http://www.oracle.com/us/products/networking/overview/index.html

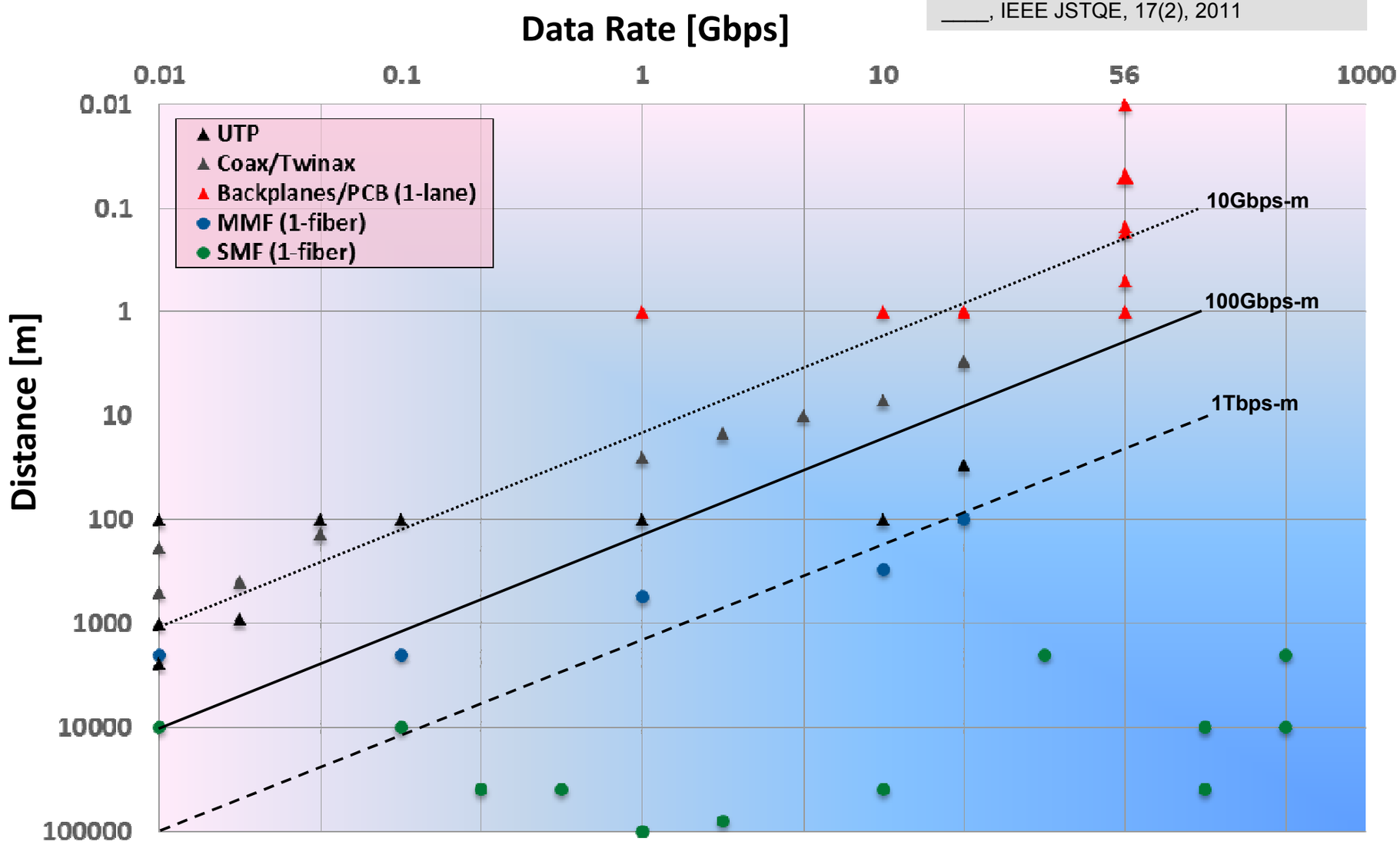
Optical Interconnect Packaging for the Cloud

