Intermediate SMF reach option for 100GE

John Petrilla: Avago Technologies

March 2012

Contributors & Supporters

Contributors:

Jon Anderson, Opnext

Tom Palkert, Luxtera

I-Hsing Tan, Avago Technologies

Doug Coleman, Corning

Paul Kolesar, CommScope

Supporters:

Nic Bustamante, Microsoft

Mike Hughes, USConec

Sharon Lutz, USConec

Scott Kipp, Brocade

Tom Issenhuth, Microsoft

Max Lawler, Microsoft

Outline

This presentation provides :

- Conclusion
- Intermediate Reach Overview
  Issue – Consequence – Solution - Status
- Avago-Opnext-Luxtera Parallel SMF (PSM) Intermediate Reach Proposal
- Relative Cost Analysis
- Data Center Acceptance of Parallel SMF
- Summary & 5 Criteria

- PSM is used as an abbreviation for Parallel Single Mode.
- PSM4 is used as an abbreviation for 4-lane Parallel Single Mode

Conclusion

A 100 Gb/s (4 lane) PMD capable of operation over 2 km of multi-lane SM fiber (G.652) with BER < 10^{-12} based on 1310 nm lasers, NRZ modulation and 64b/66b encoding offers significant module and total system cost advantages relative to 100GBASE-LR4 for reaches of at least 1 km.

100G Ethernet Status

802.3ba, in 2010, provided definitions enabling 100G Ethernet. However several of the defined interfaces defined in 802.3ba are 10-lane-10G/lane interfaces that are now seen as a limiting factor for increasing front panel port density and reducing power consumption and cost. A shift to 4-lane-25G/lane interfaces may relieve the 10-lane-10G/lane constraints and may be feasible in the appropriate time frame.

To address the 10-lane-10G/lane constraints, the 802.3bj task force was established to address backplanes and copper cable assemblies and the Next Generation 100Gb/s Optical Ethernet Study Group was established to consider chip-to-chip and chip-to-module electrical interfaces and short-wavelength-short-reach and long-wavelength-intermediate-reach optical interfaces.

Issue-Consequence-Solution-Status (1)

Issue: For 100G Ethernet electrical interfaces only 10-lane-10G/lane electrical interfaces are defined.

Consequences:

Host IC IO channels, PCB routing, transceiver electrical connectors have to be sized to accommodate 10-lane channels impacting density, power and cost

10:4 gearboxes are required to interface 10-lane-10G/lane channels with 4-lane-25G/lane channels impacting power and cost

Solution: Define 4-lane-25G/lane chip-to-chip and chip-to-module electrical interfaces relieving host IC port density constraints, PCB routing complexity, eliminating 10:4 gearboxes, lowering power consumption, enabling smaller form factors

Status: Accepted as objective by the Next Generation 100Gb/s Optical Ethernet Study Group

Issue-Consequence-Solution-Status (2)

Issue: For 100G Ethernet short-wavelength interfaces, only 10-lane-10G/lane short-wavelength-short-reach optical interfaces are defined.

Consequences: Optical fiber plant has to be sized to accommodate 10-lane channels impacting density, power and cost

Solution: Define 4-lane-25G/lane optical interfaces relieving front panel port density constraints, fiber costs and routing complexity, lowering power consumption, enabling smaller form factors

Status: Ongoing discussion in Next Generation 100Gb/s Optical Ethernet Study Group - See 100G SR4 presentations

Issue-Consequence-Solution-Status (3)

Issue: Beyond the 150 m reach supported by 100GBASE-SR10, the next 100G Ethernet defined reach is 10 km provided by 100GBASE-LR4. The differences in density, power consumption and cost structures between SR10 and LR4 are seen as too large. That is, data center reaches less than 1 km do not have an 100G Ethernet solution where the density, power consumption and cost structures scale with reach from SR solutions as expected based on prior Ethernet generations.

Consequences: 100G Ethernet adoption in the access market (switch-to-switch) may be delayed.

Solution: Adopt 4-lane-25G/lane electrical interfaces to reduce the power consumption associated with 10-lane-10G/lane electrical interfaces. Define intermediate reach optical interfaces to lower the power consumption and costs associated with 10 km optical reach support.

Status: Ongoing discussion in Next Generation 100Gb/s Optical Ethernet Study Group

Intermediate Reach: PSM4 Proposal

Avago-Opnext-Luxtera Intermediate Reach Proposal: A 100 Gb/s (4 lane) PMD capable of operation over 2 km of multi-lane SM fiber (G.652) with BER < 10^{-12} based on 1310 nm lasers, NRZ modulation and 64b/66b encoding offers significant module and total system cost advantages relative to 100GBASE-LR4. While technically capable of supporting a 2 km reach, a 1 km reach objective is proposed for economic considerations.

References: anderson_01_1111_NG100GOPTX, palkert_01_1111_NG100GOPTX, petrilla_01_1111_NG100GOPTX, anderson_01_0112_NG100GOPTX, anderson_02_0112_NG100GOPTX, kipp_01_0112_NG100GOPTX , petrilla_02a_0112_NG100GOPTX

