



Next-gen 400 and 200 Gb/s PHYs  
over Fewer MMF Pairs  
Call For Interest Consensus  
Presentation

May 12, 2017

Draft 1.1 for NEA Ad Hoc Review

# Agenda

- **Overview Discussion**
  - Presenter 1
- **Presentations**
  - Market Drivers
    - Presenter 2
  - Technical Feasibility
    - Presenter 3
  - Why Now?
    - Presenter 4
- **Straw Polls**

# Introductions for today's presentation

- Presenter and Expert Panel:

# CFI objectives

- To gauge the interest in next-gen 400 and 200Gb/s PHYs over fewer MMF pairs .
- We do not need to:
  - Fully explore the problem
  - Debate strengths and weaknesses of solutions
  - Choose a solution
  - Create a PAR or 5 Criteria
  - Create a standard
- Anyone in the room may vote or speak

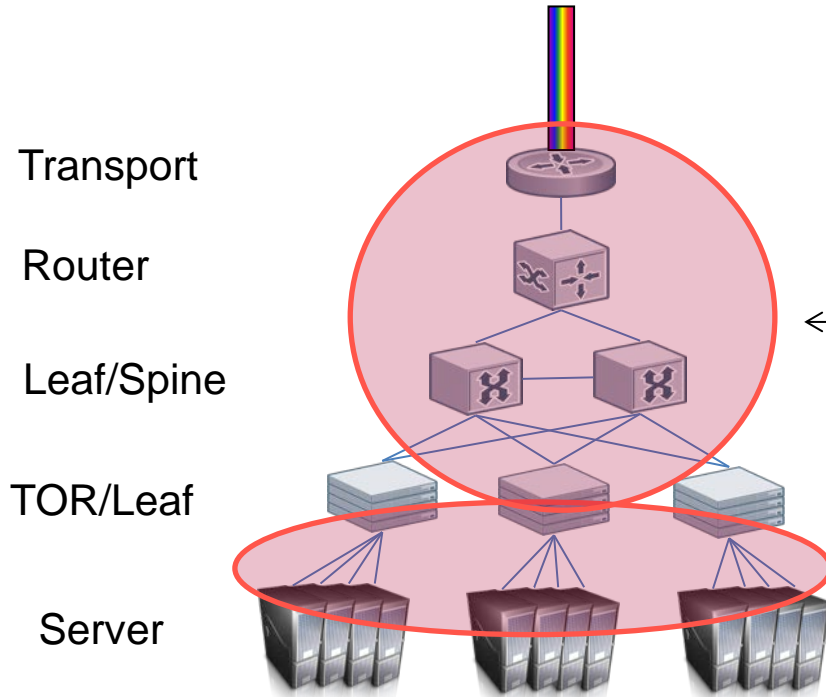
# Overview: motivation

Leverage technologies currently under development to create cost-optimized lower fiber count solutions over installed base, as well as greenfield MMF cabling, for 200 and 400 Gb/s

Global web-scale data centers and cloud based services – as well as the largest enterprise datacenters - are presented as leading applications.

Synergy with broader enterprise networking extends the application space and potential market adoption.

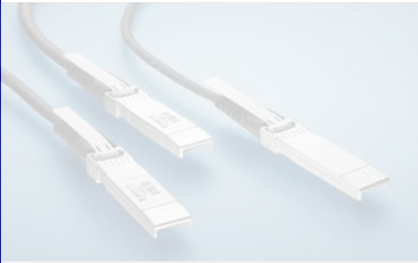
# What are we talking about?



Leading application space for next generation MMF PMDs

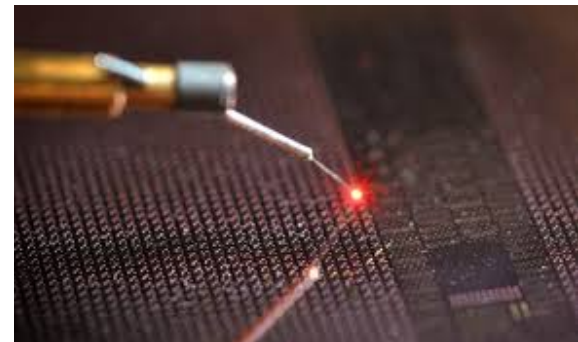
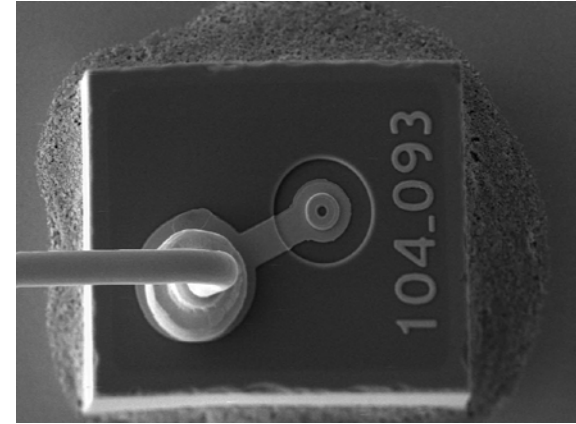
- Switch-to-switch & switch-to-router or router-to-transport connectivity
- Breakout of 400G to 100G may be used for high density 100G or breakout to 100G servers

