Cl **FM** *SC* **FM** *P* **5** *L* **21** # 4

Dawe, Piers Nvidia

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

Bad use of "may not", and contradictory to the meaning two paragraphs later. "The word may is used to indicate a course of action permissible within the limits of the standard (may equals is permitted to)."

SuggestedRemedy

Encourage IEEE staff to follow their own rules. "Statements made by volunteers may not represent..." should be changed to "Statements
made by volunteers do not necessarily represent...".
See another comment for another instance.

Proposed Response *Response Status* **O**

Cl **FM** *SC* **FM** *P* **6** *L* **39** # 5

Dawe, Piers Nvidia

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

Superscript 3 for footnote with URL for IEEE Xplore is in the wrong place

SuggestedRemedy

Have the staff move it from "contact IEEE." to "IEEE Xplore".

Proposed Response *Response Status* **O**

Cl **1** *SC* **1.1.3.2** *P* **31** *L* **13** # 7

Dawe, Piers Nvidia

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

This says about the 800GMII: "While conformance with implementation of this interface is not necessary to ensure communication, it allows flexibility in intermixing PHYs and DTEs at 800 Gb/s speeds. The 800GMII is a logical interconnection intended for use as an intra-chip interface. No mechanical connector is specified for use with the 800GMII. The 800GMII is optional." which is much the same as item d, GMII. As the current interfaces of choice for "allowing flexibility in intermixing PHYs and DTEs at 800 Gb/s speeds" are AUIs not MIIs, the first sentence quoted is misleading old cruft.

SuggestedRemedy

Delete the sentence "While conformance with implementation of this interface is not necessary to ensure communication, it allows flexibility in intermixing PHYs and DTEs at 800 Gb/s speeds."

Proposed Response *Response Status* **O**

Cl **1** *SC* **1.1.3.2** *P* **31** *L* **17** # 27

Dawe, Piers Nvidia

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

This says "only an 8-lane version of 800GAUI-n (800GAUI-8) is defined" while actually, two versions of 800GAUI-8 are defined.

SuggestedRemedy

Change "For the P802.3df project only an 8-lane version of 800GAUI-n (800GAUI-8) is defined. However, it is anticipated that in subsequent 800GbE projects other widths, e.g., a four-lane version (800GAUI-4), will be defined."
to "For the P802.3df project only 8-lane versions of 800GAUI-n (800GAUI-8) are defined. However, it is anticipated that in subsequent 800GbE projects other widths, e.g., four-lane versions (800GAUI-4), will be defined."

Proposed Response *Response Status* **O**

Cl **1** *SC* **1.1.3.2** *P* **31** *L* **17** # 8

Dawe, Piers Nvidia

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

This text "The 800GAUI-n is a physical instantiation of the PMA service interface... While conformance with implementation of this interface... The 800GAUI-n is intended... For chip-to-chip interfaces and for chip-to-module interfaces, one width of 800GAUI-n is defined: an eight-lane version (800GAUI-8) in Annex 120F and Annex 120G. No mechanical connector is specified for use with the 800GAUI-n. The 800GAUI-n is optional." reads as if there is only one kind of 800GAUI-n, and its specification is spread over two annexes. This is wrong; 800GAUI-n C2M and 800GAUI-n C2C are distinct, not interchangeable, and not intended to interoperate with each other. There is not "a version". Also, "the PMA service interface" is inaccurate; there can be more than one PMA service interface per MAC. Note the definition 1.4.184h uses "A" not "The".

SuggestedRemedy

Change the paragraph to: x) 800 Gb/s Attachment Unit Interface (800GAUI-n). An 800GAUI-n is a physical instantiation of a PMA service interface to extend the connection between 800 Gb/s capable PMAs. While conformance with implementation of 800GAUI-n is not necessary to ensure communication, it is recommended, since it allows maximum flexibility in intermixing PHYs and DTEs at 800 Gb/s speeds. 800GAUI-n C2C is intended for use as a chip-to-chip and 800GAUI-n C2M is intended as a chip-to-module interface. One width of 800GAUI-n is defined for chip-to-chip interfaces and one for chip-to-module interfaces: eight-lane 800GAUI-8 C2C in Annex 120F and eight-lane 800GAUI-8 C2M in Annex 120G. No mechanical connector is specified for use with a 800GAUI-n. A 800GAUI-n is optional.

