40GBASE-SR4 & 100GBASE-SR10 PMD Service Interface Update

John Petrilla, Rita Horner,
Brian Misek, Piers Dawe
Avago Technologies
July 2008

Supporters

- Ali Ghiasi, Broadcom Corporation
- Petar Pepeljugoski, IBM
- Tom Palkert, Luxtera
- Jonathan King, Finisar
- Kenneth Jackson, Emcore Fiber Optics
- Mike Dudek, JDS Uniphase
- Gourgen Oganessyan, Quellan
- Ryan Latchman, Gennum
- Frank Chang, Vitesse

Outline

Focus will be on the Host IC to PMD channel

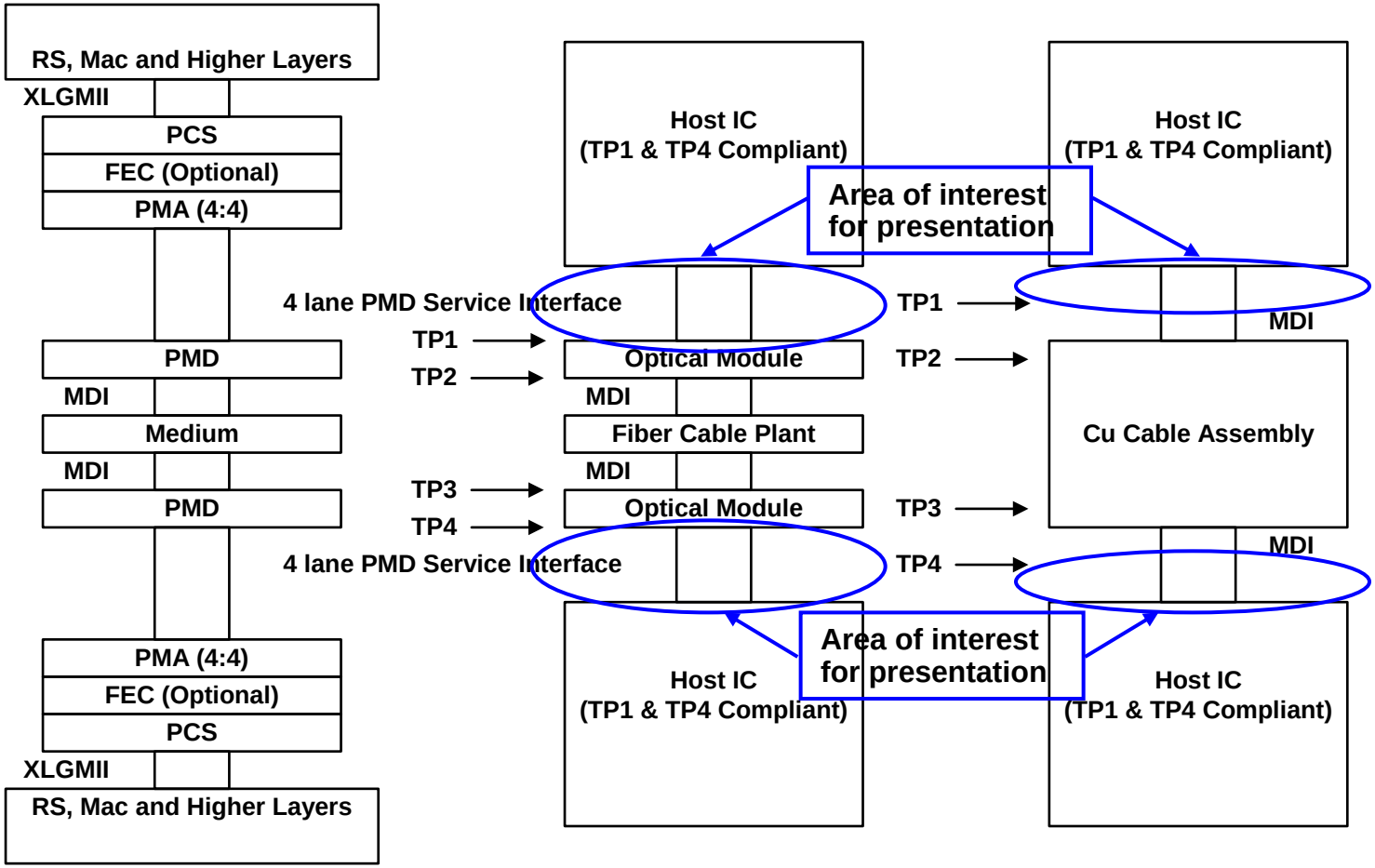
- 802.3ba Alignment
- Elements for success
  - Opportunities for and issues with common form factors
- 40GBASE-SR4 & 100GBASE-SR10 Proposal
  - PMD service interface jitter and electrical characteristics
- Conclusions, Recommendations & Next Steps
- References
- Backup

802.3ba Alignment
802.3ba Objectives Addressed in Presentation

- Support MAC data rates of 40 Gb/s & 100 Gb/s
- Support a BER better than or equal to 1E-12 at the MAC/PLS service Interface
- Provide … at least 100 m on OM3 MMF
- Provide … at least 10 m over a copper cable assembly

