Guidance for partially-populated MDIs and breakout

(comments #86, #87, #107)

Adee Ran, Cisco

Background and goal

- Annex 179C specifies MDIs for CR electrical PHYs
  - 179C.1 It includes a recommendation (“should”) for the case where the PHY does not fully utilize the MDI connector
- In 802.3dj we created Annex 180A to similarly specify MDIs for DR optical PHYs
  - 180A.4.1 and 180A.4.2 include normative requirements (“shall”) for PHYs that do not “support” the full width of the MDI connector
- These recommendations/requirements are intended to ensure interoperability
  - But they do not
- The presentation suggests more common language and better information for the reader

Comments against Annex 180A

Cl **180A** SC **180A.2** P**901** L**29** # 419

Ran, Adee Cisco Systems

Comment Type **TR** *Comment Status* **R** *MDI breakout (O)*

Table 180A-1 (and this whole Annex) are based on the idea that DR modules can be used in a breakout configuration or with multiple PMDs per connector. But this concept is not mentioned.
The sentence "Table 180A-1 shows the number of PMDs supported by each MDI type" is odd - typically an MDI is the interface of a single PMD to its medium, and the term "MDI type" (which is apparently something else) is only used here and has never been defined.
The reader should be informed that having multiple PMDs that share one connector requires proper configuration of the host to match the PMDs with their respective link partners.

SuggestedRemedy

Add a paragraph that describes the concept of an MDI connector (which can include multiple MDIs, depending on the PHY type). This paragraph should not include a requitement from a host to support any possible combination of MDIs.

Change "MDI type" to "MDI connector" (or "MDI receptacle" if it's more suitable) in the text and in the table.

Add cross-references in the first column to 180A.3.1 and 180A.3.2.

Add an informative NOTE about the need to configure the host when multiple PMDs share a connector.

Implement with editorial license.

Response *Response Status* **W**

REJECT.

The suggested remedy does not provide sufficient detail to implement. Significant changes have been agreed for the annex and the commentor is encouraged to review the updated draft.

Comment #419 against D2.1 (unsatisfied)

Cl **180A** SC **180A.4.1** P**939** L**33** # 87

Ran, Adee Cisco Systems

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

The situations described in the text, of MDI connectors that are not fully utilized (some lanes not connected to a PMD) or are used with multiple PMDs (breakout), are not detectable by a link partner that is connected to the other side of the cable plant.

In such situations, the link partner needs to be configured by management to the corresponding PMD combination. This should be noted for readers.

Also in 180A.4.2.

SuggestedRemedy

Add the following informative note:
NOTE—The PMD types on both sides of the fiber need to match. When the MDI is used for multiple PMDs or for PMDs with lower number of lanes than the MDI supports, appropriate configuration is required. The means for selecting the appropriate configuration are beyond the scope of this standard.

Add a similar note in 180A.4.2.
Implement with editorial license.

Follow up comment on the same topic – a detailed proposal.

Comments against Annex 179C

Cl **179C** *SC* **179C.1** *P* **921** *L* **4** # 107

Dudek, Mike — Marvell

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

Annex 180A provides normative requirements for which fibers should be used when connectors are not fully utilized. Whereas for the equivalent situation for CR there is just a "recommendation" with the use of "should"

SuggestedRemedy

Change "When an MDI connector is not fully utilized the lower PMD numbers in Table 179C–2 should be used." to "When an MDI connector is not fully utilized the lower PMD numbers in Table 179C–2 shall be used". Or "When all the lanes of an MDI connector do not have signals connected the lower PMD numbers in Table 179C–2 shall be used.e.g. if a QSFP224 connector is used for a single 400GBASE-CR2 connection then PMD 0 and 1 are used."

The comment suggests making a normative requirement instead of a recommendation in 179C.

Cl **179C** *SC* **179C.1** *P* **921** *L* **3** # 86

Ran, Adee — Cisco Systems

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

"When an MDI connector is not fully utilized the lower PMD numbers in Table 179C–2 should be used"

The MDI is part of the PHY so "not fully utilized" means the host does not have transmit and receiver functions for all lanes of the MDI. This is an unlikely situation, and even if it happens, following the recommendation does not guarantee interoperability, since in most cases the link partner needs to be configured accordingly.

