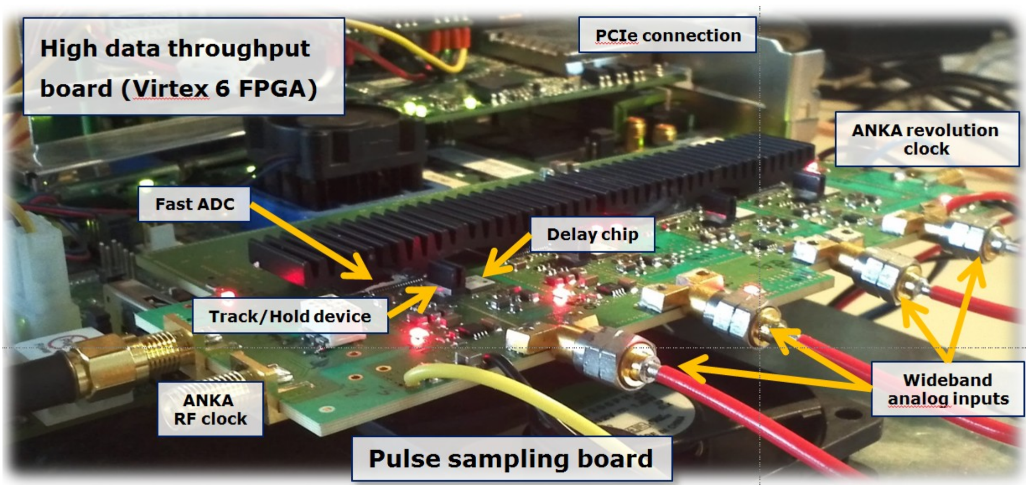
Computing Infrastructure for Online Monitoring and Control of High-throughput DAQ Electronics

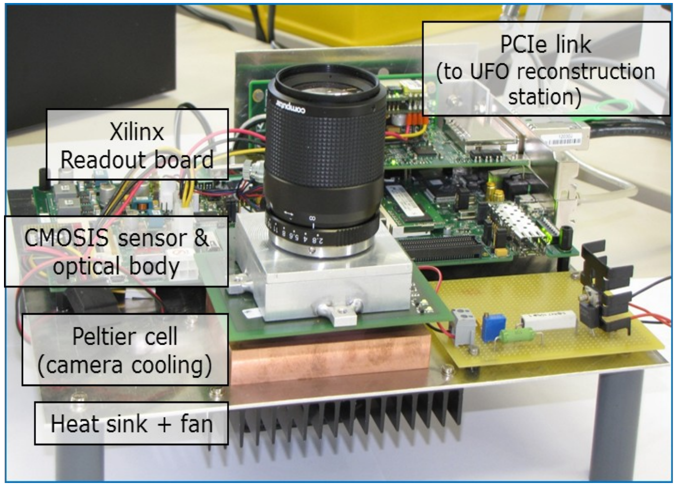
Karlsruhe Institute of Technology

S. Chilingaryan, M. Caselle, T. Dritschler, T. Farago, A. Kopmann, U. Stevanovic, M. Vogelgesang

Hardware, Software, and Network Organization



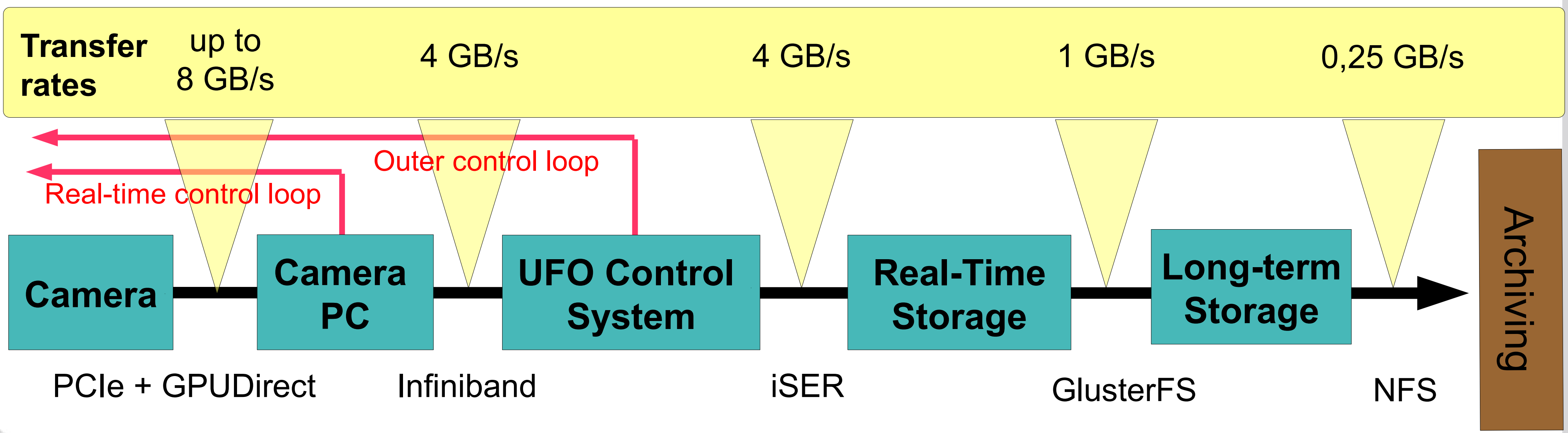
Picosecond Sampling Electronics for Terahertz Synchrotron Radiation



Prototype of Streaming PCIe Camera Developed in house

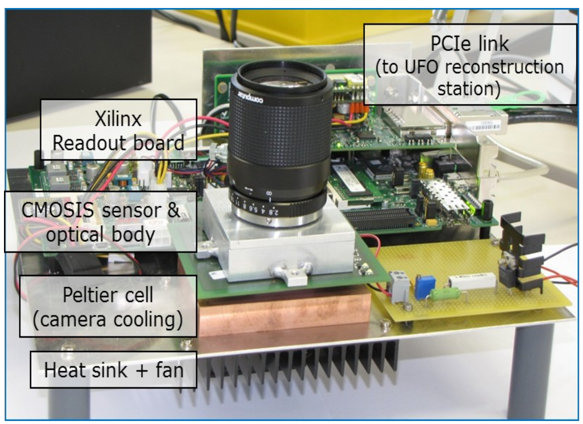
Requirements

- **Handling of Sensors with data rates up to ~ 8 GB/s (8-12 bit)**
- **Real-time control loop based on 2D Images + Online compression**
  - In-flow 8 GB/s, unpacked up to 16 GB/s (16 bit)
- **Slow control loop based on 3D Tomographic Images**
  - In-flow 4 GB/s, unpacked 32 GB/s (single-precision floating-point)
- **Raw data storage at full speed, i.e. 4 GB/s**
- **Long-term storage at 1 GB/s**
- **Integration with Tango Control System**
- **Low administrative effort**



