Network and Memory Abstractions on FPGAs for Distributed Applications.

Dario Korolija, Zhenhao He, Gustavo Alonso

Systems Group

Department of Computer Science

ETH Zurich, Switzerland

The tutorial

- Slides available at:

https://systems.ethz.ch/research/data-processing-on-modern-hardware/hacc.html

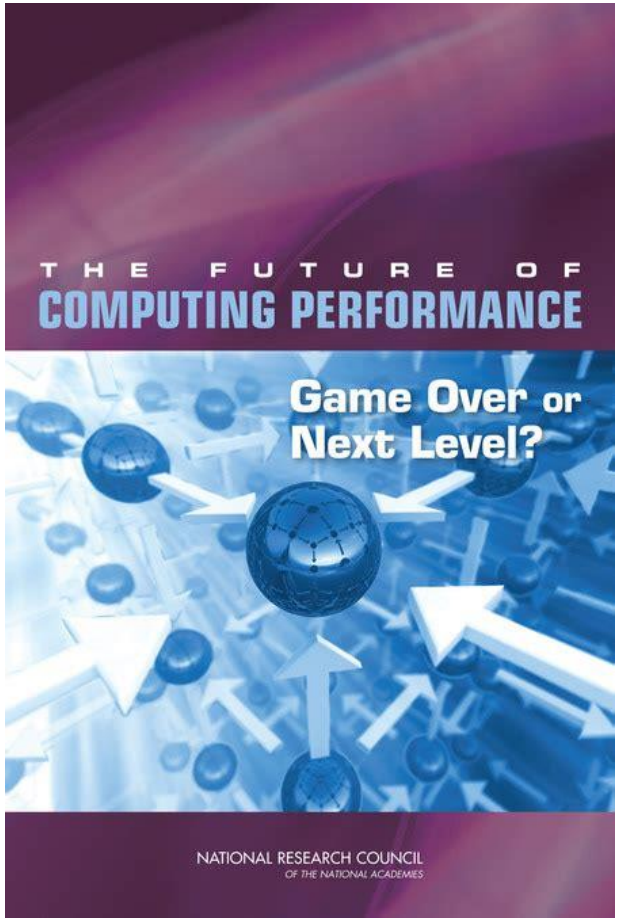
- More tutorials available covering diverse use cases and technologies around FPGAs
  - FPGA'23 tutorial on networking on FPGAs
  - SIGMOD'23 tutorial on data processing on FPGAs
- Also information on the use of the HACC cluster, research papers, etc.

Schedule

- Introduction and Motivation
- Coyote: an open shell for FPGAs
- EasyNet: an open, 100 Gbs TCP/IP network stack
- ACCL: collective network communication for FPGA clusters
- Farview: Smart Disaggregated Memory
- Distributed inference: on recommendation systems