Proposed Response *Response Status* **O**

Cl **1** *SC* **1.4.184h** *P* **33** *L* **37** # 10

Dawe, Piers Nvidia

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

This says that 800GAUI-n is used for chip-to-chip or chip-to-module electrical interfaces. It says that an eight-lane version when in fact, two versions are defined.

SuggestedRemedy

Change: 800 Gb/s Attachment Unit Interface (800GAUI-n): A physical instantiation of the PMA service interface to extend the connection between 800 Gb/s capable PMAs over n lanes, used for chip-to-chip or chip-to-module electrical interfaces. For chip-to-module interfaces and for chip-to-chip interfaces, one width of 800GAUI-n is defined: an eight-lane version (800GAUI-8). (See IEEE Std 802.3, Annex 120F and Annex 120G.)
to: 800 Gb/s Attachment Unit Interface (800GAUI-n): A physical instantiation of the PMA service interface to extend the connection between 800 Gb/s capable PMAs over n lanes, used for chip-to-chip or chip-to-module electrical interfaces. One width of 800GAUI-n is defined for chip-to-chip interfaces and one for chip-to-module interfaces: eight-lane 800GAUI-8 C2C and eight-lane 800GAUI-8 C2M. (See IEEE Std 802.3, Annex 120F and Annex 120G.)

Proposed Response *Response Status* **O**

Cl **1** *SC* **1.4.184k** *P* **34** *L* **2** # 11

Dawe, Piers Nvidia

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

Tautology: "PCS Sublayer" and "RS sublayer"

SuggestedRemedy

Delete Sublayer and sublayer, or spell out PCS and RS

Proposed Response *Response Status* **O**

Cl **1** *SC* **1.4.461** *P* **34** *L* **19** # 12

Dawe, Piers Nvidia

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

Difficult to parse "carried on a physical lane together at the..."

SuggestedRemedy

Change to "carried together on a physical lane at the..." or "carried on a single physical lane at the..." or "carried together on a different number of physical lanes at the...".

Proposed Response *Response Status* **O**

Cl **45** *SC* **45.2.1.7.4** *P* **42** *L* **16** # 1

Dudek, Mike Marvell

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

The separation between 400GBASE-KR4 and 400GBASE-KR4 should be a comma, not a period

SuggestedRemedy

Fix it.

Proposed Response *Response Status* **O**

Cl **45** *SC* **45.2.3.25.2** *P* **60** *L* **20** # 2

Dudek, Mike Marvell

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

The editor's note has served its purpose

SuggestedRemedy

Delete it

Proposed Response *Response Status* **O**

Cl **124** *SC* **124.3.1** *P* **104** *L* **14** # 13

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

The delay for 800GBASE-DR8 or 800GBASE-DR8-2 PMD including 2 m of fiber in one direction should be the same 20.48 ns as 400GBASE-DR4 and all other 200GBASE-R and 400GBASE-R optical PMDs (see tables 116-6 and 7). It was changed "because modern PMDs contain DSP": but this is not correct; reading all of 116.3.1 Inter-sublayer service interface, and 120.1.3 Summary of functions "the PMA ... Provide per input-lane clock and data recovery" and P802.3cw 156.2.1.2.1 Semantics of the primitive "The PMD_UNITDATA.indication primitive conveys four analog signals, representing the in-phase (I) and quadrature (Q) components for each of the polarizations...", it is clear that the PMD does optical to electrical conversion, and may provide some continuous-time equalization (which adds very little latency), and the PMA does clock recovery, A to D and any DSP. For a typical retimed module, the PMA-PMD interface is internal so it doesn't matter much, but as linear and co-packaged optics become more popular, the interface is accessible.
Also note that a 32:8 or 8:30 PMA is "a SerDes" but a 8:8 PMA may be implemented as two SerDes back to back, with additional delay.

SuggestedRemedy

Revert the PMD allowance to 16,384 bit times (32 pause_quanta or 20.48 ns) for all 8x100G optical, consistent with all 1/2/4x100G optical. With the new way of accounting for PMA delay, as modified by another comment, this gives a module with one PMD and one PMA 20.48+81.92 = 102.4 ns. vs. D2.1 40.96+46.08 = 87.04 ns and 802.3-2018 20.48 + 92.16/2 (maybe) = 66.56 ns which seems to be tight for some DSP.