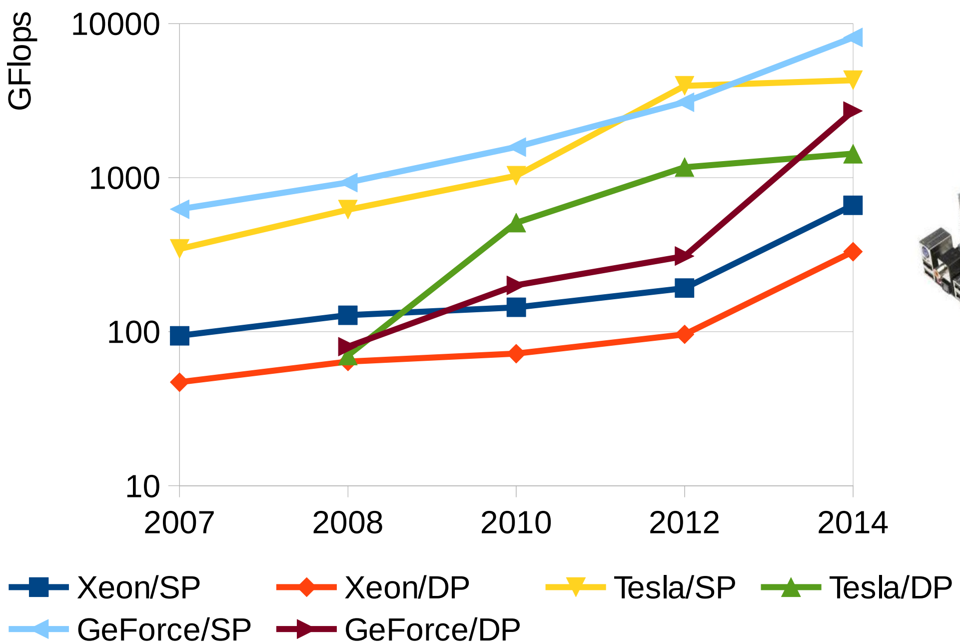
Concepts

- Programmable DAQ electronics with PCI-express interface
- Distributed control system based on Infiniband interconnects
- GPU-based computing
- Multiple levels of scalability
- Cheap off-the-shelf components



Prototype of Streaming PCIe Camera Developed in house

Poster FPI02



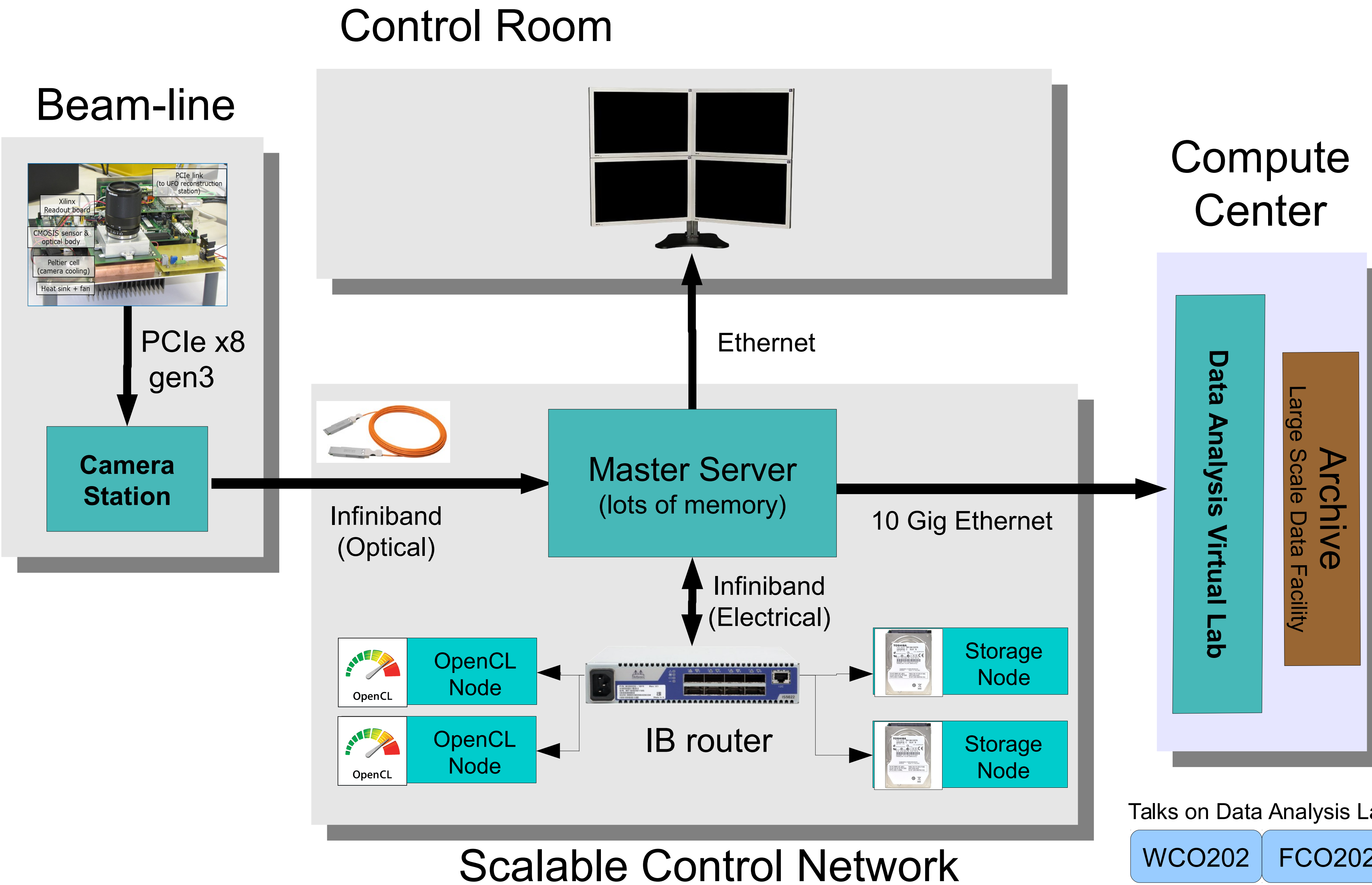
Historical trends of CPU and GPU performance

Easily scalable Quad-SLI
35 Tflops
for ~ 5000 EUR

S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

UFO Control Network



Talks on Data Analysis Lab: WCO202 FCO202

S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Master Server

Cameras

FDR Infiniband

56 Gbit/s

Ethernet

10 Gb/s

Storage

LSDF

Large Scale Data Facility

Internal
PCO.edge
PCO.dimax
….

<u>**SuperMicro 7047GR-TRF**</u> (Intel C602 Chipset)
CPU: 2 x Xeon E5-2680v2 (total 20 cores at 2.8 Ghz)
GPUs: 7 x NVIDIA GTX Titan
Memory: 256 GB (512GB max)
Network: Intel 82598EB (10 Gb/s)
Infiniband: 2 x Mellanox ConnectX-3 VPI
Storage: Areca ARC-1880-ix-12 SAS Raid
8 x Samsung 840 Pro 510 (Raid0)

- High amount of memory
- Fast SSD-based Raid for overflow data

S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Master Server

Cameras

Internal
PCO.edge
PCO.dimax
….

