Karlsruhe Institute of Technology

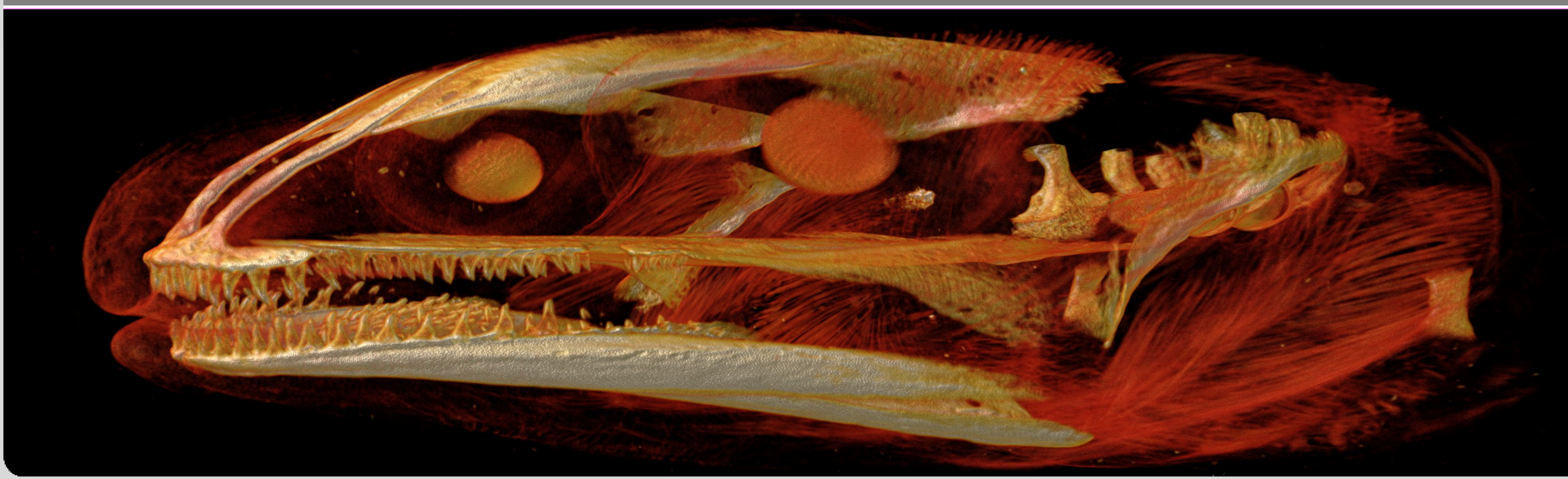
GPUs for Synchrotron Applications

Experiences and plans

Andreas Kopmann (KIT)