Proposed Response *Response Status* **O**

Cl **124** *SC* **124.8.1** *P* **117** *L* **8** # 17

Dawe, Piers Nvidia

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

"or valid 400GBASE-R signal or 800GBASE-R signal": it doesn't make sense that the 400GBASE-R signal has to be valid and the 800GBASE-R one doesn't (even though we don't define a non-valid 400GBASE-R signal so the word isn't needed, but it is there in the base text). Compare Table 167-11 "3, 4, 5, 6, or valid 100GBASE-R, 200GBASE-R, 400GBASE-R, or 800GBASE-R signal".

SuggestedRemedy

Change "3, 4, 5, 6, or valid 400GBASE-R signal or 800GBASE-R signal" to "3, 4, 5, 6, or valid 400GBASE-R or 800GBASE-R signal" (i.e. put "or 800GBASE-R" before the first (pre-existing) "signal" and delete the second one).

Proposed Response *Response Status* **O**

Cl **124** *SC* **124.8.5b** *P* **119** *L* **28** # 18

Dawe, Piers Nvidia

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

The definition of overshoot and undershoot in 140.7.7 was done in a hurry and the 1e-2 hit ratio allows a surprising amount of overshoot beyond the limit (because only a fraction of 1 UI in every 8 UI "takes part in the measurement")

SuggestedRemedy

Change to 3e-3 as in Clause 167. The limits can be adjusted to keep the effect of the spec the same. Similarly for 124.8.5c Transmitter power excursion.

Proposed Response *Response Status* **O**

Cl **124** *SC* **124.11a** *P* **124** *L* **23** # 19

Dawe, Piers Nvidia

Comment Type **ER** *Comment Status* **X**

It would be bad economics to fragment the market for 400GBASE-DR4-2 modules into those that can interoperate with 400GBASE-DR4 and those that can't, when there is no cost to being interoperable. D2.0 comment 86. As 400GBASE-DR4 is well established but 400GBASE-DR4-2 is new, and as having a lower power for the higher performance PMD is counter-intuitive, the draft 400GBASE-DR4-2 should be brought into line.

SuggestedRemedy

Delete "and the 400GBASE-DR4-2 transmitter average power is greater than or equal to the value for average launch power (min) for 400GBASE-DR4 in Table 124-6." In Table 124-6, change the Average launch power, each lane (min) from -3.1 dBm to -2.9 dBm, same as 400GBASE-DR4.
Similarly for 800GBASE-DR8-2.

Proposed Response *Response Status* **O**

Cl **124** *SC* **124.12.4.4** *P* **128** *L* **21** # 20

Dawe, Piers Nvidia

Comment Type **ER** *Comment Status* **X**

This use of + is used in several clauses in this draft. It is not defined in 21.6.2, but it is useful.

SuggestedRemedy

In 21.6.2, add: <item1>+<item2>: OR-predicate condition, the requirement has to be met if either or both optional items are implemented

Proposed Response *Response Status* **O**

Cl **162** *SC* **162.1** *P* **130** *L* **20** # 6

Dawe, Piers Nvidia

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

Bad use of "may not", and contradictory to the meaning at Table 167-6. "The word may is used to indicate a course of action permissible within the limits of the standard (may equals is permitted to)." This issue is fixed in 162A.1. Missing word "associated". Also, see style guide 10.1.2 That and which.

SuggestedRemedy

Change "information on parameters with test points that may not be testable in an implemented system" to "parameters associated with test points which might not be testable in an implemented system", aligning with 162A.1.

Proposed Response *Response Status* **O**

Cl **162** *SC* **162.8.1** *P* **137** *L* **8** # 21

Dawe, Piers Nvidia

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

Ambiguous sentence "The PMDs on both ends of the link have connected ground references." The PMDs are connected to ground? to each other? the lanes in a PMD are connected together? What does "ground reference" (as opposed to "ground") mean? If this sentence means the PMDs are connected to each other, is it telling the implementer to arrange such a connection (through mains earth?) Are Signal shield and/or Link shield in Fig 162-2 involved?

SuggestedRemedy

This phrase appears four times in this draft. It is base text so it may have to go to maintenance, but this is the ideal group to advise what it is trying to say.

