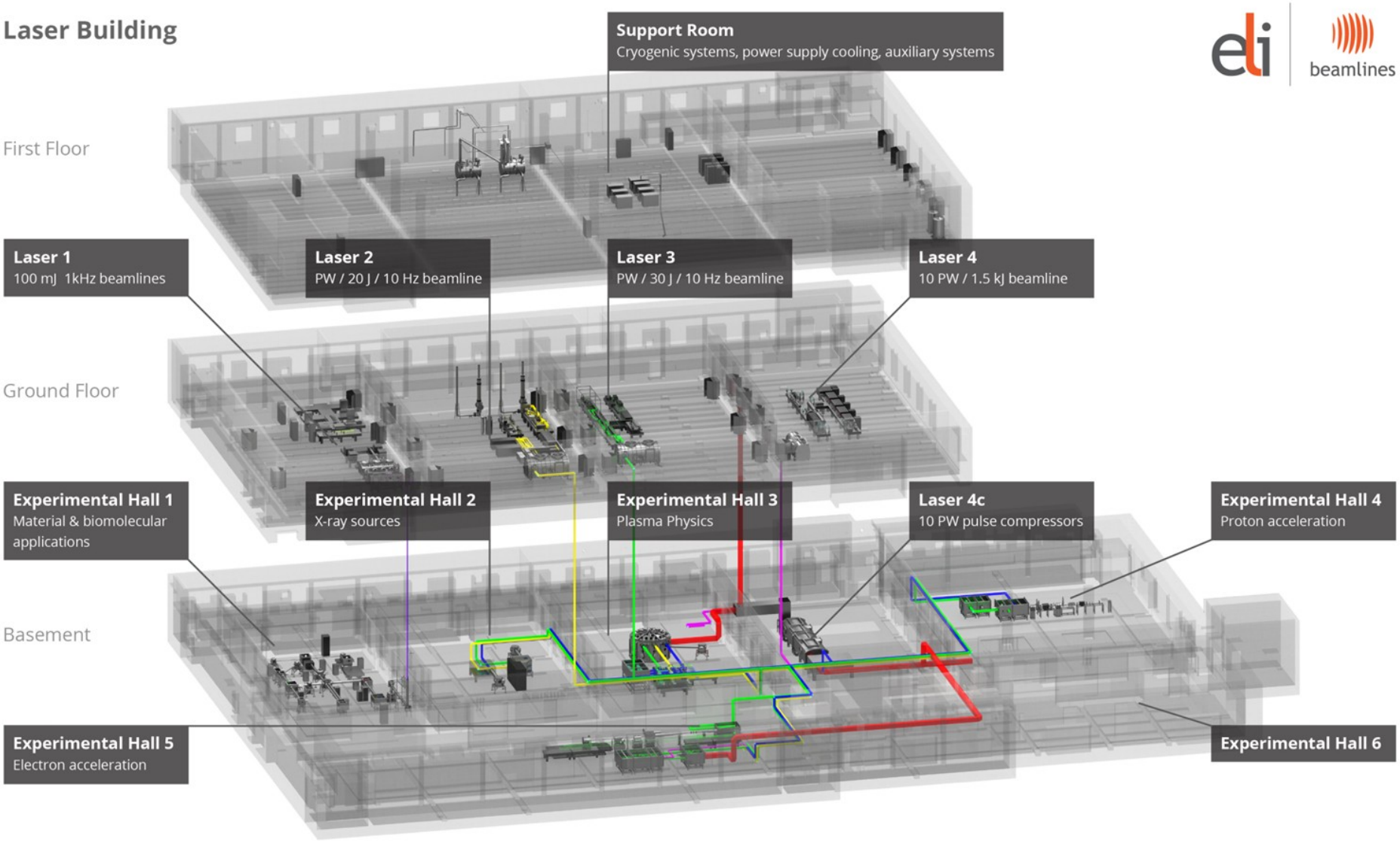
Ultra Fast Data Acquisition System in ELI Beamlines

