802.3ck D2.1 Comment Resolution 162, 162A/C/D

Howard Heck, Intel Corporation
Chris Diminico, MC Communications/PHY-SI LLC/Panduit

Clause 162, 162A/C/D (Howard, Chris)

Clause	Topic	Comments
162	host/CA IL	92, 93
162	vf method/value	30, 107
162	SNDR test response	78
162	EOJ method	109
162	TP5 specifications	111
162	RLdc/RLcd graphs	115
162	broadband noise	113
162	CA RLcc	79, 94
162	CA COM DFE RSS	96
162A	host PCB IL	108
162C/162D	MDI names	64, 57
162C	MDI pins	63, ghiasi_02

Legend: [##,##,##] = related comments, ## = pivot comment, [##,##,author_nn] = related presentation

Host/CA IL
92, 93

RECAP OF DRAFT 2.0 comment 1.66

Dawe, Piers Nvidia

Comment Type **TR** *Comment Status* **R** *CR port type*

The draft loss budget wastes over 3 dB in nearly every case.
The recommended maximum insertion loss allocation for the host traces plus BGA footprint and host connector footprint, of 6.875 dB, compares very poorly with C2M's host insertion loss up to 11.9 dB, making passive copper expensive and unattractive for a switch, while a full range of NICs can be made within only 3.75 dB. Server-switch links will get made with an asymmetric loss budget, so it would be better for the standard to regularise what will happen anyway. By the way, many server-switch links will be asymmetric anyway (different form factors at server and switch ends), and that's already allowed in this draft.
This change would also benefit CR switch-switch links because the shortest ports would get credit for their low loss.

Response *Response Status* **U**

REJECT.

The following presentation was reviewed by the task force:
https://www.ieee802.org/3/ck/public/adhoc/apr28_21/dawe_3ck_adhoc_01_042821.pdf

The suggested remedy would require two or three different CR port types.

The assymetric-port approach was discussed early in this project.
Straw Poll #1 from the July 2018 Task Force meeting indicated strongest support for the current specification.
https://www.ieee802.org/3/ck/public/18_07/minutes_3ck_0718_approved.pdf

Based on discussion and straw poll 6 and 7, there is interest in exploring this proposal further. However, the proposal is not sufficiently complete at this time. A complete proposal and consensus is required.

Straw poll #6 (direction, chicago rule)
Straw poll #7 (direction, pick one)
I would support a new pair of CR port types with reduced host insertion loss limit on one end (e.g., NIC) and increased host loss limit on the other end (e.g., switch) similar to slide 7 of dawe_3ck_adhoc_01_042821.

Strawpoll #6
A: Yes 27
B: No 13
C: Need more information 29
D: Abstain 7

Straw poll #7
A: Yes 22
B: No 11
C: Need more information 11
D: Abstain 6

Host/CA IL 92, 93

Cl **162** *SC* **162.9.3** *P* **163** *L* **18** # 92

Dawe, Piers — Nvidia

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

The draft CR loss budget wastes over 3 dB in nearly every case. The relative range of host losses, 6.875/2.3 = 3:1, is too small for switch layout yet not needed for NICs.
The recommendation for the host traces plus BGA footprint and host connector footprint, 6.875 dB, compares very poorly with C2M's host insertion loss up to 11.9 dB, making passive copper to this draft expensive and unattractive for a switch, yet a full range of NICs can be made with only 3.75 dB. Server-switch links are asymmetric in form factor (e.g. QSFP-DD to 2 x QSFP) and will get made with an asymmetric loss budget, so it would be better for the standard to regularise what will happen anyway. C2M already has short and long ports.
This change would also benefit CR switch-switch links because the shortest ports would get credit for their low loss.
The symmetric budget is used for some designs under way and may be useful in future for LOM, so it is kept here, and the better way added.

SuggestedRemedy

3 classes of CR ports, host loss allocations of A 10, B 6.875, C 3.75 dB. B is as D2.1.
A connects to C, B to B or C, C to A, B or C.
Use 2 bits in Clause 73 Auto-Negotiation Link codeword Base Page to advertise A, B or C to the other end. In the Priority Resolution function, an A port ignores a 100G/lane Technology Ability Field bit from an A or B port, a B port ignores a 100G/lane Technology Ability Field bit from an A port.
In Table 162-10, add limits A and C for linear fit pulse peak ratio (min). Change text in 162.9.3.1.2 to refer to the table.
In Table 162-14, add columns for Test 2 (high loss), A and C, with test channel insertion loss: A: 6.875-3.75 = 3.125 dB lower (20.5 dB to 21.5 dB), and C: 10-6.875 = 3.125 dB higher (26.75 dB to 27.75 dB). No change needed for Test 1.
In 162A.4, add equations for IL_PCBmax and ILHostMax A and B and show them in Fig 162A-1 and 2. In 162A.5, add Value columns A, C in Table 162A-1 (ILChmin and ILMaxHost differ). Adjust figures 162A-3 and 4.

