802.3df D3.0
Comment Resolution

P802.3df editorial team
Introduction

- This slide package was assembled by the 802.3df editorial team to provide background and detailed resolutions to aid in comment resolution.
Clause 167
Clause 167, parameter labels
Comments 13 and 14

167.8.2 Multi-lane testing considerations

Receiver sensitivity and stressed receiver sensitivity are defined for an interface at the BER specified in 167.1.1. The interface BER is the average of the BERs of the receive lanes when they are stressed. Measurements with Pattern 3 (PRBS31Q) allow lane-by-lane BER measurements. Measurement with Pattern 5 (scrambled idle encoded by RS-FEC) gives the interface BER if all lanes are stressed at the same time.

If each lane is stressed in turn, the BER is diluted by the unstressed lanes, and the BER for that stressed lane alone is found, e.g., by multiplying by four for 400GBASE-SR4 if the unstressed lanes have low BER. In stressed receiver sensitivity measurements, unstressed lanes may be created by setting the power at the receiver under test well above its sensitivity and/or not stressing those lanes with ISI and jitter, or by other means. Each receive lane is stressed in turn while all are operated. All aggressor lanes are operated as specified. To find the interface BER, the BERs of all lanes when stressed are averaged.

Understanding that the interface BER is the average of the BERs, these comments should be rejected and the “each lane” label not added.
Clause 169
Clause 169, delay wording (part 1)
Comments  52, 99, 100, 101

The text says that bit time and pause quanta are "for 800 Gigabit Ethernet".
The title of Table 169-4 has "800GBASE", and footnotes a and b start with "For 800GBASE-R".

Although 800GBASE-R is currently the only defined PHY family, it may not be so in the future; bit time and pause quanta are independent of the PHY type, so the footnotes should not be restricted to one PHY family.

Note that the addition of such footnotes started in Clause 80 in which there were two data rates, so it was required. It isn't required in clauses that define a single data rate, such as Clause 105. If it is anticipated that Clause 169 also introduces 1.6 Terabit Ethernet, then the distinction will be required; otherwise, the data rate can be removed from the footnotes.

The table title should be consistent with the text.

Suggested Remedy
In the table title, change "800GBASE" to "800 Gigabit Ethernet".

In footnotes a and b, either change "For 800GBASE-R" to "For 800 Gigabit Ethernet", or delete these words.

Proposed Response Response Status W
PROPOSED ACCEPT IN PRINCIPLE.
Clause 169 is specific to 800 Gigabit Ethernet or 800GBASE Physical Layer implementations, so the qualifier "(800GBASE)" in the table title is not necessary. The bit times and pause quanta are relevant to any 800GBASE Physical Layer implementation, so the related footnotes should not be specific to 800GBASE-R. However, it is helpful to be unambiguous that these numbers are specific to 800GBASE in general.

In Table 169-4 title delete "(800GBASE)".
In Table 169-4 footnotes 1 and 2, change "800GBASE-R" to "800BASE".

This text "Predictable operation of the MAC Control PAUSE operation ... concatenation of devices " looks like it was copied from 24.6 (for 100BASE-X) when a MAC bit was about 2 m long, the largest nominal reach was 2 km (1000 bits on the line) and there were repeaters. At 800G, a MAC bit is 0.25 mm long and we expect 40 km in P802.3ad (16 bit) bits on the line, 200,000 ns. So the medium can dominate, and one should not expect all PAUSE implementations to tolerate such long links. And, no one talks about repeaters now.

In the proposed change, the NOTE is copied from earlier clauses.

Suggested Remedy
Update and simplify this text, e.g. "The delay limits for each sublayer are relevant to the MAC Control PAUSE operation (Clause 31, Annex 31B)."

NOTE—The physical medium interconnecting two PHYs introduces additional delay in a link.

