100 G Active Optics Cables

Karim Tatah
Cray Inc.
ktatah@cray.com
715-726-4678

IEEE 802 Plenary Meeting
High Speed Study Group
Atlanta, Georgia
November 14th, 2007

100 G Active Optics Cables

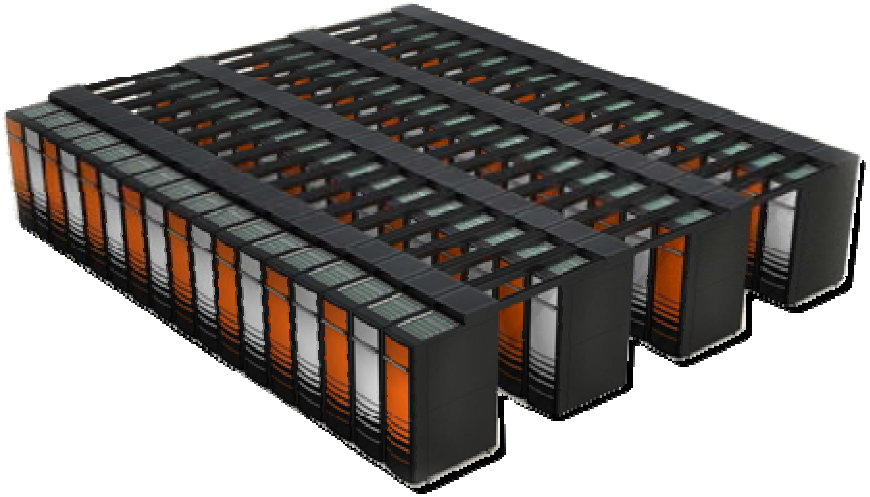
- System interconnect in Cray machines
- Cable reach dependence on network architecture
- Cray HPC interconnect approach
- What is the active cable advantage?
- The 12x duplex channels at 10+ Gb/s per channel active cable option description
- VCSEL or Silicon Photonics options
- Advocate 12x active optics cable for the “at least” 100 meters on OM3 physical layer objective:
  - Need to consider 12x duplex active optics @10+ Gb/s per channel and form an MSA early to meet HPC timeline
  - Need to define a 12x active cable electrical connector early

Cray Technology

Multiple Processing Technologies	Vector, scalar, massive multi-threading and application accelerators
Network Communications	Custom interconnect and communications network
Systems Administration & Management	Software and tools to manage thousands of processors as a single system
Packaging	Very high density, upgradeable, liquid and air-cooling
Adaptive Supercomputing	Single integrated system

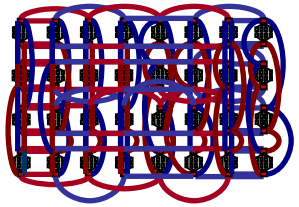
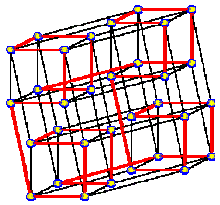
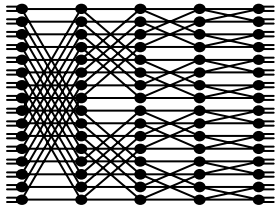
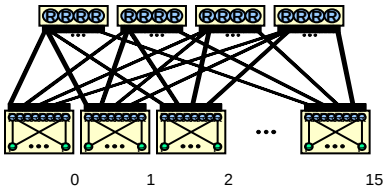
Cray System Network Bandwidth: Examples

- An example of large bandwidth system XT3/XT4 scalar system:
  - Processing: ~ 100 TFLOP
  - Bandwidth: 48 duplex channels x 3.125 Gb/s = 150 Gb/s per point-to-point cable link



- An example of large bandwidth system XT5/XT5$_h$ vector-scalar hybrid system:
  - Performance: ~ 20 GFLOP per node
  - Bandwidth: 24 duplex channels x 5 Gb/s = 120 Gb/s per point-to-point cable link