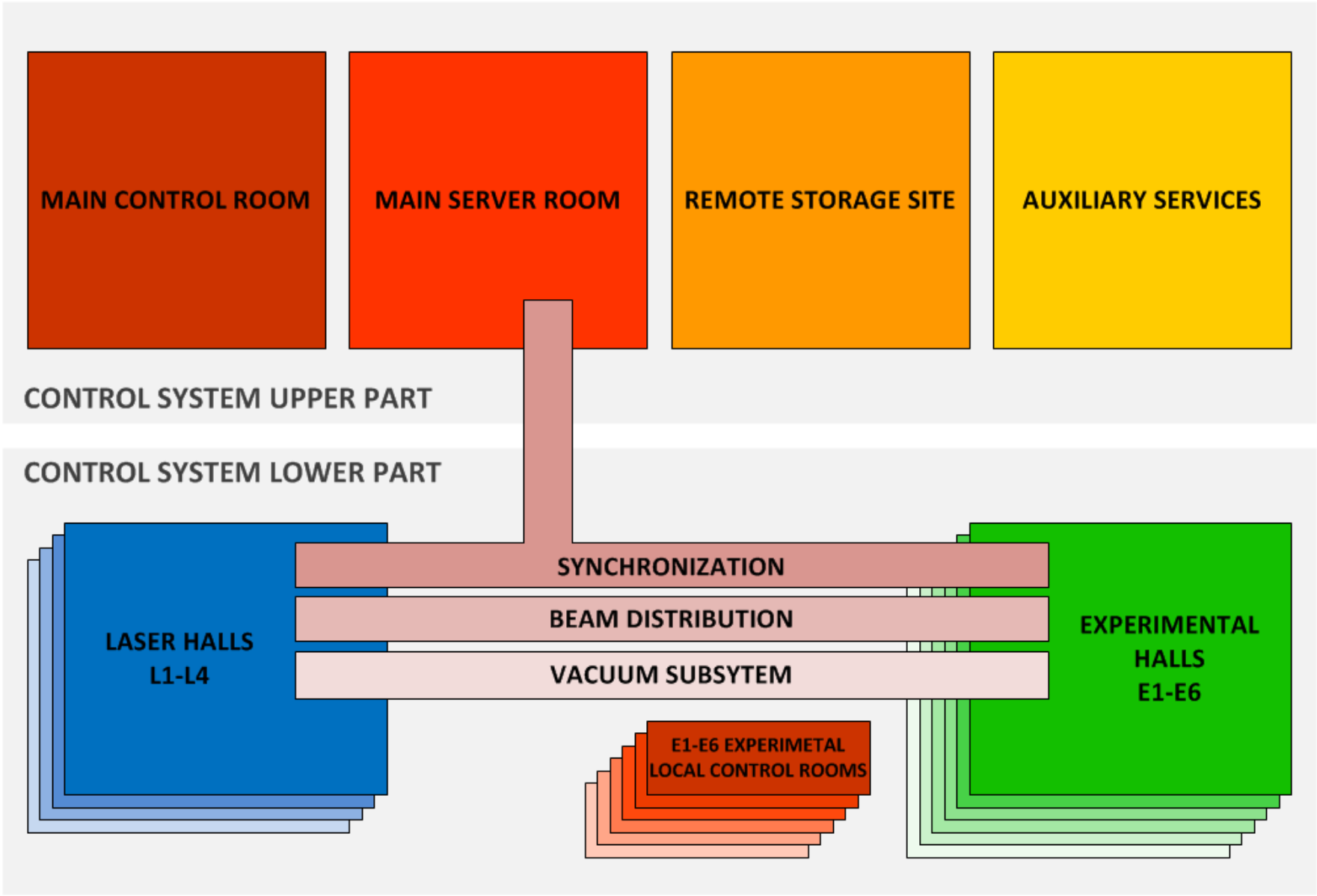
Pavel Bastl ELI Beamlines/Institute of Physics, Prague, Czech Republic

FZÚ
EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ MLÁDEŽE A TĚLOVÝCHOVY

ELI Bemlines



CS Structure

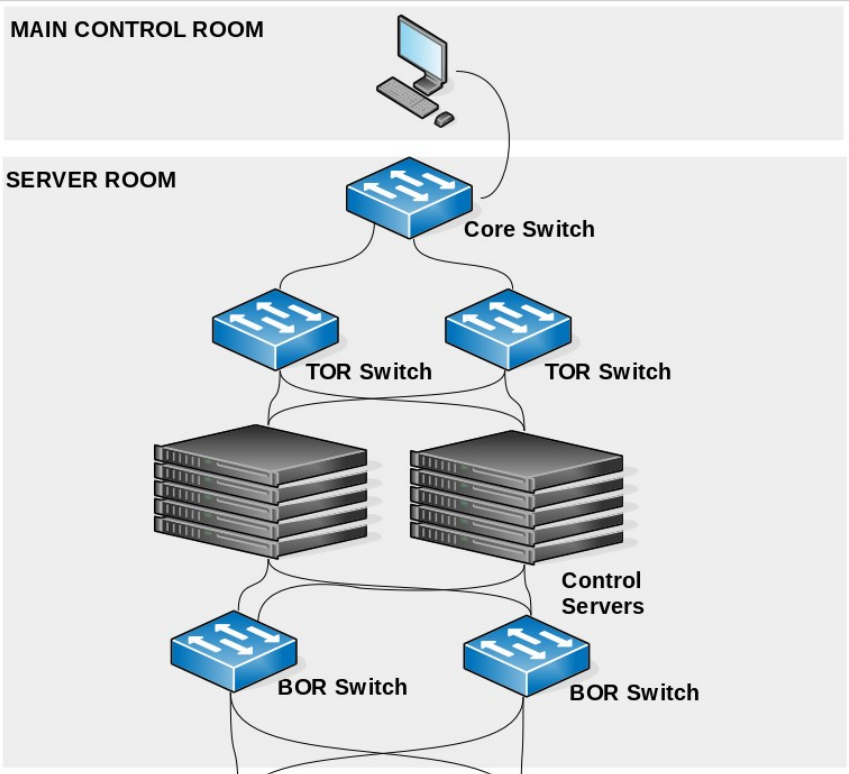


EUROPEAN UNION
European Structural and Investing Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION
YOUTH AND SPORTS

