Fast Readout Links

F. Vasey, CERN, PH-ESE

- Optical interconnects by network application
- Comparison with developments for LHC and HL-LHC
- What to expect for future High Energy Physics experiments

francois.vasey@cern.ch

Optical Link Metrics

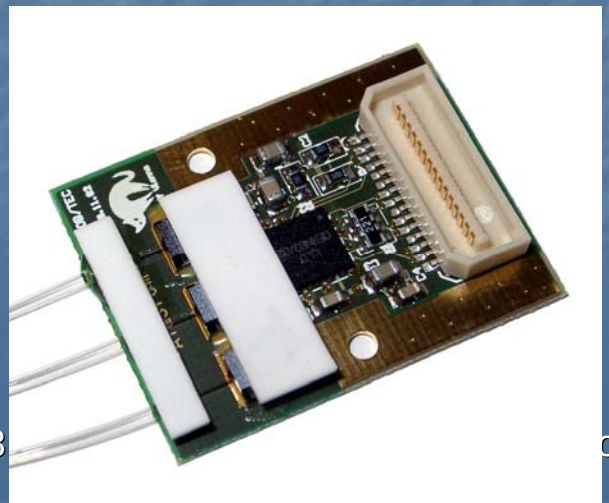
- Bandwidth
- Energy (pJ/bit, mW/Gbps)
  - 1mW/Gbps = 1pJ/bit
    - 1 google search = 1kJ → 1 Pbits (@1pJ/bit) !!!
- Density (mm^2/Gbps)
- Cost ($/Gbps)
- Environmental Resistance

Use and Misuse of Scaled Metrics

- mW/Gbps, mm^2/Gbps, $/Gbps
  - Systematically favour very high bandwidth systems
  - Cannot be used to extrapolate likely figures for bandwidth-distributed systems
    - Per device (TRx) scaling more appropriate
    - mW/TRx, mm^2/TRx, $/TRx
  - Nevertheless very useful to compare systems

Existing Optical Links in CMS (1)

- CMS Tracker AOH
  - 39k links
  - 40MS/s x 8 bits = 320 Mbps
  - 100 mW, 3 channels (dependent on laser bias)
  - 30mm x 23mm, 3 channels
  - 300 CHF / channel
- CMS ECAL GOH
  - 10k links
  - 800 Mbps
    - GOL operates up to 1600 Mbps



Aggregate bandwidth ~ Tk: 12.5Tbps
ECAL: 8Tbps

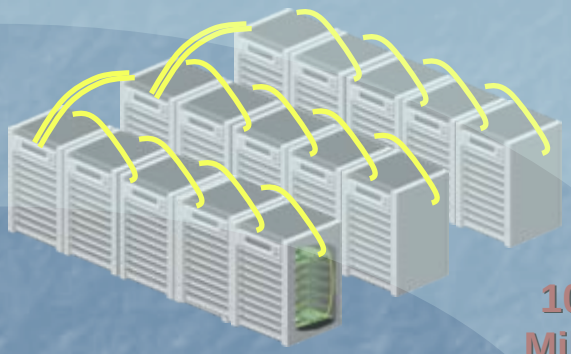
Existing Optical Links in CMS (2)

	AOH 1997
Bandwidth Gbps	0.32
Reach m	100
Energy pJ/bit, mW/Gbps	150 * extrapolating to bidi link
Density mm^2/Gbps	1000 * extrapolating to bidi link
Cost $/Gbps	900

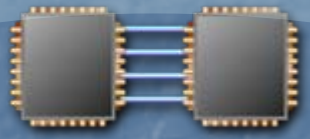
Interconnects by Network Application

Metro & Long Haul (0.1-80km)
SONET
WDM
Ethernet

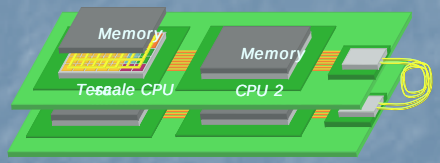
Data Center (Rack to Rack) 1-100 Meters
Ethernet
InfiniBand
Fibre Channel / SAS
Proprietary