The Hardware Era

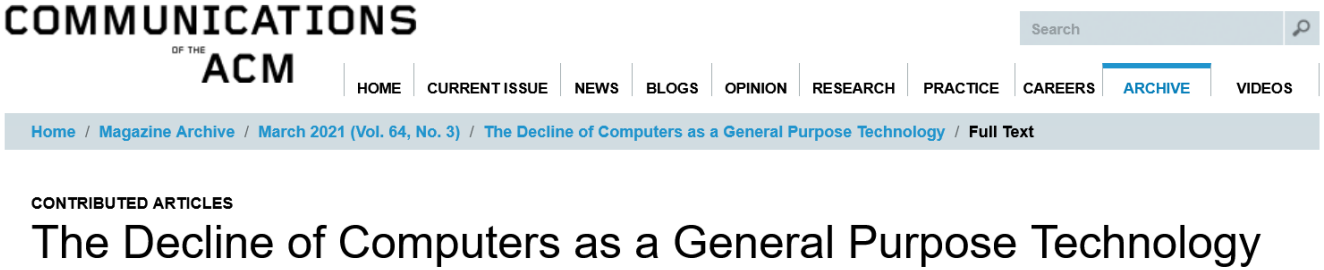
Not a new concept …

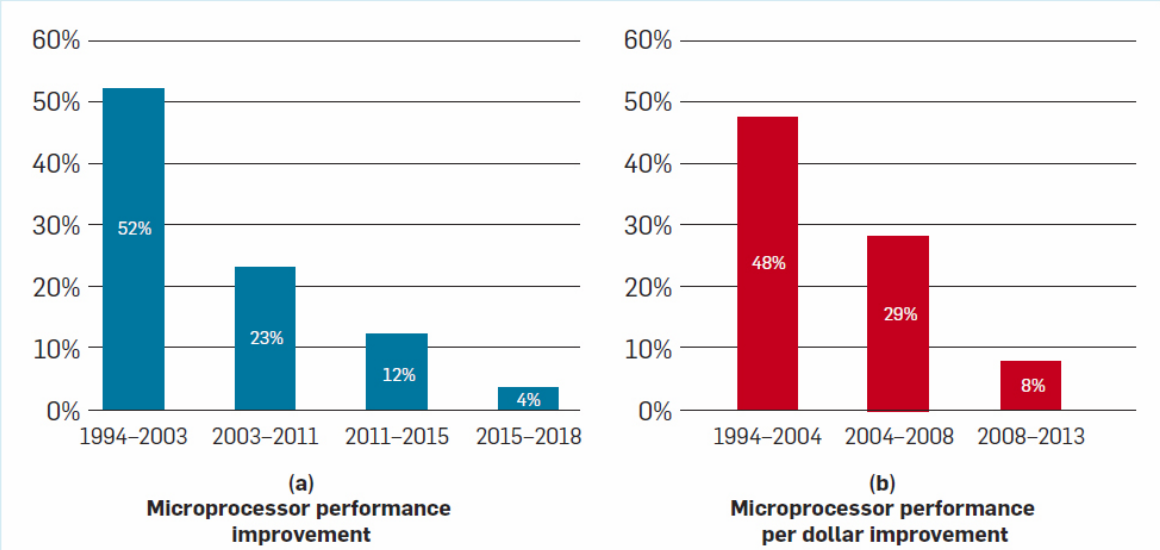


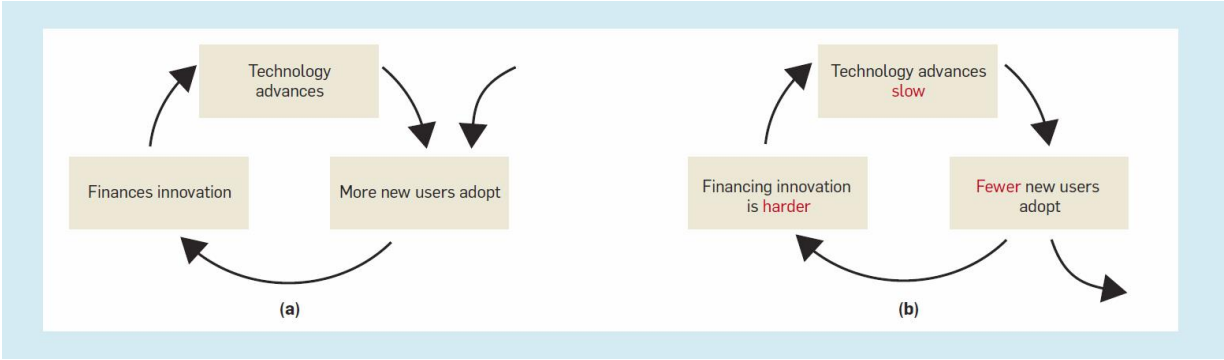
- 2011 Report
- Exponential growth for several decades
- Exponential growth no longer possible
- Switch to multicore and parallelism
  - Energy consumption becomes an issue
  - Multicore introduces parallelism that we do not know how to exploit well
- Situation will not change in near future
- Alternative is specialization
- Either somebody comes up with a new great invention or there is a problem

