 20\log_{10}\frac{f}{3\times10^6} & f < 3MHz \\ 0 & f \geq 3MHz \end{cases}$ • N = number of spurs
6.1.215	Tx output power at ProgOutputPowerMax†	Tx Output Range	A	dBm	-2		Transmit output power over wavelength, temperature, and aging. ***ProgOutputPowerMax*** registers advertise linear capability in CMIS Page 04h
			B	dBm	-7		
			C ‡	dBm	-13		
6.1.218	Minimum provisionable range of transmit output power†	Tx Output Range	A	dBm	-8	-1	***ProgOutputPowerMin*** and ***ProgOutputPowerMax*** registers advertise linear capability in CMIS Page 04h. Range extension beyond one of the required minimum ranges can be optionally advertised. The transmit output power is controlled using the ***TargetOutputPowerTx*** registers in CMIS Page 12h. Assuming the absolute accuracy given by 6.1.224.
			B	dBm	-12	-6	
			C ‡	dBm	-15	-12	
6.1.220	Default transmit output power†	Tx Output Range	A	dBm	-8		The target Tx output power will be set to this value by default for each Tx Output Range.
			B	dBm	-12		
			C ‡	dBm	-15		
6.1.221	Total output power with TX disabled			dBm		-20	***OutputDisableTx*** == true
6.1.222	Total output power during wavelength switching			dBm		-20	
6.1.223	Transmit output power stability			dB	-0.5	0.5	Output short term power stability when operating at a fixed wavelength and temperature. Measurement condition: 1ms averaging time for 1 minute accumulation.

6.1.224	Transmit output power control absolute accuracy			dB	-1.0	1.0	Absolute accuracy of delivered transmit output power relative to the ***TargetOutputPowerTx*** power setting. When operating at any temperature or wavelength within the transmitters specified operating range.
6.1.230	Inband (IB) OSNR†	Tx Output	A	dB / 12.5 GHz	35		Inband OSNR is defined as the Tx signal power between the -20 dB Tx Spectral Mask frequency points, referenced to an optical noise bandwidth of 12.5 GHz (0.1nm @ 193.7 THz) at the Tx signal peak frequency.
			B	dB / 12.5 GHz	35		
			C ‡	dB / 12.5 GHz	40		
6.1.231	Out-of-band (OOB) OSNR†	Tx Output	A	dB / 12.5 GHz	20		Out-of-Band OSNR is defined as the Tx signal power between the -20 dB Tx Spectral Mask frequency points, referenced to the maximum optical noise power within any optical bandwidth of 12.5 GHz (0.1 nm @ 193.7 THz) outside of the -20 dB Tx Spectral Mask.
			B	dB / 12.5 GHz	20		
			C ‡	dB / 12.5 GHz	40		
6.1.240	Transmitter reflectance			dB		-20	Looking into the Tx
6.1.241	Transmitter back reflection tolerance			dB		-24	Light reflected relative to Tx output power back to transmitter while still meeting Tx optical performance requirements.
6.1.250	Transmitter polarization dependent power			dB		1.5	Power difference between X and Y polarization.
6.1.260	X-Y Skew			ps		5	
6.1.270a	DC I-Q offset (mean per polarization)			dB		-26	See definition and equation in 6.4.7
6.1.270b	I-Q instantaneous offset			dB		-20	Same formula definition as 6.1.270a, however, any averaging period shall be ≤ 1us to be consistent with the timescales of Rx DSP operations. Specification applies at any point in time. Allows for modulator bias controls/errors.
6.1.271	Mean I-Q amplitude imbalance			dB		1	
6.1.272	I-Q phase error			deg	-5	+5	
6.1.273	I-Q skew			ps		0.75	

§Multiple methods are available to measure the TX phase noise on the optical output. The IA does not define the method to be used.

† Three Tx Output Range specification sets are normatively defined for 800ZR modules corresponding to the three App Codes as listed in Table 21. The specific Tx output power ranges supported by a module are determined by

the ***ProgOutputPowerMin*** and ***ProgOutputPowerMax*** capability advertisement in CMIS Page 04h. A module must comply with all the specifications (6.1.215,6.1.218,6.1.220,6.1.230,6.1.231) for at least one Tx Output Range.

[‡] Tx Output Range C may require an engineered link to achieve full-fill DWDM deployments.

6.1.3 Receiver Optical Specifications

The receiver optical tolerance specifications include margin for Tx and line impairments.

Ref	Parameter	Min	Max	Unit	Conditions/Comments
6.1.300	Frequency offset between received carrier and LO	-3.6	3.6	GHz	Acquisition Range
6.1.310	Input power range	-9	0	dBm	Signal power of the channel for the OSNR tolerance defined in (6.1.330).
6.1.320	Input power damage threshold	10		dBm	Instantaneous balanced dual polarization signal
6.1.330	OSNR Tolerance		27.0	dB/ 12.5 GHz	See definition in Section 6.4.1. The OSNR tolerance is referenced to an optical bandwidth of 12.5 GHz (0.1 nm @ 193.7 THz). Measured back-to-back with short optical channel.
6.1.340	Optical return loss	20		dB	At Rx connector input.
6.1.341	Chromatic dispersion (CD) Tolerance	2400		ps/ nm	Tolerance to Chromatic Dispersion
6.1.342	CD OSNR tolerance penalty		0.5	dB	OSNR tolerance penalty over (6.1.330) due to chromatic dispersion (6.1.160).
6.1.350	PMD tolerance (DGD, SOPMD)	10		ps	Tolerance to PMD with ≤0.5 dB penalty against OSNR tolerance (6.1.330) when change in SOP is ≤1 krad/s 10 ps of average PMD (DGD, SOPMD) corresponds to: • 33 ps of DGD_{max} when SOPMD = 0 ps^2 • 272 ps^2 of SOPMD when DGD_{max} = 23.3 ps Due to the statistical nature of PMD the DGD_{max} to DGD_{mean} ratio is calculated at 3.3 (4.1×10^{-6} probability that instantaneous DGD exceeds DGD_{max}).
6.1.351	Peak PDL tolerance	3.5		dB	Tolerance to peak PDL with ≤ 1.8 dB penalty to OSNR tolerance (6.1.330) when change in SOP is ≤ 1krad/s. Test configuration: PDL emulator applied before noise loading.

Ref	Parameter	Min	Max	Unit	Conditions/Comments
6.1.352	Tolerance to change in SOP	50		krad/s	Tolerance to change in SOP with ≤ 0.5 dB additional OSNR penalty over all PMD and PDL values defined in (6.1.350) and (6.1.351).
6.1.353	Optical input power transient tolerance	±2		dB	Tolerance to change in input power with ≤0.5 dB penalty to OSNR tolerance (6.1.330). Received power during transient shall be within the range defined by input power range (6.1.310). OSNR penalty is referenced against the steady state required OSNR tolerance at the minimum power of the input transient. The 20% to 80% rise/fall times for the input power change shall be no faster than 50 microseconds, equivalent to a ≤24 mdB/microsecond maximum slew rate.
6.1.360	Adjacent Channel Crosstalk OSNR Tolerance penalty	DWDM system dependent		dB	OSNR tolerance penalty due to crosstalk interference from neighboring channels. Back-to-back through the implemented mux/demux filters. Neighboring channels conforming to the worst-case Tx spectrum of the modules deployed in the system. Neighboring channels have +X dB higher Tx output power relative to the channel under test and are at worst case frequency offsets. X is DWDM system implementation dependent.
6.1.370	Intra-Channel filtering penalty	DWDM system dependent		dB	Due to filtering effects with the implemented 150 GHz mux/demux filters.

6.2 Module Requirements TX (Informative)

Ref	Parameter	Min	Max	Unit	Conditions/Comments
6.2.100	Transmitter disable time	1	100	ms	Time from setting ***OutputDisableTx*** bit until the transmitter optical output fails below the Tx output power given by (6.1.221). The transmitter remains tuned and ready for enable. Rx shall remain locked and thus LO must remain enabled.
6.2.110	Transmitter turn-up time: Warm start		180	s	The maximum time from ***ModuleLowPwr*** to ***DataPathActivated*** state. This includes ***ModulePwrUp***, ***DPInit***, and ***DPTxTurnOn***.
6.2.111	Transmitter turn-up time: Cold start		200	s	The maximum time from de-assertion of ***ResetS*** to ***DataPathActivated*** state with ***LoPwrS*** == false. Allows an additional 20s beyond 6.2.110 for stabilization from cold.

6.2.120	Transmitter wavelength switching time		180	s	The maximum time to change transmitter wavelength including the turn-up time. The transmitter output power during tuning is given by (6.1.222).

6.3 Module Requirements RX (Informative)

Ref	Parameter	Default	Min	Max	Unit	Conditions/Comments
6.3.110	Receiver Acquisition Time			10	s	Time to fully acquire signal after ***Rx_LOS*** de-assert, with Data Path already in the ***DataPathActivated*** state. Valid 800ZR optical Rx input signal present.
6.3.111	Receiver turn-up time: Cold start			200	s	Time to fully acquire signal after module reset. Valid 800ZR optical Rx input signal present.
6.3.130	Input total power monitor accuracy		-4.0	4.0	dB	Over the Rx input power range (6.1.310)
6.3.131	Input Channel power monitor accuracy		-4.0	4.0	dB	The module reports the received Rx channel power. Conditions are the same as 6.3.130.
6.3.132	***OpticalPowerLowAlarmFlagRx*** Assert Threshold	-13			dBm	Rx Channel Power. RxLOS is independent and defined by ***RxLOSType***=011b
6.3.133	***OpticalPowerLowAlarmFlagRx*** Hysteresis	1.0			dBm	***OpticalPowerLowAlarmFlagRx*** cleared.

6.4 Optical Parameter Definitions

6.4.1 Receiver Optical Signal-to-noise Ratio Tolerance

The DUT receiver OSNR tolerance is defined as the minimum value of OSNR referenced to 12.5 GHz (0.1 nm @ 193.7 THz) that can be tolerated while maintaining the maximum post-FEC BER defined in 6.1.120. This must be met back-to-back (heterodyne) between two modules at all Rx powers given by 6.1.310. The contribution from the inband OSNR (6.1.230) of the upstream transmitter shall not be removed from the measurement.

The DUT receiver OSNR tolerance should be tested with both the DUT and the upstream transmitter operating with TX output power at the maximum output power compliant to 6.1.215.

The receiver OSNR tolerance does not have to be met in the presence of multiple channels, chromatic dispersion, non-linear effects, PMD, PDL, Mux/Demux optical filtering, or reflections from the optical path. These effects are specified separately but contribute to total optical path OSNR penalty. System integrators need to account for these path penalties when evaluating network performance.

6.4.2 Optical return loss at S_S

Reflections are caused by refractive index discontinuities along the optical path. If not controlled, they can degrade system performance through their disturbing effect on the operation of the optical source, or through multiple reflections which lead to interferometric noise at the receiver. Reflections from the optical path are controlled by specifying the:

- minimum optical return loss of the cable plant at the source reference point (S_s), including any connectors; and
- maximum discrete reflectance between source reference point (S_s) and receive reference point (R_s)

Reflectance denotes the reflection from any single discrete reflection point, whereas the optical return loss is the ratio of the incident optical power to the total returned optical power from the entire fiber including both discrete reflections and distributed backscattering such as Rayleigh scattering.

6.4.3 Discrete reflectance between S_S and R_S

Optical reflectance is defined to be the ratio of the reflected optical power present at a point, to the optical power incident to that point. The maximum number of connectors or other discrete reflection points which may be included in the optical path must be such as to allow the specified overall optical return loss to be achieved.

6.4.4 Differential Group Delay (DGD)

Differential group delay (DGD) is the time difference between the fractions of an optical signal transmitted in the two principal states of polarization. For distances greater than several kilometers, and assuming random (strong) polarization mode coupling, DGD in a fiber can be statistically modelled as having a Maxwellian distribution.

Due to the statistical nature of polarization mode dispersion (PMD), the relationship between maximum DGD (DGD_{max}) and mean DGD (DGD_{mean}) can only be defined probabilistically. The probability of the instantaneous DGD exceeding any given value of DGD_{max} can be inferred from its Maxwellian statistics.

For purposes of this IA the ratio of DGD_{max} to DGD_{mean} is defined as 3.3, corresponding to a 4.1×10^{-6} probability of the instantaneous DGD exceeding DGD_{max}.

6.4.5 Polarization Dependent Loss (PDL)

The polarization dependent loss (PDL) is the difference (in dB) between the maximum and minimum values of the channel insertion loss (or gain) of the black link from point S_s to R_s due to a variation of the State of Polarization (SOP) over all states of polarization.

6.4.6 Polarization rotation speed

The polarization rotation speed is the rate of rotation in Stokes space of the two polarizations of the optical signal at point R_s measured in krad/s.

6.4.7 I-Q offset

I-Q offset is measured separately on each polarization and is calculated using the following formula:

$$P_{excess} = \frac{\mathrm{I}^2_{\mathrm{mean}} + \mathrm{Q}^2_{\mathrm{mean}}}{\mathrm{P}_{\mathrm{Signal}}}$$

$$IQ_{\mathrm{offset}} = 10\log_{10}(\mathrm{P}_{\mathrm{excess}})$$

Instantaneous I-Q offset is measured with an averaging period ≤ 1 µs to be consistent with the timescales of receiver DSP operations.

6.4.1 Tx Spectral Masks

800ZR DWDM transmitters are on a 150 GHz grid and are therefore required to limit their spectral content by complying with minimum and maximum masks measured using an optical spectrum analyzer. The spectral masks at zero frequency shift relative to the transmitter center frequency are approximated by a root-raised- cosine (RRC) roll-off factor of 0.4 for the upper limit mask, and 0.05 for the lower limit mask.

At baseband frequency, the measured spectrum is normalized relative to the average measured power from the transmitter over a ± 20 GHz window (excluding DC frequency). Four piece-wise linear lines define the normative Upper Mask in the Figure 25, with the 3 lower points falling on a RRC curve with an 0.4 roll-off factor (shown in blue). Three piece-wise linear lines define the normative Lower Mask in the Figure 25, with the middle point falling on an RRC curve with a 0.05 roll-off factor (in green).

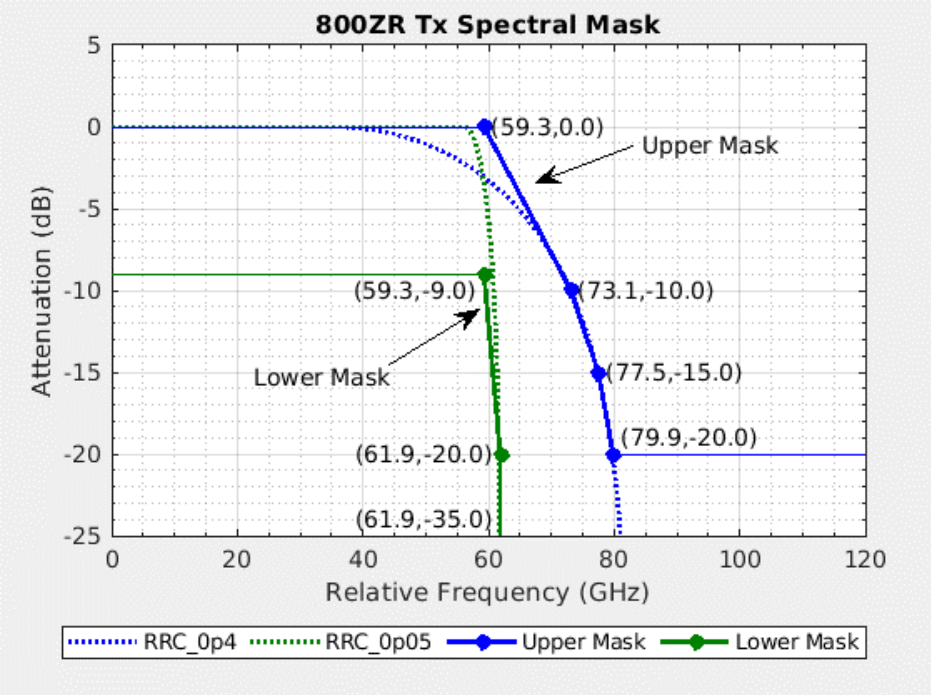


Figure 25: Transmit Spectral Masks (Min and Max)

7 Interoperability Test Methodology, Definitions

Interoperability is achievable by complying with all required aspects of this IA. Digital data path verification is achieved through a combination of interoperability test vectors and the use of common sets of test generators and checkers. The generators and checkers can be configured using loopback pairs for self-testing or in a cross-linked configuration. These test features are described in more detail in sections 7.1 and 7.2.

Recommended performance monitoring features are described in section 7.3.

Optical interworking is achieved through strict adherence to the discrete Tx/Rx optical specifications defined in Section 6.

To verify the design for interoperability, test vectors are made available to OIF member companies. Lower-level diagnostic capabilities in the form of loopbacks and insertion points for test generators/checkers is also described in Section 7.1.

Recommended diagnostic loopbacks and test patterns are described in Section 7.1 and 7.2 for compliance testing purposes. Generators and checkers can be used in conjunction with the loopbacks for self-diagnostics, or with external test equipment to verify the data path.

7.1 Loopbacks

An 800ZR module must be capable of minimally supporting one of the following loopback pairs. The CMIS supported loopback modes are shown in **bold.** Each pair has 1 Rx path and 1 Tx path.

- Modem Tx loopback + Modem Rx Loopback
- Modem Tx Loopback + **Host Side Rx Loopback**
- **Media Side Tx Loopback** + Modem Rx Loopback
- **Host Side Tx Loopback** + **Host Side Rx Loopback**

The specific loopback mode enabled must be coordinated at each end of the link by each host. The defined loopback modes are shown in Table 23 with the CMIS supported modes shown in ***italic***. The locations of each of these loopbacks is shown in Figure 26.

The recommended replacement signals for the various loopbacks are shown in Table 23. The replacement signals are optional and are intended to replace the signal downstream from the loopback location. For the Tx Loopbacks, the replacement signal would be transmitted on the media interface. For Rx Loopbacks, the replacement signal would be transmitted on the host interface.

Loopback Mode	Description	Replacement
Host Side Tx Loopback	Loopback after Alignment lock, lane de-skew, lane reordering and RS-FEC decoding → RS-FEC encoding, PMA sublayer. Host loop timed	Ethernet LF
Modem Tx Loopback	Loopback after client mapping to 800ZR frame → Client Demapping from 800ZR frame. Data re-transmitted relative to local clock	800ZR-LCK

Media Side Tx Loopback	Loopback after Tx DSP processing blocks → Before Rx DSP processing blocks	
Media Side Rx Loopback	Loopback after DSP. After polarity split, symbol de-interleave, OFEC decoding → OFEC encoding, symbols mapper/symbol interleave. Media loop timed.	LF
Modem Rx Loopback	Loopback after client demapping from 800ZR frame → Client mapping to 800ZR frame. Data retransmitted relative to local clock	LF
Host Side Rx Loopback	Loopback after distribution/interleaving block on host ingress path → Before lane alignment/reorder and deinterleave	

Table 23: 800ZR Loopbacks

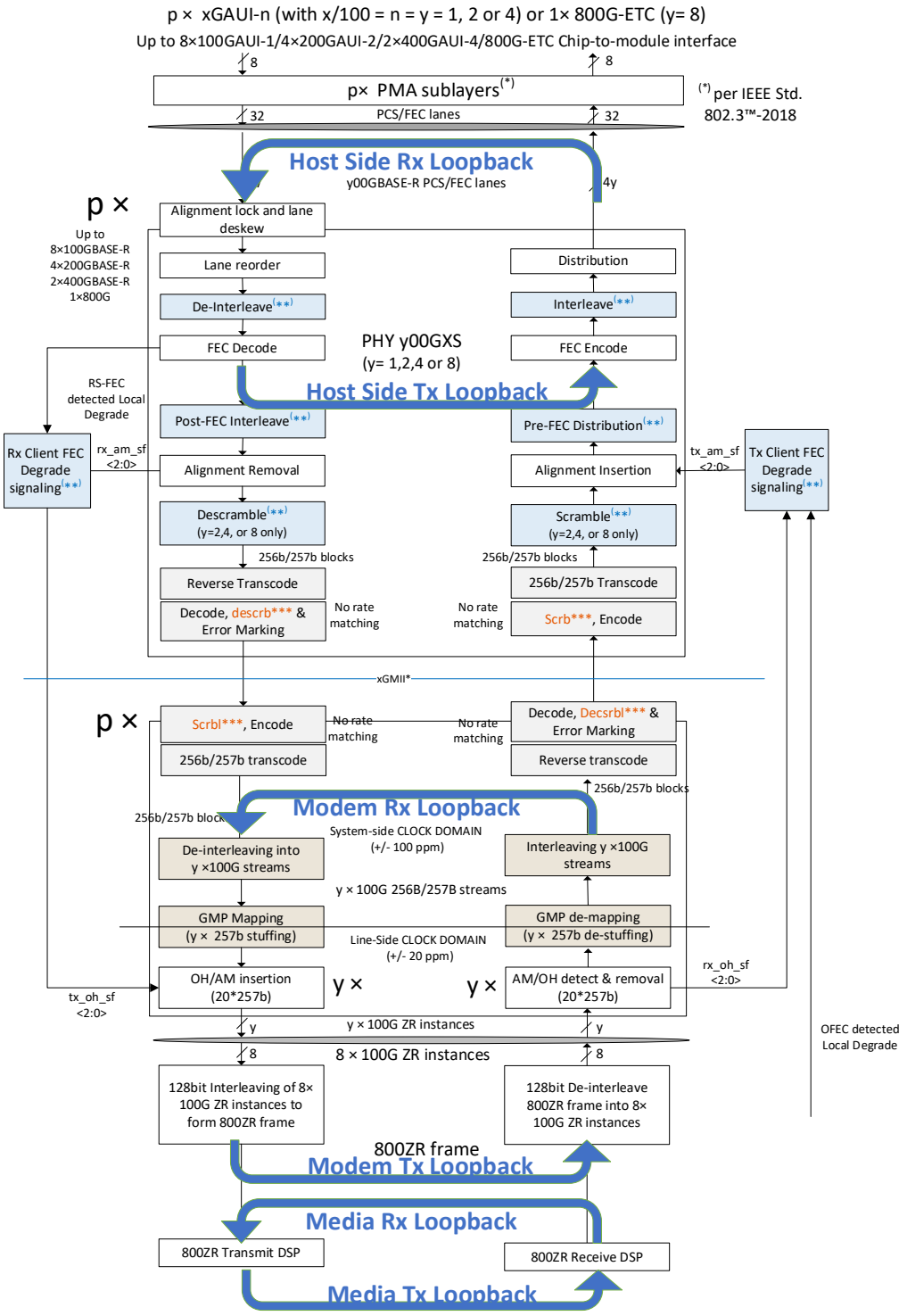


Figure 26: 800ZR Loopback Locations

7.2 Interoperability Test Vectors

The interoperability test vectors are used during design development to check for interoperability.

7.2.1 TV PRBS

The TV PRBS is used for validating the 800ZR FEC/DSP framing, symbol mapping and FAW/TS/PS insertion. The required PRBS31 is per IEEE 802.3 with initial state being all 1's and is shown in Figure 27.

- Generation/checking is to/from the media interface which is shown as the rightmost interface in Figure 14
- The PRBS test vector generator is inserted in the TX data path and replaces the entire 800ZR frame (including overhead and payload) shown as the leftmost interface of Figure 14.

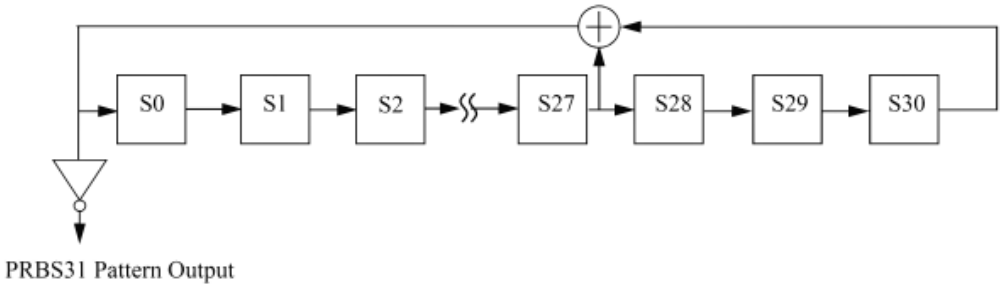


Figure 27: Test vector PRBS31 generator

The TV PRBS test vector files are listed in Table 24.

Order	Polynomial	Seed value	Test Vector File
31	$Z^{31}+Z^{28}+1$	Initialized to all 1's	https://www.oiforum.com/bin/c5i?mid=4&rid=7&gid=0&k1=52179&k2=18&k3=3

Table 24: Test vector PRBS files

7.2.2 GMP PCS test vectors

The GMP PCS test vectors are used for validating the processing of clients to form the 800ZR frame. The GMP PCS test vector files are listed in Table 25.

Description	Test Vector File
1×800GBASE-R vector	https://www.oiforum.com/bin/c5i?mid=4&rid=7&gid=0&k1=52179&k2=18&k3=3
2×400GBASE-R vector	
4×200GBASE-R vector	
8×100GBASE-R vector	

Table 25: GMP PCS test vector files

7.3 Performance Monitors for Interoperability

Table 26 provides the list of C-CMIS Rx signal quality performance monitors (PMs), the definitions for which are given in Sections 7.3.1-7.3.5. These are optional advertised PMs in C-CMIS. All the Rx signal

quality PMs (EVM_{xx}, MER, eSNR) are implementation dependent and not representative of the absolute signal quality at the optical input. They are a measure of the electrical signal quality at the decision device.

Only one of the signal quality PMs should be used for alarming purposes. eSNR is the recommended metric since it is the only signal quality PM defined in this section that is a direct, vendor- and implementation independent function of pre-FEC BER. This property also enables the definition of an SNR Margin against eSNR at the OFEC BER threshold.

C-CMIS/CMIS PM	Data Type	LSB Scaling	Unit	Range	Accuracy
EVM[1]	U16	100/65,535	%		
MER	U16	0.1	dB		
eSNR	U16	0.1	dB	12.7 to 17.2 dB	±0.1 dB
SNR Margin[2]	S16	0.1	dB	-0.5 to 4.5 dB	±0.1 dB

Table 26: C-CMIS Signal Quality Performance Monitors

[1]C-CMIS PM EVM reports EVM_{RMS}.

[2]SNR Margin is not currently a VDM observable type defined in C-CMIS/CMIS. SNR Margin is defined in OIF-C-CMIS-01.3.

Table 27 provides the list of C-CMIS/CMIS optical link performance monitors (PMs).

C-CMIS/CMIS PM	Data Type	LSB Scaling	Unit	Range	Accuracy
CD-high granularity, short link (800ZR recommended)	S16	1	ps/nm		
CD-low granularity, long link	S16	20	ps/nm		
DGD	U16	0.1	ps	0 to 50 ps	±5 ps
SOPMD	U16	0.01	ps^2		
PDL	U16	0.1	dB	0 to 5 dB	±1 dB
SOP rate of change (ROC)	U16	1	krad/s		
OSNR	U16	0.1	dB	OSNR at threshold of uncorrectable blocks (UCB) to 29 dB	Accuracy depends on Tx implementation noise

Table 27: C-CMIS Optical Link Performance Monitors

Table 28 provides the list of C-CMIS Tx/Rx signal performance monitors (PMs).

C-CMIS PM	Data Type	LSB Scaling	Unit	Range	Accuracy♦
Tx Total Power	S16	0.01	dBm	Tx Output Power Range A, B or C	±1 dB (6.1.224)
Rx Total Power	S16	0.01	dBm	-9 to 0 dBm	±4 dB (6.3.130)
Rx Channel Power	S16	0.01	dBm	-9 to 0 dBm	±4 dB (6.3.131)
Carrier Frequency Offset (CFO)	S16	1	MHz	+/-3.6 GHz (6.1.300)	

Table 28: C-CMIS Tx/Rx Performance Monitors

♦ Accuracy is valid over the specified range

Table 29 provides the list of C-CMIS modulator bias performance monitors (PMs).

C-CMIS PM	Data Type	LSB Scaling	Unit	Range	Accuracy
Modulator Bias X/I	U16	100/65,535	%		
Modulator Bias X/Q	U16	100/65,535	%		
Modulator Bias Y/I	U16	100/65,535	%		
Modulator Bias Y/Q	U16	100/65,535	%		
Modulator Bias X Phase	U16	100/65,535	%		
Modulator Bias Y Phase	U16	100/65,535	%		

Table 29: C-CMIS Modulator Bias Performance Monitors

7.3.1 EVM_{MAX}

EVM_{MAX}, is defined as a ratio of the root mean square (RMS) value of all the error vectors (averaged over N symbols) to the **maximum magnitude** of all the reference constellation points.

EVM_{MAX} can be calculated per pol P as:

$$EVM^{P}_{MAX,lin} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|S_{ref,i} - S^{P}_{meas,i}\right|^2}}{C_{MAX}}$$

Where,

$$P \in \{X, Y\}$$

$$S_{ref,k} = I_{ref,k} + Q_{ref,k} \cdot j$$

$$C_{MAX} = \max_{k}\left|S_{ref,k}\right|$$

$S_{ref,k}$ are the K complex reference constellation points used by the equalizer (MMSE, ZF, etc.) and C_{MAX} is the maximum reference constellation magnitude. $S^{P}_{meas,i}$ are the N measured constellation points at the output of the equalizer on pol P, and $S_{ref,i}$ are the N reference constellation points nearest to these measured points (assuming ML detection), i.e.

$$S_{ref,i} = \arg\max_{S_{ref,k}} p\left(S^P_{meas,i} \middle| S_{ref,k}\right) = \arg\min_{S_{ref,k}} \left|S_{ref,k} - S^P_{meas,i}\right|$$

Combining the per-pol measurements and converting the linear units to percent results in the final EVM_{MAX} metric:

$$EVM_{MAX} = \sqrt{\frac{{EVM^X_{MAX,lin}}^2 + {EVM^Y_{MAX,lin}}^2}{2}} \times 100\%$$

7.3.2 EVM_{RMS}

EVM_{RMS} is defined the same as EVM_{MAX}, except that the **RMS** value of the reference constellation point magnitudes is used for normalization instead of the maximum magnitude.

EVM_{RMS} can be calculated per polarization P as:

$$EVM^P_{RMS,lin} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|S_{ref,i} - S^P_{meas,i}\right|^2}}{C_{RMS}}$$

Where,

$$P \in \{X, Y\}$$
$$S_{ref,k} = I_{ref,k} + Q_{ref,k}.j$$
$$C_{RMS} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left|S_{ref,k}\right|^2}$$

$S_{ref,k}$ are the K complex reference constellation points used by the equalizer (MMSE, ZF, etc.) and C_{RMS} is the RMS value of the reference constellation magnitudes. $S^P_{meas,i}$ are the N measured constellation points at the output of the equalizer on polarization P, and $S_{ref,i}$ are the N reference constellation points nearest to these measured points (assuming maximum-likelihood detection).

$$S_{ref,i} = \arg\max_{S_{ref,k}} p\left(S^P_{meas,i} \middle| S_{ref,k}\right) = \arg\min_{S_{ref,k}} \left|S_{ref,k} - S^P_{meas,i}\right|^2$$

Combining the per-polarization measurements and converting the linear units to percent results in the final EVM_{RMS} metric:

$$EVM_{RMS} = \sqrt{\frac{{EVM^X_{RMS,lin}}^2 + {EVM^Y_{RMS,lin}}^2}{2}} 100\%$$

7.3.3 MER

The Modulation Error Ratio, MER, is defined as a ratio of the mean squared (MS) value of the reference constellation point magnitudes to the MS value of all the error vectors (averaged over N symbols). It is essentially equal to the squared inverse of $EVM_{RMS,lin}$. However, before converting the ratio to dB, the SNR bias introduced by scaling in the equalizer is compensated. This makes MER an accurate estimate of the SNR at the input to the equalizer.

The biased MER, $\overline{MER}^P_{lin}$, can be calculated per polarization P as:

$$\overline{MER}_{lin}^{P} = \frac{{C_{RMS}}^2}{\frac{1}{N}\sum_{i=1}^{N}\left|S_{ref,i} - S_{meas,i}^{P}\right|^2}$$

Where,

$$P \in \{X, Y\}$$
$$S_{ref,k} = I_{ref,k} + Q_{ref,k}.j$$
$$C_{RMS} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left|S_{ref,k}\right|^2}$$

All variables are as defined in Section 7.3.2.

The average biased MER over both polarizations is:

$$\overline{MER}_{lin} = \frac{\overline{MER}_{lin}^{X} + \overline{MER}_{lin}^{Y}}{2}$$

The bias introduced by equalizer scaling can be derived as follows: Let S_i be the *i*-th transmitted symbol and R_i the corresponding received symbol. The two variables are expected to be equal except for a random additive noise term N_i and an equalizer scaling factor g:

$$R_i = g.(S_i + N_i)$$

Under the assumption that both, the transmitted symbols, and the noise, are uncorrelated and have zero mean, the (unbiased) SNR of this system at the input of the equalizer is

$$SNR = \frac{\sigma_S^2}{\sigma_N^2}$$

Where σ_S^2 and σ_N^2 are the signal- and noise variance, respectively.

On the other hand, the result produced by the $\overline{MER}_{lin}$ calculation can be derived as follows: The estimated error in the *i*-th symbol, E_i, is

$$\begin{aligned} E_i &= R_i - S_i \\ &= g \cdot (S_i + N_i) - S_i \\ &= (g-1) \cdot S_i + g \cdot N_i \end{aligned}$$

and the corresponding variance (i.e., the mean squared error), ${\sigma_E}^2$:

$${\sigma_E}^2 = E\{E_i^2 - \mu_E\} = (g-1)^2{\sigma_S}^2 + g^2{\sigma_N}^2$$

Therefore

$$\overline{MER}_{lin} = \frac{{\sigma_S}^2}{{\sigma_E}^2} = \frac{{\sigma_S}^2}{(g-1)^2{\sigma_S}^2 + g^2{\sigma_N}^2}$$

Unless $g = 1$, the biased MER estimate is not equal to the SNR at the equalizer input.

$$\overline{MER}_{lin} \neq SNR$$

The required compensation is a function of the equalizer scaling factor g and depends on the h/w implementation.

For a common LMS equalizer which minimizes the mean squared error (MMSE) ${\sigma_E}^2$, the optimum gain g_{mmse} can be derived analytically:

$$g_{mmse} = \arg\min_{g}(g-1)^2{\sigma_S}^2 + g^2{\sigma_N}^2$$

The solution is:

$$g_{mmse} = \frac{{\sigma_S}^2}{{\sigma_S}^2 + {\sigma_N}^2}$$

Substituting g_{mmse} into the biased MER equation reveals the appropriate bias compensation:

$$\begin{aligned}
\overline{MER}_{lin} &= \frac{{\sigma_S}^2}{(g_{mmse}-1)^2{\sigma_S}^2 + g_{mmse}^2{\sigma_N}^2} \\
&= \frac{{\sigma_S}^2}{\left(\frac{{\sigma_S}^2}{{\sigma_S}^2+{\sigma_N}^2}-1\right)^2{\sigma_S}^2 + \left(\frac{{\sigma_S}^2}{{\sigma_S}^2+{\sigma_N}^2}\right)^2{\sigma_N}^2} \\
&= \left(\left(\frac{-{\sigma_N}^2}{{\sigma_S}^2+{\sigma_N}^2}\right)^2 + \frac{{\sigma_N}^2{\sigma_S}^2}{({\sigma_S}^2+{\sigma_N}^2)^2}\right)^{-1} \\
&= \left(\frac{{\sigma_N}^2\cdot({\sigma_S}^2+{\sigma_N}^2)}{({\sigma_S}^2+{\sigma_N}^2)^2}\right)^{-1} \\
&= \left(\frac{{\sigma_N}^2}{{\sigma_S}^2+{\sigma_N}^2}\right)^{-1} = \frac{{\sigma_S}^2+{\sigma_N}^2}{{\sigma_N}^2} = \frac{{\sigma_S}^2}{{\sigma_N}^2}+1 \\
&= SNR+1
\end{aligned}$$

Consequently, the unbiased MER, MER_{lin}, can be calculated as follows:

$$MER_{lin} = SNR = \overline{MER}_{lin} - 1$$

Unbiased MER is reported in [dB]:

$$MER = 10\cdot\log_{10}(MER_{lin})\,dB$$

NOTE 1: MER and EVM_{RMS} are related as follows:

$$MER = 10\cdot\log_{10}\left(\left(\frac{100\%}{EVM_{RMS}}\right)^2 - 1\right)$$

7.3.4 eSNR

eSNR is defined as the effective Signal to Noise ratio at the decision sampling point in dB. The method for determining the eSNR is to use an estimate of the input BER to the OFEC decoder and use the following equation to determine the eSNR,

$$eSNR_{Linear} = 10 * \left(erfcinv\left(\frac{8}{3} * ber\right)\right)^2$$

$$eSNR\ dB = 10 * log_{10}(eSNR_{Linear})$$

This formula assumes AWGN, with no other Channel, Rx or Tx impairments. Inaccuracy due to the use of the eSNR equation is ≤0.0012 dB in the SNR range of OFEC.

Recommend Range is BER 4.7e-4 to 2e-2, which is 17.2 to 12.7 dB eSNR.

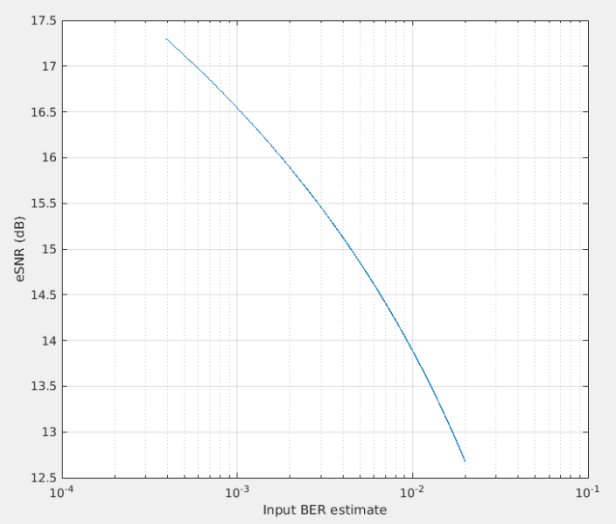


Figure 28: eSNR dB versus input BER estimate

7.3.5 SNR Margin

SNR Margin is defined as the difference between the measured eSNR, and the required eSNR (ReSNR) at the theoretical OFEC BER threshold in dB.

ReSNR is 12.71 dB assuming DP-16QAM and a theoretical OFEC BER threshold of 2%.

$$SNR\ Margin = eSNR - ReSNR = eSNR - 12.71dB$$

A positive *SNR Margin* does not necessarily guarantee post-FEC error-free operation of the modem. The OFEC BER threshold at the onset of uncorrectable blocks is implementation specific and therefore vendor dependent. The implementation specific required eSNR at onset of uncorrectable blocks ($ReSNR_{IMPL}$) can be determined for linear links by soaking the modem to find the maximum BER for post-FEC error-free operation as the theoretical BER threshold is approached.

Effective SNR Margin could be defined as the difference between the measured eSNR and $ReSNR_{IMPL}$.

$$Effective\ SNR\ margin = eSNR - ReSNR_{IMPL}$$

The *Effective SNR Margin* allows a direct comparison of the relative performance between different modems on the same link. A better performing modem will have a higher *Effective SNR Margin* given the same optical input conditions compared to a worse performing modem.

Note *Effective SNR Margin* and *SNR Margin* are only applicable near the FEC threshold where AWGN noise terms dominate.

8 Operating Frequency Channel Definitions

8.1 Normative 32 x 150 GHz DWDM Application Channels

The 800ZR application defines 32 frequencies at 150 GHz channel spacing as shown in Table 30, based on the 25GHz fixed grid setting defined in ITU-T G.694.1 Section 6 "Fixed grid nominal central frequencies for Dense WDM systems." The allowed channel frequencies (in THz) are defined by 193.1 + (*n*+3) × 0.025 where *n* is a positive or negative integer divisible by 6, including 0. For 800ZR modules, *n* = 114 to -72.

index	*n* (from ITU-T G.694.1)	freq. [THz]
1	114	196.025
2	108	195.875
…	…	…
31	-66	191.525
32	-72	191.375

Table 30: 150 GHz Channel Spacing Indices and Frequencies

9 Summary

This IA specifies the requirements for an 800ZR module that provides timing and code-word transparent transmission of 100GE/200GE/400GE/800GE Ethernet clients over a single carrier optical interface with end-to-end Ethernet bit error rates of less than 1.0E-15 guaranteed by FEC. The module's coherent optical interface uses non-differential DP-16QAM modulation and the line side frame format uses OFEC to protect the client data. This interface is designed to support up to 80-120km, amplified, point-to-point, DWDM noise limited links with 150 GHz spacing.

No restrictions are placed on the physical form factor by this IA. This 800ZR IA builds upon the work of 400ZR IA within OIF and the work of other standards bodies including IEEE 802.3 and ITU-T SG15.

References

[C-CMIS] Implementation Agreement for Coherent C-CMIS, IA # OIF-C-CMIS-01.2, March 2022

[CMIS] Common Management Interface Specification (CMIS), Rev 5.2, April 2022

[IEEE 802.3] Standard for Ethernet: IEEE Std 802.3™-2022

[IEEE 802.3ck] IEEE Std 802.3ck™-2022, IEEE Standard for Ethernet, Amendment 4: Physical Layer Specifications and Management Parameters for 100 Gb/s, 200 Gb/s, and 400 Gb/s Electrical Interfaces Based on 100 Gb/s Signaling

[IEEE P802.3df] IEEE Std 802.3df™-2024, IEEE Standard for Ethernet, Amendment 9: Media Access Control Parameters for 800Gb/s and Physical Layers and Management Parameters for 400Gb/s and 800Gb/s Operation

[ITU-T G.709] Recommendation ITU-T G.709/Y.1331 (06/2020) Amd 3 (2024), Interfaces for the optical transport network.

[ITU-T G.709.1] Recommendation ITU-T G.709.1 (2024) Ed. 3.0, Flexible OTN common elements

[ITU-T G.709.6] Recommendation ITU-T G.709.6 (2024) Ed 1.0, Flexible OTN B400G long-reach interfaces.

[ITU-T G.798] Recommendation ITU-T G.798 (09/2023) Amd 1 (2024), Characteristics of optical transport network hierarchy equipment functional blocks

[ETC-800G] Ethernet Technology Consortium 800G Specification r1.1, 8/6/2021

Annex A: GMP Parameters

This annex lists the parameters of the Generic Mapping Procedure (GMP) used to map the ethernet client data to 100G ZR instances and is an integral part of this IA.

Ref	GMP Parameter	Formula	Value	Units
f_{client}	Nominal client information bit rate	100GE clients: 100 Gbit/s × 257/256 × 16383/16384	100,384,497,642.517	bit/s
		200GE/400GE/800GE clients: y × 100 Gbit/s × 257/256 × 20479/20480	y × 100,385,723,114.014	
Δf_{client}	Client bit rate tolerance	100GE/200GE/400GE/800G-ETC clients	100	ppm
		800GE clients	50	
f_{server}	server nominal bit rate f		y × 100,622,438,327.432	bit/s
Δf_{server}	server bit rate tolerance		20	ppm
T_{server}	period of server multi-frame	B_{server}/ f_{server}	26.154	μs
B_{server}	number of bits per server multi-frame		y × 2,631,680	bits

Ref	GMP Parameter	Formula		Value	Units
O_{server}	number of overhead bits per server multi-frame			y × 5,140	bits
P_{server}	maximum number of bits in server payload area	B_{server}- O_{server}		y × 2,626,540	bits
$f_{p,server}$	nominal server payload bit rate	f_{server}× P_{server}/ B_{server}		y × 100,425,910,127.574	bit/s
m	GMP data/stuff granularity			y × 257	bits
M	m and n ratio	m/n		32	
$P_{m,server}$	maximum number of (m bit) data entities in the server payload area	P_{server}/m		10220	y × 257b blocks
C_m	number of client m-bit data entities per server multi-frame				
$c_{m,nom}$	c_m value at nominal client and server bit rates	(f_{client}/$f_{p,server}$)× $P_{m,server}$	100GE clients	10215.785	
			200GE/400GE/800GE clients	10215.910	
$c_{m,min}$			100GE clients	10,214.559	

Ref	GMP Parameter	Formula		Value	Units
	c_m value at minimum client and maximum server bit rates	$(1-\Delta f_{client})/(1+\Delta f_{server}) \times c_{m,nom}$	200GE/400GE/800G-ETC clients	10,214.684	y × 257b blocks
			800GE clients	10,215.195	
$c_{m,max}$	c_m value at maximum client and minimum server bit rates	$(1+\Delta f_{client})/(1-\Delta f_{server}) \times c_{m,nom}$	100GE clients	10,217.011	
			200GE/400GE/800G-ETC clients	10,217.136	
			800GE clients	10,216.625	
$C_{m,min}$	rounded down value of $c_{m,min}$	$\lfloor c_{m,min} \rfloor$	100GE/200GE/400GE/800G-ETC clients	10214	
			800GE clients	10215	
$C_{m,max}$	rounded up value of $c_{m,max}$	$\lceil c_{m,max} \rceil$	100GE/200GE/400GE/800G-ETC clients	10218	
			800GE clients	10217	
n	GMP justification accuracy, n bit data entity			y × 8.03125	bits
$P_{n,server}$	maximum number of (n bits) data entities in the server payload area	P_{server}/n		327040.000	y × 8.03125b blocks
C_n	number of client n-bit data entities per server multi-frame				
$c_{n,nom}$	c_n value at nominal client and	$(f_{client}/f_{p,server}) \times P_{n,server}$	100GE clients	326,905.139	
			200GE/400GE/800GE clients	326,909.130	

Ref	GMP Parameter	Formula		Value	Units
	server bit rates				
$c_{n,min}$	c_n value at minimum client and maximum server bit rates	$(1-\Delta f_{client})/(1+\Delta f_{server}) \times c_{n,nom}$	100GE clients	326,865.911	y × 8.03125 blocks
			200GE/400GE/800G-ETC clients	326,869.902	
			800GE clients	326,886.247	
$c_{n,max}$	c_n value at maximum client and minimum server bit rates	$(1+\Delta f_{client})/(1-\Delta f_{server}) \times c_{n,nom}$	100GE clients	326,944.368	
			200GE/400GE/800G-ETC clients	326,948.360	
			800GE clients	326,932.014	
c_{nD}	remainder of C_n and C_m	$c_n-((m/n) \times c_m)$		Variable	n-bit
C_{nD}	integer value of c_{nD}	$C_n-((m/n) \times C_m)$		Variable	
ΣC_{nD}	accumulated value of C_{nD}			0-31	

Table 31: GMP parameter values

Where,

- Client information rate is the y00GBASE-R rate (with y=1,2,4 or 8) after RS(544,514) FEC and AM removal with f_{client} nominal bit rate and Δf_{client} bit rate tolerance
- Server is the aligned 4-frame multi-frames (both payload and overhead) of y 100G ZR instances with f_{server} nominal bit rate, Δf_{server} bit rate tolerance and B_{server} number of bits per server 4-frame multi-frame.
- Server payload is the aligned 4-frame multi-frame payloads (before AM/PAD/OH insert) of y 100G ZR instances (y=1,2,4 or 8) with $f_{p,server}$ nominal bit rate, Δf_{server} bit rate tolerance and P_{server} number of bits per server 4-frame multi-frame payload area.
- The maximum number of m (=y × 257) bit GMP data entities per y × 4-frame multi-frame payload is $P_{m,server}$ (=10220).
- For 800ZR, we use n=[m/32] = [y × 257-bit]/32 = [y × 8.03125] UI that is used as a phase unit "n-bit equivalent" for c_n parameter. c_n indicates the number "n-bit equivalent" of the 100G client stream per 100G 4-frame multi-frame server payload. It can be used as a finer phase indicator to encode the client clock at the GMP mapper.
- So, $c_{n,nom} = 32*c_{m,nom}$; $c_{n.min}=32*c_{m,min}$; $c_{n,max}=32*c_{m,max}$
- $C_m = P_{m,server} \times$[client_bit_rate/Server_Payload_bit_rate]

- C_m is an integer value indicating the number of m-bit client blocks carried in the y 100G 4-frame server multi-frame payload = int($P_{m,server}$ × [client_bit_rate/Server_Payload_bit_rate]). This value must be duplicated and have the same value in every 100G ZR instance carrying 100G streams from the same client.
- $C_m \leq P_{m,server}$ and is a value varying between $C_{m,min}$ and $C_{m.max}$ for the given client and payload type, due to client and payload bit rate tolerance range (±100ppm or ±50ppm for client and ±20ppm for line).
- ⌊⬚⌋ and ⌈⬚⌉ denote the floor and ceiling operators respectively.

Appendix A: Glossary

Abbreviation	Definition
AM	Alignment Marker or Alignment Mechanism
AWGN	Additive White Gaussian Noise
BCH	Bose-Chaudhari-Hocquenghem
BER	Bit Error Ratio
BOH	Basic Overhead Field
BOL	Beginning Of Life
CD	Chromatic Dispersion
C-CMIS	Coherent Common Management Interface Specification
CRC	Cyclic Redundancy Check
DGD	Differential Group Delay
DP-mQAM	Dual Polarization – m state Quadrature Amplitude Modulation
DSP	Digital Signal Processing
DUT	Device Under Test
DWDM	Dense Wavelength Division Multiplexing
EOL	End of Life
eSNR	Effective Signal-to-Noise Ratio
EVM	Error Vector Magnitude
FAW	Frame Alignment Word
FEC	Forward Error Correction
FFS	For Future Study
GMP	Generic Mapping Procedure
IA	Implementation Agreement
LD	Local Degrade
LO	Local Oscillator
LOS	Loss of Signal
MER	Modulation Error Ratio
MFAS	Multi-frame Alignment Signal
NA	Not Applicable
OFEC	Open FEC

Abbreviation	Definition
OSNR	Optical Signal-to-Noise Ratio
PDL	Polarization Dependent Loss
PM	Performance Monitor
PMD	Polarization Mode Dispersion
QAM	Quadrature Amplitude Modulation
R_s	Single-Channel Reference point at the DWDM network element tributary output
ReSNR	Required Effective SNR
RMS	Root Mean Square
S_s	Single-Channel Reference point at the DWDM network element tributary input
SNR	Signal-to-Noise Ratio
SOP	State of Polarization
SOPMD	Second Order Polarization Mode Dispersion
TPID	Tributary Port Identifier

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

Accelight Technologies, Inc.
Accton Technology Corporation
Adtran Networks SE
Advanced Fiber Resources (AFR)
Advanced Micro Devices, Inc.
AIO Core Co., Ltd
Alibaba
Alphawave Semi
Amazon
Amphenol Corp.
Anritsu
Applied Optoelectronics, Inc.
Arista Networks
Astera Labs
Ayar Labs
BitifEye Digital Test Solutions GmbH
BizLink Technology, Inc.
Broadcom Inc.
Cadence Design Systems
Casela Technologies USA
Celestica
China Information Communication Technologies Group
China Telecom
Ciena Corporation
Cisco Systems
Coherent
ColorChip LTD
Cornelis Networks, Inc.
Corning
Credo Semiconductor (HK) LTD
Dai Nippon Printing Co., Ltd.
Dell, Inc.
Dexerials Corporation
DustPhotonics
EFFECT Photonics B.V.

Eoptolink Technology
Epson Electronics America, Inc.
Ericsson
EXFO
Foxconn Interconnect Technology Ltd
Fujikura
Fujitsu
Furukawa Electric Co., Ltd.
Global Foundries
Google
H3C Technologies Co., Ltd.
Hakusan Inc
Hewlett Packard Enterprise (HPE)
HGGenuine Optics Tech Company
Hirose Electric Co. Ltd.
Hisense Broadband Multimedia Technologies Co., LTD
Huawei Technologies Co., Ltd.
Infinera Corporation
InfiniLink
InnoLight Technology Limited
Integrated Device Technology
Intel
Juniper Networks
Kandou Bus
KDDI Research, Inc.
Keysight Technologies, Inc.
KYOCERA Corporation
Lessengers Inc.
Lightmatter
Linktel Technologies Co., Ltd.
Lumentum
Lumiphase AG
LUXIC Technology Co
Luxshare Technologies International, Inc.
MACOM Technology Solutions
Marvell Semiconductor, Inc.
MaxLinear Inc.
MediaTek

Meta Platforms
Microchip Technology Incorporated
Microsoft Corporation
Mitsubishi Electric US, Inc.
Molex
Multilane Inc.
NEC Corporation
Nokia
NTT Corporation
Nubis Communications, Inc.
NVIDIA
O-Net Technologies (Shenzhen) Group Co., Limited
Omniva LLC
Optomind Inc.
Orange
PETRA
Point2 Technology
Precision Optical Technologies
Quantifi Photonics USA Inc.
Quintessent Inc.
Ranovus
Retym
Rosenberger Hochfrequenztechnik GmbH & Co. KG
Samsung Electronics Co. Ltd.
Samtec Inc.
SCINTIL Photonics
Semtech Canada Corporation
Senko Advanced Components
SeriaLink Systems Ltd.
Sicoya GmbH
SiFotonics Technologies Inc.
Silith Technology
Socionext Inc.
Source Photonics, Inc.
Spirent Communications
Sumitomo Electric Industries, Ltd.
Sumitomo Osaka Cement
Synopsys, Inc.

TE Connectivity
Tektronix
Telefonica S.A.
TELUS Communications, Inc.
Teramount
TeraSignal, LLC.
US Conec
Viavi Solutions Deutschland GmbH
Wilder Technologies, LLC
Wistron Corporation
Xphor Ltd.
Yamaichi Electronics Ltd.
ZTE Corporation

1

2

3