Backplane



Chip to Chip 1 – 50 cm



Board to Board ~0.5 -1.5M

Intra-Rack
Ethernet
InfiniBand
Fibre Channel / SAS
PCIe
Proprietary

100s of Thousands

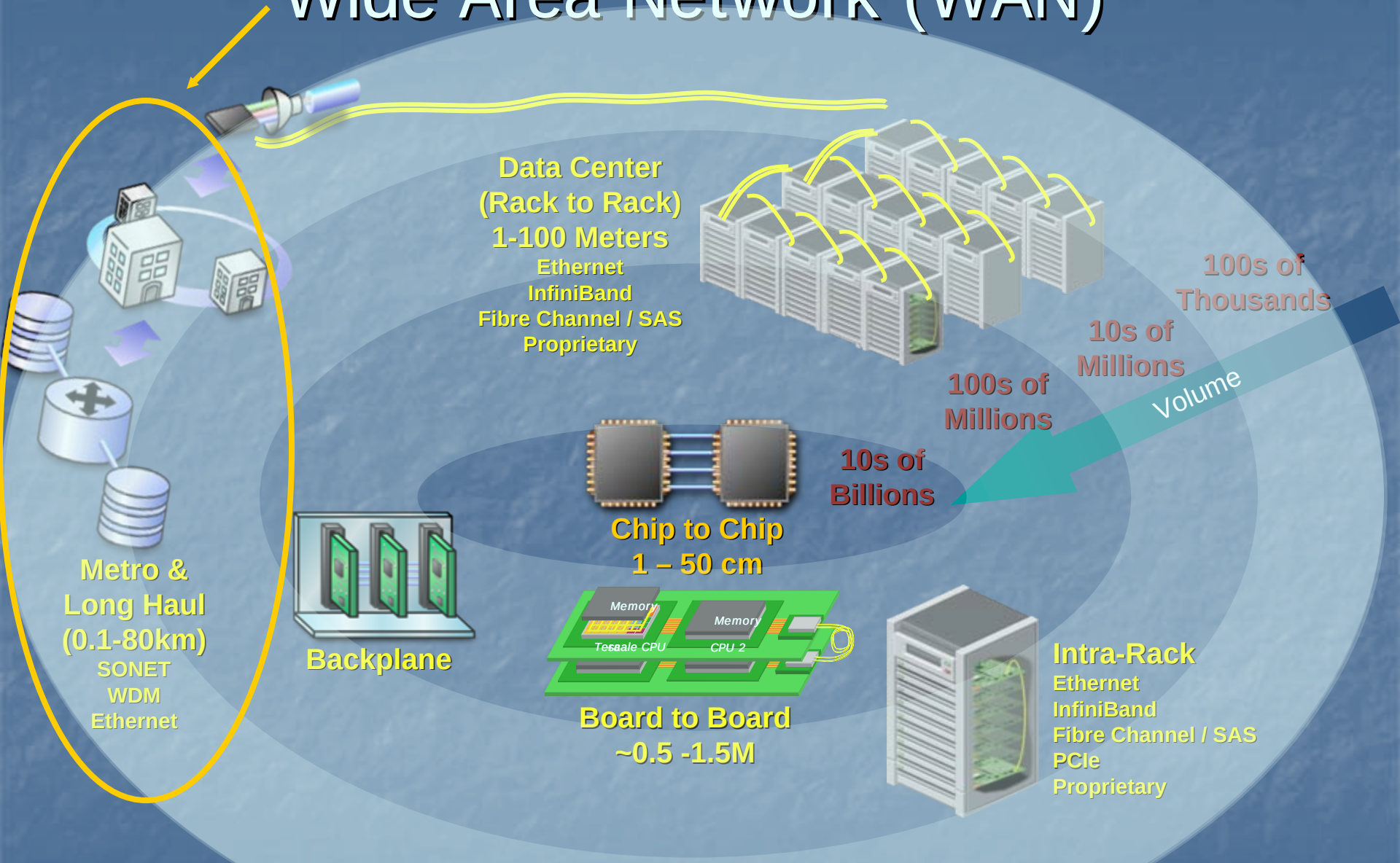
10s of Millions

100s of Millions

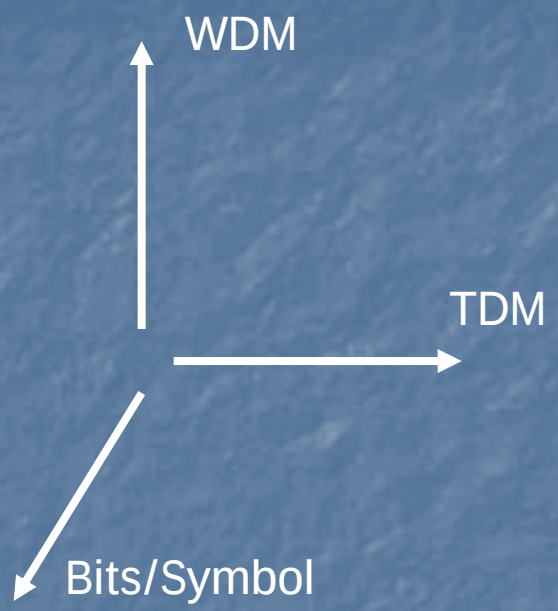
10s of Billions

Volume

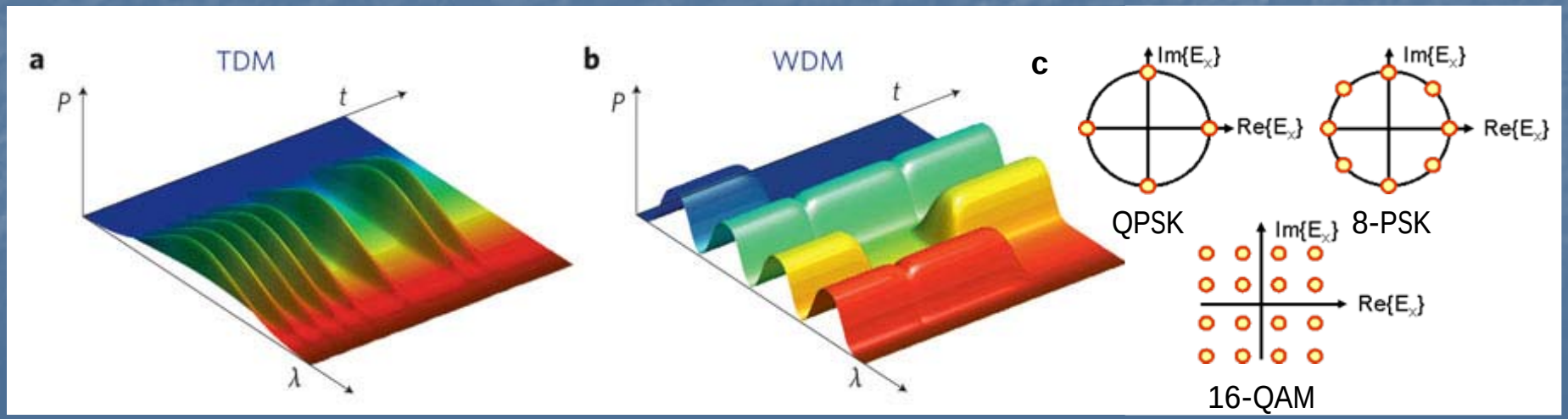
Wide Area Network (WAN)



Bandwidth and Capacity



- Bandwidth
  - #Symbols/s x bits/symbol x #λ
  - Only one fibre considered here
- Capacity = Reach x Bandwidth
  - Reach not an issue in typical detector applications (~100m)
  - Only bandwidth optimization considered here



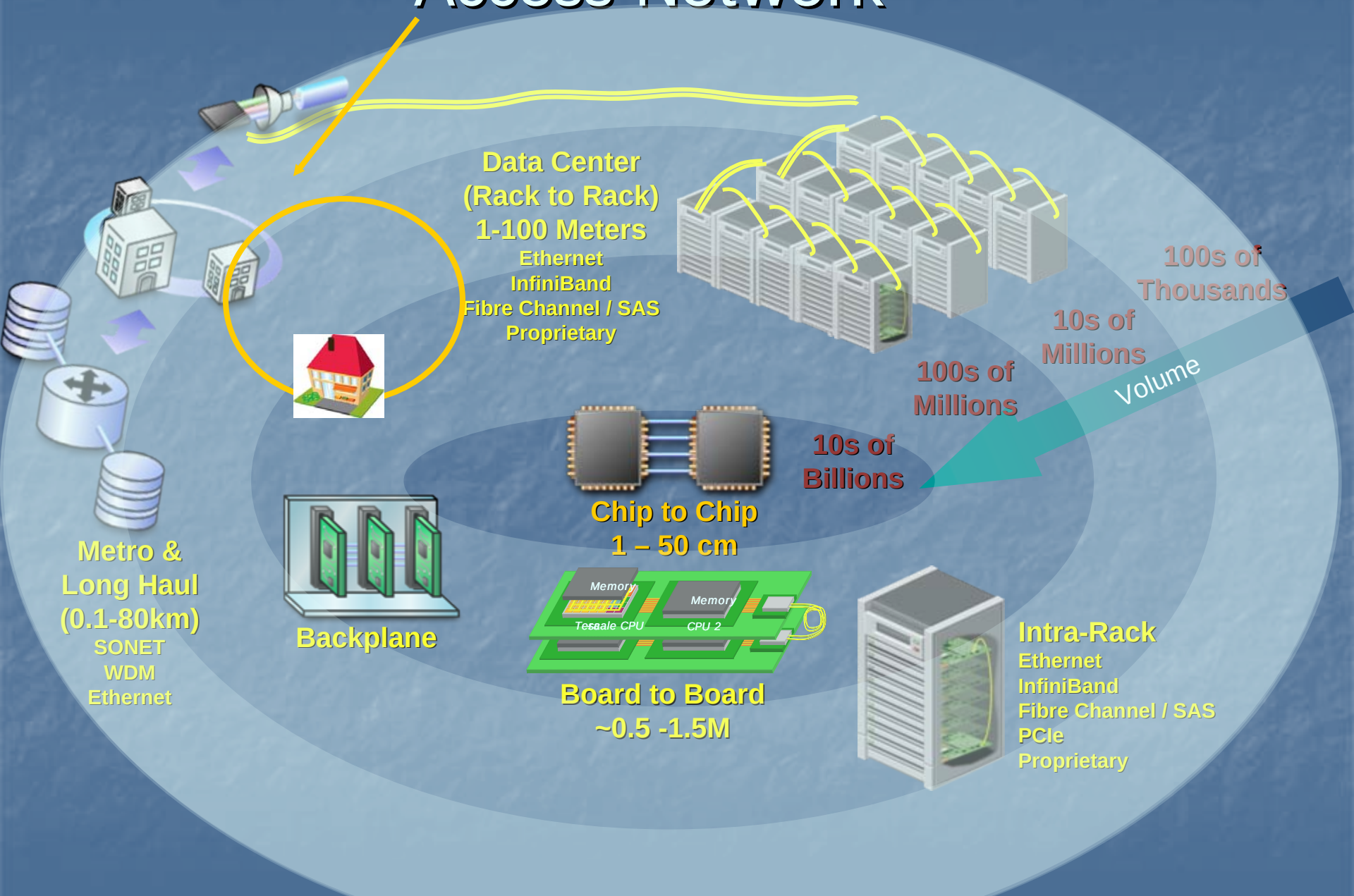
WAN Network features

- Ultimate Capacity
  - Design for minimum received energy
  - Exploit TDM, WDM, Advanced modulation formats
- Single Mode
- Electronic Mitigation Techniques
  - Ultra fast electronics operating at line rate (f.i. 50GS/s ADC)
  - FEC, equalizers, ...