Workshop on GPUs in High Energy Physics 15.-16.6.2013, DESY

Institute for Data Processing and Electronics



Outline

1 GPUs for tomographic reconstruction
   – Optimizing performance

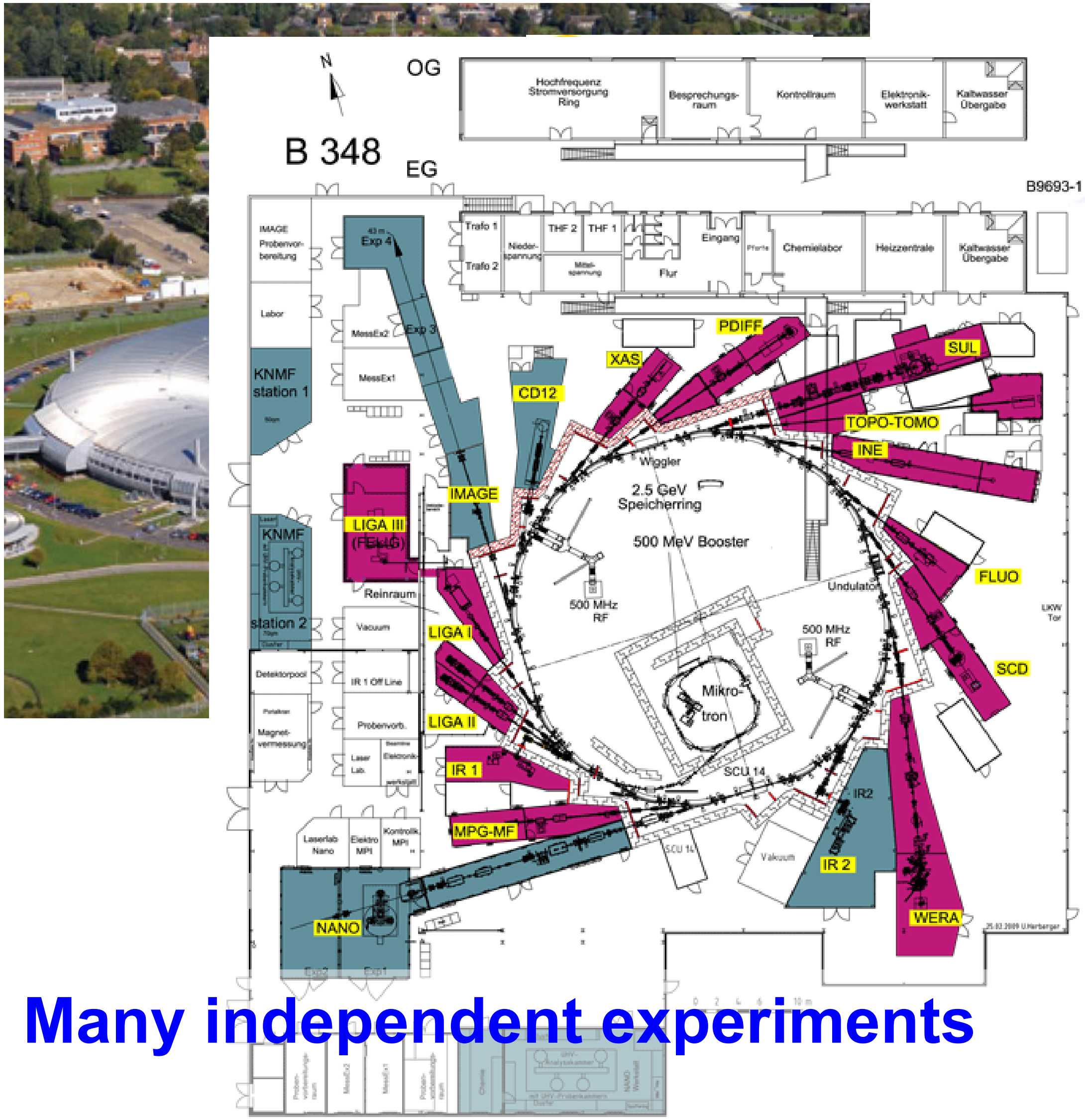
2 Processing of data streams
   – A parallel computing framework

3 System integration
   – What else is needed to become fast?

Synchrotron Applications?

Synchrontrons might look quite beautiful
Scientific infrastructure in many countries

