802.3df D1.1 Comment Resolution

P802.3df editorial team

Introduction

- This slide package is put together by the 802.3df editorial team to provide background and detailed resolutions to aid in comment resolution.

Clause 167

Fiber optics cabling
Comments 13, 14, 116 (part 1)

- Three comments were received highlighting issues with the clause 167 “Characteristics of the fiber optic cabling and MDI” PICS table
  - Comments 13 , 14, and 116

Cl **167** *SC* **167.11.4.6** *P* **158** *L* **31** # 13

Ran, Adee — Cisco

Comment Type **T** *Comment Status* **X**

The status of items OC15 through OC20 includes "AFI:", which makes them conditional on an angled fiber interface. However, the reference 167.10.3.4 also specifies flat fiber interfaces.

The value/comment needs to be different for angled and flat.

SuggestedRemedy

Add or change PICS items for 167.10.3.4 as appropriate.

Proposed Response *Response Status* **O**

Cl **167** *SC* **167.11.4.6** *P* **158** *L* **37** # 14

Ran, Adee — Cisco

Comment Type **T** *Comment Status* **X**

The value/comment for OC18 includes "or per ANSI/TIA-604-18-A designation FOCIS 18 A-1-0 or FOCIS 18 R-1x16-1-0-1-2-0".

These do not appear in the referenced subclause 167.10.3.4.

Also in OC19.

SuggestedRemedy

Align the value/comment and the subclause text.

Proposed Response *Response Status* **O**

Cl **167** *SC* **167.11.4.6** *P* **158** *L* **13** # 116

Dawe, Piers — Nvidia

Comment Type **E** *Comment Status* **X**

These PICS need work to align them to the clause

SuggestedRemedy

Removing Option A will make this task simpler

Proposed Response *Response Status* **O**

Fiber optics cabling
Comments 13, 14, 116 (part 2)

D1.1 167.11.4.6 PICS Table

Item	Feature	Subclause	Value/Comment	Status	Support
…					
OC5	MDI layout for 400GBASE-VR4 and 400GBASE-SR4	167.10.3.1 a	Optical lane assignments per Figure 167–8	(VR4 or SR4):M	Yes [ ] N/A [ ]
OC5a	MDI layout for 800GBASE-VR8 and 800GBASE-SR8 option A	167.10.3.1 a	Optical lane assignments per Figure 167–8a	(VR8 or SR8):M	Yes [ ] N/A [ ]
OC5b	MDI layout for 800GBASE-VR8 and 800GBASE-SR8 option B	167.10.3.1 a	Optical lane assignments per Figure 167–8a	(VR8 or SR8):M	Yes [ ] N/A [ ]
OC6	MDI mating, 100GBASE-VR1 and 100GBASE-SR1, with duplex optical fiber connector	167.10.3.2	MDI optically mates with plug on the cabling, performance grade Bm/2m	(VR1 or SR1):M	Yes [ ] N/A [ ]
…					
OC15	MDI requirements, with multifiber connector	167.10.3.3	Per IEC 63267-1, performance grade Bm/1m	INS and (VR1, SR1, VR2, SR2, VR4, or SR4):AFI:M	Yes [ ] N/A [ ]
OC16	MDI mating, with multifiber connector	167.10.3.4	MDI optically mates with plug on the cabling, performance grade Bm/2m	(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC17	MDI dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-1 interface 7-1-3 or interface 7-1-10	(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC18	MDI dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-2 interfaces 7-2-3 or 7-2-10, or per ANSI/TIA-604-18-A designation FOCIS 18 A-1-0 or FOCIS 18 R-1x16-1-0-1-2-0	(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC19	Cabling connector dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-2 interface 7-2-4 or ANSI/TIA-604-18-A designation FOCIS 18 P-1x16-1-0-2-2-0	INS*(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC20	MDI requirements, with multifiber connector	167.10.3.4	Per IEC 61753-1 and IEC 61753-022-2, performance grade Bm/2m	INS*(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]

Fiber optics cabling
Comments 13, 14, 116 (part 3)

