Multilane MM Optics: Considerations for 802.3ba

John Petrilla

Avago Technologies

March 2008

Acknowledgements & References

- pepeljugoski_01_0108

Outline

- 802.3ba Alignment
- Elements for success & related challenges
- Eye safety category: Class 1 or Class 1M?
- Link Model: Example power & jitter budgets
- Module form factor, lane order & density
- Summary/Conclusions/Recommendations

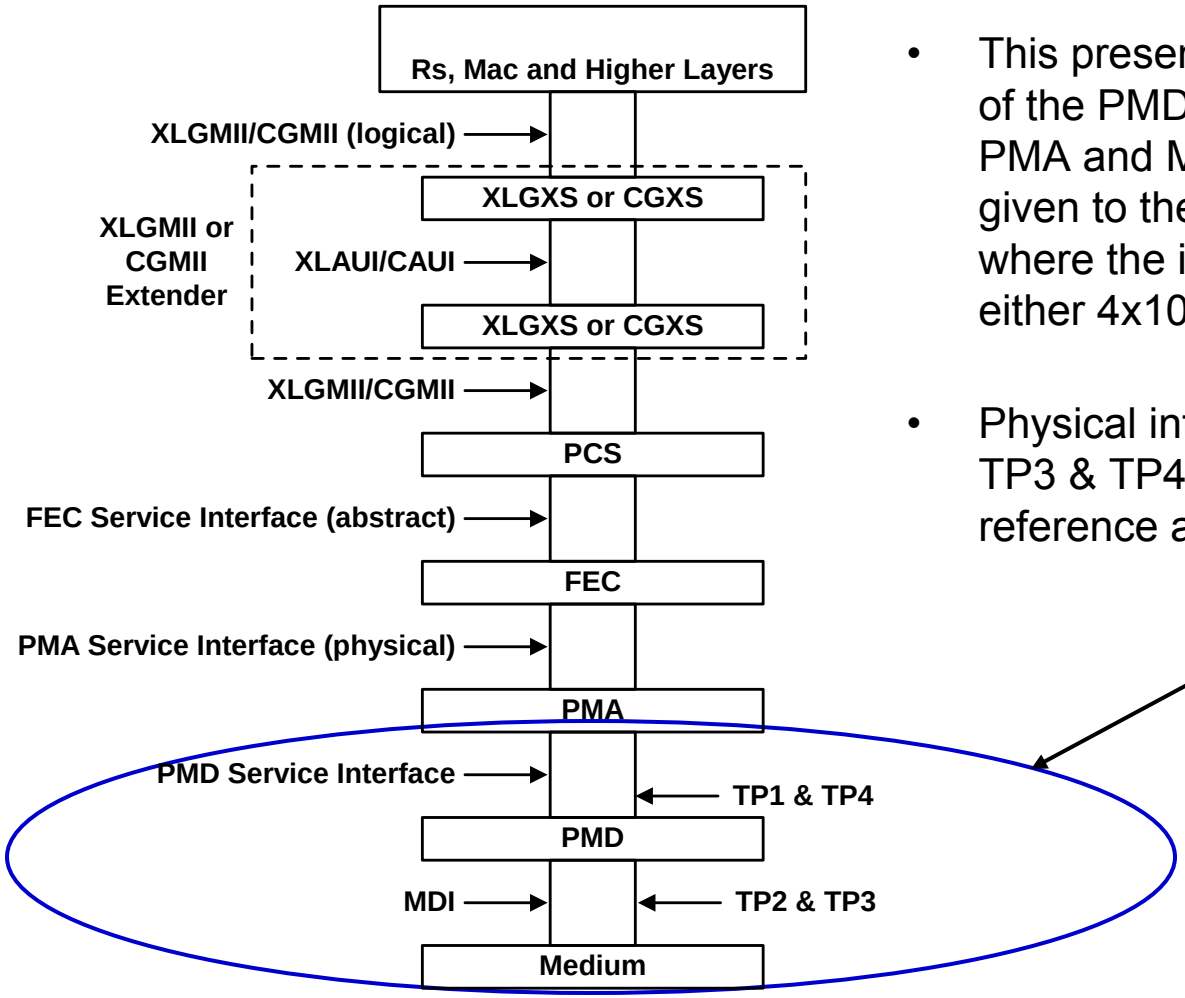
- The intention of this presentation is to work toward agreement on a suitable link budget and jitter allocation. Specifications can flow from this agreement.