Proposed Response *Response Status* **O**

Cl **169** *SC* **169.4** *P* **182** *L* **16** # 23

Dawe, Piers Nvidia

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

colocated (twice)

SuggestedRemedy

FWIW, 55B has co-located

Proposed Response *Response Status* **O**

Cl **169** *SC* **169.4** *P* **182** *L* **28** # 9

Dawe, Piers Nvidia

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

The delay allowance for a 8:8 PMA is too low, and the allowance for an optical PMD is too high and out of step with other optical PMDs. (The allowance for CR or KR PMD+AN may be wrong too, but it doesn't matter much as they are always combined with PMAs.)

SuggestedRemedy

Change "800GBASE-R PMA" to "32:8 or 8:32 800GBASE-R PMA". Add a row "8:8 800GBASE-R PMA,65,536 BT, 128 PQ, 81.92 ns. Revert the VR8, SR8, DR8 and DR8-2 PMD allowances to 16,384 BT, 32 PQ, 20.48 ns.

Proposed Response *Response Status* **O**

Cl **169** *SC* **169.4** *P* **182** *L* **28** # 22

Dawe, Piers Nvidia

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

It's clear that in Clause 120, there is one "PMA sublayer" in a stack for a port, which is how "layers" are usually used, but it could contain up to four "PMA stages". In this draft, we have up to four "instances of the 800GBASE-R PMA", and according to 173.5.4, the numbers for the PMA row apply to an instance not a sublayer.

SuggestedRemedy

Write something like "Each instance of a PMA" in the Notes column. Change the heading of the left column to "Sublayer or instance".

Proposed Response *Response Status* **O**

Cl **169** *SC* **169.5** *P* **185** *L* **34** # 16

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

D2.0 comment 96: As discussed, the Skew Variation limits were based on a digital clock rate that is
slow by modern standards, and they were heavily sandbagged. It is important to sort this out for 800G so that the future 200G/lane-based Ethernet is not locked into decisions made long ago for technology that doesn't apply in this case. This draft has better Skew numbers but Skew Variation needs more investigation.

SuggestedRemedy

Continue the investigation, revise the numbers according to relevant technology, take out some of the padding.

Proposed Response *Response Status* **O**

Cl **169** *SC* **169.6** *P* **185** *L* **51** # 24

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

This says "... FEC degrade functionality is identical to that defined ... in 116.6." But 116.6 is just non-normative introduction, it contains no definition and not even any cross-references.

SuggestedRemedy

Change "Optional FEC degrade functionality is identical to that defined for 200 Gigabit Ethernet and 400 Gigabit Ethernet in 116.6." to "Optional FEC degrade functionality is as described for 200 Gigabit Ethernet and 400 Gigabit Ethernet in 116.6. For the 800GBASE-R PCS, it is defined in 172.2.5.3 (see 119.2.5.3), 172.2.5.3 (see 119.2.5.3) and 172.2.6 (see 119.2.6.2). For the 800GMII Extender, see 171.2, 118.2.1, 171.3, 118.2.2, 171.6, and 118.2."
In 116.6, add "For the 200GBASE-R or 400GBASE-R PCS, it is defined in 119.2.5.3, 119.2.5.3, and 119.2.6.2. For the 200GMII Extender and 400GMII Extender, see 118.2.1, 118.2.2, and 118.2."

Proposed Response *Response Status* **W**

Cl **170** *SC* **170.1.2** *P* **188** *L* **29** # 26

Dawe, Piers Nvidia

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

This says "This logical interface [the 800GMII] is used to provide media independence so that an identical media access controller may be used with supported PHY types". It's not really media independence; the common PCS and PMA provide that. It would allow an identical media access controller to be used with different PCSs, if the 800GXS were not used. This is unlikely.

SuggestedRemedy

As it is not needed, delete the sentence

Proposed Response *Response Status* **O**

Cl **171** *SC* **171.1.1** *P* **195** *L* **39** # 25

Dawe, Piers Nvidia

Comment Type **ER** *Comment Status* **X**

Marketing-speak - change to standards language

SuggestedRemedy

Change "leverages" to "contains", "includes" or "uses", or "has the same functions as".

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4** *P* **211** *L* **10** # 34

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

There is an informative Annex 119A, 200GBASE-R and 400GBASE-R PCS FEC codeword examples.