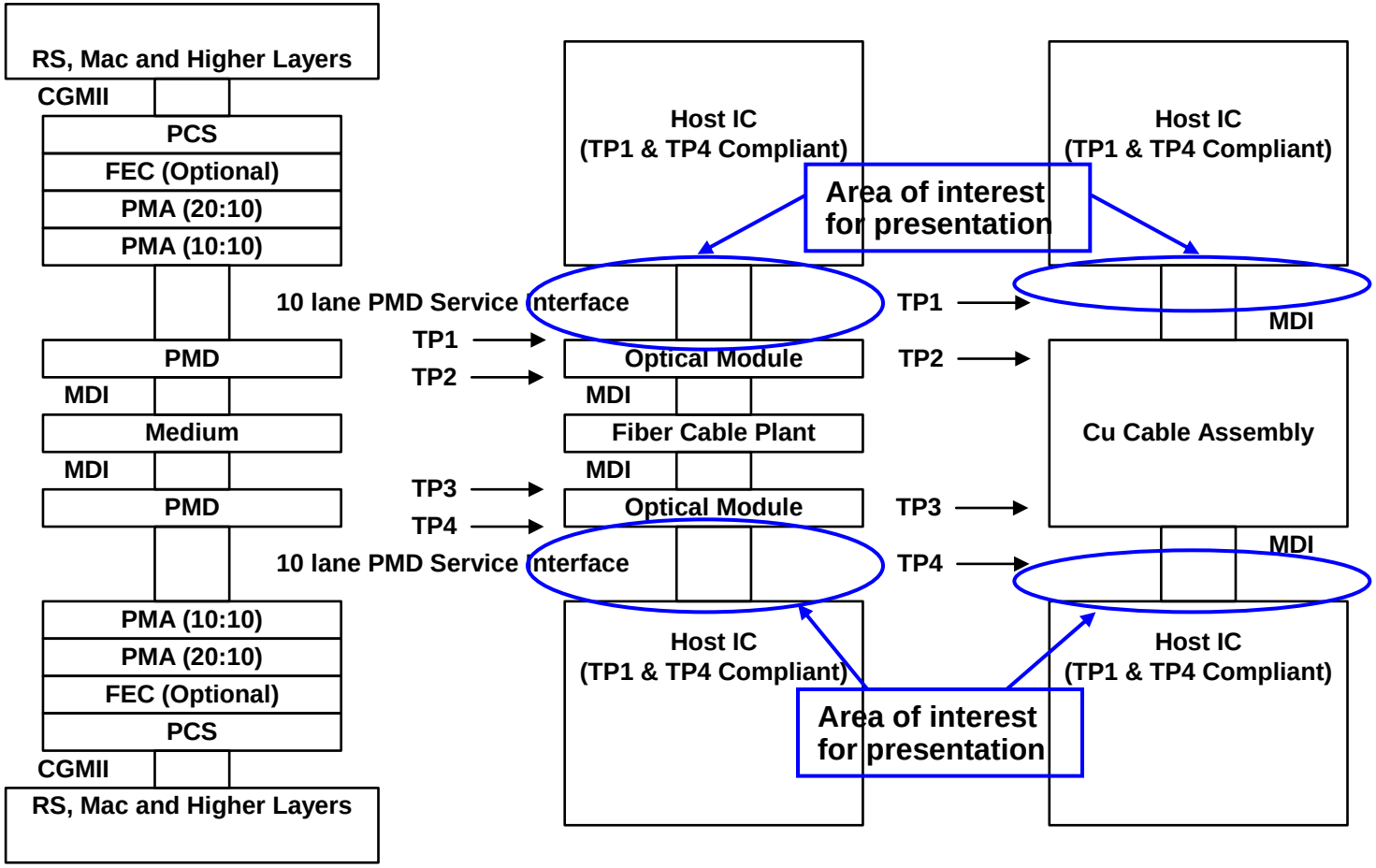
802.3ba Alignment

802.3ba Architectural Layers & Interfaces – 40G



802.3ba Alignment

802.3ba Architectural Layers & Interfaces – 100G

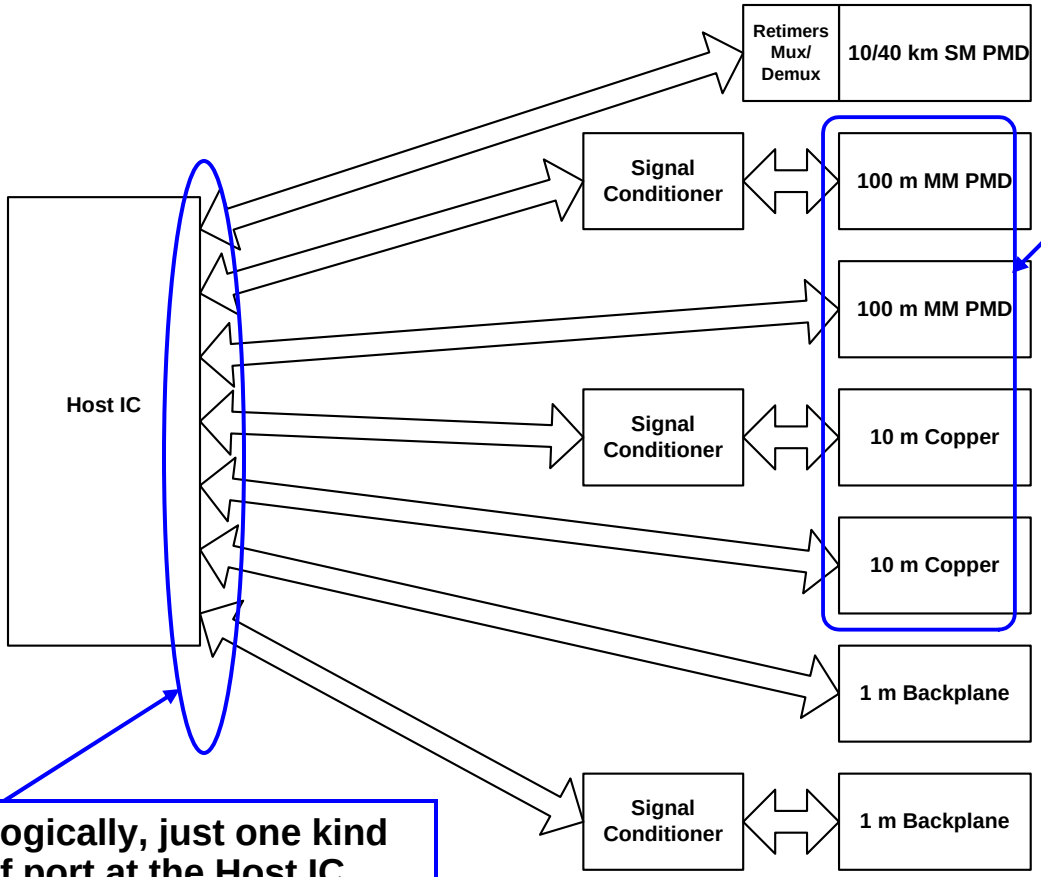


Elements for Success
Overview

- Total cost: less than four/ten 10 GbE solutions
  - A pluggable form factor that is common to multiple variants, e.g. copper cable and MM fiber, leads to lower costs by sharing piece parts and footprints and by accelerating market acceptance and increasing volumes. This enables a single build standard for DTE that can be connected by a choice of the two PMD/media types that will dominate the data center. These advantages are compelling and the opportunities should not be overlooked.
- Power consumption: less than four/ten 10 GbE solutions
- High module density: higher than 10 GbE solutions
- Cable plant: 100 m of OM3 & up to 4 intermediate connections
- Reliability: better than ten/four 10 GbE solutions
- Appropriate design points
  - Support ~6 dBe channel insertion loss (SDD21 at Nyquist frequency) between the host IC and pluggable module without an intermediate connector for 6" to 8" of PCB signal traces. This loss occurs twice, once on each end.
  - Use experience gained in SFP+ (SFF-8431) and 8GFC (FC-PI -4) as well as 802.3ap developments to guide choices for electrical attributes.

May be too aggressive. Making smart choices is essential.

Opportunities for a Common Form Factor



Common form factor

- A common pluggable form factor can support the multilane MM and copper cable variants due to similarities in size, contact count and power consumption. By comparison, the SM modules are expected to have ~4x the power consumption of the MM modules.
- For some applications the host environment will permit direct connection between a Host IC and the pluggable module or backplane connector. For other environments some form of signal conditioner will be required. Both cases are shown.
- Here the term signal conditioner is used to cover all forms of link extenders, CDRs, equalizers, etc.

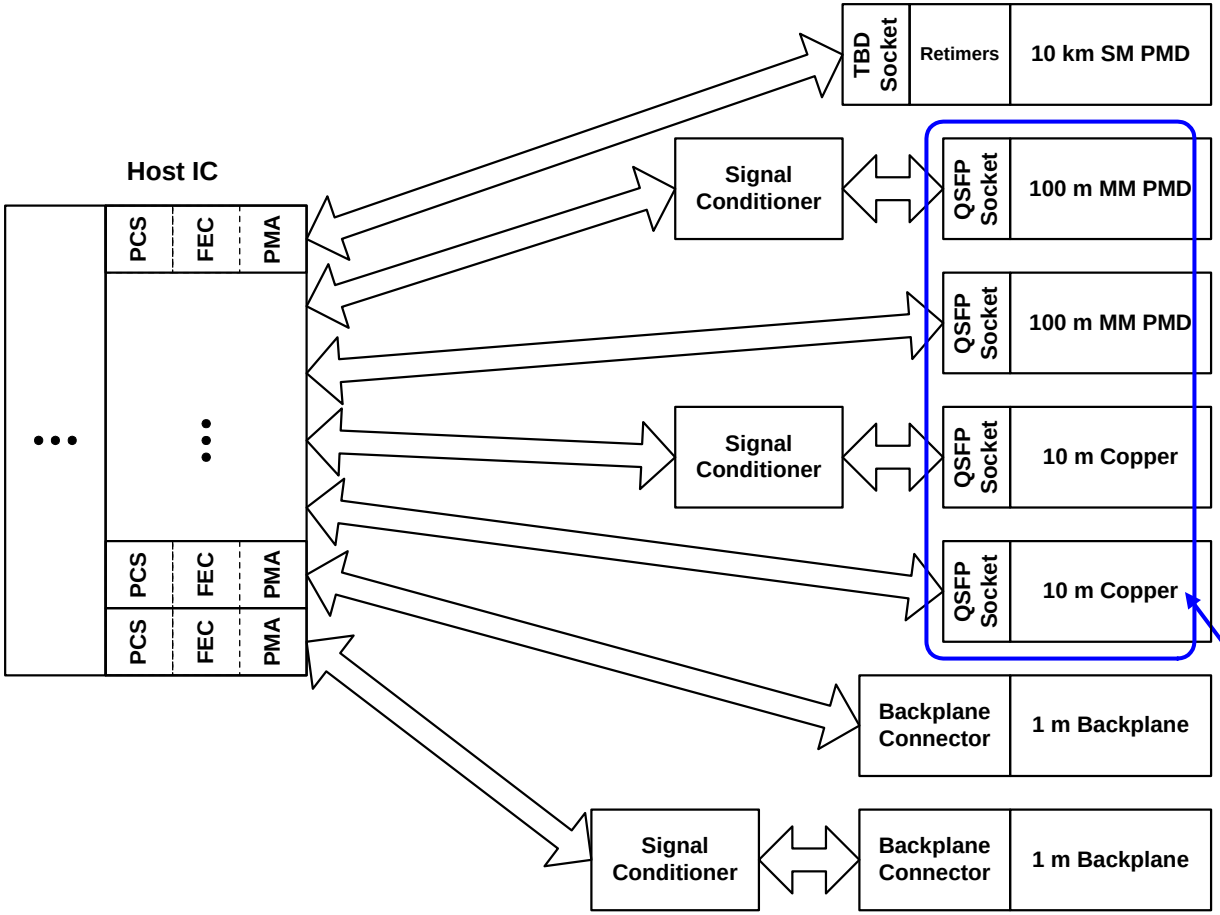
Logically, just one kind of port at the Host IC