General purpose computing

Slow improvements lead to specialization

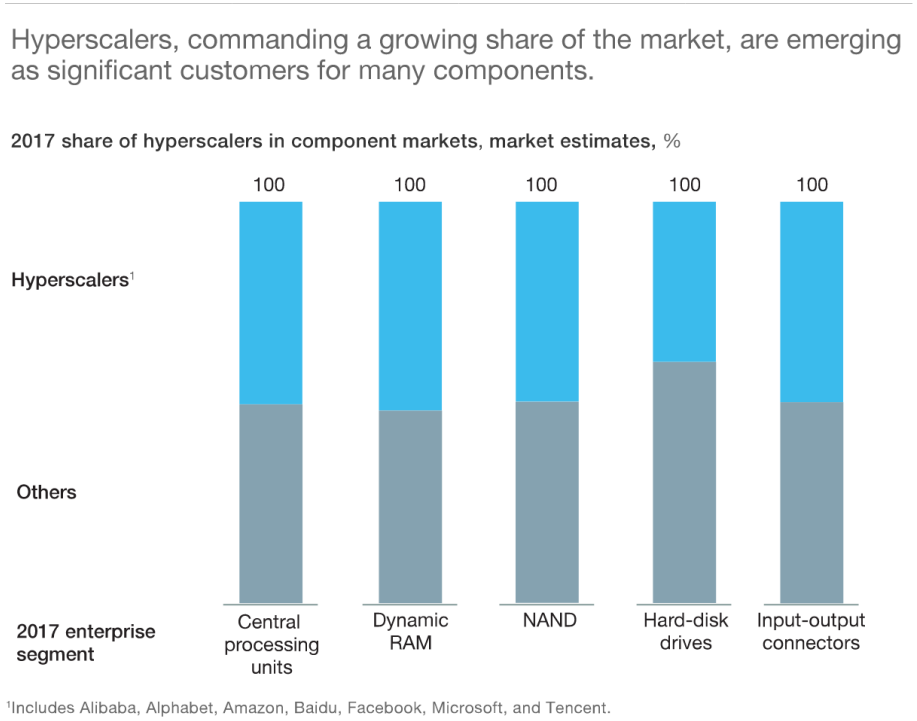






Driving specialization

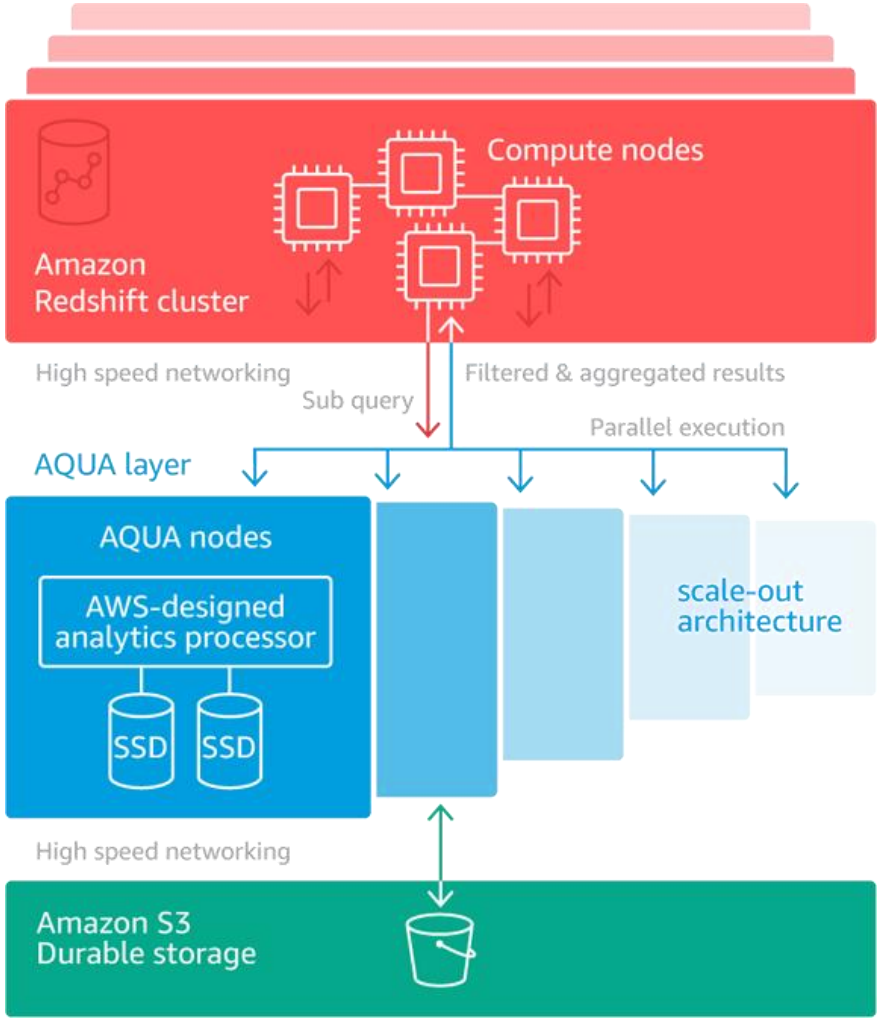
- The cloud is the big game changer:
  - New business model
  - Economies of scale
  - Very large workloads
- Every hyper scaler is its own “Killer App”
  - The scale makes many things feasible
  - The gains have a very large multiplier



https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/how-high-tech-suppliers-are-responding-to-the-hyperscaler-opportunity

Large deployment of FPGAs in the cloud – examples

- FPGAs as smart accelerator for disaggregated resources
- Amazon AQUA
  - *https://aws.amazon.com/blogs/aws/new-aqua-advanced-query-accelerator-for-amazon-redshift/*
    - Analytic engine with FPGAs
    - Pushing computation closer to data
    - Reduce CPU compute requirement
    - Reduce network traffic



Data Compression (Microsoft Zipline/Corsica)

https://azure.microsoft.com/en-us/blog/improved-cloud-service-performance-through-asic-acceleration/

Emerging themes

- Reduced CPU utilization
- Accelerate common operations
- Accelerate the infrastructure supporting the system
- Processing data on the fly
- Near data processing (memory, storage, …)
- On demand servers and functionality
- …

HACC cluster at ETH Zurich

Infrastructure – HACC cluster

- The Heterogeneous Accelerated Compute Clusters (HACC) program is a unique initiative to support novel research in adaptive compute acceleration for data center settings and high-performance computing (HPC).