SuggestedRemedy

As the Clause 172 PCS is subtly different to Clause 119, with partly different alignment markers and the block distribution and synchronised alignment marker groups of the two flow method, there are new opportunities for ambiguity and misunderstanding that 119A won't catch. So, please prepare a similar annex for Clause 172. Add text here and at the beginning of 172 and and 169.2.3 mentioning it. Revise the amendment description on page 14.
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.

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.1** *P* **211** *L* **11** # 36

Dawe, Piers Nvidia

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

Mixed parts of speech: Encode, State-diagram encoder, Stateless encoder, Rate matching, Block distribution, 64B/66B to 256B/257B transcoder and so on

SuggestedRemedy

Change Encode to Encoder or Encoding. Similarly in the title of 172.2.5.9, change Decode to Decoder or Decoding.

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.1** *P* **216** *L* **11** # 28

Dawe, Piers Nvidia

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

This wording causes confusion: "The portion of the figure above the “64B/66B to 256B/257B transcoder” is excluded." Which figure? How can they be excluded, it won't work!

SuggestedRemedy

Change to "The portion of Figure 119-11 above the “64B/66B to 256B/257B transcoder” is not used, as a similar process is done before distribution to the two flows (see Figure 172-4)."

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.1.1** *P* **211** *L* **19** # 35

Dawe, Piers Nvidia

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

"state-diagram decoder" (a tool to understand state diagrams) is something I would like to have. Would a "state-diagram encoder" turn a state diagram into code? That would be useful. If the alternative encoder needs to know the previous block as well as the one it is encoding, calling it "stateless" is borderline. So these names are not ideal.

SuggestedRemedy

Change to "Method A", "Method B" unless someone has a better suggestion.

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.5** *P* **212** *L* **19** # 29

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

"the two scramblers should be set to different states": this is too weak. The consequence of getting this wrong is much more than the bad spectrum or correlation issues we have seen elsewhere.

SuggestedRemedy

Change should to shall or is

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.6** *P* **212** *L* **35** # 31

Dawe, Piers Nvidia

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

In "and finally a unique pad per PCS lane...", "finally" is unfortunate, as the UPs don't come last. As it is only rhetorical, it can be left out.

SuggestedRemedy

Delete "finally"

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.6** *P* **212** *L* **36** # 32

Dawe, Piers Nvidia

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

172.2.4.6, Alignment marker mapping and insertion, incorporates 119.2.4.4, Alignment marker mapping and insertion, with exceptions. 119.2.4.4 is part of 119.2.4, Transmit. It says "The unique pad (UP0 to UP2) within the alignment markers and the PRBS9 pad at the end of the alignment maker group are ignored on receive."
172.2.5, Receive function > 172.2.5.1, Alignment lock and deskew, points to 119.2.5, Receive function. 119.2.5.1, Alignment lock and deskew, uninformatively says "It obtains lock to the alignment markers as specified by the alignment marker lock state diagram shown in Figure 119–12." 119.2.6.2.2, Variables, refers back to 119.2.4.4.
But I did not find anything more about the unique pads. What are they for?

SuggestedRemedy

Please add a few words here explaining why the unique pads are present. Please add a sentence in 172.2.5.1 saying which of CMs, UMs and UPs are used, for what: something like: "The state diagram in Figure 119-12

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.6** *P* **212** *L* **38** # 30

Dawe, Piers Nvidia

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

D2.0 comment 105 (accepted in principle): Add an informative NOTE saying what is common among these lanes, what is the same for the two flows, *and what is the same in 400G*.

SuggestedRemedy

To address the last point, please add something that gives the information in shrikhande_3df_01a_221004 slide 13:
CM0-CM5 and UP0-UP2 are unchanged from 400GbE CL119
UM0/UM3 for Flow lanes 0-15 are inverted from 400GbE
UM1/UM2/UM4/UM5 for Flow lanes 16-31 are inverted from 400GbE
e.g.:
The unique markers in flow 1 are bit-wise inversions of the ones in flow 0.
NOTE--CM0 to CM5 and UP0 to UP2 are the same as for 400GBASE-R (see Table 119–2). UM1, UM2, UM4, UM5 for flow 0 and UM0 and UM3 for flow 1 are are the same as for 400GBASE-R.