FDR Infiniband
56 Gbit/s

Ethernet
10 Gb/s

Storage

LSDF
Large Scale Data Facility

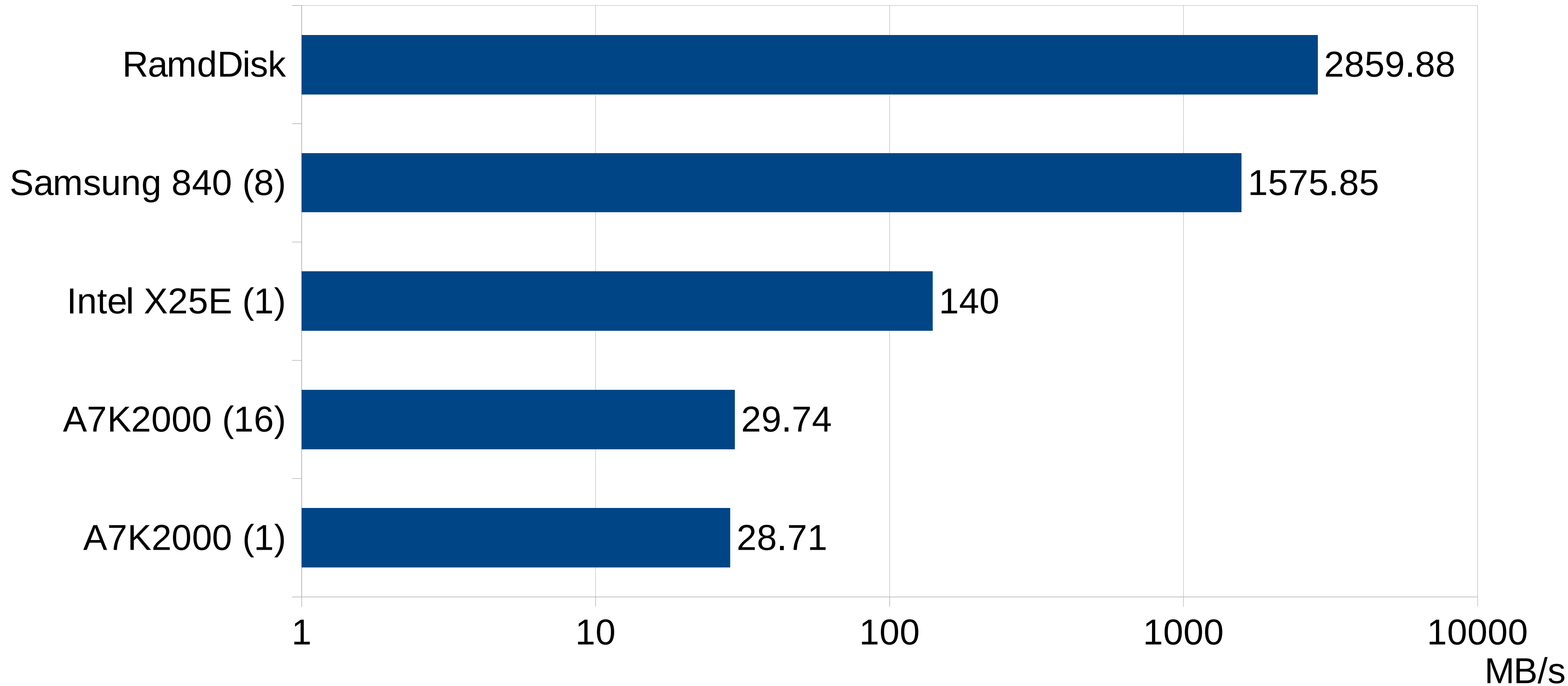
External PCIe x16 (16 GB/s)

SFF8088 (2.4 GB/s)

<u>SuperMicro 7047GR-TRF</u> (Intel C602 Chipset)
CPU: 2 x Xeon E5-2680v2 (total 20 cores at 2.8 Ghz)
GPUs: 7 x NVIDIA GTX Titan
Memory: 256 GB (512GB max)
Network: Intel 82598EB (10 Gb/s)
Infiniband: 2 x Mellanox ConnectX-3 VPI
Storage: Areca ARC-1880-ix-12 SAS Raid
8 x Samsung 840 Pro 510 (Raid0)
16 x Hitachi A7K200 (Raid6)

- High amount of memory
- Fast SSD-based Raid for overflow data
- Easy scalability with external PCI express and SAS

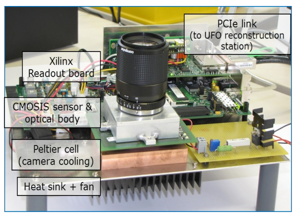
Caching large data sets



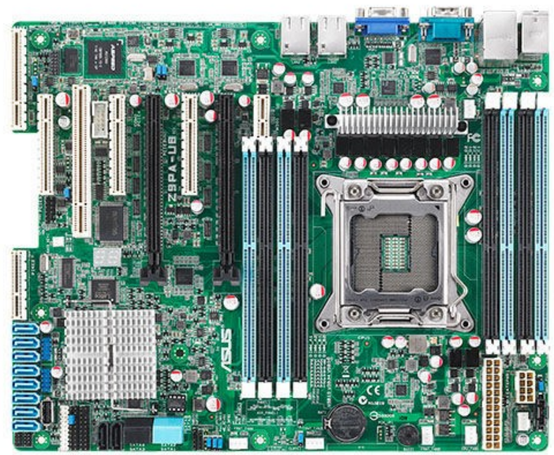
Using SSD drives may significantly increase random access performance to the data sets which are not fitting in memory completely. The big arrays of magnetic hard drives will not help unless multiple readers involved.

Camera Station

Camera Link
850 MB/s



PCI express (x8)
Up to 8 GB/s



Infiniband
56 Gbit/s

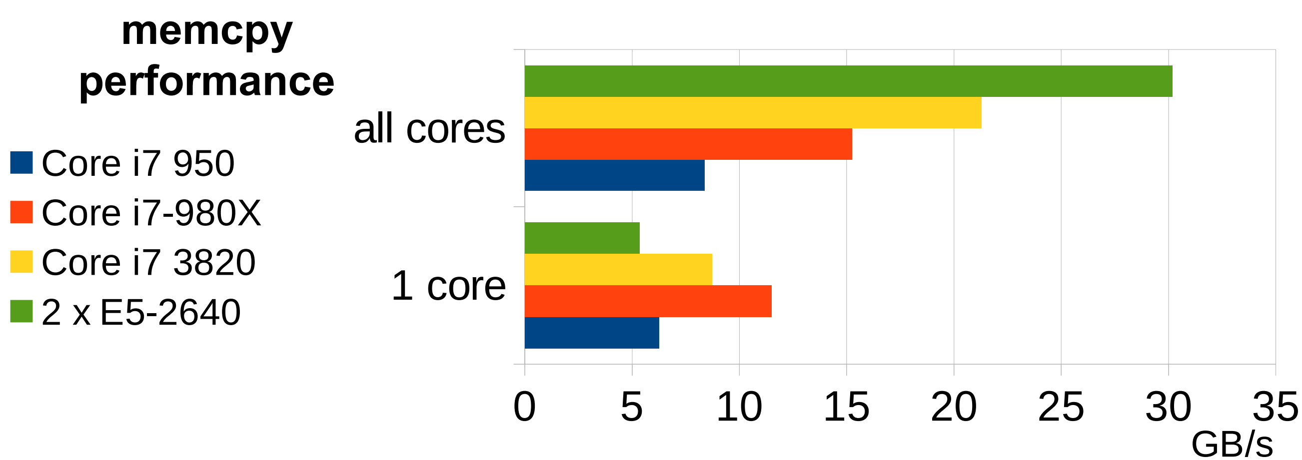
<u>Asus Z9PA-U8</u> (Intel C602 Chipset)
CPU: Xeon E5-1620v2 (total 4 cores at 3.7 Ghz)
GPUs: NVIDIA GTX Titan
Memory: 32 GB (128GB max)
Infiniband: Mellanox ConnectX-3 VPI