A Simplified View of Supercomputing Networks

	Network Graph	Bandwidth	Latency or Hops	Cost	~ Max Cable Reach (m)
Torus (XT3 and XT4)		☺	☺	☺☺☺☺	7
Hypercube		☺☺	☺☺	☺☺☺	8
Butterfly		☺☺☺	☺☺☺☺	☺☺	40
Fat Tree or Folded Clos (XT5)		☺☺☺☺	☺☺☺☺	☺	20

Links Speeds and Feeds

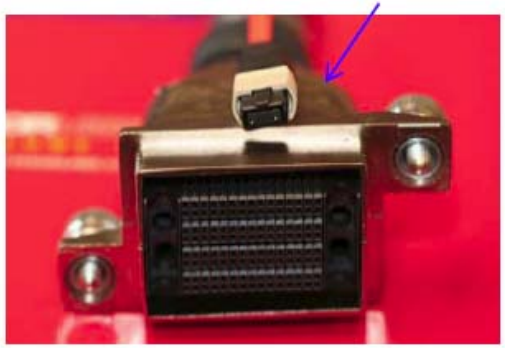
- GEthernet (GE), InfiniBand (IB), and Fibre Channel (FC) and Cray's point-to-point network links bandwidth

Channel (#) Rate (Gb/s)	1x	4x	8x	10x	12x	24x	48x
2		8 G FC	16 G FC				
2.5		10 GE 10 G-IB			30 G-IB		
3.125							150G-Seastar (XT4)
4	4G-FC						
5		20 G-IB			60 G-IB	120G-Potter (XT5)	
6.25		25G					
8	8G-FC						
10	10 GE	40 GE 40 G-IB 40 FC		100 GE	120G-IB 120-FC		
25		100 GE					

High Density Copper Cables

- 96 differential pairs connector

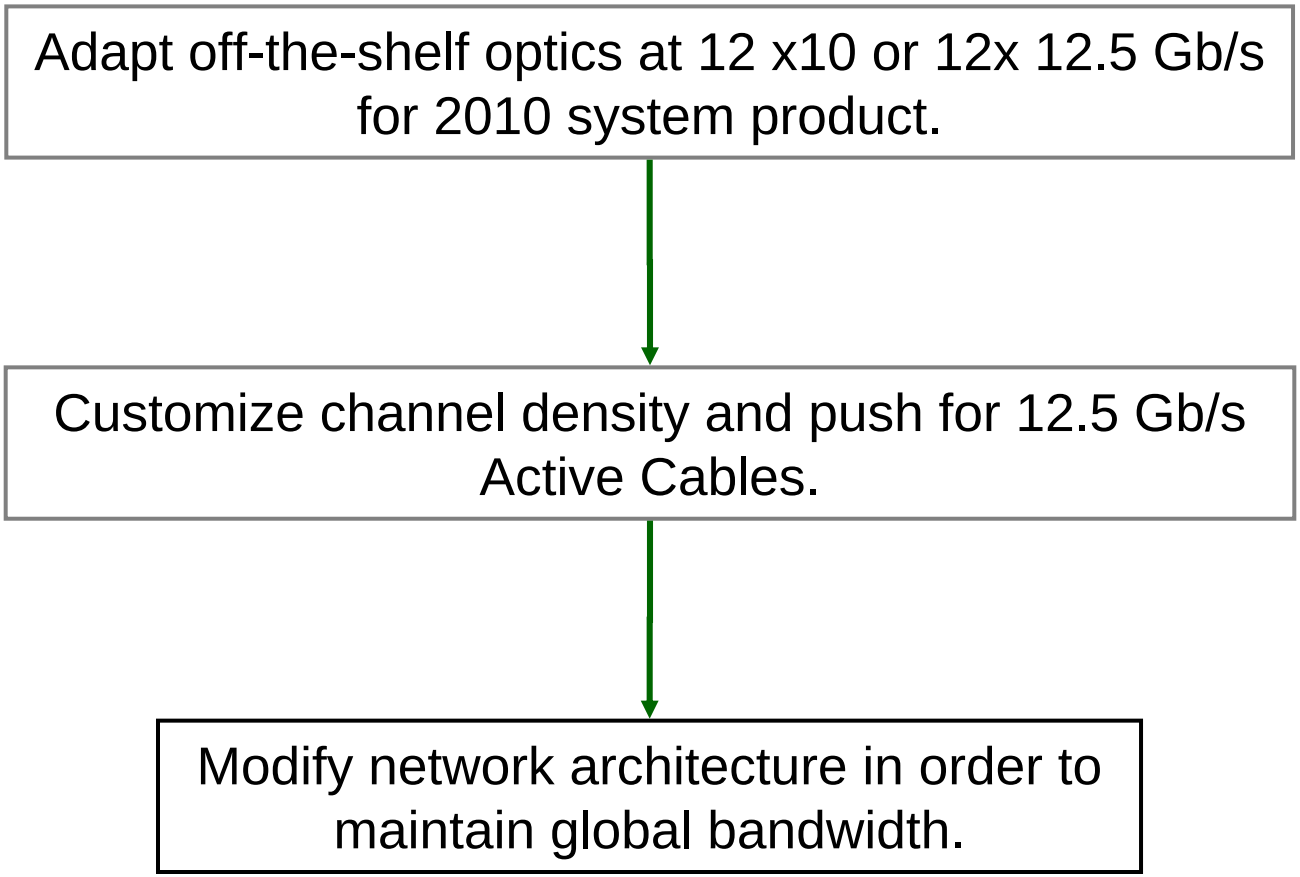
24- fiber MTP shown for scaling purposes