# Market Drivers



# Historically VCSEL-MMF links have been seen by many as the lowest cost short-reach interconnect

- Relaxed alignment tolerances
  - Several microns vs. sub-micron
  - Allows passive alignment in module
  - Better cost/loss trade-off for connectors
- Connectors more resilient to dirt
  - Cleaning SMF connectors is common issue
- Lower drive currents
  - 5-10mA vs. 50-60mA
- On-wafer testing

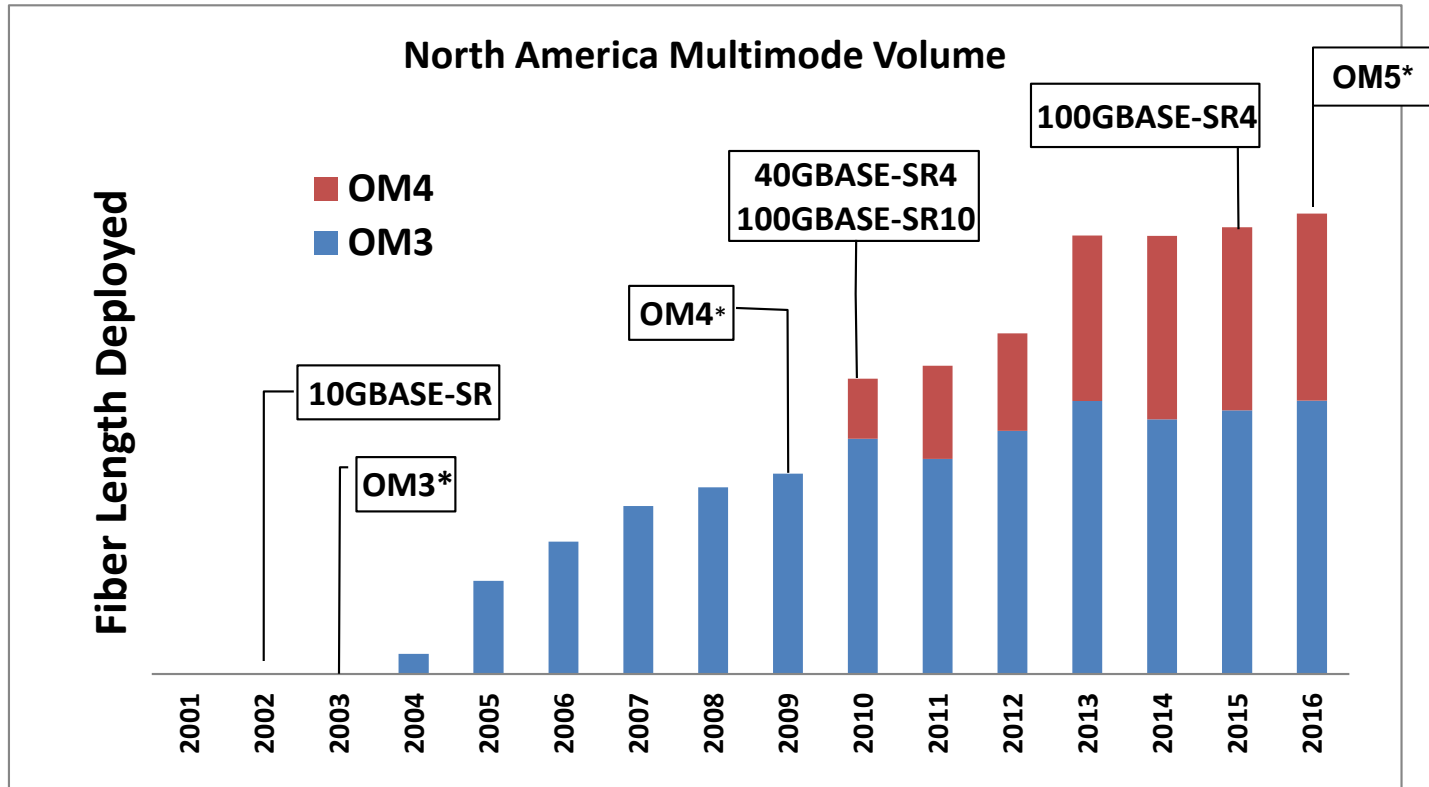




# 10/40/100G have been widely deployed over MMF

- Large installed base of duplex OM3/OM4 MMF deployed for 10GBASE-SR
- Large installed base of parallel OM3/OM4 MMF deployed for 40GBASE-SR4 and 100GBASE-SR4
- Industry investment in MMF cabling continues, including wideband OM5, now standardized

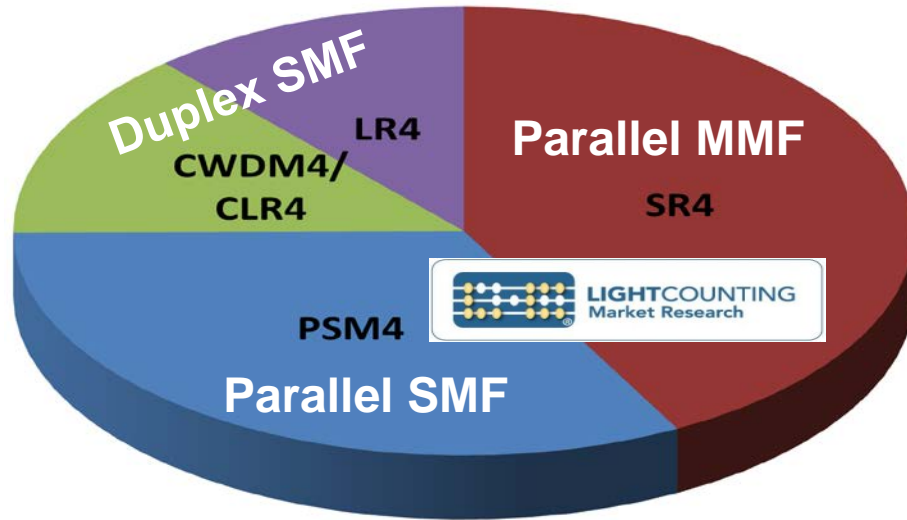
# Continued deployment & growth of OM3/OM4 MMF. Transition from OM3 to OM4. Ready for OM5.