Proposed Response Response Status W
PROPOSED ACCEPT IN PRINCIPLE.
The introduction as written is relevant and should not be pared down. The delays specified for the backplane (KR8) and copper cable (CR9) PMDs include allocation for the medium (14 ns or ~3 m) between Physical Layers. However, for delays specified for optical PMDs (VR8, SR8, DR8, DR8-2) include only 2 m (~10 ns) allocation for the medium between Physical Layers.
Add the following sentence after Table 169-4:
"The physical medium interconnecting two optical PHYs introduces additional delay in a link."
Clause 169, delay wording (part 2)

Comments 52, 99, 100, 101

---

**Clause 169, SC 169.4, P 182, L 16**

**Comment Type:** T

**Comment Status:** D

**Dawe, Piers J G** NVIDIA

Instead of "colocated", Clause 45 uses terminology like "instantiated within the same package" and "the definition of the term package is vendor specific and could be a chip, module, or other similar entity." We should use language consistent with Clause 45 if it is the same concept, as it appears to be. I suppose the key here could be whether the sublayers are the responsibilities of different parties or whether the interface between the sublayers is accessible for measurement. Also, this uses the spelling "colocated" (twice) while the base document uses "co-located" (twice in 55B). Spelling should be consistent.

**Suggested Remedy**

Change the criterion to say that the delay for the sublayers within a single implementation, which might be a PCB, package, chip or module, is constrained by the sum of constraints for all of the sublayers within it.

If the word "colocated" is kept, reconcile the spelling with the base document.

**Proposed Response**

**Response Status:** W

**PROPOSED ACCEPT IN PRINCIPLE**

The use of the word "colocated" was (incorrectly) intended to make use of established terminology for this situation. Instead, the term "in the same package" is used for this purpose. See examples in 802.3ck-2022 120.1.4, 135.1.4, 162.9.0.4, 162.9.5.1, 163.9.2, 163.9.3.1, 120F.3.1, 120G.3.1 and 802.3-2022 45.2 (many instances).

Change: "The delay for a set of colocated sublayers may be constrained by the sum of constraints for all of the colocated sublayers."

To: "The delay for a set of sublayers within the same package may be constrained by the sum of constraints for the set of sublayers."
Clause 169, delay wording (part 3)

“Package” examples

Each MMD contains registers 5 and 6, as defined in Table 45-2. Bits read as a one in this register indicate which MMDs are instantiated within the same package as the MMD being accessed. Bit 5.0 is used to indicate that Clause 22 functionality has been implemented within a Clause 45 electrical interface device. Bit 6.13 indicates that Clause 22 functionality is extended using the Clause 45 electrical interface through MMD 29. The definition of the term package is vendor specific and could be a chip, module, or other similar entity.

45.2.1 PMA/PMD registers

For devices operating at 25 Gb/s or higher speeds, the PMA may be instantiated as multiple sublayers (see 83.1.4, 109.1.4, and 120.1.4 for how MMD addresses are allocated to multiple PMA sublayers for the respective speeds). A PMA sublayer that is packaged with the PMD is addressed as MMD 1. More addressable instances of PMA sublayers, each one separated from lower addressable instances, may be implemented and addressed as MMD 8, 9, 10, and 11 where MMD 8 is the closest to the PMD and MMD 11 is the farthest from the PMD. The addresses and functions of all registers in MMD 8, 9, 10, and 11 are defined identically to MMD 1, except registers m.5 and m.6 as defined in Table 45-2.
Clause 169, delay wording (part 4)
Proposed results

169.4 Delay constraints

Predictable operation of the MAC Control PAUSE operation (Clause 31, Annex 31B) demands that there be an upper bound on the propagation delays through the network. This implies that MAC, MAC Control sublayer, and PHY implementations conform to certain delay maxima, and that network planners and administrators conform to constraints regarding the cable topology and concatenation of devices. Table 169–4 contains the values of maximum sublayer delay for each instance of a sublayer (sum of transmit and receive delays at one end of the link) in bit times as specified in 1.4 and pause quantum as specified in 31B.2 for 800 Gigabit Ethernet. If a PHY contains an Auto-Negotiation sublayer, the delay of the Auto-Negotiation sublayer is included within the delay of the PMD and medium. The delay for a set of colocated sublayers within the same package may be constrained by the sum of constraints for all of the colocated sublayers.

The physical medium interconnecting two optical PHYs introduces additional delay in a link.

See 80.4 for the calculation of bit time per meter of fiber or electrical cable.

See 31B.3.7 for PAUSE reaction timing constraints for stations at operating speeds of 800 Gb/s.

