

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 169 SC 169.1.2 P177 L41 # R1-1
 Brown, Matthew Alphawave
 Comment Type E Comment Status X
 Figure 169-1 is relevant to any 800GBASE PHY, not just 800GBASE-R PHY types.
 SuggestedRemedy
 Under the medium block change "800GBASE-R" to "800GBASE".
 Proposed Response Response Status O

Cl 90A SC 90A.3 P251 L44 # R1-4
 Marris, Arthur Cadence Design Systems, Inc.
 Comment Type E Comment Status X
 There is crossed out text "Annex_" that should not be there
 SuggestedRemedy
 Change "See Annex_90A.3" to "See 90A.3" on line 44.
 Proposed Response Response Status O

Cl 169 SC 169.4 P184 L13 # R1-2
 Brown, Matthew Alphawave
 Comment Type E Comment Status X
 The units bit times and pause_quanta are defined twice in this subclause. First in the opening paragraph and again in the table footnotes.
 SuggestedRemedy
 Change: "Table 169-4 contains the values of maximum delay (sum of transmit and receive delays at one end of the link) for each instance of a sublayer in bit times (as specified in 1.4.215) and pause_quanta (as specified in 31B.2) for 800 Gigabit Ethernet."
 To: Change: "Table 169-4 contains the values of maximum delay (sum of transmit and receive delays at one end of the link) for each instance of a sublayer."
 Proposed Response Response Status O

Cl 173 SC 173.5.4 P244 L37 # R1-5
 Brown, Matthew Alphawave
 Comment Type E Comment Status X
 Reference to "169.4 and its references" is unnecessarily verbose.
 SuggestedRemedy
 Change "169.4 and its references" to "169.4".
 Proposed Response Response Status O

Cl 169 SC 169.4 P184 L14 # R1-3
 Brown, Matthew Alphawave
 Comment Type E Comment Status X
 For a description of bit times the paragraph points to the definition in 1.4.215 while the description of pause_quanta points to a reference in 31B.2, even though there is a definition for pause_quantum in 1.4.459 which refers to 31B.2.
 SuggestedRemedy
 Change the reference for pause_quanta description from 31B.2 to 1.4.459.
 Proposed Response Response Status O

Cl 31B SC 31B.4.6 P255 L50 # R1-12
 Marris, Arthur Cadence Design Systems, Inc.
 Comment Type T Comment Status X
 Need to add PICS item TIM17 for 800 Gbps
 SuggestedRemedy
 Add new PICS item at end of 31B.4.6
 TIM17 Measurement point for station at 800 Gb/s 31B.3.7 Delay at MDI ≤ 1810 pause_quanta MIIp: M Yes
 Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 172A SC 172A P288 L10 # R1-13

He, Xiang Huawei Technologies Co., Ltd

Comment Type T Comment Status X

There were errors for AM portion in tx_scrambled_am<i>:j> tables for both flows. To be more precise, row 2-8 (<257:2055>) of Table 172A-1 and 172A-4.

SuggestedRemedy

Change the AM portion in rows 2-8 of Table 172A-1 and Tble 172A-4 to the correct values as shown in the contribution discussed during the .3dj & .3df joint ad hoc on Nov. 2.

Proposed Response Response Status O

Cl 171 SC 171.3.3 P195 L36 # R1-14

Slavick, Jeff Broadcom Inc

Comment Type T Comment Status X

The PHY 800GXS is the same as the 800GMII that is defined in Clause 170, so the wording is a bit odd. Follow the wording used in 172.1.5.1

SuggestedRemedy

Change "The service interface below the PHY 800GXS is defined as the 800GMII in Clause 170, with some exceptions and additional signals as follows:"

to
"The service interface below the PHY 800GXS is the 800GMII defined in Clause 170, with the following exceptions and additional signals:"

Proposed Response Response Status O

Cl 173 SC 173.5.2.1 P237 L15 # R1-15

Slavick, Jeff Broadcom Inc

Comment Type T Comment Status X

In 173.4.1 we state that the Tx bit multiplexing function is restricted and Rx is unrestricted for the 32:8 PMA. In 173.5.2.1 we state the PMA provides bit-multiplexing for Tx and Rx and then repeat the transmit bit-multiplex is done over these lanes and then magically convert from general bit-multiplexing phrase to "restricted bit multiplexing".