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.6** *P* **213** *L* **8** # 41

Dawe, Piers Nvidia

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

In the text above, CM0 to CM5, UM0, UP0 and so on are in regular text while in the tables, the numbers are subscripts. The subscripts are inconvenient.

SuggestedRemedy

Change the subscripts to regular text in these two figures

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.6** *P* **213** *L* **10** # 33

Dawe, Piers Nvidia

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

These table(s) of alignment markers could be put on the web in machine-readable format at https://standards.ieee.org/downloads/

SuggestedRemedy

Please prepare a plain-text file with the alignment markers (without cell straddling) for convenient reading into a program. Post it on the project web site for review with future drafts.

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.11** *P* **216** *L* **43** # 42

Dawe, Piers Nvidia

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

"is accessible through the register": which register?

SuggestedRemedy

is accessible through the BASE-R PCS test-pattern control register ?

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.4.11** *P* **216** *L* **44** # 43

Dawe, Piers Nvidia

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

Table 172-5

SuggestedRemedy

This is not a hotlink.

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.5.1** *P* **216** *L* **54** # 40

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

There is a new exception for the alignment lock and deskew process

SuggestedRemedy

The 800GBASE-R PCS receive function shall support a maximum Skew of 152 ns between PCS lanes.
(Editorial: "support" is lame, this should be tolerate.)

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.5.2** *P* **217** *L* **3** # 44

Dawe, Piers Nvidia

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

"PCS lanes can be received on different lanes of the service interface from which they were originally
transmitted." They aren't usually received on the service interface from which they were originally
transmitted, that's loopback. Lanes on lanes??

SuggestedRemedy

Signals can be received at the PCS with the lanes in a different arragement to that at the service interface from which they were originally transmitted. ?

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.5.2** *P* **217** *L* **10** # 45

Dawe, Piers Nvidia

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

the original stream of two FEC codewords - surely not just two codewords?

SuggestedRemedy

the original two streams of FEC codewords ?

Proposed Response *Response Status* **O**

Cl **172** *SC* **172.2.5.9** *P* **217** *L* **49** # 46

Dawe, Piers Nvidia

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

The receive PCS shall use the decoding method defined in either 172.2.5.9.1 or 172.2.5.9.2.

SuggestedRemedy

The receive PCS shall use one of two decoding methods, which are defined in 172.2.5.9.1 and 172.2.5.9.2.

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.2** *P* **232** *L* **54** # 47

Dawe, Piers Nvidia

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

The new optional squelch feature should be mentioned here. And, the word "squelch" should be used so readers will recognise it.

SuggestedRemedy

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.2.1** *P* **238** *L* **20** # 48

Dawe, Piers Nvidia

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

"the function": what or which function? Compare lines 31, 39, 46

SuggestedRemedy

Add words such as "bit-level multiplexing" at least here, the first time. e.g. "8:32 bit-level multiplexing" would be better.

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.2.1** *P* **238** *L* **28** # 37

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

"with two lanes from ... followed by two lanes from ..." isn't right. Lanes exist coninuously, they can be in parallel but cannot follow.

SuggestedRemedy

Bits from the four PCSLs are multiplexed in temporal order with one bit from each of two lanes from PMA client lanes i = 0 to 15 followed by one bit from each of two lanes from PMA client lanes i = 16 to 31. ?
Similarly in 173.5.2.2.

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.2.3** *P* **239** *L* **22** # 38

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

"except for possible swapping of each bit pair": bit pair is not specified, but maybe it means the pair of bits in a PAM4 symbol. Then, what is "swapping of each bit pair"? Swapping a PAM4 pair with another? Swapping the two bits within a PAM4 symbol? With or without Gray mapping? "except for possible" sounds like an anti-recommendation in ususual wording - is that meant? The reference points to 120.5.7.1, Gray mapping for PAM4 encoded lanes, it doesn't answer these questions.

SuggestedRemedy

The 4 PCSLs received on an input lane shall be mapped to one output lane. It is recommended that the Gray mapped PAM4 symbol sequence (see 173.5.7.1) on the output lane is identical to the Gray mapped PAM4 symbol sequence
on the input lane. Alternately, the the Gray mapped PAM4 symbol sequence on the output lane is [whatever is meant].

