

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

CI **FM** SC **0** P **1** L **35** # **6**
 Geoff Thompson GraCaSI S.A.

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

The text on the front page does not describe where this draft is in the process. Instead of: "Draft D2.2 is prepared for Working Group Ballot." it would have been more appropriate to say: "Draft D2.2 is prepared for Working Group 2nd recirculation ballot."

SuggestedRemedy

Use the expanded description in the front page description on all future drafts.

Proposed Response Response Status **O**

CI **001** SC **1.4** P **39** L **3** # **1**
 Robert Grow RMG Consulting

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

The inserted definitions do not follow the dictionary sort order to be used for Std 802.3. When consulting with IEEE editors on a previous project, they consider changes to subclass numbering and related editing instructions to be non-substantive. (Specifically, they would be willing to make the changes during publication preparation without need for Sponsor ballot recirculation.) Applying this principal, there is ample precedent for making similar changes during WG ballot or in preparation for Sponsor Ballot. Based on current plan, this project will be an amendment to the now in process revision of IEEE Std 802.3-2015. Because the sort order for 1.4 both in 802.3-2015 and its approved and anticipated amendments has no consistency, it does not make sense to spend much time on working for this amendment until the inserted definitions are included and the complete set of definitions are resorted to a consistent dictionary sort order.

SuggestedRemedy

When the content of P802.3 has all amendments that will be included in the revision merged into the P802.3 draft, update the editing instructions and subclass numbers in P802.3cd to be based on that revision draft. (Hopefully before the initial Sponsor ballot on P802.3cd.)

Proposed Response Response Status **O**

CI **134** SC **134.5.3.1** P **156** L **12** # **8**
 Dawe, Piers Mellanox

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

After alignment marker lock is achieved the two FEC lanes

SuggestedRemedy

on both FEC lanes (as in 134.5.3.2)? on each FEC lane? on the two FEC lanes?

Proposed Response Response Status **O**

CI **135** SC **135.1** P **172** L **5** # **9**
 Dawe, Piers Mellanox

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

Missing text.

SuggestedRemedy

Add some text in for the overview explaining what this clause is about. Mention all the annexes briefly, in the style of 136.1.

Proposed Response Response Status **O**

CI **135** SC **135.1** P **172** L **20** # **10**
 Dawe, Piers Mellanox

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

defined in Clause 135B through Clause 135G. ... in Clause 135D through Clause 135G.

SuggestedRemedy

defined in Annex 135B through Annex 135G. ... in Annex 135D through Annex 135G.

Proposed Response Response Status **O**

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

Cl 135 SC 135.1.3 P 172 L 20 # 11
 Dawe, Piers Mellanox

Comment Type E Comment Status X

We have added another function, precoding. This isn't the same as Gray mapping, which is part of PAM4 coding - a PMA might do precoding but not PAM4 coding.

SuggestedRemedy

add item k, In some circumstances, perform precoding for PAM4.
 Add full stop to item j.

Proposed Response Response Status O

Cl 135 SC 135.5.7.2 P 184 L 12 # 12
 Dawe, Piers Mellanox

Comment Type T Comment Status X

Because a lane can run through PMAs or PMDs, this text is ambiguous: does an indirect connection count? In the first paragraph we have "PMA lanes connected to" and in the last two paragraphs we have "PMA lanes adjacent to".
 Also, per 120D.1, "The... C2C link is described in terms of a ... C2C transmitter, a ... C2C channel, and a ... C2C receiver." So a PMA lane connected to a C2C link (not part of the link) might be further up or down the chain.

SuggestedRemedy

Change "For PMA lanes connected to a 50GAUI-1 C2C or 100GAUI-2 C2C link, or to the PMD service interface of a 50GBASE-CR, 50GBASE-KR, 100GBASE-CR2, or 100GBASE-KR2 PMD, the PMA shall provide $1/(1+D) \bmod 4$ precoding capability on each output lane and may optionally provide $1/(1+D) \bmod 4$ decoding capability on each input lane."
 to "A PMA shall provide $1/(1+D) \bmod 4$ precoding capability on each output lane that is part of a 50GAUI-1 C2C or 100GAUI-2 C2C transmitter, or is adjacent to the PMD service interface of a 50GBASE-CR, 50GBASE-KR, 100GBASE-CR2, or 100GBASE-KR2 PMD. A PMA may optionally provide $1/(1+D) \bmod 4$ decoding capability on each input lane that is part of a 50GAUI-1 C2C or 100GAUI-2 C2C receiver, or is adjacent to the PMD service interface of a 50GBASE-CR, 50GBASE-KR, 100GBASE-CR2, or 100GBASE-KR2 PMD."
 In the penultimate paragraph, change "For PMA inputs and outputs adjacent to a 50GBASE-CR PMD" to "For PMA lanes adjacent to a 50GBASE-CR PMD".
 In the last paragraphs, change "For PMA lanes adjacent to a 50GAUI-1 C2C" to "For PMA inputs and outputs that are part of a 50GAUI-1 C2C".