Identical functional, logical, digital and management spec

Compatible analog and auto-negotiation

Opportunities for a Common Form Factor

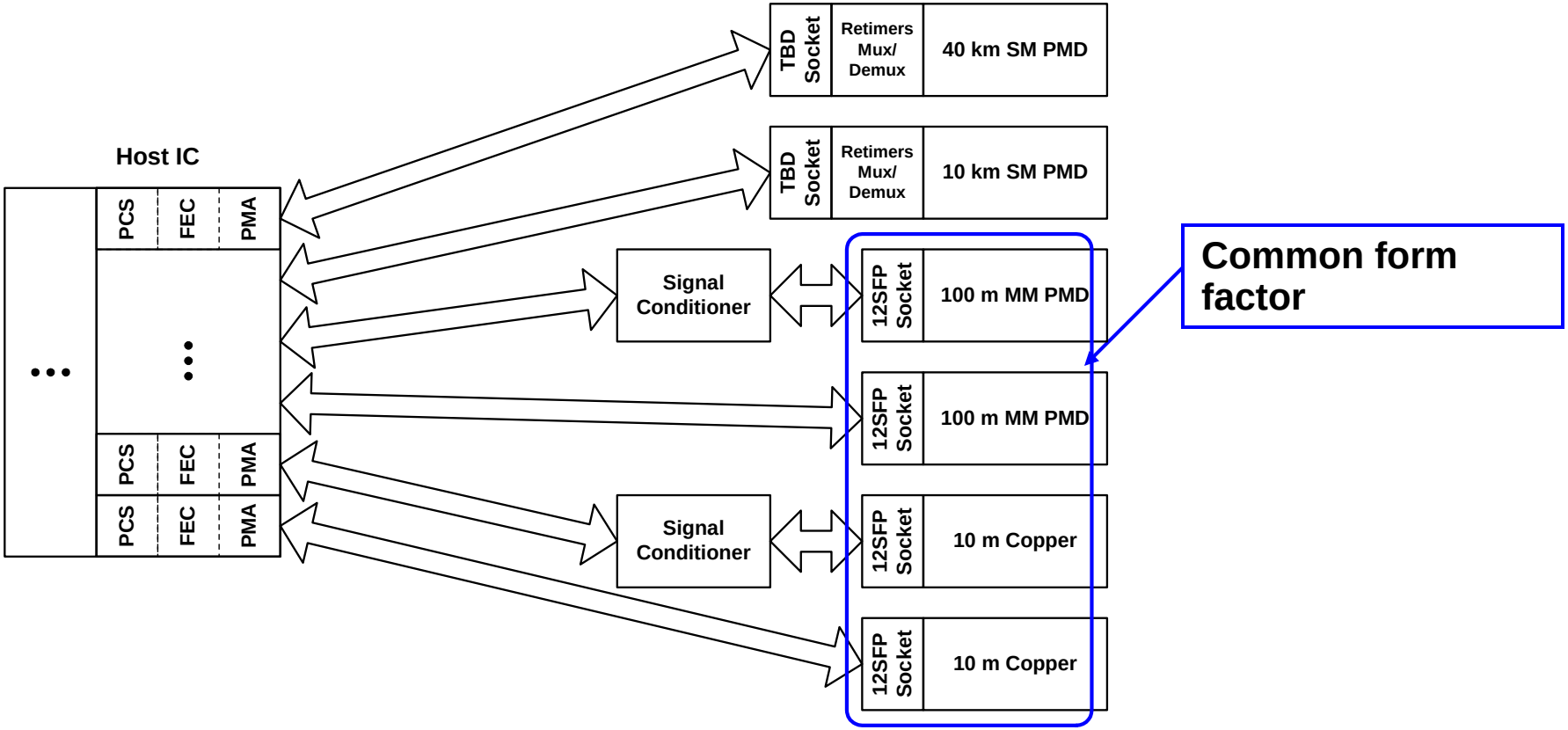
40G Variants & QSFP



- For the 40G variants, the QSFP form factor looks to be a good fit for pluggable 100m MM and 10m Copper variants.
- One issue to address is defining a channel between the Host IC and QSFP socket that supports both the 100m MM and 10m Copper variants. Here the existing terminology differences between copper and optics can be confusing.

Common form factor

Opportunities for a Common Form Factor
100G Variants & 12SFP



- For the 100 m MM PMD and 10 m copper variants, a common pluggable form factor seems reasonable. While there is currently no popular form factor, reasonable 10 lane or 12 lane form factors appear plausible. Here a 12 lane form factor, 12SFP for short, is presumed. As above, an issue to address is defining a channel between the Host IC and 12SFP socket that supports both the 100 m MM and 10 m Copper variants and not being deterred by existing terminology.

Service Interface
Definition Rationale

- For the 100 m OM3 variant, pluggable, multilane, non-retimed, limiting fiber optic modules are expected to provide the lowest power, highest density and lowest cost solution.
  - Direct connection between the module and host IC without intermediate signal conditioners is required to maximize the power, density and cost advantages.
- For the 10 m copper cable assembly variant, direct connection between the host IC and cable assembly offers similar cost, density and power savings as for the 100m OM3 variants.
- For the 1 m backplane variant, direct connection between the host IC and backplane connector offers similar cost, density and power savings as for the 100 m OM3 variants.
- To enable direct connection for the 100 m OM3 variant, the approach and characteristics on the following pages are proposed for consideration.
- Requirements for the 10 m cable assembly and 1m backplane variants are being included as they become apparent.

Beginning to look challenging

PMD Service Interface

Prologue – Jitter Allocation Targets

	GbE	8GFC	10G SFP+ SFF-8431	**40/100G Targets**
TP1 DJ, UI	0.100	0.170	0.100*	**0.150**
TP1 DDPWS, UI		0.110		
TP1 TJ, UI	0.240	0.310	0.280	**0.300**
TP4 DJ, UI	0.462	0.420	0.420	**0.400**
TP4 DDPWS, UI		0.360		
TP4 TJ, UI	0.749	0.710	0.700	**0.700**

* SFF-8431 specifies DDJ instead of DJ.

- Perhaps the most critical attributes of TP1 & TP4 are the jitter allocation. The above table provides information regarding previous choices.
- The targets were chosen to provide a better design point than those from either 8GFC or 10GBASE-SR as implemented in SFF-8431. With respect to 10G SFP+ in SFF-8431, significant relief is expected.

PMD Service Interface

Goal and Approach

Goal: Establish an approach and specifications that enable direct connection and inter-operability between host ICs and pluggable, multilane, non-retimed, limiting fiber optic modules for a reasonable range of equipment designs

- The proposed approach does not place explicit requirements on the host IC, but, rather, provides far-end characteristics over worst case channels.
- The far-end characteristics for the host IC, TP1 & TP4, are based on experience gained from SFP+ (SFF8431) and 8GFC (FC-PI-4) developments. Included are jitter, signal levels, and reflection coefficients.
- The proposed channel is intended to support 150 mm to 200 mm of PCB traces without an intermediate connector and is defined by an SDD21 template.
- TP1 requirements are defined in an input characteristics table and TP4 requirements are defined in an output characteristics table that follow.
- Compliance, as with SFF-8431, is based on use of compliance test boards to improve measurement accuracy and reproducibility.

May be too aggressive

Multilane channels are challenging

The intention is to maintain the same TP1 & TP4 requirements for 40GBASE-SR4 and 100GBASE-SR10.

PMD Service Interface
Host IC – PMD Channel

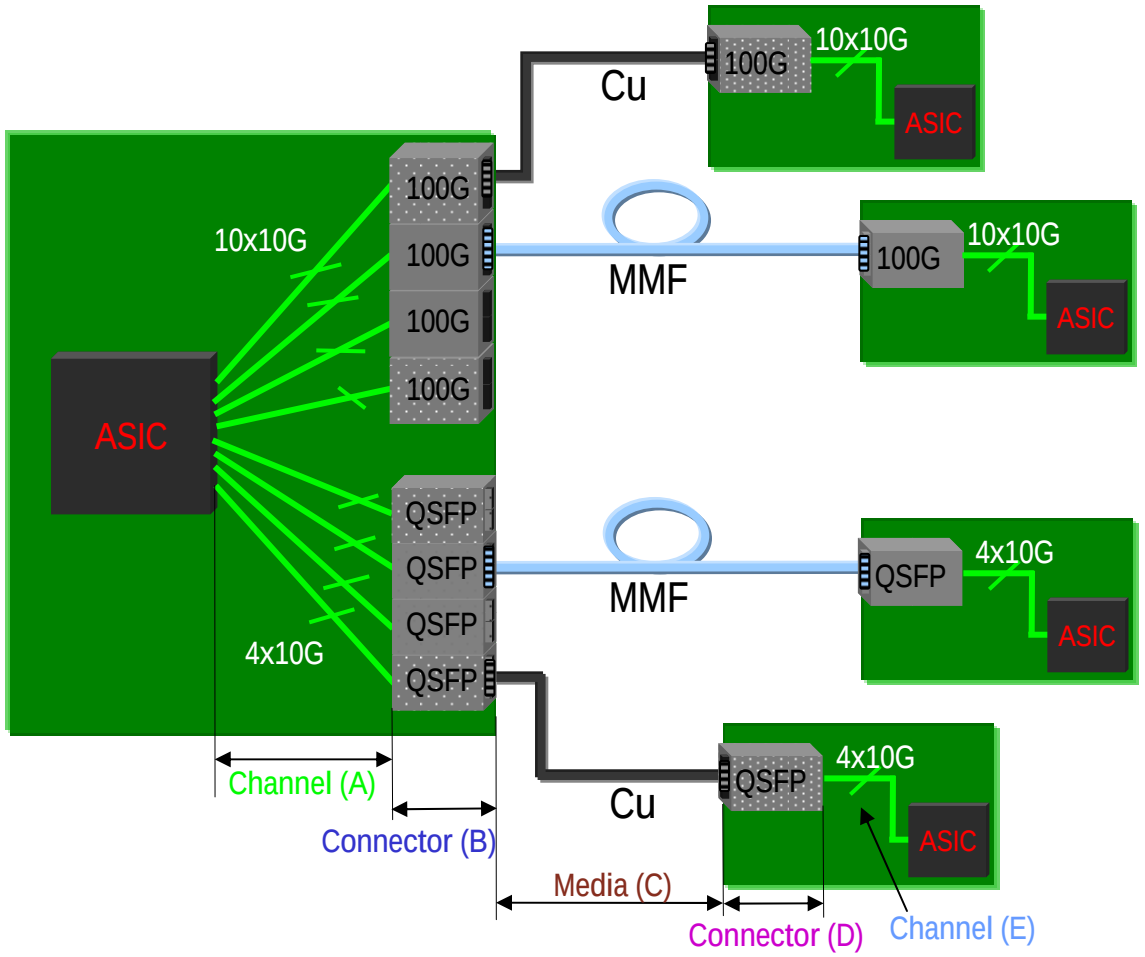
Goal: To define a common port and channel that accommodates both pluggable 100 m MM fiber and 10 m cable assemblies from a single ASIC

- For cable assemblies using 24 AWG, SDD21 loss at the Nyquist frequency is expected ~ 22 dBe. Accommodating another 6 dBe between the port socket and the ASIC on each side may be un-necessarily challenging for the host IC.
- Three issues will be presented.
  - PCB trace lengths that may be required in end use to support multi-port equipment
  - PCB trace lengths that may be required in evaluation boards and/or compliance test boards.
  - Appropriate compliance points for host IC, i.e. where to locate TP0 and TP5.