Proposed D1.2 167.11.4.6 PICS Table

Item	Feature	Subclause	Value/Comment	Status	Support
…					
OC5	MDI layout for 400GBASE-VR4 and 400GBASE-SR4	167.10.3.1	Optical lane assignments per Figure 167–8	(VR4 or SR4):M	Yes [ ] N/A [ ]
OC5a	MDI layout for 800GBASE-VR8 and 800GBASE-SR8 option A	167.10.3.1a	Optical lane assignments per Figure 167–8a	(VR8 or SR8):M	Yes [ ] N/A [ ]
OC5b	MDI layout for 800GBASE-VR8 and 800GBASE-SR8 option B	167.10.3.1a	Optical lane assignments per Figure 167–8a	(VR8 or SR8):M	Yes [ ] N/A [ ]
OC6	MDI mating, 100GBASE-VR1 and 100GBASE-SR1, with duplex optical fiber connector	167.10.3.2	MDI optically mates with plug on the cabling, performance grade Bm/2m	(VR1 or SR1):M	Yes [ ] N/A [ ]
…					
OC15	MDI requirements, with multifiber connector	167.10.3.3	Per IEC 63267-1, performance grade Bm/1m	INS and (VR1, SR1, VR2, SR2, VR4, or SR4):AFI:M	Yes [ ] N/A [ ]
OC16	MDI mating, with multifiber connector	167.10.3.4	MDI optically mates with plug on the cabling, performance grade Bm/2m	(VR8 or SR8):!AFI:M	Yes [ ] N/A [ ]
OC17	MDI mating, with multifiber connector	167.10.3.4	MDI optically mates with plug on the cabling, performance grade Bm/1m	(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC18	MDI dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-2 interface 7-2-3 or interface 7-2-10	(VR8 or SR8):!AFI:M	Yes [ ] N/A [ ]
OC19	MDI dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-4 interface 7-4-7 or interface 7-4-9	(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC20	Cabling connector dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-2 interface 7-2-4	INS*(VR8 or SR8):!AFI:M	Yes [ ] N/A [ ]

Item	Feature	Subclause	Value/Comment	Status	Support
OC21	Cabling connector dimensions, with multifiber connector	167.10.3.4	Per IEC 61754-7-4 interface 7-4-1	INS*(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]
OC22	MDI requirements, with multifiber connector	167.10.3.4	Per IEC 61753-1 and IEC 61753-022-2, performance grade Bm/2m	INS*(VR8 or SR8):!AFI:M	Yes [ ] N/A [ ]
OC23	MDI requirements, with multifiber connector	167.10.3.4	Per IEC 63267-1, performance grade Bm/1m	INS*(VR8 or SR8):AFI:M	Yes [ ] N/A [ ]

Fiber optics cabling
Comments 13, 14, 116 (part 4)

Proposed D1.2 PICS Table Updates

-
- The value/comments of OC16 through OC23 now align with the updated text in 167.10.3.4
- The use of !AFI for flat connectors and AFI for angled connectors is consistent with 802.3db-2022

OC14	MDI requirements, with multifiber connector	167.10.3.3	Per IEC 61753-1 and IEC 61753-022-2, performance grade Bm/2m	INS and (VR1, SR1, VR2, SR2, VR4, or SR4):!AFI:M	Yes [ ] N/A [ ]
OC15	MDI requirements, with multifiber connector	167.10.3.3	Per IEC 63267-1, performance grade Bm/1m	INS and (VR1, SR1, VR2, SR2, VR4, or SR4):AFI:M	Yes [ ] N/A [ ]

- The resolution of comment #115 could affect the final PICS table

Clause 173

PCSL Grouping Comment #84

Cl **173**	*SC* **173.1.3**	*P* **212**	*L* **51**	# 84

Dawe, Piers — Nvidia

Comment Type **T** *Comment Status* **D** *(bucket1)*

Adapt the PCSL (PCS lane) formatted signal to the appropriate number of abstract or physical lanes

SuggestedRemedy

Adapt the PCSL (PCS lane) formatted signal to the appropriate number and grouping of abstract or physical lanes

Proposed Response *Response Status* **W**

PROPOSED ACCEPT IN PRINCIPLE.
The constrained grouping of lanes is part of the "adapt" process and does not need to be listed as a detail here. Instead, this detail is specified in 173.4. The proposed change is not necessary.
However, the acronym PCSL is not properly introduced in this clause.
Change "PCSL (PCS lane)" to "PCS lane (PCSL)".

