

European Open
Compute **A**rchitecture for **P**owerful **E**dge

Making Composable Edge Infrastructure Interoperable: Open Standards and Interfaces in CAPE

Martin Kaiser, Lennart Tigges, Christian Klarhorst, René Griessl, Jens Hagemeyer

The CAPE project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101189899.

Jens Hagemeyer, Bielefeld University

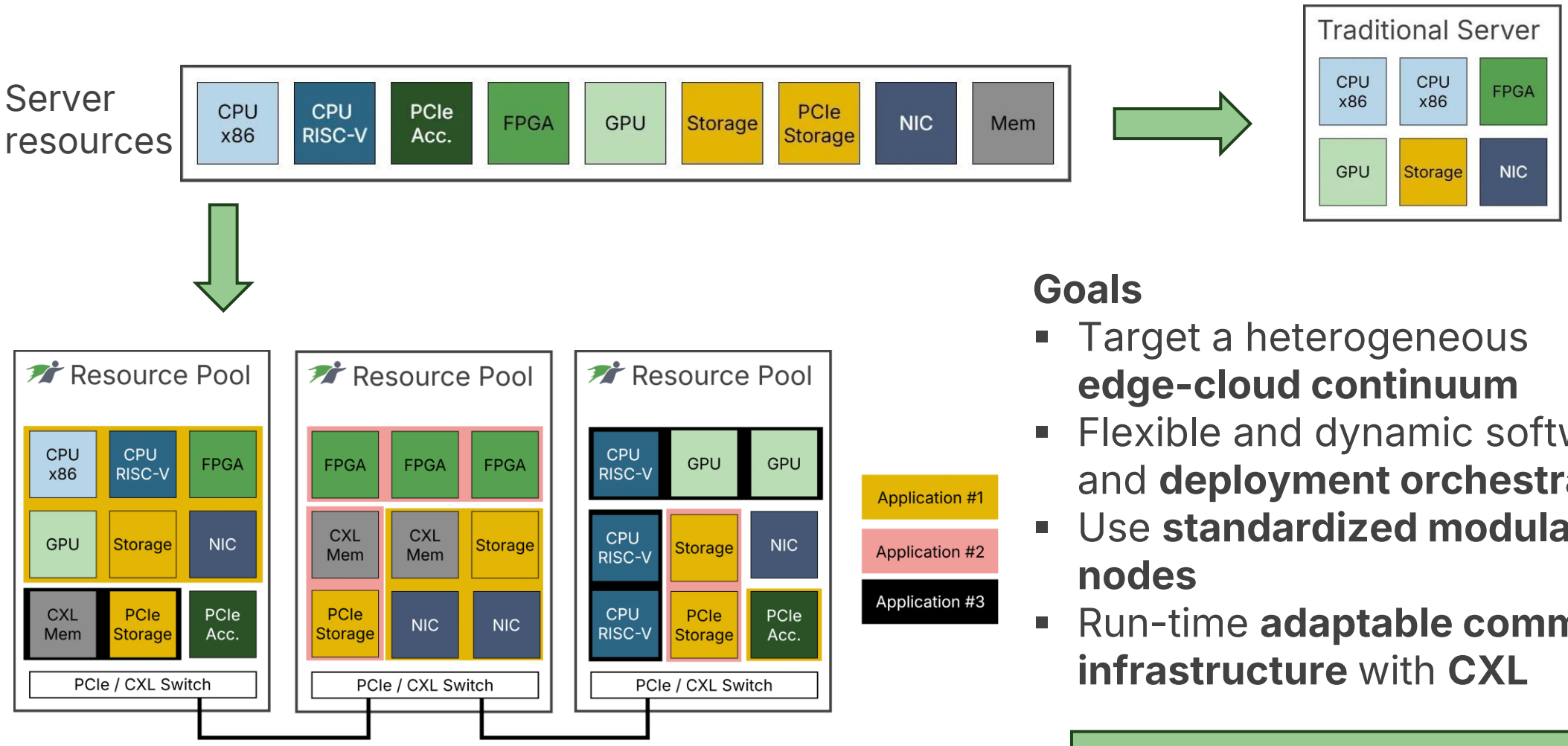
CAPE - European Open **C**ompute **A**rchitecture for **P**owerful **E**dge

- EU-Call: „Open Source for Cloud/Edge to support European Digital Autonomy" (Horizon CL4 2024 Digital Emerging 01-21)
- Start: 12/2024
- Consortium
  - 7x Industry (6x SME)
  - 4x Academic
  - 1x Standardization organization
- Coordinator: Bielefeld University



* associated participants

CAPE's Vision
towards a Composable Server Infrastructure



Resource Pools → Composable Infrastructure

Goals

- Target a heterogeneous **edge-cloud continuum**
- Flexible and dynamic software stack and **deployment orchestration**
- Use **standardized modular compute nodes**
- Run-time **adaptable communication infrastructure** with **CXL**

CXL (Compute eXpress Link) is based on the PCIe transport layer and provides for cache-coherency and memory pooling

Big Picture of CAPE

Use Cases

Smart Grid

Edge AI

Satellite Payload Processing

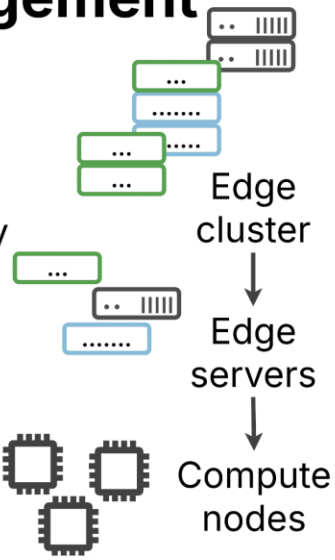
Standardization

OFEIA
(**O**pen **F**ederated **E**dge-Cloud **I**nfrastructure **A**lliance)

COM HPC

Open Edge Server Management

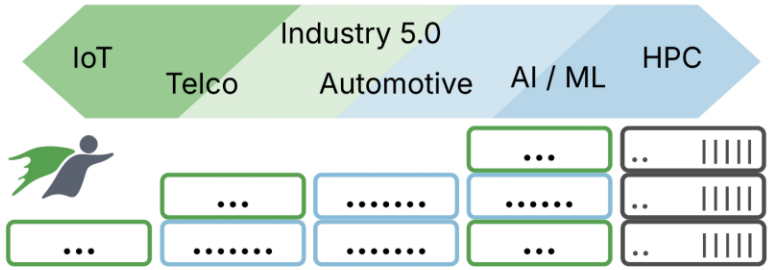
- Web-based, out-of-band configuration
- Run-time topology reconfiguration
- Live power monitoring
- Support for heterogeneous computing nodes



Software Services

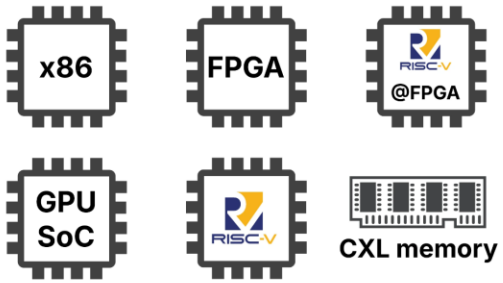
- **I**nfrastructure **f**rom **C**ode (IfC)
- AI-assisted orchestration and service deployment
- Resource clustering: dynamic node allocation & composition

Edge-Cloud Continuum



Hardware

Modular, Heterogeneous Compute Nodes



Supported Edge Server

eHPS ... (up to 3 nodes)
embedded **H**igh-**P**erformance **S**erver

EMDC ..... (up to 7 nodes)
Edge **M**icro **D**ata **C**enter

Third party on-premises or cloud

Environment Integration



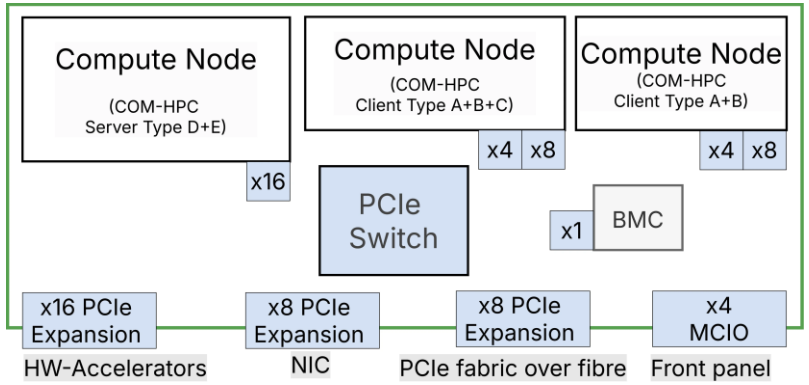
- Wall mount & 19" rack
- Hybrid: passive and active cooling
- Waste heat recovery