CS Overview

TOP LEVEL



Main Control room

CORE Switch – Cisco Nexus 7700, 10/40Gb/s, 100Gb/s ready

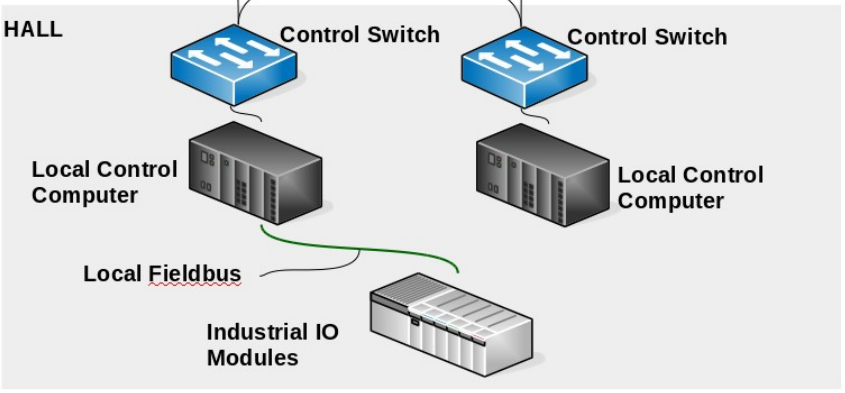
TOR Switch – Cisco Nexus 5672, 10/40Gb/s

Top level control – Server Room

- **Control servers** – Butch of 10 servers
- **Lenovo Systems x3650m5** 2U servers
  - 24 cores
  - 256GB RAM
- **Virtualization** – Private Cloud

BOR Switch – Cisco Nexus 56128, 10/40Gb/s

LOCAL LEVEL



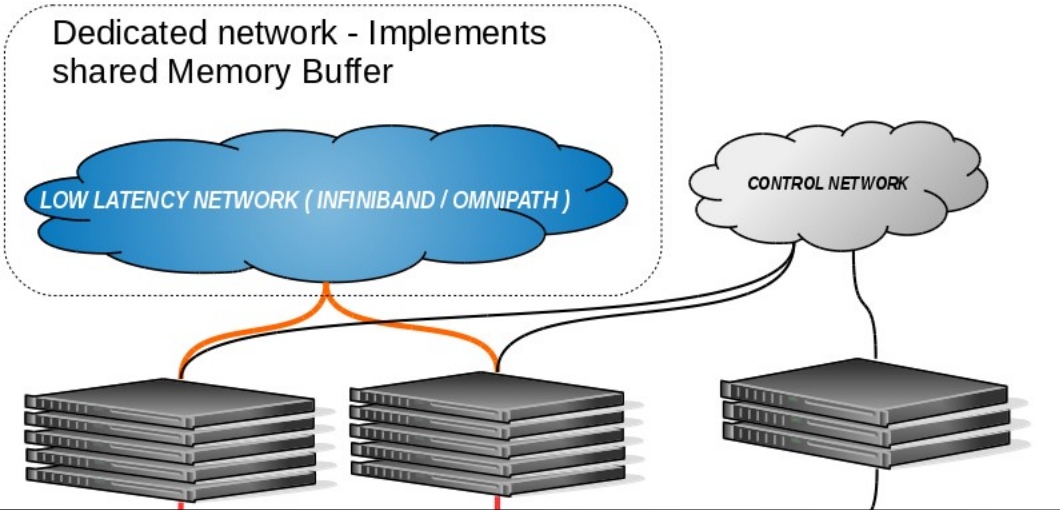
CONTROL Switch – Cisco Catalyst 2960X, 1/10Gb/s

Local level control – Halls&Plant rooms

- **Industrial control**, undemanding applications
- **Advanced control**, challenging applications, with high demands on data rates

DAQ Overview

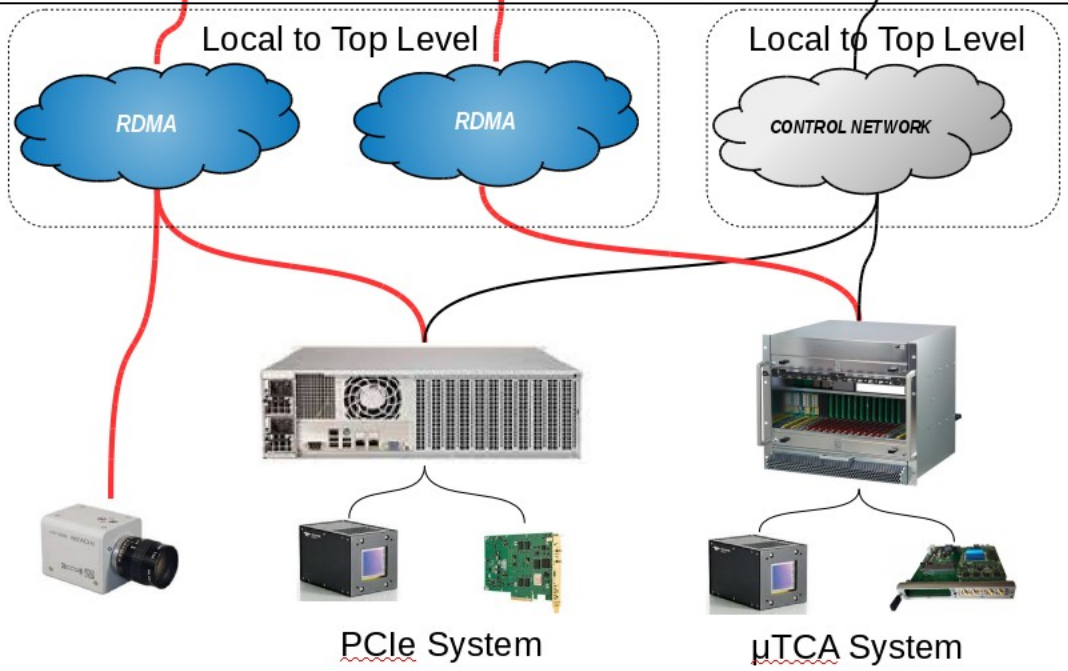
TOP LEVEL



Top Level of DAQ System

- 2x Blade Server
  - 10Gb/s ethernet (mezzanine) each
  - 56Gb/s Infiniband (mezzanine) each
  - 24 cores each Blade
  - 768GB RAM each Blade
  - 4x PCIe slot (2 slots x8, 2 slots x16)