Need for Well Defined Port Side Interface Specification

Optics and Cu

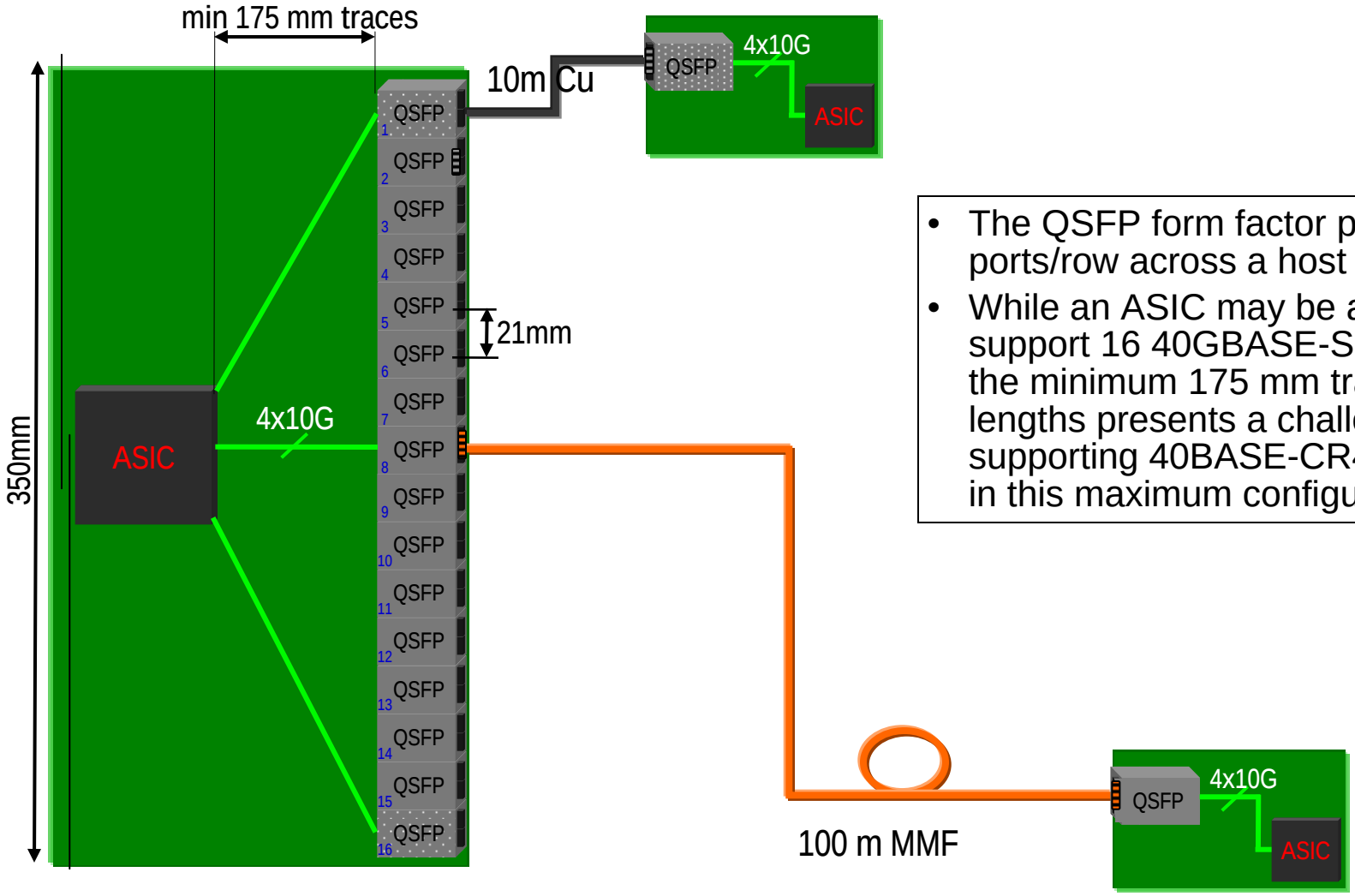
- 100 m MMF links appear jitter limited.
- 10 m Copper links appear insertion loss or ICR limited.
- Can a common test point (e.g. TP0 or TP1) with a common channel (e.g. Channel (A)) be defined that covers both variants?



- For 100 m MMF Total SerDes Link Jitter = ASIC Generation + Channel Contribution (A) + MMF Tx Contribution (B) + MMF Contribution (C) + MMF Rx Contribution (D) + Channel Contribution (E) + ASIC Tolerance
- For 10 m Cu Total SerDes Link Loss = Channel Loss (A) + Connector Loss (B) + Interconnect Media Loss (C) + Connector Loss (D) + Channel Loss (E)

40G SerDes Port Side Connections to QSFP

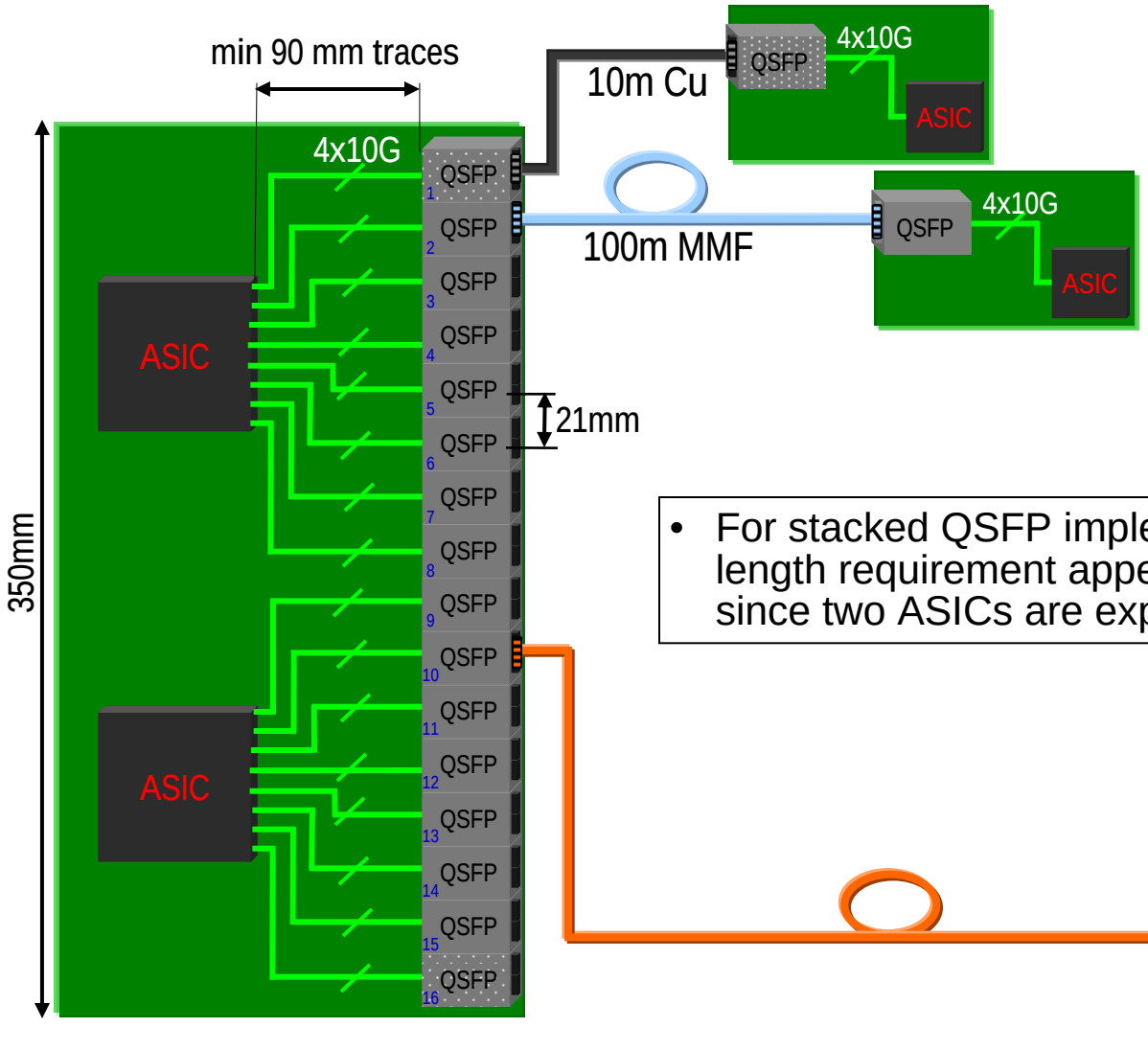
Single ASIC – Max Non-Stacked Port Case



- The QSFP form factor permits 16 ports/row across a host PCB.
- While an ASIC may be able to support 16 40GBASE-SR4 ports, the minimum 175 mm trace lengths presents a challenge when supporting 40BASE-CR4 variants in this maximum configuration.

