



Consensus building discussion on next-generation MMF PMDs

Presented by Robert Lingle, Jr.

IEEE 802.3 Interim

Huntington Beach, CA

January 9th, 2017

Contributors and Collaborators

Supplied materials for the deck and/or contributed to market & technology discussions

- Dale Murray, LightCounting
- David Piehler, Dell-EMC
- Jeff Maki, Juniper
- Steve Swanson, Corning
- Brett Lane, Panduit
- Jonathan Ingham, FIT
- John Petrilla, FIT
- Jonathan King, Finisar
- John Johnson, Broadcom
- John Kamino, OFS
- Paul Kolesar, CommScope
- Rick Pimpinella, Panduit
- Jose Castro, Panduit
- Kenneth Jackson, Sumitomo
- David Lewis, Lumentum
- Frank Chang, Inphi
- Gary Nichol, Cisco
- Richard Mei, CommScope
- Paul Vanderlaan, Nexans
- Adrian Amezcua, Prysmian
- Alexander Umnov, Corning

Agenda

- Purpose
- Market needs
- Draft roadmap for next-gen MMF PMDs
- Technical feasibility
- Next steps
- Poll

Key take-aways

- There is a need for decreasing the cost of parallel MMF channels by reducing the number of fiber pairs. Examples:

Speed	Existing PMD and media	Possible media
100G	100GBASE-SR4 over 4f pairs w/ 12f MPO	1f pair w/ duplex LC
400G	400GBASE-SR16 over 16f pairs w/ 32f MPO	4f pairs w/ 12f MPO

- Technologies exist to do this, including use of multiple wavelengths over MMF
- Hope to build consensus towards a CFI in the near future

Rationale for MMF and VCSELs

- Historically MMF cabling + VCSEL transceivers have been the most popular paradigm in datacenters for many years due to:
 - Lower link cost
 - Lower power consumption
 - Better resistance to dirt at connections
 - Less sensitivity to reflections in short links

MMF remains the most popular media in the largest portion of the DC market

- Enterprise DCs remain majority of the market
- MMF comprises vast majority of optical links in Enterprise DCs
- Enterprise DCs are very diverse in size and link speeds
- The popularity of 10GBASE-SR and 40GBASE-SR4 created an embedded base of both duplex LC and 12f MPO-terminated MMF cabling in enterprise DCs

SMF transceivers are popular in the Cloud, but MMF persists also

- Use of MMF has declined in NAR hyper-scale datacenters due to desire for $\leq 500\text{m}$ reach
- MMF remains popular for
 - China cloud companies
 - Tier 2 cloud datacenters
 - $\leq 100\text{m}$ links in some hyper-scale architectures

Duplex and parallel MMF cabling each have their place and purpose

- Duplex (SR with duplex LC)
 - There is a large embedded base of single fiber pair cabling initially used for 10GBASE-SR
 - End-users expect a new speed to work over single fiber pair as technology matures, even if parallel was first used.
- Parallel (SR4 with 12f MPO)
 - Parallel may be the fastest time-to-market for a new speed
 - Popularity of 40GBASE-SR4 created a large embedded base.
 - Early applications are heavy on breakout in downlinks

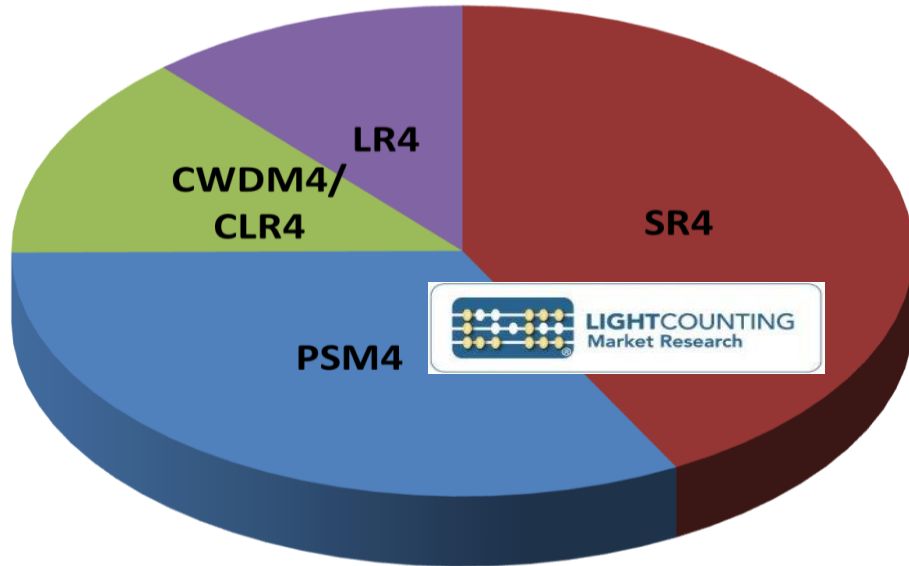
Breakout applications

- Breakout has been a very popular use of SR4
 - 40GBASE-SR4 is four high-density 10G links
 - 100GBASE-SR4 is four high-density 25G links
 - 200GBASE-SR4 can be four high-density 50G links
- For 400G, we could consider breakout to 50G and/or 100G server I/O
 - 400GBASE-SR4 breaking out to 4x100G
 - 400GBASE-SR8 breaking out to 8x50G or 4x100G

When will new MMF PMDs be needed?

- Largest enterprise DCs tend to adopt higher speeds within a year of hyper-scale DCs
 - Bulk of market will come later, but true for leading edge users
- Cloud companies in NAR & China continue to use SR4
- Input from switch makers
 - Sales of 100GBASE-SR4 were robust in 2016
 - 100G transceivers for duplex MMF could be sold now if offered
 - There is active discussion of 400GBASE-SR4 for ~2019
 - To support breakout, may consider 400GBASE-SR8 in QSFP-DD

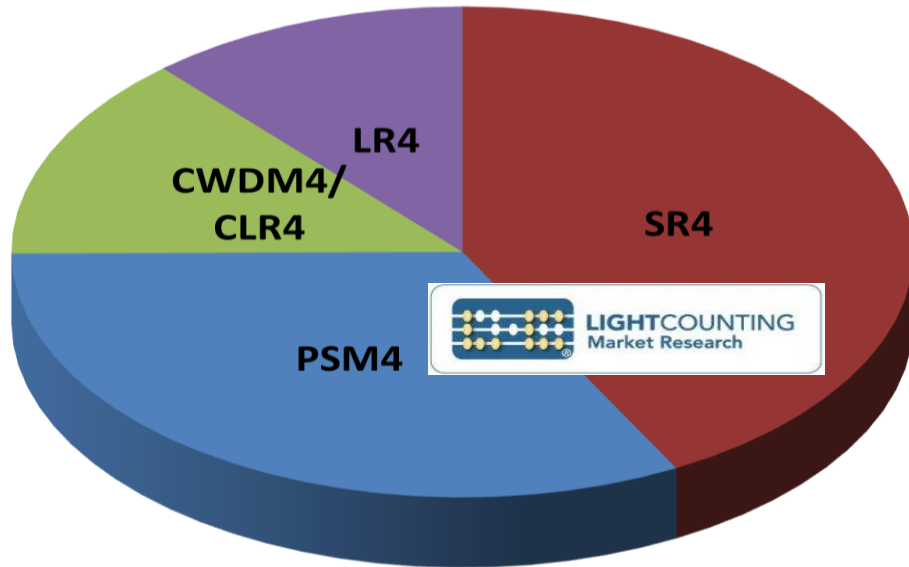
100GbE QSFP28 Consumption in 2016



- Taken together, SMF modules together have majority share
- But short-reach SR4 modules had the greatest individual contribution to 2016 shipments of QSFP28 modules

Slide courtesy of Dale Murray, LightCounting

100GbE QSFP28 Consumption in 2016



Point of Emphasis:
100GBASE-SR4 modules were sold in large quantities the first year that 100Gb/s Ethernet switches were commercially deployed in datacenters

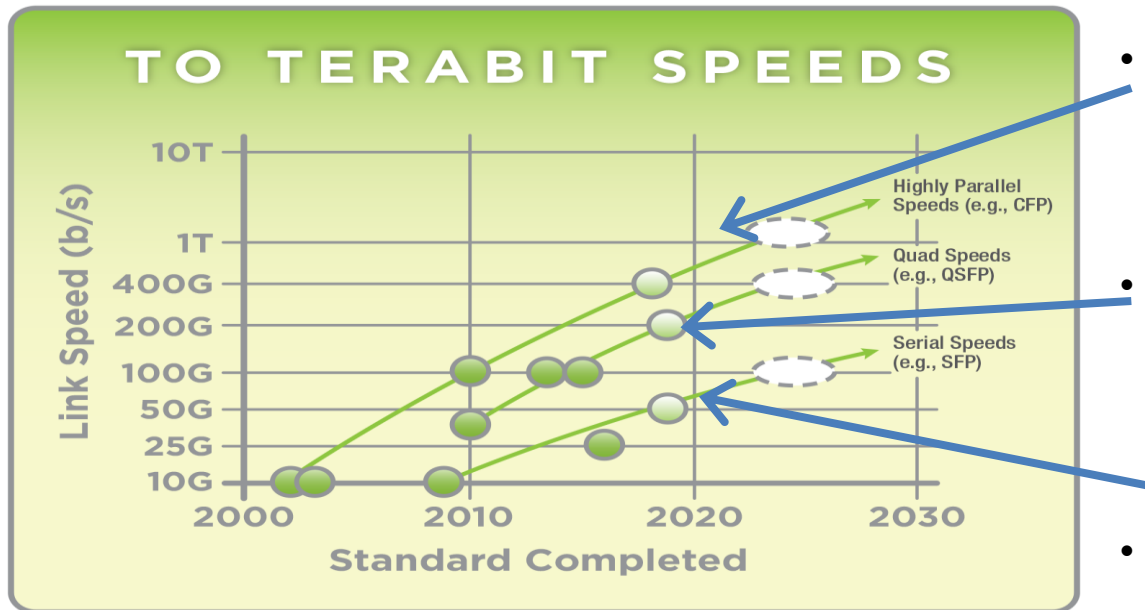
contribution to 2016 shipments of QSFP28 modules

Slide courtesy of Dale Murray, LightCounting

Methodology for drafting roadmap for MMF speeds

- New MMF speeds are needed within a year after new DC switch speeds introduced
 - SR4 first
 - Duplex soon after
- Follow the Ethernet Alliance roadmap for SR4
- Put duplex transceivers a year later
- Transceivers requiring 100G lanes are delayed to 2023 when they should be technical feasible
- Consider both 200G & 400G switching will be used in enterprise DCs (though this is a matter of debate)

Ethernet Alliance roadmap was the starting point



- **Client Optics** – needed soon to connect IP routers to 400G DWDM gear for long haul transport
- **Inter-Switch Ports** – 100G data center switches deployed in 2016 in hyper-scale, but 40G will remain common in enterprise; **4X server I/O**
- **Server I/O** – 25G and 50G in cloud, but 10G in enterprise; **advances by 2X**

Enterprise shifted by 1 year from cloud



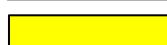

DC Type	Link	Speeds (Gb/s) Ramping																	
Enterprise	Switch	40			100			200				400			800			1600	
	Server	10			25			50				100			200			400	
Cloud	Switch			100			200				400			800			1600		
	Server			25			50				100			200			400		
Year First Application		2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	

- Year is when first applications are needed, not high volume demand.
- EA roadmap assumed to represent cloud needs
- Early adopter, large enterprise needs shifted one year later

A draft roadmap for MMF transceivers

Speed (Gb/s)		25	40	50	100	200	400	800	1600
Parallel (telecom)	24/32f MPO SR10 SR16				2010 10G lanes		2019 25G lanes		
Parallel (octal breakout)	24f MPO SR8						2019 50G lanes	2023 100G lanes	
Parallel (quad breakout)	12f MPO SR4		2010 10G lanes		2015 25G lanes	2019 50G lanes	2019 WDM 25 or 50G lanes	2023 WDM 100G lanes	2028 WDM 200G lanes
Duplex	LC SR	2016 25G lanes	2013 BiDi SWDM4 10&20G lanes	2018 50G lanes	2017 WDM 25&50G lanes	2020 WDM 50G lanes	2023 WDM 100G lanes	2028 WDM 200G lanes	
Speed (Gb/s)		25	40	50	100	200	400	800	1600






- “WDM” refers to any use of multiple wavelengths, whether BiDi, SWDM2, or SWDM4. "Lanes" in the chart refers to optical lanes with rate in Gb/s, whether wavelengths or parallel fibers. Speed divided by lane rate yields # of Tx or Rx required.
- Parallel SR4 modules needed on the EA timeline to support some cloud and large enterprise needs
- Duplex SR4 modules 1 year later, or when technically feasible
- Would SR8 be useful for an eight-way breakout application?

	Existing IEEE standard
	In progress in 802.3bs
	In progress in 802.3cd
	Proprietary

Possible work for a Study Group

Speed (Gb/s)		25	40	50	100	200	400	800	1600
Parallel <i>(telecom)</i>	24/32f MPO SR10 SR16				2010 10G lanes		2019 25G lanes		
Parallel <i>(octal breakout)</i>	24f MPO SR8						2019 50G lanes	2023 100G lanes	
Parallel <i>(quad breakout)</i>	12f MPO SR4		2010 10G lanes		2015 25G lanes	2019 50G lanes	2019 WDM 25 or 50G lanes	2023 WDM 100G lanes	2028 WDM 200G lanes
Duplex	LC SR	2016 25G lanes	2013 BiDi SWDM4 10&20G lanes	2018 50G lanes	2017 WDM 25&50G lanes	2020 WDM 50G lanes	2023 WDM 100G lanes	2028 WDM 200G lanes	
Speed (Gb/s)		25	40	50	100	200	400	800	1600

- The PMDs outlined **in red** could be vetted in a Study Group
- Do not anticipate changing PCS at this time
- Note that proprietary modules for 100G based on multiple wavelengths (both 4-lambda NRZ & 2-lambda PAM4) are expected in the market in 2017.

	Existing IEEE standard
	In progress in 802.3bs
	In progress in 802.3cd
	Proprietary
	Possible Study Group

Multiple wavelength technology can help realize the PMDs deemed necessary




- 25GBaud VCSELs are in production, used with NRZ & PAM4, but 50GBd VCSELs supporting PAM4 remain several years in the future.
- Multiple wavelengths can help beyond:
 - 50 Gb/s for one pair of MMFs (LC terminated)
 - 200 Gb/s for four pairs of MMFs (MPO terminated)

Multiple wavelengths over MMF is technically feasible and commercially viable

- Two wavelength 40G modules (BiDi) have been shipping since 2013
- Four wavelength 40G modules (SWDM4) began shipping in 2016
- Both two & four-wavelength 100G modules are in advanced stages of development now
- Multi-wavelength modules work over OM3/4/5 MMF

MMF PMD Roadmap – all options

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		

	Existing IEEE standard
	In progress in 802.3bs
	In progress in 802.3cd

Proposed SWDM Nomenclature
 SRx.y x = # fiber pairs
 y = # wavelengths

Next steps

- Regular teleconferences under the auspices of the New Ethernet Applications process will occur
- A CFI deck will be drafted starting with the next teleconference
- Input from a broad range of 802.3 participants is desired