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

July 11, 2017

Draft 1.3 for NEA Ad Hoc Review

Berlin Plenary

# Agenda

- **Overview Discussion**
  - Presenter 1
- **Presentations**
  - Market Drivers
    - Dale Murray, LightCounting (tentative)
  - Technical Feasibility
    - Presenter 3
  - Why Now?
    - Presenter 4
- **Straw Polls**

# Introductions for today’s presentation

- Presenter and Expert Panel:
  - David Piehler, Dell EMC (tentative)

# CFI objectives

- To gauge the interest in next-gen 200 and 400Gb/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

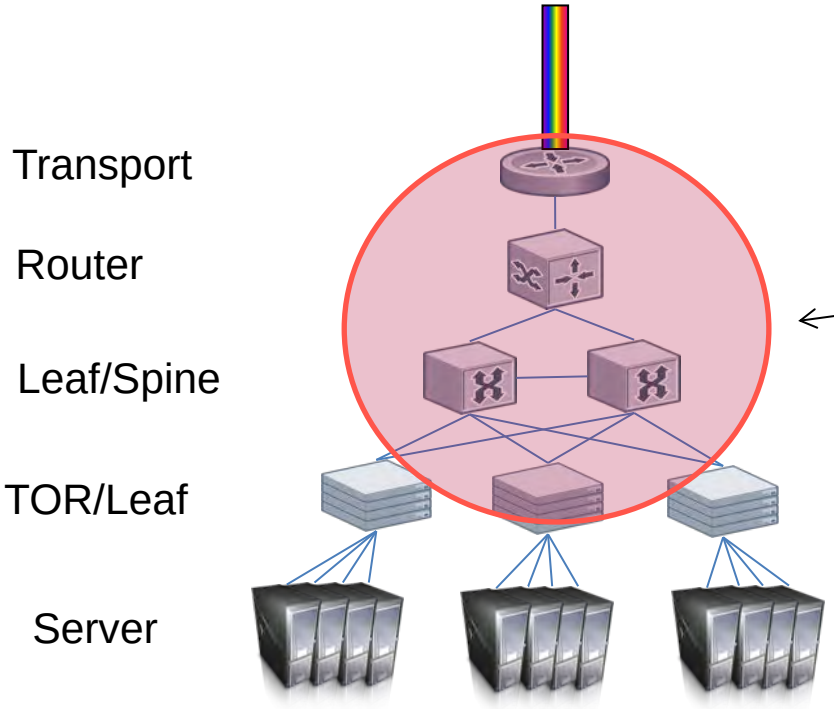
# Overview: motivation

Leverage technologies in advanced stages of development to create cost-optimized lower fiber-count solutions, over installed base and 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?

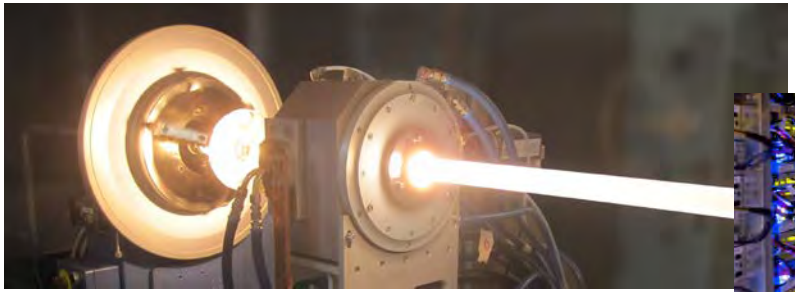


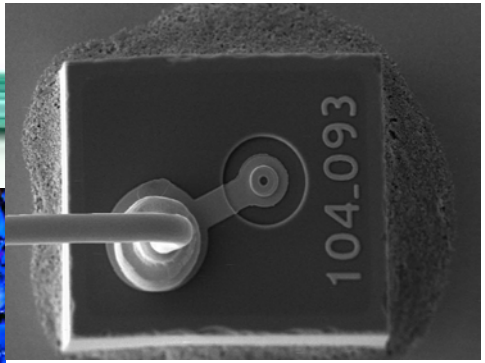

Applications for early adoption of next-generation MMF PMDs include connectivity in web2.0 and *largest* enterprise data centers for

- switch-to-switch
- switch-to-router
- router-to-transport

Other applications may arise later when the broad enterprise market needs higher speeds