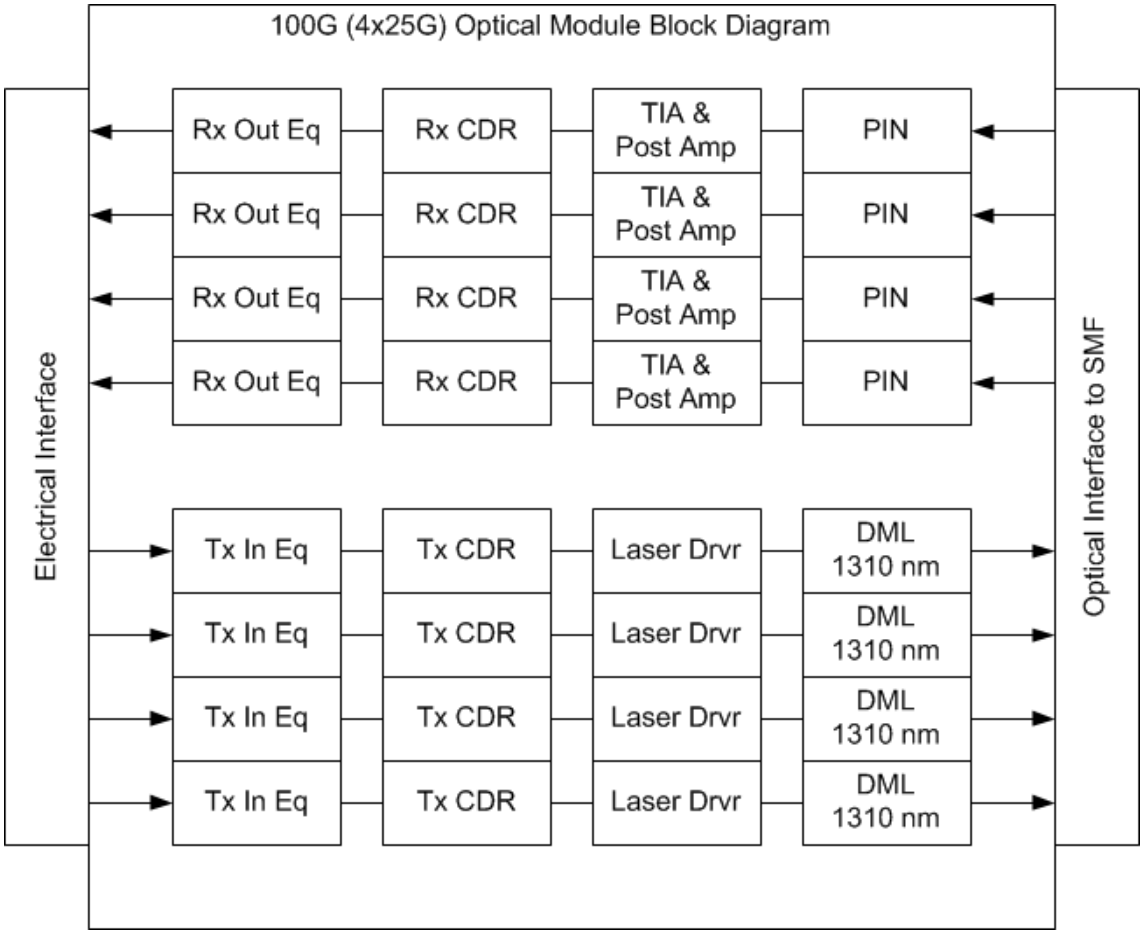
A 4-lane PSM solution trades off a lower cost module with a higher cost/m fiber plant. This tradeoff will be explored in subsequent pages.

Relative cost estimates, see references, for PSM4/LR4 range from 0.15x to 0.5x

The Kolesar Kalculator (Kolesar_Kalculator_2012_02_29.xls) was used to provide comparative total data center costs for various scenarios of relative module costs and supported reaches. Results follow.

Working with Corning, various cable plant scenarios were considered to explore crossover reach points for the existing and proposed 100G PMDs. Results follow.

100G 25G/Lane PSM4 Transceiver: Description [1/2]



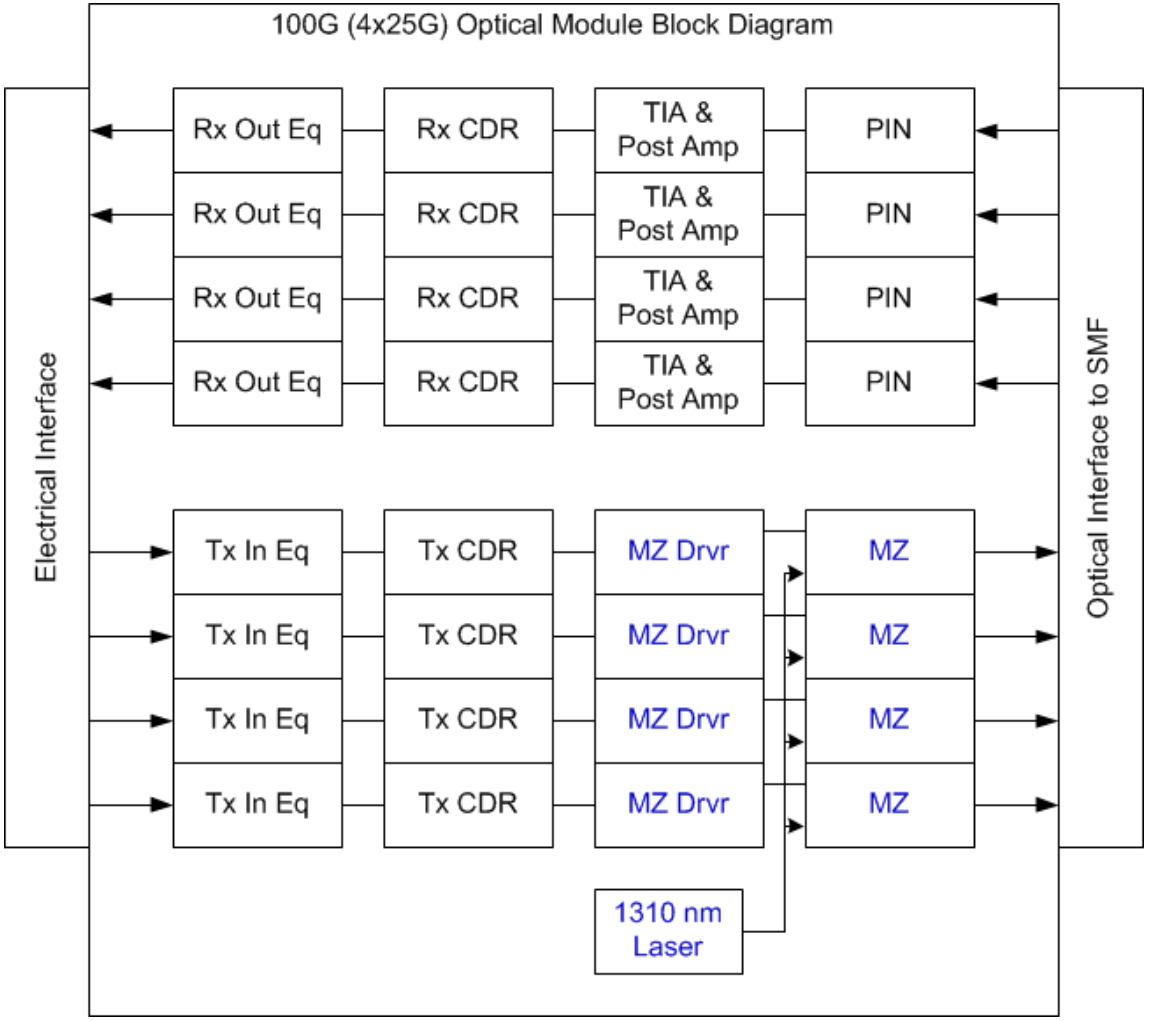
MPO connectors are assumed for the optical interfaces.
NRZ modulation and 64b/66b encoding is expected for both electrical and optical signals.
Other block diagrams are possible.