802.3ba Alignment
802.3ba Objectives Addressed in Presentation

- Support MAC data rates of 40 Gb/s & 100 Gb/s
- Achieve better than or equal to 1E-12 BER at the MAC/PLS Service Interface
- 100 m on OM3 MMF

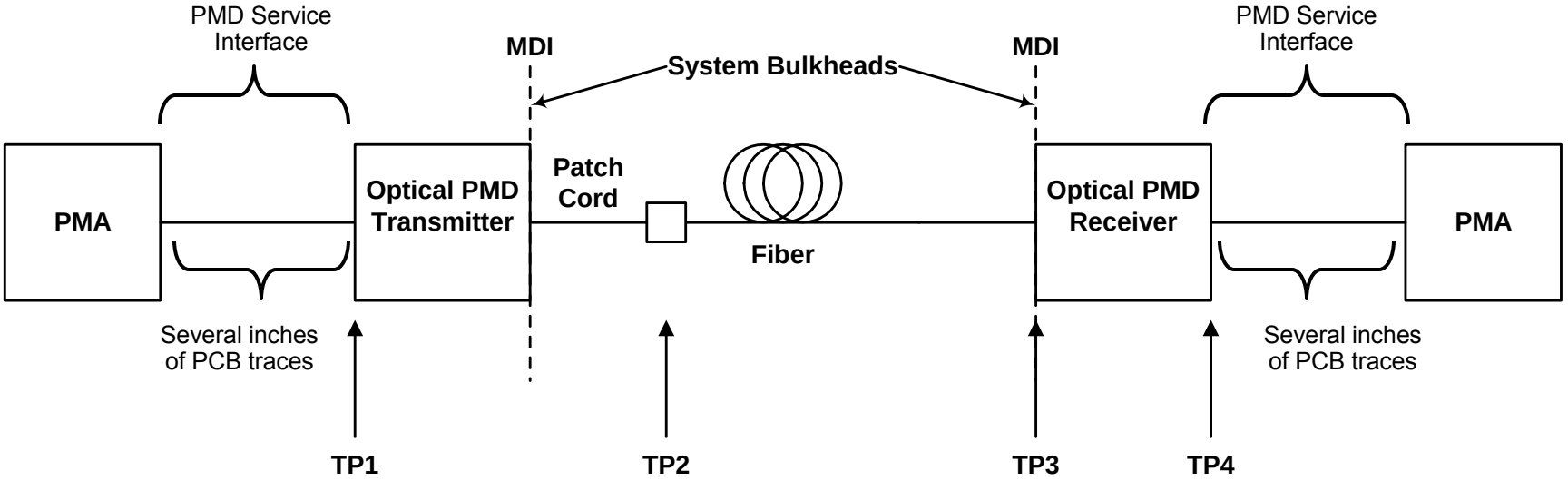
802.3ba Alignment

802.3ba Architectural Layers & Interfaces



- This presentation addresses aspects of the PMD and its interfaces to the PMA and Medium with emphasis given to the high speed signal paths where the interfaces are multilane, either 4x10G or 10x10G.
- Physical interface points TP1, TP2, TP3 & TP4 are identified for future reference and further defined below.

802.3ba Alignment
802.3ba PMD Block Diagram



- The above block diagram shows relevant elements and interfaces for a link between two PMAs. The patch cord is included for the definition of TP2. Otherwise intermediate fiber connectors are not shown.
- TP1, TP2, TP3 and TP4 are traditional 802.3 labels for interfaces of a fiber optics link. Two physical interfaces, TP1 on the left and TP4 on the right, are shown between the PMD and PMA elements. Here the PMA may be a host ASIC and the PMD may be a fiber optics module.
- P802.3ba should fully specify the signals at TP2 and TP3 for optical media and, at least, the jitter allocations at TP1 and TP4 for robust design, as in Gigabit Ethernet.

Elements for Success
Overview

- Total cost: less than ten/four 10 GbE solutions
- Power consumption: less than ten/four 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

Elements for Success
Cost

- 10GbE & 100GbE solutions will use the same technology but multilane devices have ~ 1.4 dB/lane lower Class 1 eye safety category limits than single lane 10GBASE-SR.
- Therefore, unless specifications are changed, test costs are not expected to improve from the unfortunate 10GBASE-SR cost point and will actually degrade.
  - Test costs (which includes calibration and programming) can easily dominate.
  - Relaxed parameters are needed to simplify testing and improve yield.
  - 10GBASE-SR is not a satisfactory basis for multilane specifications.
- Module Costs are strongly impacted by testing and yield.
  - Targets requiring narrow operating ranges force over temperature tests.
  - Relaxed spectral width specs when compared with 10GBASE-SR can improve VCSEL array yield.
  - Moving to Class 1M from Class 1 opens a narrow operating range and reduces module cost by saving on testing.
- Costs and time-to-market will benefit from reuse of 10GbE technology, e.g. VCSELs, 64b/66b style encoding & physical-layer IC technology, and by sharing technology and development cost with InfiniBand.
- Low costs are not achieved by setting specifications based on best-of-breed but based instead on industry-wide capabilities.

Elements for Success
Power

- 10GbE & 100GbE solutions will use the same device technology.
- Therefore, unless specifications are changed, power consumption is not expected to change.*
- Optical transmitter power consumption is driven by input sensitivity, output transition times and jitter requirements.
  - Relaxed* transition times and jitter specs would constrain power consumption.
  - Multilane interfaces add crosstalk, thereby increasing jitter contribution and leading to a need for greater allocation of jitter to the transmitter. Unfortunately, the same applies to the upstream device.
- Optical receiver power consumption is driven by output signal level amplitude, bandwidth, input dynamic range requirements, and any additional features such as signal detect, maintaining a linear signal path and inclusion of equalization or CDR functions.
  - Multilane interfaces add crosstalk, thereby reducing RX sensitivity, increasing RX jitter contribution and leading to a need for greater allocation of jitter to the receiver. Unfortunately, the same applies to the downstream device.

