OCP SUMMIT

March 20-21
2018
San Jose, CA

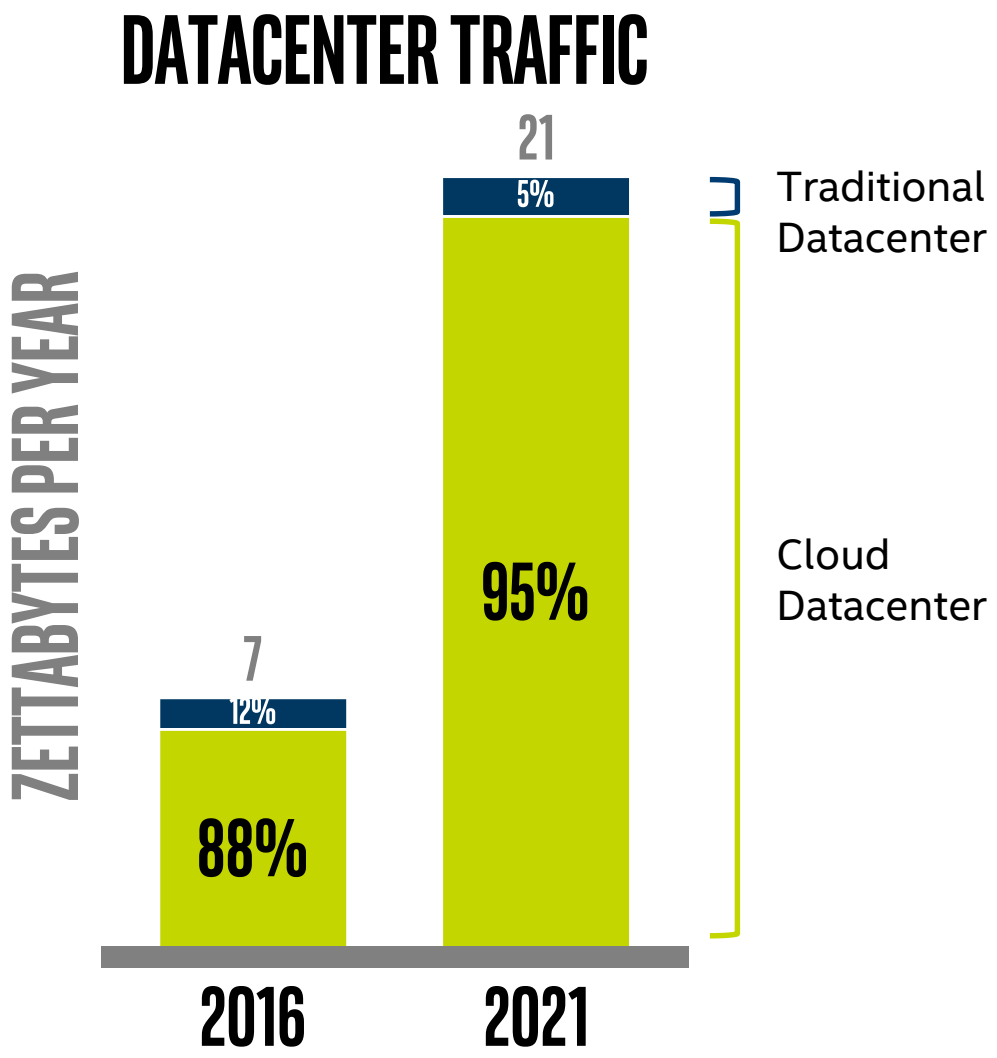
OPEN. FOR BUSINESS.

NEW WORKLOADS AND THE EVOLVING NETWORK

Uri Cummings
Connectivity Group Chief Technical Officer, Data Center Group
Intel Corporation

uri.cummings@intel.com

By 2021,
95% of all data center traffic will be based in the cloud.

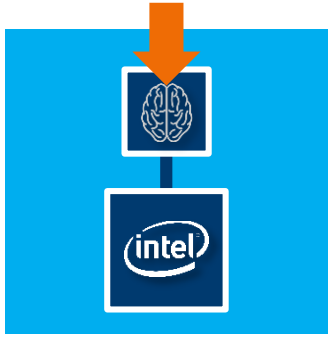


Source:: Cisco Global Cloud Index, 2016-2021

Data center network inflections

Critical workloads redefine the network

Cloud Scale: the big get bigger



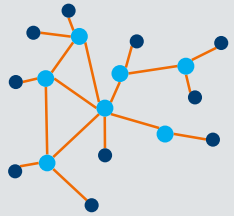
Pervasive offloads Distributed intelligence