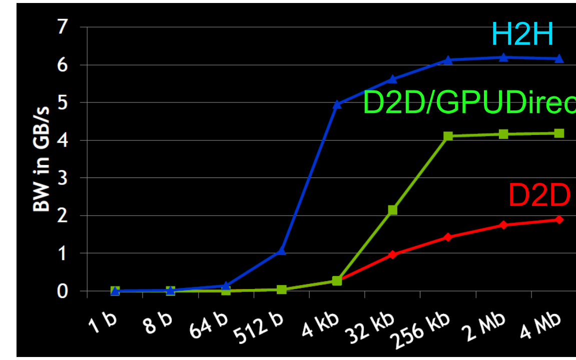
- High-speed 4-channel memory
- IPMI-based remote control
- Optional fast SSD-based storage
- 4x high speed PCI express slots

S. Chilingaryan et. all

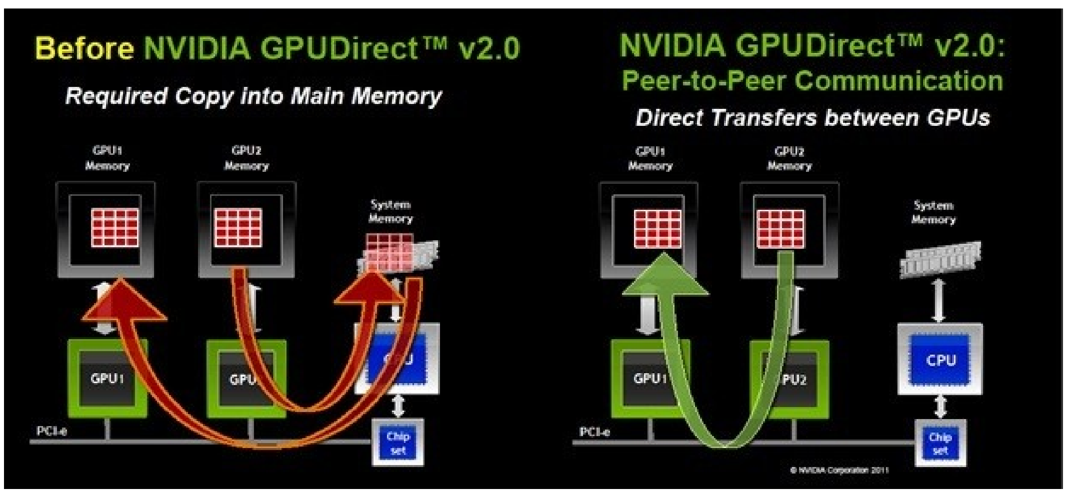
Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

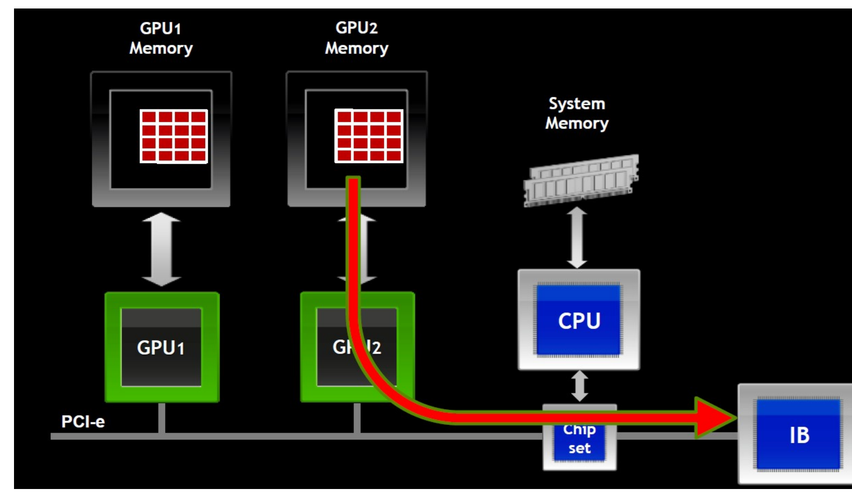
NVIDIA GPUDirect





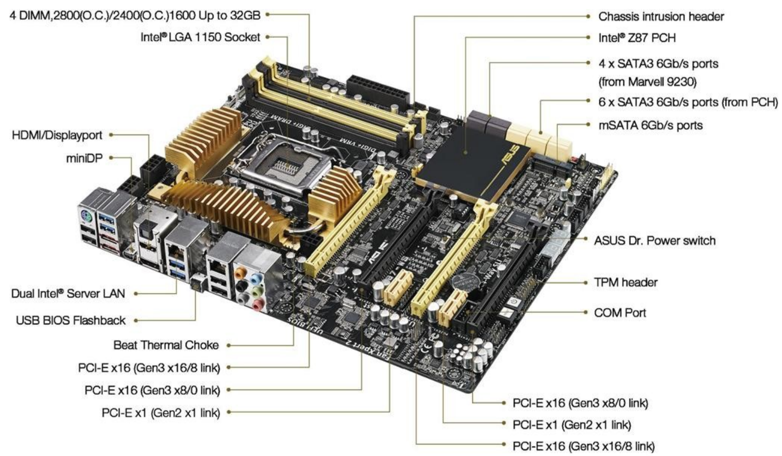
MVAPICH2 1.9b
throughput





Direct communication between GPUs, Network,
and other devices on PCI express bus

Cluster Node



Asus Z87-WS (Intel Z87 PCH Chipset)
CPU: Core i5-4670 (total 4 cores at 3.4 Ghz)
GPUs: 3 x NVIDIA GTX Titan
Infiniband: Mellanox ConnectX-3 VPI
Memory: 16 GB (32GB max)

NVIDIA GTX Titan
Memory: 6 GB at 288 GB/s
Single-precision Gflops: 4500
Double-precision Gflops: 1500

