The Open Cloud *FPGA* Testbed – Supporting Experiments on Emerging Datacenter Configurations*

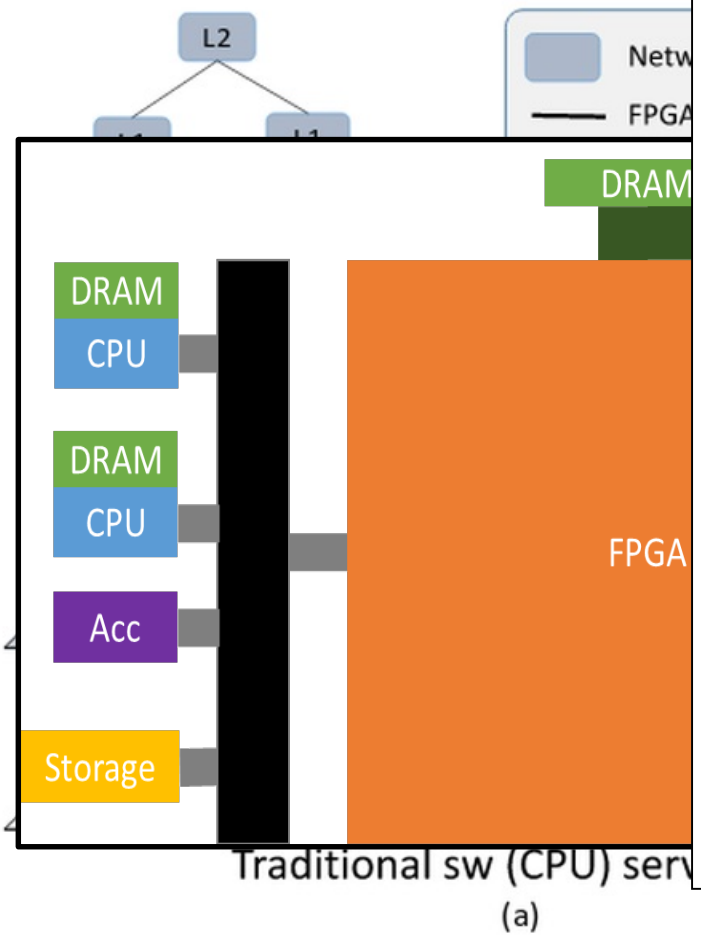
Martin Herbordt **Miriam Leeser**

BOSTON UNIVERSITY

Northeastern University

* Funded by the National Science Foundation through the Computer Community Research Infrastructure **CCRI** Grand Program

Motivation (1/3) – Millions of FPGAs in the Cloud for provider use – *Microsoft Catapult*