Figure 173–1—800GBASE-R PMA relationship to the ISO/IEC Open Systems Interconnection (OSI) reference model and IEEE 802.3 Ethernet model

173.1.3 Summary of functions

The following is a summary of the principal functions implemented (when required) by the PMA in both the transmit and receive directions:

- Adapt the PCSL (PCS lane) formatted signal to the appropriate number of abstract or physical lanes
- Provide per input-lane clock and data recovery
- Provide bit-level multiplexing

- Provide clock generation
- Provide signal drivers
- Optionally provide local loopback to/from the PMA service interface
- Optionally provide remote loopback to/from the PMD service interface
- Optionally provide test-pattern generation and detection
- Tolerate Skew Variation
- Perform PAM4 encoding and decoding
- Provide receive link status information in the receive direction

173.1.4 PMA sublayer positioning

An implementation may use one or more PMA sublayers to adapt the number and rate of the PCSLs to the number and rate of the PMD lanes. The number of PMA sublayers required depends on the partitioning of functionality for a particular implementation.

Figure 173–2 shows examples of the PMA sublayer positioning for implementations with an 800GMII

PCSL Grouping Comment #87

Cl **173**	*SC* **173.4**	*P* **217**	*L* **6**	# 87

Dawe, Piers — Nvidia

Comment Type **T** *Comment Status* **D** *(bucket1)*

PMA:IS_UNITDATA_0:31.request would be better shown as PMA:IS_UNITDATA_0:15.request and PMA:IS_UNITDATA_16:31.request as in Figure 172-2. The PMA doesn't really know lane numbers, it doesn't read alignment markers, but it needs to know the two groups to apply the restricted bit muxing rules.
The output lanes can stay as one group.

SuggestedRemedy

Show two groups of 16 input lanes, PMA:IS_UNITDATA_0:15.request and PMA:IS_UNITDATA_16:31.request.
Similarly for the 32 PHY_XS:IS_UNITDATA_0:31.indication lanes in Figure 173-4, 8:32 PMA functional block diagram.

Proposed Response *Response Status* **W**

PROPOSED REJECT.
There are 32 PCS lanes represented by PMA:IS_UNITDATA_0:31. Figure 172-2 shows the two groups, one from 0:15 and the other from 16:31, to show how the lanes from each flow map to the set of 32 PCS lanes. Showing the separation of the two groups of lanes in this PMA diagram is not helpful. Since the PMA is connected directly to the PCS (colocated), the lane numbers are known by the PMA.

Figure 173-3
(PMA Clause)

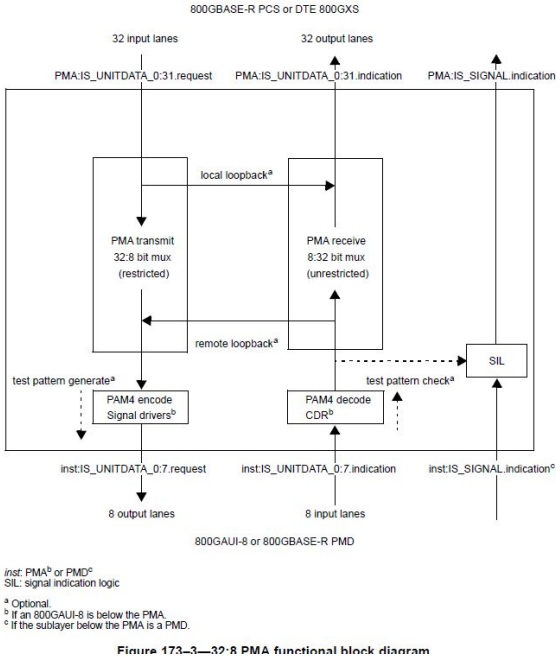


Figure 173–3—32:8 PMA functional block diagram

Figure 172-2
(PCS Clause)