* When compared with 10GBASE-R.

Elements for Success
Density

High density brings challenges for heat transfer (power dissipation) and signal integrity (inter-lane crosstalk). Some existing form factors may be useful or offer an example.

Form Factor Attributes and Projections			
	SFP(+)	QSFP	Plausible 10 Lane FF
Module + Cage Dimension	15 mm W x 12 mm H x 59 mm D	19 mm W x 14 mm H x 79 mm D	25.4? mm W x 17? mm H x 79? mm D
Horizontal Port Pitch, mm	16.25 mm and stackable	21.0 mm and stackable	27? mm, stackable?
Power Level	1.0 W Power Level 1	2.0 W Power Level 2	5.0 W
Optical Connector	Dual LC	MPO: 1 x 12	MPO: 2 x 12
Electrical Connector	20 contact single-edge, double-sided	38 contact single-edge, double-sided	84? contact 2-edge stack, double-sided

Relative to SFP horizontal port pitch, QSFP offers a density advantage of 3.1x GBd/mm and the Plausible 10 Lane FF offers a 6.0x GBd/mm advantage. Both seem attractive

Elements for Success
Reliability

- 10GbE & 100GbE solutions will use the same device technology.
- Consequently, means to improve operating life and/or reliability may be limited to higher levels of integration and by constraining device temperature.
- Specifications that permit low power consumption will also benefit operating life.

Elements for Success
Cable Plant

- Link Length: 100 m is more than sufficient to cover all distances in HPC environment, almost 100% of Enterprise Data Center Client-to-Access Channels, >90% of Enterprise Data Center Access-to-Distribution Links and almost 85% of Enterprise Data Center Distribution-to-Core Channels. [flatman_01_0108]
  - OM3 minimizes cost.
  - Longer distances should not come by using tighter specs on the modules but better fibers (i.e. OM4) and connectors and/or other means such as FEC.

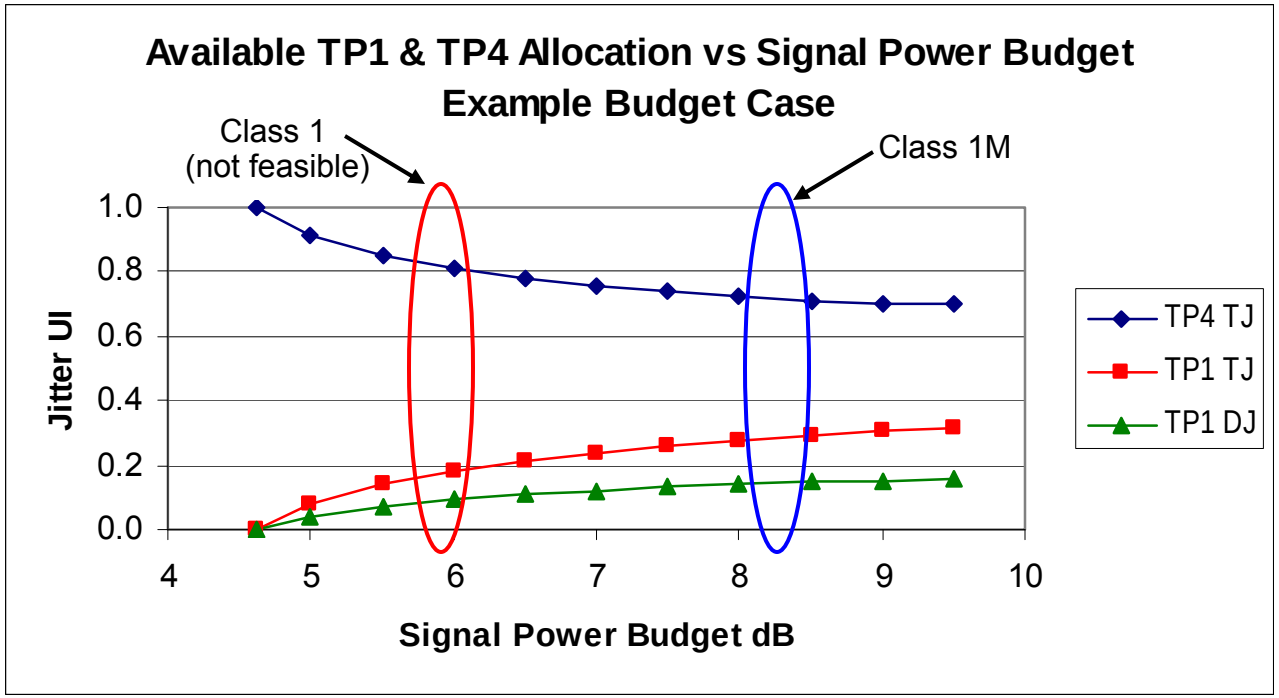
- Number of Intermediate Connector Pairs: up to 4 (twice as many as the reference topology for previous 802.3 specifications)
  - Connectors interfacing to modules are not counted.
  - Parallel connectors have higher loss than single connectors.
  - It is necessary to limit both aggregate loss and individual loss at each connection to constrain impact on modal noise and link performance.