\* Dates are ANSI/TIA standardization dates, not ISO/IEC

Used with permission: Matthew Burroughs North America Multimode Reports

# New 100G switching speed was applied to MMF in 1st year



- Chart shows units shipped
- Taken together, modules for SMF cabling have majority share
- But short-reach SR4 modules had the greatest individual contribution to 2016 shipments of QSFP28 modules

## 100GbE QSFP28 Consumption in 2016

Chart courtesy of Dale Murray, LightCounting

## Comments on a switch vendor's experience of demand for 100Gb/s MMF optics, from David Piehler (Dell EMC)

- Sold 100GBASE-SR4 into large enterprise DC space in 2016
- Could have sold 100G duplex MMF transceivers in 2016 had they been commercially available
- There is demand for MMF solutions with the highest speeds, and the lowest fiber counts
- He expects this trend to play out again for >100 Gb/s speeds

# How have 40 and 100Gb/s optics been used with MMF?

- 40G SR4
  - ~ 50% is breakout, to servers as well as a means to creating larger 10G switch fabrics
  - ~ 50% is switch-to-switch links
- 40G BiDi and SWDM4
  - Proprietary single fiber-pair solutions used in switch-to-switch connections
- 100G SR4
  - The 100G SR4 modules deployed in 2016 represent switch-to-switch and switch-to-router connections in Cloud and Largest enterprise DCs
  - Less likely to be used for breakout in cloud or largest enterprise DCs (the early adopters), where breakout to servers often done with DAC cables from TOR, and 25G is not expected to be a popular switch fabric speed there
  - 100G Breakout may be popular in smaller enterprise DCs and campus networks (later adopters), where “25G may be the new 10G”
- 100G duplex over MMF
  - Proprietary solutions; not yet in market; two sources expected in 2017; could have been sold in 2016 if available

# Market applications of 400G short reach

- Earliest use for low-cost router-transport and laboratory development applications in telecom and the cloud
- Initial volume applications in switch-router & switch-switch connections
  - in the cloud
  - largest enterprise DCs
- Breakout of 400G to 100G may become popular, similar to the case for breaking out 40G to 10G

# Comments on prospects for 400GBASE-SR4, by Chongjin Xie, Sr. Director of Infrastructure Service at Alibaba

- Alibaba uses 100GBASE-SR4 heavily for 100m switch-switch connections now
- 100GBASE-SR4 links over MMF cabling are lower cost for Alibaba today than PSM4 or CWDM4 links over SMF cabling
- Alibaba expects to deploy 100G switching for approximately three years, perhaps moving to 400G in 2019
- He hopes to have 400GBASE-SR4.n available in 2019 and strongly supports its standardization in IEEE

# Existing 400GBASE-SR16 does not fulfill the needs of the datacenter market

- 400GBASE-SR16 was envisioned as a lower-cost, fast time-to-market solution for router-transport & development needs
- 400GBASE-SR16 may not be a high-volume datacenter module
  - CFP8 will not be a common front panel port in datacenter switches
  - 32-fiber link with atypical connector will offset the low-cost nature of the transceiver.
  - Restricted to 16x25G electrical interface (400GAUI-16)
    - No path to 400GAUI-8 without reverse gearbox
- A lower fiber-count MMF solution is expected to have lowest cost for short-reach 400G



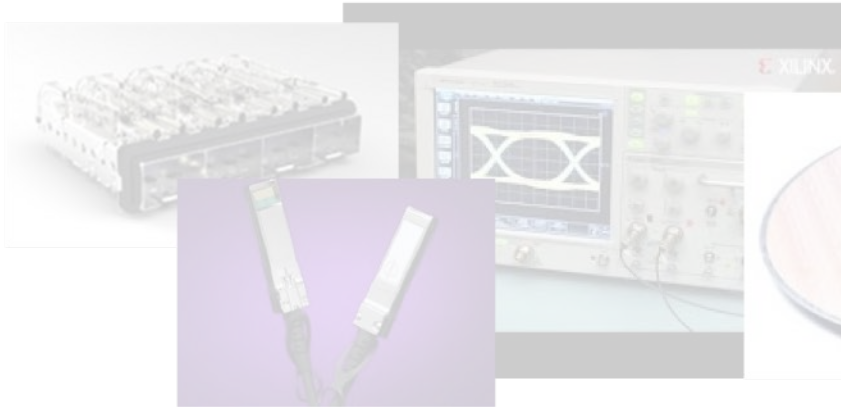
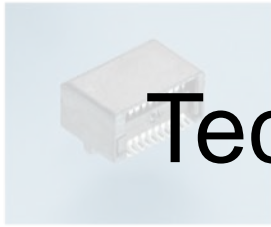
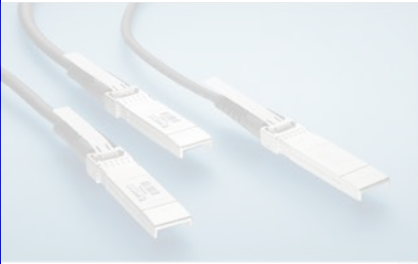
# Benefits of 400GBASE-SR4 over 400GBASE-SR16 for the datacenter market