More information in dawe_3ck_adhoc_01a_071421
Straw poll at 7/14 ad hoc (during plenary):

Straw Poll #1:

At this time, I support a new pair of CR port types with reduced host insertion loss limit on one end (e.g., NIC) and increased host loss limit on the other end (e.g., switch) similar to dawe_3ck_adhoc_01_071421.

Yes 28
No 26
Abstain 18

Offline discussions and consensus building

For task force discussion. Straw poll needed to close.

Cl **162** *SC* **162.11** *P* **177** *L* **29** # 93

Dawe, Piers — Nvidia

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

The poor max cable loss makes CR unattractive, while all NICs and some ports on any switch have host loss going to waste. Enabling longer cables on a minority of links is needed.
In the remedy, each host knows the other host's loss class through AN and the cable's loss class from its I2C compliance code, so the situation is just like any other CR scenario, no extra management features needed in the spec for the long cable class.

SuggestedRemedy

2 classes of cable, which could be called "short" (19.75 dB, as today) and "long", 19.75+2*(6.875-3.75) = 19.75+6.25 = 26 dB max (achievable cable length 3 m). Long cables connect port types C (see another comment) at both ends, short cables connect a valid combination of A, B, C.
In 162.11.2, cable assembly insertion loss, change text to refer to Table 162-17.
In 162.11.7.1.1, add zp = 30.7 mm for the "short" cable.
In Table 162A-1, add a column for the A-short-A scenario (ILCamax differs).

162 Vf method/value 30, 107

Cl 162 SC 162.9.3.1.2 P 166 L 4 # 30

Ran, Adee Cisco systems

Comment Type **TR** *Comment Status* **D** *vf method*

"The steady-state voltage vf is defined in 136.9.3.1.2, and is determined using Nv=200 and the linear fit pulse peak ratio calculated by the procedure in 162.9.3.1.1"

It is determined _from_ the linear fit pulse, and the _peak ratio_ is irrelevant here.

Also, 162.9.3.1.1 does not use the parameter Nv - it has Np which is 13. This is the subject of another comment.

SuggestedRemedy

Change this sentence to
"The steady-state voltage vf is defined in 136.9.3.1.2, and is determined from the linear fit pulse peak ratio calculated by the procedure in 162.9.3.1.1 with the exception that Np is replaced by Nv=200" or "with Np=200".

Proposed Response *Response Status* **W**

PROPOSED ACCEPT IN PRINCIPLE.

This comment does not apply to the substantive changes between IEEE P802.3ck D2.1 and D2.0 or the unsatisfied negative comments from the initial ballot.
Hence it is not within the scope of the recirculation ballot.

However, the proposed change is an improvement to the draft.

Implement the suggested remedy.

The steady-state voltage v_f is defined in 136.9.3.1.2, and is determined using N_v=200 and the linear fit pulse peak ratio calculated by the procedure in 162.9.3.1.1. The steady-state voltage shall be greater than or equal

The steady-state voltage vf is defined in 136.9.3.1.2, and is determined from the linear fit pulse peak ratio calculated by the procedure in 162.9.3.1.1 with Np=200.

162 Vf method/value
107

Cl 162 SC 162.9.3.1.2 P 166 L 5 # 107

Dawe, Piers Nvidia

Comment Type T *Comment Status* D *vf value*

Redundantly stating normative requirements is bad practice. Table 162-10 is normative.

SuggestedRemedy

Change "The steady-state voltage shall be greater than or equal to 0.387 V and less than or equal to 0.6 V" to "The steady-state voltage shall be within the limits given in Table 162-10", "meet the requirements specified in Table 162-10", or similar.

Proposed Response *Response Status* W

PROPOSED ACCEPT IN PRINCIPLE.
Change "The steady-state voltage shall be greater than or equal to 0.387 V and less than or equal to 0.6 V" to "The steady-state voltage shall meet the requirements specified in Table 162-10"

162 SNDR test response 78

CI 162 SC 162.9.3.3 P 167 L 31 # 78

Dudek, Mike Marvell

Comment Type T *Comment Status* D *SNDR test response*

The measurement method for SNDR in 120D.3.1.6 uses a 33MHz filter bandwidth, which would take precedence over the statement that for Transmitter electrical characteristics "A test system with a fourth-order Bessel-Thomson low-pass response with 40 GHz 3 dB bandwidth is to be used for all transmitter signal measurements, unless otherwise specified as it is "otherwise specified". This was probably not intended and there is potential ambiguity here that should be removed. However as the Rx is only expected to have approximately the Nyquist bandwidth measuring SNDR to 40GHz may be excessive and we should consider using a narrower bandwidth.