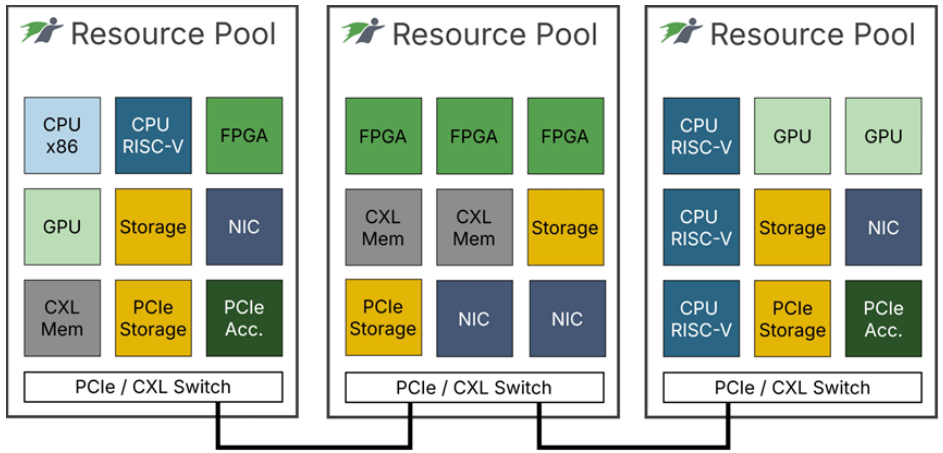
Hardware Platforms

UNIVERSITÄT BIELEFELD · CHRISTMANN INFORMATIONSTECHNIK + MEDIEN · parXent efficiency matters



eHPS: embedded **H**igh **P**erformance **S**erver



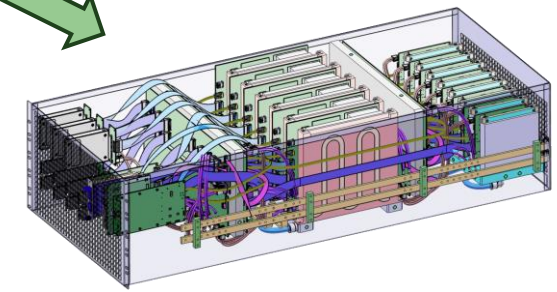


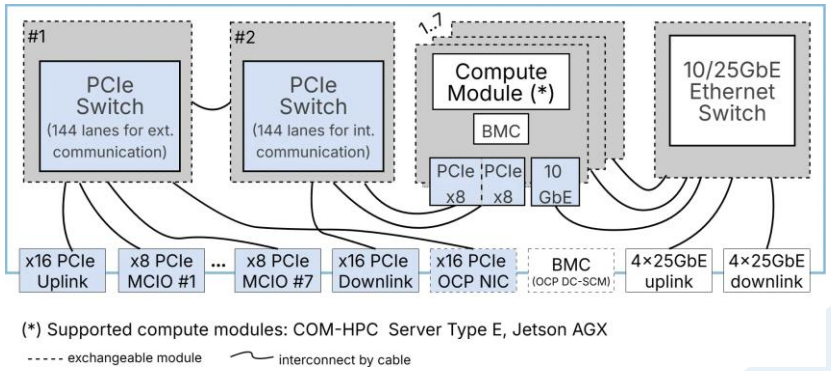
Heterogeneous edge servers

Two server platforms

High-density computing, hybrid cooling & building integration

EMDC: Edge **M**icro **D**ata **C**enter





CAPE's fully Composable Infrastructure

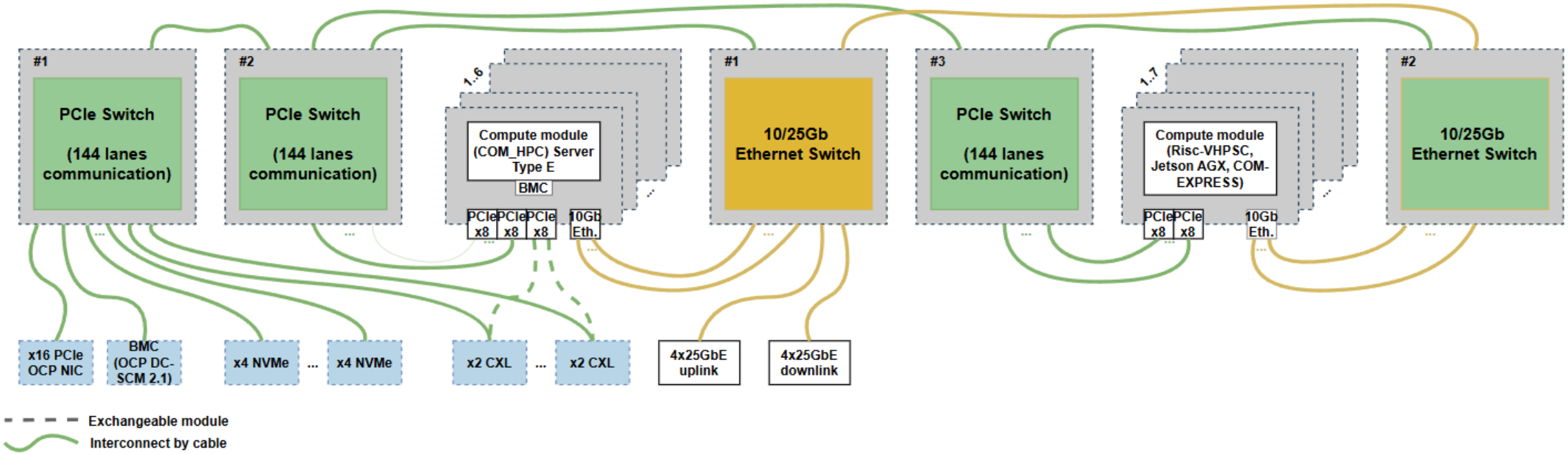
PCB DESIGN

EPFL Cadterr

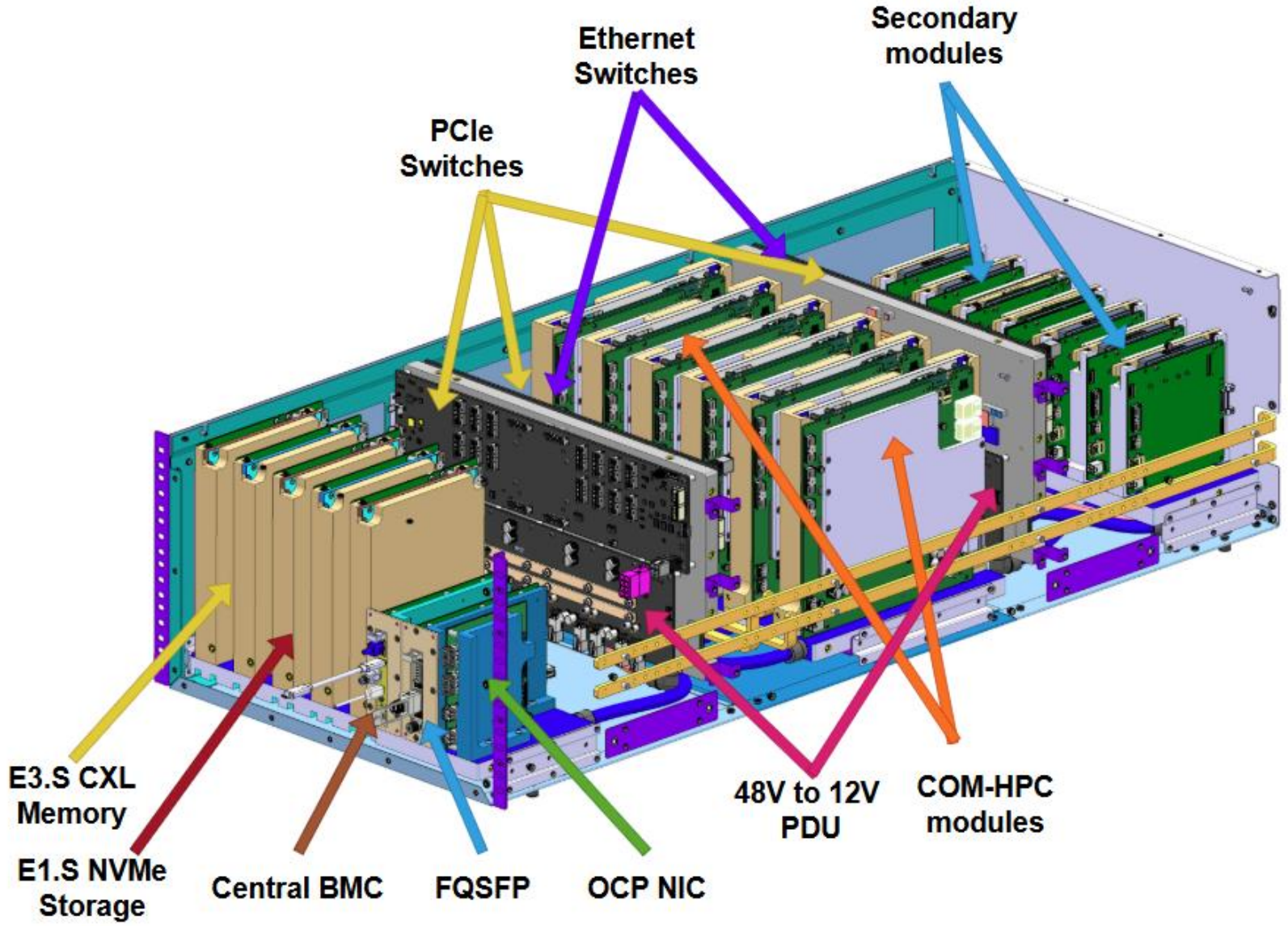
CAPE EMDC (**E**dge **M**icro **D**ata **C**enter)