SuggestedRemedy

In the second paragraph. Change "The restricted bit-level multiplexing function is identical"
To: "This is a restricted bit-level multiplexing function that is identical"

In the third paragraph. Change "The unrestricted bit-level multiplexing function is identical"
To: "This is an unrestricted bit-level multiplexing function that is identical"

Proposed Response Response Status O

Cl 173 SC 173.5.2.2 P237 L37 # R1-16

Slavick, Jeff Broadcom Inc

Comment Type T Comment Status X

In 173.4.2 we state that the Tx bit multiplexing function is unrestricted and Rx is restricted for the 8:32 PMA. In 173.5.2.2 we state the PMA provides bit-multiplexing for Tx and Rx and then repeat the transmit bit-multiplex is done over these lanes and then magically convert from general bit-multiplexing phrase to "unrestricted bit multiplexing".

SuggestedRemedy

In the second paragraph. Change "The unrestricted bit-level multiplexing function is identical"

To: "This is an unrestricted bit-level multiplexing function that is identical"

In the third paragraph. Change "The restricted bit-level multiplexing function is identical"
To: "This is a restricted bit-level multiplexing function that is identical"

Proposed Response Response Status O

Cl 173 SC 173.5.2.3 P238 L15 # R1-17

Slavick, Jeff Broadcom Inc

Comment Type T Comment Status X

In 173.4.3 we state that the Tx and Rx bit multiplexing function is restricted for the 8:8 PMA. In 173.5.2.3 we state the PMA provides bit-multiplexing for Tx and Rx and then state transmit bit-multiplex is done over these lanes and then magically convert from general bit-multiplexing phrase to "restricted bit multiplexing".

SuggestedRemedy

In the third paragraph. Change "The restricted bit-level multiplexing function is identical"
To: "This is a restricted bit-level multiplexing function that is identical"

Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 172A SC 172A P282 L24 # R1-18

Slavick, Jeff Broadcom Inc

Comment Type T Comment Status X

Just before "the" 257-bit block was scrambled is not quite correct since it doesn't truly specify which of the 32 257-bit blocks in each flow the seeds applies to, but it is the first one

SuggestedRemedy

Change: "just before the 257-bit block was scrambled"
To: "prior to scrambling the first 257-bit block"

Proposed Response Response Status O

Cl 172A SC 172A P282 L30 # R1-19

Slavick, Jeff Broadcom Inc

Comment Type T Comment Status X

The scrambling and mapping processes have produced a state of the tx_scrambled_am variable which are shown in the tables.

SuggestedRemedy

Change: "the variable tx_scrambled_am is produced as shown in "
To: "the state of the variable tx_scrambled_am is shown in"

Proposed Response Response Status O

Cl 172A SC 172A P288 L4 # R1-20

Marris, Arthur Cadence Design Systems, Inc.

Comment Type T Comment Status X

There are errors in the "tx_scrambled_am i;j Flow <f>" table values.

My understanding is that the values in the tables incorrectly used the following coding:

```
For all k=0 to 11
For all j=0 to 7
if even(k)
am_mapped<160k+20j+ 9:160k+20j > = am_{2j }<10k+9:10k>
am_mapped<160k+20j+19:160k+20j+10> = am_{2j+1}<10k+9:10k>
else
am_mapped<160k+20j+ 9:160k+20j > = am_{2j+1}<10k+9:10k>
am_mapped<160k+20j+19:160k+20j+10> = am_{2j }<10k+9:10k>
```

when it should have used the following coding:

```
For all k=0 to 11
For all j=0 to 7
if even(k)
am_mapped<160k+20j+ 9:160k+20j > = am_{2j }<10k+9:10k>
am_mapped<160k+20j+19:160k+20j+10> = am_{2j+1}<10k+9:10k>
else
am_mapped<160k+20j+19:160k+20j+10> = am_{2j+1}<10k+9:10k>
am_mapped<160k+20j+ 9:160k+20j > = am_{2j }<10k+9:10k>
```

SuggestedRemedy

Please correct the example coding tables in Annex 172A

Proposed Response Response Status O

Cl 172 SC 172.1.5.1 P211 L47 # R1-21

Ran, Adeo Cisco Systems, Inc.