LOCAL LEVEL



Local Level of DAQ System

- Local PCIe Server
  - 4x 10Gb/s ethernet with RDMA
  - 24 cores
  - 128GB RAM each
  - 10x PCIe slots x8
- µTCA.4
  - 12x AMC, 12x RTM
  - 4x 10Gb ethernet to RDMA
  - MCH-PHYS-80 PCIe on Fat-Pipes, PCIe connection with PCIe local server through optical fibers

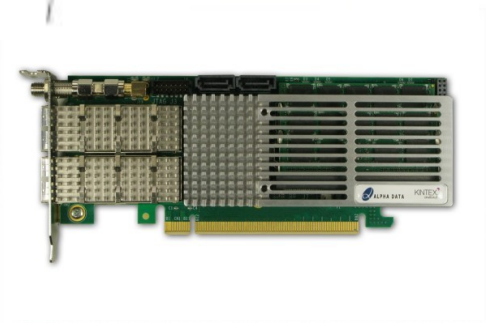
DAQ Top Level

Blade Server With PCIe Slots

Infiniband and Ethernet Switches are integrated inside the chassis.

+

NIC – Our 'standard'
- 2x 10GBASE-X
- RDMA Support



FPGA XCKU060
- 2x QSFP
- SDAccell support
- CAPI support
- 8GB DDR RAM

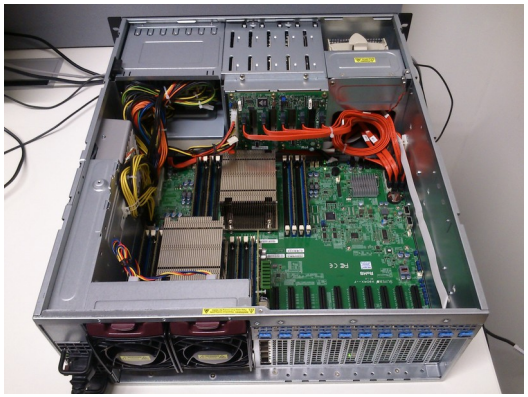
FPGA+NIC
- 1X QSFP (4xSFP+)
- Connect-X NIC Offload chip
- RDMA Support
- XCKU060
- 2GB DDR RAM

DAQ Local Level

PCIe System

PCIe Local Server

- 24 core
- 128GB RAM
- 10x PCIe x8



NIC AOC-STG-b4S

- 4x 10GBASE-X
- RDMA support

FMC Carrier

Kintex UltraScale XCKU085

- Acquisition and local processing
- VITA 57.4 compatible with 57.1
- 16GB DDR RAM
- 1x SFP+

µTCA.4 System

µTCA.4 Chassis

- 12 AMC
- 12 RTM
- WR Support
- MCH-PHYS-80
- PCIe interconnection

FMC Carrier

Artix/Kintex XC7A200, XC7K325

- Acquisition and local processing

Compatible RTM Module

- 8x SFP+

NIC Vadatech AMC211

- 2x 10GBASE-X

Online Data Processing

REASONS FOR ONLINE DATA PROCESSING

- Reduction of Data to be stored → online data compression and selection
- Online results for users → data processing based on OpenCL
- Control system feedback → online results for control system feedback