- Modular
- Composable
- Heterogeneous
- Rackmount

- Workload -> Select resources -> Compose system
- Examples:
  - X86 CPU + GPU + Storage + Network
  - CPU + RISC-V + CXL Memory + Network



CAPE EMDC (**E**dge **M**icro **D**ata **C**enter)



eHPS – embedded High Performance Server Overview

- High performance Edge Computing applications
- Heterogeneous computing with GPU SoC and FPGA SoC hardware accelerators
- Applications with high-bandwidth and low-latency requirements demanding PCIe Gen6/CXL
- Commodity air-cooled 19” rack server housings

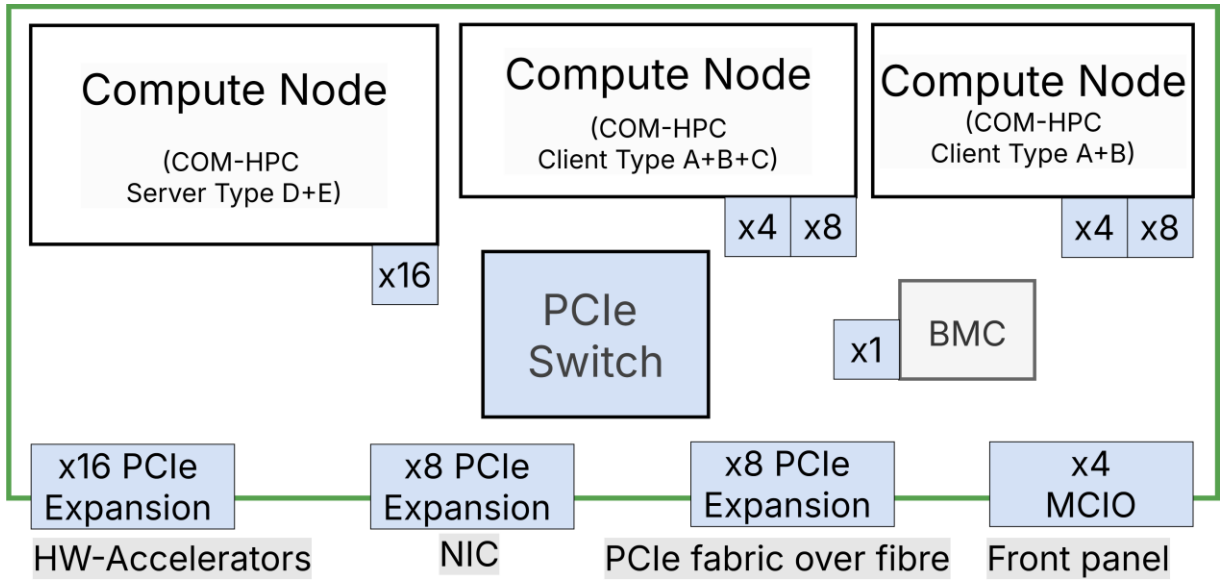


3 height units 19“ server

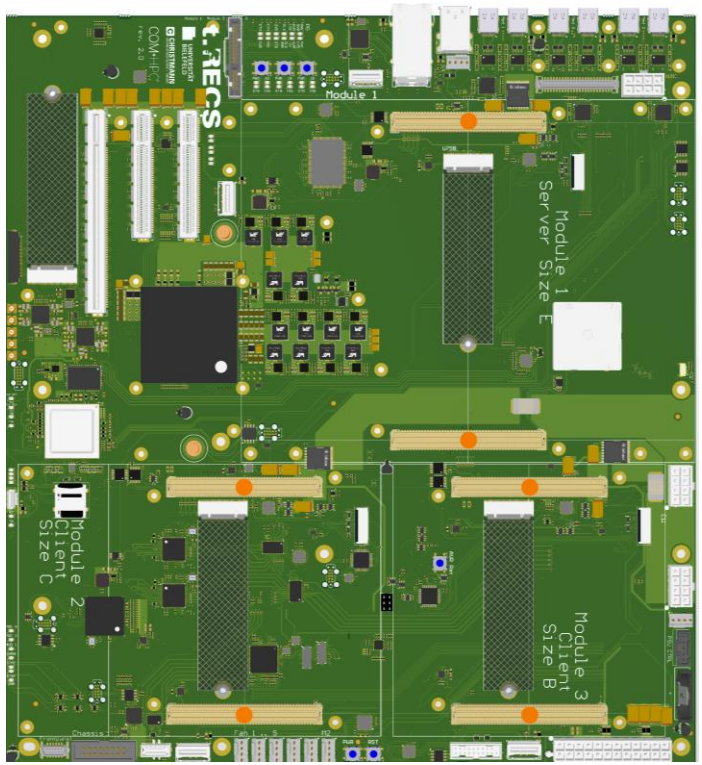


eHPS – embedded High Performance Server Architecture

- 3x COM-HPC compute modules
- PCIe-based communication infrastructure
- 1 central Board Management Controller (BMC)
- DisplayPort / USB 3.2 for edge appliances



eHPS
PCIe-centric
architecture



Rendering of eHPS PCB

What is CXL Compute Express Link ?

CXL focuses on efficiently sharing, caching, coherency, and memory expansion

- Alternate protocol across PCIe physical layer
  - CXL 1.x + 2.0 : PCIe Gen. 5 (32 GT/s)
  - CXL 3.x : PCIe Gen. 6 (64 GT/s)
  - CXL 4.x : PCIe Gen. 7 (128 GT/s)
- Topologies are CPU centric

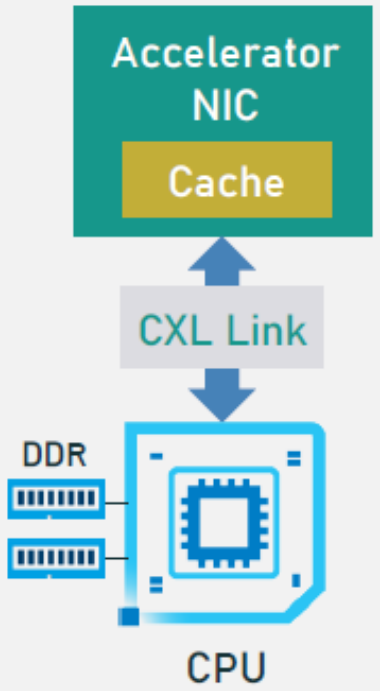
Protocols:
CXL.io: Similar to PCIe, non-coherent
CXL.cache: Extends coherence, allows to cache host memory, reducing latency
CXL.mem: Shared, coherent memory space, enabling mem expansion and pooling