- 48 differential pairs connector

- Space congestion due to the large number of cables and their bulk leaving a chassis or cabinet present cable managements challenges and repair-ability issues such as the limited access to the network boards.

Cray HPC Interconnect Approach

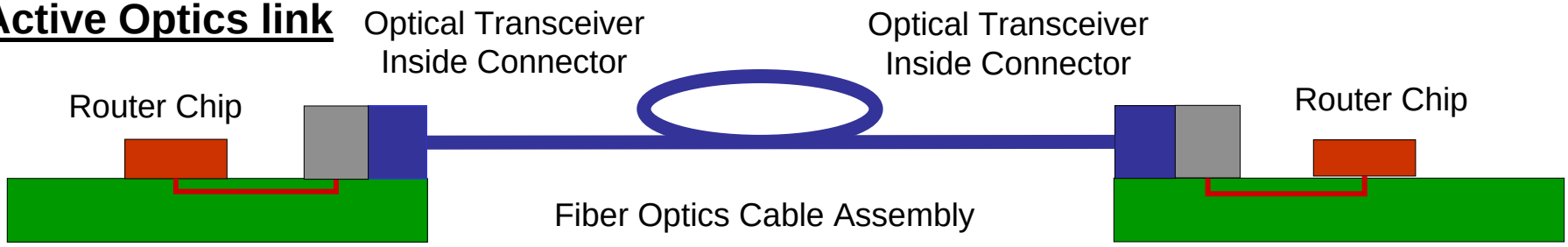


Interconnect Options

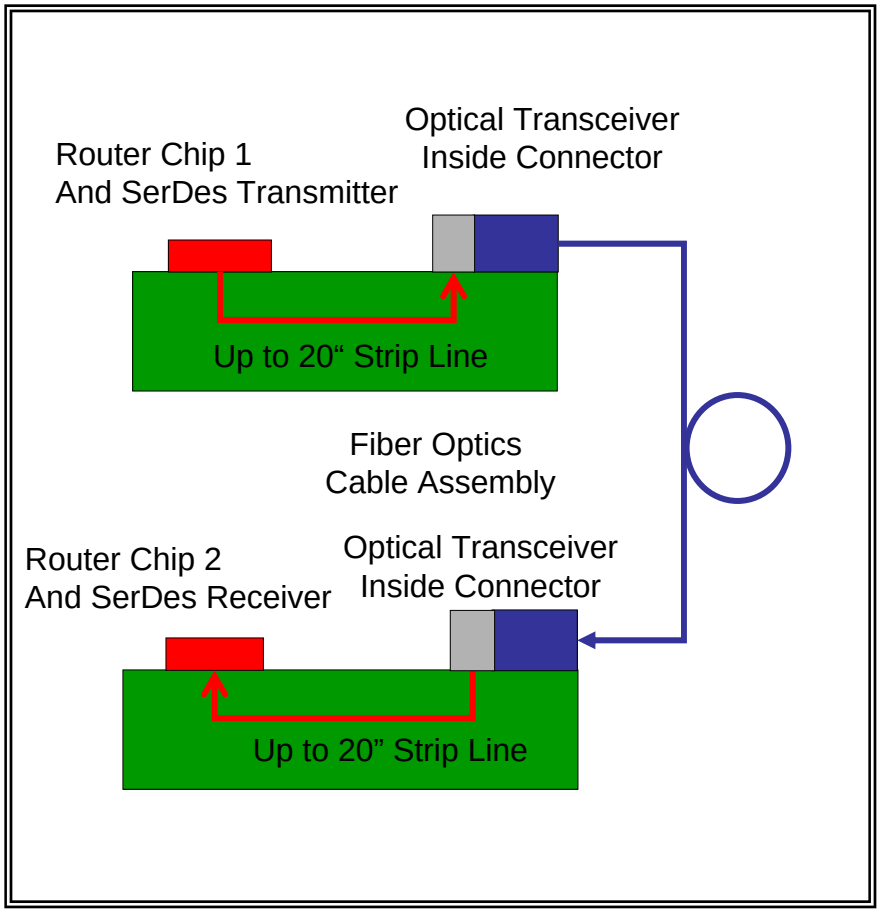
Copper with no repeaters