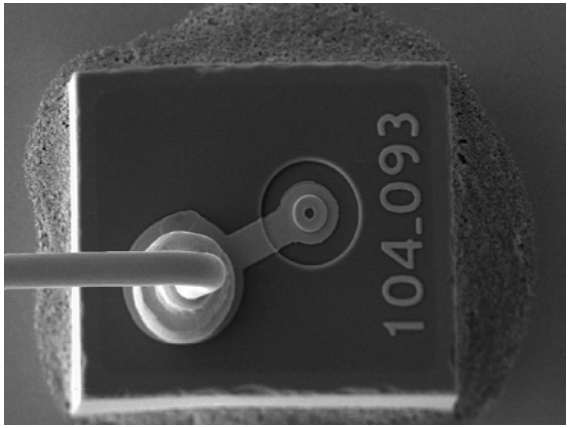
# Market Drivers

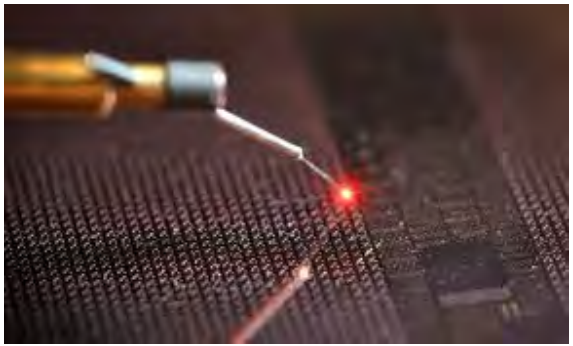




# Historically VCSEL-MMF links have been seen by many as the lowest cost and power short-reach interconnect

- 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

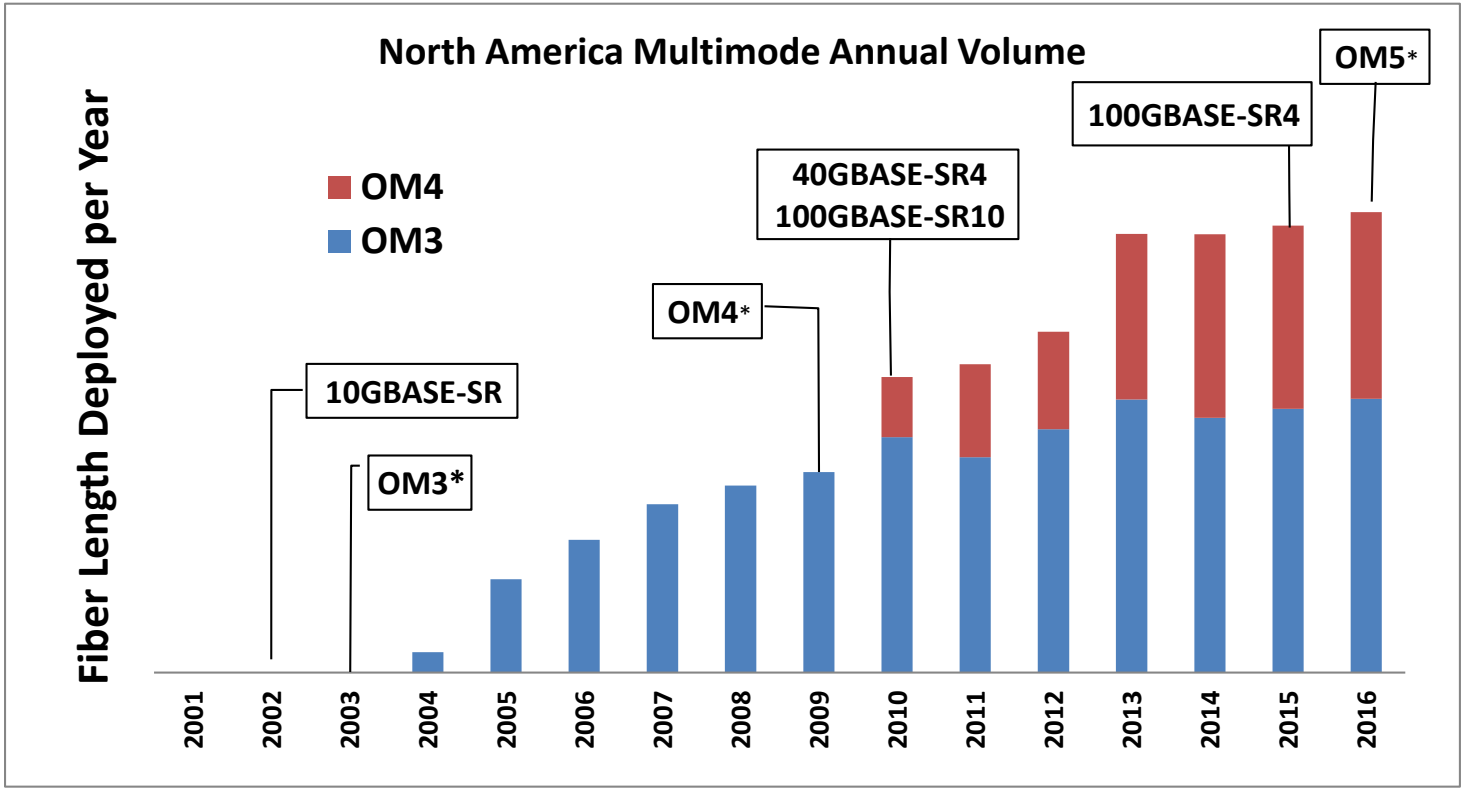


# 10/40/100G have been widely deployed over MMF

- 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

# Deployment of OM3 MMF ramped up after standards were complete, with OM4 ramping up next after standards issued

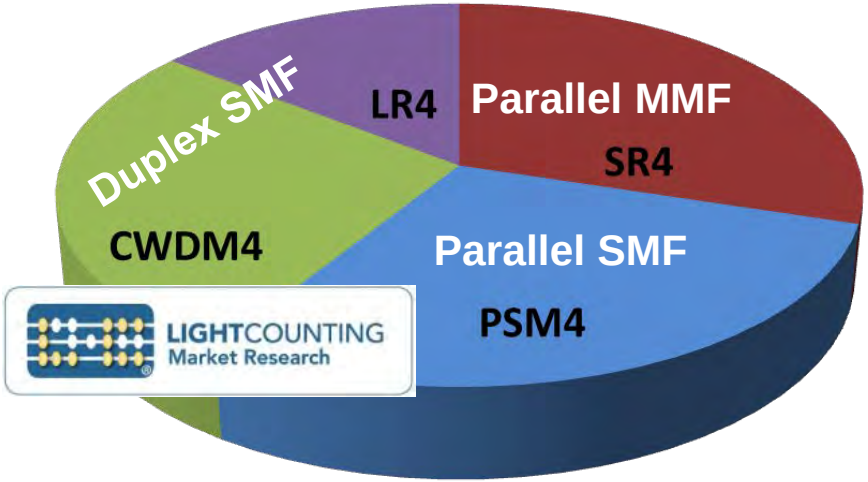



* Dates are ANSI/TIA standardization dates, not ISO/IEC

Used with permission: Matthew Burroughs North America Multimode Reports

# 100GBASE-SR4 in QSFP28 was required by web2.0 and largest enterprise data centers as soon as 100G switches entered the market

**Unit Shipments 2016-2017: QSFP28 Modules**




**100GbE QSFP28 Consumption**

Chart courtesy of Dale Murray, LightCounting

- Chart shows units shipped
- Taken together, modules for SMF cabling have majority share
- However short-reach SR4 modules are comparable in volume to each of the SMF module types

# 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

# 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 also 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

# 40 & 100Gb/s MMF optics have been used in volume for switch-to-switch connections in data centers

- 40G SR4 for 4 pairs MMF
  - ~ 50% used in switch-to-switch links
- 40G BiDi & SWDM4 for 1 pair MMF
  - Proprietary solutions used in switch-to-switch connections
- 100G SR4 for 4 pairs MMF
  - The 100GBASE-SR4 links deployed in 2016 comprise switch-to-switch and switch-to-router connections in cloud and largest enterprise DCs
- 100G BiDi & SWDM4 for 1 pair MMF
  - Proprietary solutions; expected in 2017; market demand existed in 2016

# Market applications of 400G short reach

- Smaller volume used 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

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

- 400GBASE-SR16 was originally 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

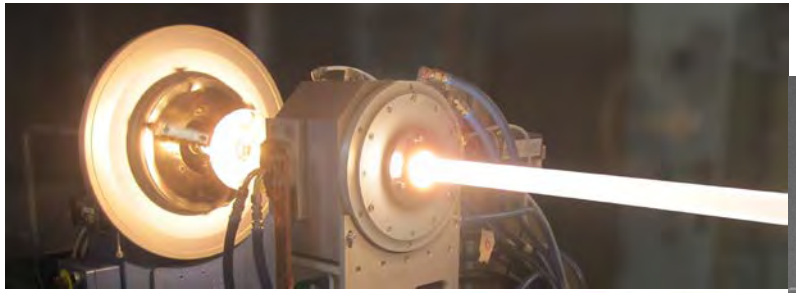
# Benefits of 400GBASE-SR4 (for example) over 400GBASE-SR16 for the datacenter market