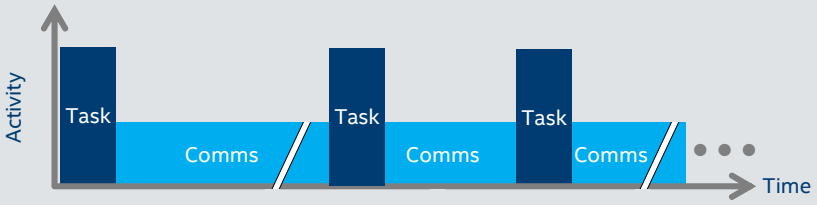
Emerging workloads redefining the network

Networks reconfigured for east-west, jitter reduction, and accelerators

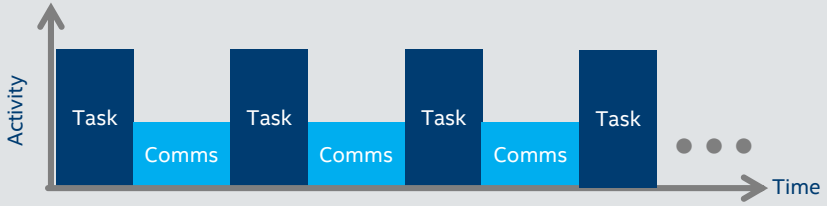
TRADITIONAL NETWORKS AND DATACENTER FABRICS

Network: Flexibly-connected topology that enables endpoints to share data (e.g., arbitrary non-cyclic topologies)





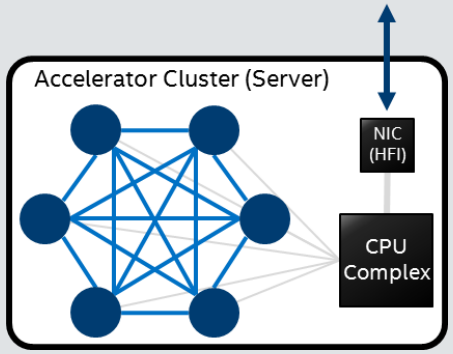
Scale-Out Fabric: Specialized form of a network, with a fully-connected topology, that is engineered and tuned to sustain performance & latency at high scale while reducing jitter (e.g., Cloud Ethernet and HPC Fabric)



DATA AND THE EMERGENCE OF ACCELERATOR NETWORKS

Node-focused:

- Light-weight network that interconnects a cluster of local accelerators with high-bandwidth, low-latency links
- Lowest power/bit
- Code & connectivity optimized for locality
- Emerging Need: Data-lakes



Accelerator scaling fabric:

- Used for high-bandwidth memory sharing/coordination at larger scale
- Allows separation of accelerator vs. general communications, minimizing interference in the scaling fabric.