- Operate on same cabling as previous SR4 modules
  - No special connector
- Suitable for all 400G form factors
  - CFP8, QSFP-DD, OSFP
- No reverse gearbox with 400GAUI-8 interface

# Market need for 200G module for duplex MMF

- 200G switching is expected to find acceptance in parts of the cloud and enterprise DC networking space on same time frame as 400G
- 200GBASE-SR4 is already being standardized in 802.3cd to support parallel MMF cabling
- There is no reason to believe that the early demand for 100G duplex MMF optics will not be replicated for 200G optics

# Technology Feasibility



# Technologies for next-gen MMF PMDs

- PMDs for 400G over 4 MMF pairs and 200G over < 4 MMF pairs will require several technologies currently in advanced stages of development
  - VCSELs supporting 50Gb/s PAM4 signaling
  - Multiple wavelengths over MMF
- OM5 provides longer reach when using multiple wavelengths over MMF, but is not required

# Technical options for next-gen MMF PMDs

| Technology (per fiber) | 1 fiber pair | 2 fiber pairs | 4 fiber pairs | 8 fiber pairs  | 16 fiber pairs |
|------------------------|--------------|---------------|---------------|--|----------------|
| 25G- $\lambda$ NRZ     | 25G-SR       |               | 100G-SR4      |  | 400G-SR16      |
| 50G- $\lambda$ PAM4    | 50G-SR       | 100G-SR2      | 200G-SR4      | 400G-SR8   |                |
| 2x50G- $\lambda$ PAM4  | 100G-SR1.2   | 200G-SR2.2    | 400G-SR4.2    | Technology options for 200 & 400 Gb/s links over fewer MMF fiber pairs |                |
| 4x25G- $\lambda$ NRZ   | 100G-SR1.4   | 200G-SR2.4    | 400G-SR4.4    |  |                |
| 4x50G- $\lambda$ PAM4  | 200G-SR1.4   | 400G-SR2.4    | 800G-SR4.4    |  |                |



Existing IEEE standard

In progress in 802.3bs, cd

Multi-Wavelength Nomenclature

SRm.n

m = # fiber pairs

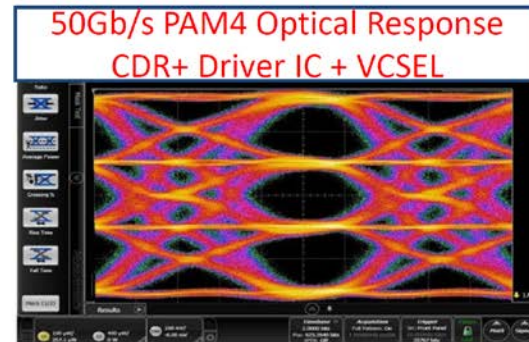
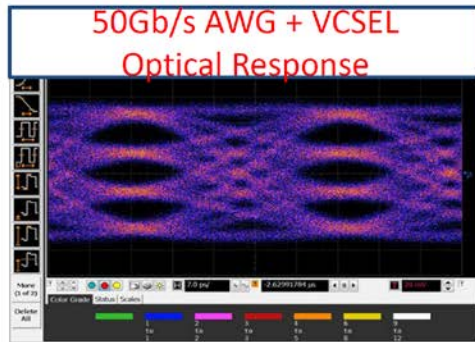
n = # wavelengths

# 50Gb/s PAM4 over MMF in IEEE 802.3cd

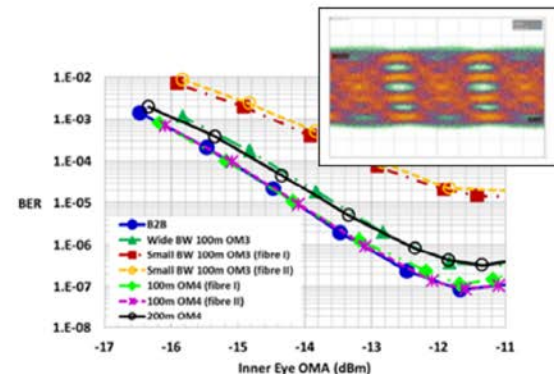
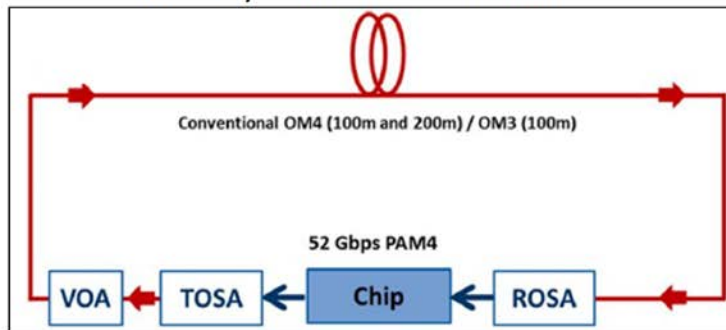
- 802.3cd has an objective to “Define single-lane 50 Gb/s PHY for operation over MMF with lengths up to at least 100m”
- 26.5625 GBd signaling with PAM4 modulation was selected to implement 50 Gb/s
- Could re-use RS(544,514,10) FEC from clauses 134 (50G), 91 (100G), and 119 (200G and 400G) if appropriate in this project