Instead, it would be helpful for readers to know that in some cases, such as breakout cables, the combination of PMDs types on both sides of the cable can require management to create matching configurations

SuggestedRemedy

Delete the quoted sentence.
Add the following informative note:
NOTE—The PMD types on both sides of the cable assembly need to match. When the MDI is used for multiple PMDs or for PMDs with lower number of lanes than the MDI supports, appropriate configuration is required. The means for selecting the appropriate configuration are beyond the scope of this standard.

The comment suggests an informative note as suggested in 180A.
Also, it suggests removing the recommendation since it refers to an “unusual situation”.

What are we talking about

- Both Annex 179C and Annex 180A specify MDIs.

Annex 179C

(normative)

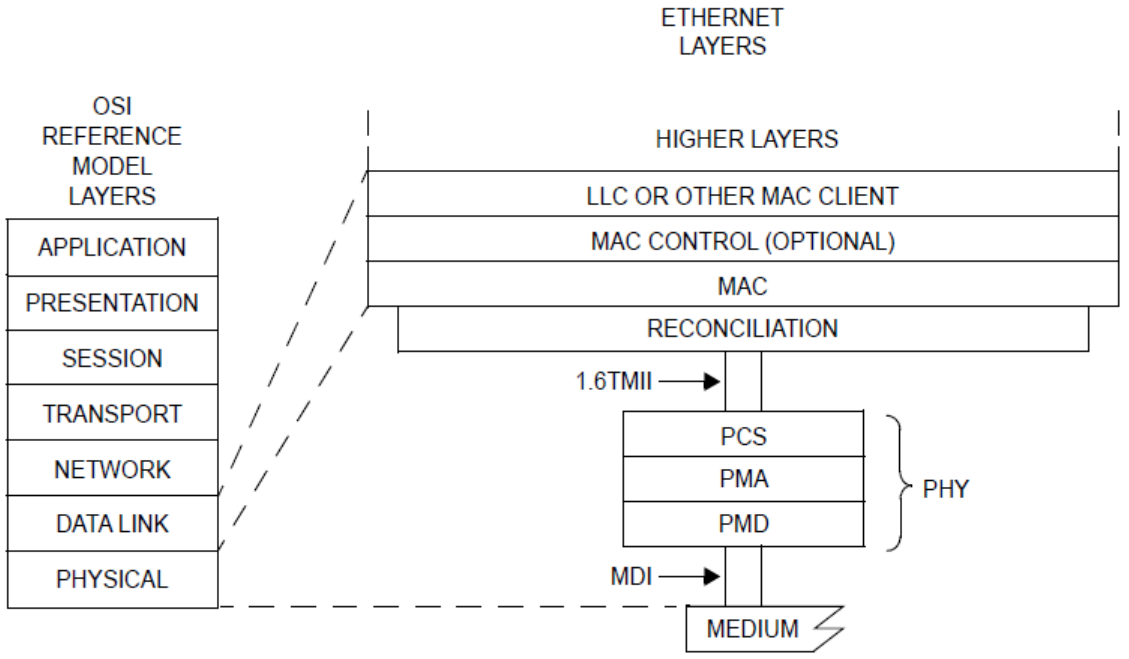
MDIs for 200GBASE-CR1, 400GBASE-CR2, 800GBASE-CR4, and 1.6TBASE-CR8

Annex 180A

(normative)