<u>100GBASE-LR4 Comparison</u>
- No optical MUX
- No optical DMUX
- No tight wavelength requirements
- No TEC

Since this device can be viewed as a simplified and/or reduced cost 100GBASE-LR4 for a reduced reach, technical feasibility has been demonstrated by the LR4.

- <u>100GBASE-SR4 Comparison</u>
- 1310 nm DFB lasers replace 850 nm VCSELs
- PIN Photodiode responsivity shifts from 850 nm to1310 nm
- No optical link equalization required
- No FEC required

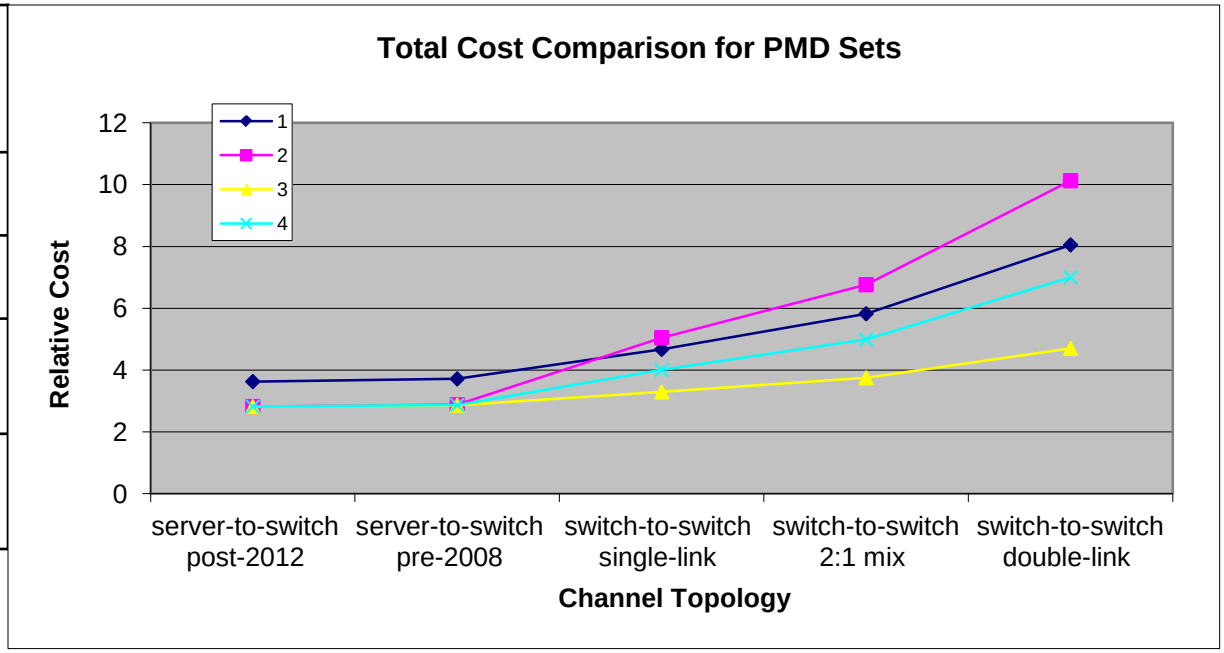
100G 25G/Lane PSM4 Transceiver: Description [1/2]



- Alternative block diagrams are possible, e.g. single CW laser and optical modulators combination, that may lead to lower cost and power consumption.

Kolesar Kalculator Analysis

Case #	Case	Rel. 100G Ch. Cost	Reach OM3	Reach OM4	Reach OS2
1	SR10 LR4	1 10	100 m	150 m	 10 km
2	SR4 LR4	1 10	70 m	100 m	 10 km
3	SR4 PSM4 LR4	1 2 10	70 m	100 m	 1 km 10 km
4	SR4 PSM4 LR4	1 5 10	70 m	100 m	 1 km 10 km



The above chart shows data center costs comparisons generated with the Kolesar Kalculator (Kolesar_Kalculator_2012_02_29.xls) for the reach distributions built into the calculator. Four 100G/Ch. cases were analyzed. See above table for definitions.