The motivation driving acceleration and offload

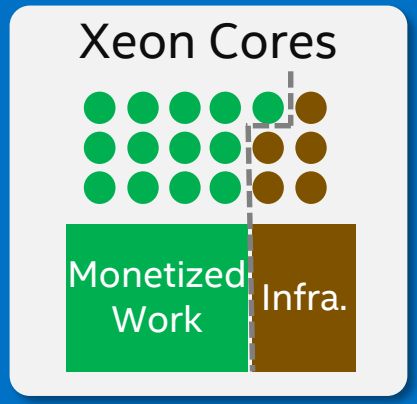
Agility

Continuous Integration

Cloud optimizes at the speed of Software

Infrastructure Acceleration

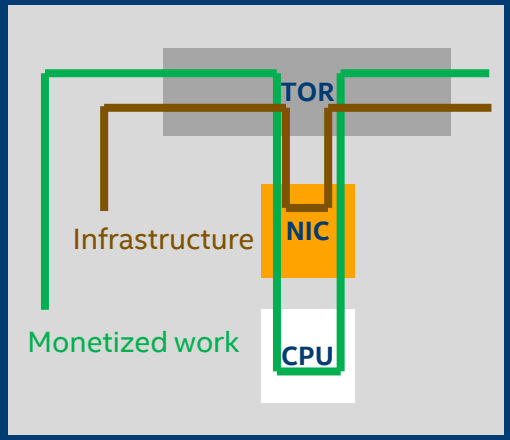
Lower Datacenter Tax



Reduce utility cycles in CPU

Application Acceleration

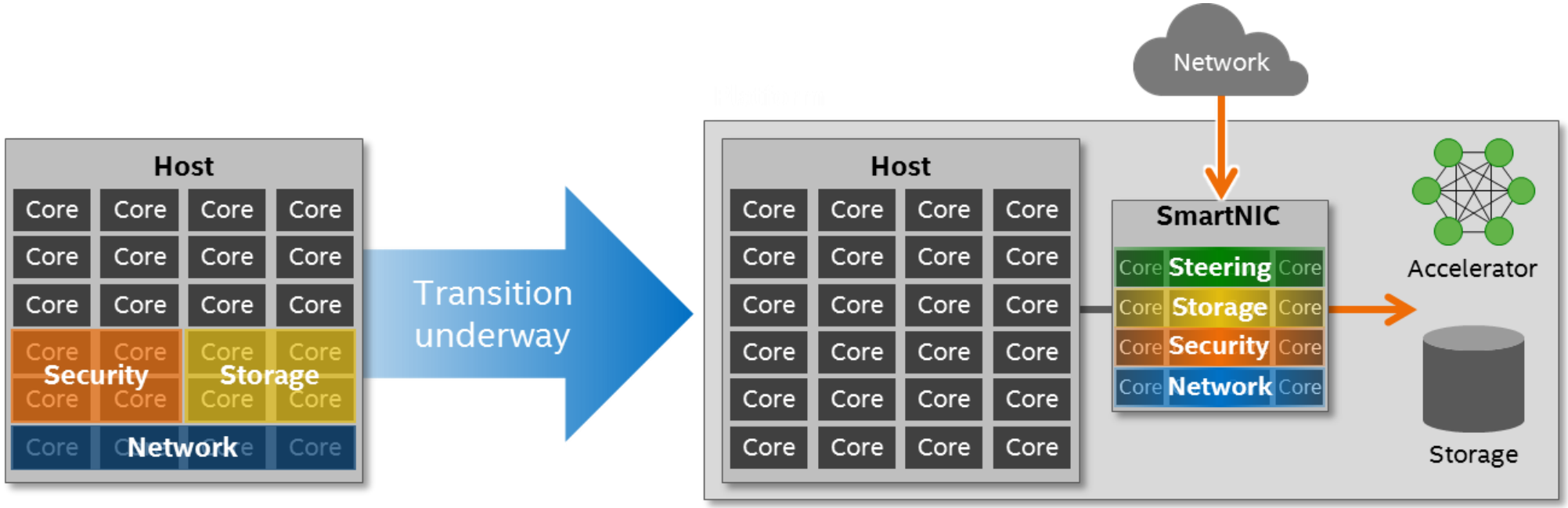
Enhance Growing Applications



Specialized infrastructure

Objective: Grow acceleration capabilities for network-related workloads

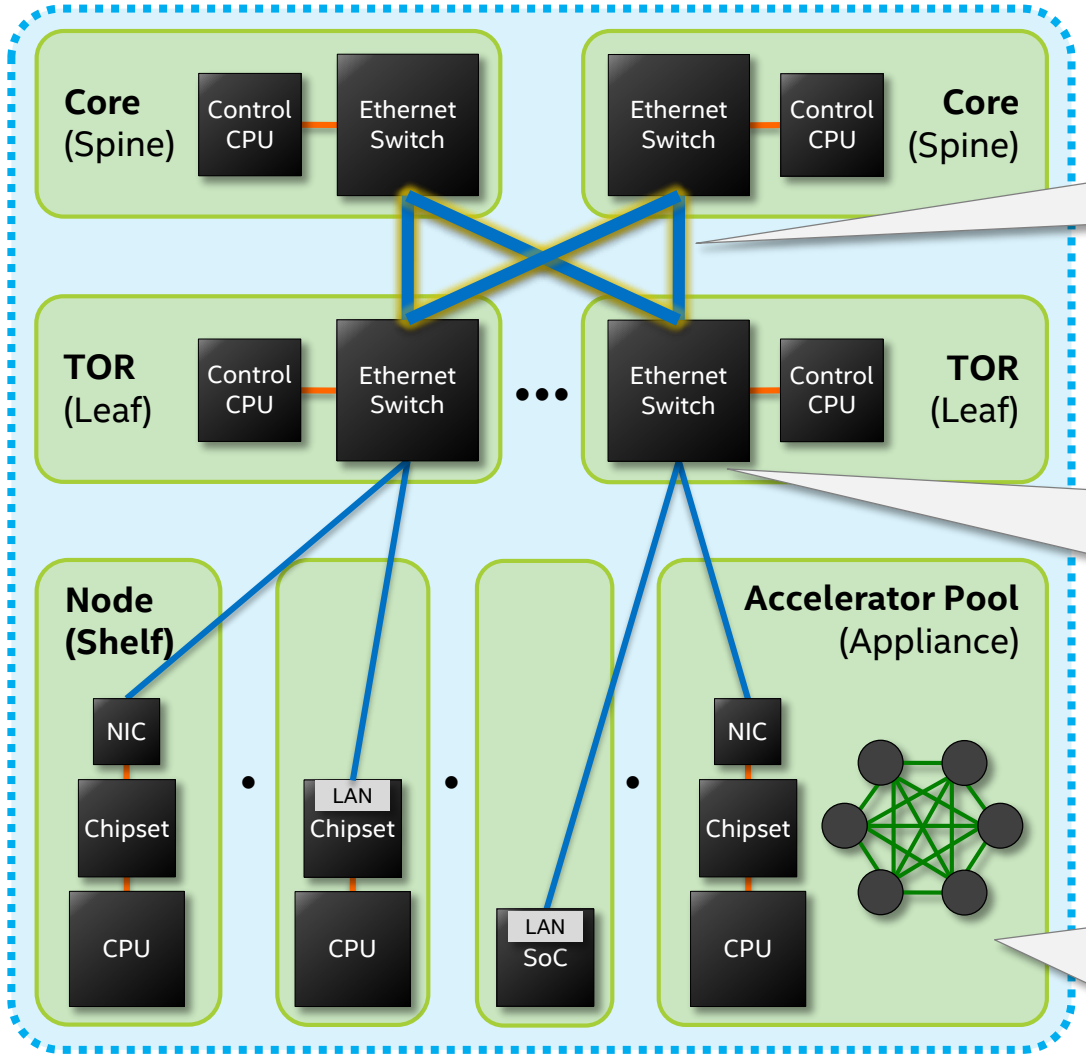
Emerging use case : bare-metal server



Bare metal servers

- Offer new usage models—full server rentals
- Improve infrastructure security in cloud
- Bolster performance, reduce cost/bit by removing bottlenecks

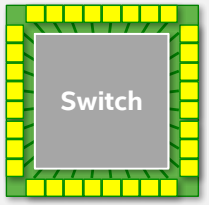
Cloud-era data center network



Single mode becoming ubiquitous

<25G Copper Optics >100G

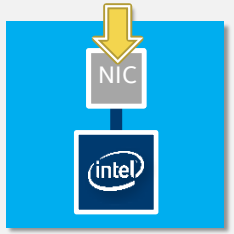
Switch defines network architecture



- Bandwidth and radix
- Power efficiency
- Reliability

Bandwidth to the node is exploding

- East-West communication