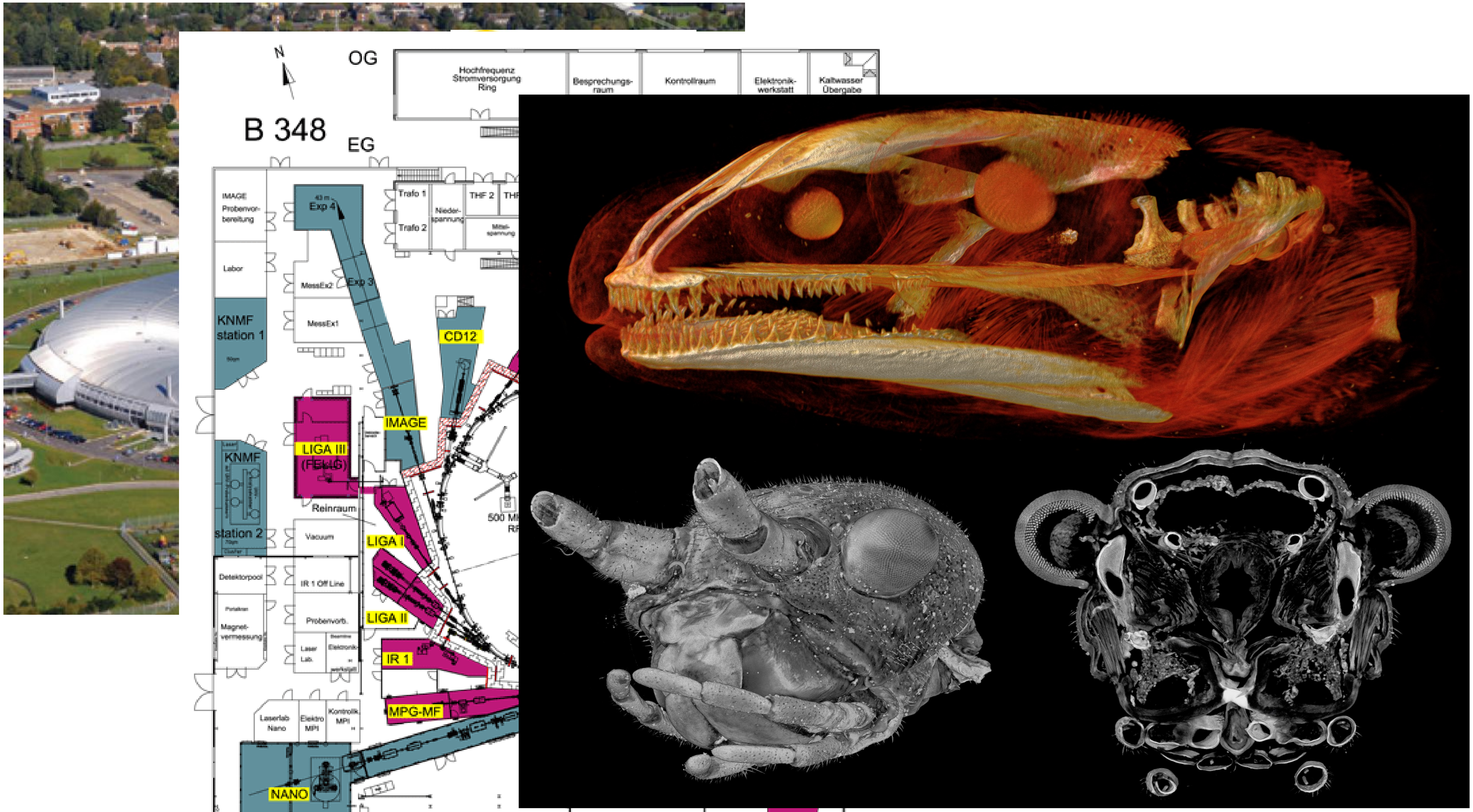
Synchrotron Applications?



Foorplan ANKA

Many independent experiments

Synchrotron Applications?