<table>
<thead>
<tr>
<th>Sublayer</th>
<th>Maximum (bit time)</th>
<th>Maximum (pause quantum)</th>
<th>Maximum (us)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>800G MAC, RS, and MAC Control</td>
<td>196 808</td>
<td>384</td>
<td>245.76</td>
<td>See 170.1.4.</td>
</tr>
<tr>
<td>800GBASE-R PCS or 800GX2S</td>
<td>640 000</td>
<td>1250</td>
<td>800</td>
<td>See 172.5.</td>
</tr>
<tr>
<td>800GBASE-R PMA</td>
<td>36 864</td>
<td>72</td>
<td>46.08</td>
<td>See 173.5.4.</td>
</tr>
<tr>
<td>800GBASE-ER PMD</td>
<td>32 768</td>
<td>64</td>
<td>40.96</td>
<td>Includes allocation of 14 ms for one direction through backplane medium. See 163.5.</td>
</tr>
<tr>
<td>800GBASE-ER PMD</td>
<td>32 768</td>
<td>64</td>
<td>40.96</td>
<td>Includes allocation of 14 ms for one direction through cable medium. See 162.5.</td>
</tr>
<tr>
<td>800GBASE-VR3 PMD</td>
<td>32 768</td>
<td>64</td>
<td>40.96</td>
<td>Includes 2 m of fiber. See 167.3.1.</td>
</tr>
<tr>
<td>800GBASE-SR5 PMD</td>
<td>32 768</td>
<td>64</td>
<td>40.96</td>
<td>Includes 2 m of fiber. See 167.3.1.</td>
</tr>
<tr>
<td>800GBASE-DR5 PMD</td>
<td>32 768</td>
<td>64</td>
<td>40.96</td>
<td>Includes 2 m of fiber. See 124.3.1.</td>
</tr>
<tr>
<td>800GBASE-DR3-2 PMD</td>
<td>32 768</td>
<td>64</td>
<td>40.96</td>
<td>Includes 2 m of fiber. See 124.3.1.</td>
</tr>
</tbody>
</table>

* For 800GBASE, 1 bit time (BT) is equal to 1.25 ps. (See 14.2.15 for the definition of bit time.)
* For 800GBASE, 1 pause quantum is equal to 640 ps. (See 31B.2 for the definition of pause quantum.)
* Should there be a discrepancy between this table and the delay requirements of the relevant sublayer clause, the sublayer clause prevails.
* If an implementation includes the 800GMII Extender, the delay associated with the 800GMII Extender includes two 800GXS sublayers.
Clause 172
Comments # 110, 62 propose changes to 172.2.4.5

172.2.4.5 Scrambler

The scrambler in each flow is identical to that specified in 119.2.4.3. Although there is no requirement on the initial value of each scrambler, if an implementation sets the scrambler state to a fixed value (e.g., when reset is asserted), the two scramblers should be set to different states.

Cl: 172 SC: 172.2.4.5 P: 212 L: 19 #: I-62

Ran, Adee
Cisco Systems, Inc.

Comments # 110, 62 propose changes to 172.2.4.5

SuggestedRemedy

The recommendation to "set to different states" deserves further explanation.

Cl: 172 SC: 172.2.4.5 P: 212 L: 19 #: I-110

Dawe, Piers J G
NVIDIA

Comment Type: TR

The two scramblers should be set to different states": this is too weak, and readers do not understand the importance of this. The consequence of getting it wrong is much more than the bad spectrum or correlation issues we have seen elsewhere.

SuggestedRemedy

Change should to shall or is.
Add a sentence: This is because before the link can carry traffic, the 66-bit blocks in the two flows have the same content.

Proposed Response

PROPOSED ACCEPT IN PRINCIPLE.

Resolve using the response to comment #I-110.
Scrambler, 172.2.4.5
Comments 110, 62

Options to resolve comments 110 and 62:

A. Resolve using suggested remedy in comment 62 (add informative note only)

B. Resolve using suggested remedy in comment 110 (change “should” to “shall” or “is” and add a note)
Clause 173
Clause 173, skew wording (part 1)
Comments 69, 128

---

Ran, Adee
Cisco Systems, Inc.