Provider *system* use

- SDN
- Instrumentation and Metering

Provider *internal applications*

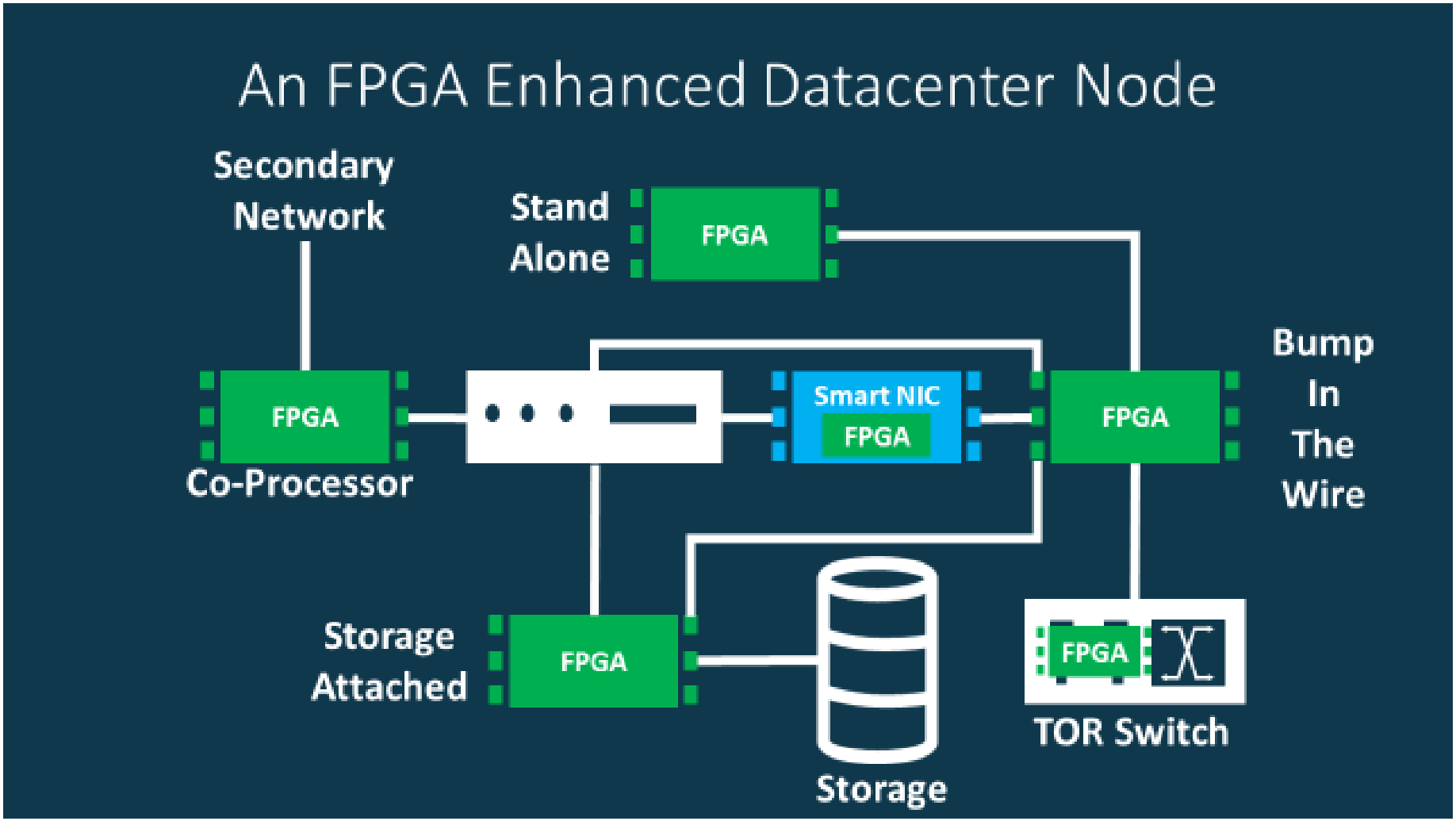
- Compression
- Encryption

Provider *external applications*

- Machine Learning
- Other big-data analytics

Fig. 1. (a) Decoupled Programmable Hardware Plane, (b) Server + FPGA schematic.

Motivation (2/3) - *FPGAs Everywhere in the Datacenter – Academic Research Programs*

