Considerations related to ppm and interop (Comments #188, #432 against D2.1)

Gary Nicholl - Cisco
Eugene Opsasnick - Broadcom
David Ofelt - Juniper Networks / HPE
Kapil Shrikhande - Marvell

IEEE 802.3dj, Minnesota, September 2025

Overview

- Comments #122 and #432 against 802.3dj D2.1 highlight a potential multi-generational interop issue between a legacy AUI/PMD (running at 25Gb/s, 50Gb/s or to a lesser extent100Gb/s) and a new AUI/PMD (running at 200Gb/s), due to ppm tolerance mismatch.
- This contribution reviews the issue and proposes some updates to the draft.
- Notes:
  - The issue is limited to 200GbE and 400GbE solutions (as these can contain a mix of interfaces operating at 100ppm and 50ppm).
  - There is no issue for 800GbE and 1.6TbE solutions (as all interfaces are required to be 50ppm).
  - It is important to point out that the spec is not broken, but for 200GbE and 400GbE there is a subtle interoperability issue of a sort that we have never introduced previously and highlighting this with an informative note seems appropriate.

Comments #188 and #432

Cl **120** *SC* **120.1.4** *P* **194** *L* **15** # 188

Ofelt, David — Juniper Networks / HPE

Comment Type **TR** *Comment Status* **D** *ppm (L)*

We have changed the ppm tolerance of the 200Gb/s SERDES to be 50ppm in all cases. This leads to interoperability issues when plugging an older PMD (generated with 25Gb/s or 50Gb/s SERDES) into a new 200Gb/s SERDES-based receiver or when a new 802.3dj PMD is plugged into an older box using 25Gb/s or 50Gb/s SERDES due to the fact one end of those links generates data at 100ppm and the receive side can only handle 50ppm. The solution is to insert an XS to do rate matching. At the moment, I believe this interop issue is not called out anywhere in the draft nor is the fact that adding the required XS will also cause the PTP accuracy to suffer. Note that this was not an issue in the 100Gb/s SERDES because they were specified to tolerate 100pm at the receiver, so there were no multi-generational interop issues. This is also not a problem when 100Gb/s source and 200Gb/s sourced PMDs are connected because the 100Gb/s SERDES are specified to have transmitters that are 50ppm.

The set of footnotes in this subclause attempt to provide the full set of rules for managing ppm, but the details are incomplete for the cases mentioned here.

As it stands, the spec is not broken, but this is a subtle interoperability issue of a sort that we've never introduced previously, therefore a helpful note seems appropriate.

SuggestedRemedy

Add some additional informative information to the ppm guideline footnotes in 120.1.4 to clarify the subtle 100/50ppm intereop cases that need an XS as well as a comment that this will degrade PTP accuracy.

A supportning presentation will be forthcoming.

Proposed Response *Response Status* **W**

PROPOSED ACCEPT IN PRINCIPLE.

Pending review of the following presentation and CRG discussion.
<URL>nicholl_3dj_02_2509.pdf

Cl **120** *SC* **120.1.4** *P* **194** *L* **10** # 432

Nicholl, Gary — Cisco Systems

Comment Type **TR** *Comment Status* **D** *ppm (L)*

List items (7) and (9) essentially mean that the only way to support a 200G/400G PHY which includes 200G/lane technology in a legacy host with AUIs running at 100ppm is to use an Extender. The Extender would convert beteen the two ppm rates, allowing the existing AUIs to continue to run at 100ppm and the new 200G/400G PHY to run at 50 ppm.

But the consequence of this is that two types of optical module are required, a simple one which can be used in hosts with AUIs that are running at 50ppm and a more complex one (which includes a PHY XS and PCS) which can be used in legacy hosts where the AUIs are running at 100ppm.

But the question is how does an end user know what rate (50ppm or 100ppm) the AUIs on his host are running at, and therefore which version (simple or complex) of optical module is required ?

List items 7 and 9 essentially create two different versions of 200G/400G AUIs (one running at 50ppm and one running at 100ppm), with no obvious way to identify the different versions.

SuggestedRemedy

A presentation will be provided to further discuss the issue and provide some possible solutions.

Proposed Response *Response Status* **W**

PROPOSED ACCEPT IN PRINCIPLE.

Resolve using the response to comment #188.