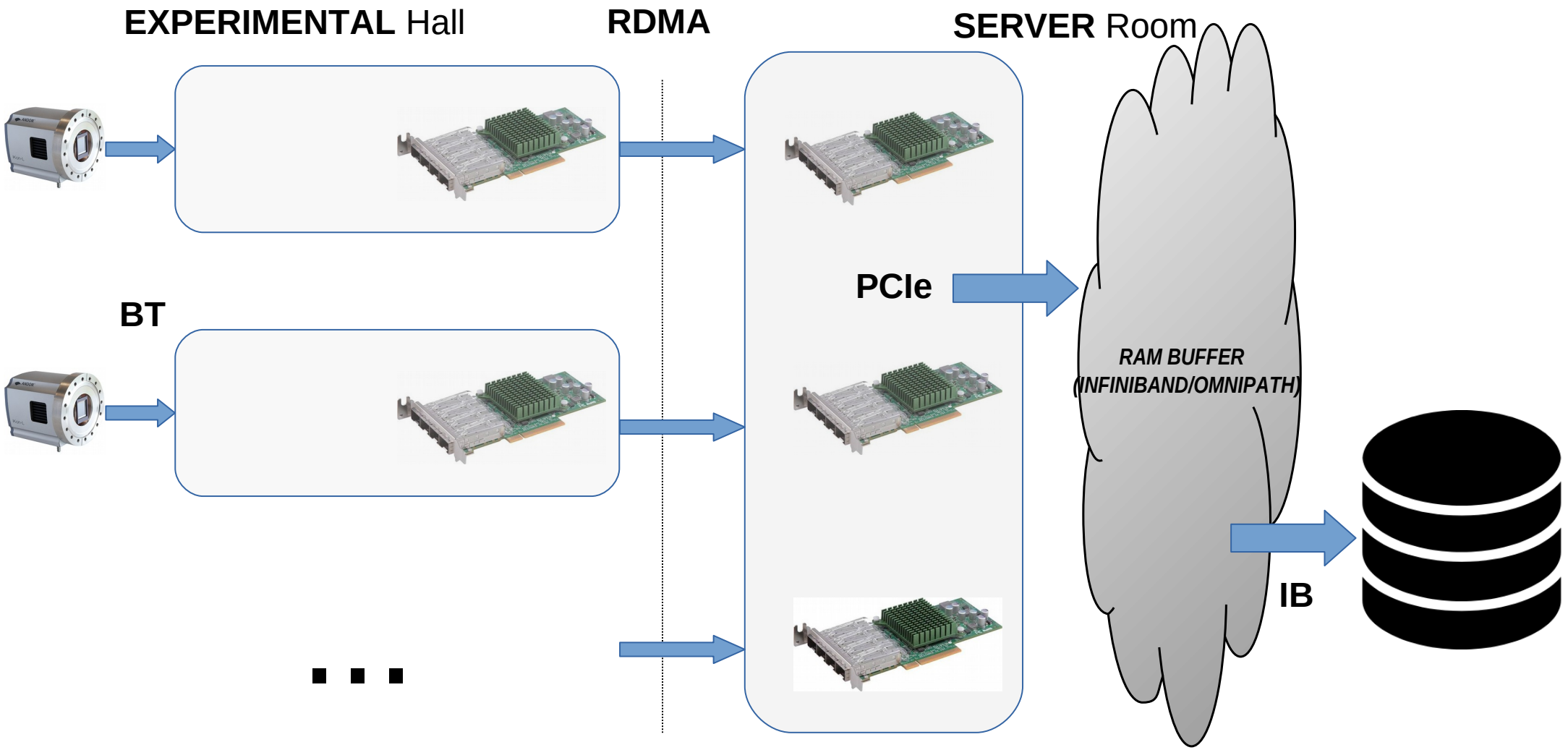
ONLINE DATA PROCESSING CASES

- Data storage
  - Distributed sources
  - Data saved and shared in RAM buffer

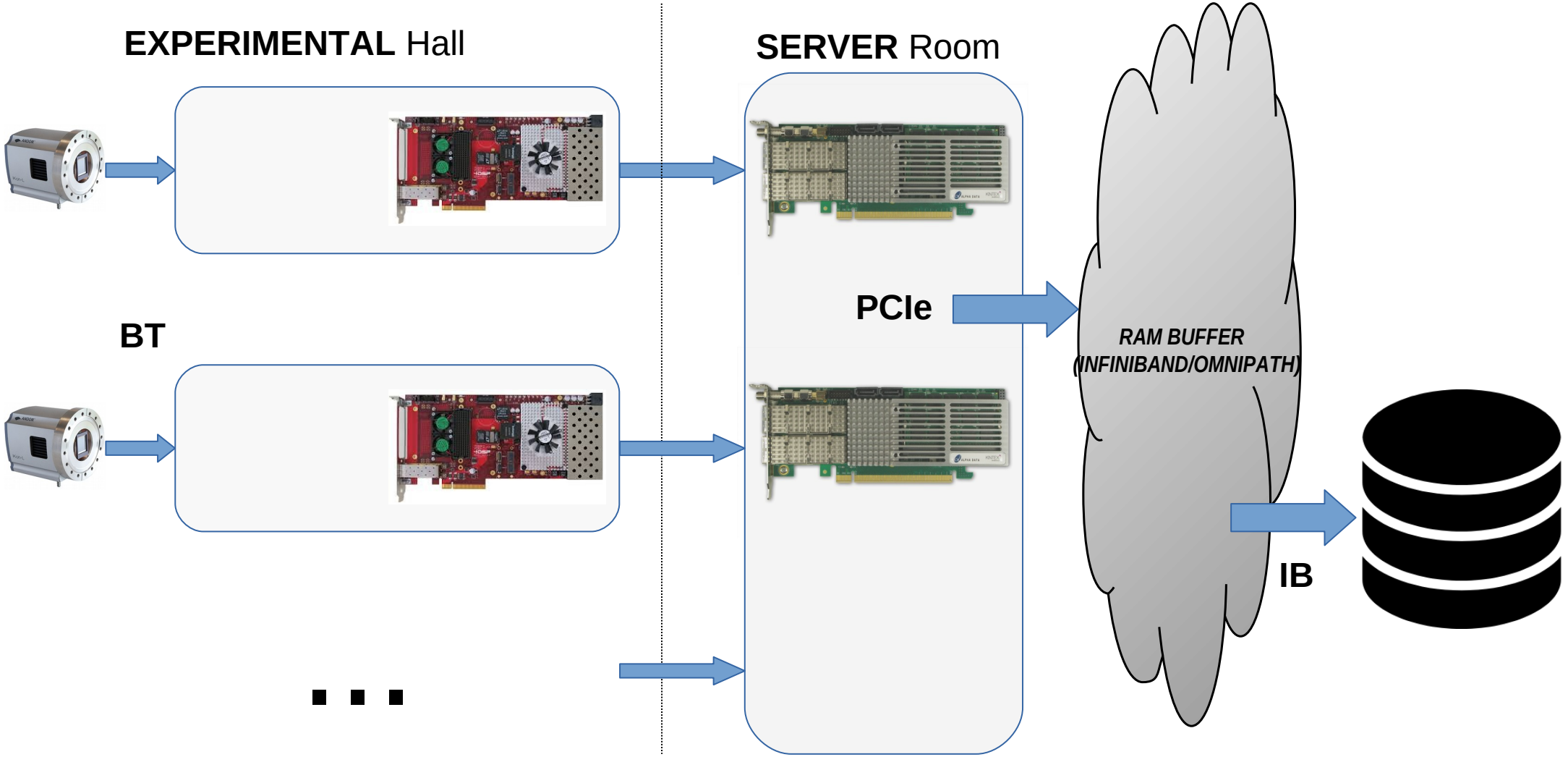
- Control system feedback
  - Dedicated communication line
  - Communication optimized for feedback