# Finisar demonstration of 50G PAM4 over MMF from king\_GE\_NGOATH\_01\_0116

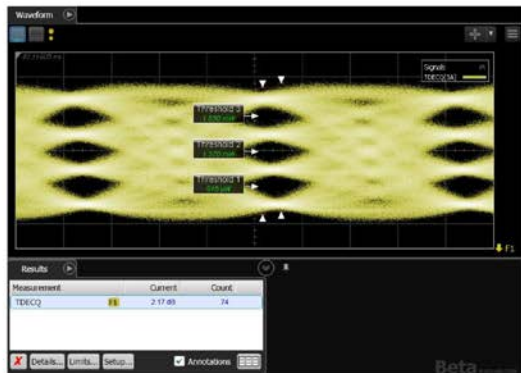
Bench top PAM4 experiments using 25Gb/s VCSELs



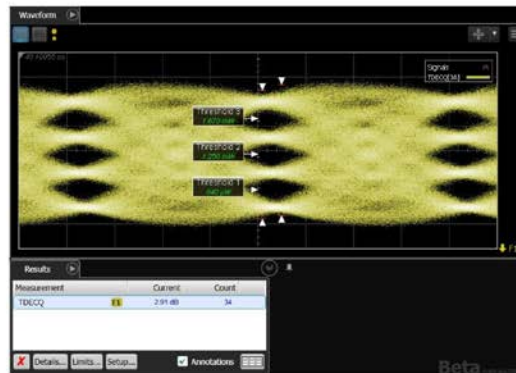
and early PAM4 PHY evaluation...



- Transmission of 26.5625 GBd PAM4 over 70 m worst-case OM3 MMF
- TDECQ measured for 855 nm and 908 nm VCSEL-based transmitters



855 nm Tx  
TDECQ: 2.2 dB



908 nm Tx  
TDECQ: 2.9 dB

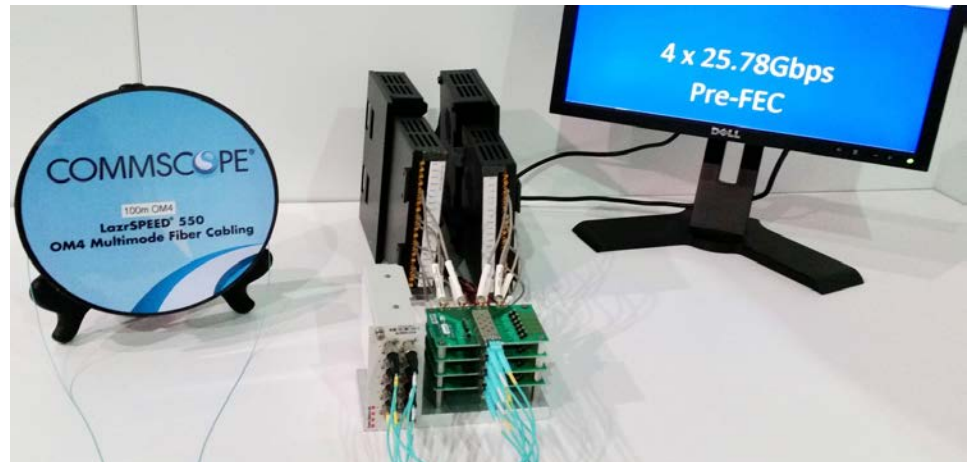
- For both wavelengths, transmission over worst-case MMF results in TDECQ values within the 4 dB requirement for MMF PMDs in P802.3cd



# Technical feasibility of four wavelengths over MMF

Finisar announced a VCSEL-based 100 Gb/s SWDM4 product, with technical demonstration at OFC 2015.

- Error free operation over 150 m on OM4, 275 m on a sample OM5
- Same 30 nm channel spacing as 40 Gb/s, centered at 850, 880, 910, 940 nm
- Balances cost and performance of mux/demux optics, VCSEL wavelength pass-bands, and fiber wavelength range over which modal bandwidth is critical

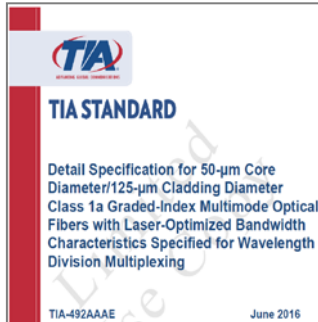


## Adding wavelengths & PAM4 to MMF modules will not negate historical relative cost advantage vs. SMF modules

- Tolerances for mux/demux are significantly more relaxed in the case of MMF than SMF
- More costly circuits are needed to implement PAM4 for both fiber types
- Reduction of laser RIN for PAM4 is not more difficult for VCSELs than for DFBs
- Packaging for VCSEL sources at 50Gb/s PAM4 is based on known technology, whereas packaging for 1310nm sources at 100 Gb/s per lane PAM4 has required significant development