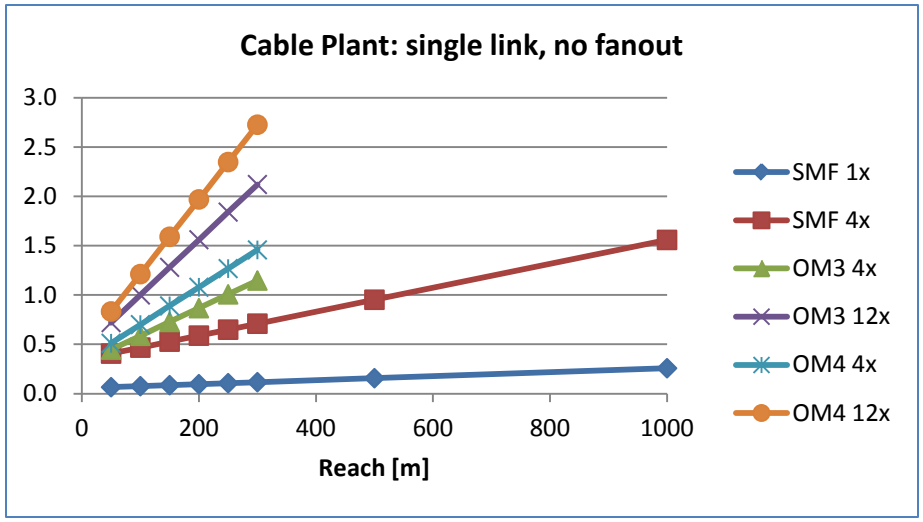
Case 1 is the incumbent case where only 100GBASE-SR10 and 100GBASE-LR4 exist.

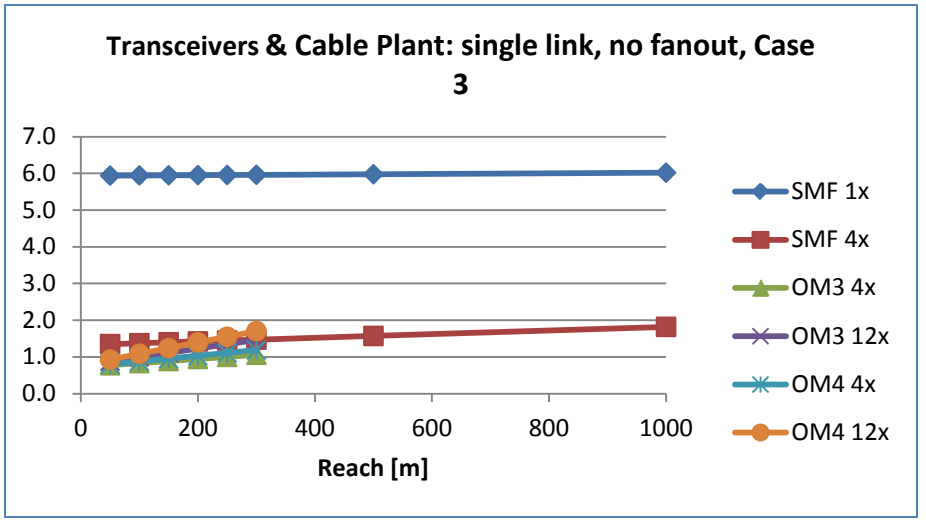
Case 2 replaces SR10 in Case 1 with SR4. SR4 OM3 and OM4 reaches are in discussion. The values used here are for example.

Cases 3 & 4 add PSM4 to Case 2 with two different relative cost assumptions for PSM4: Case 3 (2x SR10 for 4 MZ) and Case 4 (5x for 4 DML).

Inclusion of PSM4 offers an advantage relative to the incumbent (SR10 LR4) case as well as the SR4 LR4 case where reaches beyond 100 m are expected for relative PSM4/LR4 costs in the range of 0.2 to 0.5.

100G Crossover Analysis – Relative Cable Plant & Transceiver Cost

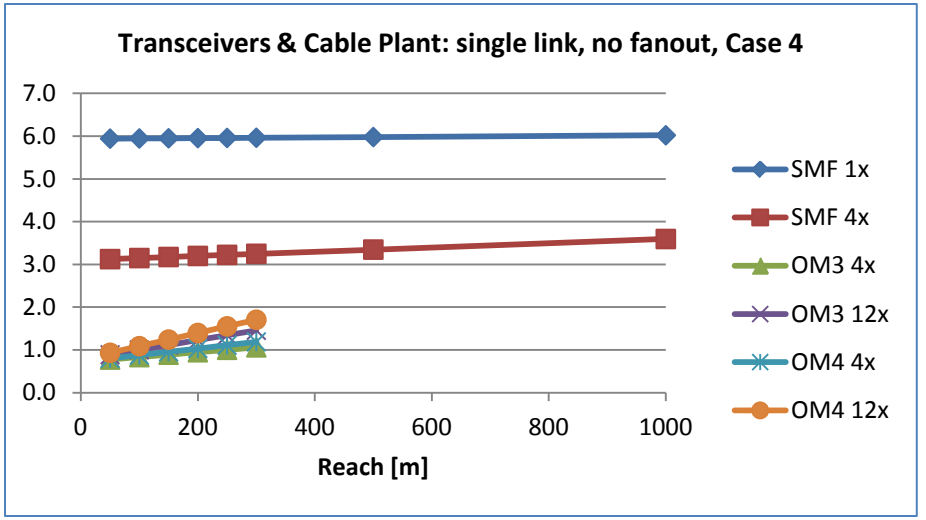




The charts shows relative cable and transceiver costs for a 100G channel based on estimates from Corning and petrilla_02a_0112_NG100GOPTX (page 12)

This scenario reflects the simplest cable plant structure using the same relative transceiver cost assumptions as used in the Kolesar Kalculator analysis.