40G SerDes Port Side Connections to QSFP

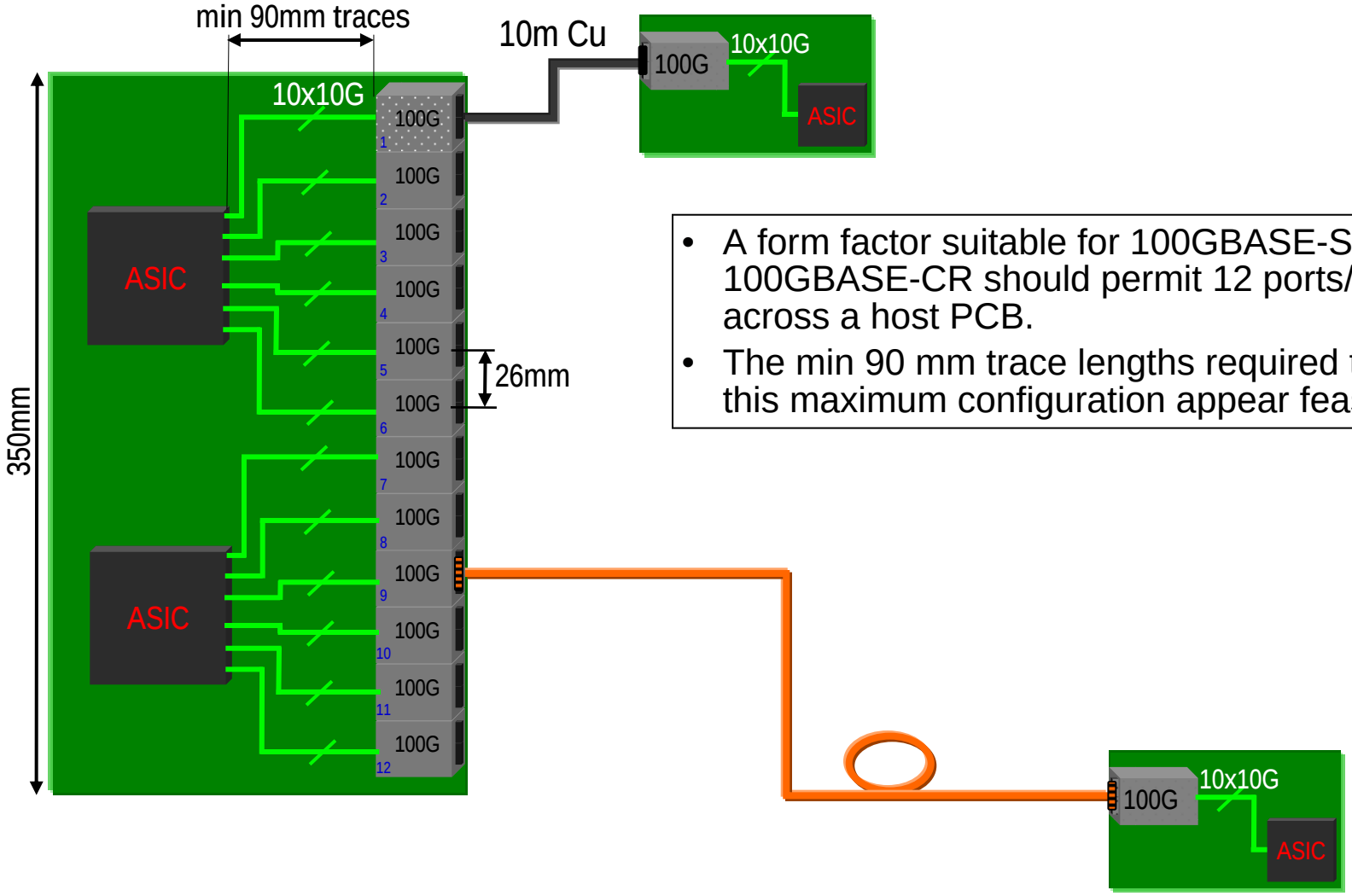
Maximum Stacked Port Case



- For stacked QSFP implementations the trace length requirement appears more supportable since two ASICs are expected.

100G SerDes Port Side Connections

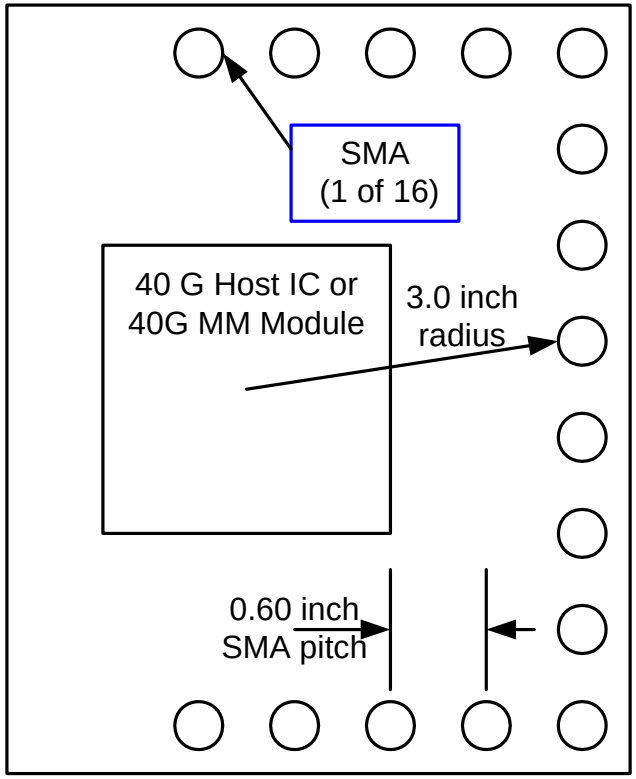
Two ASICs – Max Port Case



- A form factor suitable for 100GBASE-SR10 and 100GBASE-CR should permit 12 ports/row across a host PCB.
- The min 90 mm trace lengths required to support this maximum configuration appear feasible.

TP1 & TP4 PMD Service Interface

4 Lane Compliance and/or Evaluation Boards

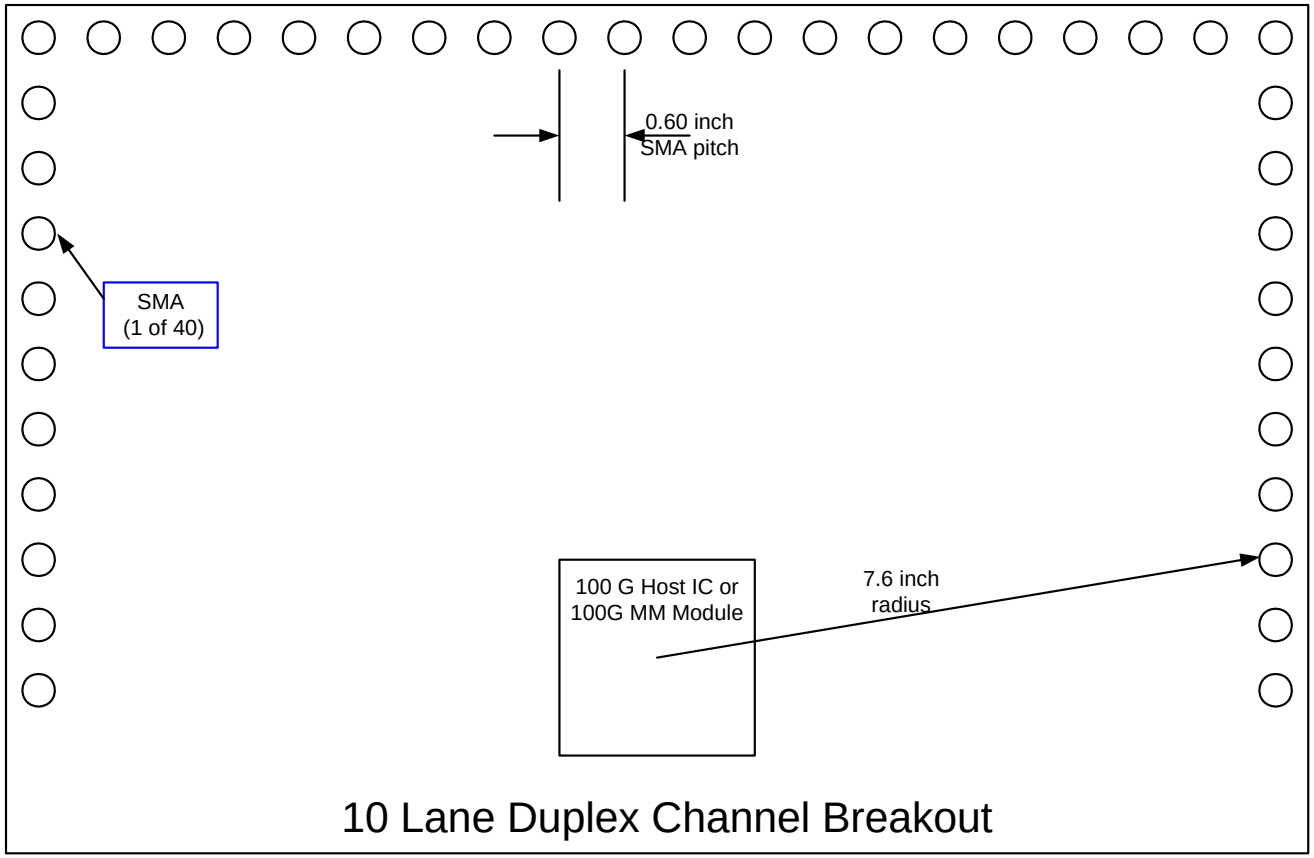


4 Lane Duplex Channel Breakout

- Evaluation and/or compliance test cards for multilane channels face the challenge of breaking out all channels. In the figure all eight differential pairs of a full duplex, four lane channel are made available with SMA connectors. For an SMA pitch of 0.60 inch and semicircular pattern, a minimum radius of 3.0 inches is required. This is larger than the trace length in SFF-8431 module compliance boards but may be tolerable.
- The pattern in the drawing is rectangular for simplicity.