Elements for Success
Summary – The right design point is crucial

- Users expect improvement in cost, power, density, and reliability per GBd relative to 10GBASE-SR in the SFP+ form factor.
- Multilane devices will have ~ 1.4 dB/lane lower Class 1 eye safety category limits.
- Since the same device technology will be used in 10GbE and 100GbE multilane products, 10GBASE-SR specs are not a satisfactory basis for multilane specification.
  - Unless specs are changed*, test and yield costs per Gb/s will not improve.
  - Unless specs are changed*, power consumption per Gb/s will not improve.
- Inter-lane crosstalk can adversely impact signal quality and appearance.
  - Impact on link cost and power consumption can be avoided with suitable allocations in signal budgets and proper attention in design.
- Operating life and/or reliability can be adversely impacted by device temperature.
  - Specifications that permit low power consumption will also benefit operating life.
- While port density is often limited by IO count, QSFP appears to be an attractive form factor for 40GbE MMF transceiver variants and there's an attractive plausible form factor for 100GbE MMF transceivers.
- Low costs are not achieved by setting specifications based on best-of-breed but based instead on industry-wide capabilities. Keep it simple!

*When compared with 10GBASE-R specs

Eye Safety Category
Class 1 or Class 1M?



- Existing parallel modules are normally Class 1M.
- For parallel optics Class 1 eye safety max limits can be reduced by 1.4 dB compared to single source modules.
- For more detail see petrilla_02_0308.

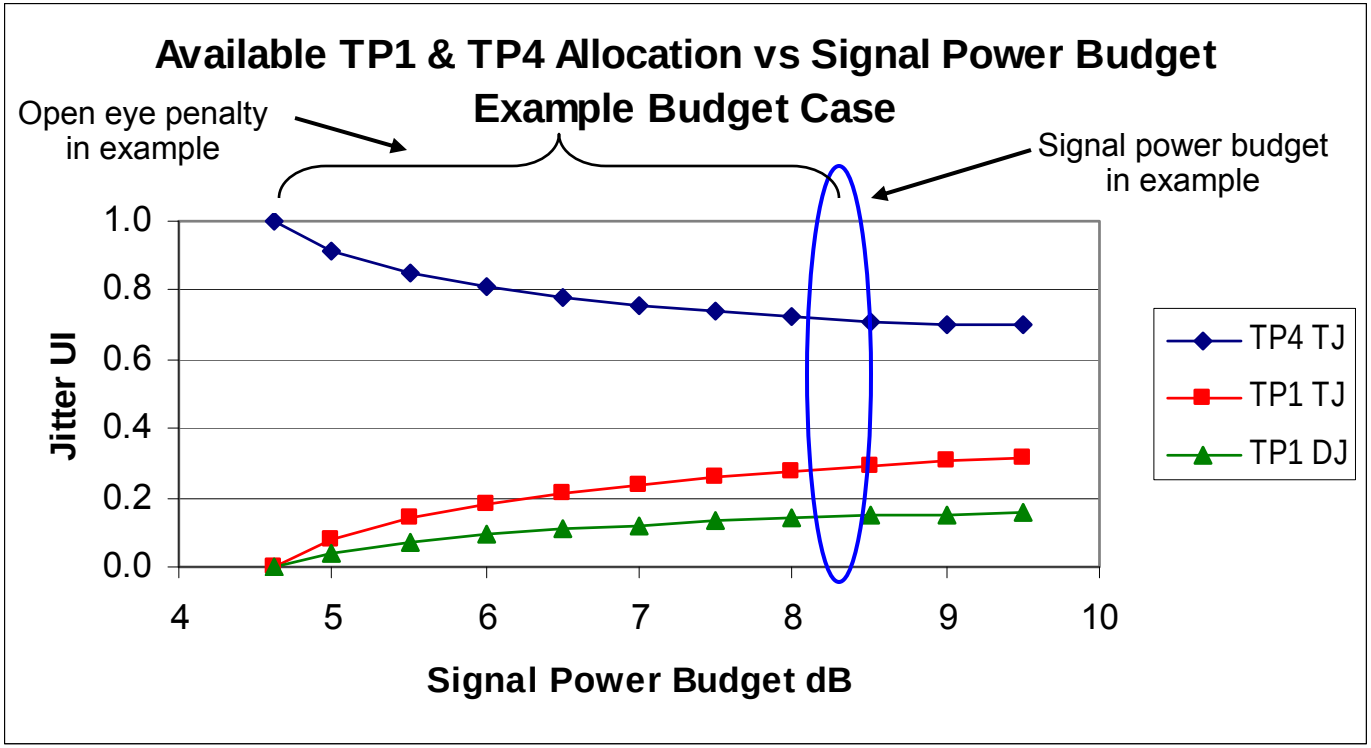
- The above chart shows the tradeoff expected between possible jitter allocations at TP1 and TP4 and the available link signal power budget. The available allocation is arbitrarily split evenly between TP1 and TP4.
- Existing Class 1 eye safety category maximums limit the signal power budget to less than 6 dB requiring unrealistic jitter allocations for the host at TP1 & TP4 as shown in the above chart. Class 1M permits more practical budgets.