# Standardized Wideband MMF/OM5 improves performance with multiple wavelengths

- OM5 MMF extends the 850nm performance of OM4 out to 953nm
- Drop-in replacement for OM4 at 850nm. Fully backward-compatible with previous IEEE standards
- Accommodates at least four wavelengths on economical grid spacing
- Standards:
  - Fiber: TIA-492AAAE (2016), IEC 60793-2-10 ed. 6 (1Q17)
  - Cabling: ANSI/TIA-568.3-D (2016), ISO/IEC 11801 ed. 3 (target 4Q17)
  - Application (WBMMF/OM5 operating at 850nm only): IEEE 802.3bs draft (2016), IEEE 802.3cd draft (2016), Fibre Channel FC-PI-7/64GFC & 256GFC (target 2Q17)



Fiber Standards



Structured Cabling Standards



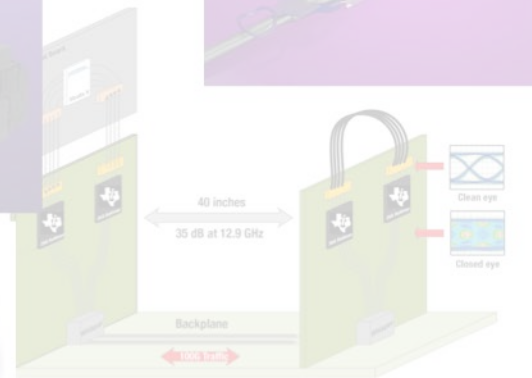
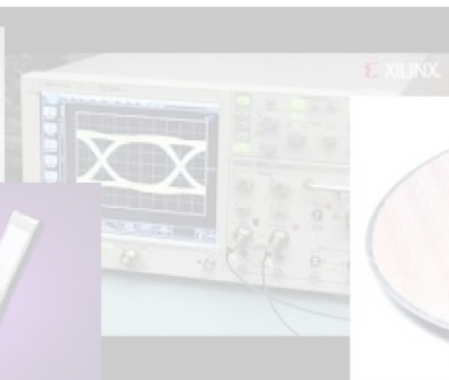
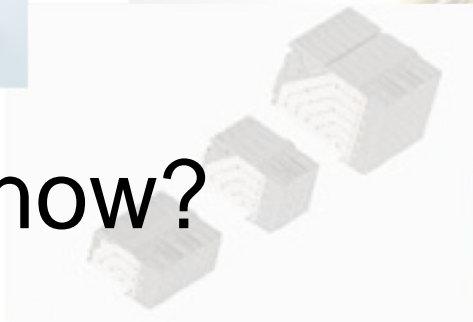
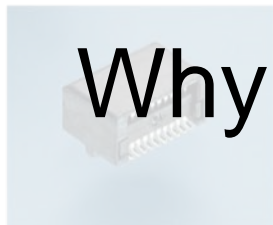
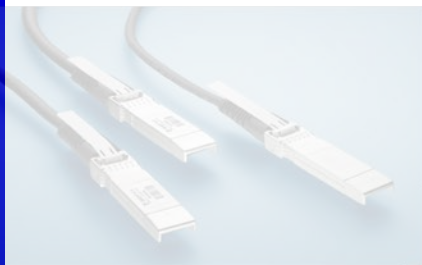
# Potential merits of OM5 cabling

- Supports all legacy multimode applications to at least the same reach as OM4
  - Compliant to OM4 and OM3 specifications
  - Interoperable with legacy media
- Optimizes reach with multi-wavelength transceivers
  - Reaches of 70/100/150m proposed in ingham\_3cd\_01a\_0916 for two wavelengths of 50 Gb/s PAM4 at 857 & 908nm operating over OM3/OM4/OM5 MMF, respectively.
  - Reaches of 75/100/150m proposed in forthcoming 100G SWDM MSA for 4 wavelengths of 25G NRZ at 850, 880, 910, 940 nm operating over OM3/OM4/OM5 MMF, respectively.

# The VCSEL optical lane rate roadmap will support PMDs beyond those contemplated here

| Bit Rate per Optical Lane (Gb/s) | Source             | Year | Speed for SR1.4 module | Speed for SR4 module | Speed for SR4.2 module |
|----------------------------------|--------------------|------|------------------------|----------------------|------------------------|
| 25                               | 802.3bm            | 2015 | 100G                   | 100G                 |                        |
| 50                               | 802.3cd            | 2018 | 200G                   | 200G                 | 400G                   |
| 100                              | Consensus estimate | 2023 | 400G                   | 400G                 | 800G                   |