TP1 & TP4 PMD Service Interface

10 Lane Compliance and/or Evaluation Boards

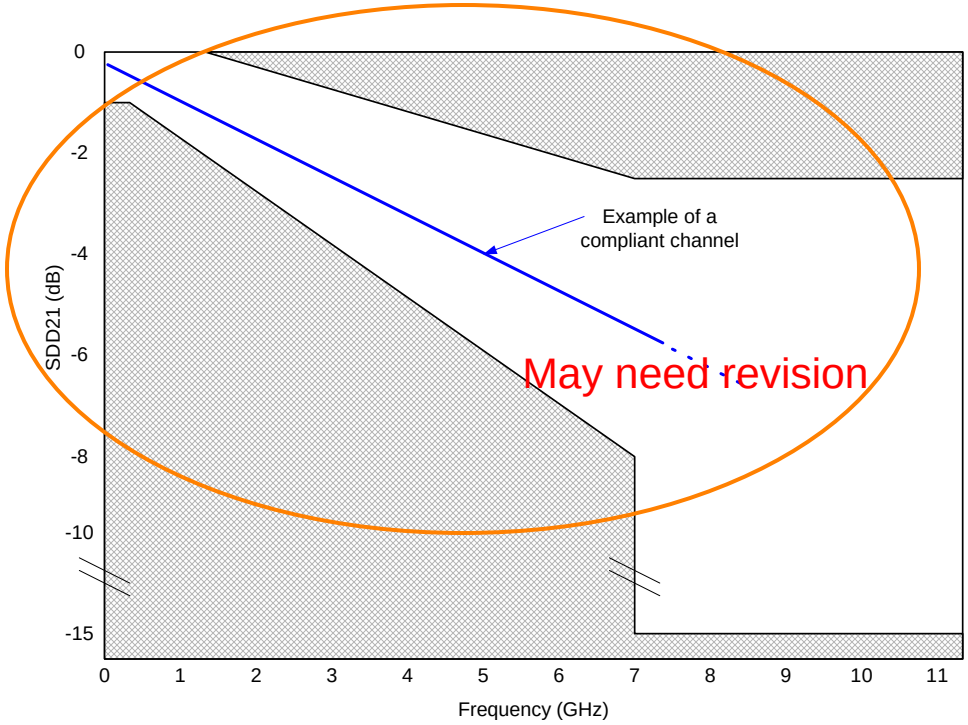


- In the figure all 20 differential pairs of a full duplex, ten lane channel are made available with SMA connectors. For an SMA pitch of 0.60 inch and semicircular pattern, a minimum radius of 7.6 inches is required. This trace length is significantly longer than trace lengths in SFF-8431 module compliance boards and presents a significant challenge.

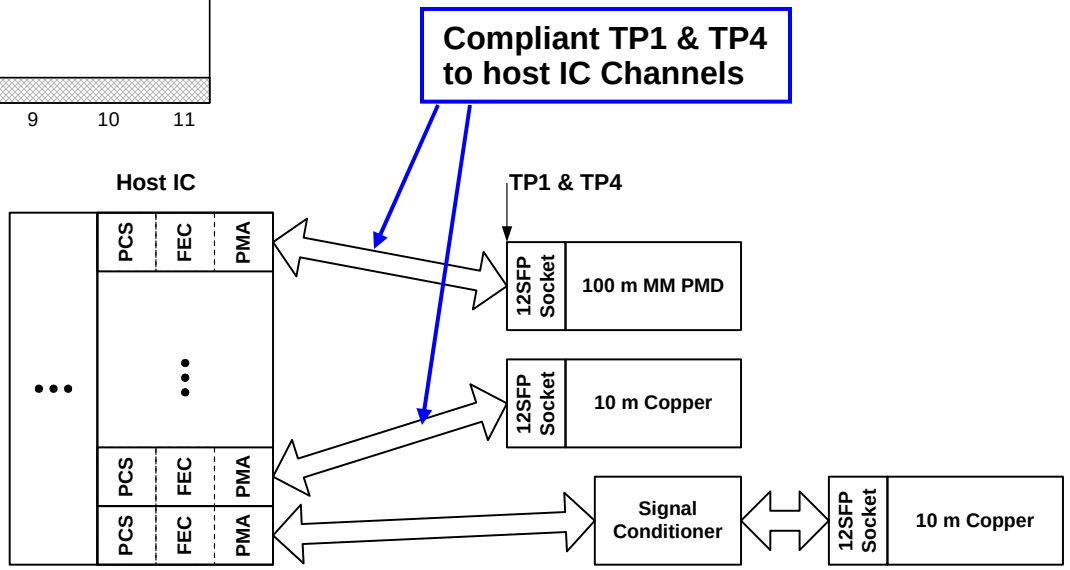
PMD Service Interface
Host IC – PMD Channel Summary

- While for 100 m MMF variants, jitter appears to the limiting factor, for 10 m Copper links, insertion loss (or insertion loss crosstalk ratio) appears to be the limiting factor.
  - Finding a way to reduce the combined insertion loss without compromising the needed reaches will be key.
- While supporting some possible maximum port population cases requires a minimum 175 mm of PCB traces, most can be supported with a minimum of 90 mm.
- The short trace lengths found in SFF-8431 compliance cards may not be feasible for 10-lane channels. Connectors offering a higher packing density and only breaking-out a subset of the lanes should be explored.
- Finding optimum compliance points for the PMD Service Interface that account for the needs of the ASIC, 100 m MMF pluggable modules and the 10 m Copper pluggable assemblies as well as pragmatic consideration of compliance test cards requires further attention. For the ASIC, a single set of test points that are common to copper and MM optics variants is important.

TP1 & TP4 to Host IC Channel SDD21 Compliance Template

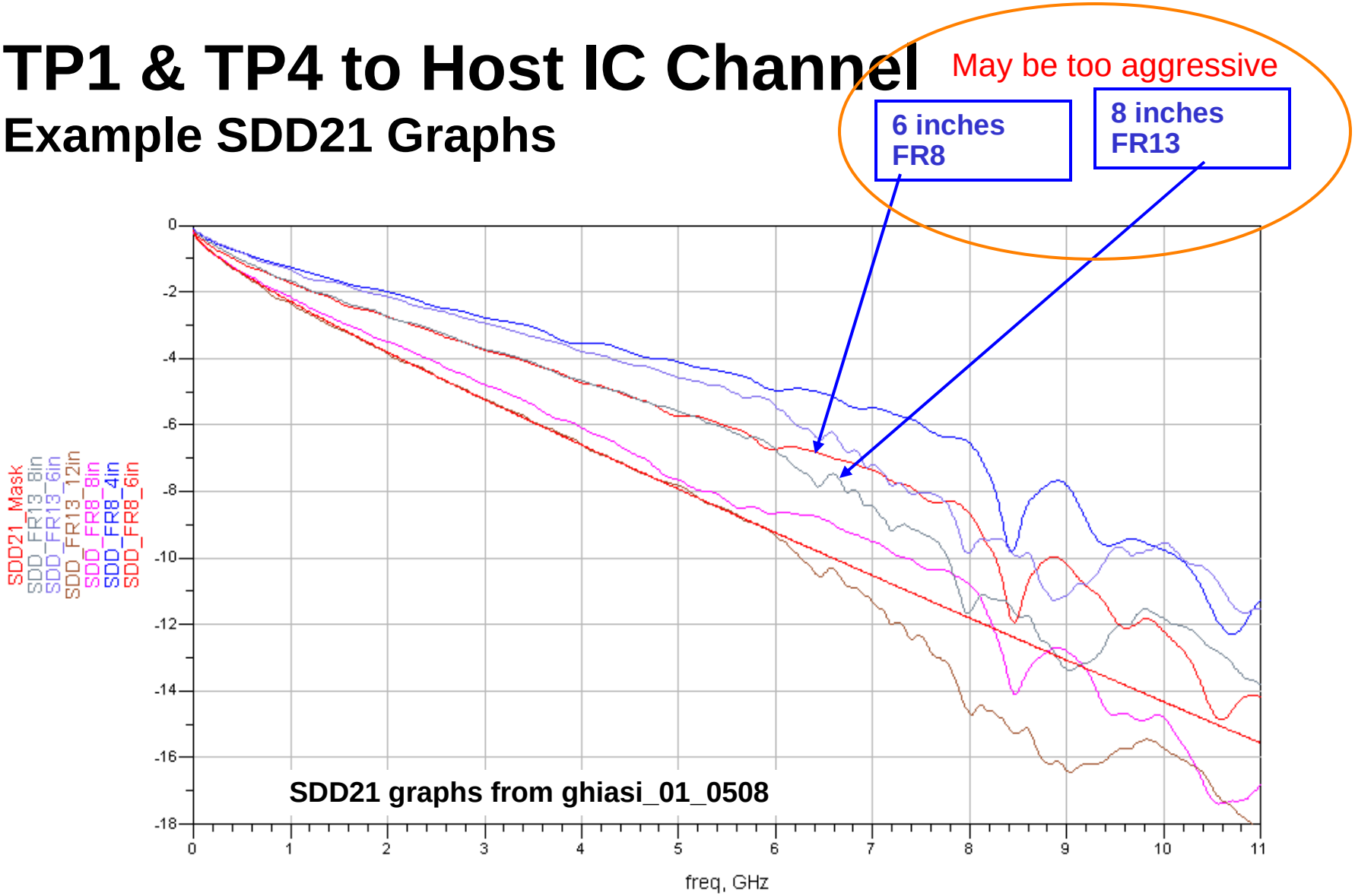


- Characteristics are for each lane individually and are normative except where noted.
- All values are provisional, shown for example, and will benefit from additional study.