Online Data Processing

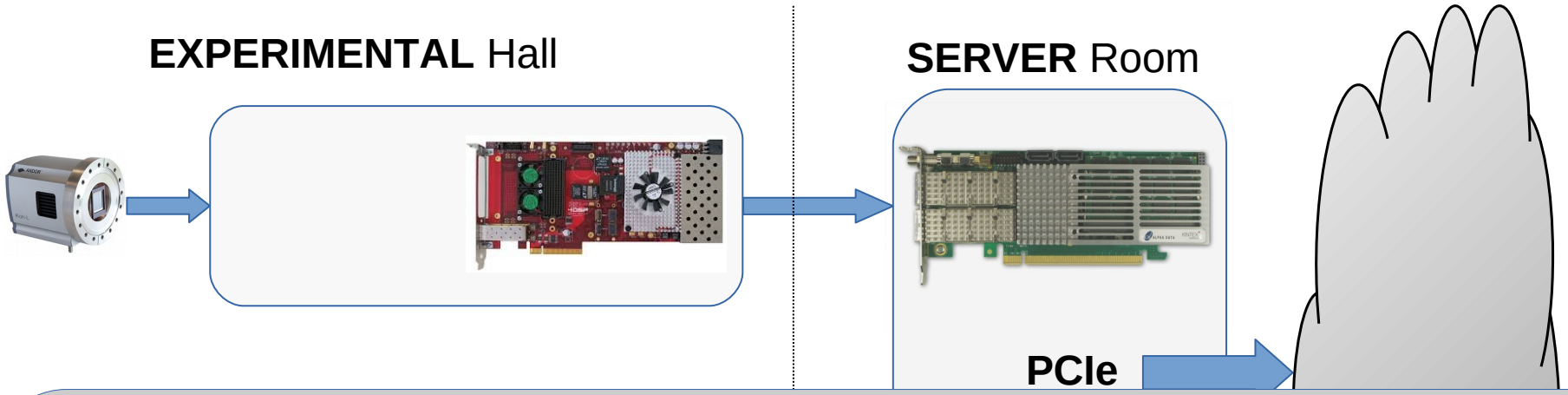
GENERAL SCHEME



Data Reduction scheme



Data Reduction scheme



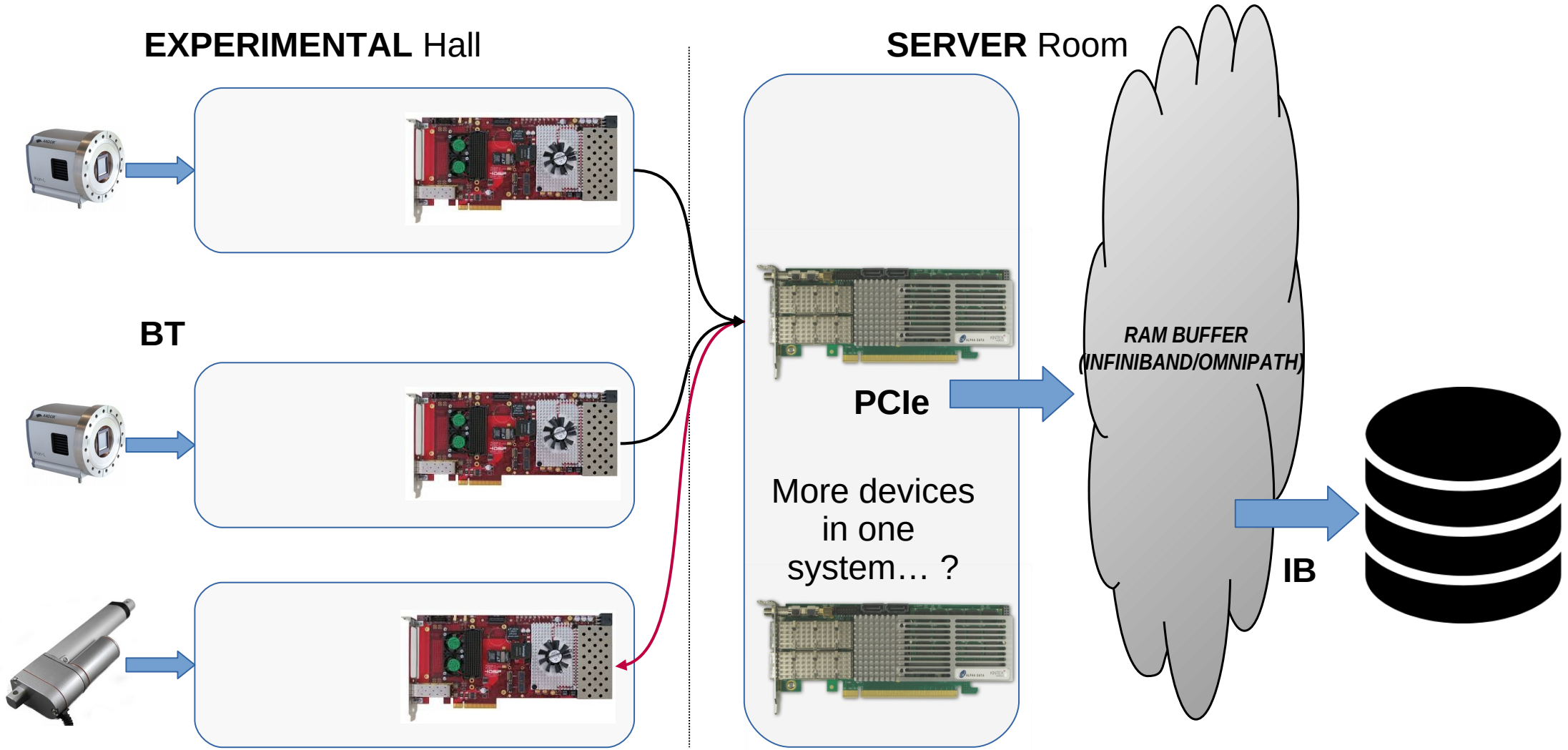
ADVANTAGES OF FPGA

- Support of OpenCL → C like programming
- Communication interfaces directly connected to FPGA

APPLICATIONS

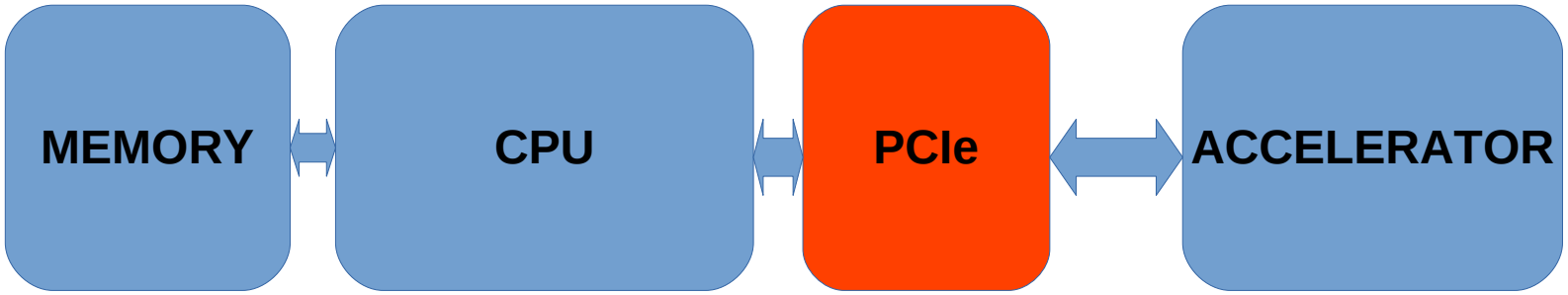
- User data acquisition
- **Control system loop → thanks to connected communication interfaces**

Feedback scheme



Online Data Processing

TRADITIONAL DATA ACCESS



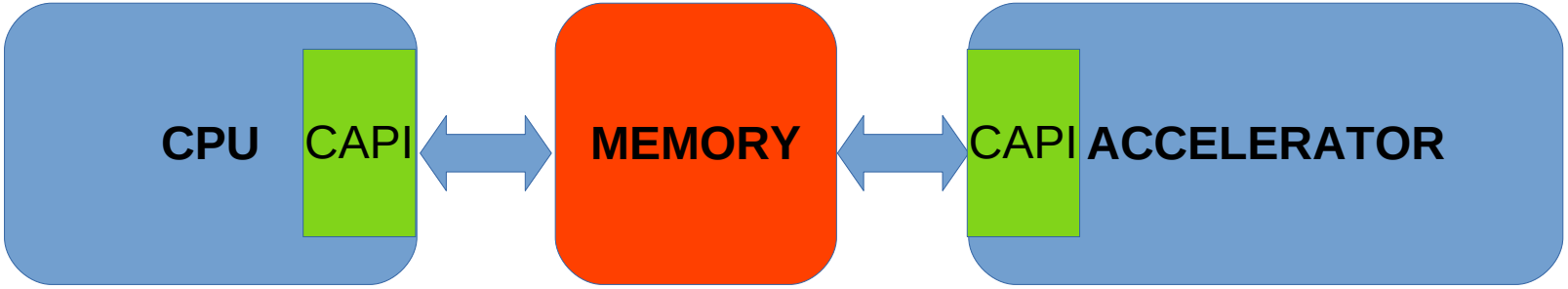
Data transfers are bottleneck

- PCIe data transfer has significant influence to acceleration and Data Acquisition
- There are SW drivers necessary for data transfers
- FPGA can keep data on local memory but…
  - The data have to be transferred for further processing or usage by other components of the Control / Data acquisiston system
- Complex system cooperation where one FPGA accelerator is not enough → interconnection is required

Online Data Processing

LATEST TECHNOLOGIES AVAILABLE

- Gen-Z (Independent memory semantic protocol)
- CCIX (Cache Coherent Interconnect for Accelerators)
- **CAPI (Coherent Accelerator Processor Interface)**



CAPI

- Enables to share Memory (coherently) between CPU and Accelerator
- Currently uses PCIe connector → No special HW is necessary

Functional unit in CPU (CAPP – Coherent Acc. Processor Proxy)
Functional unit in Accelerator (PSL – Power Service Layer)