Comment Type E Comment Status X

"The PCS service interface is the 800GMII in Clause 170"
(twice, line 47 and line 50)

Similar references to xGMII clauses in the base document use the word "defined". For example see 149.3.1.

SuggestedRemedy

Change to "The PCS service interface is the 800GMII defined in Clause 170", twice.

Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 172 SC 172.1.5.1 P212 L1 # R1-22

Ran, Adeo Cisco Systems, Inc.

Comment Type E Comment Status X

"The TXRD and TXLD status signals indicate..."

These are not referred to as "status signals" elsewhere. The subsequent two paragraphs describe RXRD and RXLD without the word "status".

The last paragraph has "The PCS_status signal indicates..." but in this case "status" is part of the signal name - this adds confusion.

SuggestedRemedy

Change to "The TXRD and TXLD signals indicate..."

Proposed Response Response Status

Cl 172 SC 172.2.4.6 P216 L38 # R1-23

Ran, Adeo Cisco Systems, Inc.

Comment Type T Comment Status X

"tx_am_sf<2:0> = {FEC_degraded_SER + rx_local_degraded,0,0}"

The "+" sign apparently means logical-or here, but it is used in two other places in this subclause and in Figure 172-3 with the meaning of numerical addition. It can also be interpreted as addition modulo 2 (XOR) as used in other contexts.

The text should be made unambiguous.

Also applies to 171.6.1, although there are no additional + signs there.

SuggestedRemedy

Add "and + denotes logical OR" after "where FEC_degraded_SER and rx_local_degraded are defined in 172.2.6.2.2".

Add a similar statement in 171.6.1, including references to the variable definitions in 172.2.6.2.2.

Proposed Response Response Status

Cl 172 SC 172.2.5.2 P221 L12 # R1-24

Ran, Adeo Cisco Systems, Inc.

Comment Type T Comment Status X

"Within a flow, the data from the 16 PCS lanes is de-interleaved to reconstruct the original two streams of FEC codewords"

The similar statement in 119.2.5.2 is "the two FEC codewords are de-interleaved to reconstruct the original stream of two FEC codewords". And indeed this is a single stream of (pairs of) codewords, not two (independent) streams, that should be reconstructed.

The wording of 119.2.5.2 may be improved by changing "the original stream of two FEC codewords" to "the original stream of FEC codewords", or alternatively "of FEC codeword pairs" if the CRG prefers.

SuggestedRemedy

Change "Within a flow, the data from the 16 PCS lanes is de-interleaved to reconstruct the original two streams of FEC codewords" to "Within a flow, the data from the 16 PCS lanes is de-interleaved to reconstruct the original stream of FEC codewords".

Proposed Response Response Status

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 173 SC 173.2 P237 L8 # R1-25

Ran, Adee Cisco Systems, Inc.

Comment Type T Comment Status X

"The 8:32 and 8:8 PMAs may optionally provide signal status information to the PMA client by disabling (squelching) one or more of the PAM4 symbol streams sent to the PMA client (PMA:IS_UNITDATA_0:7.indication), see 173.5.8.2 and 173.5.8.3

This sentence is technically inaccurate - it is the output lane (AUI transmitter) that is squelched, not the PAM4 symbol streams; a squelched transmitter does not correspond to any PAM4 symbol stream. Indeed, the text in 173.5.8.2 and 173.5.8.3 uses different wording.

It is also is not directly related to the subject of this subclause, PMA service interface. Since signal detect is defined in other subclauses, this level of detail is not necessary here.

Similarly for the 4th paragraph in 173.3.

SuggestedRemedy

In 173.2, change the quoted sentence to "The 8:32 and 8:8 PMAs may optionally provide signal status information to the PMA client as described in 173.5.8.2 and 173.5.8.3" and make it a separate paragraph.