TP1 & TP4 to Host IC Channel

Example SDD21 Graphs



- The above examples for FR13- 8in, FR13- 6in, FR8-6in and FR8-4in all meet the proposed template.

TP1

40GBASE-SR4 & 100GBASE-SR10 Input Characteristics

Parameter Description	Value	Units	Conditions
Single ended input voltage tolerance range	-0.3 to 4.0	V	Ref'd to module signal common
AC common mode input voltage tolerance (min)	15	mV (RMS)	
Differential input reflection coefficient, SDD11 (max)	See Template A	dB	0.01 to 11.1 GHz
Reflected differential to common mode conversion, SCD11 (max)	-12	dB	0.01 to 11.1 GHz
Total jitter tolerance	0.30	UI	At BER = 1E-12
Deterministic jitter tolerance	TBD	UI(p-p)	
Eye mask coordinates: X1, X2, Y1, Y2	0.15, TBD, 90, 350		TBD

Characteristics are for each lane individually and are normative except where noted.

All values are provisional, shown for example, and will benefit from additional study.

As requirements for 100 m MMF and 10 m Cu are harmonized, DJ jitter tolerance may be relaxed and the maximum input signal amplitude tolerance increased depending on the maximum supported PCB trace length. Alternatively, the ASIC could generate different signal amplitudes for copper and fiber

TP4

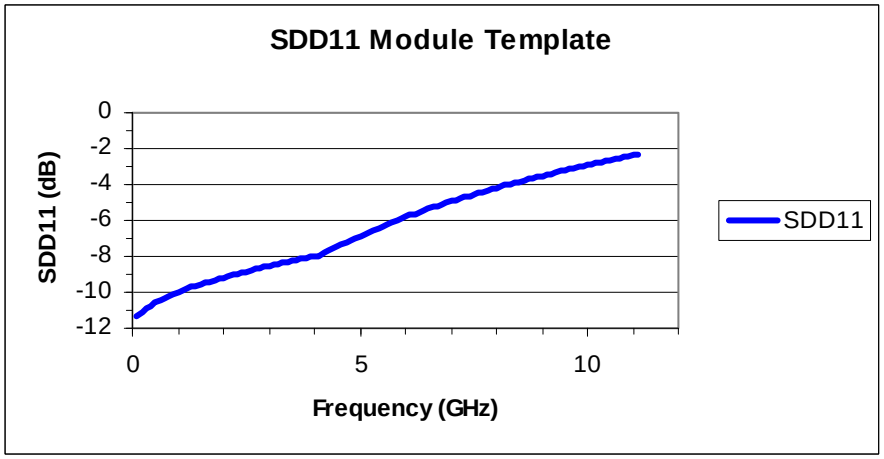
40GBASE-SR4 & 100GBASE-SR10 Output Characteristics

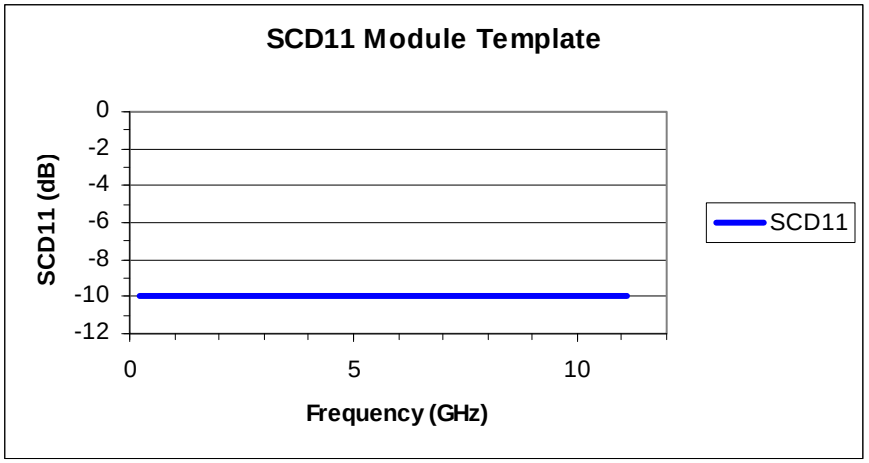
Parameter Description	Value	Units	Conditions
Single ended output voltage tolerance range	-0.3 to 4.0	V	Ref'd to module signal common
AC common mode output voltage (max)	7.5	mV (RMS)	
Termination mismatch at 1 MHz	5	%	
Differential output reflection coefficient, SDD22(max)	See Template B	dB	0.01 to 11.1 GHz
Common mode output reflection coefficient, SCC22 (max)	See Template C	dB	0.01 to 11.1 GHz
Output transition time, 20% to 80%, (min)	28	ps	
Total jitter	0.70	UI	At BER = 1E-12
Deterministic jitter	0.40	UI(p-p)	
Eye mask coordinates: X1, X2, Y1, Y2	0.35, TBD, 150, 425		TBD

Characteristics for are each lane individually and are normative except where noted.

All values are provisional, shown for example, and will benefit from additional study.

TP1
Reflection Coefficient Characteristics





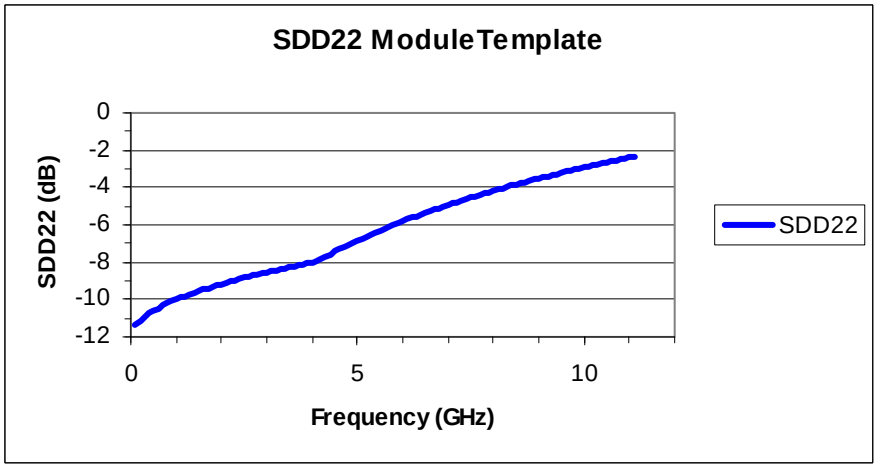
Template A
(updated for SFF-8431 r3.0)

Characteristics are for each lane individually and are normative except where noted.

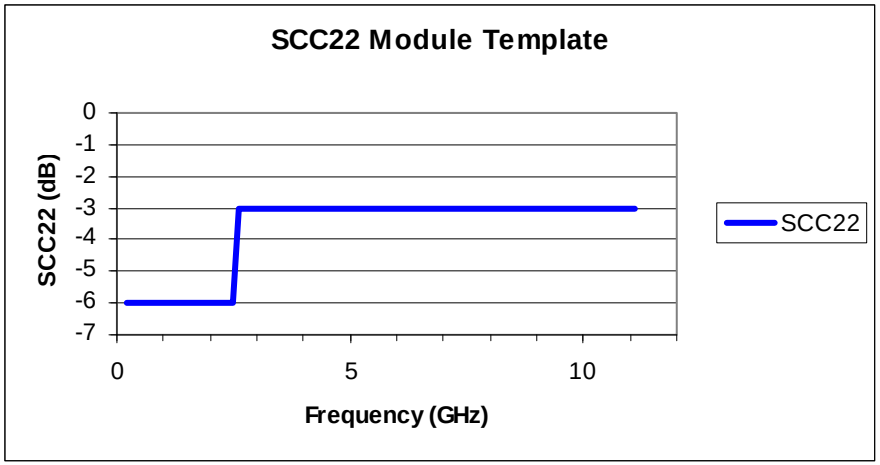
All values are provisional, shown for example, and will benefit from additional study.

TP4

Reflection Coefficient Characteristics



Template B
(updated for SFF-8431 r3.0)

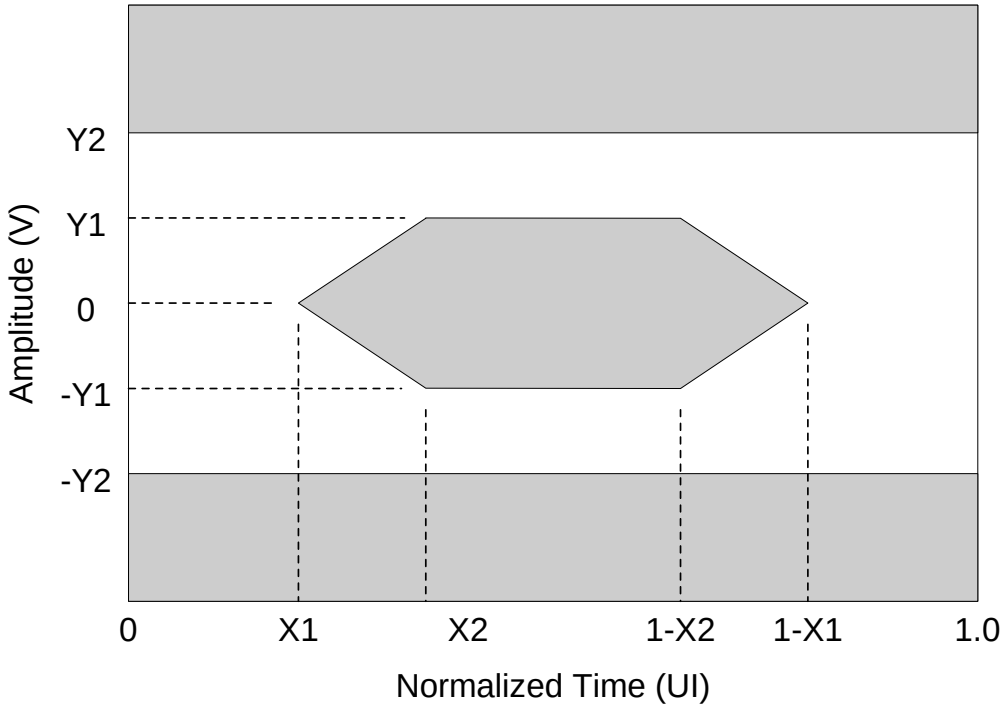


Template C

Characteristics are for each lane individually and are normative except where noted.