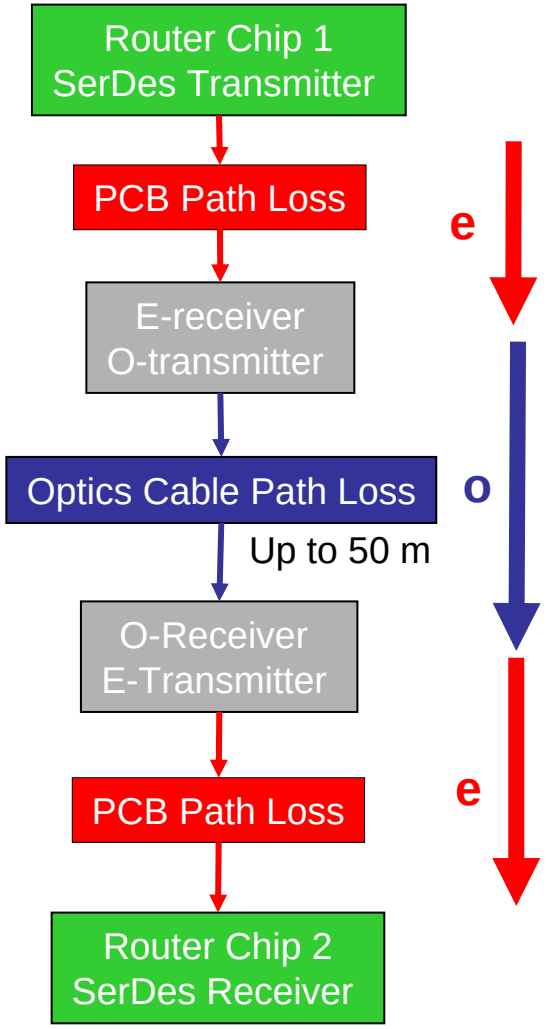
Copper with repeaters

Active Optics link



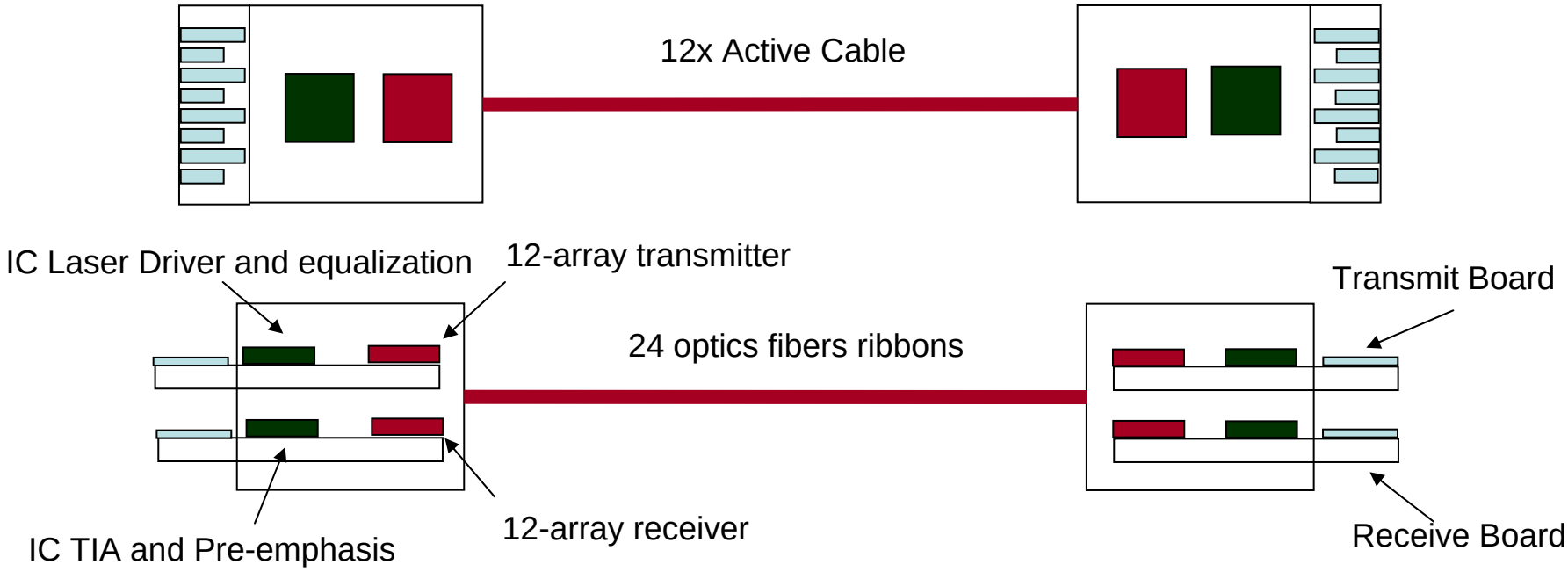
Active Optics link





What is an Active Optics Cable?

- Optics and IC Embedded In Copper Connector
- Standard copper connector ends allows interchangeability with copper cable assemblies
- Fiber Ribbons Hard Wired to OE Chips (Pigtail with no optical connectors)



Active Optics Cables

Pros:

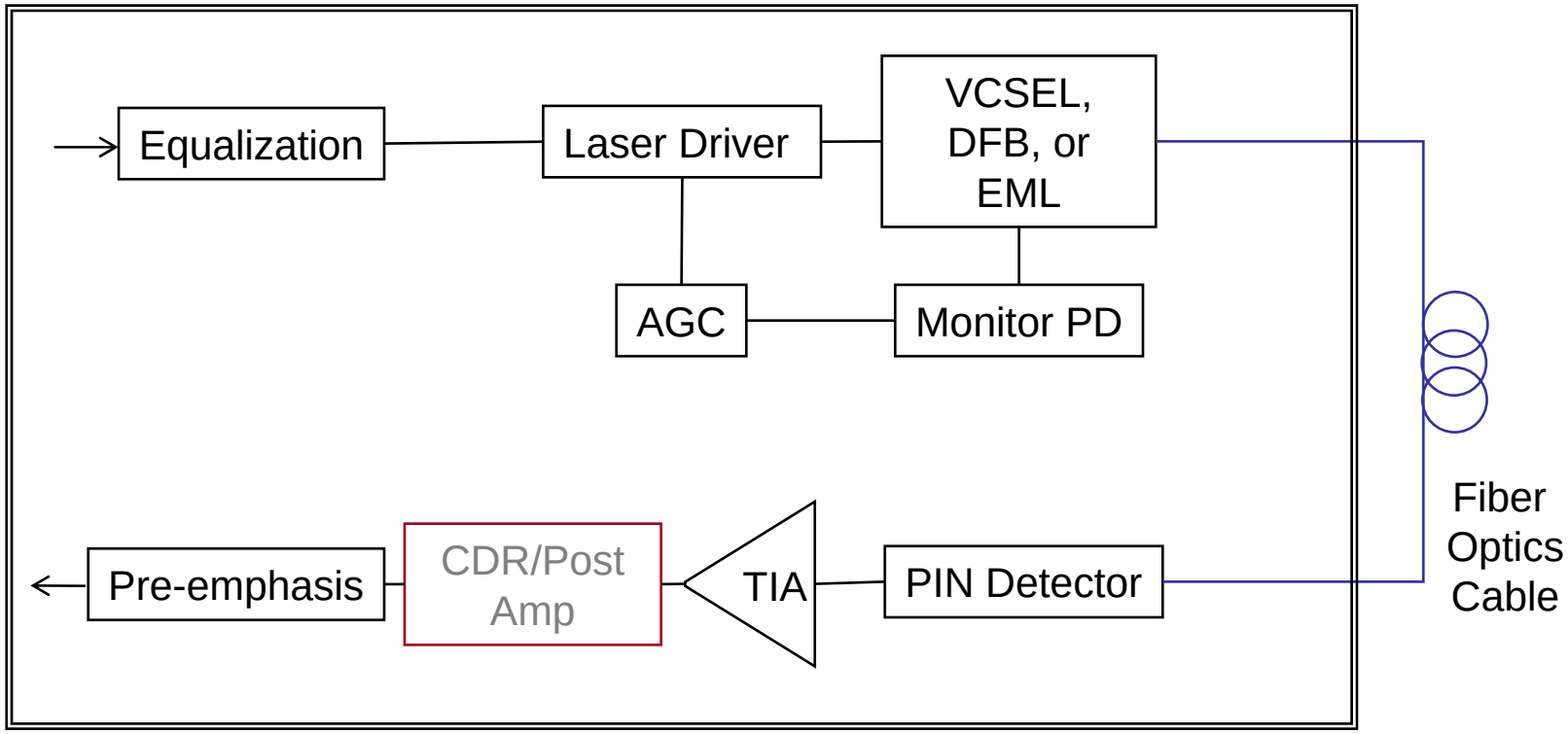
- Low cost
- Designed as source-fiber-detector matched for low cost active cable assemblies
- Compact fiber optics cables
- Optics “inside” : Optics hidden from user/customer
- No cleaning of the optics needed
- Provide capability short and long distances
- Allows for low power components
- Tight bend radii (saves cabinet space up to 4”)
- Lower weight than copper (Glass and copper densities 2.2 g/cm^3 and 8.9 g/cm^3 respectively)
- Better security (hard to tap into the optics)

Cons:

- Repair-ability
- Re-configurability

Transceiver Functional Content

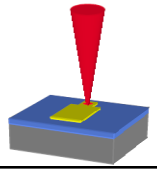
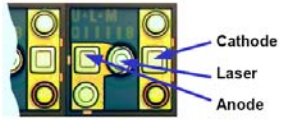
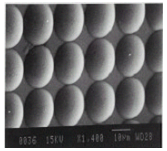
- Channel Definition



Examples of Active Optics Cable Suppliers

Manufacturer	4-channel active cable	Connector Package	Laser/Modulator Transmitter
A	4 x 5 Gbps	QSFP	VCSEL
B	4 x 3.125 Gbps	microGIGA^{CN}	VCSEL