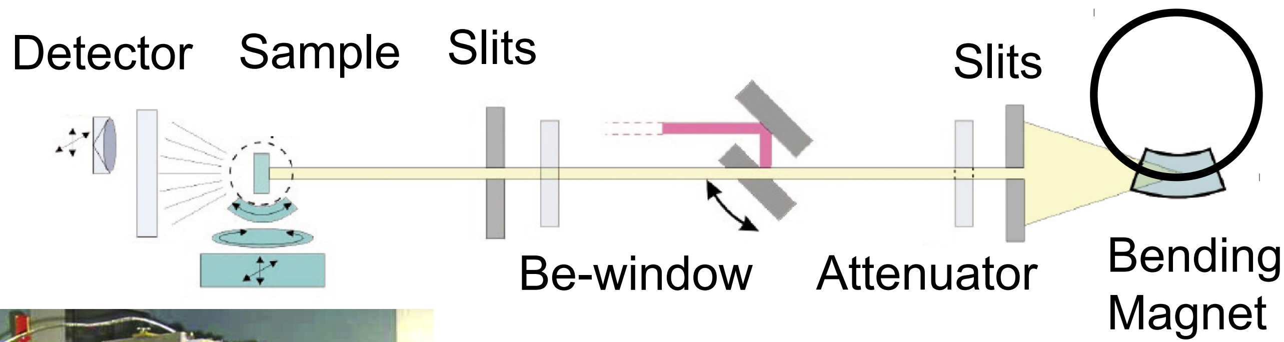


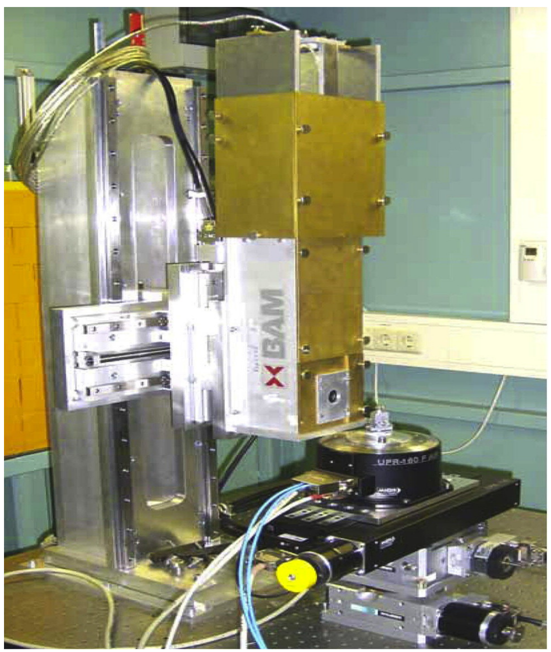
T. van de Kamp, KIT

Many independent experiments
3D imaging application w. high data rates

Tomography Beamline at ANKA Synchrotron

Experiment + Detector | Monochromator | Storage Ring

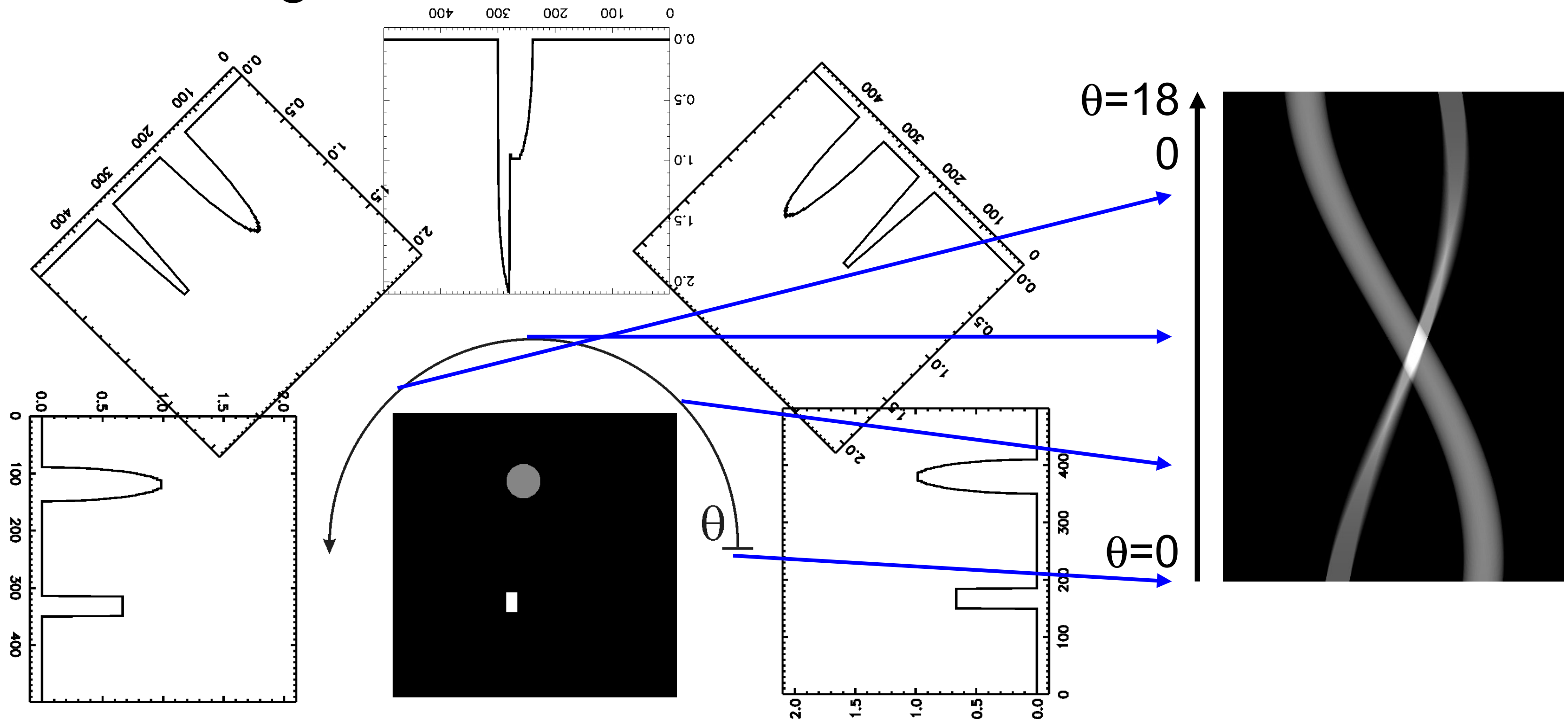




The rotating sample in front of a pixel detector is penetrated by X-rays produced in the synchrotron. Absorption at different angles is registered by camera and 3D map of sample denisity is reconstructed.

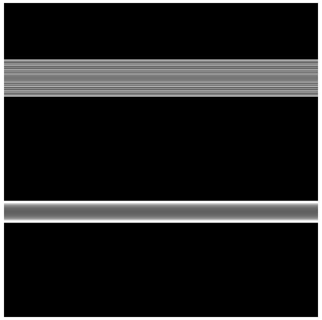
Synchrotron Tomography

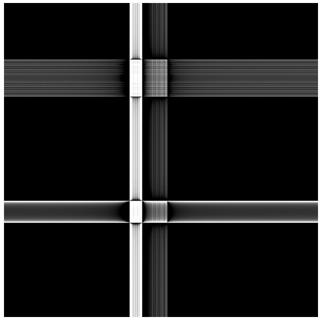