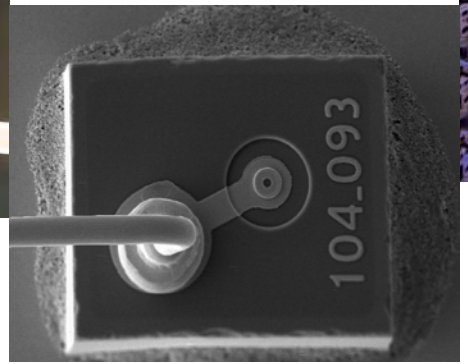
- Operates on same cabling as previous SR4 modules
  - No special connector
- Suitable for all 400G form factors
  - CFP8, QSFP-DD, OSFP
- No reverse gearbox with 400GAUI-8 interface

# 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
- The early demand for 100G duplex MMF optics suggests that large enterprise data centers will need 200G duplex optics as well
  - Believe there is value for the industry in a *standardized* solution for 200G over single pair of MMF

# Technology Feasibility






# Technologies for next-gen MMF PMDs

- PMDs for 400G over 4 MMF pairs and 200G over < 4 MMF pairs will require two technologies
  1. Multiple wavelengths on MMF – introduced in 2013
  2. VCSELs supporting 50Gb/s PAM4 signaling – sampling now

- OM5 provides longer reach when using multiple wavelengths over MMF, but is not required

# Technical options exist for 200/400G over fewer MMF pairs

| Technology (per fiber) | 1 fiber pair | 2 fiber pairs | 4 fiber pairs | 8 fiber pairs | 16 fiber pairs |
|---|---|---|---|---|---|
| 25G-λ NRZ | 25G-SR | | 100G-SR4 | | 400G-SR16 |
| 50G-λ PAM4 | 50G-SR | 100G-SR2 | 200G-SR4 | **400G-SR8** | |
| 2x50G-λ PAM4 | 100G-SR1.2 | **200G-SR2.2** | **400G-SR4.2** | | |
| 4x25G-λ NRZ | 100G-SR1.4 | **200G-SR2.4** | **400G-SR4.4** | | |
| 4x50G-λ PAM4 | **200G-SR1.4** | **400G-SR2.4** | 800G-SR4.4 | | |

**Technology options for 200 & 400 Gb/s links over fewer MMF fiber pairs**

Existing IEEE standard

In progress in 802.3bs, cd

Multi-Wavelength Nomenclature

SRm.n m = # fiber pairs

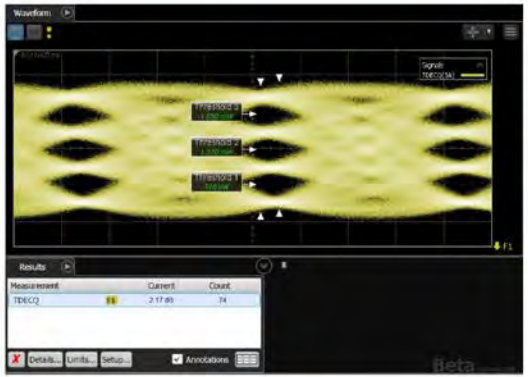
n = # wavelengths

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

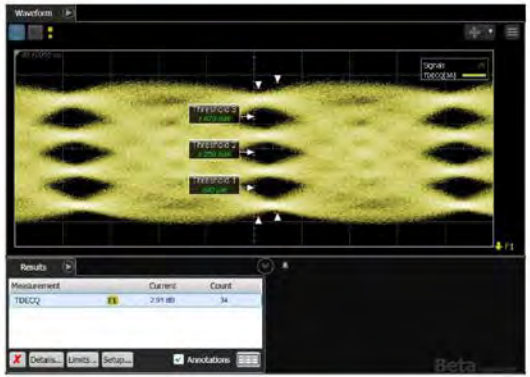
- 802.3cd has an objective to “Define single-lane 50 Gb/s PHY for operation over MMF with lengths up to at least 100m”
- 26.5625 GBd signaling with PAM4 modulation was selected to implement 50 Gb/s
- Could re-use RS(544,514,10) FEC from clauses 134 (50G), 91 (100G), and 119 (200G and 400G) if appropriate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