SuggestedRemedy

Add a sentence. A test system with a fourth-order Bessel-Thomson low-pass response with 40 GHz 3 dB bandwidth should be used.

Proposed Response *Response Status* W

PROPOSED ACCEPT.

162.9.3.3 Output SNDR

The transmitter SNDR is defined by the measurement method described in 120D.3.1.6 with the exception that the linear fit procedure in 162.9.3.1.1 is used.

A test system with a fourth-order Bessel-Thomson low-pass response with 40 GHz 3 dB bandwidth should be used

162 EOJ Method
109

Cl **162**	*SC* **162.9.3.4**	*P* **167**	*L* **47**	# 109

Dawe, Piers — Nvidia

Comment Type **TR** *Comment Status* **D** *EOJ method*

Allowing 4 different ways to measure the same thing, admitting that they will give different results yet not ranking them, is too indecisive, and forces people to do all four tests in borderline cases. Worse, "lower than 4 MHz" is open-ended and introduces yet more uncertainty.

SuggestedRemedy

Pick one pattern and CRU corner as definitive, the others can be "if it passes/fails this it would have passed/failed".

Proposed Response *Response Status* **W**

PROPOSED REJECT.
The suggested remedy is not sufficiently complete to implement.

162 TP5 Specifications
111

Cl **162** SC **162.9.4** P **170** L **29** # **111**

Dawe, Piers Nvidia

Comment Type **E** *Comment Status* **D** *TP5 specifications*

The receiver specifications at TP5 are provided informatively in 162A.3: that's not what 162A.3 says.

SuggestedRemedy

The *recommended* receiver specifications at TP5 are... Also change the title of 162A.3, Receiver characteristics at TP5, to Recommended receiver characteristics at TP5.

Proposed Response *Response Status* **W**

PROPOSED ACCEPT IN PRINCIPLE.
The suggested change in wording in 162.9.4 is an improvement to the draft. However, if this text is changed, then similar text in 162.9.3 Transmitter characteristics should be updated.
It is not necessary to update the title for subclauses 162A.2 and 162A.3 since Annex 162A is informative and the text introduces the specifications as recommended.
In 162.9.3…
change "The transmitter characteristics at TP0 are provided informatively in 162A.2."
to "Recommended transmitter specifications are provided in 162A.2."
In 162.9.4…
change "The receiver specifications at TP5 are provided informatively in 162A.3."
to "Recommended receiver specifications are provided in 162A.3."
[Editor's note: CC: 162, 162A]

162.9.4

The receiver specifications at TP5 are provided informatively in 162A.3.

↓

The recommended receiver specifications at TP5 are provided informatively in 162A.3.

162.9.3

The transmitter characteristics at TP0 are provided informatively in 162A.2.

↓

The recommended transmitter specifications at TP0 are provided informatively in 162A.3.

162 RLdc/Rlcd graphs
115

Cl **162**	*SC* **162.9.4.6**	*P* **176**	*L* **11**	# 115

Dawe, Piers — Nvidia

Comment Type **ER** *Comment Status* **D** *RLdc/RLcd graphs*

Don't waste the reader's time.

SuggestedRemedy

Combine the graphs for Transmitter common mode to differential return loss and Receiver differential to common-mode return loss.

Proposed Response *Response Status* **W**

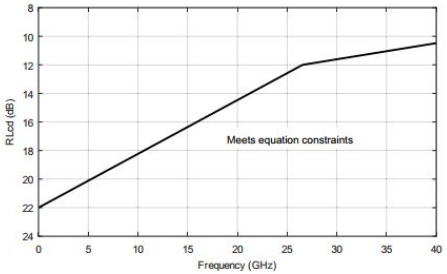
PROPOSED REJECT.

This comment does not apply to the substantive changes between IEEE P802.3ck D2.1 and D2.0 or the unsatisfied negative comments from the initial ballot.
Hence it is not within the scope of the recirculation ballot.

The two graphs represent requirements for different components, which happen in this case to have identical responses.
[Editor's note: Change page from 175.]