- ETH Zurich HACC *https://systems.ethz.ch/research/data-processing-on-modern-hardware/hacc.html*

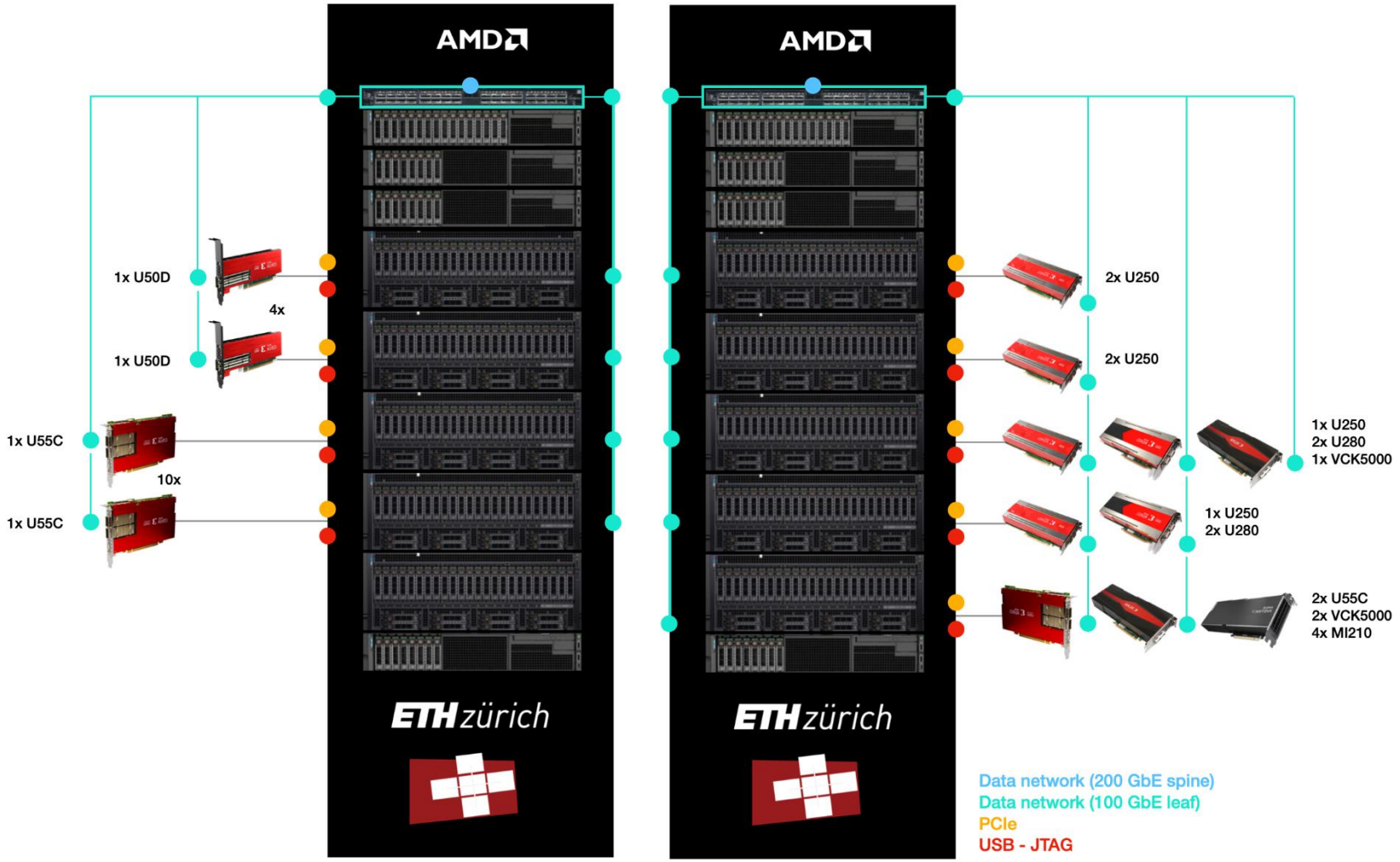
HACC Cluster

- Target: Facilitate practitioners and researchers exploring distributed applications on FPGA clusters