Link Models

- The 10GbE link model, 10GEPBud3_1_16a, available at http://www.ieee802.org/3/ae/public/adhoc/serial_pmd/documents/ was used to determine link attributes presented in the following pages.
- Jitter in the following pages follows dual-Dirac methodology and, where used, DJ is intended to be dual-Dirac DJ.
- Since the 10GbE link model is open, available to all and reasonably well-regarded, it is a useful tool for comparing various proposals and tradeoffs among attributes.

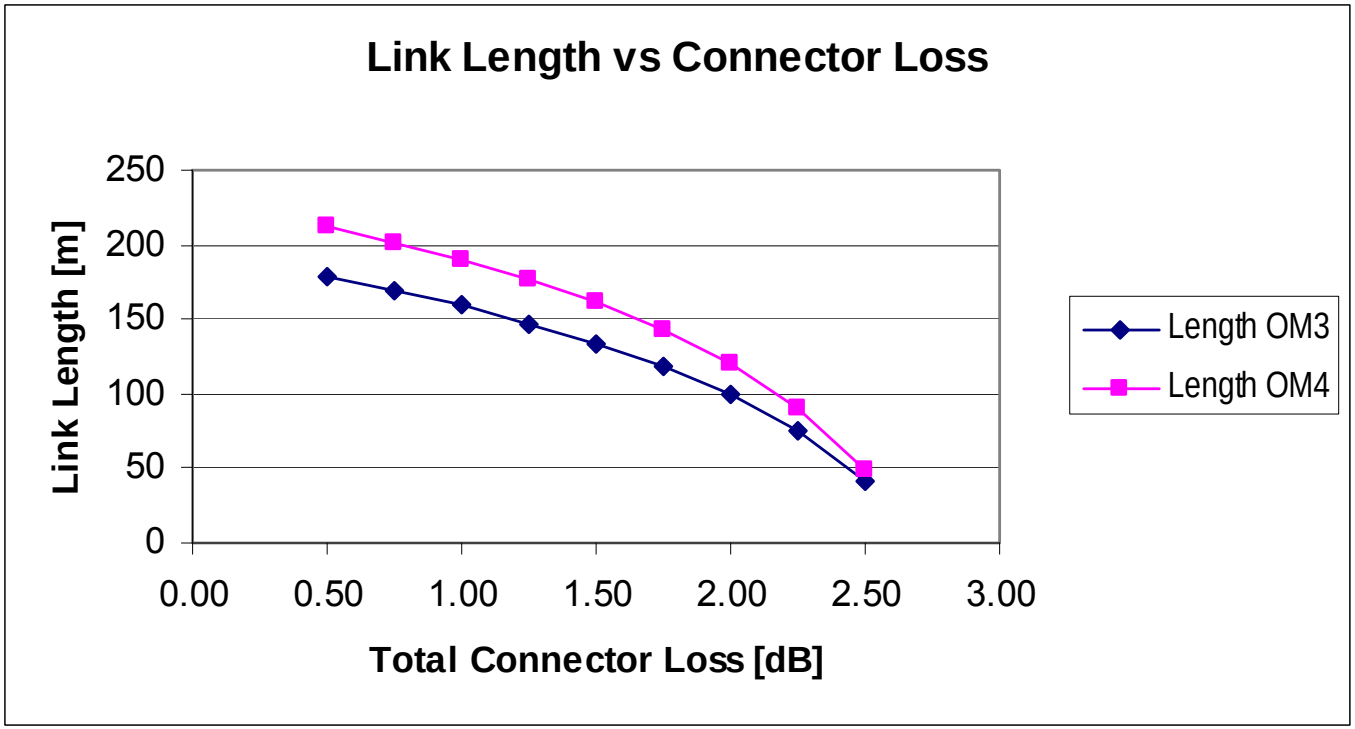
Link Model Results

TP4 Eye Opening



- The example budget provides a 8.3 dB signal power budget and can support jitter allocations of TP1(DJ) = 0.143 UI, TP1(TJ) = 0.286 UI and TP4(TJ) = 0.716 UI.
- Use of FEC may provide some relief in the struggle to allocate jitter properly.

Link Model Results
Link Length & Connector Loss



- The example budget supports 100 m of OM3 with 2.0 dB connector loss.
- Longer reaches are possible with reduced connector loss budgets, improved fiber and/or use of FEC.