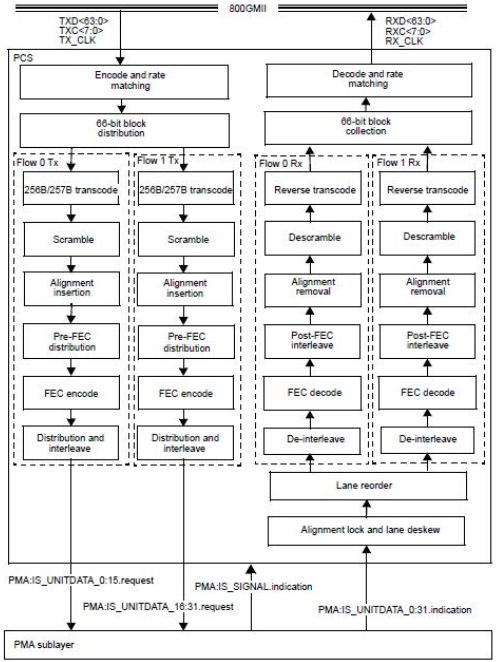


Figure 172–2—Functional block diagram

Link Status Comment #85

Cl 173 SC 173.1.3 P 213 L 10 # 85

Dawe, Piers Nvidia

Comment Type T *Comment Status* D *(bucket1)*

In common cases (800GAUI-8) receive link status information may be used but isn't forwarded.
"Provide receive link status information in the receive direction": do we need another bullet, that when connected to a PHY XS, it provides link status information in the transmit (egress) direction?

SuggestedRemedy

Per comment

Proposed Response *Response Status* W

PROPOSED REJECT.
The opening sentence in 173.1.3 states "The following is a summary of the principal functions implemented (when required) by the PMA in both the transmit and receive directions:" The phrase "when required" implies that some of the funcitons listed are conditional upon the PMA type. The requirement for each of the functions listed is specified per PMA type in 173.4.

Section 171.1.3

173.1.3 Summary of functions

The following is a summary of the principal functions implemented (when required) by the PMA in both the transmit and receive directions:

— Adapt the PCSL (PCS lane) formatted signal to the appropriate number of abstract or physical lanes
— Provide per input-lane clock and data recovery
— Provide bit-level multiplexing

— Provide clock generation
— Provide signal drivers
— Optionally provide local loopback to/from the PMA service interface
— Optionally provide remote loopback to/from the PMD service interface
— Optionally provide test-pattern generation and detection
— Tolerate Skew Variation
— Perform PAM4 encoding and decoding
— Provide receive link status information in the receive direction

Clause 172

Test pattern control Comment #79 (part 1)

Cl **172** *SC* **172.2.4.9** *P* **202** *L* **52** # **79**

Dawe, Piers — Nvidia

Comment Type **T** *Comment Status* **D** *test pattern*

This mentions the test-pattern control register (bit 3.42.3). But does 3.42.7 Scrambled idle test-pattern apply also?

SuggestedRemedy

Please clarify, and please refer to 172.3.1 PCS MDIO function mapping

Proposed Response *Response Status* **W**

PROPOSED REJECT.

!! pulled from bucket #1 !!

The pattern selection bits were implemented for lower rate PCS specifications (e.g., 10GBASE-R) where the PCS supported more than one pattern type. For the 100GBASE-R, 200GBASE-R, 400GBASE-R, and now 800GBASE-R PCS, only one pattern is supported, so a separate bit to select a pattern type is not required. The bit 3.42.7 defined in 45.2.3.19.1 is not specified for use with any PCS in the base standard. The scrambled idle pattern is therefore enabled or disabled using bit 3.42.3 only.

172.2.4.9 Test-pattern generators

The PCS shall have the ability to generate a scrambled idle test pattern, which is suitable for receiver tests and for certain transmitter tests. When a scrambled idle pattern is enabled, the test pattern is generated by the PCS. The scrambled idle test pattern is the output of the PCS when the input to the PCS at the 800GMII is a control block with all idle characters.