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.3** *P* **239** *L* **24** # 14

Dawe, Piers Nvidia

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

Delay should come before skew, as in 116 124, 162, 169 and so on, not after as in 120.

SuggestedRemedy

Move 173.5.4 Delay constraints to before 173.5.3 Skew and Skew Variation

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.3.1** *P* **239** *L* **39** # 39

Dawe, Piers Nvidia

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

In these subclauses, skew is generated, produced or delivered. It is not clear what these terms mean. I suspect that all limits are cumulative (unlike for delay) - but then how can an implementation of e.g. the 800GAUI-8 closest to the PCS "shall deliver no more than 145 ns of Skew" when it doesn't control its input Skew?

SuggestedRemedy

Define or clean up the terminology

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.4** *P* **240** *L* **32** # 3

Rechtman, Zvi NVIDIA

Comment Type **TR** *Comment Status* **X**

The new concept of PMA 32:8, PMA8:32 and PMA8:8 together with the separation of the delay constrain for each PMA, introduce some ambiguity.
For example: 8-lanes "retimer" device can be built using two entities of PMA8:32 and PMA32:8 or single PAM8:8 entity.
Therefore, the delay constraint for such "retimer" can be considered either as 46.08 nsec (PAM8:8), or its delay constraint can be considered as 2x46.08 = 92.16 nsec (PMA8:32+PMA32:8) which is more reasonable.

SuggestedRemedy

Split the delay constrains to two usecases:
1) Delay of 92.16 nsec for PMA8:8.
2) Delay of 46.08 nsec for PAM32:8 and PMA8:32.

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.4** *P* **240** *L* **35** # 49

Dawe, Piers Nvidia

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

within a Physical Layer, which is composed of an 800GBASE-R PHY and an optional 800GMII Extender

SuggestedRemedy

within a Physical Layer, which is composed of an 800GBASE-R PHY and, optionally, an 800GMII Extender

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.4** *P* **240** *L* **35** # 50

Dawe, Piers Nvidia

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

It would avoid misinterpretation if the words to the effect of delay is the sum of transmit and receive delays at one end of the link, were reinstated.

SuggestedRemedy

Per comment

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.5** *P* **241** *L* **2** # 51

Dawe, Piers Nvidia

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

If an output lane's clock is derived from its corresponding input, it's not independent.

SuggestedRemedy

As this is only an example, changing "independent" to "separate" or "its own" would be enough to fix it

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.8.2** *P* **242** *L* **13** # 52

Dawe, Piers Nvidia

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

It is hard work reverse engineering this: "In the *transmit* direction ... The SIGNAL_OK parameter is set to OK when data is being *received*...

SuggestedRemedy

Change "when data is being received on all 8 input lanes (PMA:IS_UNITDATA_0:7.request)." to "when data is being received by this PMA from the PMA sublayer above on all 8 transmit lanes (PMA:IS_UNITDATA_0:7.request) ?

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.8.3** *P* **242** *L* **18** # 53

Dawe, Piers Nvidia

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

Name this feature by its familiar name so readers can find it.

SuggestedRemedy

by disabling (squelching) one or more output lanes
Same in next subclause

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.5.8.3** *P* **242** *L* **19** # 54

Dawe, Piers Nvidia

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

Two dumb cross-references, and two more at line 29.

SuggestedRemedy

Make them hot links

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.6.4** *P* **240** *L* **46** # 15

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **X**

This new delay allocation per PMA-instance may be OK where a PMA is packaged with a PCS, XS or PMD, but it is tight for a standalone PMA (e.g. "on-board retimer"). It is unlikely that a PMA will be packaged with an exposed 32x25G PMA interface except in a prototype.

SuggestedRemedy

Increase the allowance for the 8:8 PMA only, from 36,864 BT, 72 PQ, 46.08 ns to 65,536 BT, 128 PQ, 81.92 ns. No need to change the delay allocation for 32:8 and 8:32 PMA.

Proposed Response *Response Status* **O**

Cl **173** *SC* **173.7.7** *P* **248** *L* **37** # 55

Dawe, Piers Nvidia

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

If the two loopback abilities aren't in the major options table, there is no point having separate PCS for "PMA local loopback" and "PMA local loopback implemented". Nothing else depends on "LBL".

SuggestedRemedy

Combine the two pairs

Proposed Response *Response Status* **O**