# Feasibility of two-wavelength PAM4 over MMF

- 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



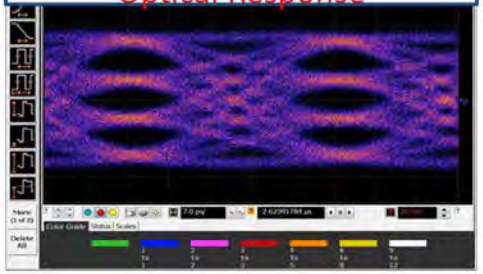

908 nm Tx
TDECQ: 2.9 dB

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

# Finisar demonstration of 50G PAM4 over MMF from king_GE_NGOATH_01_0116

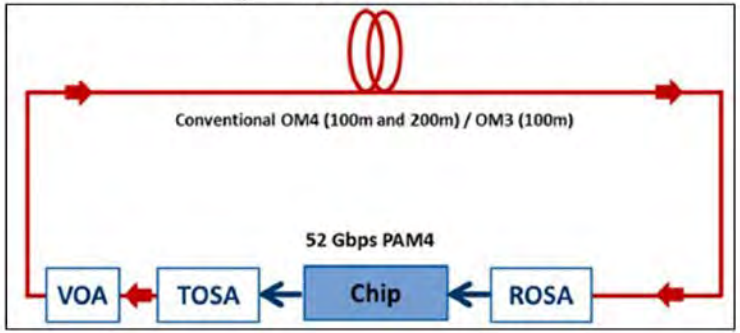
Bench top PAM4 experiments using 25Gb/s VCSELs

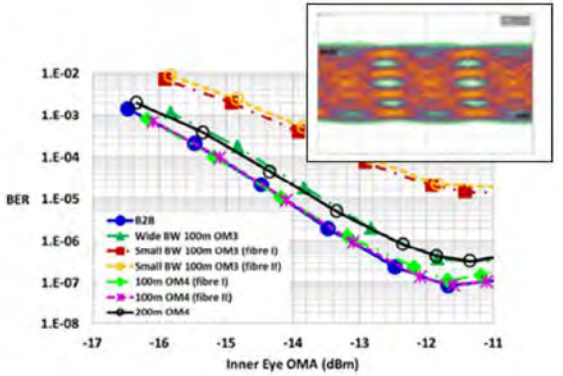
50Gb/s AWG + VCSEL
Optical Response



50Gb/s PAM4 Optical Response
CDR+ Driver IC + VCSEL

and early PAM4 PHY evaluation….

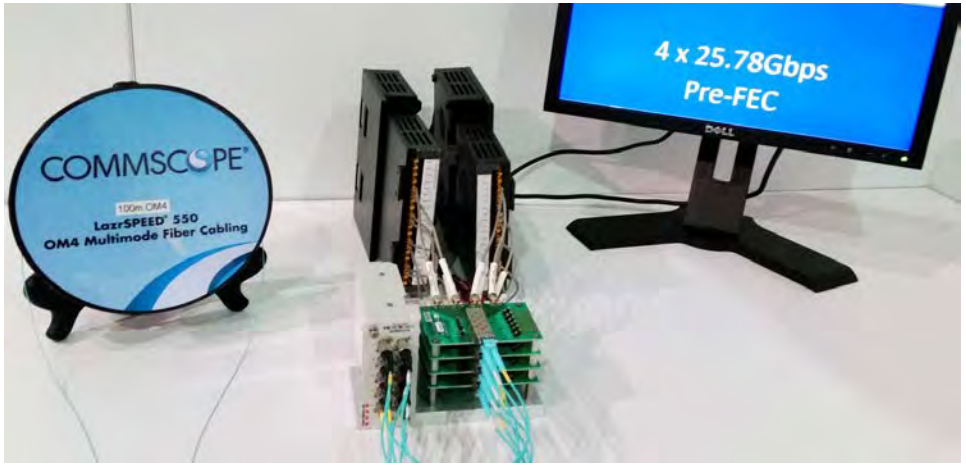


# Technical feasibility of four wavelengths over MMF

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 preserves the historical cost & power advantage over SMF modules

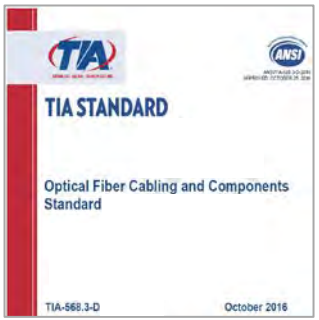
- Tolerances for mux/demux are significantly more relaxed in the case of MMF than SMF
- More costly circuits are needed to implement PAM4 for both fiber types
- Reduction of laser RIN for PAM4 is not more difficult for VCSELs than for DFBs
- Packaging for VCSEL sources at 50Gb/s PAM4 is based on known 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 (2Q17)