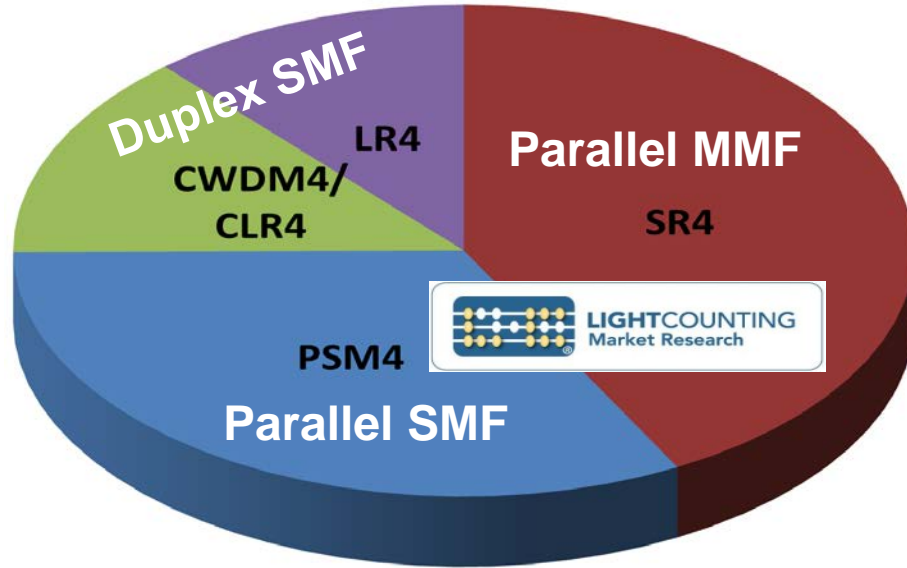
# Why now?



# Rationale for beginning now

- Recent history shows that higher speeds over MMF are needed in the first year that new switch speeds are commercially available
- The existing 400GBASE-SR16 solution will not meet that need
- There is no 200Gb/s duplex MMF PMD in existing IEEE standards
- These PMDs are needed in the market commercially in 2019

# 100GbE QSFP28 Consumption in 2016

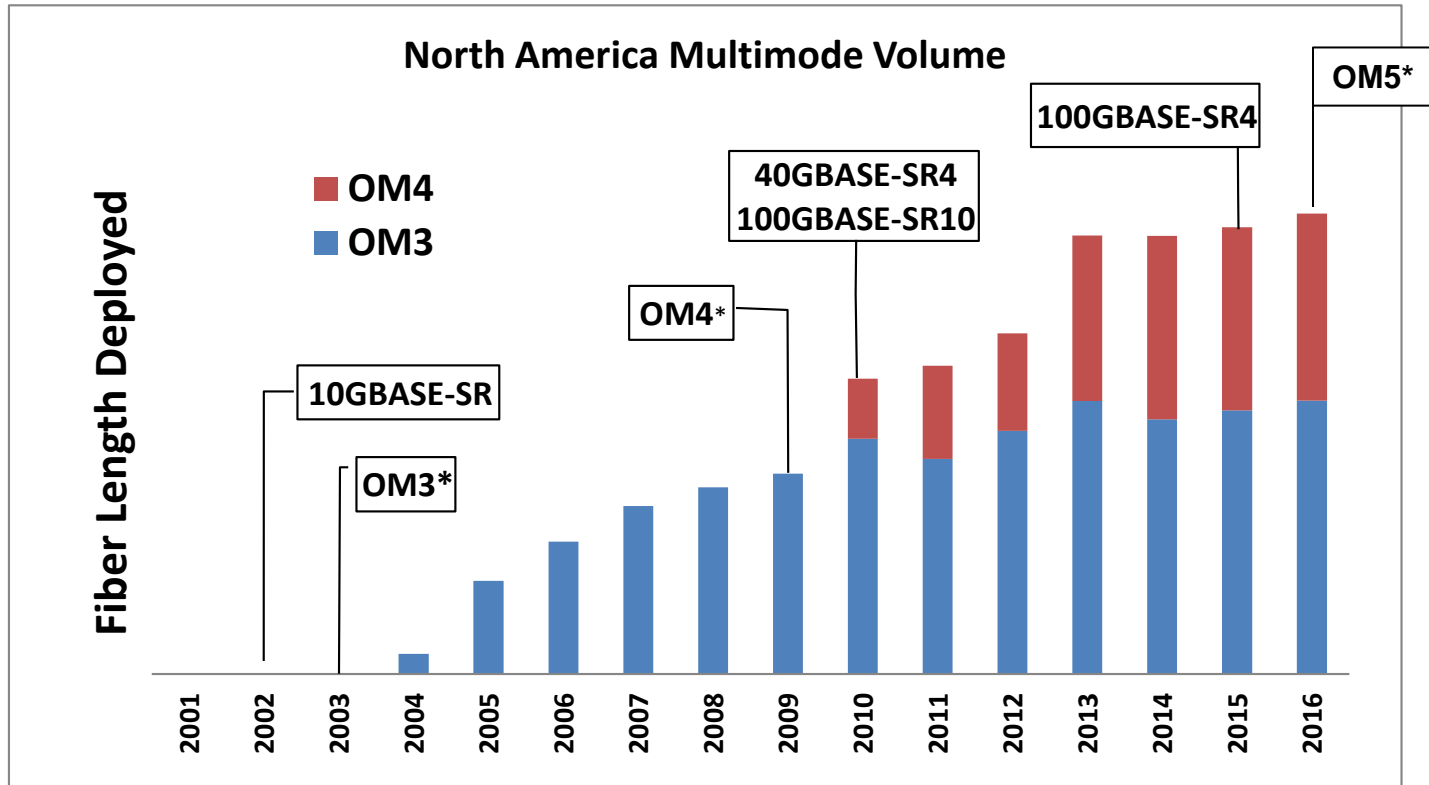


- 100G datacenter switches ramped up in 2016
- 100GBASE-SR4 modules sold in large quantities year one
- Duplex MMF modules would have sold in year one if generally available