IEEE ECTC 2016

A. V. Krishnamoorthy
Architect and Chief Technologist, Photonics
Oracle Networking

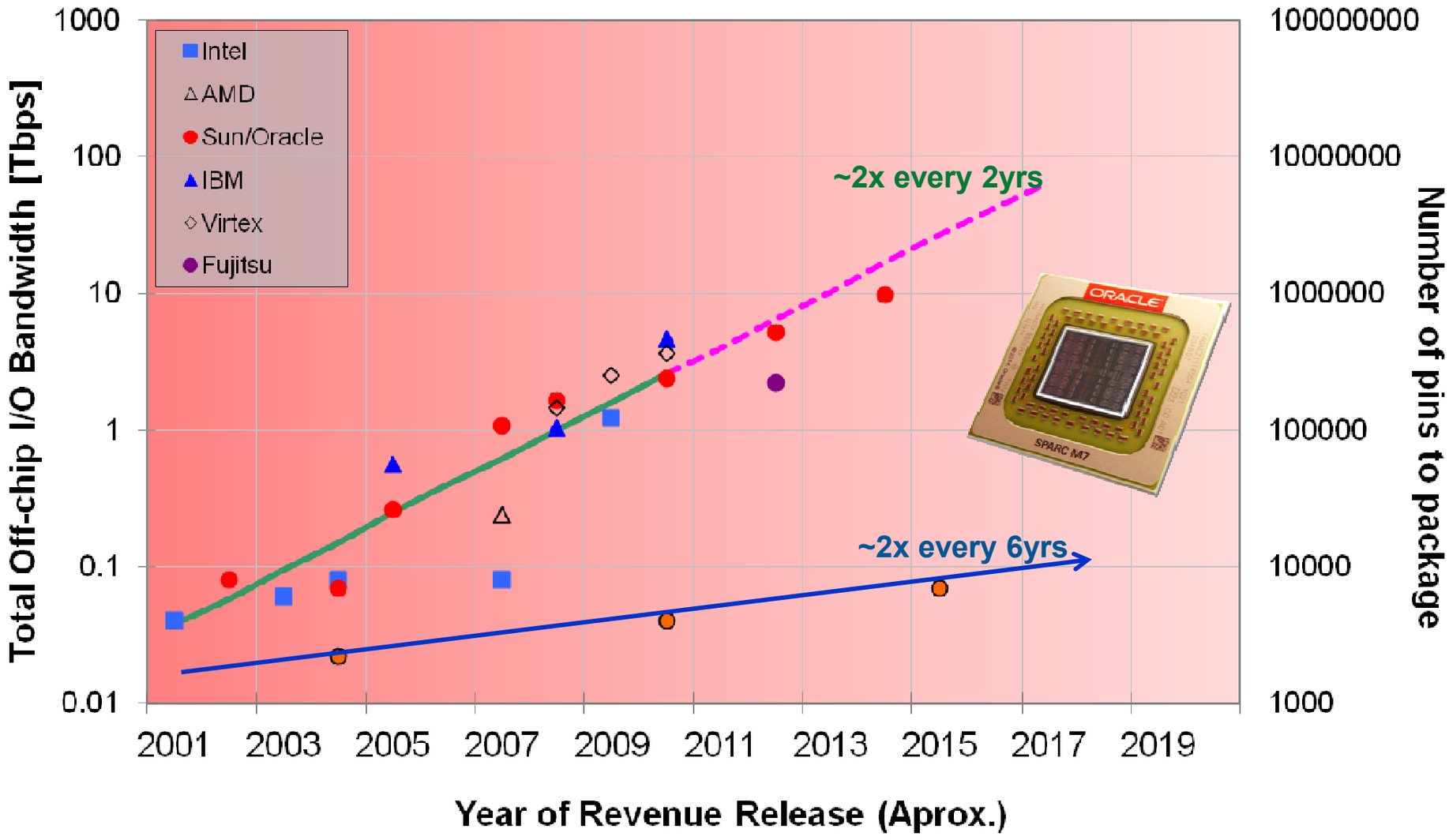
Optical Link Technology Penetration

____, IEEE JSTQE, 17(2), 2011



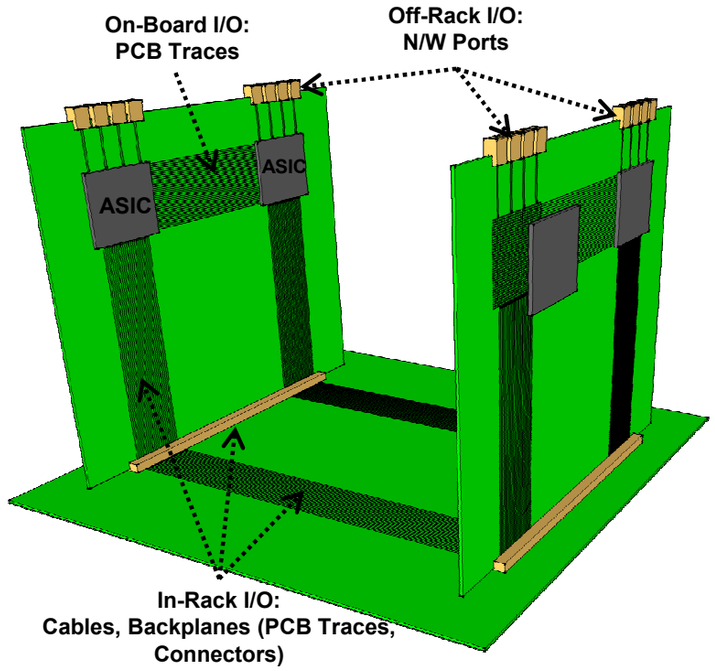
Processor IO Requirements

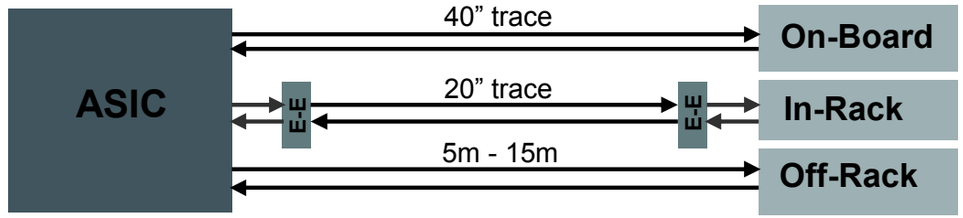
____, Proc. IEEE Optical Interconnects Conf., 2016



System Interconnect Evolution --- The Past

Mostly Electrical





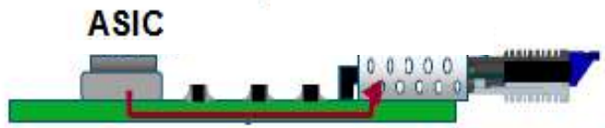
- 2.5Gbps-15Gbps/Lane
- LR SerDes
  - Re-timers required at every hop
- BW*Density Limited
  - Both ASIC and front panel
- BW*Distance Limited