In the 4th paragraph of 173.3, change "the 8:8 PMA may optionally provide signal status information by disabling (squelching) one or more of the PAM4 symbol streams sent to the sublayer below via PMA:IS_UNITDATA_0:7.request (see 173.5.8.3)" to "the 8:8 PMA may optionally provide signal status information to the sublayer below as described in 173.5.8.3".

Proposed Response Response Status O

Cl 172A SC 172A P282 L30 # R1-26

Ran, Adee Cisco Systems, Inc.

Comment Type E Comment Status X

"the variable tx_scrambled_am is produced as shown in Table 172A-1 for flow 0 and Table 172A-4 for flow 1"

and then

"The expanded codewords are shown in Table 172A-2 and Table 172A-3 for flow 0, and in Table 172A-5 and Table 172A-6 for flow 1"

This annex would be easier to read and follow if the order of the tables was such that tables 172A-1 and 172A-4 appear first, right after the text that describes them, followed by the text that describes the remaining tables, and the remaining tables. All tables would be renumbered accordingly.

SuggestedRemedy

Re-order the tables and the text per the comment.

Proposed Response Response Status O

Cl 30 SC 30.5.1.1.2 P37 L44 # R1-27

Ran, Adee Cisco Systems, Inc.

Comment Type T Comment Status X

Following the response to comment I-43:

The changes to the entries for 200GBASE PHYs are not within the scope of this project, which is "for 400 Gb/s and 800 Gb/s Operation".

The changes to the entries for existing 400GBASE PHYs (400GBASE-DR4, 400GBASE-SR4, 400GBASE-SR4.2, 400GBASE-SR8, 400GBASE-SR16, and 400GBASE-VR4) should be reconsidered as they may affect existing implementations.

SuggestedRemedy

Delete the changes related to 200GBASE PHYs.

Consider deleting the changes to existing 400GBASE PHYs and making appropriate changes to the descriptions of new 400GBASE PHYs to distinguish them from existing ones instead..

Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 171 SC 171.8 P202 L44 # R1-28
 Dudek, Michael Marvell
 Comment Type E Comment Status X
 Sentence without a verb
 SuggestedRemedy
 Change "described" to "are described"
 Proposed Response Response Status O

Cl 172A SC 172A P288 L10 # R1-29
 Nicholl, Shawn Advanced Micro Devices (AMD)
 Comment Type T Comment Status X
 There are errors in "Table 172A-1 - Example tx_scrambled with alignment marker group for 800GBASE-R PCS flow 0" table values, specifically rows 2-8. The errored values differ from the expected values based on 119.2.4.4.2.
 Similar errors exist in "Table 172A-4 - Example tx_scrambled with alignment marker group for 800GBASE-R PCS flow 1" table values.
 SuggestedRemedy
 A presentation is expected that provides the correct values.
 Proposed Response Response Status O

Cl 116 SC 116 P L # R1-30
 Ran, Adee Cisco Systems, Inc.
 Comment Type T Comment Status X
 The PHY type 400GBASE-DR4-2 introduced by this amendment is not listed in clause 116.
 The following seem to require updates:
 116.1.2 item h
 116.1.3: Table 116-2
 116.1.4: Table 116-5
 SuggestedRemedy
 Add Clause 116 into the amendment and add 400GBASE-DR4-2 in the locations listed in the comment, and elsewhere if required.
 Proposed Response Response Status O

Cl 124 SC 124.1.1 P105 L9 # R1-31
 Dawe, Piers J G NVIDIA
 Comment Type E Comment Status X
 This sentence needs more work. At present, it says that if something is not good enough to achieve an end, something else has to be better than what's needed to achieve that unachievable end.
 However, clarifying this may be out of scope.
 pdf page 100, printed page 105

SuggestedRemedy
 If the error statistics are not sufficiently random to meet the specified frame loss ratio for 64-octet frames with minimum interpacket gap "when the BER is at the limit", then the BER shall be less than the value required to meet that frame loss ratio.
 Proposed Response Response Status O

Cl 124 SC 124.5.4 P110 L11 # R1-32
 Dawe, Piers J G NVIDIA
 Comment Type T Comment Status X
 The signal detect max could be defined better, considering that the same modules are used for 400GBASE-DR4-2 and 800GBASE-DR8-2 lanes and 100GBASE-FR1.
 SD thresholds would be lower than 0.2 dB below spec-worst sensitivity, so it's OK to base the SD max on -7.1 while the average power min is -6.9 dBm.