The PSM4, 4-lane parallel, option (SMF 4x) offers a clear economic advantage to LR4 (SMF 1x) for this simple cable plant scenario.



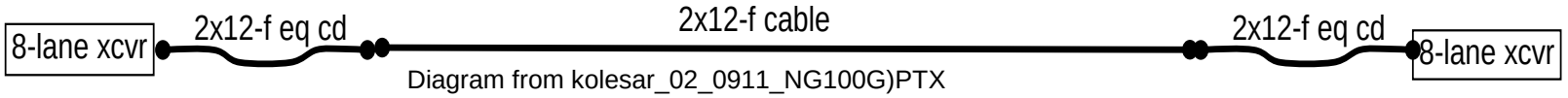
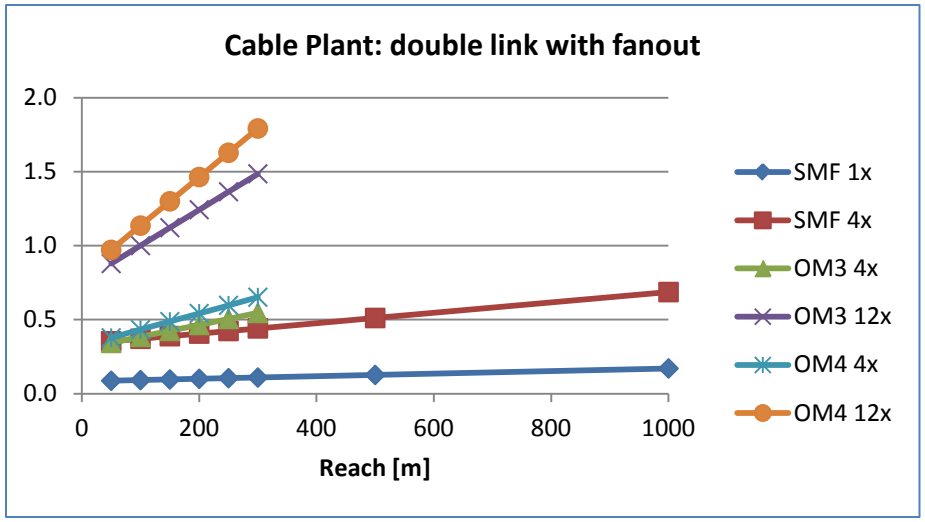
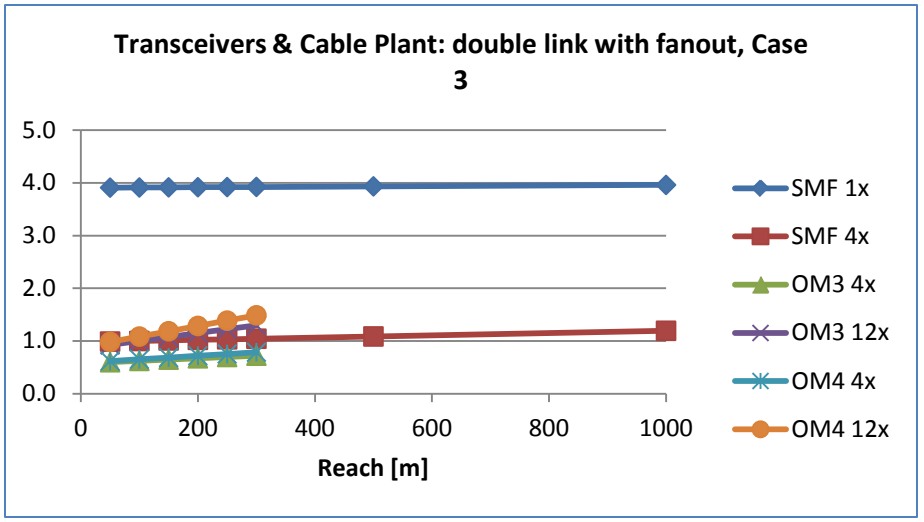


Diagram from kolesar_02_0911_NG100G)PTX

100G Crossover Analysis – Relative Cable Plant & Transceiver Cost

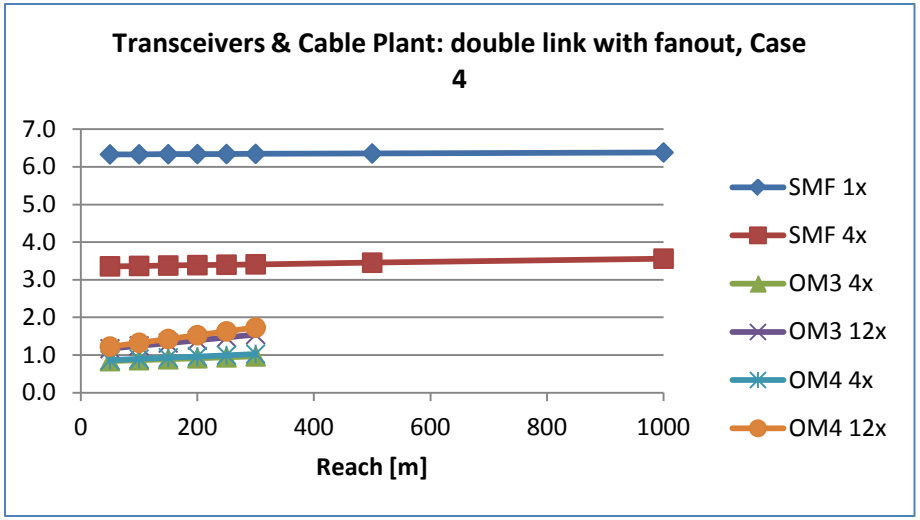




The charts shows relative cable and transceiver costs for a 100G channel based on estimates from Corning and petrilla_02a_0112_NG100GOPTX (page 12)

This scenario reflects a more complex cable plant structure using the same relative transceiver cost assumptions as used in the Kolesar Kalculator analysis.