Problem Statement (from comment #188)

- In 802.3dj we changed the ppm tolerance of the 200Gb/s serdes to be 50ppm in all cases
- This causes interoperability issues in two scenarios:
  - Plugging a legacy PMD (based on 25Gb/s, 50Gb/s or 100Gb/s lanes operating at 100ppm) into a new host with a 200Gb/s receiver (that can only tolerate 50ppm), e.g. a 400GBASE-DR4 module with a 400GAUI-2 electrical interface
  - Plugging a new 802.3dj PMD (based on 200Gb/s) into a legacy host (based on 25Gb/s, 50Gb/s or 100Gb/s lanes operating at 100ppm), e.g. 400GBASE-DR2 module with a 400GAUI-8 electrical interface
  - In both these cases the issue is that one end of the link generates data at 100ppm, whereas the receive side can only handle 50ppm
- The solution in both cases is to insert an XS to do the rate matching
  - But this will degrade PTP accuracy
- As it stands the spec is not technically broken (can always insert an XS), but this is a subtle interoperability issue of a sort that we have never introduced previously, and so multi-generational solutions are not as “plug’n’play” as they were in the past.
- A set of statements in 120.1.4 attempt to provide the full set of rules for managing ppm, but the details are incomplete for the two cases identified above
- Recommendation:
  - Add additional informative information to the ppm guidelines in 120.1.4 to clarify the subtle 100/50ppm interoperability cases that need an XS, as well as a comment that this will degrade PTP accuracy

Subclause 120.1.4

4) 200GAUI-2 is a 53.125 GBd by 2 lane PAM4 physical instantiation of the 200 Gb/s connection. 400GAUI-4 is a 53.125 GBd by 4 lane PAM4 physical instantiation of the 400 Gb/s connection.

5) For a PHY that includes a 200GAUI-2 interface, it is recommended that the signaling rate range for a 200GAUI-8 or 200GAUI-4 PMA output that is in the same package as the PCS be limited to 26.5625 GBd ± 50 ppm.

6) For a PHY that includes a 400GAUI-4 interface, it is recommended that the signaling rate range for a 400GAUI-16 or 400GAUI-8 PMA output that is in the same package as the PCS be limited to 26.5625 GBd ± 50 ppm.

7) For a PHY that includes a 200GAUI-1 interface or a 200GBASE-KR1, 200GBASE-CR1, 200GBASE-DR1, or 200GBASE-DR1-2 PMD, the signaling rate range for a 200GAUI-8, 200GAUI-4, or 200GAUI-2 PMA output that is in the same package as the PCS shall be limited to ± 50 ppm, instead of ± 100 ppm.

8) For a 200GMII Extender that includes a 200GAUI-1 interface, the signaling rate range for a 200GAUI-8, 200GAUI-4, or 200GAUI-2 PMA output that is in the same package as the DTE 200GXS or PHY 200GXS shall be limited to ± 50 ppm, instead of ± 100 ppm.

9) For a PHY that includes a 400GAUI-2 interface or a 400GBASE-KR2, 400GBASE-CR2, 400GBASE-DR2, or 400GBASE-DR1-2 PMD, the signaling rate range for a 400GAUI-16, 400GAUI-8, or 400GAUI-4 PMA output that is in the same package as the PCS shall be limited to ± 50 ppm, instead of ± 100 ppm.

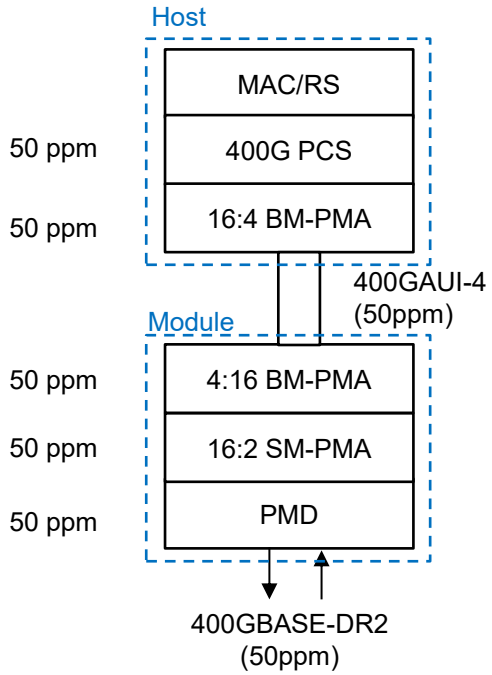
10) For a 400GMII Extender that includes a 400GAUI-2 interface, the signaling rate range for a 400GAUI-16, 400GAUI-8, or 400GAUI-4 PMA output that is in the same package as the DTE 400GXS or PHY 400GXS shall be limited to ± 50 ppm, instead of ± 100 ppm.