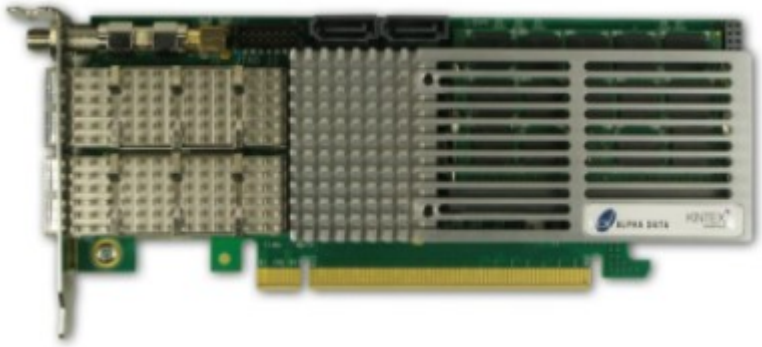
CAPI 2 → Power9 processors, uses PCIe Gen 4 connector
OpenCAPI → Consorcium
OpenCAPI 3 → Not layerd on top of PCIe inetrfaces

Further Development

POWER ARCHITECTURE + CAPI SOLUTION

- PowerN architecture → Superscalar architecture
- Currently Power8 with CAPI enabled PCIe slots
- AlphaData CAPI enabled FPGA card + CAPI IP Core
- CAPI is supported by Mellanox Infiniband EDR components

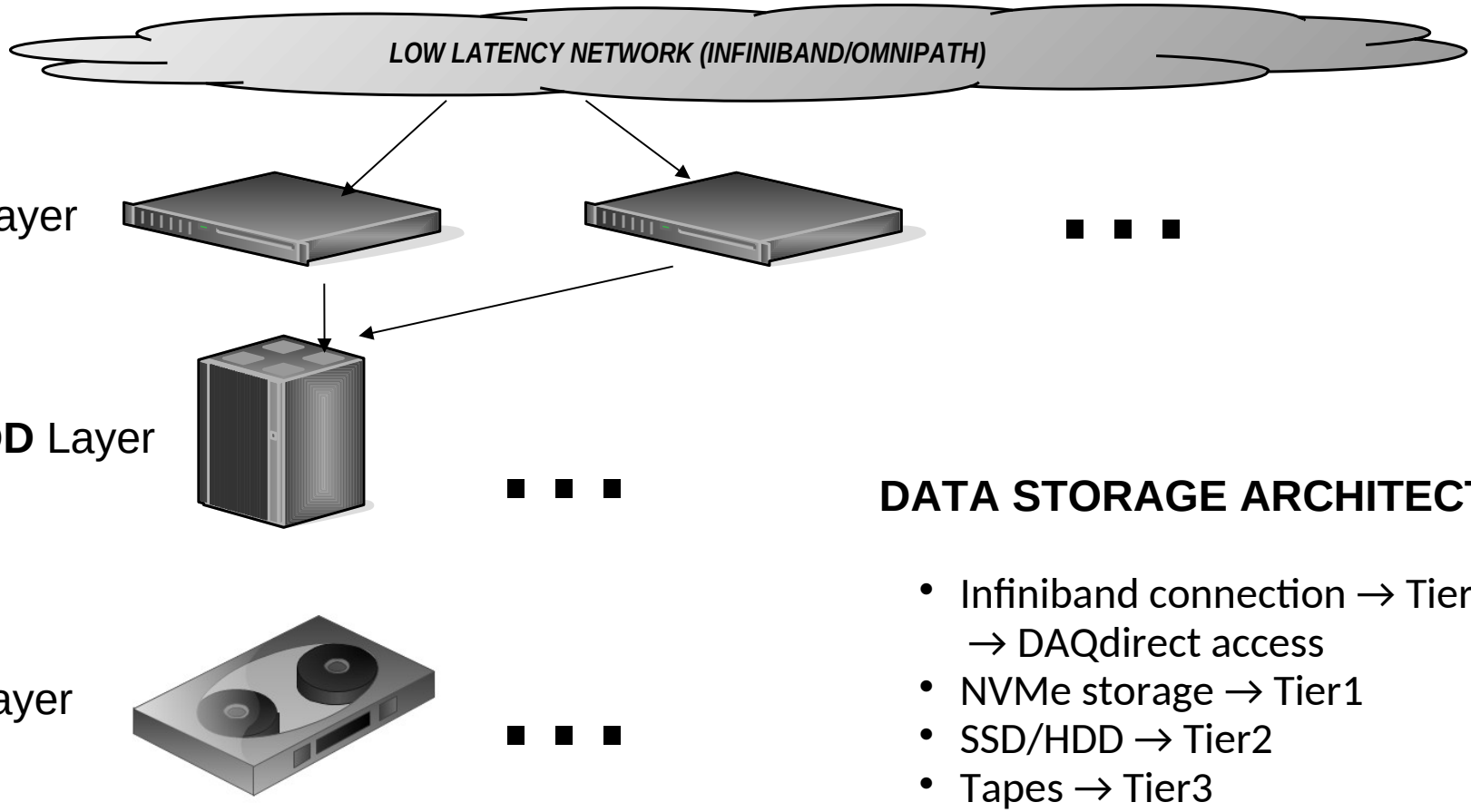
+



Data Storage

HORIZONTAL DOMAIN → Involves throughput

VERTICAL DOMAIN
→ Involves size



DATA STORAGE ARCHITECTURE

- Infiniband connection → Tier0
  → DAQdirect access
- NVMe storage → Tier1
- SSD/HDD → Tier2
- Tapes → Tier3

Thank you for attention!