SuggestedRemedy
 For 400GBASE-DR4-2 and 800GBASE-DR8-2, change the SIGNAL_DETECT Optical power at TP3 criterion from "average receive power, each lane (min) in Table 124-7" to >=-7.1 dBm.
 Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

CI 124 SC 124.5.4 P110 L12 # R1-33

Dawe, Piers J G

NVIDIA

Comment Type T Comment Status X

At present an OMA-based signal detect is required to say OK for a signal at -6.9 dBm regardless of its extinction ratio, so a signal with $-6.9-4.2+3 = -8.1$ dB OMA must be shown as OK when the intended minimum OMA at the receiver is $-0.1-4 = -4.1$ dBm. (4.2 dB is the extinction ratio penalty for 3.5 dB). ("compliant 400GBASE-R or 800GBASE-R signal" is about signalling rate, scrambling and so on.)

The proposed remedy is based on -7.1 dB average power (see another comment).

Notice that "The PMD receiver is not required to verify whether a compliant 400GBASE-DR4 signal is being received", so the receiver may reject a signal that fails any of the three criteria without checking the other two.

SuggestedRemedy

For 400GBASE-DR4-2 and 800GBASE-DR8-2, SIGNAL_DETECT should be OK when:

Optical power at TP3 ≥ -7.1 dBm; and

OMA at TP3 ≥ -4.3 dBm; and

compliant 400GBASE-R or 800GBASE-R signal input.

Proposed Response Response Status O

CI 171 SC 171.6 P201 L21 # R1-34

Dawe, Piers J G

NVIDIA

Comment Type T Comment Status X

The FEC degrade feature doesn't propagate FEC degrade conditions. It signals or reports them, and sometimes in the opposite direction, so the first "propagate" doesn't work. Is "all" telling us something (what?) or is it a rhetorical flourish? If the feature is present, it reports a lack of FEC degrade (nothing untoward detected) too.

SuggestedRemedy

Change the first sentence from "The FEC degrade feature provides the ability to detect degrade conditions at the RS-FEC decoder using FEC degrade detection and to propagate all detected FEC degrade conditions using FEC degrade signaling." to "The FEC degrade feature provides the ability to detect degrade conditions at the RS-FEC decoder using FEC degrade detection and to report FEC degrade conditions using FEC degrade signaling."

If "all" is intentional, change it to "report all three possible types of FEC degrade condition". Same in 172.1.4.

Proposed Response Response Status O

CI 172 SC 172.1.3 P211 L18 # R1-35

Dawe, Piers J G

NVIDIA

Comment Type T Comment Status X

MDIO is optional. So is any management, usually, although "it is recommended that an equivalent access is provided" (172.3).

SuggestedRemedy

Change "and informing" to "and, optionally, informing"

Proposed Response Response Status O

CI 172 SC 172.2.4.6 P216 L49 # R1-36

Dawe, Piers J G

NVIDIA

Comment Type E Comment Status X

Font size

SuggestedRemedy

Fix

Proposed Response Response Status O

CI 172A SC 172A P287 L11 # R1-37

Dawe, Piers J G

NVIDIA

Comment Type E Comment Status X

A more specific reference would make this annex easier to use

SuggestedRemedy

Change cross-reference from Clause 172 to 172.2.4

Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 172A SC 172A P287 L50 # R1-38
 Dawe, Piers J G NVIDIA
 Comment Type E Comment Status X
 These valuable tables are easier to use in plain text format. D3.0 comment 107 "Please prepare a plain-text file with the large tables for convenient reading into a program, and post it on the project web site for review with future drafts". Files have been made available.
 SuggestedRemedy
 Upload the text files, eventually to <https://standards.ieee.org/downloads/802.3/> , and include a NOTE here bringing them to the reader's attention.
 Proposed Response Response Status O