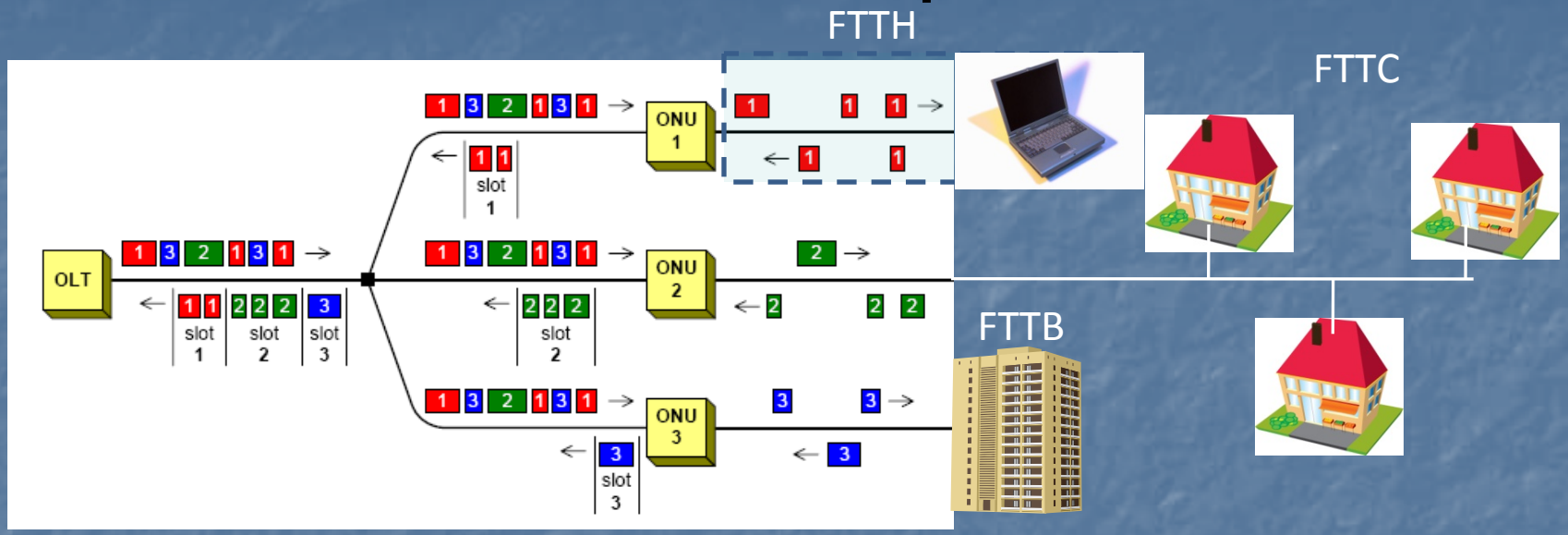
- Metrics (very coarse guess)
  - Bandwidth: 10Tbps
  - Reach: 100km
  - Energy: 10’000 pJ/bit
  - Density: 10’000 mm^2/Gbps
  - Cost: 100$/Gbps

Not in our ballpark, but watch out for emerging technologies

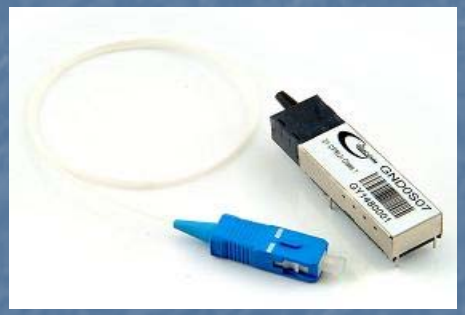
Access Network



Access over Passive Optical Network



- GPON, EPON 1.25Gbps - 2Gbps
  - Consumes 18x less energy per user than VDSL2
  - Sharing of this capacity among multiple users (time-division multiplexing)
  - Distance to central-office can go up to 60km when amplifier-assisted, not a few 100s meters from set-top box to DSLAM.



- 10GPON solutions recently proposed by system vendors
- WDM dimension can also be exploited to increase capacity,
  - And provide Point to Point connections, capacity on demand...while preserving scalability

PON Access Network features

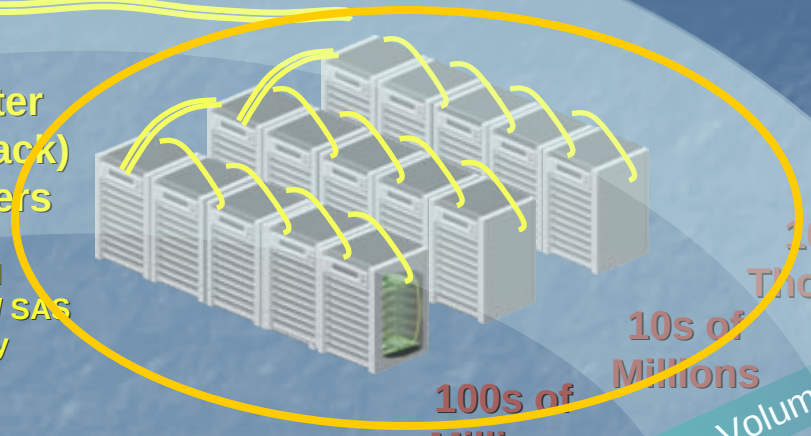
- Moderate speed, reach
- Scalability
- Shared resource
- High dynamic range
  - High split ratio
- High Volume, Low Cost
- Single Mode, bidirectional
- 10G PON coming up, WDM PON next

- Metrics (coarse guestimates)
  - Bandwidth: 2.5Gbps
  - Reach: 10km
  - Energy: 400pJ/bit
  - Density: 200mm^2/Gbps
  - Cost: 10$/Gbps

Not in our ballpark,
Traffic flows in wrong direction
But high volume and low cost model in right bandwidth range

Local Area Network (LAN)

Metro &
Long Haul
(0.1-80km)
SONET
WDM
Ethernet



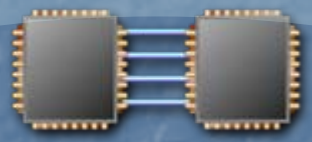
Data Center
(Rack to Rack)
1-100 Meters
Ethernet
InfiniBand
Fibre Channel / SAS
Proprietary

100s of Thousands

10s of Millions

100s of Millions

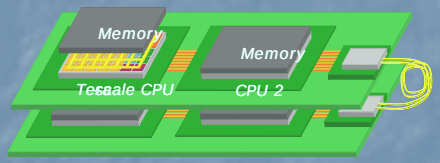
Volume



10s of Billions

Chip to Chip
1 – 50 cm

Backplane



Board to Board
~0.5 -1.5M

Intra-Rack
Ethernet
InfiniBand
Fibre Channel / SAS
PCIe
Proprietary

40/100GbE based on 10G PHYs

PHY Types to be added by 802.3ba

	40GbE 40GBASE–	100GbE 100GBASE–
1m backplane	4 x 10.3125G KR4 Optional FEC	
10m Cu Cable	4 x 10.3125G CR4 Parallel Coax Cable	*10 x 10.3125G CR10* *Parallel Coax Cable*
100m OM3 MMF	4 x 10.3125G SR4 0.8µm, Ribbon Fiber	*10 x 10.3125G SR10* *0.8µm, Ribbon Fiber*
10km SMF	4 x 10.3125G LR4 1.3µm CWDM	4 x 25.78125G LR4 1.3µm LAN-WDM
40km SMF		4 x 25.78125G ER4 1.3µm LAN-WDM