If a Clause 45 MDIO is implemented, then control of the test-pattern generation is from the BASE-R PCS test-pattern control register (bit 3.42.3).

45.2.3.19 BASE-R PCS test-pattern control register (Register 3.42)

The assignment of bits in the BASE-R PCS test-pattern control register is shown in Table 45–248. This register is only required when the BASE-R capability is supported. If both BASE-R and 10GBASE-W capability is supported, then this register may either ignore writes and return zeros for reads when in 10GBASE-W mode or may function as defined for BASE-R. PRBS9, PRBS31, pseudo random, and square wave test patterns are defined for 10GBASE-R PCS only. Scrambled idle test patterns are defined for 25/40/50/100/200/400GBASE-R PCS only. The test-pattern methodology is described in 49.2.8 and 82.2.11.

Table 45–248—BASE-R PCS test-pattern control register bit definitions

Bit(s)	Name	Description	R/W[a]
3.42.15:8	Reserved	Value always 0	RO
3.42.7	Scrambled idle test-pattern enable	1 = Enable scrambled idle test-pattern mode 0 = Disable scrambled idle test-pattern mode	R/W
3.42.6	Single Lane PHY BASE-R PRBS9 transmit test-pattern enable	1 = Enable PRBS9 test-pattern mode on the transmit path 0 = Disable PRBS9 test-pattern mode on the transmit path	R/W
3.42.5	Single Lane PHY BASE-R PRBS31 receive test-pattern enable	1 = Enable PRBS31 test-pattern mode on the receive path 0 = Disable PRBS31 test-pattern mode on the receive path	R/W
3.42.4	Single Lane PHY BASE-R PRBS31 transmit test-pattern enable	1 = Enable PRBS31 test-pattern mode on the transmit path 0 = Disable PRBS31 test-pattern mode on the transmit path	R/W
3.42.3	Transmit test-pattern enable	1 = Enable transmit test pattern 0 = Disable transmit test pattern	R/W
3.42.2	Receive test-pattern enable	1 = Enable receive test-pattern testing 0 = Disable receive test-pattern testing	R/W
3.42.1	Test-pattern select	1 = Square wave test pattern 0 = Pseudo random test pattern	R/W
3.42.0	Data pattern select	1 = Zeros data pattern 0 = LF data pattern	R/W

[a]RO = Read only, R/W = Read/Write

Test pattern control
Comment #79 (part 2)

Proposed changes...

In 45.2.3.19, with appropriate editing instructions amend the last sentence in the first paragraph to:
"Scrambled idle test patterns are defined for 25/40/50/100/200/400/800GBASE-R PCS only. The test-pattern methodology is described in 49.2.8, 82.2.11, and 172.2.4.9."

Change 172.2.4.9 as follows:

172.2.4.9 Test-pattern generator~~s~~

The PCS shall have the ability to generate a scrambled idle test pattern, which is suitable for receiver tests and for certain transmitter tests. When test pattern generation ~~a scrambled idle pattern~~ is enabled (tx_test_mode is 1), the scrambled idle test pattern is generated by the PCS. The scrambled idle test pattern is the output of the PCS when the input to the PCS at the 800GMII is a control block with all idle characters.
If a Clause 45 MDIO is implemented, then ~~control of the test-pattern generation is from the BASE-R PCS test-pattern control register (bit 3.42.3)~~ the tx_test_mode variable is accessible through the register as shown in Table 172-5.

Stateless encoder / decoder Comments #20, 24

Existing tables

Suggested remedy

CI 172 SC 172.2.4.1.1 P 198 L 40 # 20

Ran, Adee Cisco

Comment Type TR Comment Status D stateless enc-dec

Table 172-1 column "T_TYPE (tx_raw_i-1)" has cells with the strings "C + T" and "S + D". These seem to be based on the state diagram convention that "+" is a logical-OR, but this is not a state diagram, and the letters are not conditions, so it isn't very clear. Using "or" would be preferable (as in the similar Table 172–4).

In addition, for each of these two strings there are two rows with two values in "T_TYPE (tx_raw_i)" column; these can be merged with the word "or" as well.