Comment Type: E
Comment Status: D

"shall produce" here, "shall generate" in 173.5.3.3, "shall deliver" in 173.5.3.5... the title of all three has "skew generation".

In fact, the skew numbers stated are cumulative.

Since the skew at any point is not necessarily generated at that point, the proper requirement seems to be "shall have".

Suggested Remedy
Change all three "shall" statements in the comment to "shall have".

Proposed Response
PROPOSED ACCEPT IN PRINCIPLE.
Some changes to the wording would be an improvement to the draft. Appropriate change will be provided in a supporting presentation.

---

Dawe, Piers J G
NVIDIA

Comment Type: TR
Comment Status: D

In these subclauses, skew is generated, produced or delivered. It is not clear what these terms mean. I believe that all Skew limits are cumulative (unlike for delay) which has a bearing on what the terms mean.

Suggested Remedy
Write down what generated, produced and delivered mean here and what the differences are.

Proposed Response
PROPOSED ACCEPT IN PRINCIPLE.
Resolve using the response to comment #69.
Clause 173, skew wording (part 2)
Comments 69, 128

From IEEE Std 802.3df D3.0…

173.5.3.1 Skew generation toward SP1

In an implementation with one or more physically instantiated 800GAIU-8 interfaces, the PMA that sends data in the transmit direction toward the 800GAIU-8 that is closest to the PMD (SP1 in Figure 169-4 and Figure 169–5) shall produce no more than 16 ns of Skew between PCSLs toward the 800GAIU-8 and no more than 200 ps of Skew Variation.

173.5.3.3 Skew generation toward SP2

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, the PMA adjacent to the PMD service interface shall generate no more than 25 ns of Skew, and no more than 400 ps Skew Variation.

173.5.3.5 Skew generation at SP5

In an implementation with one or more physically instantiated 800GAIU-8 interfaces, at SP6 (the receive direction of the 800GAIU-8 closest to the PCS), the PMA or group of PMAs between the PMD and the 800GAIU-8 closest to the PCS shall deliver no more than 145 ns of Skew, and no more than 3.8 ns of Skew Variation between output lanes toward the 800GAIU-8 in the receive direction.

Observations:
1) As Adee points out the wording is inconsistent between 173.5.3.1 ("shall produce"), 173.5.3.3 ("shall generate") and 173.5.3.5 ("shall deliver").
2) The current wording is consistent with Clause 81 (3ba, circa 2010) and Clause 120 (3bs)
3) The current wording could be interpreted as being for just the skew introduced by the PMA itself.
4) The specification is actually for the skew at the output of the PMA, including skew introduced by the PMA itself and skew introduced by other sublayers/interfaces above the PMA.
Clause 173, skew wording (part 3)
Comments 69, 128

Change 173.5.3.1 as follows:

173.5.3.1 Skew generation toward SP1

In an implementation with one or more physically instantiated 800GAUI-8 interfaces, at the output of the PMA that sends data in the transmit direction toward the 800GAUI-8 that is closest to the PMD (SP1 in Figure 169–4 and Figure 169–5) there shall produce be no more than 16 ns of Skew between PCSLs toward the 800GAUI-8 and no more than 200 ps of Skew Variation between output lanes toward the 800GAUI-8.

Change 173.5.3.3 as follows:

173.5.3.3 Skew generation toward SP2

If the PMD service interface is physically instantiated so that the Skew at SP2 can be measured, at the output of the PMA adjacent to the PMD service interface there shall generate be no more than 25 ns of Skew between PCSLs, and no more than 400 ps of Skew Variation between output lanes toward the PMD service interface (SP2 in Figure 169–4 and Figure 169–5). In an implementation with one or more physically instantiated 800GAUI-8 interfaces, the Skew measured at the input to the PMA adjacent to the PMD service interface (SP1 in Figure 169–4 and Figure 169–5) is limited to no more than 16 ns of Skew and no more than 200 ps of Skew Variation.
Clause 173, skew wording (part 4)
Comments 69, 128

Change 173.5.3.5 as follows:

173.5.3.5 Skew generation at SP6

In an implementation with one or more physically instantiated 800GAUI-8 interfaces, at SP6 (the receive direction of the 800GAUI-8 closest to the PCS), at the output of the PMA or group of PMAs between the PMD and the 800GAUI-8 closest to the PCS, there shall be no more than 145 ns of Skew between PCSLs, and no more than 3.8 ns of Skew Variation between output lanes toward the 800GAUI-8 in the receive direction.
Clause 173, Test patterns (part 1)
Comments 66, 141

Two comments received relating to how “test pattern generate” and “test pattern check” are represented in the PMA functional block diagram.