Proposed Response Response Status O

Cl 135F SC 135F.1 P 367 L 7 # 29
 Dawe, Piers Mellanox

Comment Type T Comment Status X

This annex does not refer to Clause 135 at all, nor does it mention precoding for the data path.

SuggestedRemedy

Make reference to 135.
 Here, add sentence saying that a receiver may request precoding and a transmitter should? shall? follow the request.
 In 135F.3.1, say that in addition the C2C transmitter provides a precoding function that can be switched on and off.
 In 135F.3.2, say that in addition the C2C receiver may provide an inverse precoding function.

Proposed Response Response Status O

Cl 135G SC 135G.3.1 P 375 L 22 # 30
 Dawe, Piers Mellanox

Comment Type TR Comment Status X

As shown in http://ieee802.org/3/bs/public/adhoc/elect/05Oct_17/dawe_01b_100517_elect.pdf there is a need for an additional spec to protect the module from e.g. very noisy hosts, and a max VEC spec provides worthwhile protection.

SuggestedRemedy

Here, add a requirement for VEC, max 12 dB. In 135G.4, add definition of VEC, which was in P802.3bs D2.0 120E.4.2.1 (the AVs were illustrated in Figure 120E-13, although they could be on Fig 120E-14 and the text under what was equation 120E-3 is clear enough so we don't have to add them to the figure).
 Add PICS to 135G.5.4.1.

Proposed Response Response Status O

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

CI 135G SC 135G.3.4 P 375 L 35 # 7
 Wertheim, Oded Mellanox Technologie

Comment Type T Comment Status X

The jitter specification for the 100G per lane 100GBASE-DR1 receiver uses the same frequency corner as the 50G per lane 100GAUI-2 with the same jitter but with half the peak-to-peak jitter as the jitter mask is defined in UIs. This requires the 100GBASE-DR transceiver PMA to implement a de-jitterizer, which requires to add a PLL to handle the low frequency jitter and a jitter buffer. This adds unnecessary complexity, cost and power to the transceiver.

SuggestedRemedy

Scale the corner frequency for 100GAUI-2 to 2MHz (half the corner frequency of 100GBASE-DR).

The proposed resolution doesn't introduce constraints on future 100G per lane interfaces and provides simpler solution than previous options that were investigated.

1.Add an exception to 135G.3.4 50GAUI-1 C2M and 100GAUI-2 C2M module input characteristics:

With an exception that:

a.The reference CRU for the Module stressed input test and Host stressed input test has a corner frequency of 2MHz

b.The applied sinusoidal jitter values for 100GAUI-2 Module stressed input test and Host stressed input test shall be:

{Jitter frequency, Jitter amplitude}

Case A: {0.02, 5}

Case B: {0.66, 0.15}

Case C: {2, 0.05}

Case D: {6, 0.05}

Case E: {20, 0.05}

2.Add an exception to 135G.4 50GAUI-1 C2M and 100GAUI-2 C2M measurement methodology

With an exception that:

a.The reference CRU for the Eye width and eye height measurement method has a corner frequency of 2MHz

See presentation to follow with additional details.

Proposed Response Response Status O

CI 136 SC 136.1 P 197 L 11 # 13
 Dawe, Piers Mellanox

Comment Type E Comment Status X

"There are three associated annexes." No, there are four.

SuggestedRemedy

Change three to four. Add sentence for 136C.

Proposed Response Response Status O

CI 136 SC 136.8.11.1.3 P 212 L 1 # 4
 Hidaka, Yasuo Fujitsu Lab. of Americ

Comment Type T Comment Status X

Although the variable "n" was changed to "p", the statement starting "The polynomials for each identifier value p and" is still difficult to read, because "i" is used as the variable for the identifier in page 211 at line 30, and it is written as "p=i" in page 211 at line 46. It seems that the variable "p" is not necessary, and if we can avoid the variable "p", we can simplify the description. Or, if the variable "p" is necessary, I recommend to rephrase the statement at line 1, page 212.

SuggestedRemedy

At line 46, page 211, change "for p = i" to "for identifier i".

At line 1, page 212, change "The polynomials for each identifier value p and the default seeds" to "The polynomial and the default seed for each identifier value i".