SuggestedRemedy

Merge rows 2 and 5 to a single row with columns:
"0 | C or T | C or S | ENCODE (tx_raw_i)".
Merge rows 3 and 4 to a single row with columns:
"0 | S or D | D or T | ENCODE (tx_raw_i)".

Table 172–1—PCS stateless encoder rules

reset	T_TYPE (tx_raw_{i-1})[a]	T_TYPE (tx_raw_i)[b]	Resulting tx_coded
1	any block type	any block type	LBLOCK_T
0	C + T	S	ENCODE (tx_raw_i)
0	S + D	D	ENCODE (tx_raw_i)
0	S + D	T	ENCODE (tx_raw_i)
0	C + T	C	ENCODE (tx_raw_i)
0	any combination not listed above		EBLOCK_T

Table 172–1—PCS stateless encoder rules

reset	T_TYPE (tx_raw_{i-1})[a]	T_TYPE (tx_raw_i)[b]	Resulting tx_coded
1	any block type	any block type	LBLOCK_T
0	C or T	S or C	ENCODE (tx_raw_i)
0	S or D	D or T	ENCODE (tx_raw_i)
0	any combination not listed above		EBLOCK_T

CI 172 SC 172.2.5.8.1 P 204 L 23 # 24

Ran, Adee Cisco

Comment Type TR Comment Status D stateless enc-dec

In Table 172-4, row 3, column "R_TYPE (rx_coded_i)", the value is "S or D or T or C".

The possible R_TYPE values (based on 119.2.6.2.3) are C, LI, S, T, D, and E; LI is not valid for clause 172 (per 172.2.3, EEE and low power idle are not supported). Therefore, "S or D or T or C" is equivalent to "not E". This excludes only the combination "E | E".

However, the combination "E | E" matches the second row, and therefore results in the same rx_raw, EBLOCK_R. So having R_TYPE(rx_coded_i-1)=E with any value of R_TYPE(rx_coded_i) would result in EBLOCK_R.

This means the table can be simplified and made more readable.

SuggestedRemedy

Change the third row to the following contents:
"0 | E | any block type | EBLOCK_R".

Table 172–4—PCS stateless decoder rules

reset	R_TYPE (rx_coded_{i-1})[a]	R_TYPE (rx_coded_i)[b]	Resulting rx_raw
1	any block type	any block type	LBLOCK_R
0	any block type	E	EBLOCK_R
0	E	S or D or T or C	EBLOCK_R
0	any combination not listed above		DECODE (rx_coded_i)

Table 172–4—PCS stateless decoder rules

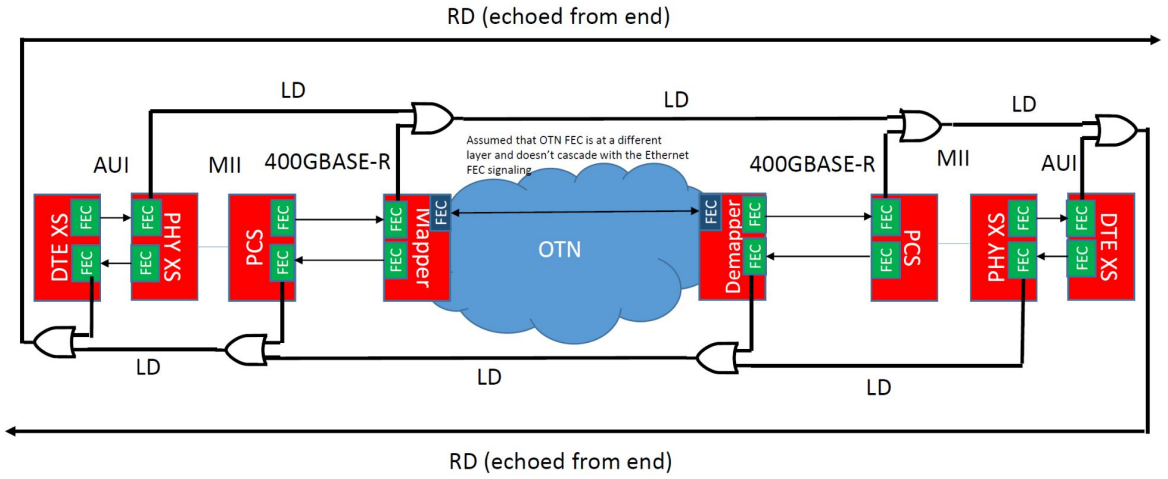
reset	R_TYPE (rx_coded_{i-1})[a]	R_TYPE (rx_coded_i)[b]	Resulting rx_raw
1	any block type	any block type	LBLOCK_R
0	any block type	E	EBLOCK_R
0	E	any block type	EBLOCK_R
0	any combination not listed above		DECODE (rx_coded_i)