Nomenclature 40G BASE – L R 4

MAC Data Rate | Media Type | PCS Type | PMD Lane

10-Lane PHY

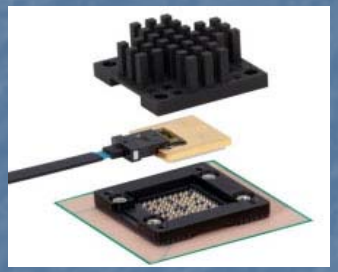
4-Lane PHY

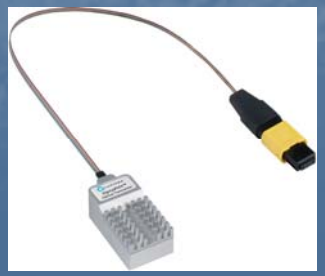
- 10G: directly modulated EE lasers or VCSELs
- 25G: externally modulated lasers
- 40G-100G: aggregation

LAN Network Features

- converging with MAN
- TDM limits reached
- Aggregation via parallelism
  - Physical
  - Wavelength
  - Inheriting WAN features
- High Density
- Low Power
- High Volume Low Cost
- Multimode OM3 & OM4 fibre
  - 10G over 100m & 300m
  - Single Mode next

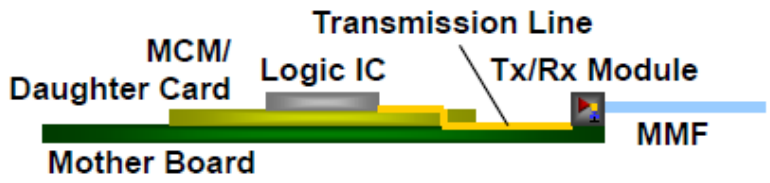
- Metrics (guestimates)
  - Bandwidth: 100Gbps
  - Reach: 100m
  - Energy: 10pJ/bit
  - Density: 2mm^2/Gbps
  - Cost: 1$/Gbps

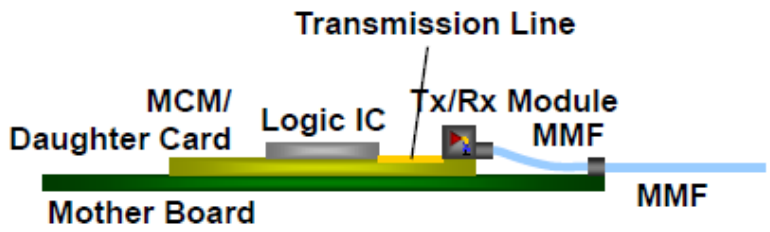




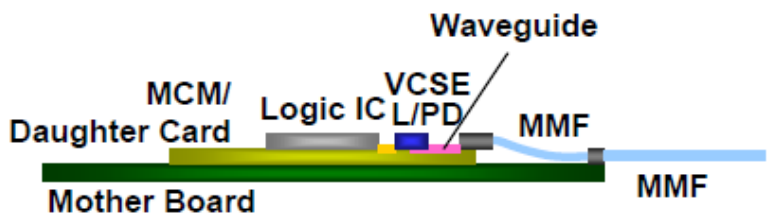
From modules to chips: packaging issues

- **Module-Based**

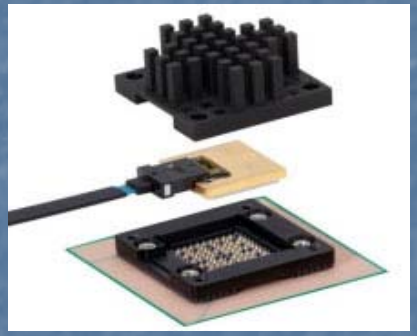


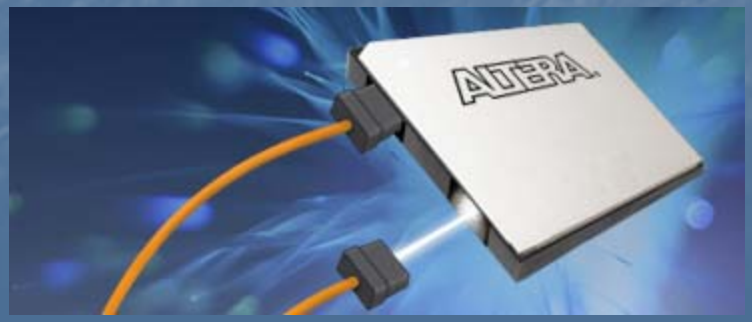


- **Chip-Level Packaging**



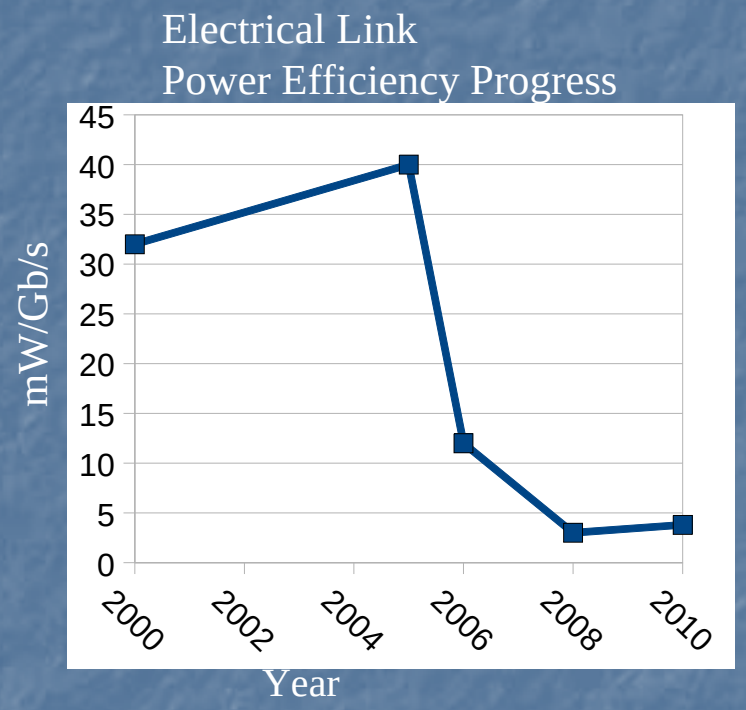
Courtesy Shigeru Nakagawa, IBM





Electrical Interconnects are still alive

- For short distances, EOE cannot beat EE
- Short range low power electrical links are the subject of intense development efforts
  - Moving target



Link Type	Power Efficiency (pJ/bit)	Distance
Electrical	3	< 1 m
EOE	7	< 100 m

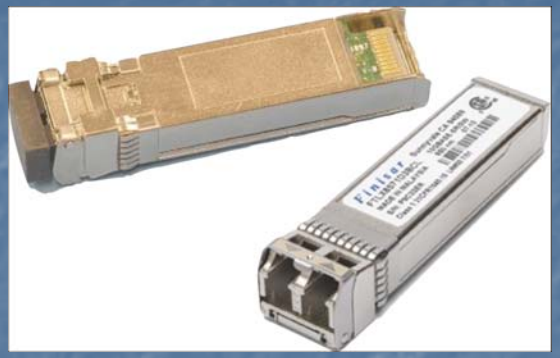
Optical network metrics
A risky comparison* (1)