In Table 136-8, change "p" to "i" in the first column, and change "Polynomial_p" to "Polynomial" in the second column.

Proposed Response Response Status O

CI 136 SC 136.8.11.1.3 P 212 L 18 # 3
 Hidaka, Yasuo Fujitsu Lab. of Americ

Comment Type E Comment Status X

120.5.11.2.3 describes SSPRQ test pattern.

SuggestedRemedy

Change 120.5.11.2.3 to 120.5.11.2.1.

Proposed Response Response Status O

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

CI 136 SC 136.9.3 P 225 L 36 # 2
Hidaka, Yasuo Fujitsu Lab. of Americ

Comment Type T Comment Status X

AC common-mode RMS output voltage (max) is specified as 30mV at TP2 normatively in 136.9.3. It is also specified as 30mV at TP0a in 137.9.2 by a reference to Table 120D-1 normatively for clause 137 and informatively for clause 136 by a reference from 136A.2. There should be some difference in these numbers in order to take account of the mode conversion from differential mode to common mode in signal propagation from TP0 to TP2. In the past clauses, the difference was often 18mV (12mV at TP0a and 30mV at TP2).

SuggestedRemedy

Change AC common-mode RMS output voltage (max) at TP2 in Table 136-11 from 30mV to 48mV.

Or, add the following exception in 137.9.2 as an exception to Table 120D-1: The AC common-mode output voltage (max, RMS) is 12mV.

Proposed Response Response Status O

CI 136 SC 136.9.3 P 226 L 10 # 14
Dawe, Piers Mellanox

Comment Type TR Comment Status X

As noted in D2.0 comment 143 and 144, and D2.1 comment 43, these TP2 Jrms and J4u limits, which are copies of the ones in Table 120D-1 (different BER, different test point) should be replaced with Jrms and J3u limits that are consistent (not the same) as the TP0a limits. Crosstalk at the connector combined with the slower edges increases J3u from TP0a to TP2.

SuggestedRemedy

Change J4u to J3u, here and in 137. Choose the limits at TP2 considering the jitter limit at TP0a, the mated compliance board crosstalk specs, and the slower edges at TP2. In 136.9.4.2.3 step e, change J4u to J3u (3 places).

Proposed Response Response Status O

CI 136 SC 136.9.4.2 P 230 L 26 # 15
Dawe, Piers Mellanox

Comment Type TR Comment Status X

The COM value in the receiver interference tolerance isn't a maximum, it's the reference value that defines what we mean by receiver interference tolerance, and it is used as a target when adjusting the injected noise. See maintenance D2.0 comments 135 and 136.

SuggestedRemedy

In Table 136-13, straddle the "Min" and "Max" columns for the "COM" row and place the contents of the "Max" column into the straddled column. Add the following table footnote to the "COM" parameter label.

"The COM value is the target value for the SNR_TX calibration defined in 136.9.4.2.3 item f). The SNR_TX value measured at the Tx test reference should be as close as practical to the value needed to produce the target COM. If lower SNR_TX values are used, this would demonstrate margin to the specification but this is not required for compliance."

Proposed Response Response Status O

CI 136 SC 136.9.4.2.2 P 230 L 42 # 16
Dawe, Piers Mellanox

Comment Type T Comment Status X

As pointed out in hidaka_3cd_01a_0517.pdf and hidaka_060717_3cd_adhoc-v2.pdf, and D2.0 comment 72, we need a spec for the test channel RL (Rx end) that's better than the regular cable RL spec given by 92.10.3, eq 92-27: 16.5-2rt.f to 4.1 GHz then 10.66-14log10(f/5.5). The comment proposed the mated test fixtures return loss limit, eq 92-38, 20-f to 4 GHz then 18-0.5f. Adopting a limit about half way between these two would be much better than doing nothing. See hidaka_3cd_01a_0517 slides 17/18 to end.

SuggestedRemedy

Insert new requirement into 136.9.4.2.2:
The test channel is the same as the one defined in 110.8.4.2.2, except that the cable assembly meets the requirements of 136.11, the differential return loss of the test channel measured at the Rx test reference (see Figure 110-3b) meets Equation (136-new)."
Eq 136-new to be about half way between eq 92-27 and eq 92-38. 18-f to 4 GHz then 16-0.5f

Proposed Response Response Status O

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

CI 136 SC 136.9.4.2.3 P 231 L 3 # 17
 Dawe, Piers Mellanox
 Comment Type T Comment Status X
 It is not likely that the frequency dependent attenuator would have 109.8 ohm impedance.
 SuggestedRemedy
 If the PCB impedance in 136.11.7.1 (referring to Table 92-12) is not changed (see another comment), add an exception here that the PCB impedance is 100 ohm.
 Proposed Response Response Status O