Network System Realization --- Magnum M9

O. Torudbakken & A. Krishnamoorthy, OFC OTu3H.1, 2013

2x 12x10Gbps AOC

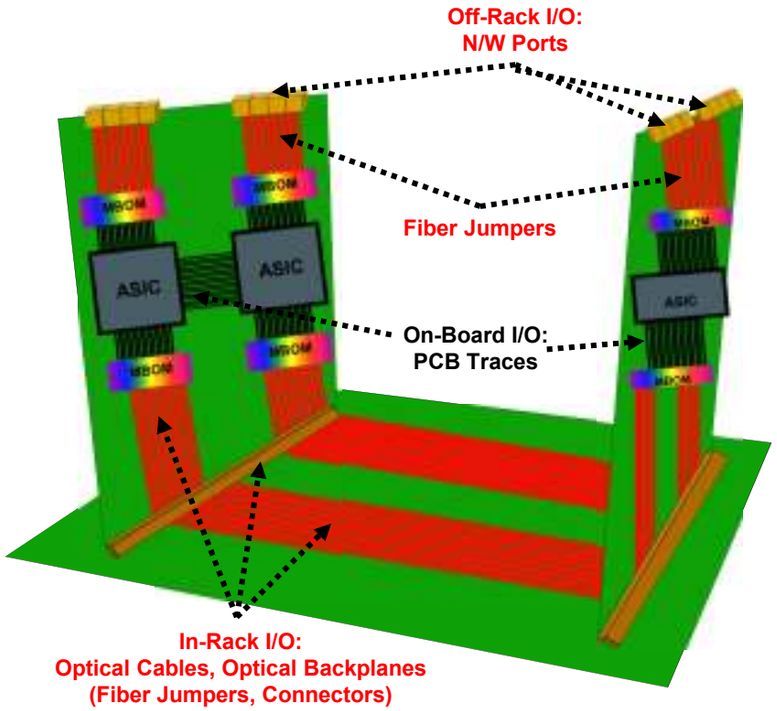
- 648 external network ports
  - 26Tb/s with AOC
  - 11RU standard chassis
- 6.5KW power consumption

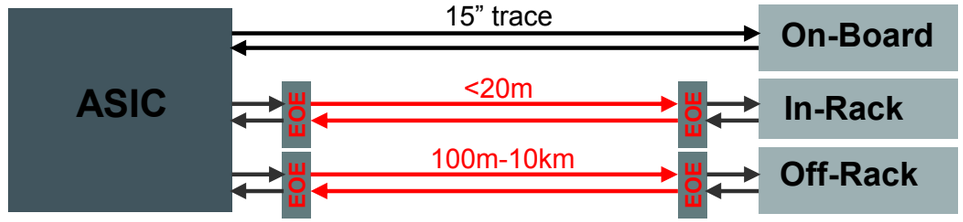


EO conversion (if any) at board's edge or in the cable (AOC)

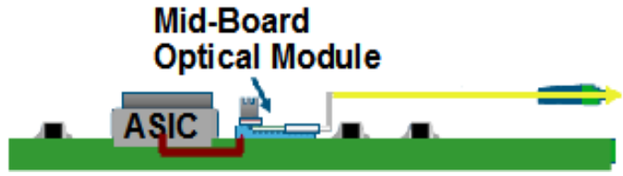
System Interconnect Evolution --- The Present

Mostly Optical





- 25+Gbps/Lane
- SR SerDes
  - Re-timers optional
- Optics replace the lossiest Cu interconnects
- Eliminates front panel bandwidth-density limit
  - At least 8x improvement
- ASIC still bandwidth-density limited
- Improved signaling across entire system