Fiber Standards





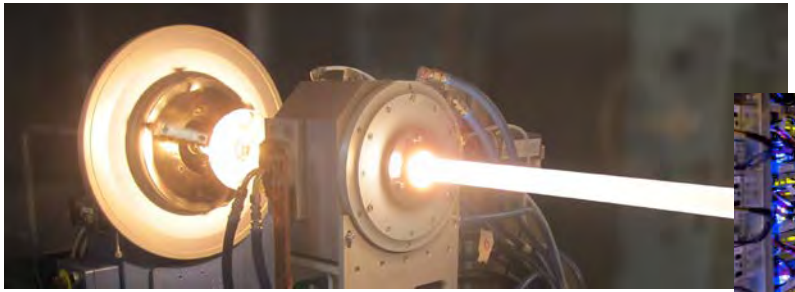
Structured Cabling Standards

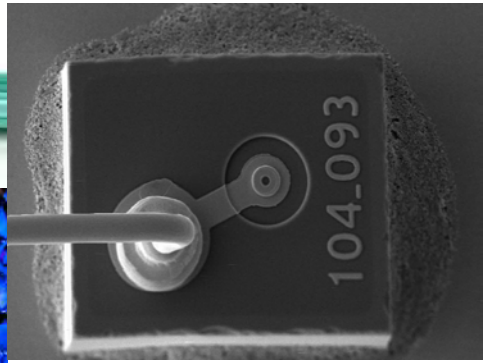
# The VCSEL optical lane rate roadmap will support PMDs beyond those contemplated here

| Bit Rate per Optical Lane (Gb/s) | Year | Speed for SR1.4 module | Speed for SR4 module | Speed for SR4.2 module |
|---|---|---|---|---|
| 25 | 2015 | 100G | 100G | |
| 50 | 2018 | 200G | 200G | 400G |
| 100 | proofs-of-concept exist | 400G | 400G | 800G |

* Several technical presentations at OFC 2017 showed research feasibility

# Why now?





# The use of VCSELs & MMF persists for shorter reach, even though SMF modules are now defined for 500m reach

- 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
  - We believe there is value for the industry in a *standardized* solution for 400GBASE-SR4.n
- There is no 200Gb/s duplex MMF PMD in existing IEEE standards
- These PMDs will be needed commercially in 2019