- 4-Way SLI
- Low Price

Storage Protocols

Network FS
NFS
Samba
SSHFS

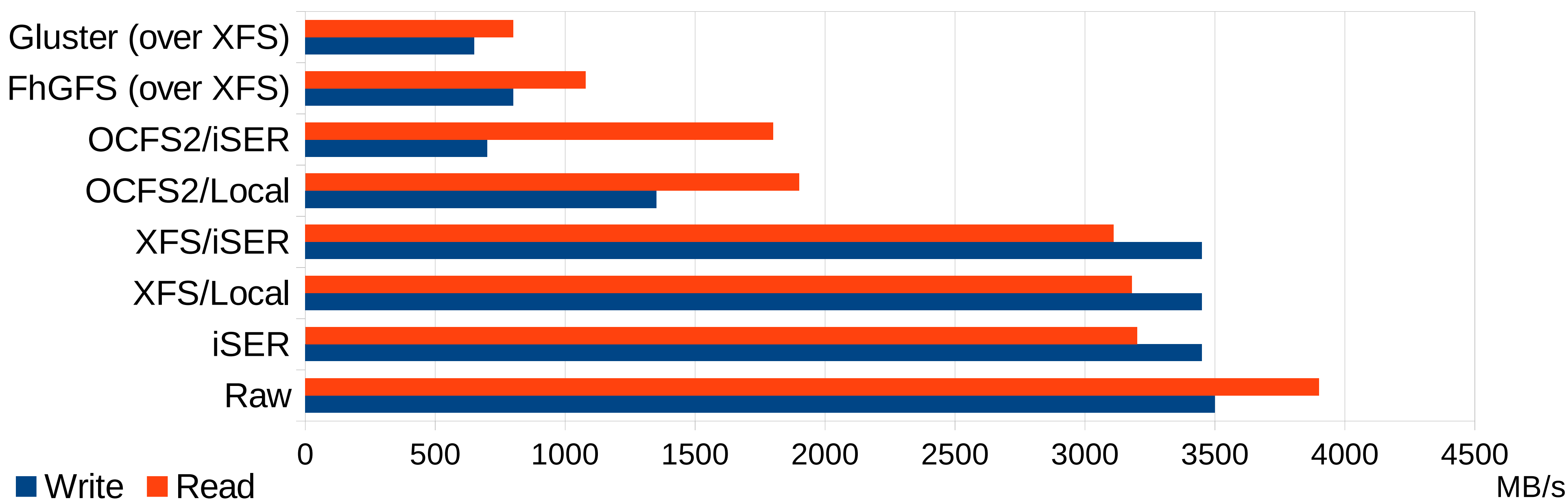
Slow

Cluster FS
Lustre (patched kernel)
Gluster
FhGFS (close-sourced)

Network Devices
iSCSI (slow)
iSER

Slow if few nodes

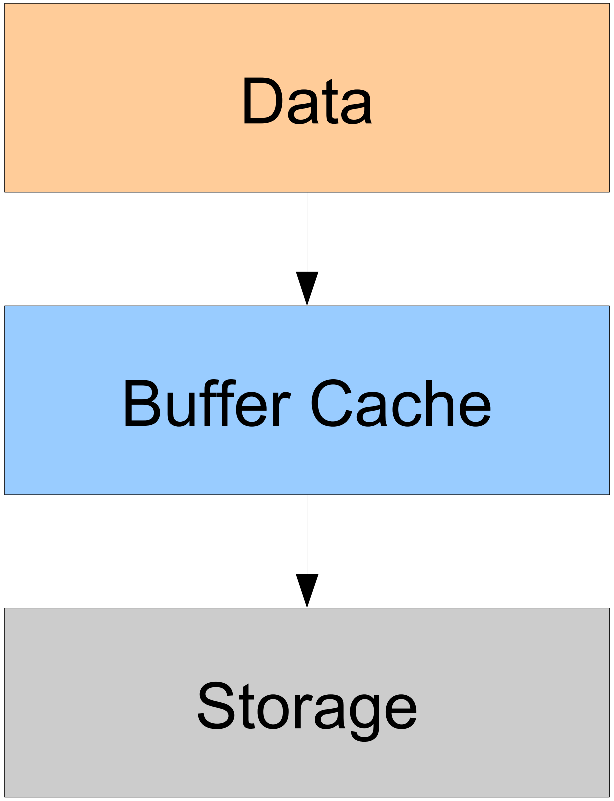
OCFS2



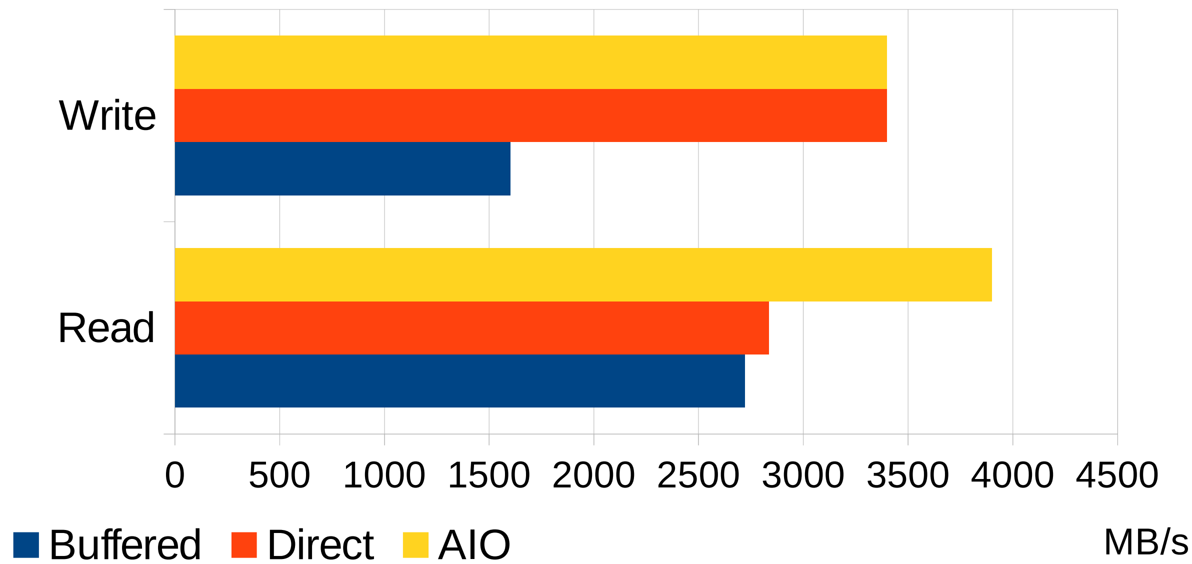
S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Handling High-speed Storage