Link Model Attributes
Example Link & Cable Plant

- Signal Rate: 10.3125 GBd ± 100 PPM
- BER: < 10^{-12}
- Signal Power Budget: 8.3 dB
- 100 m of OM3
- 2 dB connector loss allocation for 4 MPO inline connections
- Center Eye Penalties
  - Attenuation = 0.36 dB
  - Pisi = 1.43 dB
  - Pdj = 0.20 dB
  - Pmn = 0.40 dB
  - Pmpn = 0.03 dB
  - Prin = 0.24 dB
  - Pcross = 0.08 dB
- 0.284 UI Open Eye Penalty = 3.55 dB (see figure on page 19)

Link Model Attributes
Example Tx Power & Jitter Budget

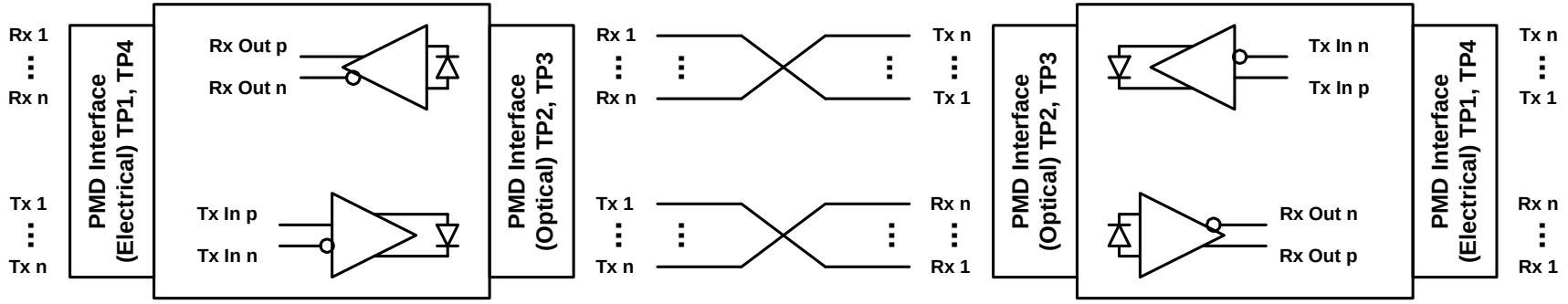
- Eye Safety Category: Class 1M
- Max Pave: 1.0 dBm
- Min OMA: -3.0 dBm
- Min ER: 3.0 dB
- Center Wavelength Range: 840 to 860 nm
- Max RMS Spectral Width: 0.65 nm
- Max Transition Time (20%, 80%): 35 ps
- Max RIN12OMA: -128 dB/Hz
- RIN Coefficient: 0.70
- Mode Partition Noise Coefficient: 0.30
- Min Tx Reflection Tolerance: 12 dB
- TP1 Jitter Allocation: DJ = 0.143 UI, TJ = 0.286 UI
- TP2 Jitter Allocation: DJ = 0.277 UI, TJ = 0.504 UI

Link Model Attributes
Example Rx Power & Jitter Budget

- Max Sensitivity: -11.3 dBm
- Max Input Power: 1.0 dBm
- Min Bandwidth: 7500 MHz
- RMS Base Line Wander: 0.010
- Max Rx Reflection: 12 dB
- TP3 Jitter Allocation: DJ = 0.282 UI, TJ = 0.544 UI, DCD = 10.0 ps
- TP4 Jitter Allocation: DJ = 0.371 UI, TJ = 0.716 UI

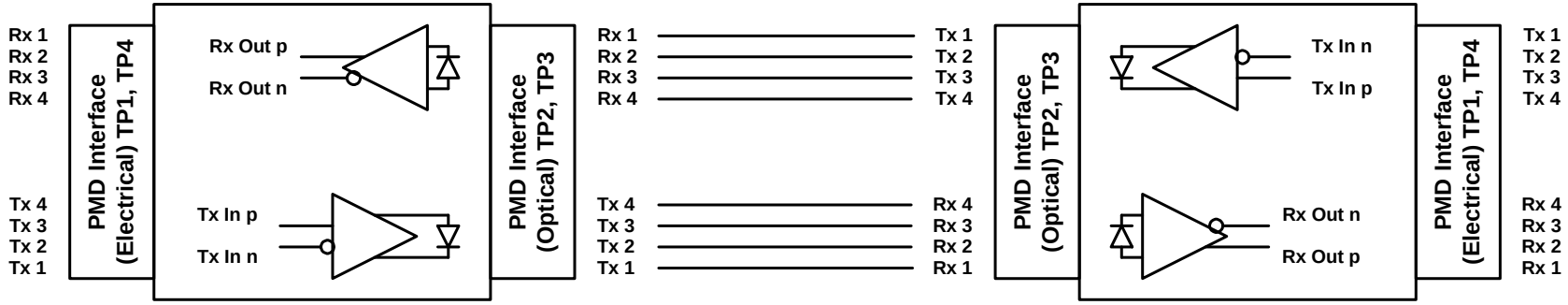
Lane Order

Multilane Variant Overview



- Some lane order choices will force lane crossovers. This can be problematic, e.g. requiring special cables, but can usually be avoided. Lane recognition and reordering at higher levels in the architecture may be part of the solution.
- QSFP and POP4 form factors provide examples of lane ordering suitable for a four lane transceiver that fits fiber cable industry practice.
- There are no popular form factors of ten lane or twelve lane optics transceivers to use as examples. Double row MPO connectors are not widely deployed. Side-by-side MPO connectors are unlikely to meet density requirements. Lane ordering for two row MPO connectors may be an item for study for 802.3 and IEC/TIA.
- Twelve and sixteen lane electrical cable connectors have been defined for InfiniBand and PCI Express and may provide some guidance for the electrical interface.