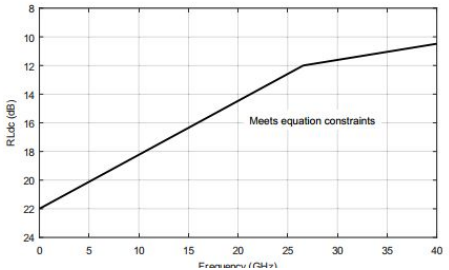
Rx RLcd

$$RLcd(f) \geq \begin{cases} 22 - 10(f/26.56) & 0.05 \leq f < 26.56 \\ 15 - 3(f/26.56) & 26.56 \leq f \leq 40 \end{cases}$$



Tx RLdc

$$RLdc(f) \geq \begin{cases} 22 - 10(f/26.56) & 0.05 \leq f < 26.56 \\ 15 - 3(f/26.56) & 26.56 \leq f \leq 40 \end{cases}$$



162 Broadband noise
113

Cl **162**	*SC* **162.9.4.3.3**	*P* **173**	*L* **38**	# 113

Dawe, Piers — Nvidia

Comment Type **E** *Comment Status* **D** *broadband noise*

"sigma_bn is the RMS broadband noise amplitude" means nothing because the text doesn't call it that.

SuggestedRemedy

Add "RMS broadband noise amplitude" to the text where sigma_bn is mentioned (step g).

Proposed Response *Response Status* **W**

PROPOSED ACCEPT IN PRINCIPLE.
Implement the suggested remedy with editorial license.

e) COM is used to calibrate the amplitude of the noise added to the signal before the Tx test reference using the definition of σ_{TX} given by Equation (162–9), Equation (162–10) and Equation (162–11). In Equation (162–9), SNR_{TX} is set to the SNDR value measured at the Tx test reference using the procedure in 120D.3.1 with the exception that the linear fit in 120D.3.1.3 is performed with a pulse length (N_p) of 29 UI and with pattern generator noise disabled. Determine the value of σ_{bn} required to achieve COM value specified in Table 162–15. The amplitude of the noise added to the signal before the Tx test reference is σ_{hp} which is derived from ~~σ_{bn}~~ as defined in Equation (162–12).

the RMS broadband noise amplitude, σ_{bn},

162 CA COM DFE RSS 96

Cl **162**	*SC* **162.11.7**	*P* **183**	*L* **40**	# 96

Dawe, Piers — Nvidia

Comment Type **TR** *Comment Status* **D** *COM DFE RSS*

The spec allows a cable (not even the whole channel) to have its COM calculated with 9 taps in the range 13 to 24 clipped at +/-0.05 - which means that the channel's pulse response could be worse than +/-0.05 for all these 9 taps. That's a very bad cable! and not likely to get made: there won't be that many reflections in the same area. (Remember, these are reference receiver limits not hard cable limits anyway; a cable can go beyond a tap limit if it makes up the COM another way, e.g. with acceptable crosstalk.)
We don't need to provide all the receiver power and complexity to cope with unreasonably bad cables.

SuggestedRemedy

Use another DFE root-sum-of-squares limit for positions 13-24. Similarly in 163, but as 163 specifies the complete channel while 162 uses clean synthetic host traces, the limit should be higher.

Proposed Response *Response Status* **W**

PROPOSED REJECT.

This comment does not apply to the substantive changes between IEEE P802.3ck D2.1 and D2.0 or the unsatisfied negative comments from the initial ballot.
Hence it is not within the scope of the recirculation ballot.

The suggested remedy is not complete.

162 CA RLcc

Comment #94, 79

Cl 162 SC 162.11.6 P 181 L 38 # 94

Dawe, Piers Nvidia

Comment Type TR *Comment Status* X

Relaxing the already very loose CM RL spec from 2 dB to 1.8 dB at all frequencies isn't justified. This draft spec becomes useless at the frequency when the MCB loss is 1.8/2 dB, which is only 8.5 GHz.

SuggestedRemedy

Use a frequency-dependent mask e.g. 1.6 + 0.01f. Similarly for Tx, Table 162-11, 162.9.3.6.

Proposed Response *Response Status* O

Cl 162 SC 162.11.6 P 181 L 38 # 79

Dudek, Mike Marvell

Comment Type T *Comment Status* X

As was pointed out in the unsatisfied comment # 177 against draft 2.0 the existing specification for common mode return loss limit effectively doesn't exist once the test fixture loss exceeds 0.9dB. The rejection however had a valid point that there is a potential issue up to 4GHz where the loss is low.