Default data flow in Linux

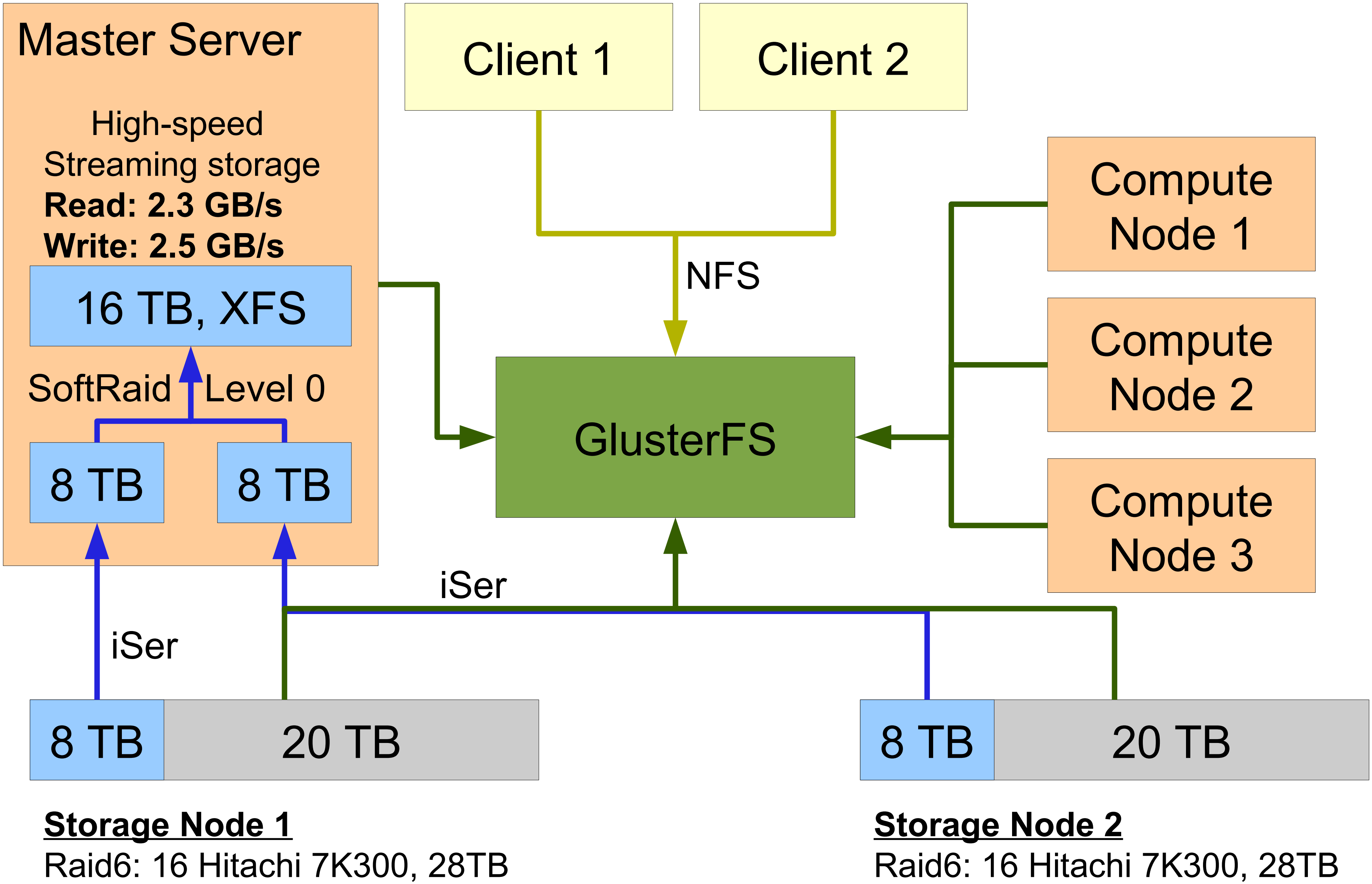


- Buffer cache significantly limits maximal write performance
- Kernel AIO may be used to program IO scheduler to issue read requests without delays

Optimizing I/O for maximum streaming performance using a single data source/receiver