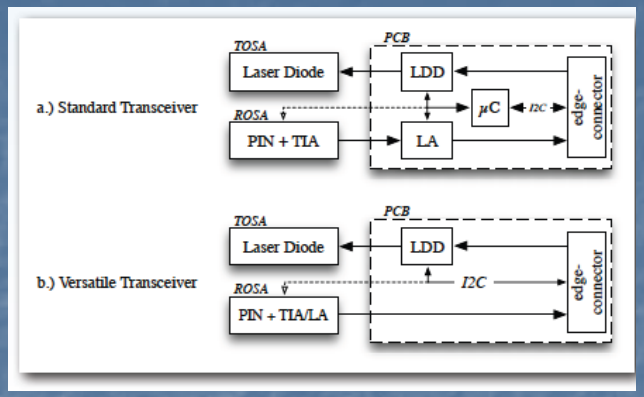
*This comparison expresses the views of the author at presentation time and do not necessarily represent an accurate vision of the current and future situation

Arrays

	AOH 1997	WAN 1990	Access 2000	LAN 2010
Bandwidth Gbps	0.32	10’000	2.5	100
Reach m	100	100’000	10’000	100
Energy pJ/bit	150	10’000	400	10
Density mm^2/Gbps	1000	10’000	200	2
Cost $/Gbps	900	100’s	10	1

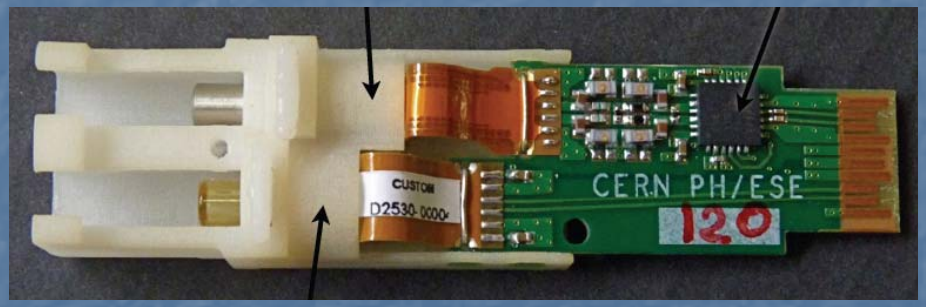
Next for HEP: Front End Versatile-TRx for HL-LHC





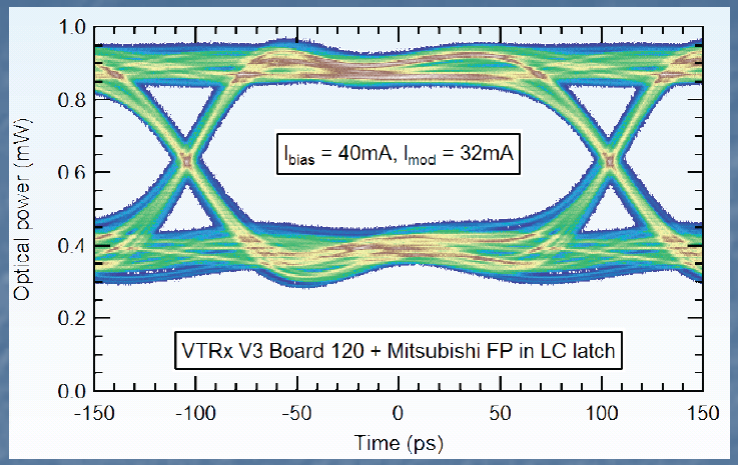
VTRx

- Low Mass & Volume
  - Minimize material, avoid metals
- Capable of operating in a magnetic field
  - Requires replacement of ferrite bead used in laser bias network
- Radiation Tolerant
- Bitrate determined by ASICs: 5 – 10 Gbps
- 850nm and 1310nm flavours

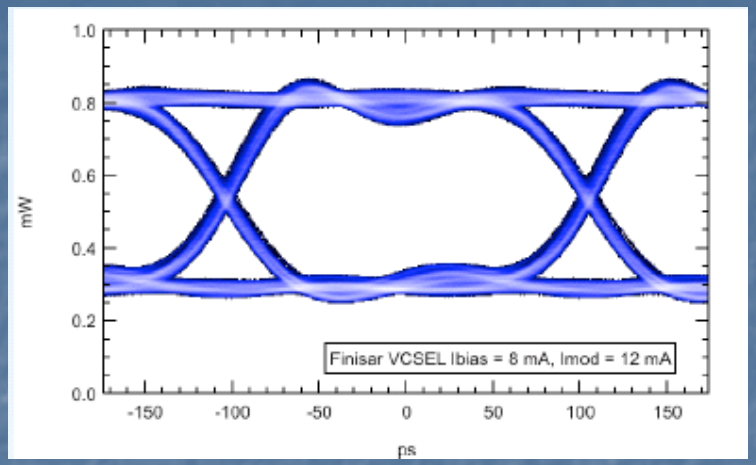


Front End Versatile-TRx Results at 5Gbps

- Tx

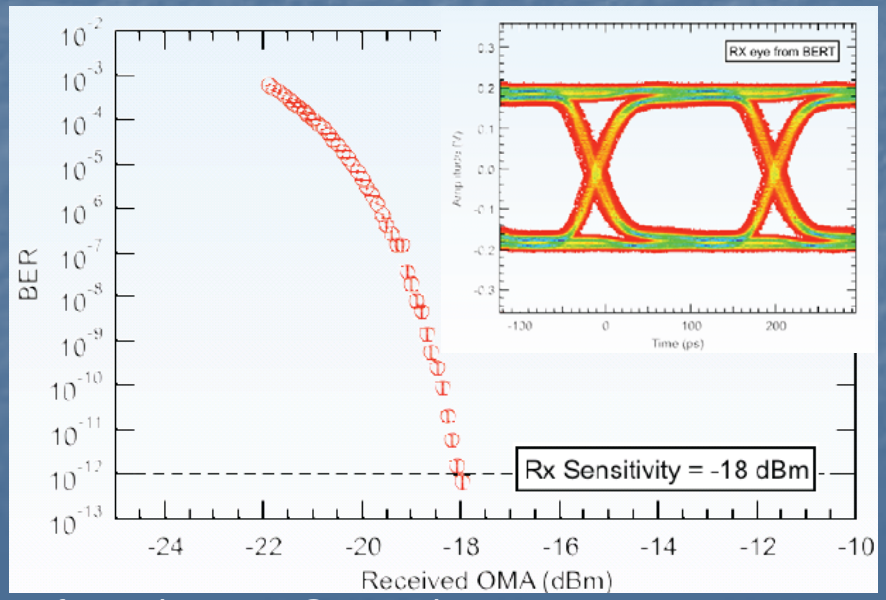


SM@1310nm



MM@850nm

- Rx



Optical network metrics
A risky comparison* (2)

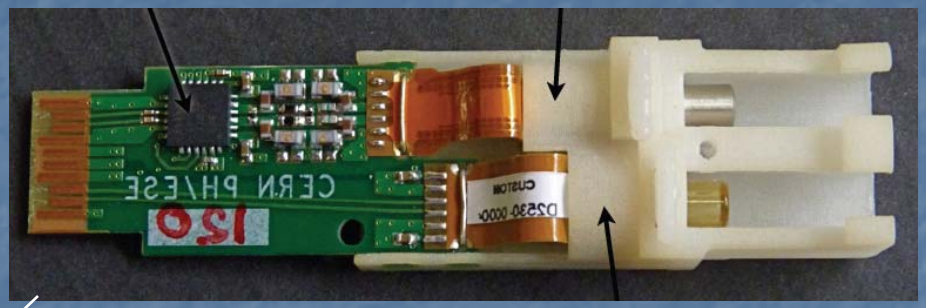
*This comparison expresses the views of the author at presentation time and do not necessarily represent an accurate vision of the current and future situation