The PSM4, 4-lane parallel ,option (SMF 4x) offers a clear economic advantage to LR4 (SMF 1x) for this complex cable plant scenario.



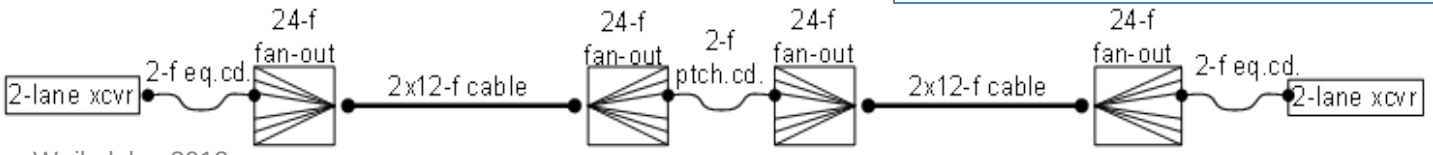


Diagram from kolesar_02_0911_NG100G)PTX

Acceptance of parallel SMF in data centers

Parallel fiber use is not new in the data center. Ethernet adopted MMF objectives for 802.3ba in July 2007 for 40GBASE-SR4 and 100GBASE-SR10 that depend on parallel fiber.

InfiniBand used 4-lane channels (4x) and 12-lane channels (12x) implementations beginning circa 2000 at 2.5G then 5.0G and 10G rates and expects to continue at 14G and 25G

Concerns associated with bending ribbon cables were resolved with the introduction of circular cross section assemblies.

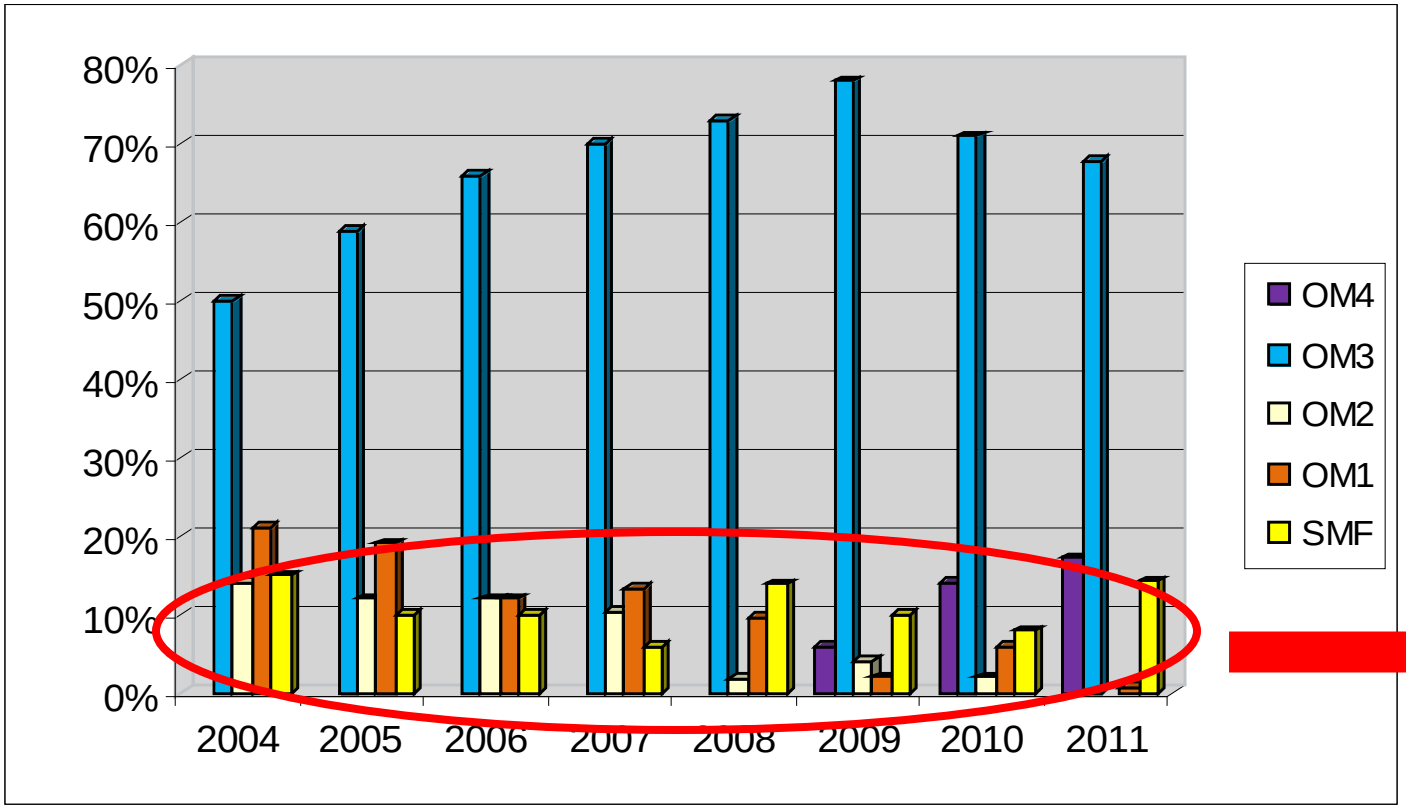
Industry estimates and examples of installed parallel SMF in data centers are found in the following pages.