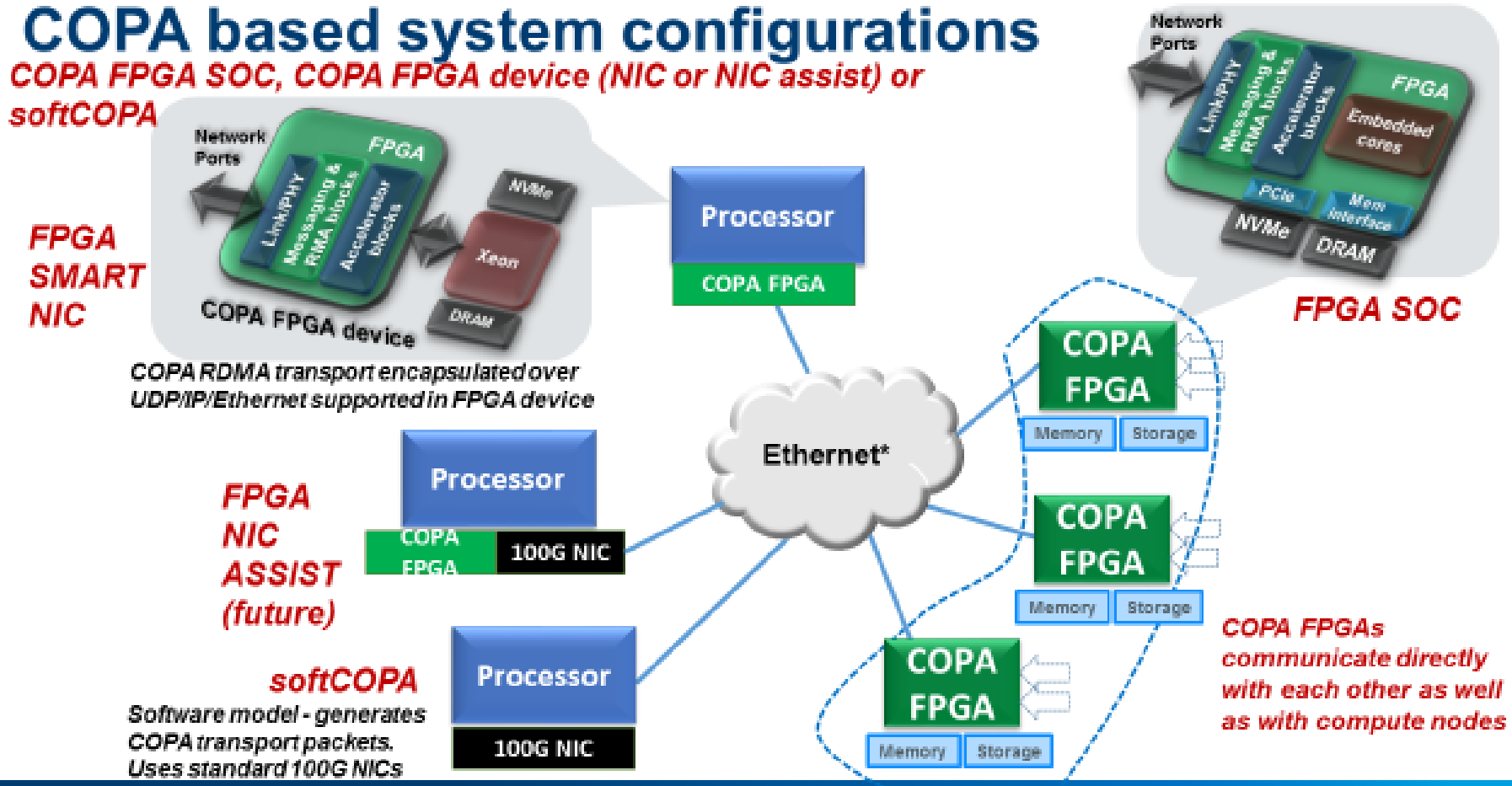


UMASS AMHERST

BOSTON UNIVERSITY

Northeastern University

Motivation (3/3) – *Potential of* Millions of FPGAs in Datacenters for HPC – *Intel COPA*



The Open Cloud Testbed

- Funded by National Science Foundation CCRI Grand Program
  - Computer Community Research Infrastructure

- Collaboration among
  - UMass Amherst
  - Boston University
  - Northeastern University

UMASS AMHERST

BOSTON UNIVERSITY

Northeastern University

Core Team

Mike Zink, PI
Orran Krieger, Co-PI, lead @ BU
Peter Desnoyers, Co-PI,
lead @ Northeastern
Miriam Leeser, Co-PI,
Northeastern
Martin Herbordt, Co-PI, BU