CI 136 SC 136.11.7.1 P 236 L 39 # 19
 Dawe, Piers Mellanox
 Comment Type T Comment Status X
 Using 109.8 ohm PCB impedance in COM could provide an incentive to build cables to that (wrong) impedance, which seems unhelpful.
 SuggestedRemedy
 Add another exception to Table 92-12: Zc = 100. In 136.11.7.1.1 and 136.11.7.1.2, delete "and the parameter values given in Table 92-12" (because that is already stated in 136.11.7.1).
 Proposed Response Response Status O

CI 136 SC 136.11 P 237 L 5 # 20
 Dawe, Piers Mellanox
 Comment Type TR Comment Status X
 This cable loss limit is based on bad reasoning (copying a number from a backplane spec, which is something that doesn't apply to this cable spec). It will be a benefit to the market if a 50GBASE-CR cable can also be a CA-25G-S cable, so the limit should be consistent with 16.48 dB, adjusted for Nyquist frequency. Setting it too high is objective creep (the objective is "copper twin-axial cables with lengths up to at least 3m"), and creates a class of 50GBASE-CR cables that aren't CA-25G-S compliant. I made an estimate of the adjustment and got 16.93 dB. This can be rounded off to 17, which is still significantly more than the 16.06 dB in D1.3.
 This comment is a refinement of D2.0 comment 44.
 SuggestedRemedy
 Change the max loss from 17.16 to 17, to be consistent with CA-25G-S, in Table 136-14, 136.11.2, Table 136A-1 and Figure 136A-1 (two places). Change the RITT losses in Table 136-13 from 15.16 and 17.16 to 15 and 17. In Table 136A-1 and Figure 136A-1, change ILChmax from 30 to 29.84.
 Proposed Response Response Status O

CI 136C SC 136C.1 P 387 L 41 # 31
 Dawe, Piers Mellanox
 Comment Type T Comment Status X
 The paragraph about AC coupling, which should be a property of and requirement on the cable not the MDI, is in the wrong place. The subclause reference in PICS CA9 is wrong.
 SuggestedRemedy
 Move this paragraph to 136.11 just before 136.11.1 (older clauses have it in the equivalent of 136.12, which is not really correct but at least it's in the clause). Update the subclause reference in PICS CA9.
 Proposed Response Response Status O

CI 136 SC 136.11.7 P 235 L 18 # 18
 Dawe, Piers Mellanox
 Comment Type TR Comment Status X
 The COM impedances should be moved towards neutral, as explained in D2.0 comment 71 and 113.
 SuggestedRemedy
 Make changes similar to D2.0 comment 71 and hidaka_3cd_01_0717
 Proposed Response Response Status O

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

CI 137 SC 137.9.2 P 251 L 23 # 21
 Dawe, Piers Mellanox

Comment Type T Comment Status X

Now that COM is defined with a near-neutral termination and package impedance, we don't expect transmitter return loss to align to the COM model any more. This RL is much tighter than CEI-56G-LR-PAM4 at low (and high) frequency (although apparently looser between 4 and 9 GHz). At low frequencies it is tighter than the channel RL. The effect of (good) RL at low frequency is much less than the less good RL at higher frequencies anyway, and there is less concern about end-to-end reflections than in C2C because the loss is higher when the receiver is challenged. So we can go back to what we had a few drafts ago.

SuggestedRemedy

Insert a new first item in the list of exceptions to Table 120D-1, create a new equation for Tx RL that is similar to the Cl.93 and the channel RL at low frequencies; $12 - 0.625f$, $8.7 - 0.075f$. Add figure to illustrate.

Proposed Response Response Status O

CI 137 SC 137.9.2 P 251 L 28 # 22
 Dawe, Piers Mellanox

Comment Type TR Comment Status X

Transmitter output residual ISI SNR_ISI (min) 36.8 dB (Clause 136) and 43 dB (Clause 137) is still too high - can barely measure the IC through the test fixture. The warning NOTE in 120D.3.1.7 shows the issue, but doesn't solve it. D2.0 comment 140, D21. comment 49.

SuggestedRemedy

Proposed Response Response Status O

CI 137 SC 137.9.2 P 251 L 29 # 23
 Dawe, Piers Mellanox

Comment Type TR Comment Status X

Signal-to-noise-and-distortion ratio (min), increased to 33.3 dB (Clause 136) and to 32.5 dB (Clause 137) for all Tx emphasis settings, is still too high. D2.0 comment 139, D2.1 comment 50.