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

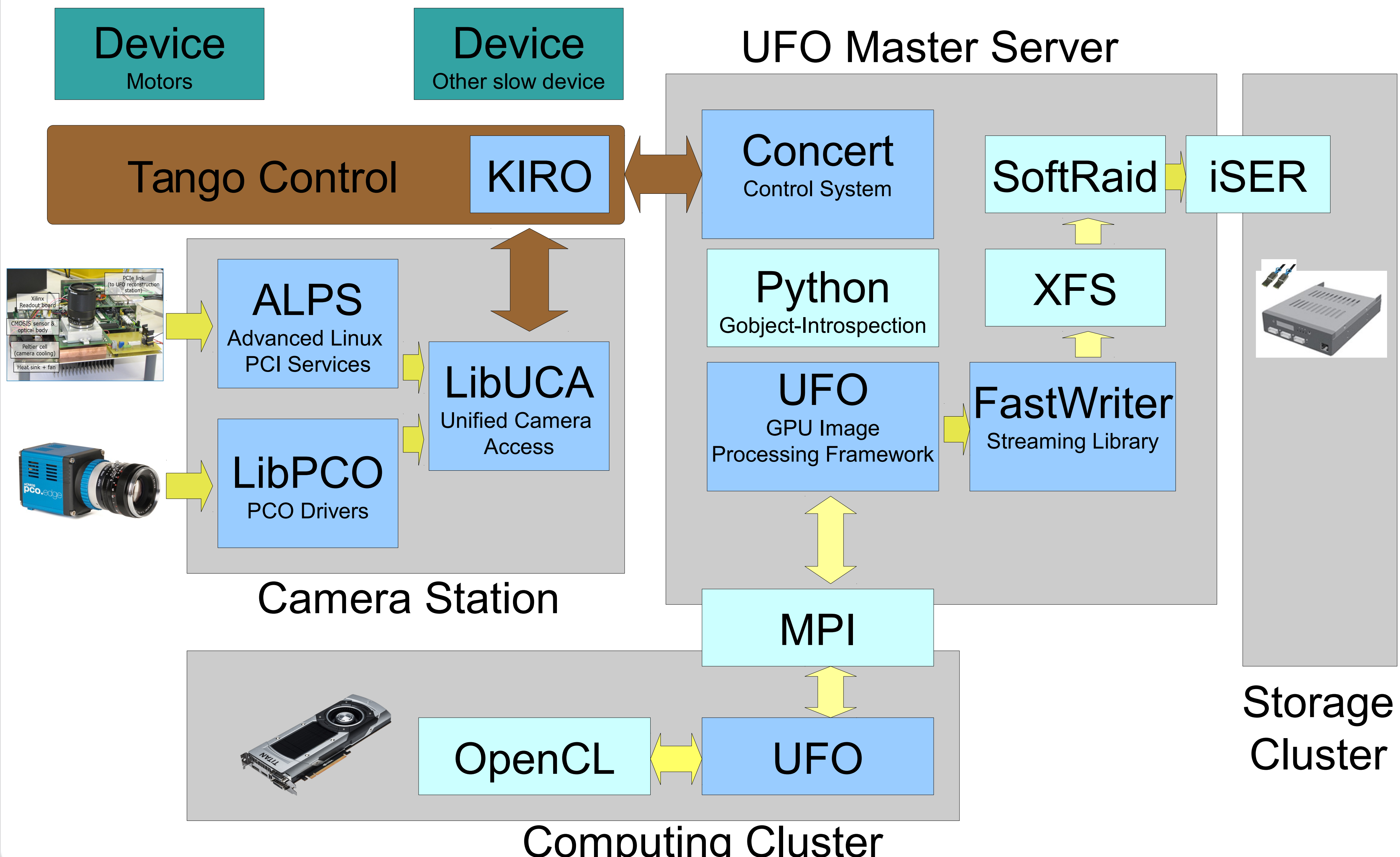
UFO Storage Subsystem



S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Software Stack

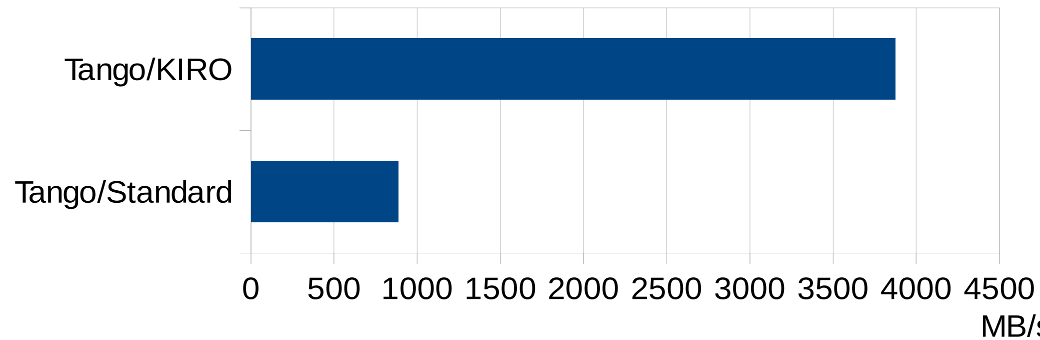


S. Chilingaryan et. all
Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

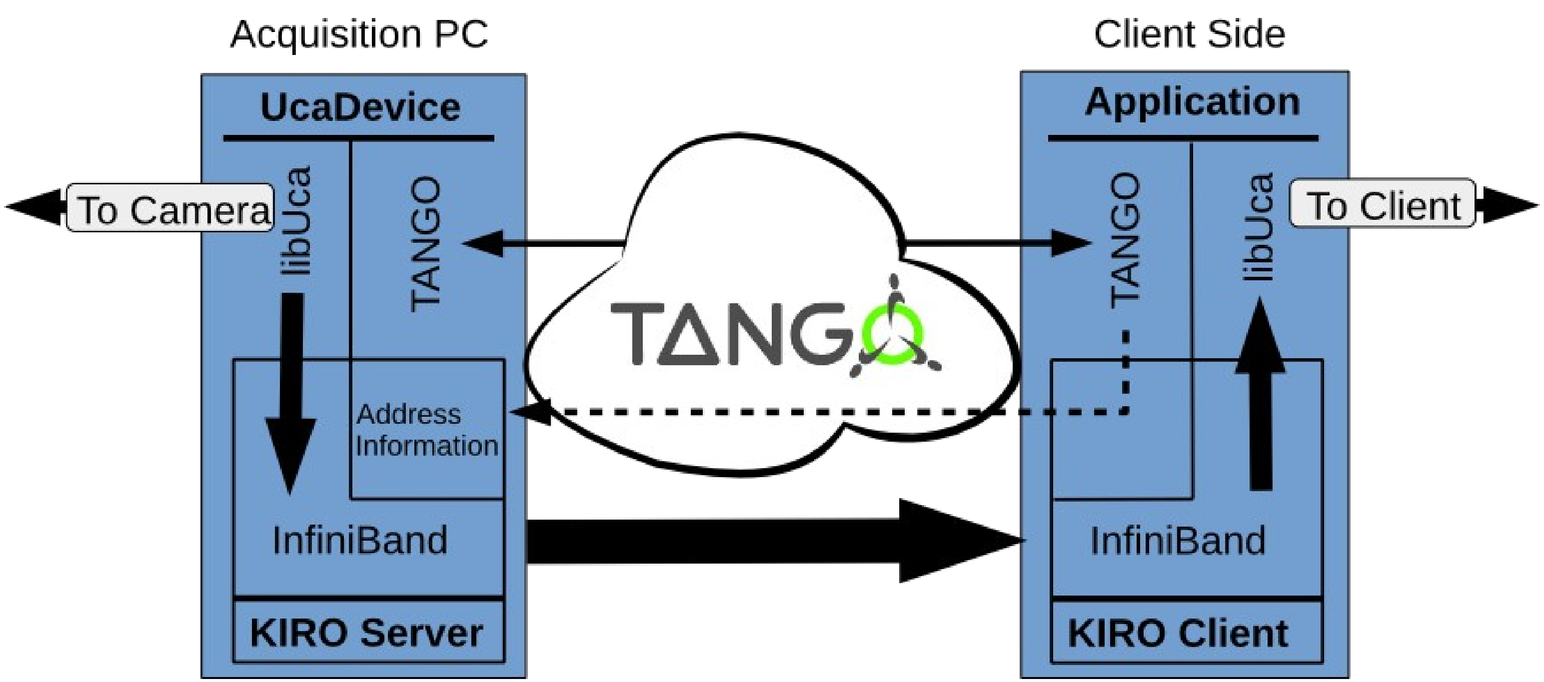
KIRO: Integration with Tango

Tango over Corba over TCP over Infiniband is slow

Poster FPI01
Using InfiniBand for High-Throughput Data Acquisition in a TANGO Environment



KIRO Architecture



S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Summary

- Only easy to get off-the-shelf components are used
- Our architecture can be easily scaled from a single PC to the cluster with performance in hundreds of teraflops.
- The reliable storage for data streaming with rates over 3 GB/s can be easily build based on 1 – 2 low-end servers.
- The electronic components can be distributed over large area and connected with high-speed using Optical Infiniband links.
- The provided software stack allows easily integrate new devices and processing algorithms.
- The Tango control system is extended to support high-speed communication over Infiniband

Check related talks and posters:

Poster FPI01 Using InfiniBand for High-Throughput Data Acquisition in a TANGO Environment	**WCO202** Data Management at the Synchrotron Radiation Facility ANKA
Poster FPI02 Picosecond Sampling Electronics for Terahertz Synchrotron Radiation	**FCO202** OpenGL-Based Data Analysis in Virtualized Self-Service Environments