David Irwin,
Community Outreach Director, UMass
Emmanuel Cecchet,
Senior Research Scientist, UMass
Jack Brassil,
Head of Advisory Board, Princeton



MOC
MGHPCC
MASSACHUSETTS GREEN HIGH PERFORMANCE COMPUTING CENTER


CloudLab
Northeastern

UMASS AMHERST

BOSTON UNIVERSITY

Northeastern University

The Open Cloud *FPGA* Testbed - *OCFT*

Tag line → An MOC-style Catapult testbed *and so much more*

- Enhanced with programmable hardware (FPGA) capabilities not present in other facilities available to researchers today

FPGAs in the Datacenter: What exists

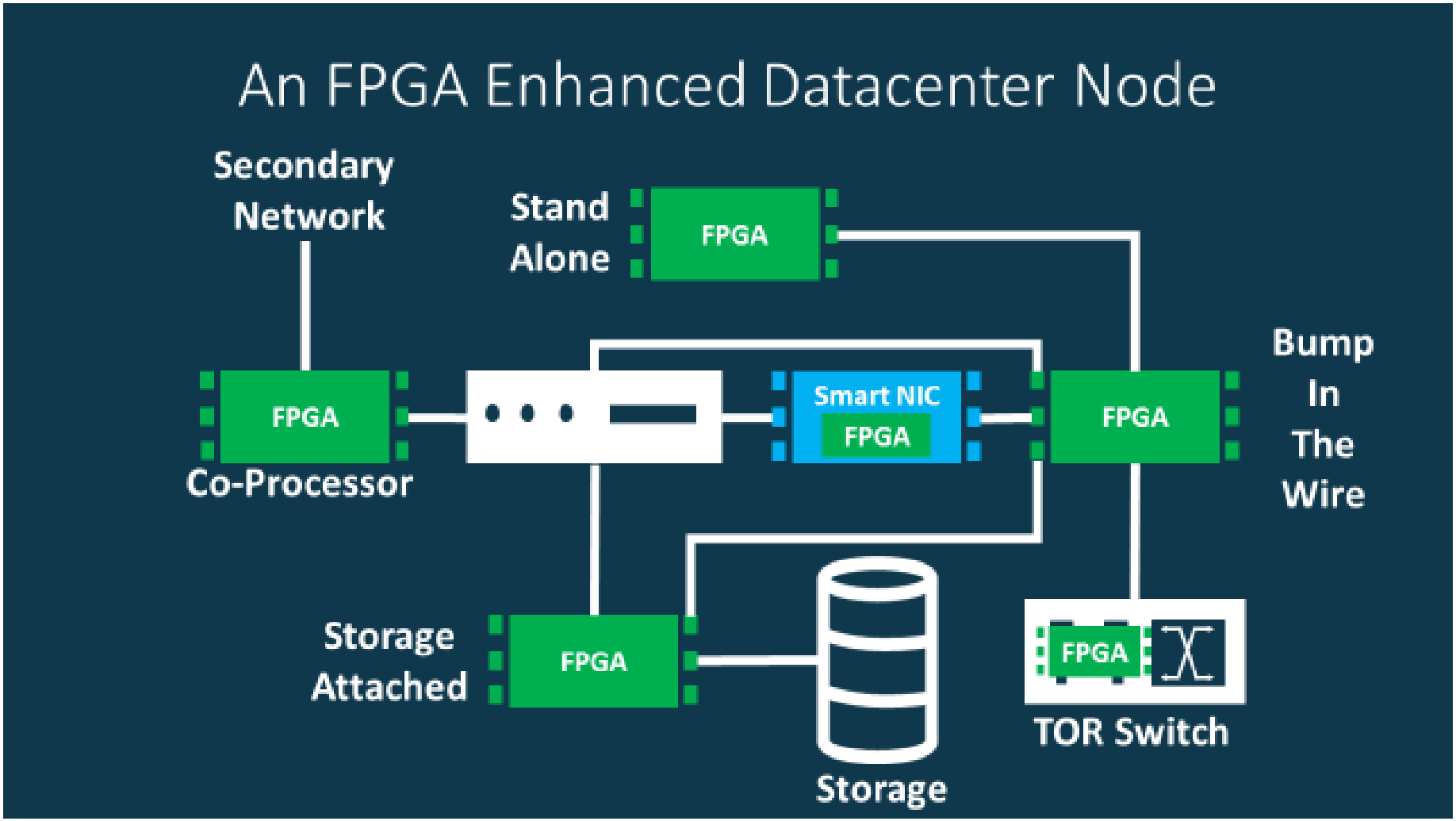
- Microsoft Catapult
  - No user access
- AWS F1 instances *>> and Baidu, Chameleon, TACC, etc.*
  - Available to users as accelerators, but interactions are restricted
- Various FPGA-centric clusters *>> BU, Paderborn, Riken, TACC, Tsukuba*
  - Very difficult to bring on line, even for a single institution
  - Even more difficult to maintain
  - HPC-specific rather than general datacenter

UMASS AMHERST

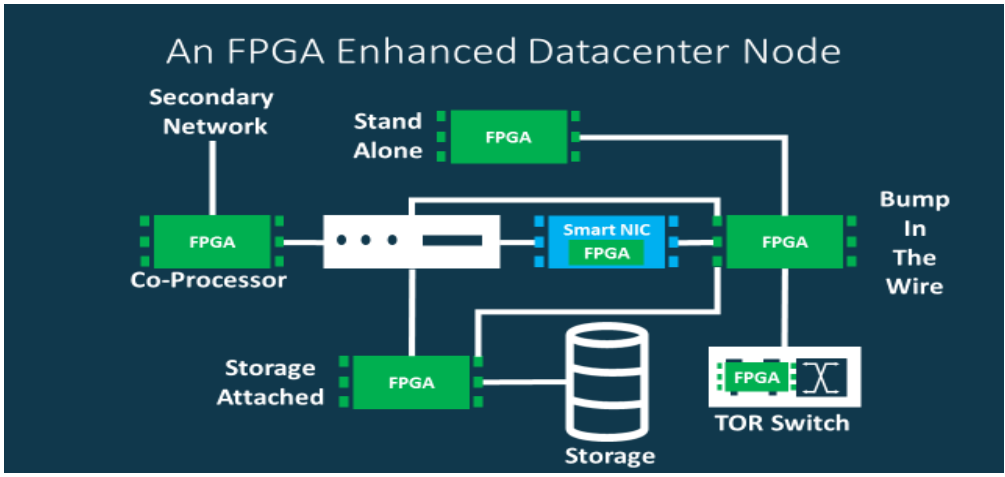
BOSTON UNIVERSITY

Northeastern University

OCFT for FPGAs in the Datacenter



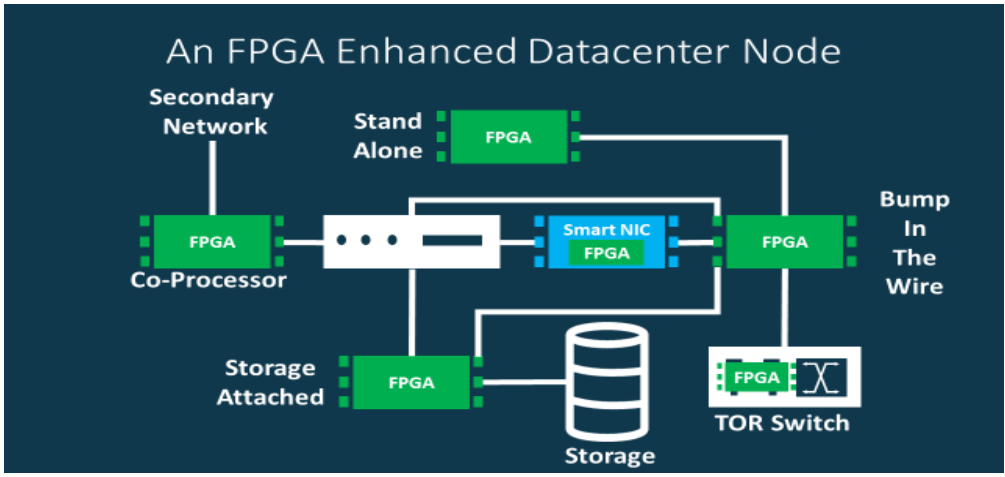
How OCFT will be used – Sample Projects



- **Hardware operating system (on the FPGAs)**
  - Drivers, Multitenancy, Handling "Pass-through" system communication
- **Development environment**
  - Enable access and programming by system and application developers
- **System applications**
  - Compression, security
- **User applications – in the node to across the datacenter**
  - Middleware offload – MPI
  - Application-aware I/O support through lossy compression
  - Massively parallel applications – Large scale physical simulations
  - Distributed machine learning

11:40 – Ahmed Sanaullah
HW OS and Sys App Development

Why OCFT will work



Funding for FPGA-specific system management and customer service

- FTE FPGA engineer

Integration into existing cloud ecosystem

Broader community will be pitching in

- Industry partners, advisory board, beta users

UMASS AMHERST

BOSTON UNIVERSITY

Northeastern University

OCFT – Beta Users

- **Alpha** cohort – Herbordt & Leeser research groups, Red Hat
- **Beta** cohort – Established FPGA/Cloud/HPC research groups. Survey is for Beta cohort.
- **Gamma** cohort – broader community with certain attributes, particularly the experience to be able to use this rather than other infrastructure.

Initial list of potential users by affiliation

Universities	Replies
Boston University	2
Brown	
BYU	
Cornell	
CMU	x
MIT	x
NCSU	x
Northeastern	
Penn	
Stevens	
Tufts	x
U. Arkansas	x
U. Alabama	x
UCSD	
U. Florida	x
U. Miami of Ohio	x
U. Massachusetts	x

Universities, cont.	Replies
UNCC	
U. Pittsburgh	
U. Tennessee	
Worcester Polytechnic	
Wash. U. St. Louis	x
W. Michigan	
Yale	

National Labs	Replies
Argonne	x
Lawrence Berkeley	
Pacific Northwest	x

Industry	Replies
AlgoLogic	
Atomic Rules	x
Comma Corp	x
Gray Research LLC	
Red Hat	x

Beta user configuration priority

	First Choice	Total
FE1: Catapult2-like – Bump-in-the-Wire	10	10
FE2: Programmable NIC	2	2
FE3: FPGA is the node	0	0
BE1: Bare-metal back-end processor	1	2
BE2: Tightly coupled back-end processor (CCIX)	2	4
BE3: Cluster of directly connected FPGAs	2	5

Beta user project types

Project Type	first choice	total
Cloud and Operating System	6	6
Middleware	2	5
FPGA systems	3	4
FPGA tools	3	6
Provider applications	1	3
Tenant applications	2	3

Miscellaneous

Enthusiasm for OCFT (17 replies)

13/17 gave as part of their answer some variation of ***very interested***

4/17 gave ***practical responses*** of what they would do with OCFT

Tools preference (17 replies)

Intel – 11 Xilinx – 12 Generic – 1

Both or would switch – 12/17

HBM? (17 replies)