SuggestedRemedy

Proposed Response Response Status O

CI 137 SC 137.9.2 P 251 L 30 # 24
 Dawe, Piers Mellanox

Comment Type TR Comment Status X

This clause with a BER of $2.4e-4$ needs a J3u spec, just as 120D with a BER of $1e-5$ uses J4u. Using J3u enables a shorter measurement as well as a more relevant, accurate one.

SuggestedRemedy

Add exception 5: the J4u limit in Table 120E-1 does not apply but the maximum J3u is 0.106 UI.

In Eq 136-7 and 136-8 and the NOTE, change J4u to J3u, $Q4=3.8906$ to $Q3=3.2905$, $Q(Q3) = 5 \times 10^{-4}$.

Jrms and its value don't change.

If wished, add an informative NOTE in 137.9.2 saying that the J3u limit here is consistent with the J4u limit in Table 120D-1.

Add a new subclause:

136.9.3.n J3u Jitter

J3u is defined similarly to J4u (see 120D.3.1.8). J3u is defined as the time interval that includes all but 10-3 of $fJ(t)$, from the 0.05th to the 99.95th percentile of $fJ(t)$.

Proposed Response Response Status O

CI 137 SC 137.9.3 P 251 L 35 # 25
 Dawe, Piers Mellanox

Comment Type TR Comment Status X

Now that COM is defined with a near-neutral termination and package impedance, receiver mismatch is the receiver designer's concern, not the standard's, unless it is very extreme, because the receiver interference tolerance test finds its effect combined with other receiver attributes. And we don't expect receiver return loss to align to the COM model any more. This RL is much tighter than CEI-56G-LR-PAM4 at low (and high) frequency (although apparently looser between 4 and 9 GHz). At low frequencies it is tighter than the channel RL. The effect of (good) RL at low frequency is much less than the less good RL at higher frequencies anyway. So we can go back to what we had a few drafts ago.

SuggestedRemedy

Insert a new first item in the list of exceptions to Table 120D-5, create a new equation for Rx RL that is similar to the Cl.93 and the channel RL at low frequencies; $12 - 0.625f$, $8.7 - 0.075f$. Add figure to illustrate or point to the figure for Tx RL (see another comment).

Proposed Response Response Status O

IEEE P802.3cd 50 Gb/s, 100 Gb/s, 200 Gb/s Ethernet 2nd Working Group recirculation ballot comments

CI 138 SC 138.5.7 P 269 L 20 # 26
Dawe, Piers Mellanox

Comment Type E Comment Status X

Function names don't have underscores like this, although functional variable names do. See maintenance D2.0 comments 139, 142, compare 136.8.6.

SuggestedRemedy

Change "PMD_global_transmit_disable function" to "PMD global transmit disable function". Similarly in 139, 140.

Proposed Response Response Status O

CI 138 SC 138.5.7 P 269 L 26 # 27
Dawe, Piers Mellanox

Comment Type E Comment Status X

Improving the language. See maintenance D2.0 comment 140.

SuggestedRemedy

Change "set the PMD_global_transmit_disable to one" to "set the PMD_global_transmit_disable variable to one" Similarly in 139, 140.

Proposed Response Response Status O

CI 138 SC 138.5.8 P 269 L 30 # 28
Dawe, Piers Mellanox

Comment Type E Comment Status X

Function names don't have underscores, don't need lane numbers. See maintenance D2.0 comment 141, compare 136.8.7.

SuggestedRemedy

Change: The PMD_transmit_disable_i (where i represents the lane number in the range 0:3) function is optional...

a) When a PMD_transmit_disable_i variable is set...
If the optional PMD_transmit_disable_i function is not...
to:

The PMD lane-by-lane transmit disable function is optional...

a) When a PMD_transmit_disable_i variable (where i represents the lane number in the range 0:3) is set...

If the optional PMD lane-by-lane transmit disable function is not...

Proposed Response Response Status O

CI 140 SC 140.7.9 P 320 L 26 # 5
Maki, Jeffery Juniper Networks

Comment Type T Comment Status X

The applied sinusoidal jitter requirements in 121.8.9 are not correct in this case. An additional exception is needed.

SuggestedRemedy

Add the exception:

-- 80 kHz is to be used instead of 40 kHz and 8 MHz instead of 4 MHz.

Note that this proposed remedy places the burden solely on this PMD to have the correct requirements and there is no need to change any of the AU1 specifications. This approach is similar to what was done originally for 100GBASE-LR4 with CAUI-10.

Proposed Response Response Status O