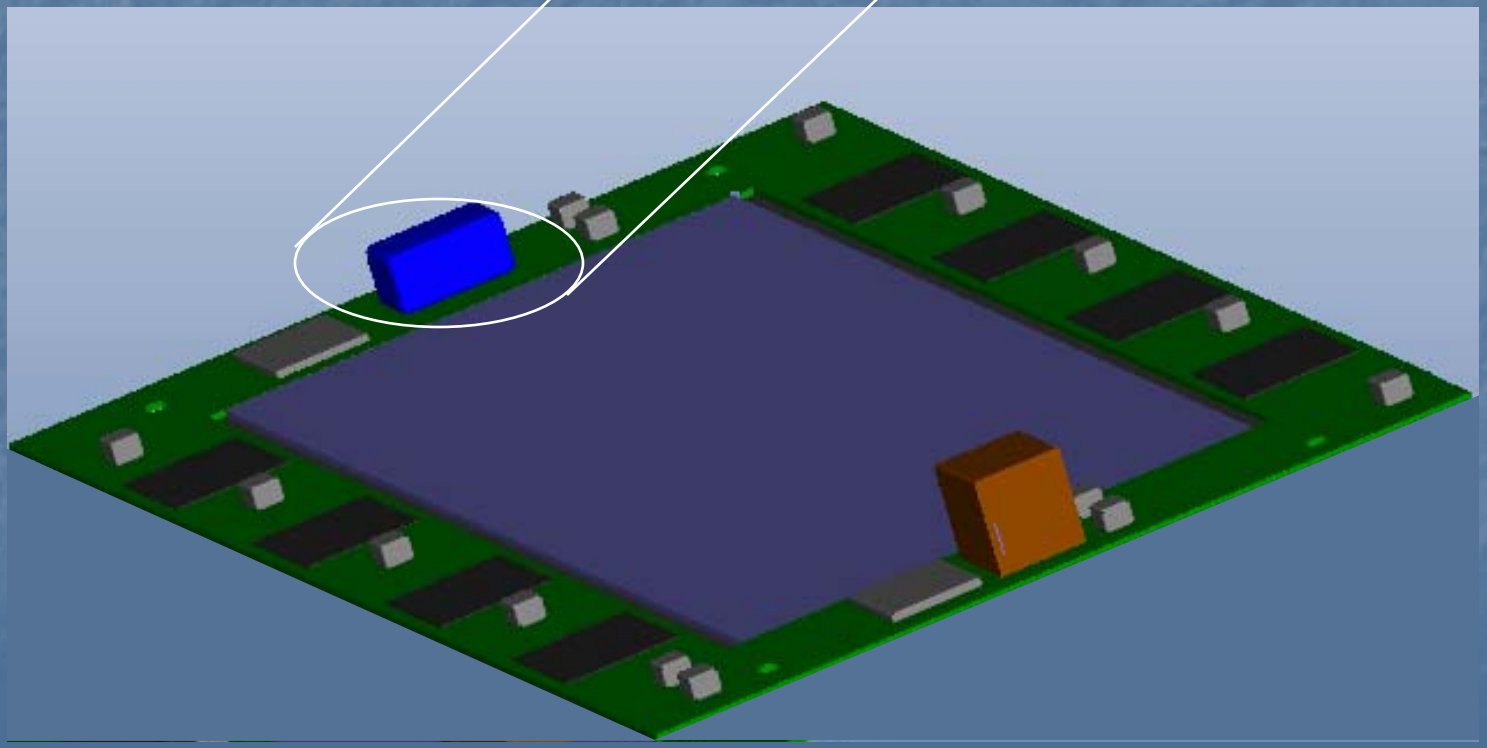
Arrays

	AOH 1997	WAN 1990	Access 2000	LAN R2R 2010	VTRx 2008
Bandwidth Gbps	0.32	10’000	2.5	100	5
Reach m	100	100’000	10’000	100	100
Energy pJ/bit	150	10’000	400	10	100
Density mm^2/Gbps	1000	10’000	200	2	100
Cost $/Gbps	900	100’s	10	1	60 ?

A TRx for an all-in-one module

- 50% scaling required
  - Size
  - power





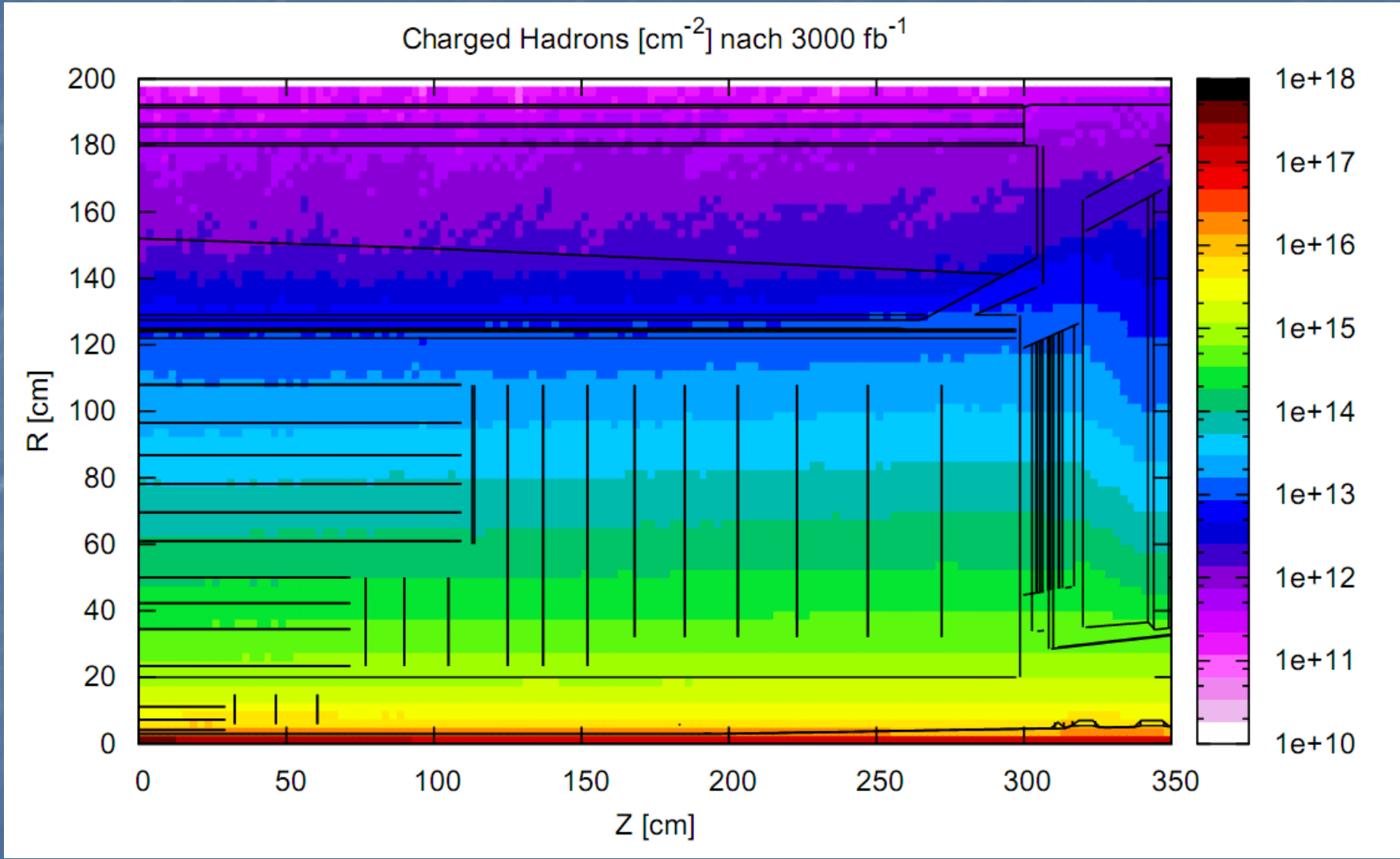
Optical network metrics
A risky comparison* (3)