- Guidelines 7) and 9) in 120.1.4, state that for any 200GbE or 400GbE PHY which include either a 200Gb/s per lane AUI or PMD (or both), then any associated 25Gb/s, 50Gb/s or 100Gb/s per lane AUI in the same PHY (all the way to the PCS) are limited to outputting at 50ppm.
- This means that an 802.3dj based PHY cannot be directly supported in a legacy host with 25G, 50G or 100G per lane AUIs running at 100ppm. It is not allowed, period.
- This essentially places an indirect requirement that to support an 802.3dj PHY in a legacy platform with AUIs running at 100ppm, the legacy AUIs must be part of an XS (cannot be part of the PHY), and the XS used to rate match between the 100ppm of the XS and the 50ppm of the PHY (see following slides).
- This addresses one of the use cases highlighted by comment #188, but the message is somewhat subtle and indirect and likely should be called out more explicitly.

Note: Issue is in transmit direction

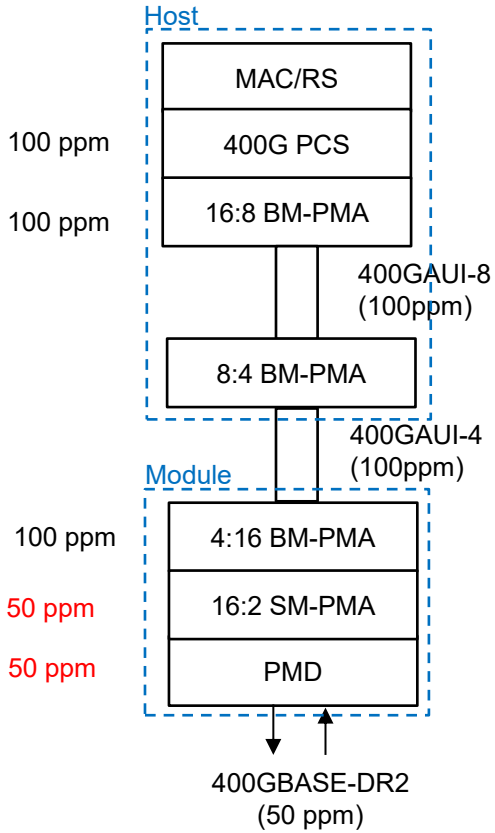
Use Case 1a: ppm interop issue with 200G/lane PMD

Legacy 100G/lane host (50 ppm)

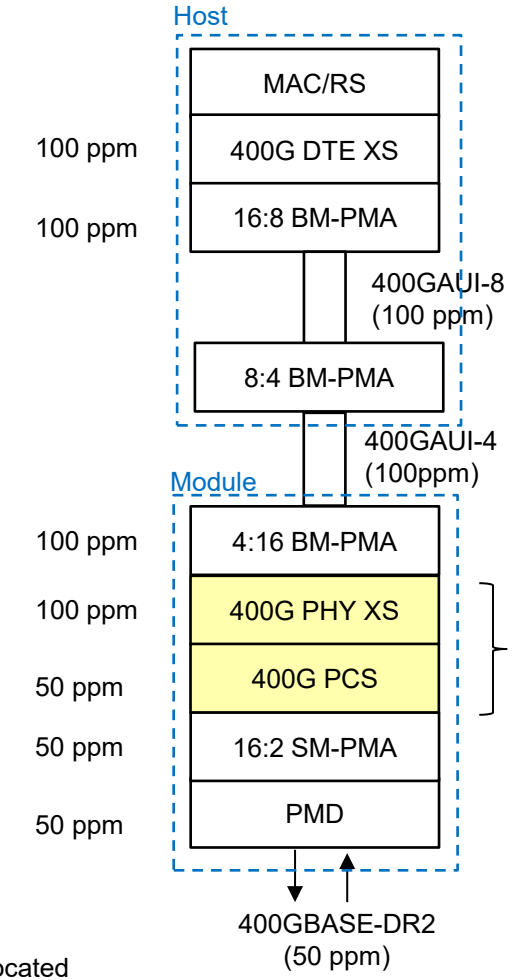


- In this case the 400GAUI-4 is co-located with PCS and required to be 50ppm (802.3ck)
- This likely accounts for the majority of 400GAUI-4 based hosts in the field
- This is stack up is compliant with guideline 9) in 120.1.4, and is allowed.