Observations:
- Both comments essentially have the same suggested remedy
- The current approach for representing “test pattern generate” and “test pattern check” has been in force since 3bs (circa 2010)
- Some updates to the block diagrams to more accurately represent test pattern generate/check may be warranted, and make the diagrams more interpretable
- Proposed updates for Figure 173-3, Figure 173-4 and Figure 173-5 are provided on the following slides

---

Comment Type T
Comment Status D
The dashed-line arrows in Figure 173-3 are not connected to the right places.

"Test pattern generate" creates bits that are encoded as PAM4 symbols and then driven by the same signal drivers. It should go into the "PAM4 encode/Signal drivers" box.

"Test pattern check" operates on a bit stream, so should take the output of "PAM4 encode/CDR".

The arrow leading to "SIL" denotes information from the CDR. It should be taken from the PAM4 decoder/CDR box.

Similarly in Figure 173-4 and Figure 173-5.

Suggested Remedy
Modified figures will be supplied

Proposed Response  Response Status W
PROPOSED ACCEPT IN PRINCIPLE.
Some updates to the diagram are warranted.
Slides will be provided to address this.
Clause 173, Test patterns (part 2)
Comments 66, 141

802.3df D3.0 - Figure 173-3

Proposed update to Figure 173-3
Clause 173, Test patterns (part 3)
Comments 66, 141

802.3df D3.0 - Figure 173-4

Proposed update to Figure 173-4
Clause 173, Test patterns (part 4)
Comments 66, 141

802.3df D3.0 - Figure 173-5

Proposed update to Figure 173-5
Clause 173, signal status (part 1)
Comment i-72

The highlighted statement has some important significance, but a few extra words would make it more interpretable. Also, a similar statement should be reinstated the 8:32 PMA transmit direction. See next slide.

<table>
<thead>
<tr>
<th>CI</th>
<th>SC</th>
<th>P</th>
<th>L</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>173</td>
<td>173.5.8.1</td>
<td>242</td>
<td>3</td>
<td>i-72</td>
</tr>
</tbody>
</table>

Ran, Adee
Cisco Systems, Inc.

Comment Type T
Comment Status D
signal status (bucket2)

The requirement that "data is being sent on all 32 output lanes (PMA:IS_UNITDATA_0:31 indication)" is unique to this PMA (32:8); the other two PMAs set the signal status only based on data being received on the appropriate interface.

In real implementations, an indication to the PCS that data is not being received by the PMA (which may be due to lack of a link partner) would likely be separate from an indication that data is not being transmitted (essentially a local fault). Specifying in the standard that it's the same indication is not helpful for readers.

Suggested Remedy
Delete the second item in the list.
Consider converting the list to regular paragraph text as in the other two subclauses.

Proposed Response
PROPOSED ACCEPT IN PRINCIPLE.

Change the SIGNAL_OK definition to the following:
"The SIGNAL_OK parameter is set to OK when data is being received on all 8 input lanes (inst.IS_UNITDATA_0:7 indication) and the SIGNAL_OK parameter of the PMD:IS_SIGNAL.indication primitive is set to OK, if there is a PMD immediately below the PMA. Otherwise SIGNAL_OK is set to FAIL."

### 173.5.8 Signal status

#### 173.5.8.1 32:8 PMA signal status

The 32:8 PMA provides signal status information to the PMA client (800BASE-R PCS or DTE 800GX5) using the PMA:IS_SIGNAL.indication(SIGNAL_OK) service interface primitive (see 173.2 and Figure 173-3).