*This comparison expresses the views of the author at presentation time and do not necessarily represent an accurate vision of the current and future situation

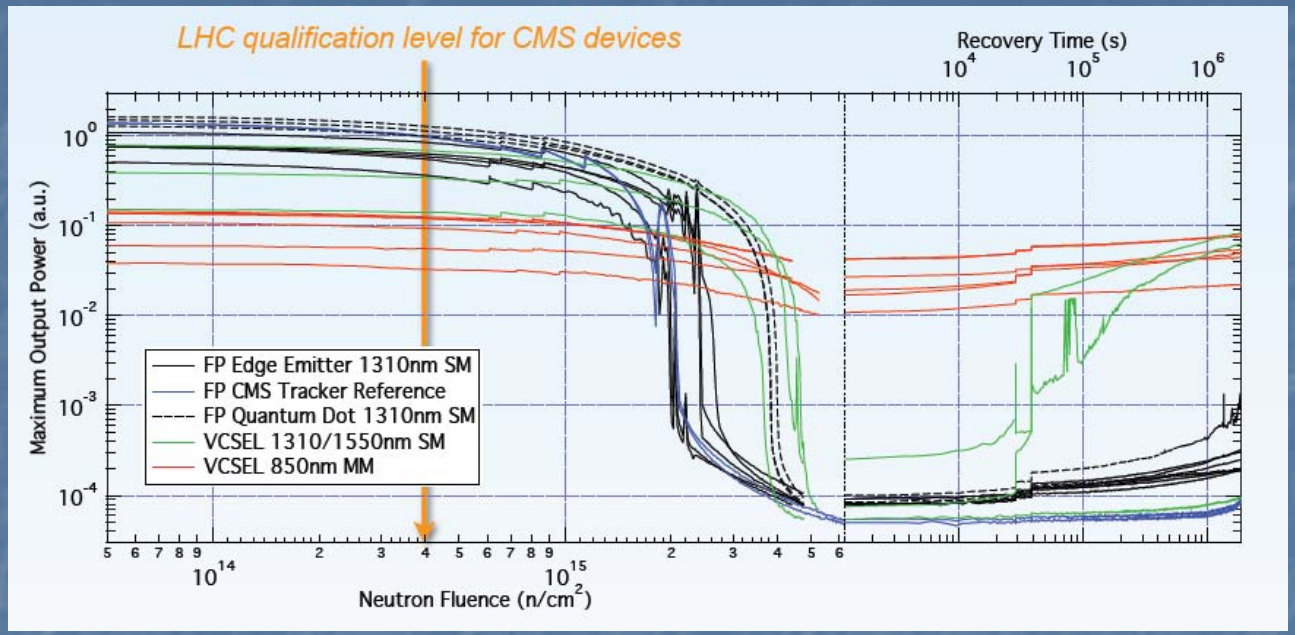
Arrays

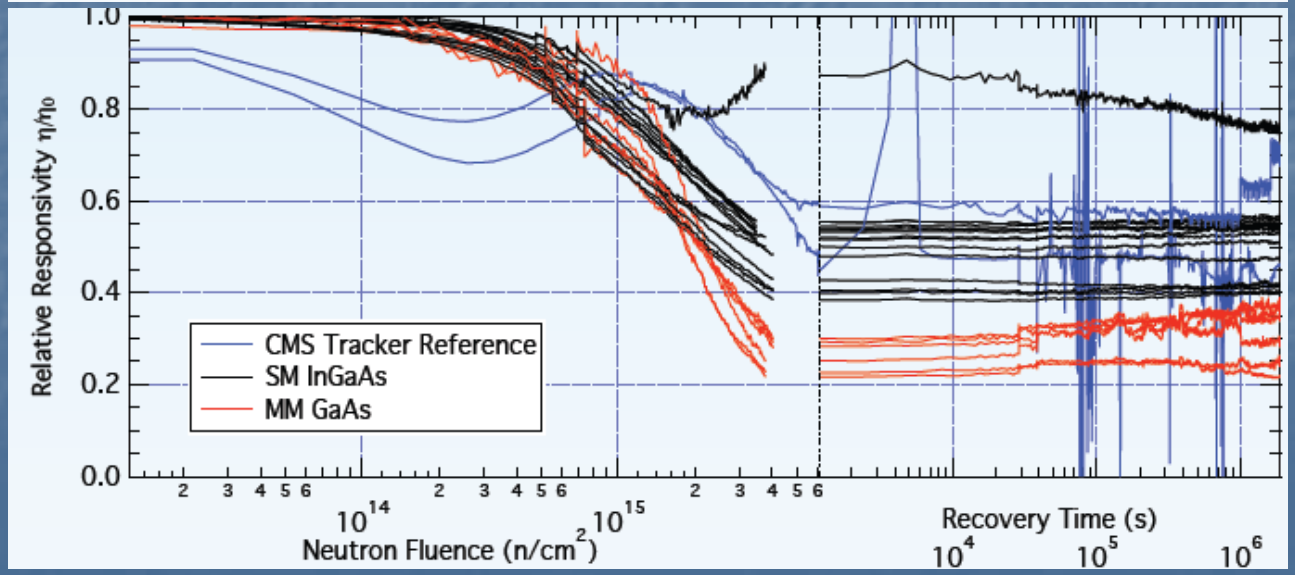
	AOH 1997	WAN 1990	Access 2000	LAN R2R 2010	VTRx 2008 →12
Bandwidth Gbps	0.32	10’000	2.5	100	5
Reach m	100	100’000	10’000	100	100
Energy pJ/bit	150	10’000	400	10	100 → 50
Density mm^2/Gbps	1000	10’000	200	2	100 → 50
Cost $/Gbps	900	100’s	10	1	

Let's not forget radiation tolerance



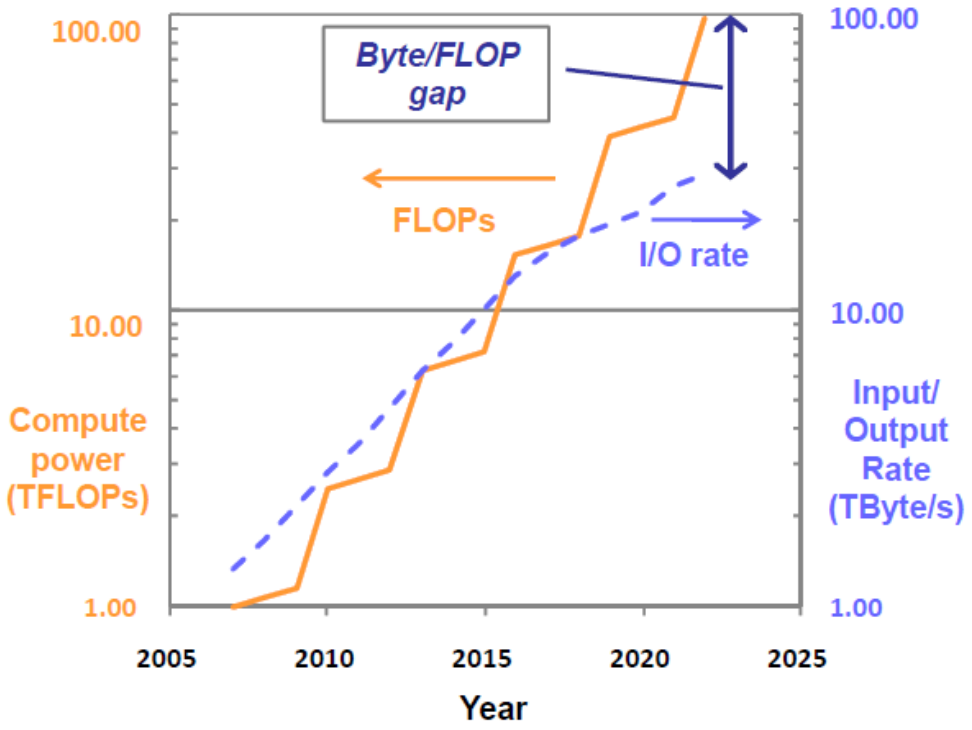
Radiation Tolerance: Is there a limit?