ENZIAN

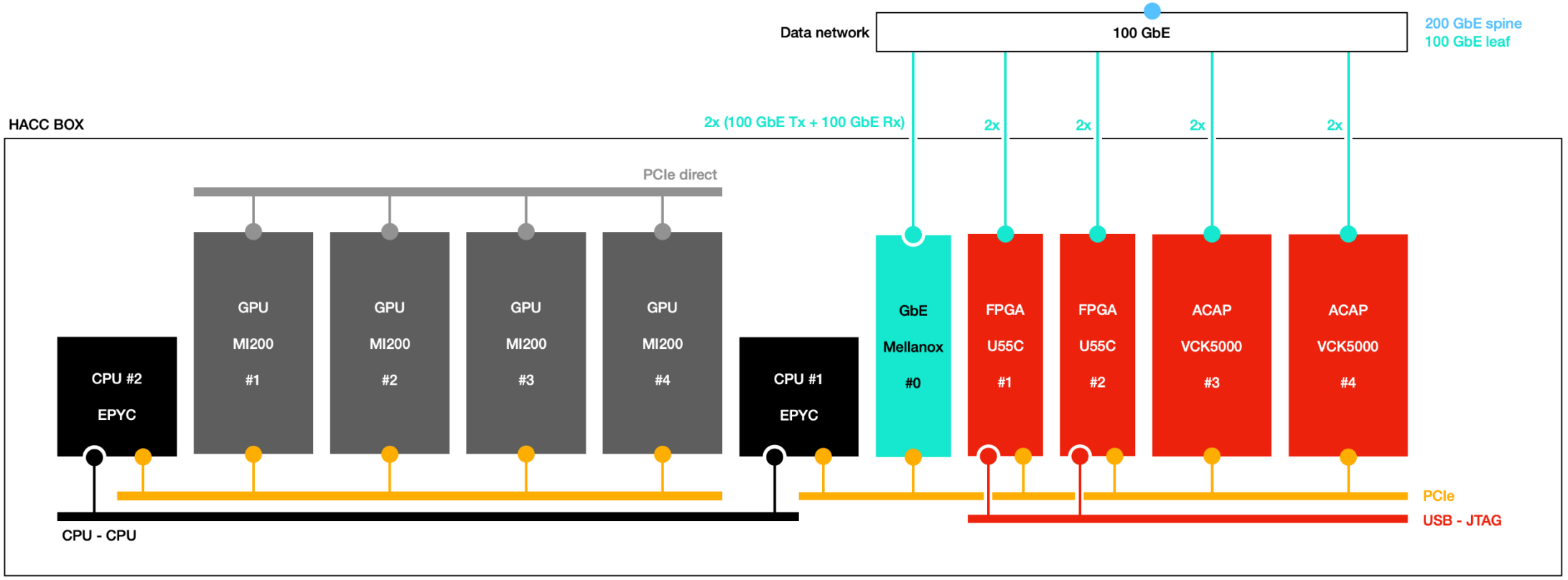
  - **FPGA clusters (HACC and Enzian)**
    - Data center standard infrastructure
  - **Frameworks and abstractions (Vitis and Coyote)**
    - Shell support and abstractions for in-network processing, disaggregated computation, distributed applications …
  - **Systems and applications built on top**