- ML, DL, application & network acceleration



NIC is innovation point for platform offloads

Motivating smart NIC

Integrating end-to-end capabilities in the platform

CORES/COMPUTE	FPGA	HOST INTERFACES	INTERFACE TECH.
• Xeon/Core • Atom • Graphics/Media • QAT/Crypto/Comp. • Special purpose (accelerators)	• Arria10 • Stratix10 • Flexible tile architecture	• PCIe Gen3, 4 and 5 • IAL (I/0, Memory, Cache) • UPI/Coherent • On-package (Rlink)	• DMA • SVM/Coherent • Hardware queue engines • Virtual device manager
NETWORK	**SOFTWARE**	**PACKAGE**	**STORAGE MEMORY**
• MAC/PHY/SerDes • RDMA (RoCE, iWARP) • OPA for HPC • Packer parsers • P4 engines • Schedulers • Switches	• Compilers, tools • WOS • OTC/Intel Clear Linux • P4 API working group • Libraries/Standards bodies	• BGA: Low-cost to large footprint • Package MCP & Interfaces (Rlink)	• NVMe

Programmable assets

- FPGA
- CPUs
- Flexible state machines

Networking assets

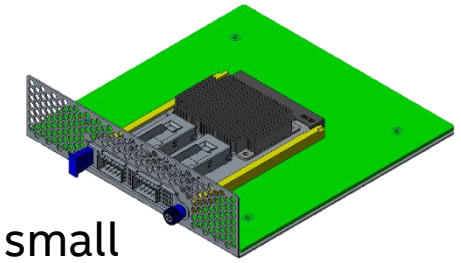
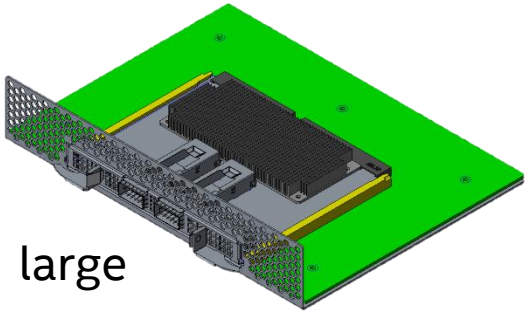
- Broad market NIC business
- Open source contribs
- DPDK

OCP NIC v3.0 enables the smart NIC

Joint effort across system/NIC/Connector suppliers and end users

Community development since Feb 2017 solved major challenges

- Improved serviceability
- Manage NIC / platform transition
- Ready for PCIe Gen4 and Gen5
- Enlarge PCB space for smart NIC use case
- Allow higher TDP ASICs
- Improve mechanical interface

Single connector	Dual connector
 small	 large
Low profile—similar profile to OCP NIC 2.0 card	Larger PCB width to support additional NIC
Up to 16 PCIe lanes	Up to 32 PCIe lanes
80W max	150W max

Significant OCP contributions in…

OCP Ethernet Network Adapters

>25 OCP-compliant NIC and PHY SKUs, covering 1GbE, 10GbE, 25GbE, and 40GbE

>1M Intel-branded ports shipped in OCP form factor (excluding partner shipments)

10GbE SFP+

25GbE SFP28

40GbE QSFP+

100G Silicon Photonics

Fully compliant with 100G CWDM4-OCP, CWDM4 MSA, QSFP, and CAUI-4 specs

500m, 2km, or 10km reach on duplex single mode fiber

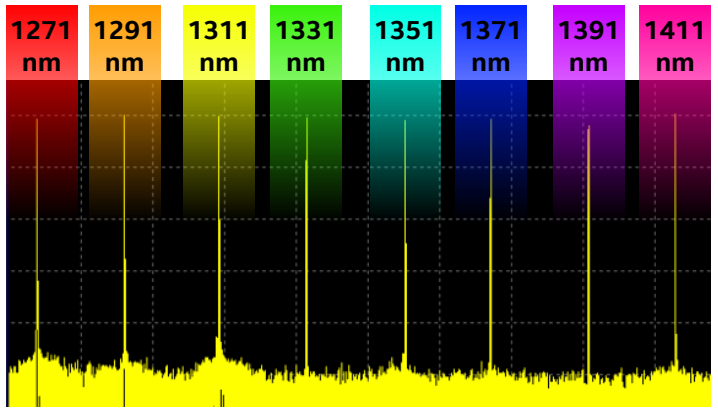
CWDM4-OCP 500m

CWDM4 2km

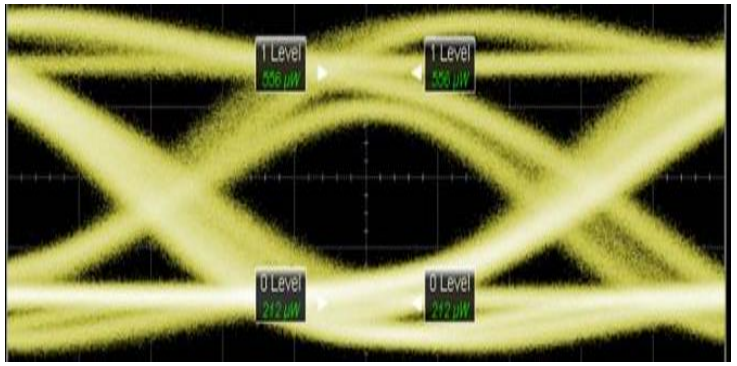
CWDM4 10km

Intel 400G CWDM8 QSFP-DD optical module

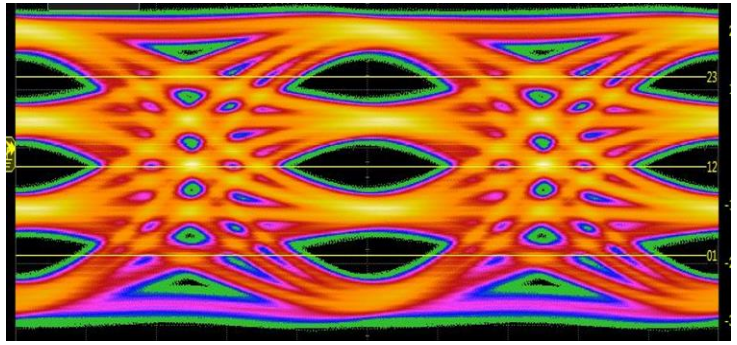
Technology for 400G is working in the lab



400G Transmitter: Optical Output



400G Receiver: Host Side Electrical Output



CWDM8 MSA

OCP SUMMIT