Chart courtesy of Dale Murray, LightCounting



# Continued deployment & growth of OM3/OM4 MMF. Transition from OM3 to OM4. Ready for OM5.



\* Dates are ANSI/TIA standardization dates, not ISO/IEC

Used with permission: Matthew Burroughs North America Multimode Reports

# Why Now?

The market for 100GBASE-SR4 over parallel MMF cabling was robust in 2016 as soon as significant deployment of 100 Gb/s switching began in the datacenter

- Cloud DCs in North America and China
- Largest enterprise DCs
- 100Gb/s duplex transceivers for MMF would have been deployed in 2016 had they been available
- Early adopters will deploy next-gen 200/400 Gb/s MMF PMDs if they are available

50 Gb/s ecosystem supporting 200/400G switch ASICs progresses towards ~2019 deployment

The enabling technologies exist to support next-gen MMF PMDs over fewer fiber pairs

- 50Gb/s PAM4 in development for 802.3cd
- Two and four wavelengths already used in proprietary duplex MMF transceivers
- Operation over installed base as well as new OM5 MMF cabling is supported

Data shows that the market continues to deploy MMF cabling

- New performance grades are accepted when they provide benefit
- Standardizing lower cost applications for MMF facilitates upgrades and improves Ethernet market

# Contributors

Dale Murray, LightCounting

Chongjin Xie, Alibaba

David Piehler, Dell EMC

Jonathan Ingham, FIT

Jonathan King, Finisar

Frank Chang, Inphi

Brett Lane, Panduit

Steve Swanson, Corning

John Kamino, OFS

Mabud Choudhury, OFS

Paul Kolesar, CommScope

Adrian Young, Leviton

# Supporters (p Individuals from q companies)

Scott Kipp, Brocade  
Jeffery Maki, Juniper  
David Piehler, Dell EMC

Rob Stone, Broadcom  
John Johnson, Broadcom  
Frank Chang, Inphi  
Mike Dudek, Cavium

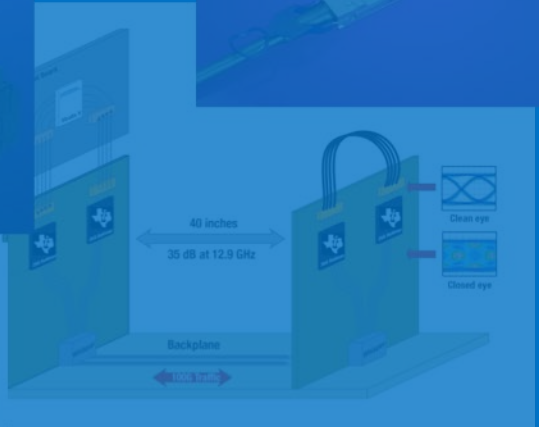
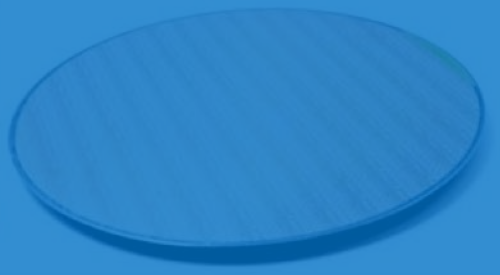
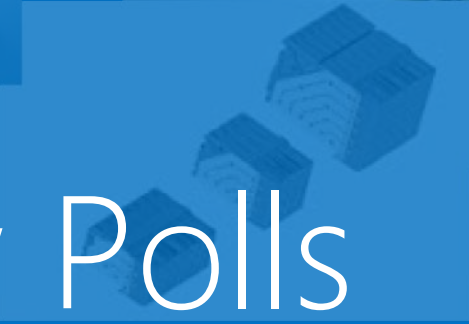
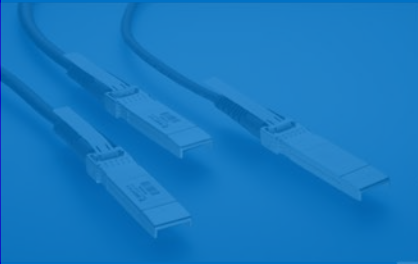
Jonathan Ingham, FIT  
Jonathan King, Finisar  
Vipal Bhatt, Finisar  
David Lewis, Lumentum

Dale Murray, LightCounting  
Chongjin Xie, Alibaba

Adrian Amezcua, Prysmian  
Alexander Umnov, Corning  
Steve Swanson, Corning  
Paul Vanderlaan, Nexans  
Rakesh Sambaraju, Nexans  
John Kamino, OFS  
Mabud Choudhury, OFS  
Paul Kolesar, CommScope  
Adrian Young, Leviton  
Qing Xu, Belden  
Phong Pham, USConec

# Supporters (2)

# Straw Polls



# Call-for-Interest Consensus

- Should a study group be formed for “Next-gen 400 and 200 Gb/s PHYs over Fewer MMF Pairs”?
- Y:            N:            A:
- Room count:

# Participation

- I would participate in a “Next-gen 400 and 200 Gb/s PHYs over Fewer MMF Pairs” study group in IEEE 802.3
  - Tally:
- My company would support participation in a “Next-gen 400 and 200 Gb/s PHYs over Fewer MMF Pairs” study group
  - Tally:



# Future Work

- Ask 802.3 at Thursday's closing meeting to form study group
- If approved:
  - Request 802 EC to approve creation of the study group on Friday
  - First study group meeting would be during Xyz 2017 IEEE 802.3 interim meeting