Cl 172A SC 172A P282 L51 # R1-39
 Dawe, Piers J G NVIDIA
 Comment Type TR Comment Status X
 Experience with Annex 172A shows us how valuable it is. But more complexity follows: twice "Mux and 10-bit symbol distribution" as in 119.2.4.8 Figure 119-11 (with an order reversal that doesn't seem to be mentioned in the text), then 32:8 bit mux as in 173.5.2.1 where the two flows get interleaved, which is a new thing and worth an example.
 SuggestedRemedy
 Show some of the 16+16-lane output of the PCS for these cxA and cxB. It may be enough to show e.g. the beginnings of lanes 1 and 31, enough to include some differences between four codewords.
 Also show some of the 8-lane output of an 32:8 bit mux from that (which could go in a NOTE in 173). Again, showing a couple of lanes would be enough to resolve most or all misinterpretations or ambiguities. Add a cross-reference from here.
 If only a few hundred bits are needed, it could go in text. But if a more complete example is preferred, tables could be added and plain-text equivalents uploaded.
 Proposed Response Response Status O

Cl 172A SC 172A P288 L10 # R1-40
 Dawe, Piers J G NVIDIA
 Comment Type TR Comment Status X
 Improved tx_scrambled_am tables and text files are available
 SuggestedRemedy
 Use the improved tables and text files
 Proposed Response Response Status O

Cl 171 SC 171.8 P202 L44 # R1-41
 Dawe, Piers J G NVIDIA
 Comment Type E Comment Status X
 Missing verb
 SuggestedRemedy
 ... are described
 Proposed Response Response Status O

Cl 169 SC 169.5 P187 L33 # R1-42
 Dawe, Piers J G NVIDIA
 Comment Type T Comment Status X
 I suspect that the "N/A" here was copied from Table 116-9 and dates from a time when there were 26.5625 GBd (50G) AUIs but not 53.125 GBd AUIs. Now that there are, the missing numbers should be filled in.
 SuggestedRemedy
 Change the three N/A to approx 11, 202, 213.
 This should be done in Table 116-9 also, and a 53.125 GBd column should be added to Table 80-9 (both out of scope).
 Proposed Response Response Status O

Cl 169 SC 169.5 P187 L1 # R1-43
 Dawe, Piers J G NVIDIA
 Comment Type E Comment Status X
 Empty lines
 SuggestedRemedy
 Removing the blank space at lines 1 and 25-26 should let the 169.6 FEC Degrade section fit on this page.
 Proposed Response Response Status O

IEEE P802.3df D3.1 1st Sponsor recirculation ballot comments

Cl 169 SC 169.5 P185 L34 # R1-44

Dawe, Piers J G NVIDIA

Comment Type T Comment Status X

D2.0 comment 96: 0.2 ns Skew Variation. This dates back to SFI-5 when it was 1.5 UI of "relative wander at up to 11.1 Gbps" (per lane, so 0.14 ns). It got rounded up to 0.2 ns or just over 2 UI "dynamic skew" (giannakopoulos_01_1108) which was unfortunate. At 53.125 GBd this is 11 UI and "dynamic skew buffer per input lane Size is 2x the max dynamic skew", so over 21 UI, very roughly four times the length of the 4-tap or 6-tap AUI equaliser.

SuggestedRemedy

Define SP0 as the first exposed AUI interface (nearest the PCS or PHY 800GXS).
Recommend a max Skew Variation 0.1 ns or about 5 UI at 53.125 GBd there.
Modify 173.5.3 accordingly.

Proposed Response Response Status O

Cl 171 SC 171.1 P196 L35 # R1-45

Dawe, Piers J G NVIDIA

Comment Type ER Comment Status X

Layout

SuggestedRemedy

Set Figure 171-1 to float and save a page.

Proposed Response Response Status O

Cl 173 SC 173.4.1 P239 L1 # R1-46

Dawe, Piers J G NVIDIA

Comment Type ER Comment Status X

Possibly, removing the blank line 1 and reducing the figure at lines 9-10...

SuggestedRemedy

would let it fit on the previous page with its subclause text.

Proposed Response Response Status O