Introduction to HACC cluster



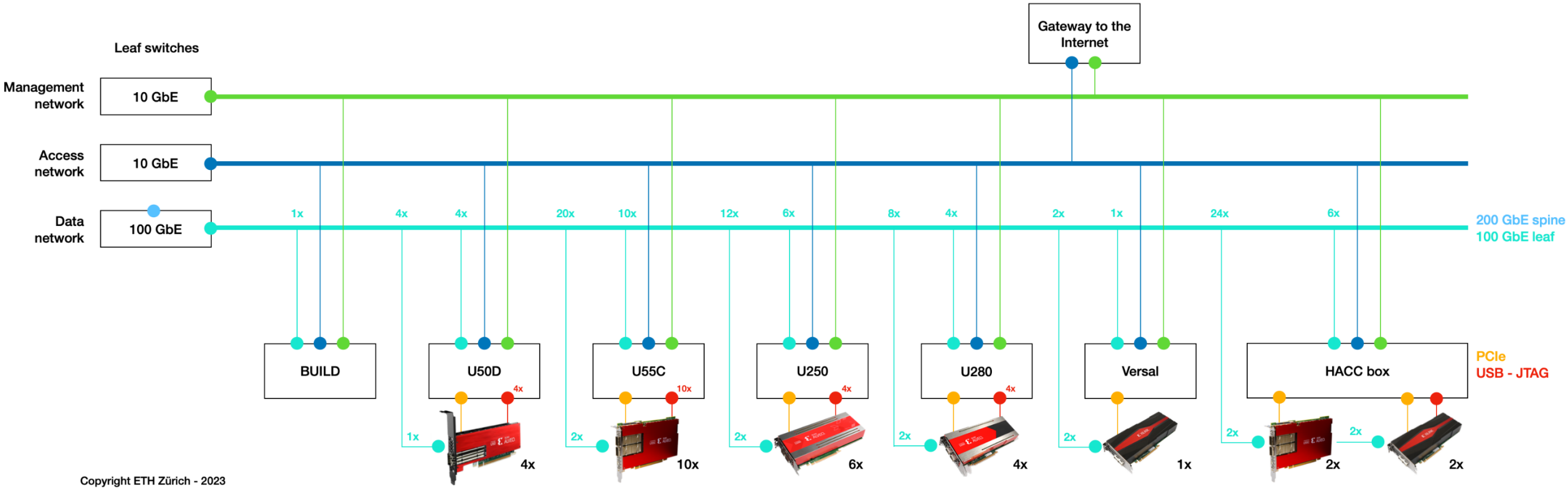
Copyright ETH Zürich - 2023

Overview (HACC heterogeneous boxes)



Copyright ETH Zürich - 2023

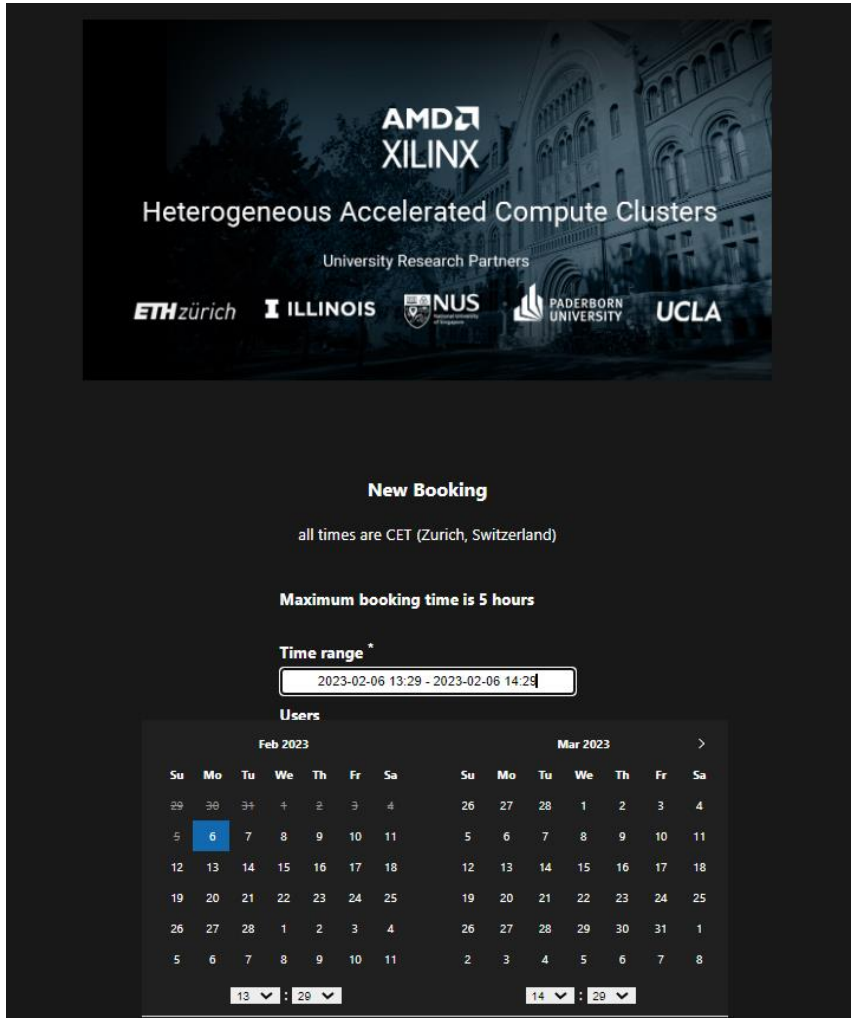
Overview



Copyright ETH Zürich - 2023

Booking system

- Reserving a specific VM/device for a specific period
  - Maximum 5 hours per reservation
- During a reservation, only the selected user can connect to the VM/device
- User can choose different workflows when login
  - Vitis workflow
  - Coyote workflow