All values are provisional, shown for example, and will benefit from additional study.

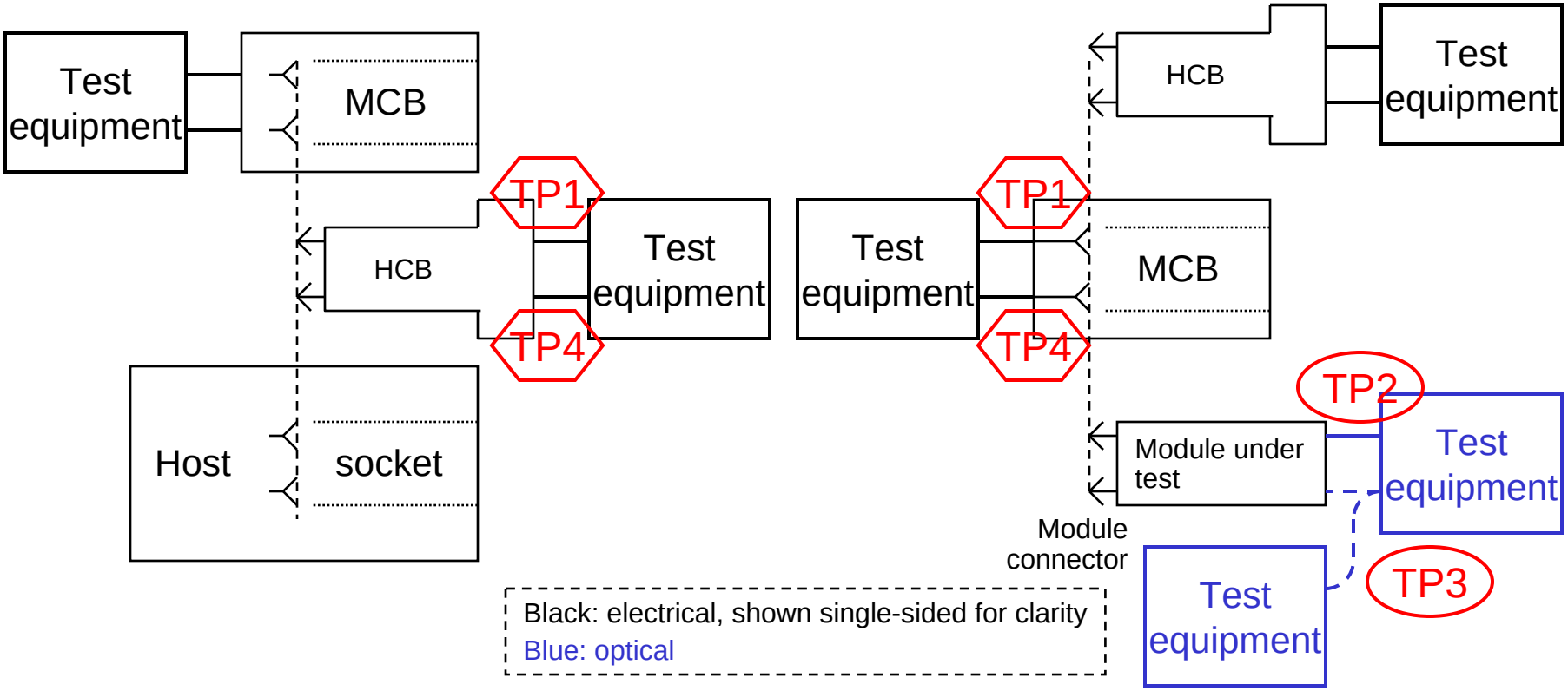
TP1 and TP4
Eye Mask



Characteristics are for each lane individually and are normative except where noted.

All items will benefit from additional study.

Use of compliance boards



- HCB (Host Compliance Board) Used to provide access at the far-end, TP1 & TP4, of Host IC signals and for calibration of module compliance test signals. Host system transmit and receive signal compliance are defined with the HCB inserted in the pluggable interface of the host.
- MCB: (Module Compliance Board) Used to provide access at TP1 & TP4 to module signals and for calibration of host compliance test signals. Module transmit and receive signal compliance are defined with the module inserted in the pluggable interface provided by the MCB.

Conclusions, Recommendations & Next Steps

Conclusions:

- Cost, power and density advantages of 100m OM3, 10m copper cable assembly and 1m backplane variants are maximized by direct connection with the host IC and a single build standard for DTE that can be connected by a choice of the two PMD/media types that will dominate the data center.
- Robust solutions, inter-operability of pluggable modules and cable assemblies and market acceptance are enabled by well chosen required interface characteristics.
- This presentation provides a set of characteristics sufficient to enable direct connection between the proposed 100m OM3 variant and a host IC.

Recommendations:

- The specification approach as outlined in the tables, pages 24 & 25, and associated templates, pages 22, 26, 27 & 28 should be considered for inclusion in 802.3ba.
- Coordinate TP1 and TP4 requirements with other PMD variants, looking for commonality in use for ASICs, 100 m MM and 10 m copper variants. Re-activation of the Test Point ad hoc should be considered.

Next Steps:

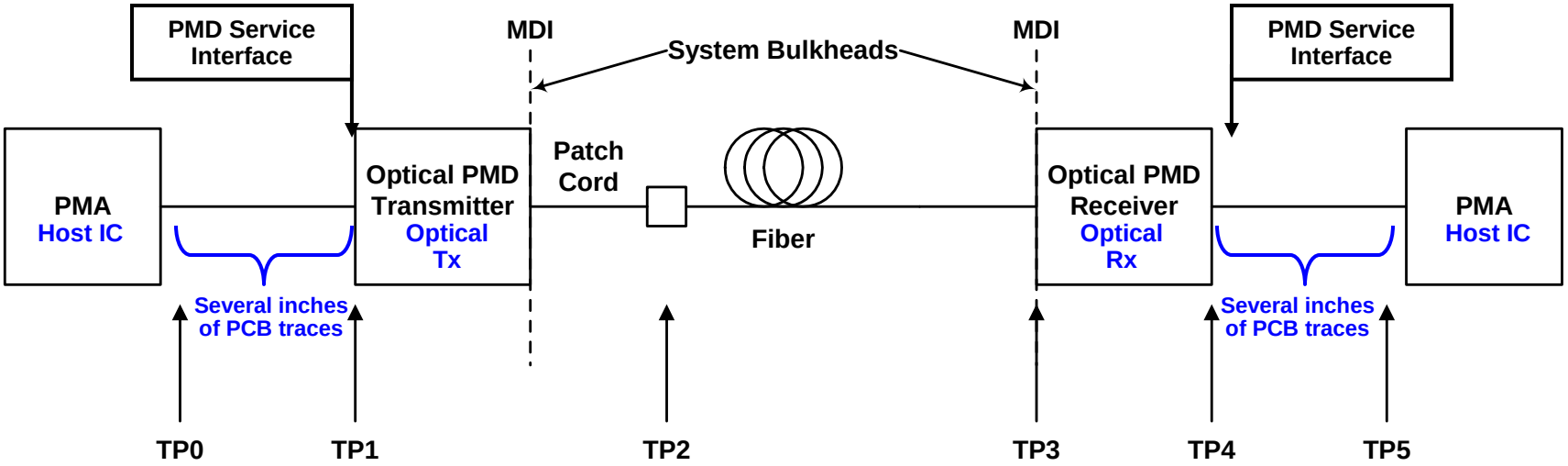
- Continue to gather and incorporate feedback regarding host IC capabilities
- Collect information regarding crosstalk and other impairments related to multilane channels, stacked connectors and modules and incorporate.
- Upgrade proposed specifications.

References

- SFF-8431 Specifications for Enhanced 8.5 and 10 Gigabit Small Form Factor Pluggable Module ”SFP+”: ftp://ftp.seagate.com/sff/SFF-8431.PDF
- FC-PI-4 Physical Interface-4 (8GFC): http://www.t11.org/index.htm
- ghiasi_01_0508
- pepeljugoski_01_0508
- petrilla_01_0508
- petrilla_01_0308

802.3ba Alignment

802.3ba 100 m MM PMD Block Diagram



- The above block diagram shows relevant elements and interfaces for an optical link between two PMAs. The patch cord is included for the definition of TP2. Otherwise intermediate fiber connectors are not shown.
- P802.3ba, in Munich, agreed to include “an optional n-lane x 10.3125GBd electrical interface for PMD service interface … as baseline”.
- Here the PMAs may be host ICs and the PMDs, fiber optic modules. TP1, TP2, TP3 and TP4 are traditional labels for interfaces of a fiber optic link. TP0 and TP5 may be used as labels of the host IC interfaces.