CXL Device Types

Type 1 Device
Caching Devices/Accelerators

Usages:
- PGAS NIC
- NIC atomics

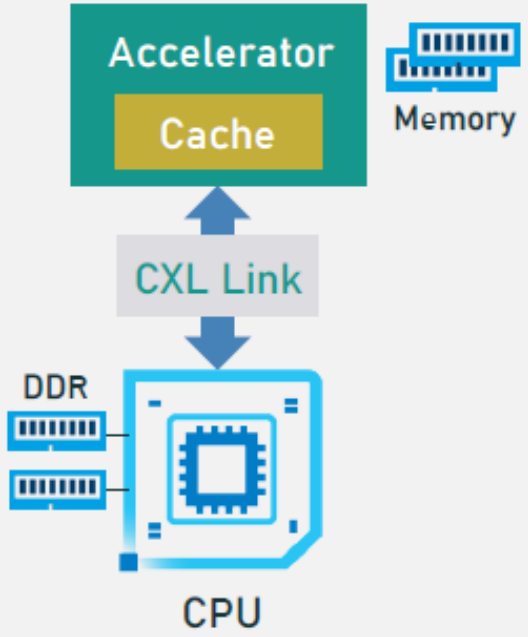
Protocols:
- CXL.io
- CXL.cache



Type 2 Device
Accelerators with Memory

Usages:
- GPU
- FPGA
- Dense Computation

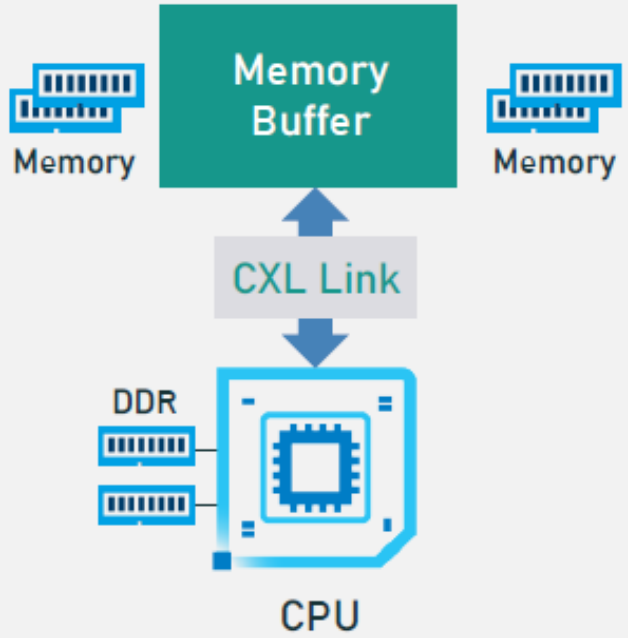
Protocols:
- CXL.io
- CXL.cache
- CXL.memory



Type 3 Device
Memory Buffers

Usages:
- Memory BW expansion
- Memory capacity expansion
- 2LM

Protocols:
- CXL.io
- CXL.mem



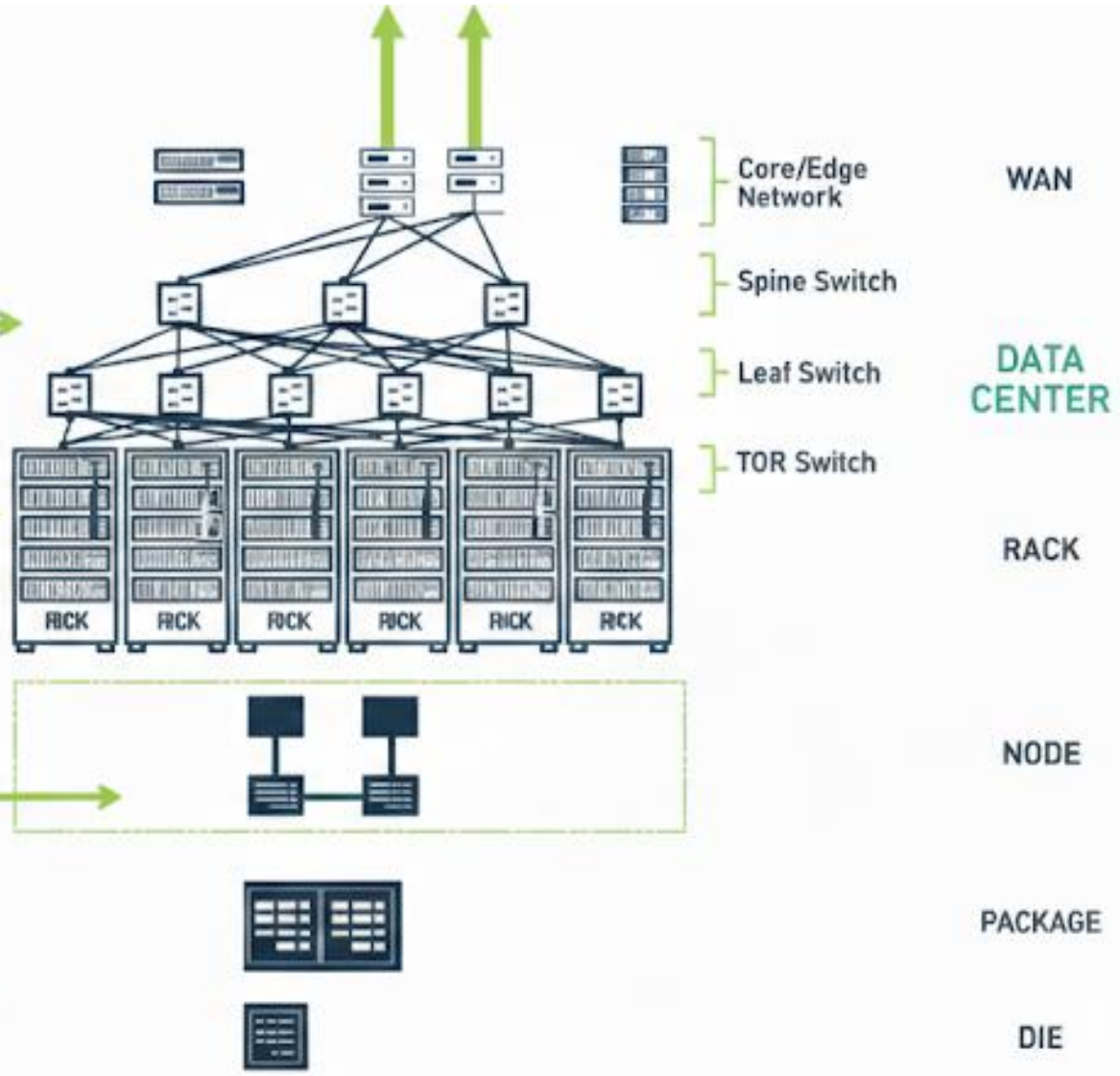
CXL Evolution

CXL 4.0
- 128 GTs
- CXL Bundled Port
- Memory RAS enhancements

CXL 3.x
- 64 GTs
- Improved security
- Composable Fabric growth for disaggregation of memory & accelerators
- Memory sharing with back invalidate

CXL 2.0
- Multiple nodes inside a Rack/Chassis supporting pooling of resources
- Switch-based memory pooling
- Global Persistent Flush (GPF)
- Link-level encryption

CXL 1.1
- 32 GTs
- Single Node coherent interconnect
- Load/store semantics
- Hardware enforced cache coherence