The SIGNAL_OK parameter is set to OK when all of the following conditions are met:
- data is being received on all 8 input lanes (inst.IS_UNITDATA_0:7.indication)
- data is being sent on all 32 output lanes (PMA:IS_UNITDATA_0:31 indication)
- the SIGNAL_OK parameter of the PMD:IS_SIGNAL.indication primitive is set to OK, if there is a PMD immediately below the PMA.

Otherwise SIGNAL_OK is set to FAIL.

#### 173.5.8.2 8:32 PMA signal status

In the transmit direction the 8:32 PMA provides signal status information to the PHY 800GX5 using the PHY:XS:IS_SIGNAL.request(SIGNAL_OK) service interface primitive (see 173.3 and Figure 173-4). The SIGNAL_OK parameter is set to OK when data is being received on all 8 input lanes (PMA:IS_UNITDATA_0:7.request). Otherwise SIGNAL_OK is set to FAIL.

In the receive direction the 8:32 PMA optionally provides signal status information to the client PMA by disabling one or more output lanes (PMA:IS_UNITDATA_0:7.request) when the PHY:XS:IS_SIGNAL.indication SIGNAL_OK parameter (see 173.3 and Figure 173-4) is FAIL.

#### 173.5.8.3 8:8 PMA signal status

In the transmit direction the 8:8 PMA optionally provides signal status to the sublayer below by disabling one or more output lanes (PMA:IS_UNITDATA_0:7.request) when data is not being received on all 8 input lanes (PMA:IS_UNITDATA_0:7.request).

In the receive direction the 8:8 PMA optionally provides signal status to the client PMA by disabling one or more output lanes (PMA:IS_UNITDATA_0:7.indication) when the PMD:IS_SIGNAL.indication SIGNAL_OK parameter (see 173.3 and Figure 173-4) is FAIL or when data is not being received on all 8 input lanes (inst.IS_UNITDATA_0:7.indication).
Clause 173, signal status (part 2)

Comment i-72

173.5.8 Signal status

173.5.8.1 32:8 PMA signal status

In the receive direction, the 32:8 PMA provides signal status information to the PMA client (800GBASE-R PCS or DTE 800GXS) using the PMA:IS_SIGNAL.indication(SIGNAL_OK) service interface primitive (see 173.2 and Figure 173–3): The SIGNAL_OK parameter is set to OK when all of the following conditions are met:

- data is being received on all 8 input lanes (inst:IS_UNITDATA_0:7.indication)
- the received data is being sent on all 32 output lanes (PMA:IS_UNITDATA_0:31.indication)
- the SIGNAL_OK parameter of the PMD:IS_SIGNAL.indication primitive is set to OK, if there is a PMD immediately below the PMA

Otherwise SIGNAL OK is set to FAIL.

173.5.8.2 8:32 PMA signal status

In the transmit direction the 8:32 PMA provides signal status information to the PHY 800GXS using the PHY_XS:IS_SIGNAL.request(SIGNAL_OK) service interface primitive (see 173.3 and Figure 173–4). The SIGNAL_OK parameter is set to OK when data is being received on all 8 input lanes (PMA:IS_UNITDATA_0:7.request) and the received data is being sent on all 32 output lanes (PHY_XS:IS_UNITDATA_0:31.request). Otherwise SIGNAL OK is set to FAIL.

In the receive direction the 8:32 PMA optionally provides signal status information to the client PMA by disabling one or more output lanes (PMA:IS_UNITDATA_0:7.indication) when the PHY_XS:IS_SIGNAL.indication SIGNAL_OK parameter (see 173.3 and Figure 173–4) is FAIL.

173.5.8.3 8:8 PMA signal status

In the transmit direction the 8:8 PMA optionally provides signal status to the sublayer below by disabling one or more output lanes (inst:IS_UNITDATA_0:7.request) when data is not being received on all 8 input lanes (PMA:IS_UNITDATA_0:7.request).

In the receive direction the 8:8 PMA optionally provides signal status to the client PMA by disabling one or more output lanes (PMA:IS_UNITDATA_0:7.indication) when the PMD:IS_SIGNAL.indication SIGNAL_OK parameter (see 173.3 and Figure 173–4) is FAIL or when data is not being received on all 8 input lanes (inst:IS_UNITDATA_0:7.indication).