C	4 x 5 Gbps	microGIGA^{CN}	VCSEL
D	4 x 10 Gbps	QSFP	DFB and Silicon Mod.
E	4 x 5 Gbps	QSFP and microGIGA^{CN}	VCSEL

WDM and Parallel Optics in a 12x Active Cable

Compts.	WDM	Parallel Optics
Sources	**Edge Emitter and Modulator** ♦ Data rate 25 Gb/s possible ♦ Bias current 10-100 mA (10x power of VCSELs)  	**Vertical Cavity Surface Emitting Laser** ♦ Data rate at 25 Gb/s not possible at present ♦ Bias current 1-10 mA (1/10 th DFB power)  
Optical Fiber	**2 single mode fibers:**  ♦ Core diameter: 8.2 µm ♦ Cladding diameter: 125 µm ♦ Numerical Aperture: 0.12 ♦ Attenuation: 0.2 dB/km at 1550 nm ♦ Alignment Tolerance: ~ 1 µm	**24 multimode fibers:**  ♦ Core diameter: 50 µm ♦ Cladding diameter: 125 µm ♦ Numerical Aperture: 0.2 ♦ Attenuation: 2.4 dB/km at 850 nm ♦Alignment Tolerance: ~ 5 um
Detectors	**Waveguide or PIN or APD**	**Surface PIN**
Optical Couplers	**Optical Mux/Demux** 	**Lens Array** 