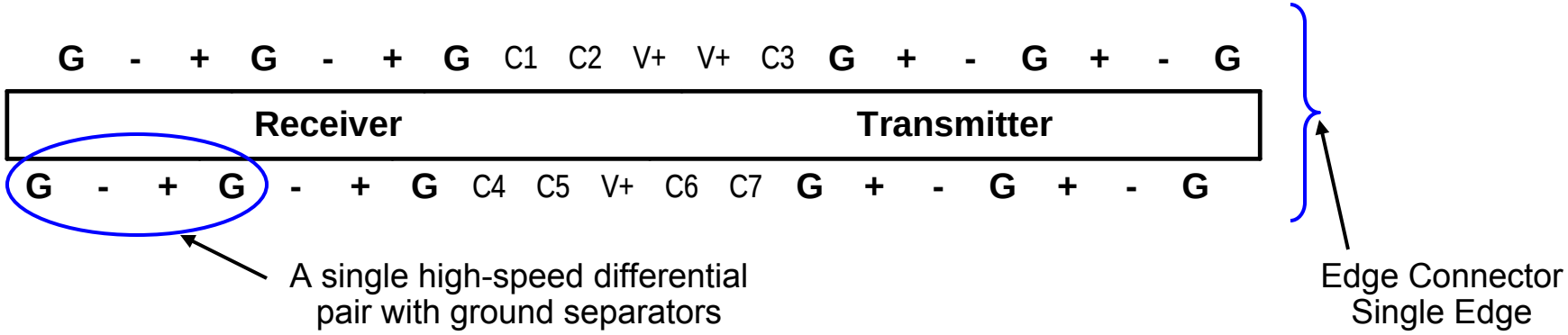
Lane Order – Four Lane Transceivers
40 GbE & QSFP



- The lane order shown above is similar to that chosen by InfiniBand and QSFP. No crossovers are required between the electrical interfaces of the modules. A net zero degree rotation in the fiber plant is required. The required fiber cable plant is currently supported in the industry.
- QSFP with a 21 mm horizontal pitch can offer 3.1x the horizontal density (GBd/mm) of an SFP form factor.

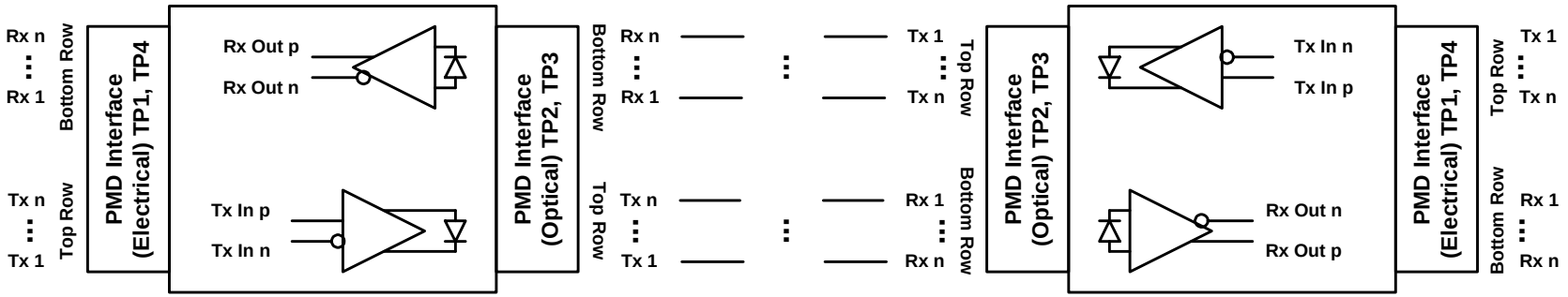
Lane Order – Four Lane Transceivers

40 GbE PMD Module Electrical Interface



- QSFP provides an example of lane ordering suitable for a four lane transceiver. Crosstalk at the electrical interface is minimized by separating the transmit lanes from the receive lanes. 38 contacts provide eight high-speed differential signal pairs, including ground separators, three supply inputs and six control signals.
- Time-to-market and market acceptance may be accelerated if there is a common electrical interface with other variants using the QSFP form factor and this should be considered.