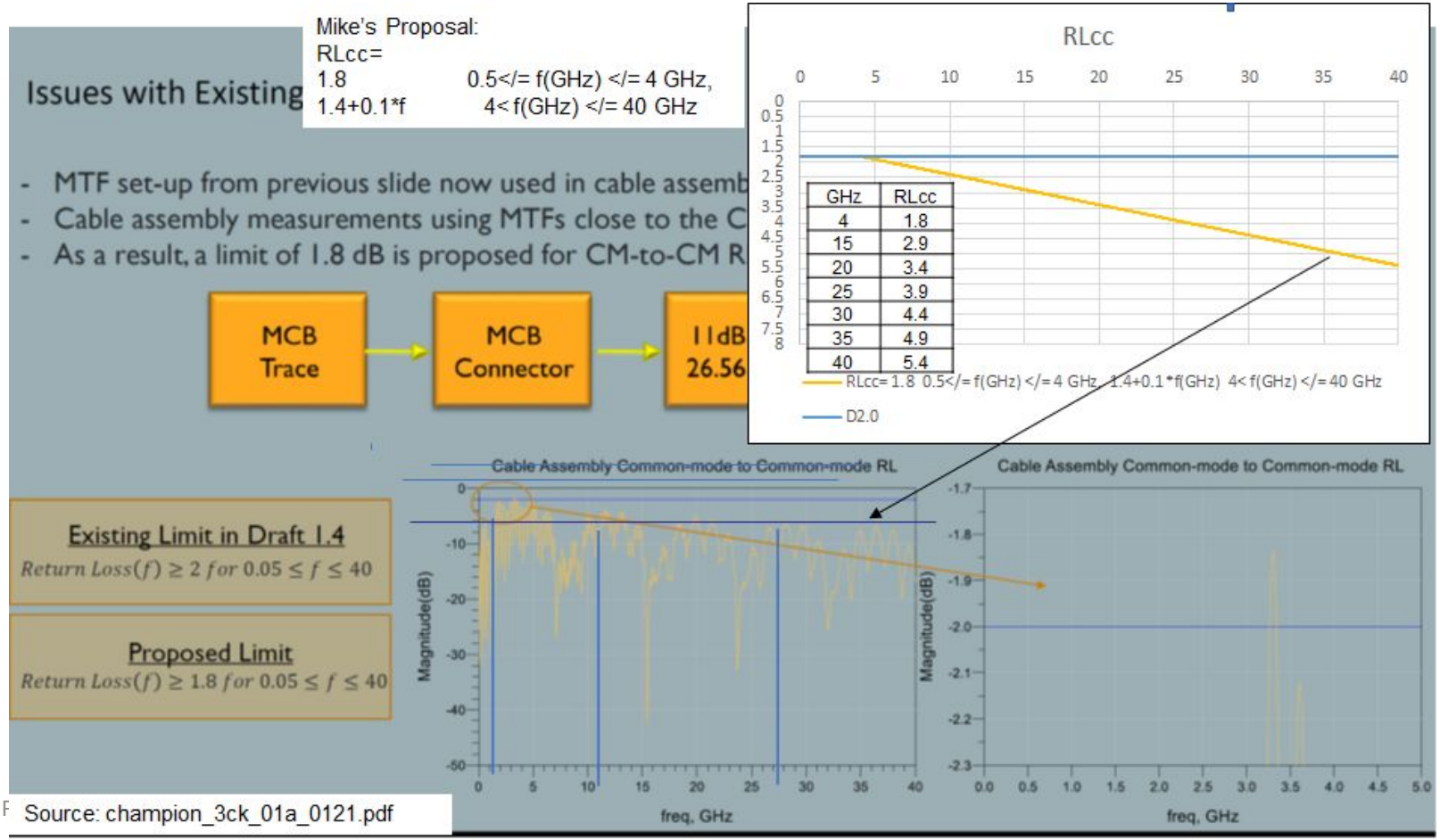
SuggestedRemedy

Change the limit to 1.8dB from 0 to 4GHz, 2.2-0.1*f from 4GHz to 40GHz.