Cross-Clause

FEC Degrade
Comments 17, [55, 56], 59

Consensus View on how FEC degrade signaling should work

OTN mapper contribution to FEC degrade signaling is for ITU to decide, but LD, RD must be propagated whether or not the FEC decoder in the OTN mapper contributes to the accumulated LD status



LD and RD passed in AMs in AUI or Ethernet link, out-of-band to adjacent sublayer across MII
Drawback in description: PCS has different logic depending on whether it is directly below an RS or XS

2

https://www.ieee802.org/3/bs/public/17_03/trowbridge_3bs_01_0317.pdf

- FEC degrade signals should be retained.
- FEC degrade signal can propagate through OTN network.
- Both FEC degrade signals and FEC bin counters can be used to determine the link quality.

Multiplexing rules
Comments [27, 89, 92] - (part 1)

- Three comments were received proposing additional PMA multiplexing constraints above and beyond what was agreed to in the adopted baseline: https://www.ieee802.org/3/df/public/22_10/22_1004/shrikhande_3df_01a_221004.pdf
- Comment #6 against Draft 1.0 made a similar proposal. Straw polls recorded in the response to comment #6 indicated favor for adopting the proposal but there were many votes for “need more information”, and more consensus building was necessary.
- A new presentation provides more information on the problem (ran_3df_01_2301) and suggests essentially the same remedy (with minor editorial changes).

Multiplexing rules
Comments [27, 89, 92] - (part 2)

- Below is the final response to D1.0 comment #6, including the results of the straw poll.

REJECT.

The current text and constrainted PCSL multiplexing requirement is consistent with the adopted baseline (see slides 17&18 in https://www.ieee802.org/3/df/public/22_10/22_1004/shrikhande_3df_01a_221004.pdf) .

The following presentation was reviewed by the task force:
https://www.ieee802.org/3/df/public/22_12/ran_3df_01a_2212.pdf

Based on straw polls #1 and #2 there is no consensus to make the proposed changes at this time. The commenter is encouraged to refine the proposal and build consensus.

Straw poll #1 (direction)
I would support the changes proposed on slides 10 in ran_3df_01a_2212.
Yes: 19
No: 6
Need more information: 27

Straw poll #2 (direction)
I would support the changes proposed on slides 11 in ran_3df_01a_2212.
Yes: 13
No: 6
Need more information: 33

Suggested remedy (modified) – part 1

173.4.2.1 32:8 PMA bit-level multiplexing

Change the second list item as shown:

– The multiplexing function has an additional constraint that each of the 8 output lanes contain two unique PCSLs from PMA client lanes i = 0 to 15 ~~and~~ followed by two unique PCSLs from PMA client lanes i = 16 to 31

173.4.2.2 8:32 PMA bit-level multiplexing

Change the second list item as shown:

– The multiplexing function has an additional constraint that each of the 8 output lanes contain two unique PCSLs from service interface lanes i = 0 to 15 ~~and~~ followed by two unique PCSLs from service interface lanes i = 16 to 31.

Suggested remedy (modified) – part 2

173.4.2.3 8:8 PMA bit-level multiplexing

Change the second list item as shown:

– The 4 PCSLs received on any input lane shall be mapped together to an output lane such that the Gray-coded PAM4 symbol sequence on the output is identical to the Gray-coded PAM4 symbol sequence on the input (see 173.4.7.1). ~~The order of PCSLs from an input lane does not have to be maintained on the output lane.~~