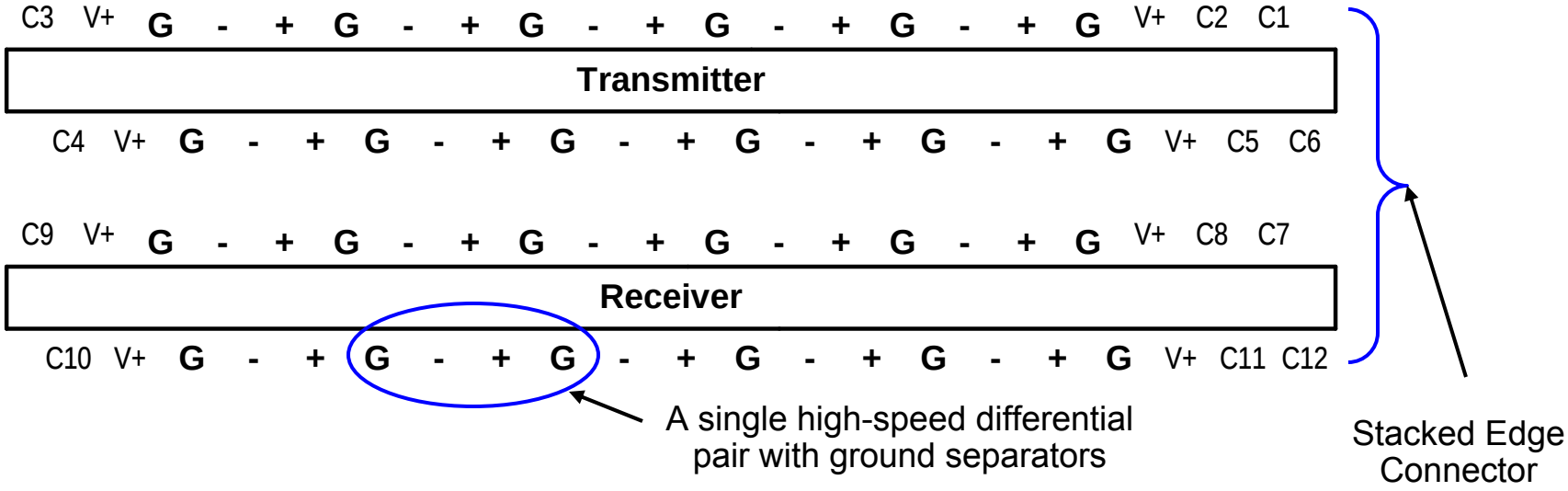
Lane Order – Ten Lane Transceivers
100 GbE



- There are no popular form factors of 10 lane or 12 lane optical transceivers to use as examples.
- Density requirements argue for a double-row MPO optical connector and a stacked edge electrical connector based module. Such a module may have a 27 mm horizontal pitch and offer 6.0x the horizontal density (GBd/mm) of an SFP form factor.
- A net 180 degree rotation is required in the cable plant if module transmitter outputs are aligned in one row and module receiver inputs are aligned in the other (e.g. if the cable is formed from two 12 lane ribbons, then both ribbons, as a pair, together rotate). Lane ordering to ensure this may be an item for study for 802.3 and IEC/TIA.

Lane Order

100 GbE PMD Module Electrical Interface



- A 84 contact stack edge connector is being considered in other applications for 12 lane copper cable assemblies. That connector provides twenty high-speed differential signal pairs, including ground separators, eight supply inputs and twelve control signals. Not all control signals may be needed and it may be possible to downsize the connector to 80 contacts.
- It is not yet certain that stacked edge connectors can support the signal integrity requirements of 10+ GBd signals and additional evaluation is needed.

Multilane Optics: Budgets & Other Considerations
Summary/Conclusions/Recommendations (1)

- Cost, power consumption, reliability and cable plant elements for success can be met if appropriate design points, especially for link length, connector loss and jitter allocation, are chosen. This also includes adoption of Class 1M Eye Safety category limits and generating a signal power budget ≥ 8.3 dB. With appropriate choices a limiting receiver without heavyweight equalizers or in-module CDR can provide satisfactory performance for reasonable lengths of PCB traces. Keep it simple!
- Density elements for success for 40GbE can be met using the QSFP form factor. A common electrical interface with other variants considering the QSFP form factor should be considered.
- Density elements for success for 100GbE appear to require a double-row MPO fiber connector and a stacked-edge electrical connector. Appropriate deployment of each requires investigation and early effort here will be valuable. A common electrical interface with other variants considering a similar form factor should be considered.
- The signals at TP2 and TP3 should be fully specified for optical media and, at least, the jitter allocations at TP1 and TP4 for robust design, as in Gigabit Ethernet.
- FEC should be considered as a means to provide jitter allocation relief.
- The 10GbE link model, 10GEPBud3_1_16a, should be used in development of 40GbE & 100GbE MMF variants' specs to provide robust physical specifications.

Multilane Optics: Budgets & Other Considerations
Summary/Conclusions/Recommendations (2)

- Lane order for 40GbE should follow the QSFP form factor.
- Lane order for 100GbE should consider a form factor based on a double-row MPO fiber connector and stacked electrical connector. Lane ordering and cable topologies should be studied in concert.
- Stacked-edge electrical connectors should be evaluated to see if they supports signal integrity requirements of the high speed electrical signals.