Survey of four top-ranked internet datacenter operators provided the following feedback .

a. 3 out of 4 agreed with the PSM4 proposal and would purchase the products when available. The one not in favor of the proposal would use PSM4 to replace SR4 if cost parity was reached.

b. Two of the surveyed companies would like to see the SMF PSM4 concept used today to aggregate four 10GE transceivers into one module and/or as an alternative implementation of 40 GE followed by the 4x25G optics in 2014 – 2015.

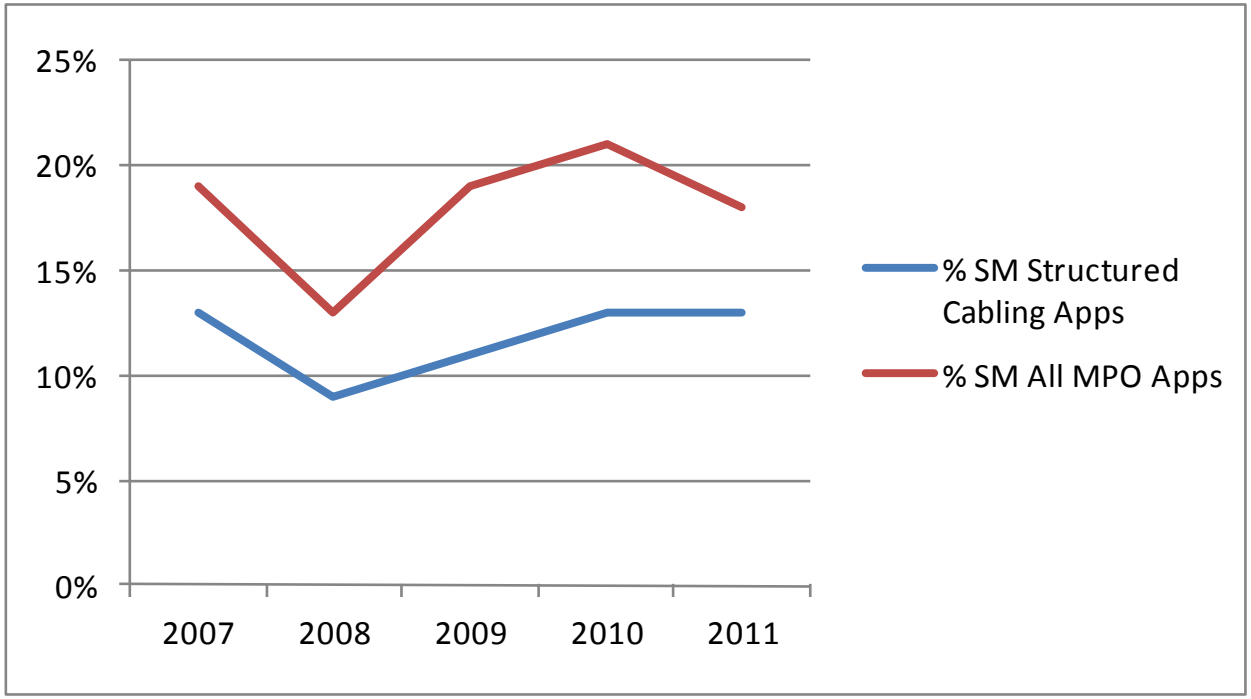
Data Center Fiber Environment



SMF =<10%

Source: Corning

Corning Estimated SMF MPO Percentage



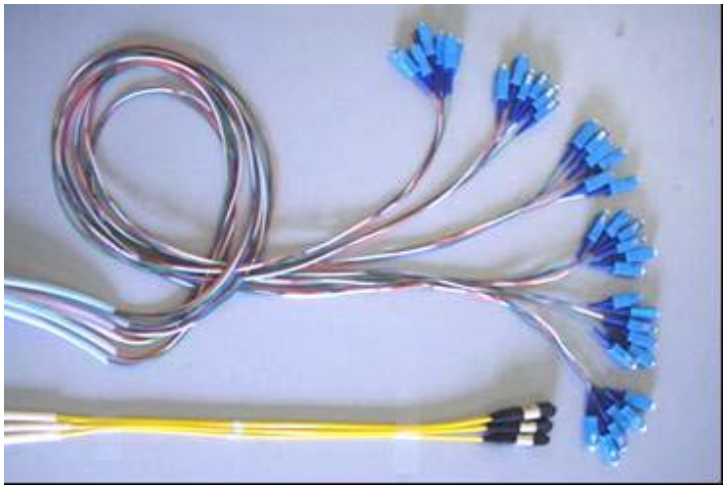
Data Center = Structured Cabling Applications

Source: Corning

SMF MPO Connectivity

- Mature Method
  - =>15 years Utilization
- Data Center Applications
  - Carrier Access
  - Ethernet
    - Extended Reach
    - Perceived Future Proof
  - Fibre Channel (SAN)
    - FICON
  - InfiniBand (HPC)