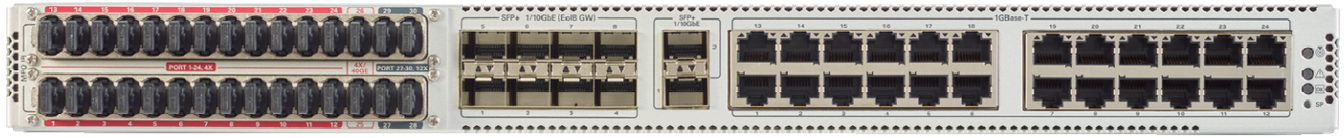


EO conversion next to ASIC package

Network System Realization --- The Present

Nano-Magnum, 2016

____, Proc. IEEE Optical Interconnects Conf., 2016



Leaf: >4Tb/s; <0.3KW

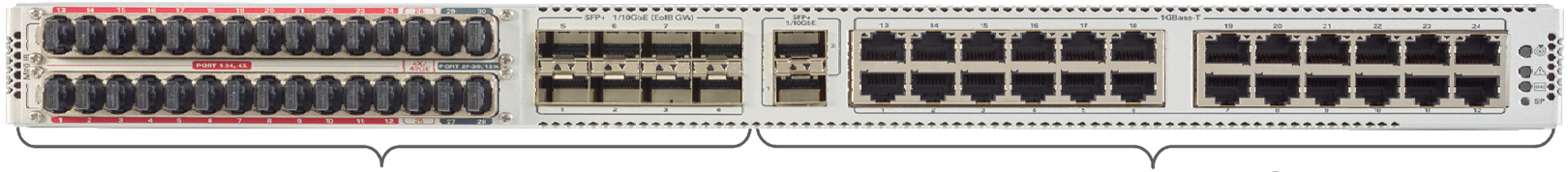
Spine: >24Tb/s; <3KW

Virtualized I/O & SDN

Oracle InfiniBand Switch IS2-46

Leaf Switch

____, Proc. IEEE Optical Interconnects Conf., 2016

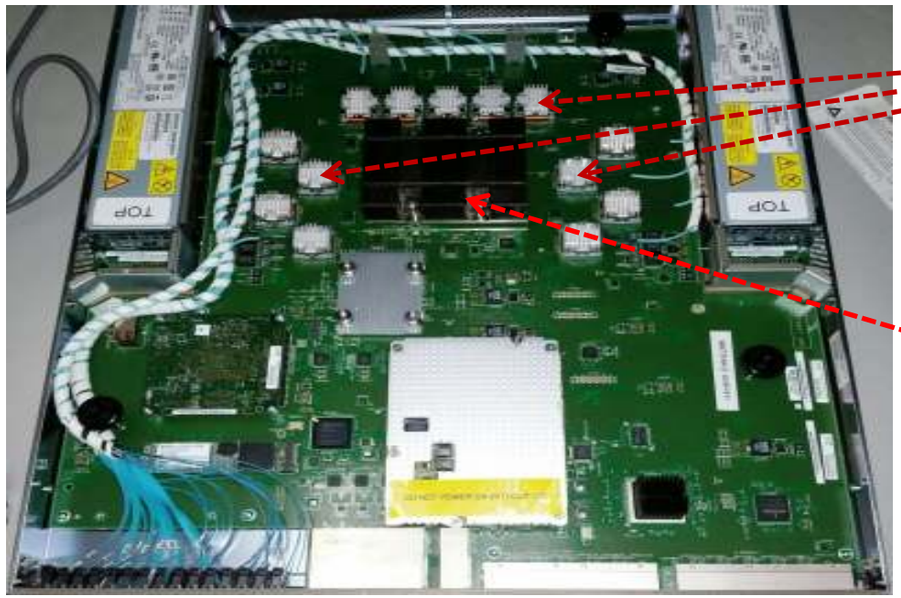


Data Plane

- IB Switch & Router
  - 4Tbps bisectional b/w
  - 36 fixed + 2 optional IB EDR ports (MPO)
- Ethernet Gateway
  - 8 SFP+ ports for 1-GE/10-GE connectivity
  - 2 optional MPO ports for 40-GE connectivity

Ethernet Management Switch

- Ethernet connectivity for server & storage management
  - 24 RJ-45 ports
  - 2 SFP+ ports