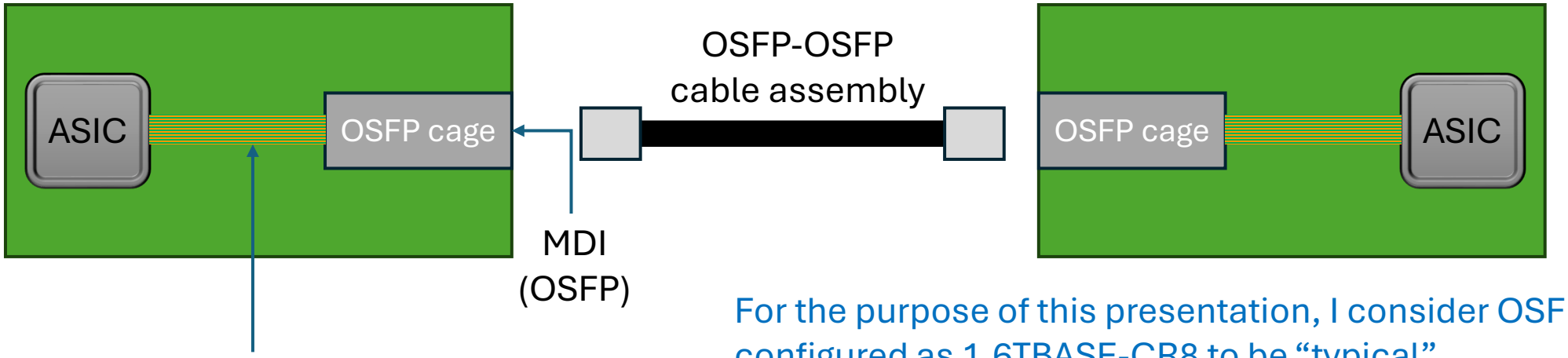
MDIs for 200GBASE-DR1, 400GBASE-DR2, 800GBASE-DR4, 1.6TBASE-DR8, 200GBASE-DR1-2, 400GBASE-DR2-2, 800GBASE-DR4-2, and 1.6TBASE-DR8-2

- The MDI is part of the Physical Layer, not part of the medium.



- All statements in these annexes pertain to the host, not the cable/fiber!
- The existing text does can be easily misinterpreted

“Typical” CR host connectivity (OSFP example)



All 8 MDI pairs are connected to PMD Tx and Rx functions in the ASIC

For the purpose of this presentation, I consider OSFP host configured as 1.6TBASE-CR8 to be “typical”

Alternatively, the host could be configured as

- Two 800GBASE-CR4
- Four 400GBASE-CR2
- Eight 200GBASE-CR1
- Or combinations thereof

I consider these as “breakout” configurations.
Breakout require different settings for all PHY sublayers (ILT in PMD, PMA, PCS, MAC), which need to be configured.

Text in 179C

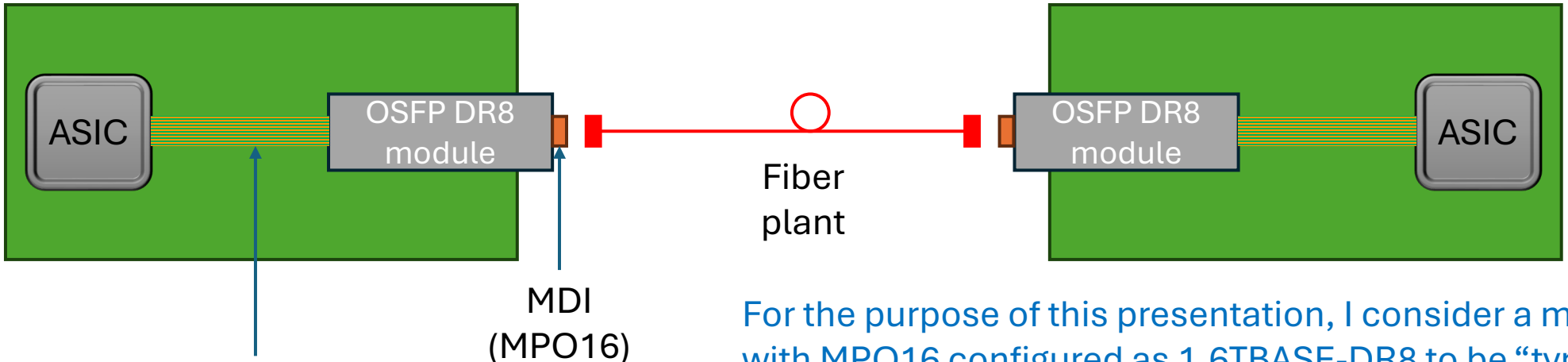
An MDI connector type may support one or more PMDs. The assignment of PMD signals to connector signals is specified in Table 179C–2, where as an example 0:DL0n refers to the DL0n signal of the first PMD; see 179.8.2 and 179.8.3 for signal naming definitions. When an MDI connector is not fully utilized the lower PMD numbers in Table 179C–2 should be used.