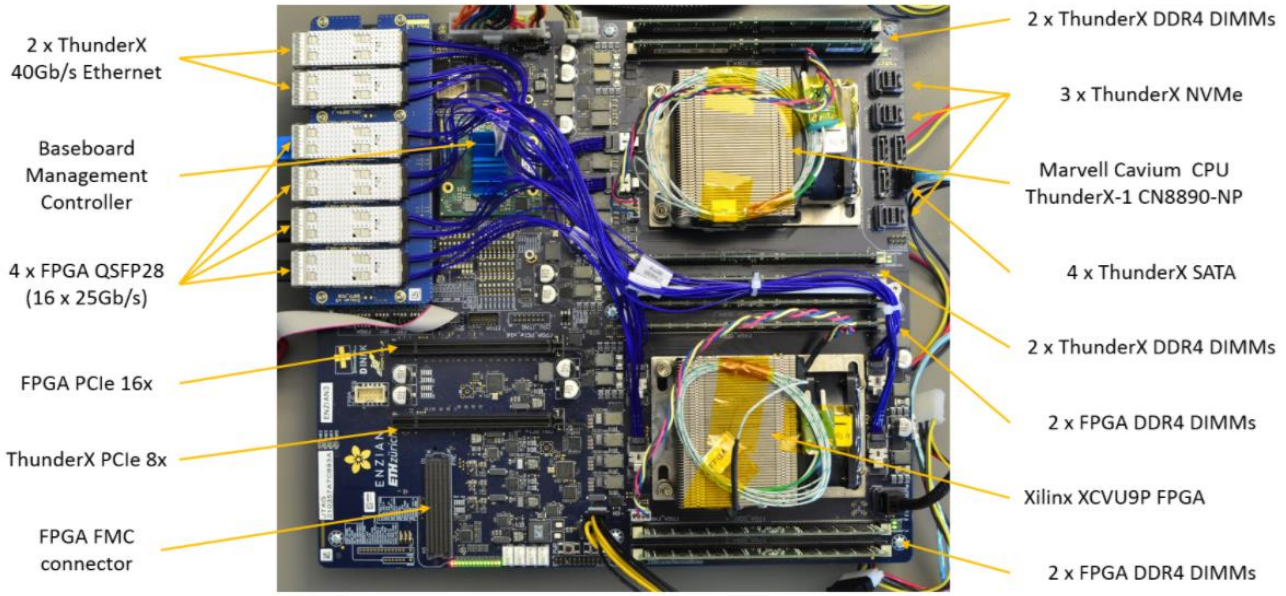
User access

- Access requires registration
  - ETH users contact Gustavo Alonso
  - All others through AMD Xilinx (HACC program)
  - Users get guest account at ETH (renewable)

Enzian

- Research computer developed within the Systems Group at ETH
- Designed for computer systems software research, deliberately over-engineered
- Big server-class CPU closely coupled to a large FPGA, with ample main memory and network bandwidth on both sides
- Cache-coherent asymmetric NUMA system

ENZIAN



http://enzian.systems
ASPLOS '22

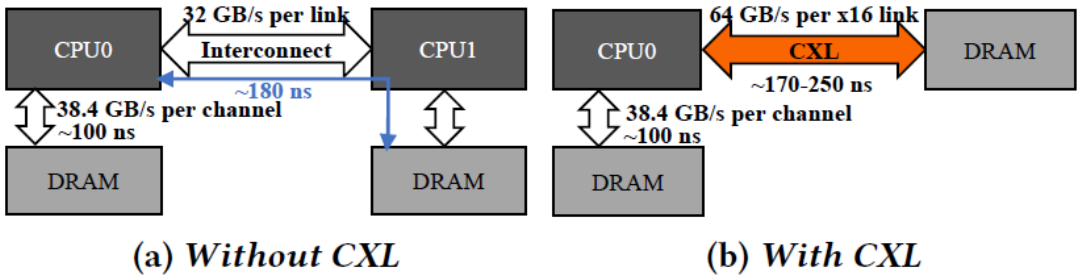
Opportunities for FPGAs

The future of accelerators

TPP: Transparent Page Placement for CXL-Enabled Tiered Memory

Hasan Al Maruf*, Hao Wang†, Abhishek Dhanotia†, Johannes Weiner†, Niket Agarwal†, Pallab Bhattacharya†, Chris Petersen†, Mosharaf Chowdhury*, Shobhit Kanaujia†, Prakash Chauhan†