# Strong 100GBASE-SR4 consumption now suggests strong 400GBASE-SR4 uptake in ~2019

**Unit Shipments 2016-2017: QSFP28 Modules**

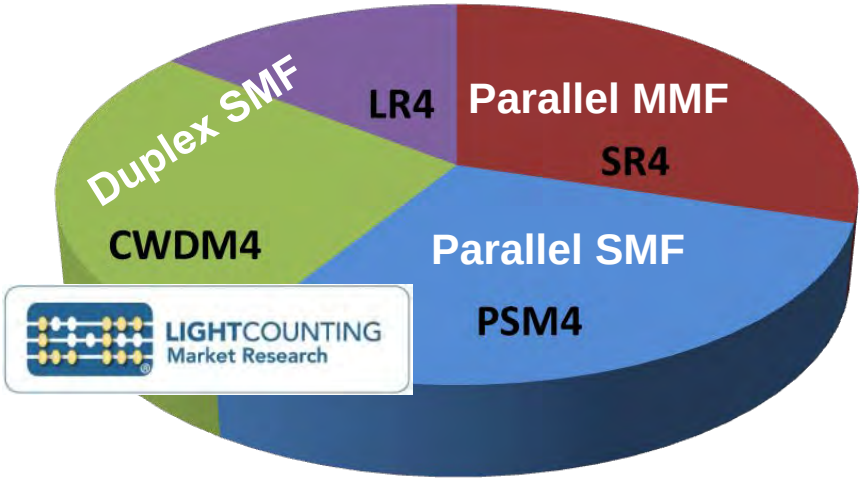
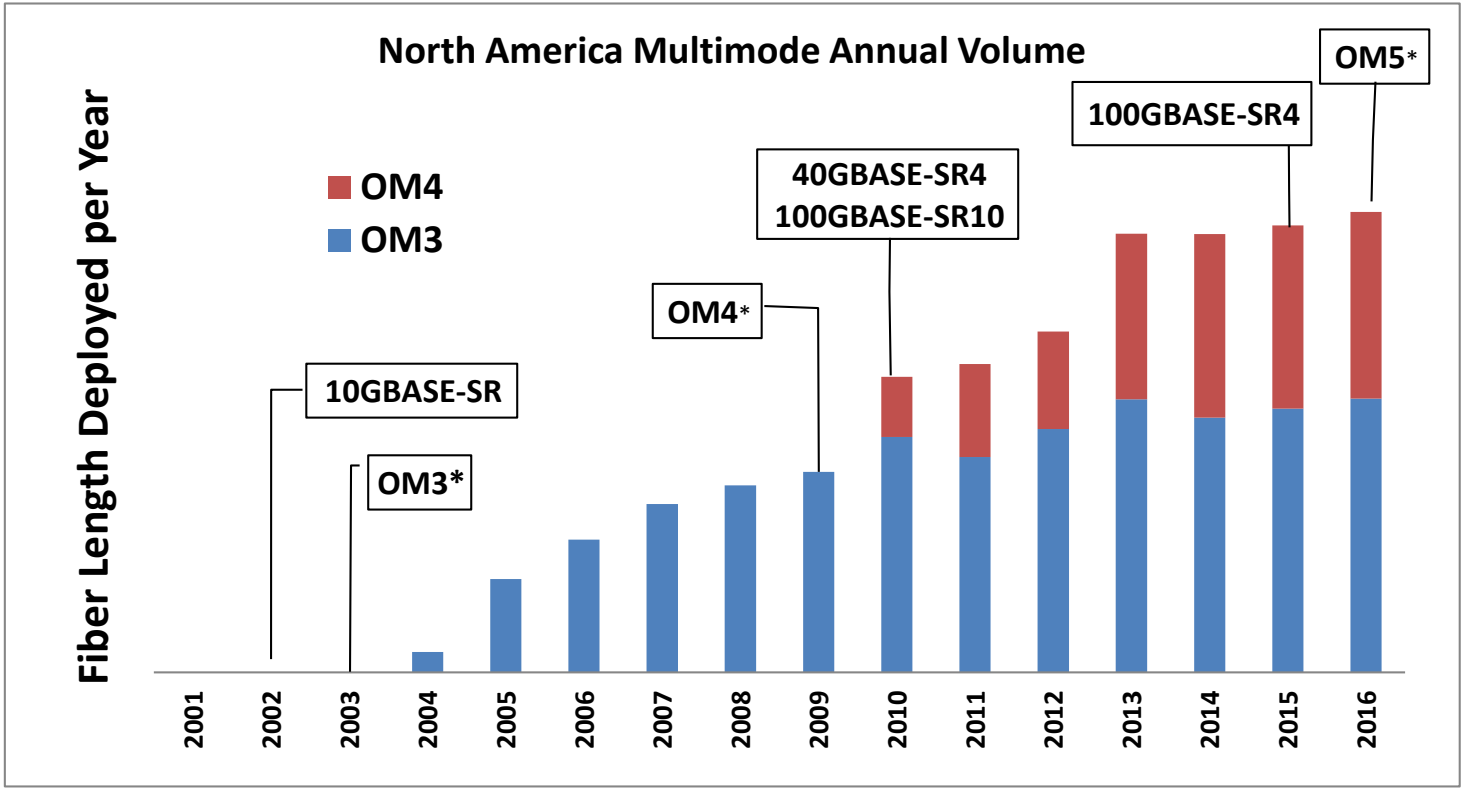



Chart courtesy of Dale Murray, LightCounting

- 100GBASE-SR4 has been used heavily by early adopters in 2016 & 2017
- The low relative cost and high technical feasibility of SR4 for short reach should be valuable to the same end-users at 400G as well

# MMF cable is being deployed at the same rate as in the past, and the Ethernet market will grow by standardizing lower cost PMDs for it




* 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; used in proprietary MMF modules sampling now
- 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

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

# Contributors

Dale Murray, LightCounting
Chongjin Xie, Alibaba
David Piehler, Dell EMC
Jonathan Ingham, FIT
Jonathan King, Finisar
Frank Chang, Inphi
Steve Swanson, Corning
John Kamino, OFS
Mabud Choudhury, OFS
Paul Kolesar, CommScope
Adrian Young, Leviton
Robert Lingle, Jr., OFS