Yes = 8 No = 1 "Nice but" = 2 No reply re HBM = 6

What board? (17 replies)

no reply = 2 no preference = 7 Xilinx = 2 Intel = 4 Both = 2

FPGA Options

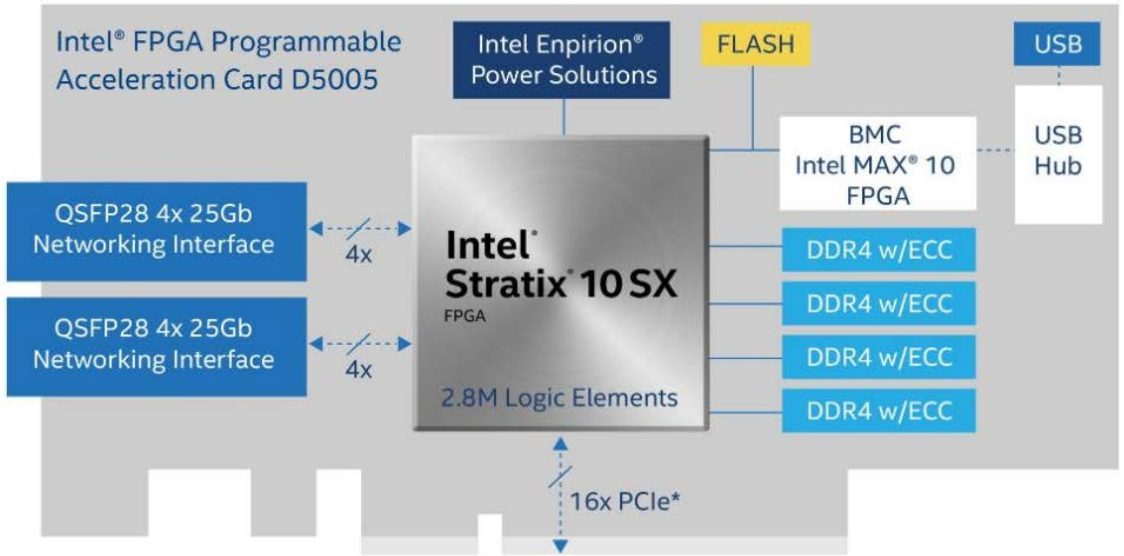
Xilinx Alveo Cards for data centers:

https://www.xilinx.com/products/boards-and-kits/alveo/u280.html#specifications

Card Specifications	U280 Production	
Thermal Cooling	**Passive**	**Active**
Compute		
INT8 TOPs (peak)	24.5	24.5
Dimensions		
Width	Dual Slot	Dual Slot
Form Factor	Full Height, 3/4 Length	Full Height, Full Length
DRAM Memory		
HBM2 Total Capacity	8GB	8GB
HBM2 Total Bandwidth	460GB/s	460GB/s
DDR Format	2x 16GB 72b DIMM DDR4	2x 16GB 72b DIMM DDR4
DDR Memory Capacity	32GB	32GB
DDR Total Bandwidth	38GB/s	38GB/s
SRAM Memory		
Internal SRAM Capacity	41MB	41MB
Internal SRAM Total Bandwidth	30TB/s	30TB/s
Interfaces		
PCI Express	Gen4x8 with CCIX	Gen4x8 with CCIX
Network Interfaces	2x QSFP28 (100GbE)	2x QSFP28 (100GbE)
Logic Resources		
Look-up Tables (LUTs)	1,079,000	1,079,000
Power		
Maximum Total Power	225W	225W

Intel D5005:

https://www.intel.com/content/www/us/en/programmable/products/boards_and_kits/dev-kits/altera/intel-fpga-pac-d5005/overview.html



Advantages and Disadvantages

- Xilinx Alveo 280
  - + High Bandwidth Memory (HBM)
  - -- Only 2 QSFP28 connections
  - Programming: Xilinx Vitis Tool
- Intel D5005
  - + 4 QSFP28 connections
  - -- No HBM
  - Intel OneAPI

For more on OCFT

Breakout session tomorrow afternoon @2PM

UMASS AMHERST

BOSTON UNIVERSITY

Northeastern University