Optics Vs Copper Actives

Parameter	Optics Actives	Copper Actives
Wavelength (cm @10 Gbps in cable)	0.57 x10^{-4}	2.3
Reach (m) (@10 Gbps)	Up to 100	5 to 10
Network Board Trace (cm)	10-50	10-50
Loss (dB/m)	0.002	2
Cable coupling per end (1 dB)	~ 1	~ 2
Cable Weight (Kg)	☺	☹
Cable Volume (cm^3)	☺	☹
BER	10^{-15}	10^{-12}
Cable Time of Flight Latency (ns/m)	4.99	4.28
Latency per Active End (ps)	~ 0.3	NA
Bend Radius (mm)	5 to 10	75
Power per channel (mW)	~ 200	~ 100
EMI/EMC	☺	☹

Insertion Loss dependence on Cable Length

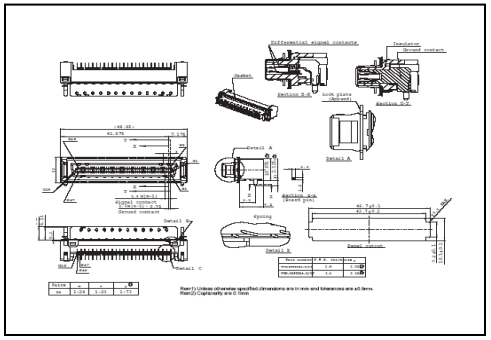
- Copper Cable - 26AWG Twinaxial @10 Gb/s:
  - Insertion loss ~ 2 dB/m

  Example: 20 dB power is lost in 10 meters cable length – it is a large power to recover

- Fiber Optics - 50 μm OM3 Fiber @10 Gb/s:
  - Insertion loss ~ 0.002 dB/m

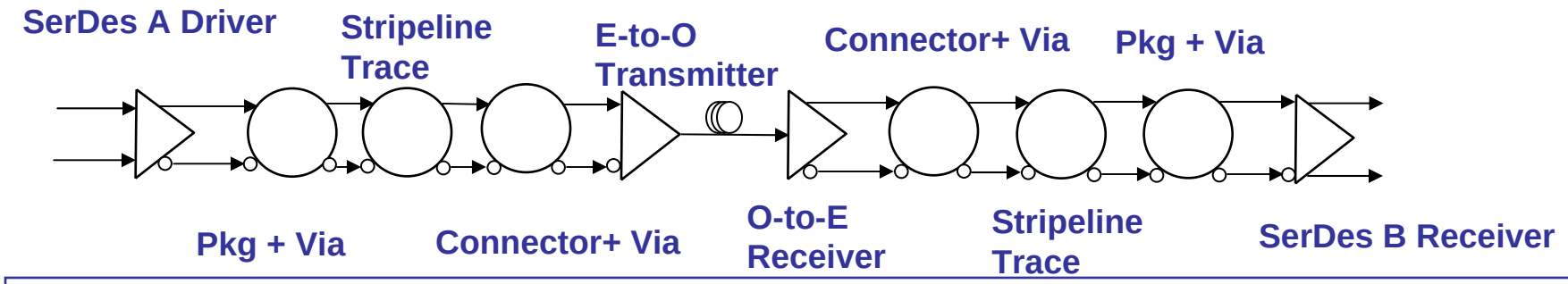
  Example: 0.02 dB power is lost in 10 meters – it is neglible power that is lost in transmission – most power expanded in the e-o and o-e conversion