    Active Optical Cable



•SMF MPO Trunk x SMF LC Trunk

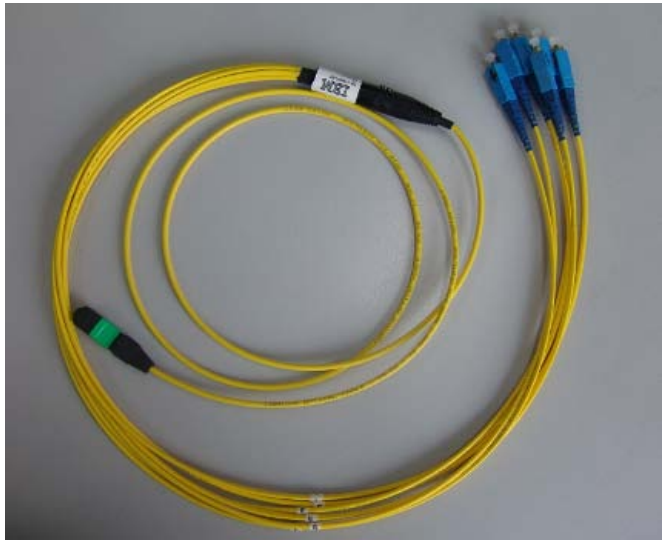
•Source: Luxtera

Source: Corning

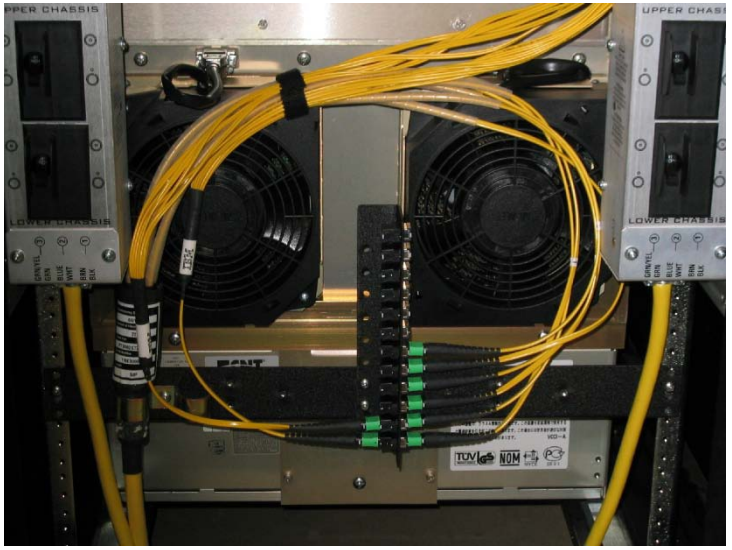
SMF MPO Connectivity

- High Density Connectivity
  - Ease Cable management
  - Scalability
  - Electronic Interface
  - Cooling Efficiency

SMF MPO Trunk Cable



SMF MPO Jumper Cable



SMF MPO Trunk Cable

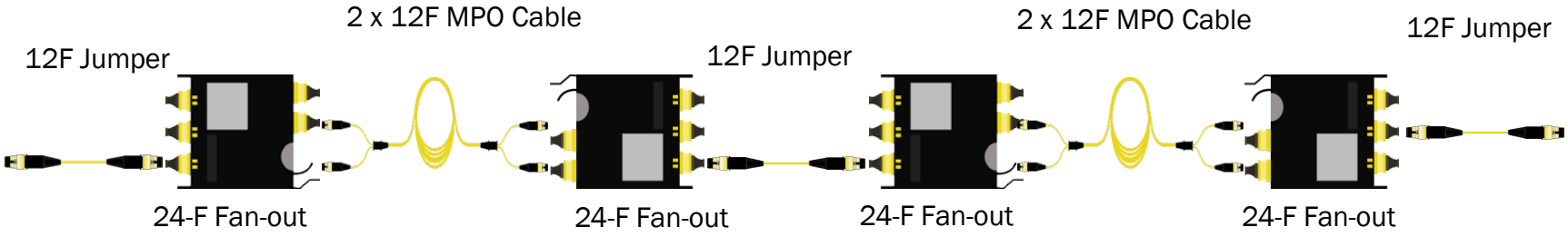
Source: IBM

Source: Corning

SMF MPO Connectivity: 100G Typical Installations

2 Channel Model for 100GBASE-NR4 with WDM

8 Channel Model for 100GBASE-NR4 no WDM



Source: Corning

Summary & 5 Criteria

Three transceiver module vendors, Avago Technologies, Opnext and Luxtera, with experience in 100G LR4 and SR10 transceivers have provided estimates used herein and have concluded, 'A 100 Gb/s (4 lane) PMD capable of operation over 2 km of multi-lane SM fiber (G.652) with BER < 10^{-12} based on 1310 nm lasers, NRZ modulation and 64b/66b encoding offers significant module and total system cost advantages relative to 100GBASE-LR4 for at least a 1 km reach', and believe that the 5 Criteria are satisfied. While technically capable of supporting a 2 km reach, a 1 km reach objective is proposed for economic considerations.

Broad Market Appeal: Several contributions speak to the interest in a mid range (MR) reach including ghiasi_01_911_NG100GOPTX, kolesar_01_0911_NG100GOPTX, & kipp_01_0112_NG100GOPTX and this presentation addresses concerns regarding acceptability of 4-lane parallel SMF in data centers

Compatibility: Changes are confined to the MR PMD and no compatibility issues exist.

Distinct Identity: The MR PMD will result in a new 802.3 clause making it straightforward for the reader to identify the relevant information. The PMD will be uniquely identified by fiber, wavelength, signal rate and reach.

Technical Feasibility: Contributions, anderson_01_1111_NG100GOPTX, anderson_01_0112_NG100GOPTX, palkert_01_1111_NG100GOPTX, petrilla_01_1111_NG100GOPTX & petrilla_02a_0112a_NG100GOPTX as well as the existence of 100GBASE-LR4 speak to the technical feasibility.

Economic Feasibility: This contribution, anderson_01_1111_NG100GOPTX, anderson_01_0112_NG100GOPTX, kipp_01_0112_NG100GOPTX, nowell_01_1111_NG100GOPTX and petrilla_02a_0112a_NG100GOPTX speak to economic feasibility.