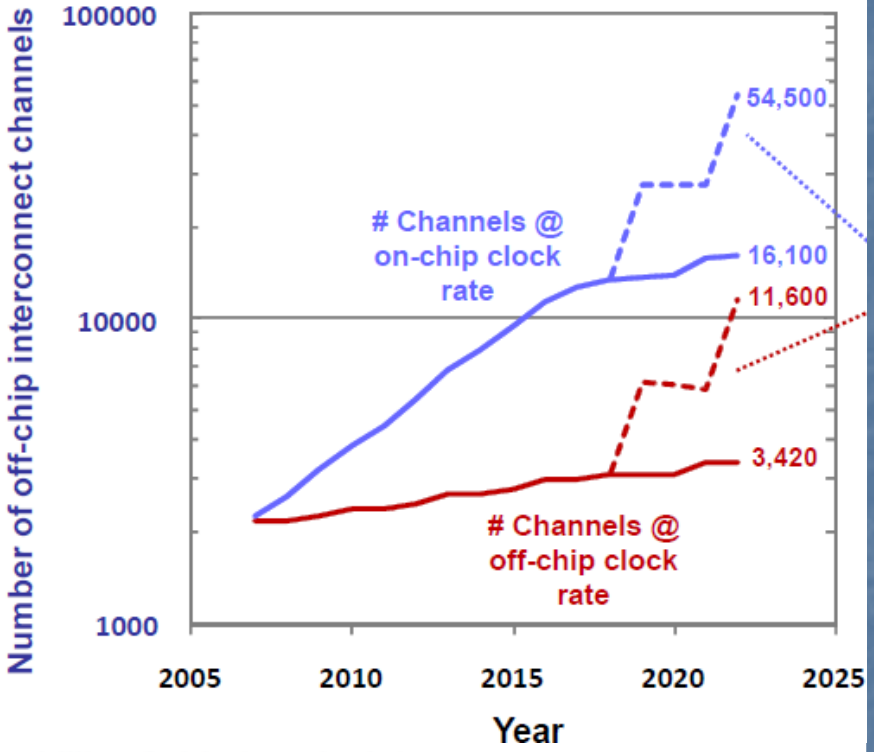




- Active opto devices tested up to now do not resist fluences beyond 10^{16} cm^{-2}
  - Still OK for SLHC Tk
  - Tighter margins
  - Are there alternatives for fluences beyond 10^{16} cm^{-2} ?
- (Re)-consider Passives ?
  - modulators

Extrapolating I/O Bandwidth Requirements





Input/Output interconnect (I/O) rate does not keep up with ability of chip to calculate

Ideal of 1 Byte of memory access for each floating point operation (FLOP) cannot be retained

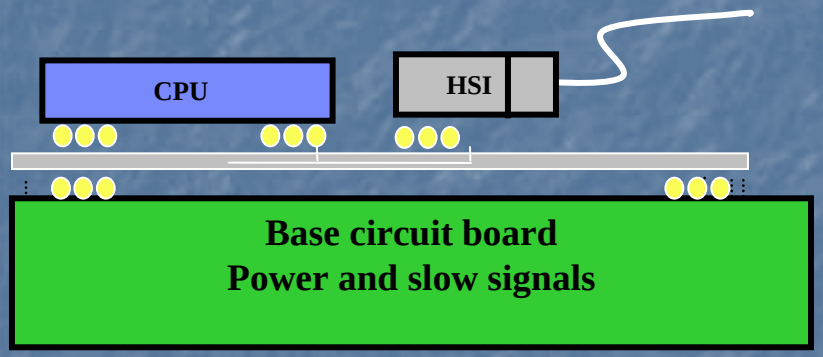
- **Byte/FLOP gap**

Need to go optical, on chip, with WDM

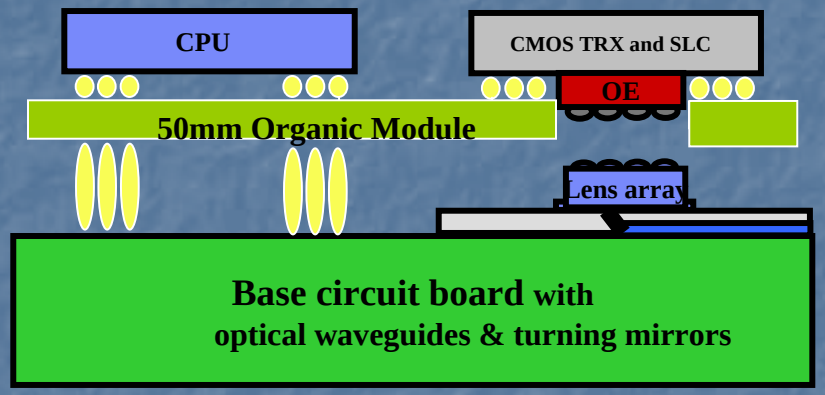
- **Si – CMOS Photonics**

Evolution of high speed interconnects 2010-2015

Electrical or Optical engines + Opto PCB MM
PIC TRx

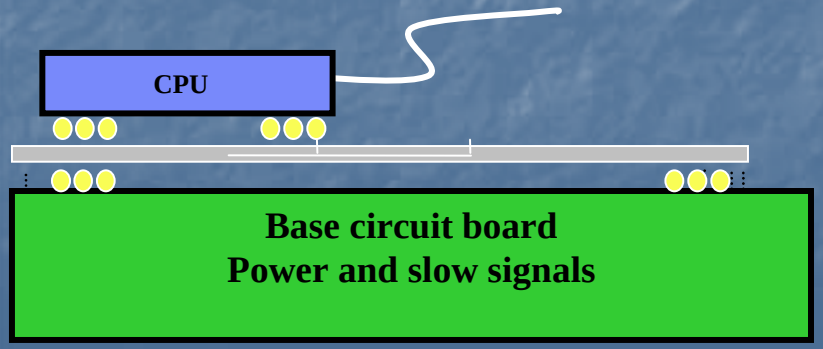


OR



2015-2020

Monolithic integration of TRx and Processing chip SM



Will this paradigm shift occur ?
Will it bring us benefits
(density, power, radition tolerance) ?

Conclusions

- Broadband optical communication is a commodity and is ubiquitous
  - Also in HEP
- Scaling bandwidth and connectivity remains a major issue at all levels of the network hierarchy
  - HEP has a distributed bandwidth need. It cannot simply aggregate.
- Current optical communication technologies will not allow to meet the challenges ahead without major breakthroughs:
  - Advanced modulation, coding and electronic mitigation schemes
  - High density parallel and wavelength multiplexed optics (optical engines)
  - Photonic integration with CMOS
- Pragmatically speaking, all incremental electrical chip to chip interconnection approaches will be exhausted before fast interconnects go optical
  - Will this kill photonics integration?
- In the short/medium term: laser-based EOE conversion such as proposed by the Versatile Link common project is the way forward
  - 10Gbps, 10's pJ/bit, 10's mm^2/Gbps
- In the long term:
  - Worth watching photonics integration
  - Embracing a new technology is resource intensive