MBOM

Switch ASIC

Oracle InfiniBand Switch IS2-254

Spine Switch

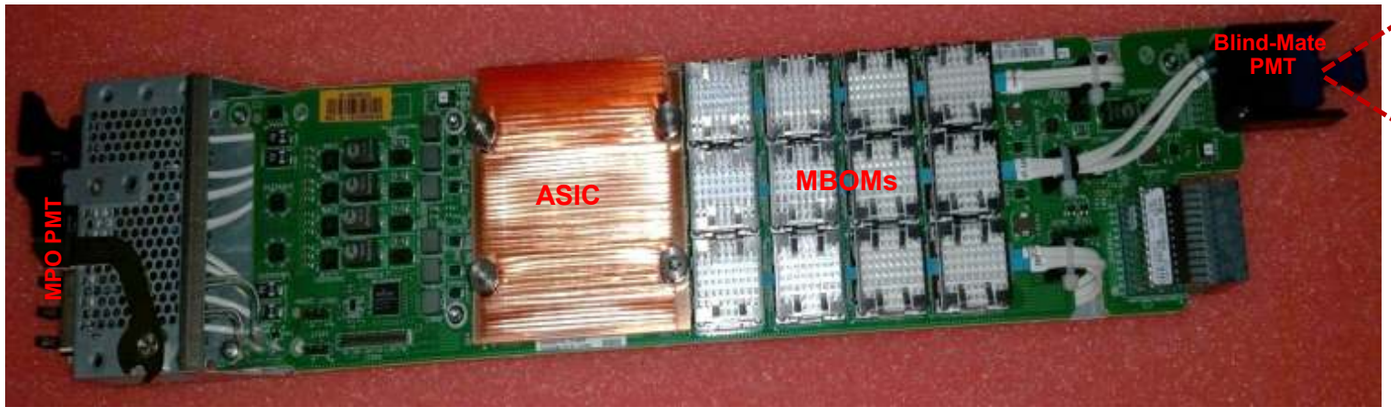
____, Proc. IEEE Optical Interconnects Conf., 2016

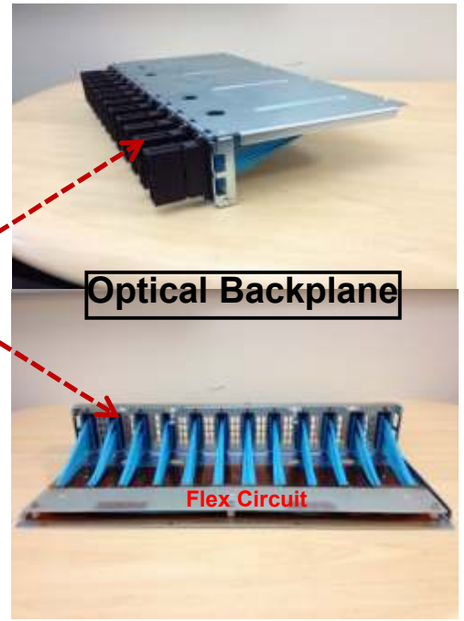
System Interfaces

- IB Switch & Router Uplinks
  - 12 EDR ports (4 x 12x)
- Ethernet Gateways
  - 2 MPO ports for 40-GE connectivity
- Ethernet Management
  - 4 RJ45 ports for 1-GE connectivity