# 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

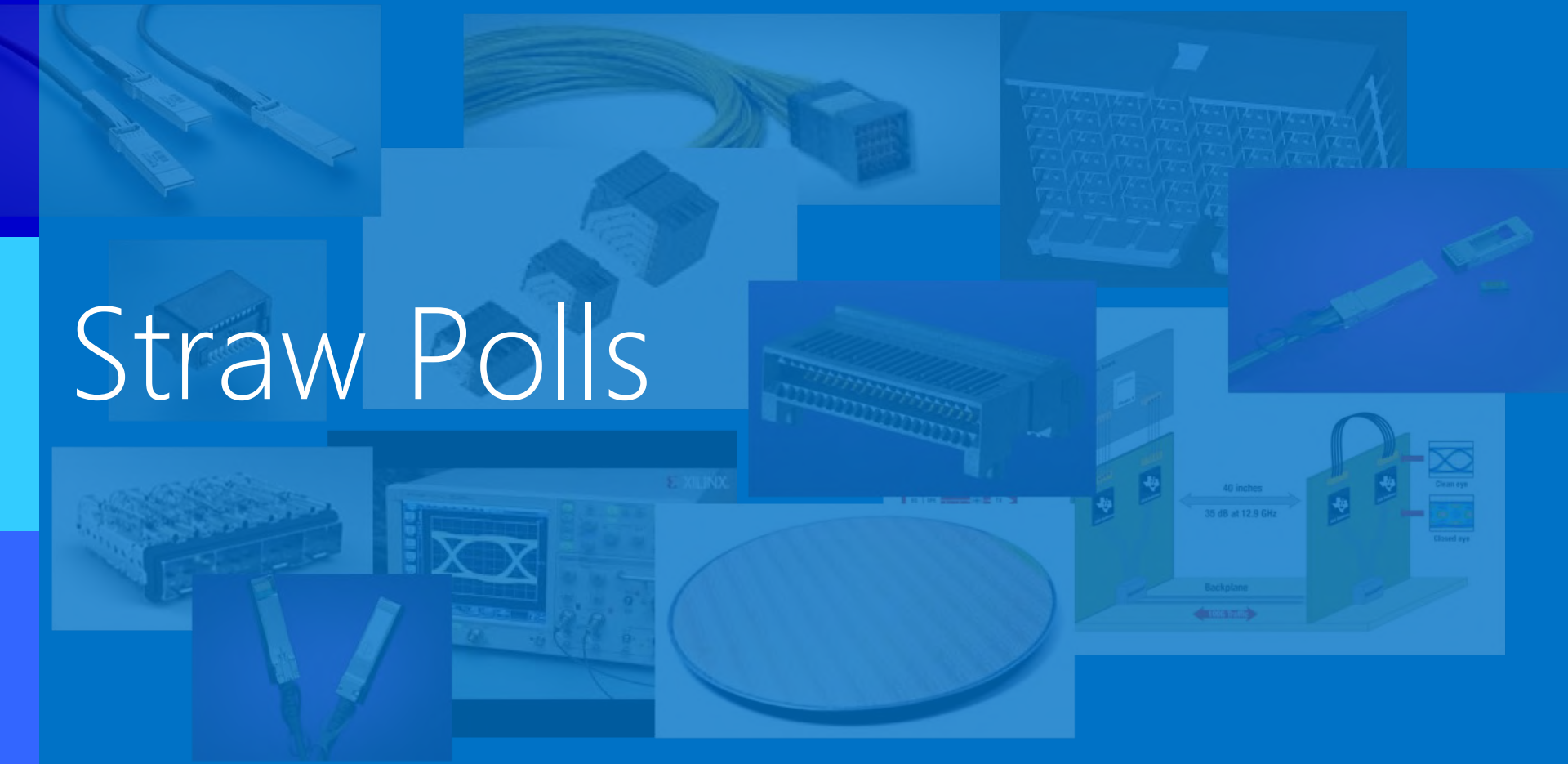
Adrian Amezcua, Prysmian
Alexander Umnov, Corning
Steve Swanson, Corning
Paul Vanderlaan, Nexans
Rakesh Sambaraju, Nexans
Robert Lingle, Jr., OFS
John Kamino, OFS
Mabud Choudhury, OFS
Paul Kolesar, CommScope
James Young, CommScope
Adrian Young, Leviton
Qing Xu, Belden
Phong Pham, USConec

Pavel Zivny, Tektronix

Ted Sprague, Infinera

# Supporters (2)

# Straw Polls



# Call-for-Interest Consensus

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

# Participation

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

- My company would support participation in a “Next-gen 200 and 400 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 January 2018 IEEE 802.3 interim meeting