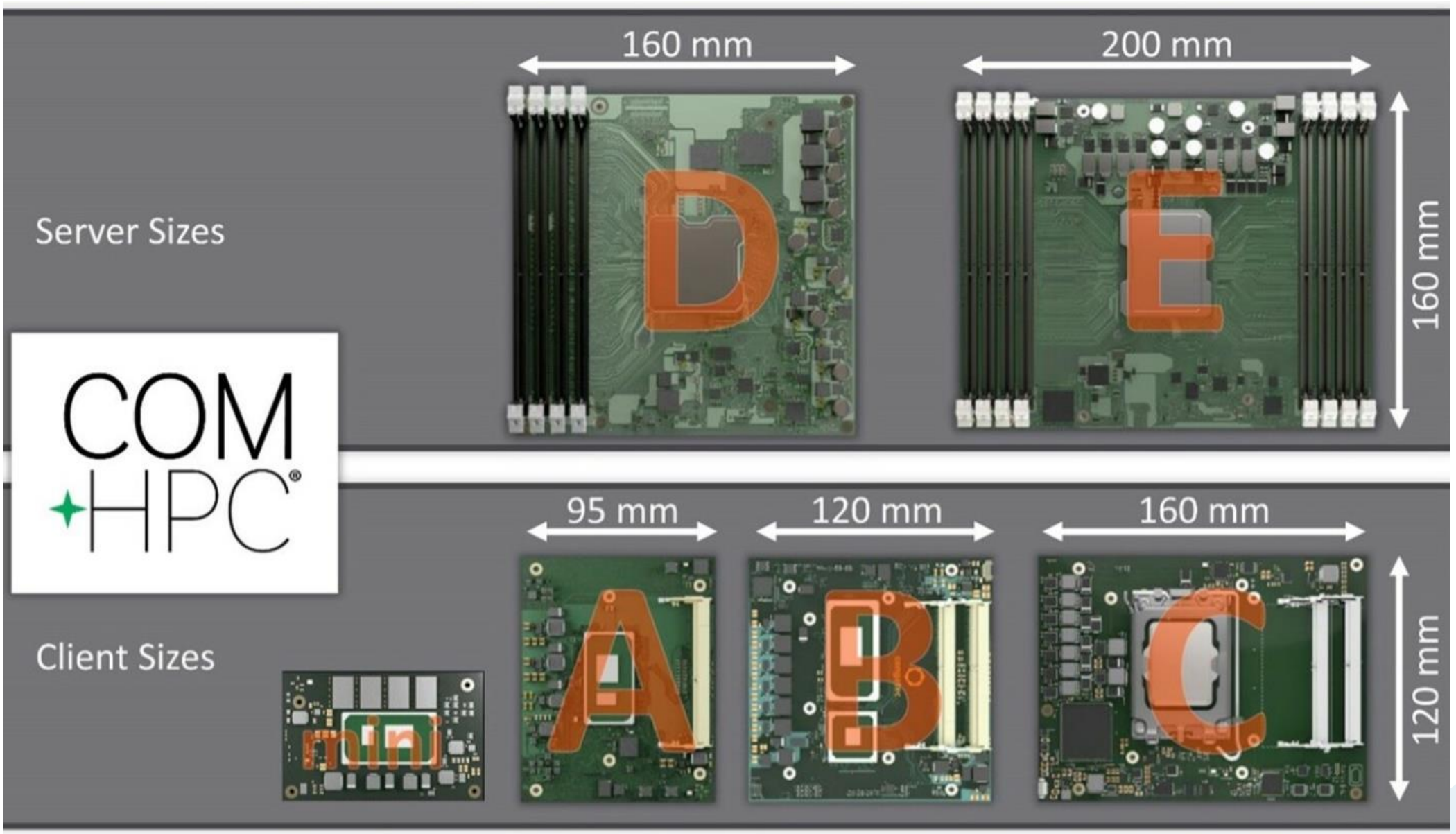


Edge Cloud hardware standards

- Standardisation is for Edge Cloud hardware is mostly pending, but required to enable industry growth
- CAPE enables standardization based on Computer on module (COM) standards
- COM-HPC has been established by PICMG, building on the legacy of from factors for embedded applications
- COM-HPC workgroup members include Intel, Kontron, Congatec, Seco, Samtec Adlink Technology, etc.

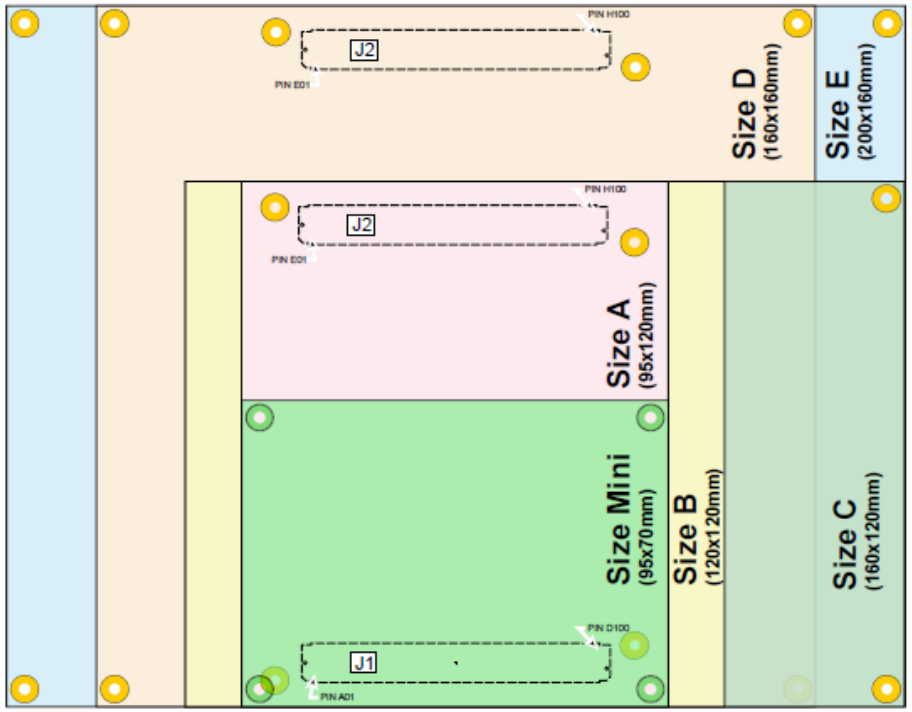


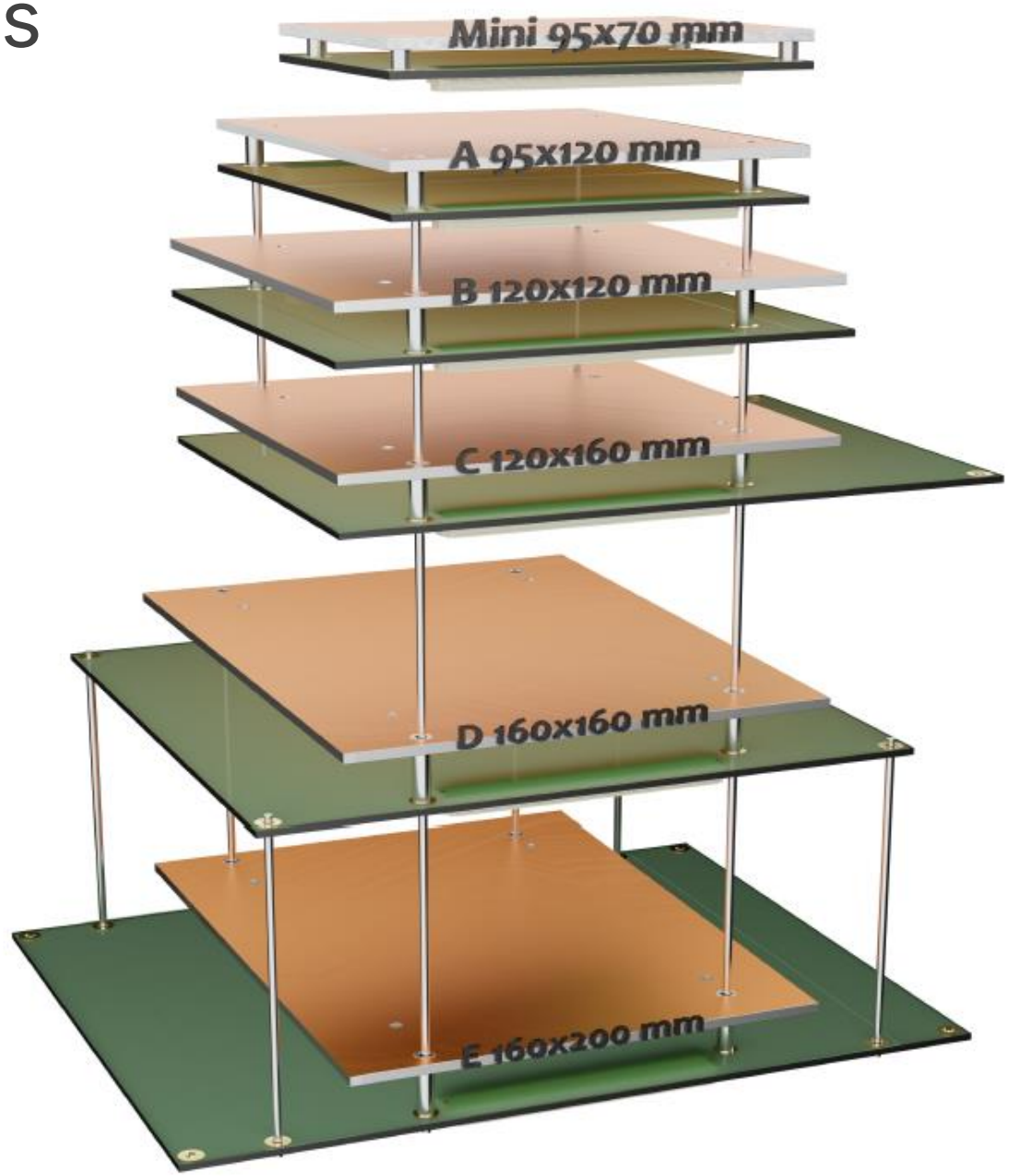
COM-HPC – Form factors



COM-HPC – Form factors

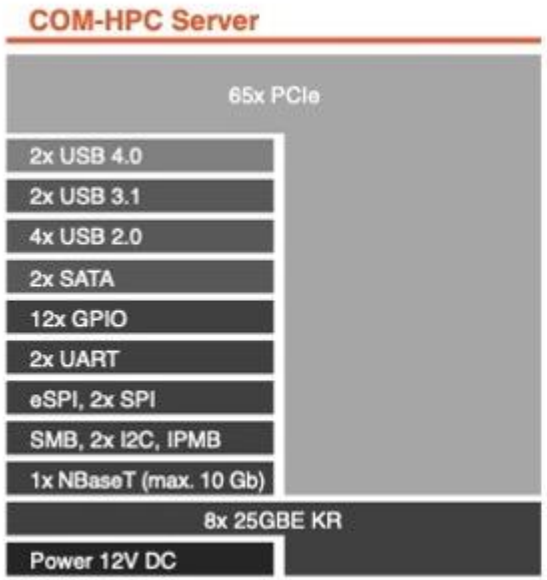
- Six form factors defined
- Client: Size A + B + C
- Server: Size D + E
- Mini: 95 mm x 70 mm