Proposed Response *Response Status* O

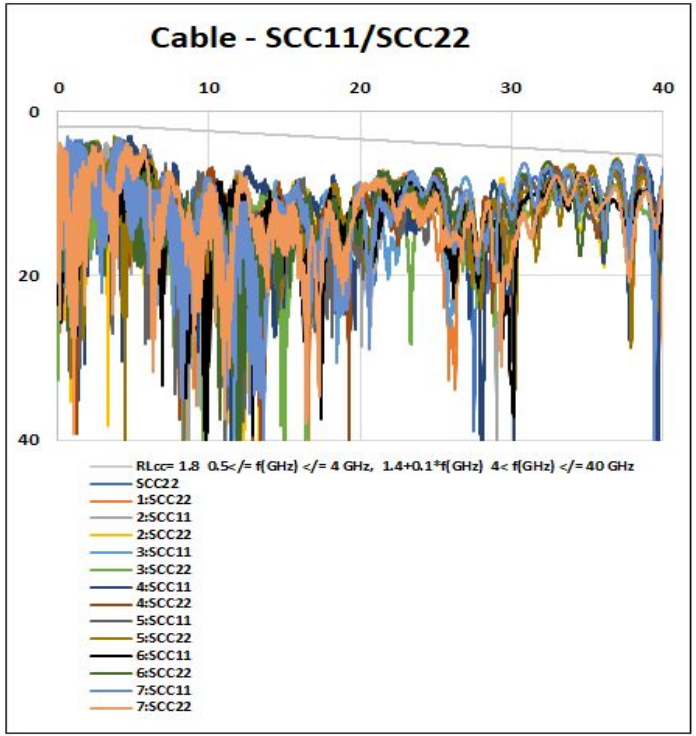
Mike's Proposal:
RLcc=
1.8 0.5</= f(GHz) </= 4 GHz,
1.4+0.1*f(GHz) 4< f(GHz) </= 40 GHz