Module Slots

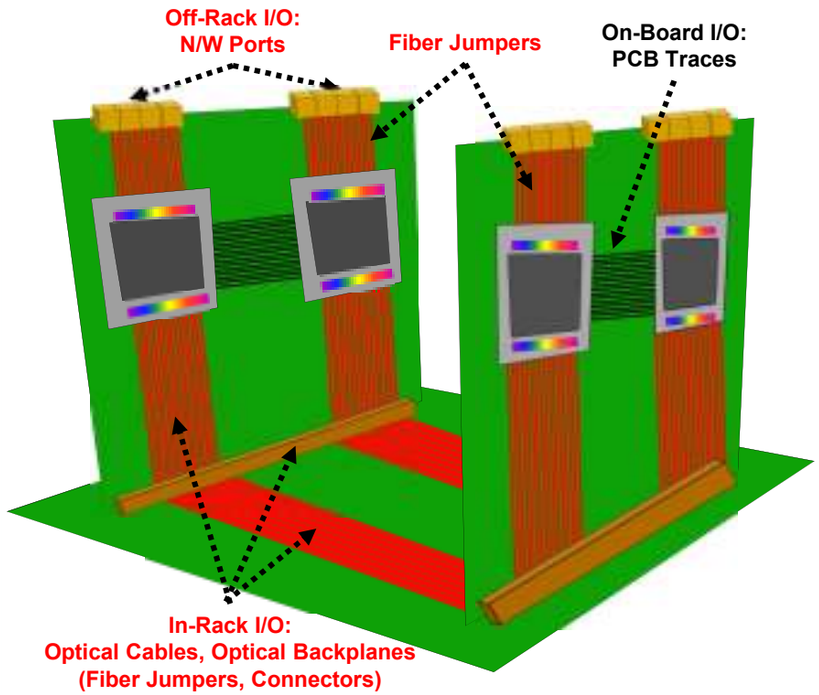
- Switch Modules
  - Line and Fabric; 24Tbps switching capacity;
- I/O Modules
  - 4x10GBase-T, 4x40-GE, 16x10-GE
  - 2x16G FC; 4xIB-EDR Extended Reach (40km)
- Network Services Modules

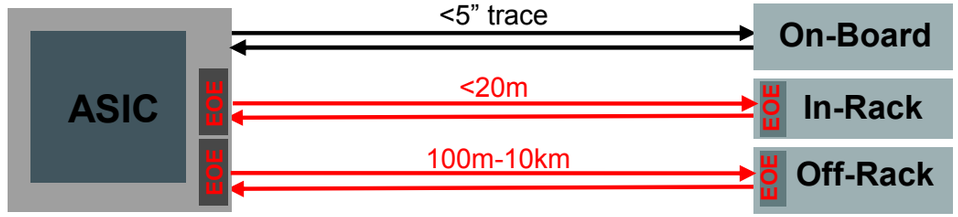




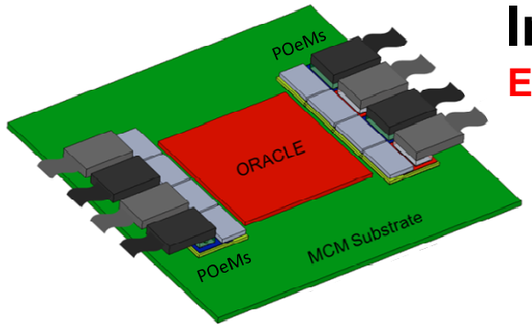
System Interconnect Evolution --- The Future

The Last 100mm





- 50+Gbps/Lane
- Integration of the high-speed interfaces at the package substrate (MCM)
  - Package Opto-Electronic Module (POEM)
- Lightweight (XSR/USR) SerDes
  - Optimal energy efficiency
  - No Re-timers
- Addresses bw*density limitations at the ASIC



Intra-Package Optics

EO conversion next to ASIC die

Power Reduction Potential: Another 3.5x?

R&D prototype 16 port switch, 20Gbps/port, GbE signaling:
<20pJ/switched bit with VCSELS and CMOS switch flip-chip co-packaged

____, IEEE JSTQE, 17(2), 2011

Take-aways

- Bandwidth-Distance (“Scale-out”) and Bandwidth-Density (“Scale-up”) requirements are driving adoption of optical technologies
- For scale-out, growth in bandwidth-distance products continue to drive optics adoption
  - 100Gbps-m for multi-mode; 1Tbps-m (or 1Gbps-km) for single-mode
- For scale-up, compute trends continue to drive single-lane data rates
  - Doubles every 3-4 years to balance bandwidth requirements to pin limitations
- The electrical channel continues to be a major contributor to silicon complexity and to the overall system power consumption
  - Bringing the optics closer to the silicon enables performance scale-up and efficiency in next generation systems
- Oracle has announced a family of performance-leading all-optically-interconnected switching platforms
  - 49.1Tbps bandwidth, 450ns latency, 2U chassis, 16X improvement in Size*Power