Legacy 100/lane host (100 ppm)



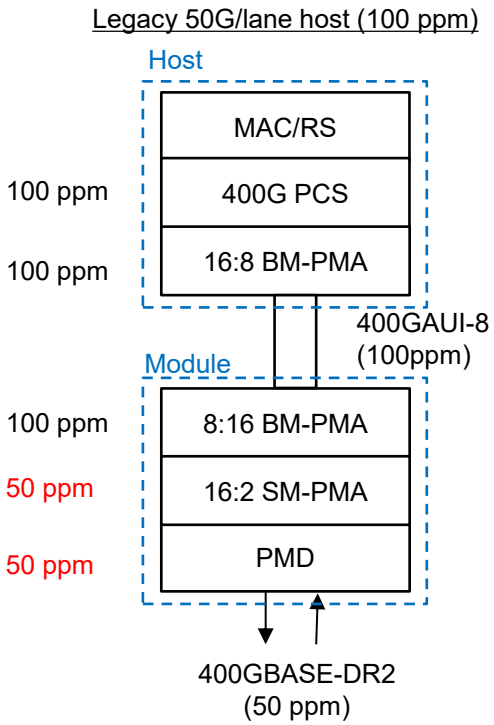
- In this case the 400GAUI-4 is not co-located with the PCS and therefore allowed to run at 100pm (802.3ck)
- This stack up is not compliant with guideline 9) in 120.1.4 and is disallowed



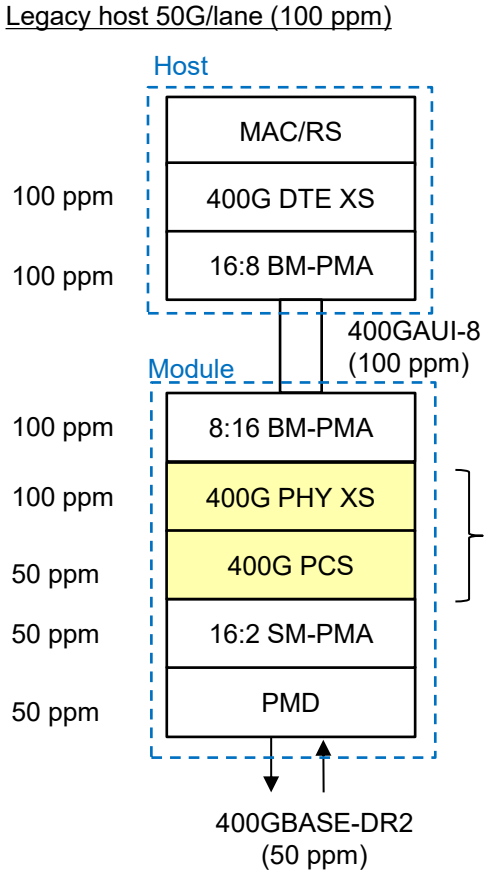
- Must use XS to convert from 100ppm to 50 ppm
- Requires extra functionality in module
- A different optical module is required to interface to a legacy host with a 100ppm 400GAUI-4 versus a legacy host with a 50ppm 400GAUI-4

Note: Issue is in transmit direction

Use Case 1b: ppm interop issue with 200G/lane PMD



- In this case the 400GAUI-8 is always 100ppm
- This stack up is not compliant with guideline 9) in 120.1.4 and is disallowed



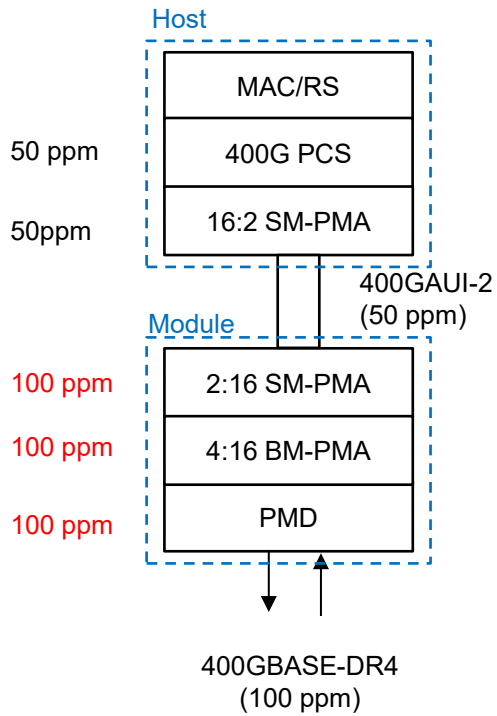
- Must use XS to convert from 100ppm to 50 ppm
- Requires extra functionality in module