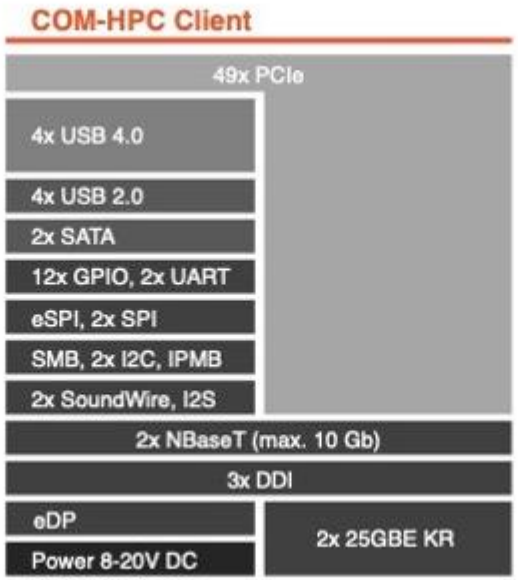


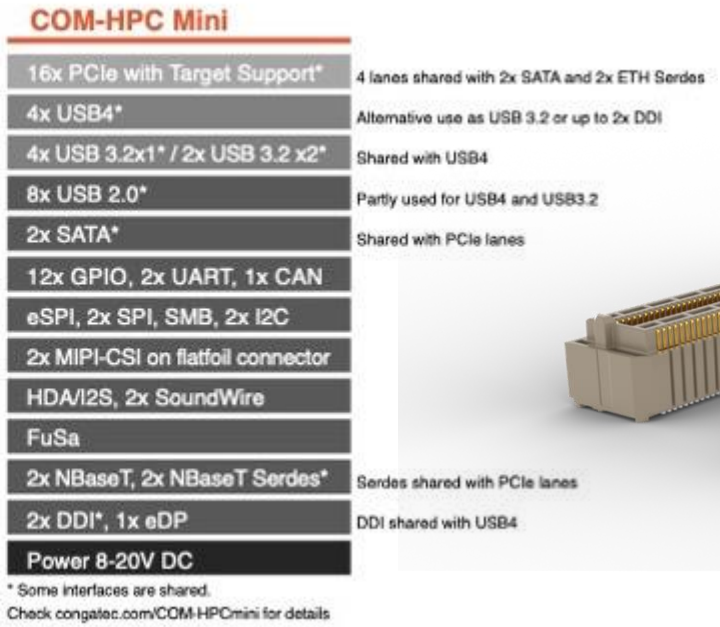


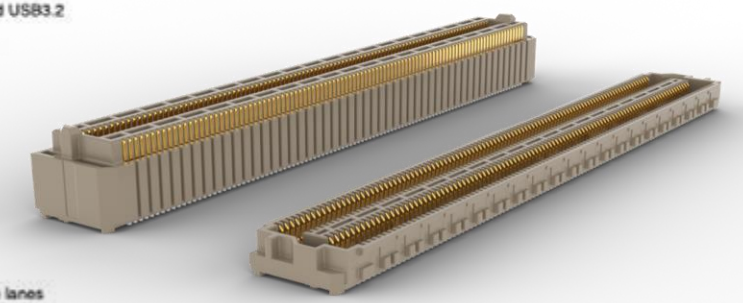
COM-HPC – Connectors

- Up to 65 PCIe Lanes (supporting CXL), providing massive communication bandwidth
- Up to 350 W of input power
- 56 Gbps NRZ/112 Gbps PAM4 performance









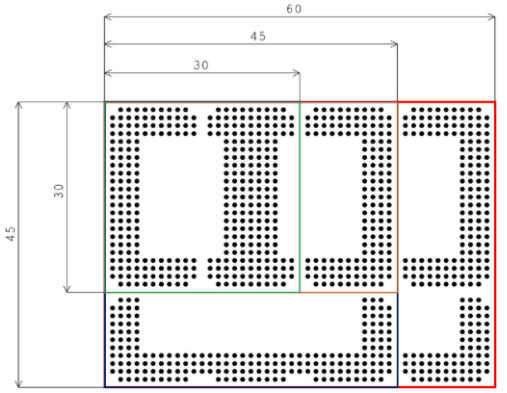
oHFM Summary

UNIVERSITÄT BIELEFELD

SGeT STANDARDIZATION GROUP FOR EMBEDDED TECHNOLOGIES

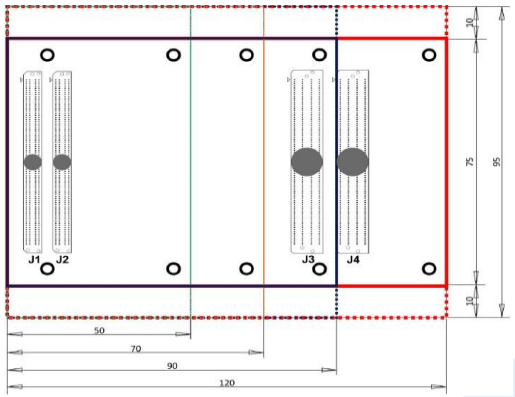
oHFM

"oHFM – Open Harmonized FPGA Module is the world's first open standard specifically designed for FPGA and SOC-FPGA modules"



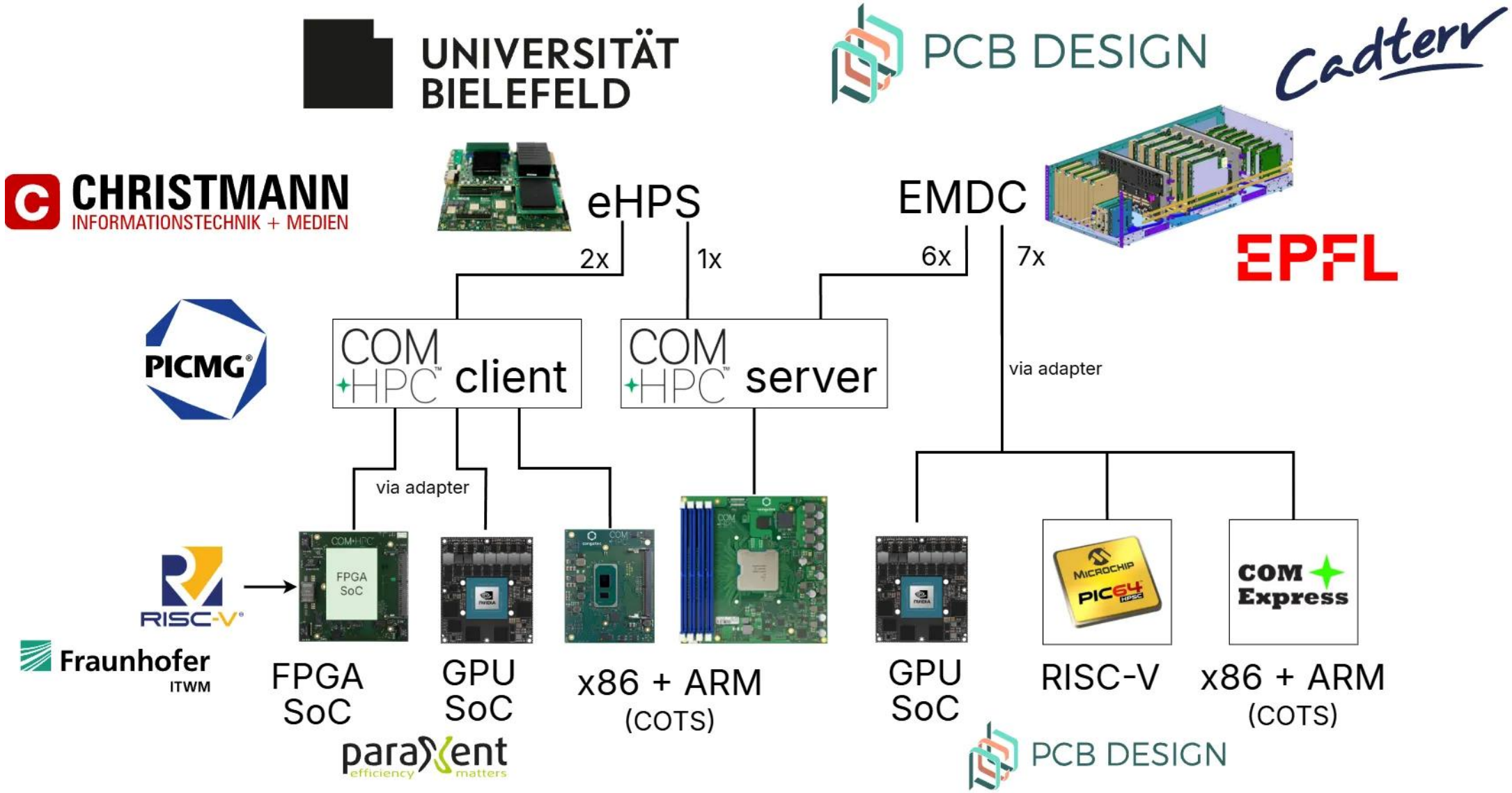
oHFM.s (solderable)