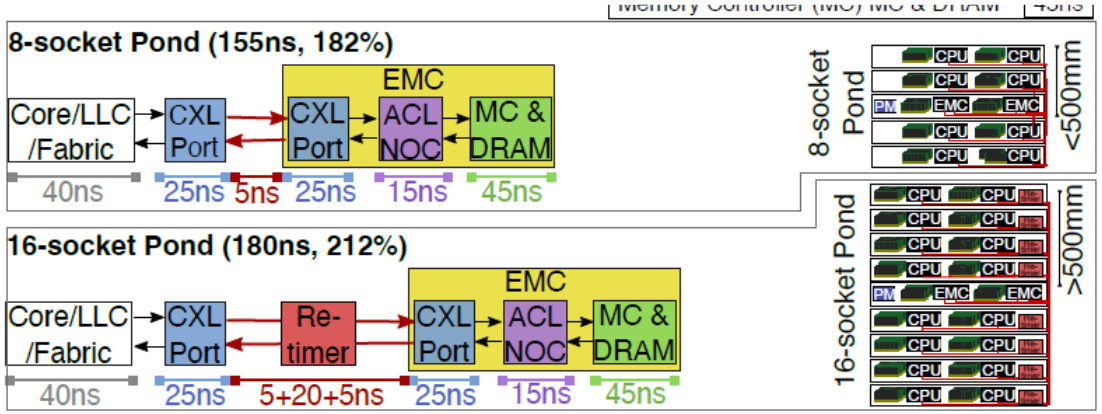
University of Michigan* Meta Inc.†



Pond: CXL-Based Memory Pooling Systems for Cloud Platforms

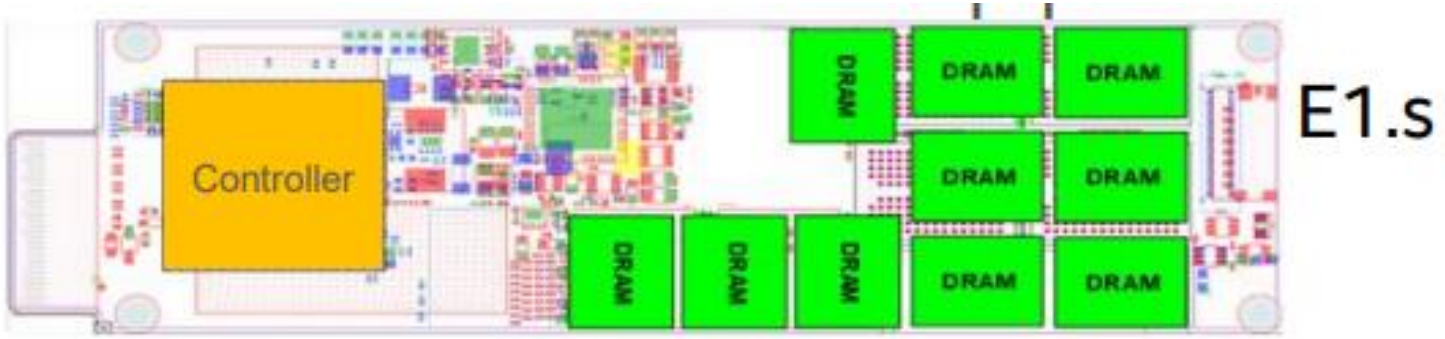
Huaicheng Li†, Daniel S. Berger*‡, Stanko Novakovic*, Lisa Hsu*, Dan Ernst*, Pantea Zardoshti*, Monish Shah*, Samir Rajadnya*, Scott Lee*, Ishwar Agarwal*, Mark D. Hill*°, Marcus Fontoura*, Ricardo Bianchini*

†Virginia Tech and CMU *Microsoft Azure ‡University of Washington °University of Wisconsin-Madison



Disaggregated memory

- CXL memory will not be just memory
- It will be a module with a controller/processor that runs the protocol and manages the memory
- The controller is a great point to add near-data processing capabilities



The tutorial in context

FPGAs in context

- We do not sell or market FPGAs
- FPGAs are the only way to explore:
  - New architectural designs (even to the CPU design level, e.g., RISC-V)
  - New computer architectures (near-memory processing, smart storage, smart NICs, accelerators, etc.)
  - Processing of data streams at line rate
- Are FPGAs difficult?
  - No, this is systems level programming, no less involved than writing your own database engine, operating system, etc.
  - Yes, the tools are not what we are used to in the software world (by a long margin)

Goals

- Showcase open source tools available for researchers
- Facilitate research in data center applications and distributed computing
- Overall: encourage the community to explore this opportunity of achieving higher efficiency in data centers in a context where the new hardware is going to be available, even if for other reasons.