- In this case the 400GAUI-8 is always 100ppm
- This stack up is compliant with guideline 9) in 120.1.4 and is allowed, but requires an XS in the module
- At least this time only a single variant of module (with XS) is required

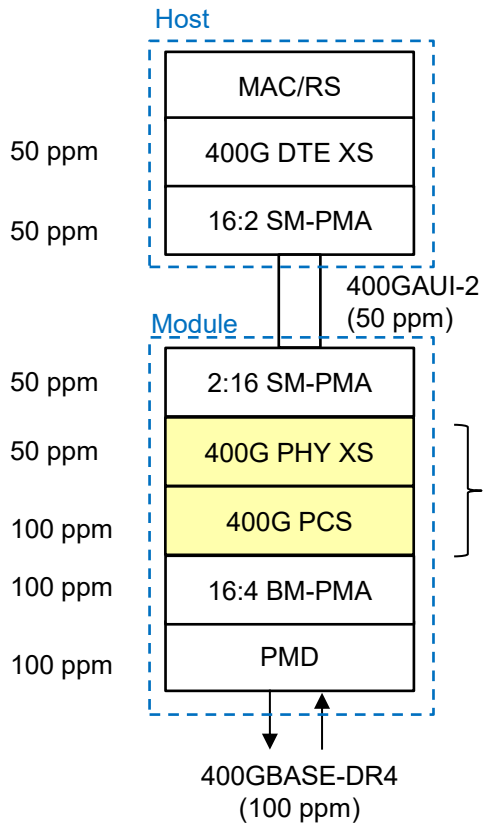
Note: Issue is in receive direction

Use Case 2: ppm interop issue with 200G/lane AUI

New 200G/lane host (50 ppm)



- In this case the 400GBASE-DR4 PMD has to be assumed to be 100 ppm worst case (it could be either 50ppm or 100ppm depending on far end host)
- This stack up does not work because of a ppm mismatch between the PMD (100ppm) and the 400GAUI-2 (only 50pm tolerant) - issue is in the receive direction only.
- Nothing in the current draft that explicitly excludes this stack up



- Must use XS to convert from 100ppm to 50 ppm
- Requires extra functionality in module

- In this case the 400GBASE-DR4 PMD has to be assumed to be 100 ppm worst case (it could be either 50ppm or 100ppm depending on far end host)
- This stack up is fine because ppm match is handled by the XS in the module
- This stack up is supported in the current draft
- Only a single variant of module is needed

Potential solution for Use Case 1

- Update the text of guideline 7) in 120.1.4 as follows:

For a PHY that includes a 200GAUI-1 interface or a 200GBASE-KR1, 200GBASE-CR1, 200GBASE-DR1, or 200GBASE-DR1-2 PMD, the signaling rate range for a 200GAUI-8, 200GAUI-4, or 200GAUI-2 PMA output that is in the same package as the PCS shall be limited to +/- 50ppm, instead of +/-100ppm, otherwise the 200GAUI-8, 200GAUI-4, or 200GAUI-2 shall be implemented within a 200GMII Extender.

- Make a similar update to the text of guideline 9) in 120.1.4

Potential solution for Use Case 2

- Needs further discussion

Summary

- Comments #122 and #432 highlighted a potential, but subtle, multi-generation interoperability issue for 200GbE and 400GbE, between legacy interfaces (either AUI or PMD) based on 25Gb/s, 50Gb/s or 100Gb/s per lane signaling and operating at 100ppm and new interfaces (either AUI or PMD) based on 200Gb/s per lane signaling and operating at 50ppm.
- As it stands, the specification is not broken, but there is a subtle interoperability issue of a sort that we have never introduced previously, and highlighting this with an informative note/text seems to be appropriate.

Thanks