The phrase “When an MDI connector is not fully utilized” is ambiguous…

- It could mean a breakout configuration, where all MDI pairs are physically connected but some lanes are disabled or there is no link partner
  - In these cases – it is possible for any lane to enabled/disabled
- It could also mean that not all pairs in the MDI are physically connected to the ASIC
  - In such case, it makes sense to specify which pairs are connected
  - This situation is probably much less common

- With any meaning of this phrase, the situation described is not “typical”
  - It requires configuration of hosts on both sides of the cable, to ensure that the PHY types on both sides match
  - Without such configuration, there will be no link, regardless of which pairs are used
  - The current text does not mention it
  - The recommendation is not required and not sufficient for interoperability

“Typical” DR host connectivity (OSFP DR8 example)



All 8 AUI pairs are connected to AUI-C2M Tx and Rx functions in the ASIC

For the purpose of this presentation, I consider a module with MPO16 configured as 1.6TBASE-DR8 to be “typical”

Alternatively, the module could be configured as

- Two 800GBASE-DR4
- Four 400GBASE-DR2
- Eight 200GBASE-DR1
- Or combinations thereof

I consider these as “breakout” configurations.
Breakout require different settings for all PHY sublayers, which need to be configured.

Text in 180A.4.2

When the optical MDI is only supporting a single 1-lane PMD (200GBASE-DR1 or 200GBASE-DR1-2), the optical lanes shall be assigned to the optical connector positions Tx1 and Rx1, as shown in Table 180A–4, regardless of whether fibers are populated in the remaining optical connector positions.

When the optical MDI is only supporting a single 2-lane PMD (400GBASE-DR2 or 400GBASE-DR2-2), the optical lanes shall be assigned to the optical connector positions Tx1, Tx2, Rx1, and Rx2, as shown in Table 180A–4, regardless of whether fibers are populated in the remaining optical connector positions.

The phrase “When the optical MDI is only supporting” (*present participle*) is ambiguous…

- It could mean a regular module in a breakout configuration, where all fibers are physically connected but some lanes are disabled or there is no link partner
  - In these cases – the user is free to enable/disable any lane
- It could also mean a special module that does not have all fiber positions in the MDI physically connected to optical Tx/Rx (the module can never use fibers in some positions)
  - In such case, it makes sense to specify which fiber positions are connected
  - This situation is probably much less common

- With any meaning of this phrase, the situation described is not “typical”
  - It requires configuration of hosts on both sides of the fiber plant to ensure that the PHY types on both sides match
  - Without such configuration, there will be no link, regardless of which pairs are used
  - The current text does not mention it
  - The normative requirement is not required and not sufficient for interoperability

Recommendations

- In 179C.1
  - In the sentence "When an MDI connector is not fully utilized the lower PMD numbers in Table 179C–2 should be used"
    - Change "When an MDI connector is not fully utilized" to "When an MDI connector is only partially connected to the host transmit and receive circuitry"
    - Alternatively, delete the sentence altogether
  - Add the following informative note:

    NOTE—The PMD types on both sides of the cable assembly need to match. When the MDI is used for multiple PMDs or for PMDs with lower number of lanes than the MDI supports, appropriate configuration is required. The means for selecting the appropriate configuration are beyond the scope of this standard.
- In 180A.4.1 and 180A.4.2
  - Change all instances of "When the optical MDI is only supporting…" to a single statement "When the MDI is only partly connected to the optical transmit and receive circuitry, the lower PMD numbers in [Table 180A-2/Table 180A-4] should be used"
    - Changing "shall" to "should"
  - Alternatively, delete these sentences altogether
  - Add the following informative note in each subclause:

    NOTE—The PMD types on both sides of the fiber plant need to match. When the MDI is used for multiple PMDs or for PMDs with lower number of lanes than the MDI supports, appropriate configuration is required. The means for selecting the appropriate configuration are beyond the scope of this standard.

Summary

- The text in 179C and 180A regarding partially utilized MDIs is ambiguous and potentially confusing
- The recommendations/specifications for lane assignments are neither required nor sufficient
- Alternative language and informative NOTEs were proposed

That’s all!

Questions?