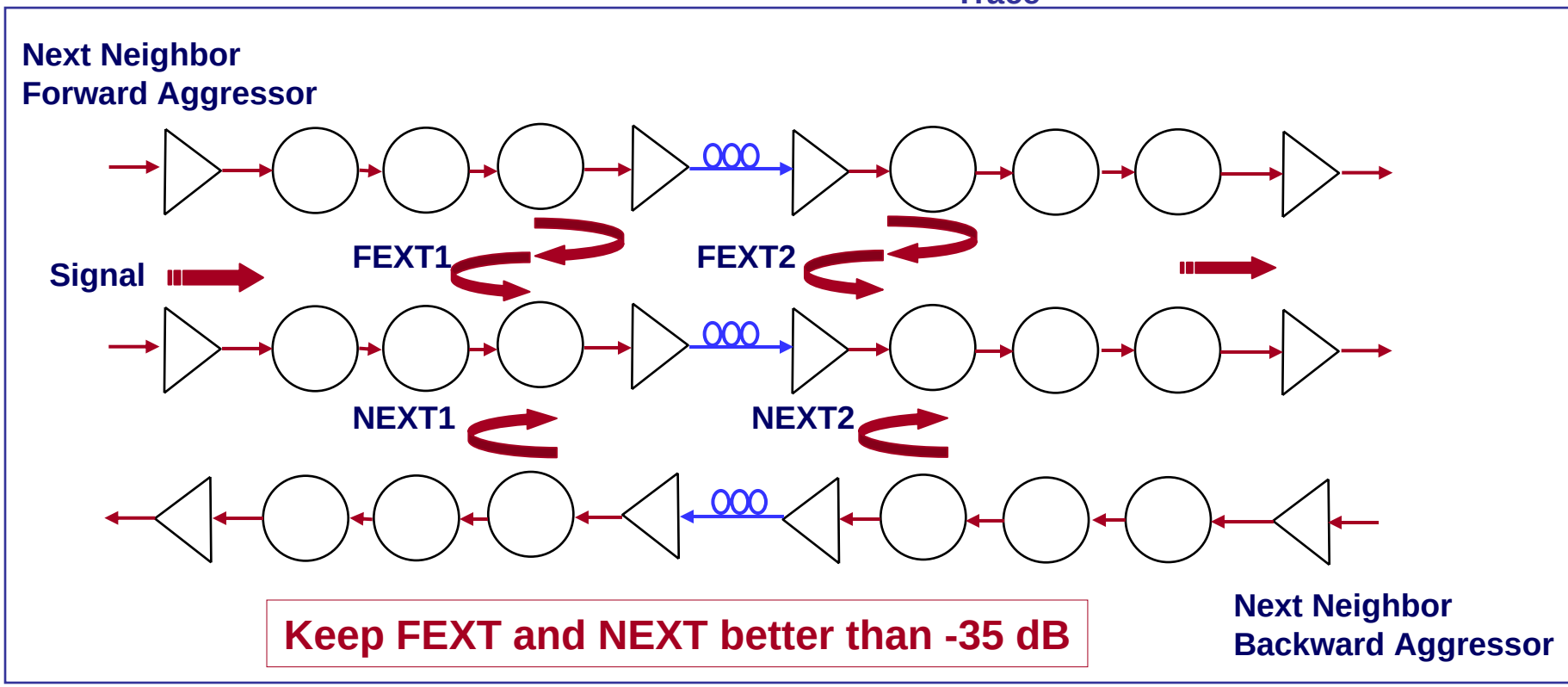
Electrical Connector Options for 12x Active Cables

Does not meet Cray X-talk budget at 10 Gbps	Does not meet Cray board edge density
12x Connector proposed at a recent IBTA ♦ ~ 1” Wide with better edge board density ♦ Thermal: Actives inboard that offers cooling advantage	microGiGa CN 12x Connector  ♦ ~ 2” Wide with worse board edge density ♦ O-E active components outboard hard to cool

Active cable industry is encouraged to propose soon a connector and form a 12x active cable MSA.

E-connector Xtalk and Signal Integrity of Complete SerDes –to-SerDes Path





Conclusion

- Advocate 12x active optics cable for the “at least” 100 meters on OM3 physical layer objective:
  - Need to consider 12x duplex active optics @10+ Gbps per channel and form an MSA early to meet HPC timeline.
  - Need to define a 12x active cable electrical connector early.

Thank You