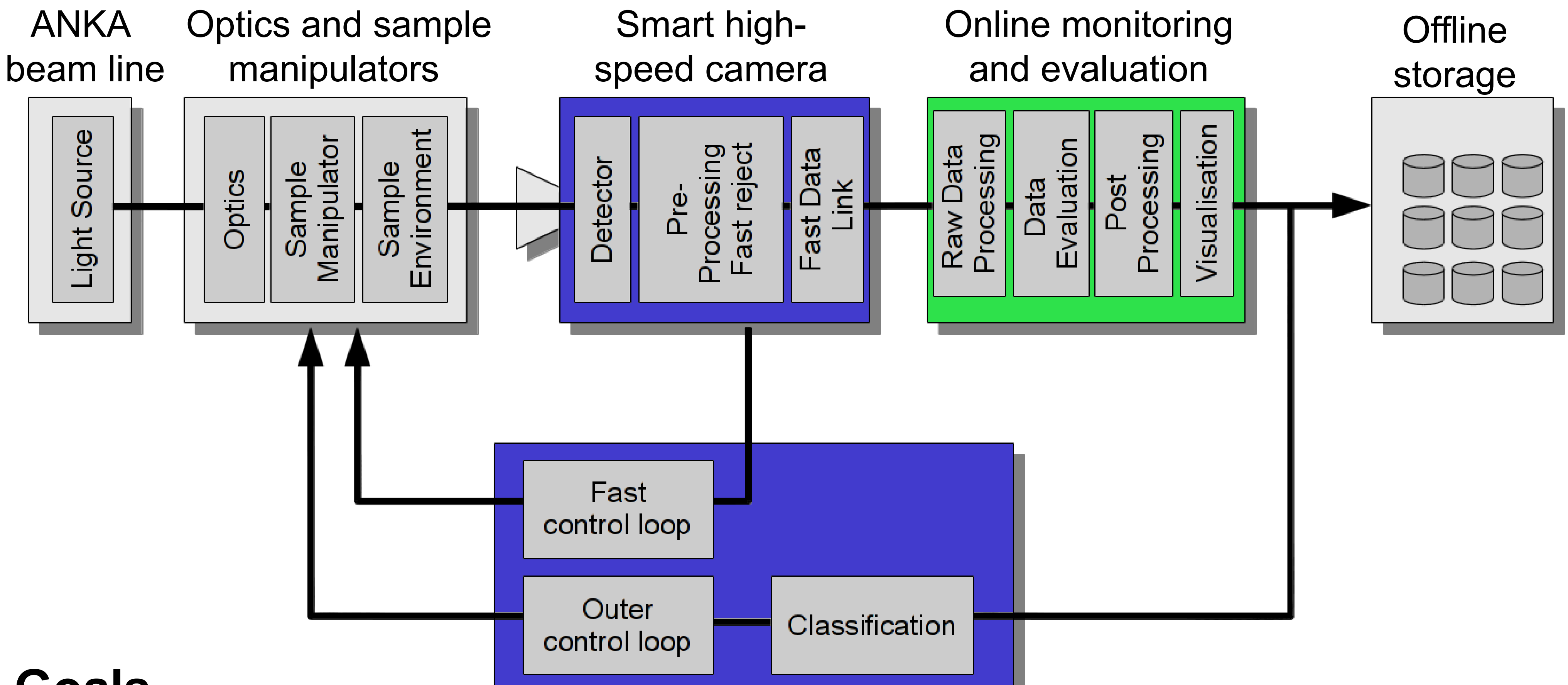
S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Extra Slides

UFO *Ultra Fast X-ray Imaging of Scientific Processes with On-Line Assessment and Data-Driven Process Control*

Karlsruhe Institute of Technology



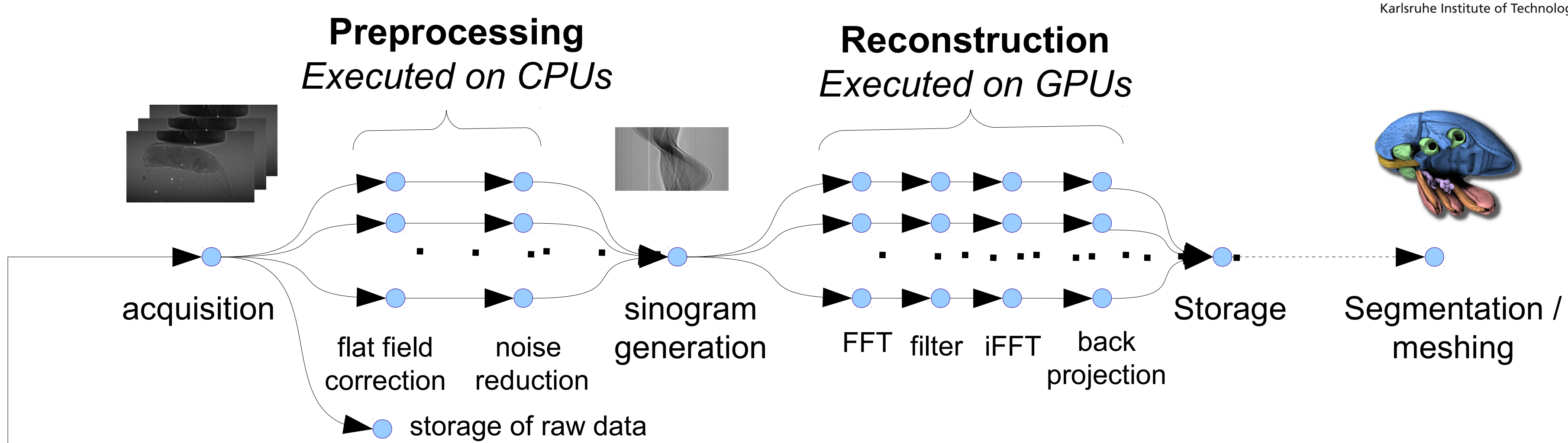
Goals

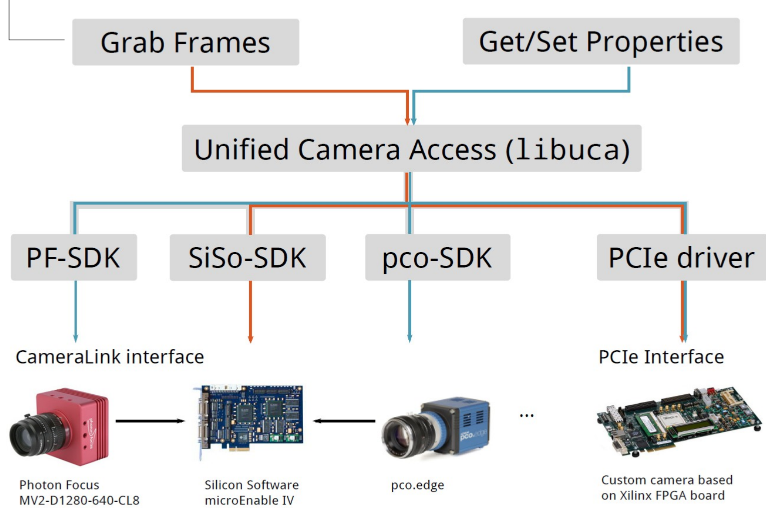
- High speed tomography
- Increase sample throughput
- Tomography of temporal processes
- Allow interactive quality assessment
- Enable data driven control
  - Auto-tunning optical system
  - Tracking dynamic processes
  - Finding area of interest

S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

UFO Image Processing Framework





OpenSource

http://ufo.kit.edu

Features

- Easy Algorithm Exchange
- Camera Abstraction
- Pipelined Processing
- Glib/GObject, scripting language support with introspection
- OpenCL + automated management of OpenCL buffers

Filters

Tomography & Laminography

FBP, DFI, Algebraic Methods

Non Local Means for Noise Reduction

Optical Flow

S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

ALPS

Scripting
Bash, Perl, Ruby, Python

LabVIEW
Control System Integration

pcitool
Command-line tool

GUI
Python/GTK User Interface

Web Service
Remote Programming API

Event Engine
IPE Camera

DMA Engine
Northwest Logic DMA

Register Access
XML Register List + Dynamic Registration API

DMA Memory Mapping

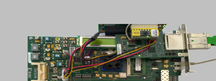
PCI Memory Access
Plain and FIFO

Access Serialization, Software Registers

IRQ Handling

Driver Access
Layer Small & Easy to Maintain

PCI / PCI Express Board
(Variety of FPGA Boards: KIT Camera, etc)



S. Chilingaryan et. all

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology