100Gb/s SMF PMDs

40Gb/s and 100Gb/s Fiber Optic Task Force

SMF Ad-Hoc

2 April 2013

Chris Cole

IEEE

Finisar

Outline

- End User Comments
- Total Link Cost Ratios
- PSM Observations

Hong Liu, et al. (Google) Comments

2011 Data Center optics requirement:

2km/5dB duplex SMF

http://www.ieee802.org/3/100GNGOPTX/public/jan12/cole_01a_0112_NG100GOPTX.pdf#page=12

2013 publication:

"C. Kachris, et al, Editors, "Optical Interconnects for Future Data Center Networks", Springer Media, New York, 2013, Chapter 2: Hong Liu, et al., "Optical Interconnects for Scale-out Data Centers", p24.

> "Parallel optical transceivers using ribbon fiber and MPO connectors is widely deployed within datacenter and HPC environments. However, the MPO connector and ribbon fiber can incur a significant portion of the entire data center network cost. Scaling bandwidth through parallelism in this manner can lead to an unmanageable volume and size in the fiber infrastructure. Thus, when longer reach interconnects are required, this approach becomes obsolete."

3/24/13 HL clarifying email:

> "Our preference for fiber cable plant is duplex SMF."

Nathan Farrington (Facebook) Comments

3/17/13 OIDA Data Center Workshop, OFC/NFOEC, as reported by Stephen Hardy, “How to sell to Facebook (and Google and Microsoft)”, Lightwave Blog:

> ‘Sure enough midway through his presentation {Nathan Farrington}, presented a list of requirements. Well, to be accurate, it was a list of requirement – a repetition of the single phrase “Looking for more bits per second for less capex and opex.” Chuckle, chuckle. But then after truly offering a wish list that included a universal transceiver based on no more than two fibers because “parallel fiber is lame, …”’
>
> http://www.lightwaveonline.com/blogs/lightwave-blog/2013/03/How-to-sell-to-Facebook-and-Google-and-Microsoft.html?cmpid=EnlDirectMarch182013

3/24/13 NF clarifying email:

> “Today it takes 2 fibers for 10G and tomorrow it will take 12 fibers for 40G. We don't want to do 8 fibers because that could preclude a future upgrade option. And we don't want to do 12 fibers because it is 6x the cost.”

Tom Issenhuth (Microsoft) Comments

3/18/13 OFC/NFOEC Workshop "How Can Optics Address Bandwidth and Latency Bottlenecks in Data Centers?", OM1F, Tom Issenhuth's last slide:

- "Add it all together and we require cheap switches with integrated 2 fiber single mode based optical interfaces to achieve the required scale
  - Integrated optics
    - Only way to achieve the switch densities required
    - Would be supported with high count MPO connectors
    - Lack of pluggability not a concern
  - Single Mode Fiber
    - Required to support 1km reaches in data center complexes
    - Required to future proof the fiber infrastructure
  - 2 Fiber interfaces
    - Fiber infrastructure costs and management challenges are large at scale
    - 2 fiber solutions require 1/6th the fiber of 8 fiber solutions
      - Effort to reclaim the 4 unused fibers is not worth it"

100GE PMD Relative Cost

SR10 CXP = 1x	SR4	PSM4	LR4 Gen1	LR4 Gen2	LR4 Gen3
Form-factor	CFP4/ QSFP28	CFP4/ QSFP28	CFP	CFP2	CFP4/ QSFP28
petrilla_02a_0112	1.2x	3x to 4x			
nicholl_01_0112		4x	14x	6.5x	
cole_02a_0312	1.2x	3x to 4x	16x	8x	5x to 6x
anderson_01_0113	<4x	4x		9.3x	
petrilla_03a_0113	1.1x	4x		12x	
welch_01a _0113		0.82x			3.5x

http://www.ieee802.org/3/bm/public/mar13/cole_01_0313_optx.pdf#page=6

Cabled Fiber Link Relative Cost

Channel Type	Double-Link (DL)		
Fiber Type	100m	300m	500m
2f OS2 SMF	1.5	2	2.5
8f OS2 SMF	6	8	10
12f OS2 SMF	9	12	15

Total Link Cost Ratio Scenarios

- Direct Module and External Cabling (2f/8f)

  http://www.ieee802.org/3/bm/public/mar13/cole_01_0313_optx.pdf#page=8

- Direct Module and Internal Cabling (2f/8f)

  http://www.ieee802.org/3/bm/public/mar13/cole_01_0313_optx.pdf#page=9

- Current Distributor Module and External Cabling (2f/8f/12f)

  (Distributor Modules are <1% of 100GbE volume)

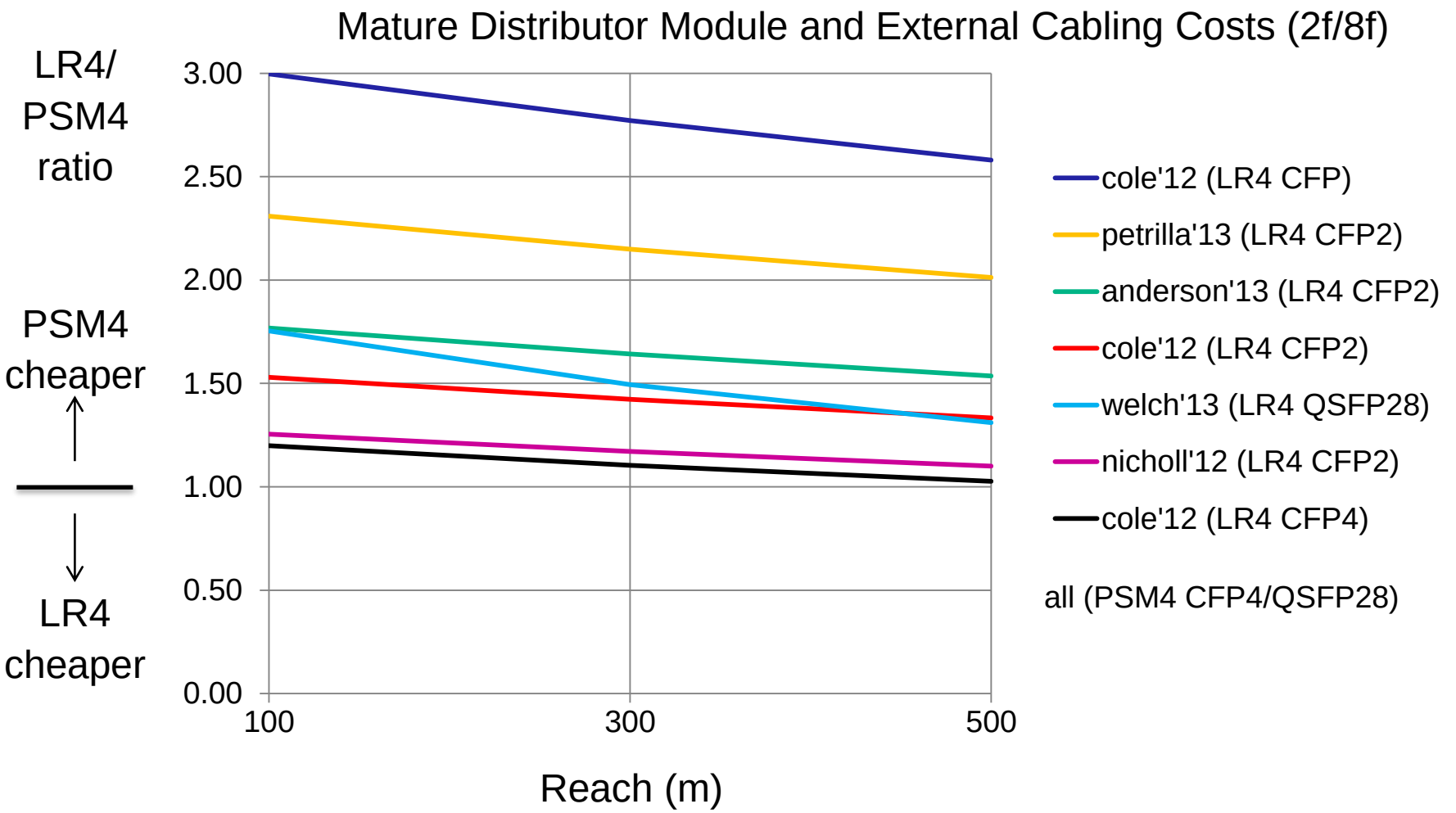
  http://www.ieee802.org/3/bm/public/mar13/kolesar_01a_0313_optx.pdf#page=13

  (right hand side graph)

- Mature Distributor Module and External Cabling (2f/8f & 12f)

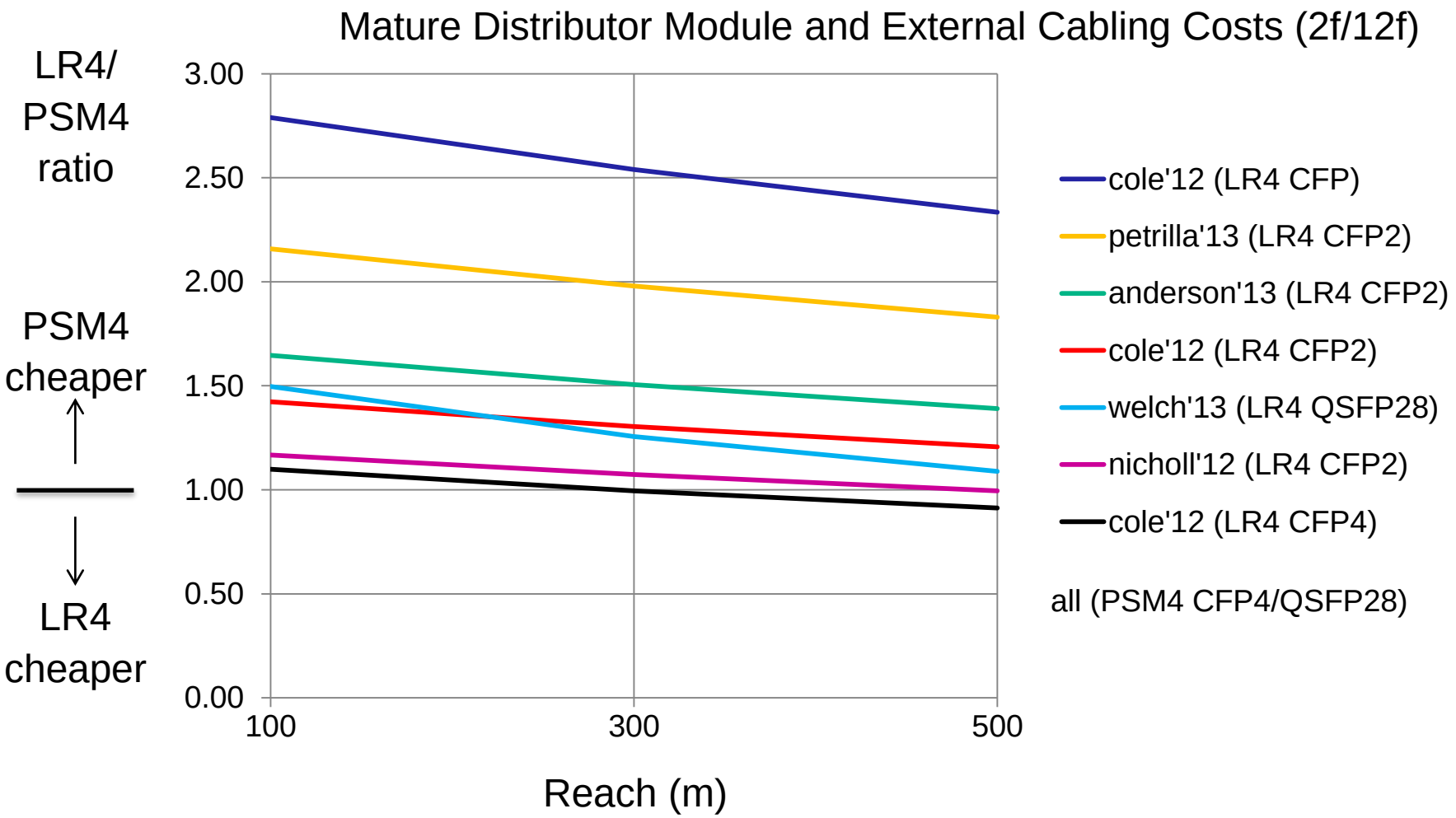
  Graphs on following pages

100GE LR4/PSM4 Total Link Cost Ratio



$$LR4/PSM4 = (2*LR4 + \alpha*DL\ 2f\ OS2)/(2*PSM4 + \beta_1*DL\ 8f\ OS2)$$

100GE LR4/PSM4 Total Link Cost Ratio



$$LR4/PSM4 = (2*LR4 + \alpha*DL\ 2f\ OS2)/(2*PSM4 + \beta_2*DL\ 12f\ OS2)$$

PSM Observations

- Applications:
  1. 400GbE Gen1 SMF PMD
  2. SMF-only Datacenter: PSM4 instead of SR4
  3. Multi-link MPO based Transceiver: Nx10/40GbE
- Some scenarios have lower cost versus WDM, but …
- 100GbE PMD Broad Market Potential has not be shown
- Niche application best standardized in a MSA
- MSA specified optics not an issue for large end users