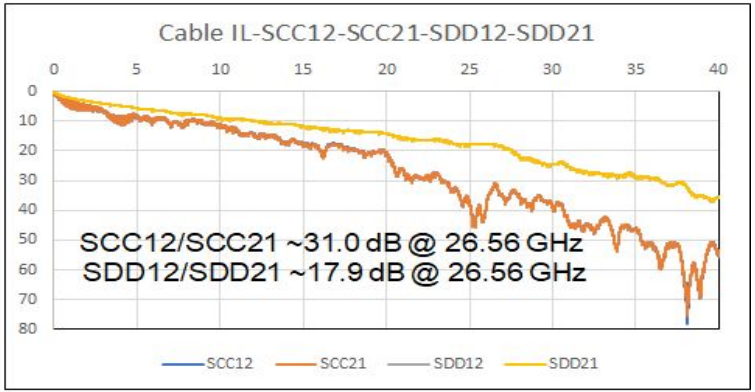
Comment 94 and 79



Comment 94 and 79

Measured OSFP 2m 25awg Cable

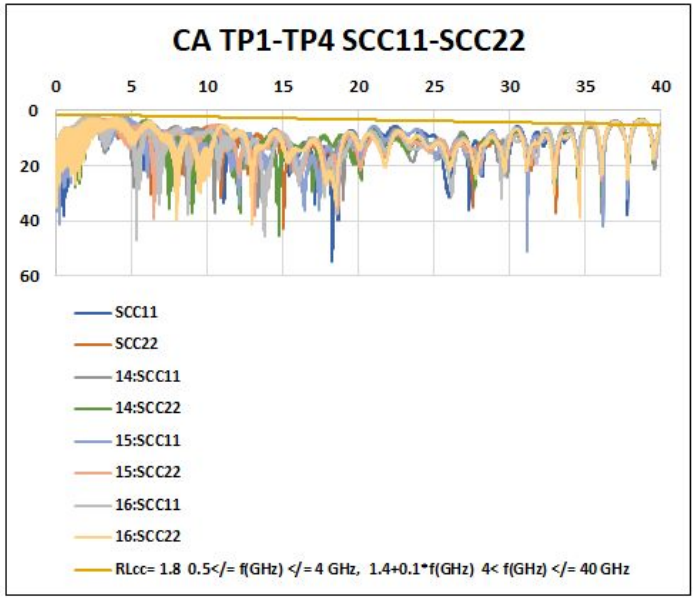


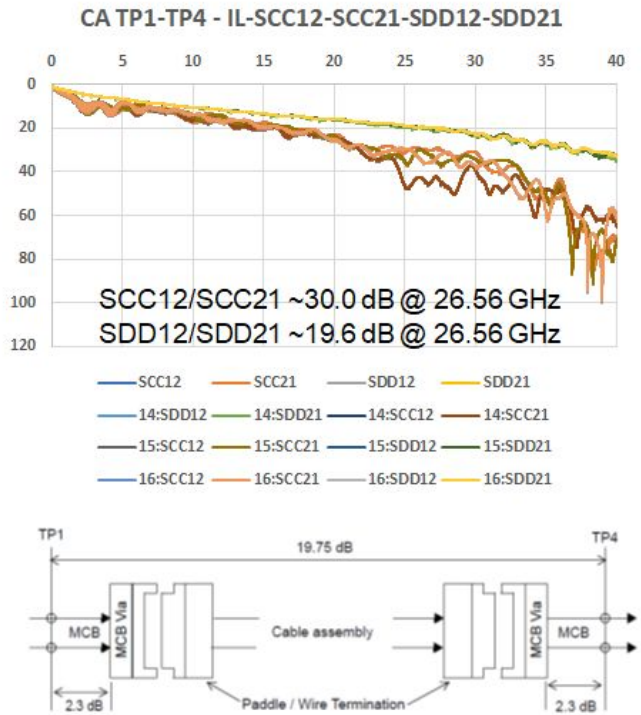


https://www.ieee802.org/3/ck/public/tools/cucable/matoglu_3ck_adhoc_01_030420_channels.zip

Comment 94 and 79

100G CR Cu Cable OSFP channels



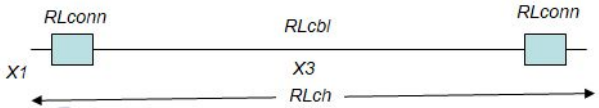


https://www.ieee802.org/3/ck/public/tools/cucable/kocsis_3ck_01_1119_channels.zip

Comment 94 and 79

Return Loss Limit Line Equations for Channel (scaled)

- **Measurements of 2 m CA (TP1-TP4) presented**
- **Measurements of CA with < 2 m of same cable will exhibit worse RL scaling with cable insertion loss**



Includes Phase Effects of connectors

$$RLcnx_k := -20 \cdot \log\left[\left|\exp\left(-2 \cdot \gamma_k \cdot x_1\right) \cdot 10^{-\left(\frac{RLconn_k}{20}\right)} + \exp\left(-2 \cdot \gamma_k \cdot x_3\right) \cdot 10^{-\frac{RLconn_k}{20}}\right|\right]$$

$$RLch_k := -10 \cdot \log\left[10^{-\left(\frac{RLcbl_k}{10}\right)} + 10^{-\left(\frac{RLcnx_k}{10}\right)}\right]$$

Envelope

$$RLcne_k := -20 \cdot \log\left[\left|\exp\left(-2 \cdot \alpha_{1k} \cdot x_1\right) \cdot 10^{-\left(\frac{RLconn_k}{20}\right)} + \exp\left(-2 \cdot \alpha_{3k} \cdot x_3\right) \cdot 10^{-\frac{RLconn_k}{20}}\right|\right]$$

$$RLche_k := -10 \cdot \log\left[10^{-\left(\frac{RLcbl_k}{10}\right)} + 10^{-\left(\frac{RLcne_k}{10}\right)}\right]$$

https://www.ieee802.org/3/bq/public/nov13/diminico_3bq_01_1113.pdf

Comment 108

Cl **162A** *SC* **162A.4** *P* **273** *L* **40** # 108

Dawe, Piers — Nvidia

Comment Type **T** *Comment Status* **D** *host PCB IL*

The recommended minimum insertion loss allocation for the transmitter or receiver differential controlled impedance PCBs, 2.3 dB, has been set the same as the 2.3 dB MCB PCB IL without evidence as to what happens with less loss. 2.3 dB is 1/3 of the maximum host trace loss (6.875 dB) which is too small a ratio to lay out a switch PCB. 92A.4 and 136A.4 use a ratio of 1/5.8 which allows more flexibility in host layout than 1/3 does. 120G has host insertion loss up to 11.9 dB (11.9/2.3 = 5.2/1, which is OK. If it wasn't wanted, the C2M max loss would not have been increased as it was).

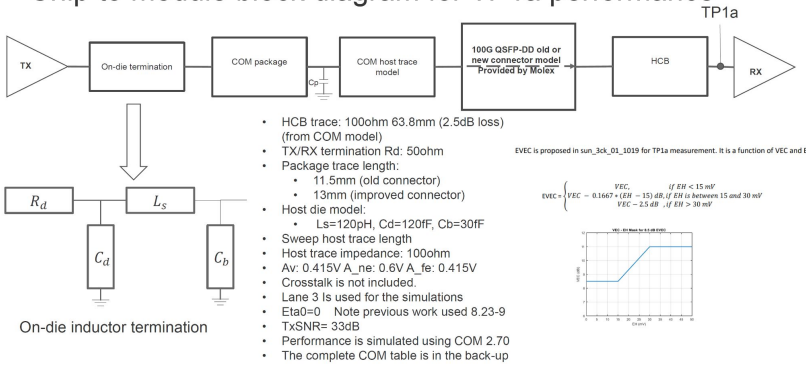
SuggestedRemedy

Reduce the recommended minimum insertion loss allocation for the CR transmitter or receiver differential controlled impedance PCBs to whatever is justified. If the reasonable limit is a strong function of host package reflection, state whether the recommendation is for a "nominal worst" package, or what. If there is no justification, remove the recommendation.

Proposed Response *Response Status* **W**

PROPOSED REJECT.
See comment response #180 D2.0 Slides 4 and 5 of the following presentation were reviewed by the task force:
https://www.ieee802.org/3/ck/public/adhoc/apr28_21/dawe_3ck_adhoc_01_042821.pdf
Slide 3 of the following presentation were reviewed by the task force:
https://www.ieee802.org/3/ck/public/21_05/diminico_3ck_04b_0521.pdf
The IL pcb min and max are derived on the basis of PCB material IL and via IL . The PCB IL assumed is 1.24 dB/in and via of 0.68 dB @26.56 GHz. With consideration for maintaining reasonable minimum length while allowing loss between TX and connector.
Ilpcb(min)=(0.76 in*1.24 dB/in)+(2*0.68) dB = ~ 2.3 dB.
The MCB PCB IL is the same to emulate min host IL.
Lower loss hosts can also create poor performance due to reflections see the following presenation: https://www.ieee802.org/3/ck/public/20_01/dudek_3ck_01_0120.pdf

Chip to module block diagram for TP1a performance



dudek_3ck_01_0120.pdf

Comment 64,57

Cl **162C** *SC* **162C.1** *P* **290** *L* **20** # 64

Ghiasi, Ali — Ghiasi Quantum/Inphi

Comment Type **TR** *Comment Status* **D** *MDI names*

Table 162C-1 should be updated with MDI that actually operate at 53.1 GBd, currenlty what is specified are MDIs that either operate at 10.3 GBd or 25.78 GBd

SuggestedRemedy

Please replace SFP+ with SFP112
http://sfp-dd.com
SFP-DD with SFP-DD112
http://sfp-dd.com
QSFP+ with QSFP112 for reference see
http://www.qsfp-dd.com/wp-content/uploads/2021/05/QSFP-DD-Hardware-Rev6.01.pdf

Proposed Response *Response Status* **W**

PROPOSED REJECT.

This comment does not apply to the substantive changes between IEEE P802.3ck D2.1 and D2.0 or the unsatisfied negative comments from the initial ballot.
Hence it is not within the scope of the recirculation ballot.

This is a restatement of comment D2.0 comment #45 with some additional information.
Comment #57 is requesting similar changes in Annex 162D.
MDI names align with 1.3 normative references in 802.3ck and the base standard.

Cl **162D** *SC* **162D.1** *P* **302** *L* **21** # 57

Ghiasi, Ali — Ghiasi Quantum/Inphi

Comment Type **TR** *Comment Status* **D** *MDI names*

Table 162D-1, 162D-2, 162D-3, and 162D-4 should be updated with MDI that actually operate at 53.1 GBd, currenlty what is specified are MDIs that either operate at 10.3 GBd or 25.78 GBd

SuggestedRemedy

Please replace SFP+ with SFP112
http://sfp-dd.com
SFP-DD with SFP-DD112
http://sfp-dd.com
QSFP+ with QSFP112 for reference see
http://www.qsfp-dd.com/wp-content/uploads/2021/05/QSFP-DD-Hardware-Rev6.01.pdf

Proposed Response *Response Status* **W**

PROPOSED REJECT.

This comment does not apply to the substantive changes between IEEE P802.3ck D2.1 and D2.0 or the unsatisfied negative comments from the initial ballot.
Hence it is not within the scope of the recirculation ballot.

Comment #57 is requesting similar changes in Annex 162C.
Resolve using the response to comment #64

Comment 63

Cl **162C**	*SC* **162C.1**	*P* **292**	*L* **5**	# 63

Ghiasi, Ali — Ghiasi Quantum/Inphi

Comment Type **TR** *Comment Status* **D** *MDI pins*

The pin map for Table 162C-3 is all messed up

SuggestedRemedy

I will include pin maps for all the MDI connectors in the ghiasi_3ck_02_0721

Proposed Response *Response Status* **W**

PROPOSED REJECT.

This comment does not apply to the substantive changes between IEEE P802.3ck D2.1 and D2.0 or the unsatisfied negative comments from the initial ballot.
Hence it is not within the scope of the recirculation ballot.

The suggested remedy does not provided sufficient information to make changes to the draft.
For task force review of the following presentation:
https://www.ieee802.org/3/ck/public/21_07/ghiasi_3ck_02_0721.pdf