- Architectural gap: FPGA manufacturers prefer different form factors than COM-HPC
- oHFM fills this gap
  - Tailored specifically for FPGA and SoC-FPGA architectures, optimized for flexible I/O banks and multi-gigabit transceivers
  - 4 scalable form factors (S,M,L,XL)
  - Modules with either connectors or solder-bumps



oHFM.c (connector)

CAPE Module overview



Overview of COM-HPC Modules

COM-HPC Type	Architecture details	PCIe/CXL
Client A	CPU/x86: Intel Panther Lake-H (Core Ultra 3)	Gen 5
Client B	GPUSoC: NVIDIA AGX Orin	Gen 4
Client B	GPUSoC: NVIDIA AGX Thor	Gen 5
Client B	FPGASoC: AMD Versal Premium Gen. 2	Gen 6 / CXL 3.1 (full)
Client B	CPU/x86: AMD EPYC Embedded 2005 (Zen 5)	Gen 5
Server D	CPU/RISC-V: VMicrochip PIC64-HPSC1000 (SiFive Intelligence X280)	Gen 3 / CXL 2.0 (non-standard)
Server E	CPU/x86: AMD EPYC Embedded 9006 “Venice” (Zen 6c)	Gen 6 / CXL 3.1 (full)
Server E	CPU/ARM: Ampere Altra / Altra Max (ARMv8.2, ARM Neoverse N1, 128 Cores)	Gen 4
Server E	CPU/x86: Intel Diamond Rapids / Panther Cove (Intel Xeon 7 P-Core)	Gen 6 / CXL 3.1 (full)



AGX Orin Devkit



Adapter:
COM-HPC <> AGX Orin/Thor

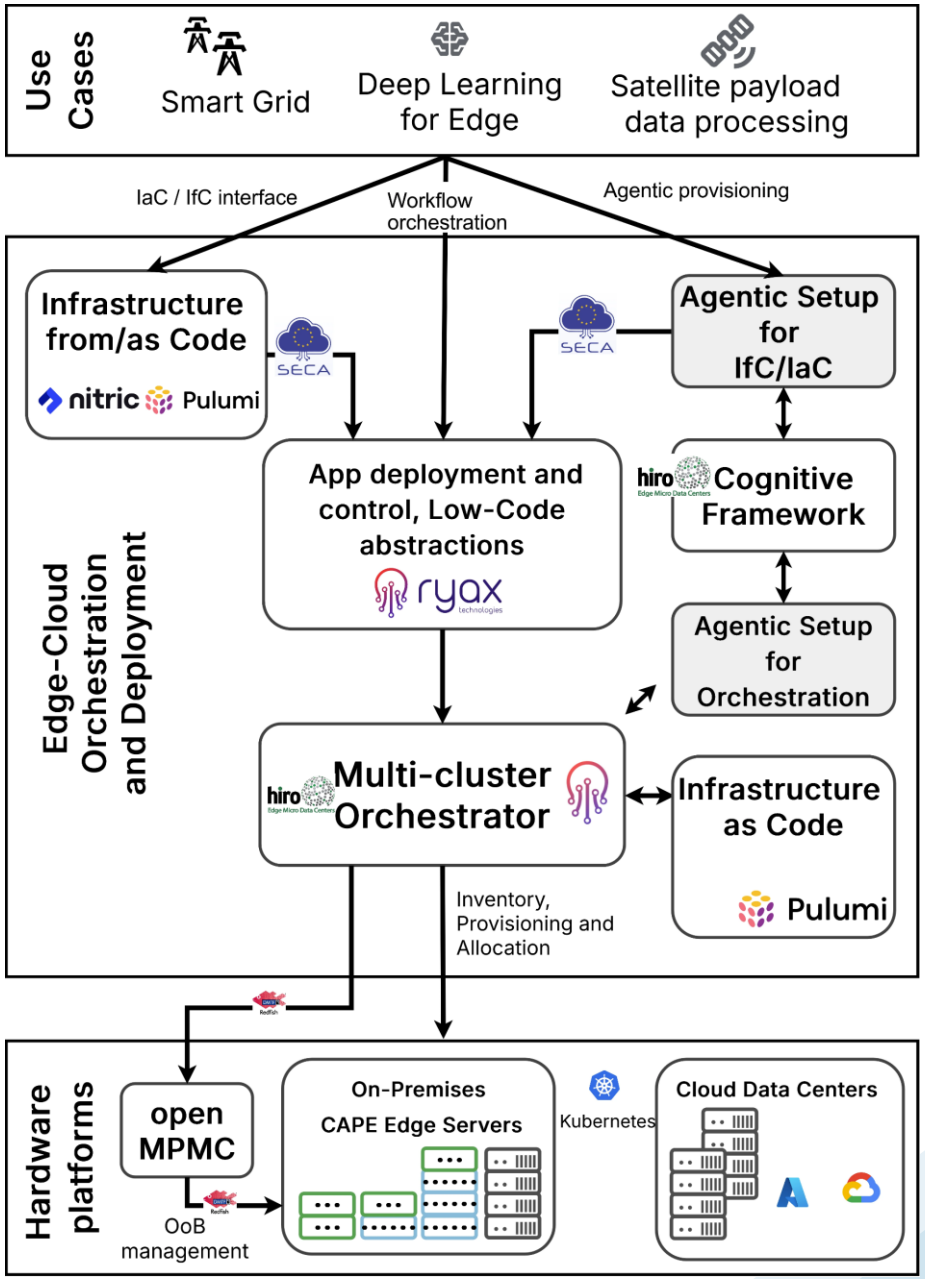
Software Architecture

Features

- Transparent deployment across edge-cloud continuum
- Heterogeneous computing resources
- Three different programming and deployment models
- Fully open source stack including (Sovereign European Cloud API)
- Redfish as interface towards hardware management

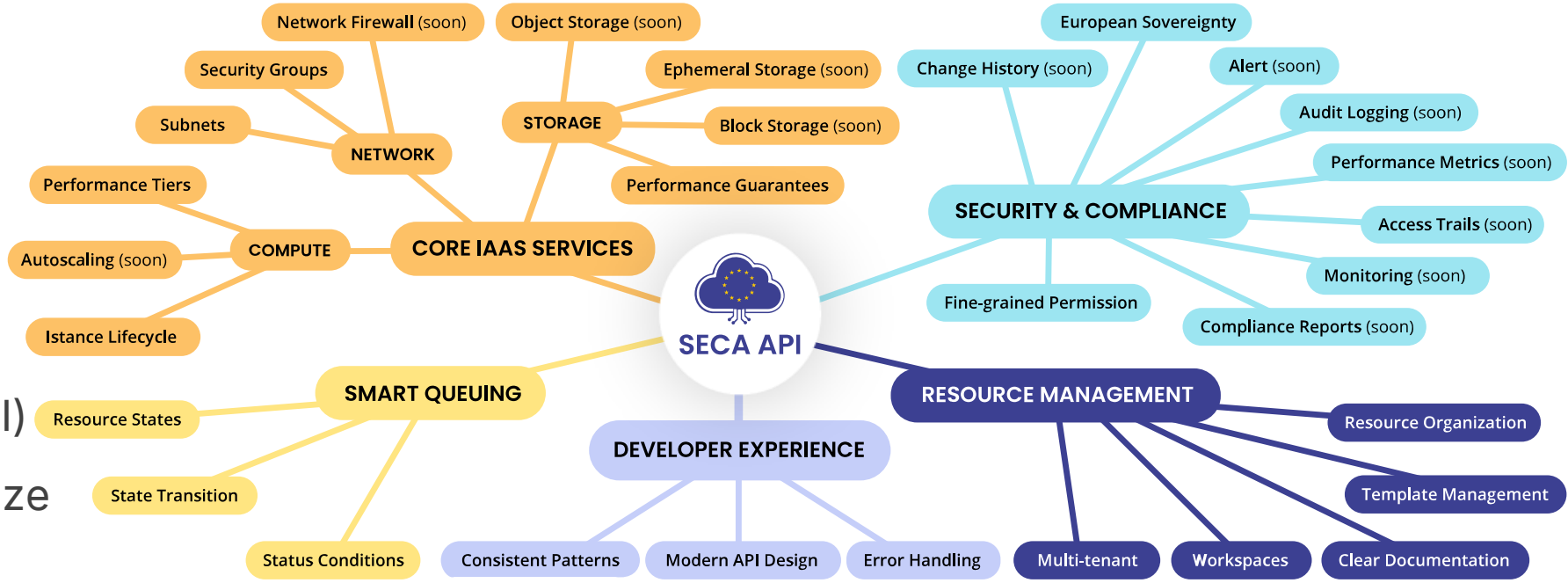
Components / tools

- Infrastructure from/as Code (IfC/IaC): Automated Software Testing
- Ryax: Deployment, resource selection and data center management
- Orchestrator: manages inventory, configures network infrastructure and compute nodes
- openMPMC (Multi-Platform Management Controller): low-level server management



SECA API Overview

- Provide **abstraction** for deploying and **provisioning resources** in the continuum
- Adoption of an EU based API specification: **SECA API** (Sovereign European Cloud API)
- European Initiative to standardize Cloud providers
- Open industry standard
- Safeguard against vendor lock-in
- CAPE extends and contributes back the support to specific resources like GPUs, FPGAs, CXL



https://spec.secapi.cloud/
https://github.com/eu-sovereign-cloud/spec

Redfish Overview

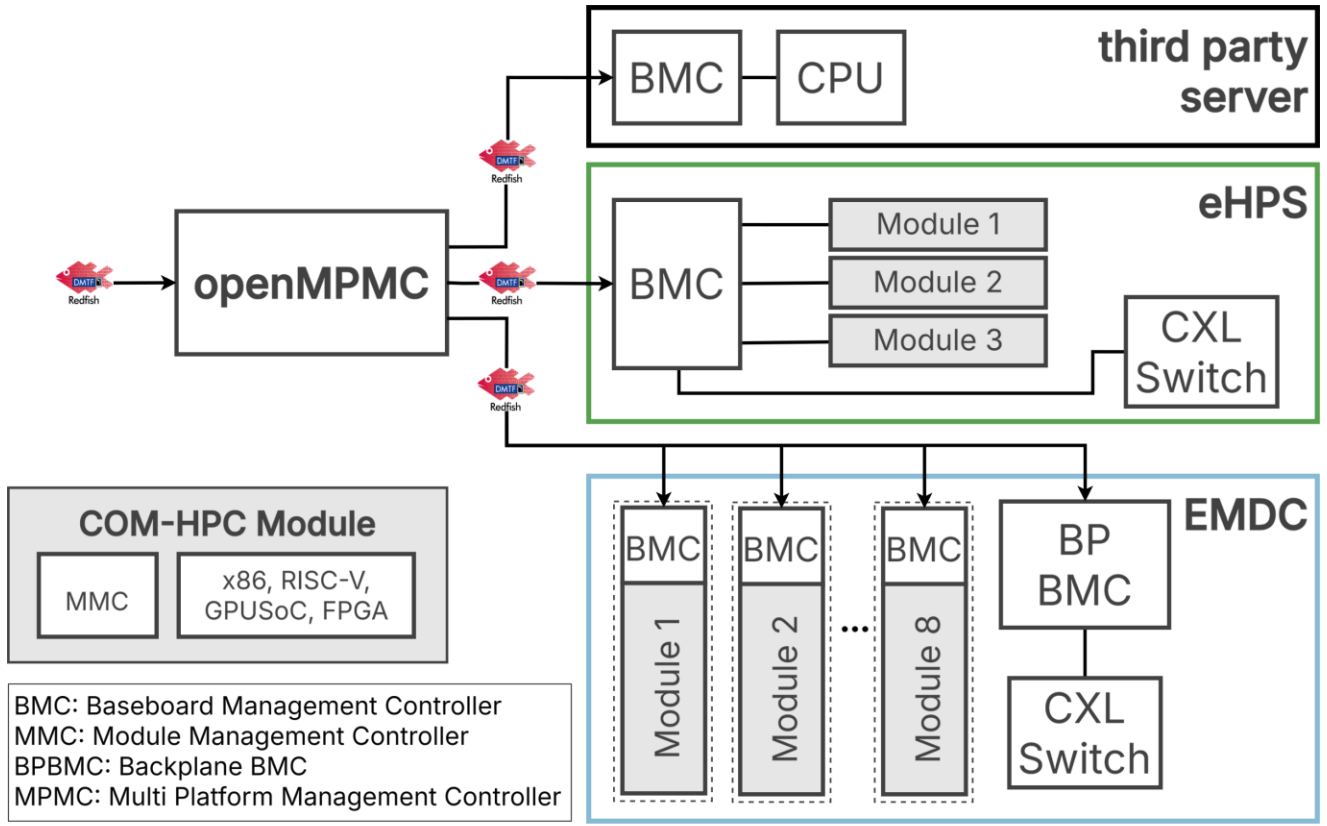
Redfish in general

- Open DMTF standard for vendor-neutral hardware management
- REST API with JSON over HTTPS; BMC access independent of the host OS

Redfish in CAPE

- openMPMC links BMCs and exposes one Redfish API to HIRO's provisioning agent
- Inventory, power, boot, virtual media and sensor monitoring across sites
- Web GUI for out-of-band management of many servers
- Multi-tenant, open source, agents for remote sites

CAPE target architecture



OFEIA

- Open Federated Edgecloud Infrastructure Alliance
- Linux Foundation collaboration
  - Further collaboration in progress
- OFEIA workshops at HIPEAC conferences
- LinkedIn page



THE LINUX FOUNDATION | Education

https://www.linkedin.com/company/open-federated-edgecloud-infrastructure-alliance/

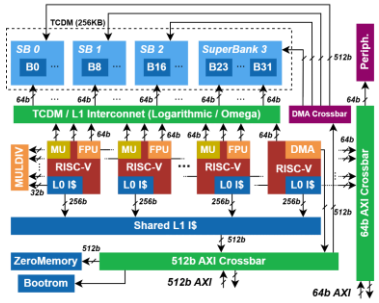
Use Case Overview



Edge AI

Fraunhofer ITWM

- Surveillance on RISC-V / FPGA with accelerators
- Optimize deep learning algorithms
- **Low-latency & high throughput**

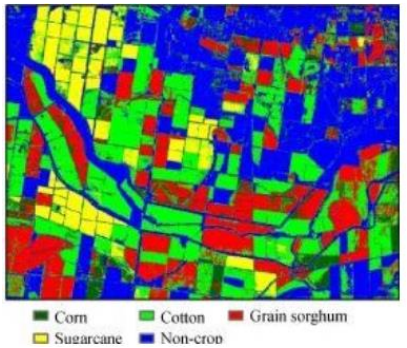




Satellite Payload Processing

PCB DESIGN

- Data-(pre-)processing
- Data clustering & fusion
- **High throughput & energy efficiency**

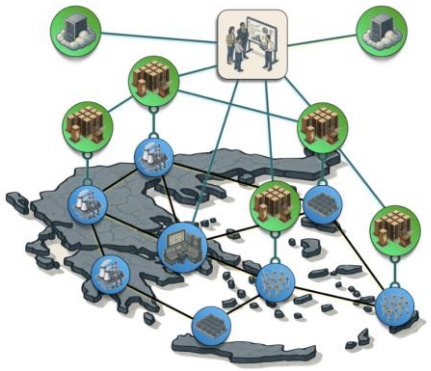




Smart Grid

ipto

- Real-time anomaly detection
- Energy forecasting & market prediction
- **Robustness & real-time capabilities**



Summary

- CAPE is involved in multiple standardization efforts
  - COM Express 3.2
  - COM-HPC 1.3
  - oHFM 1.1
  - HPM 2.0/MetaHPM
  - COM-HPC 1.4
- CAPE is using and actively driving open-source efforts
  - SECA adopted as software-API baseline
  - Using DMTF Redfish for hardware management
  - COM-HPC and oHFM implemented for hardware modules

CAPE

European Open
Compute **A**rchitecture
for **P**owerful **E**dge

The CAPE project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101189899.

Making Composable Edge Infrastructure Interoperable: Open Standards and Interfaces in CAPE

Jens Hagemeyer, jhagemey@techfak.uni-bielefeld.de