The sample is evenly rotated and the pixel detector registers series of parallel 2D projections of the sample density at different angles.

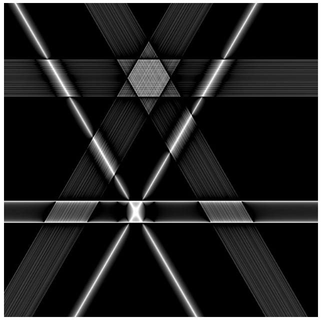


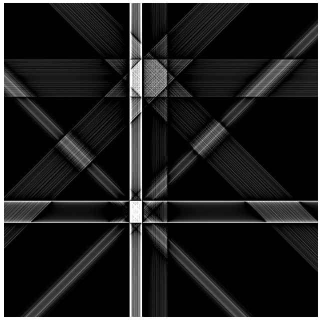
Synchrontron Tomogaphy

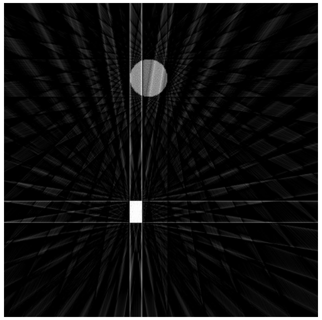
Filtered back-projection is used to produce 3D images from a manifold of two dimensional projections**.** Vertical slices are processed independently. For each slice all projections are smeared back onto the cross section along the direction of incidence yielding an integrated image.

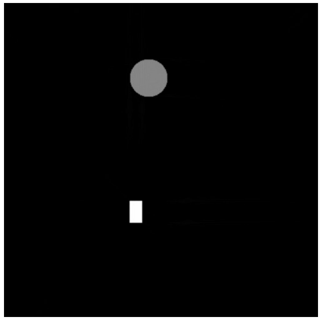

n = 1


n = 2


n = 3


n = 10


n = 50


n = 1000

Reconstruction Problem

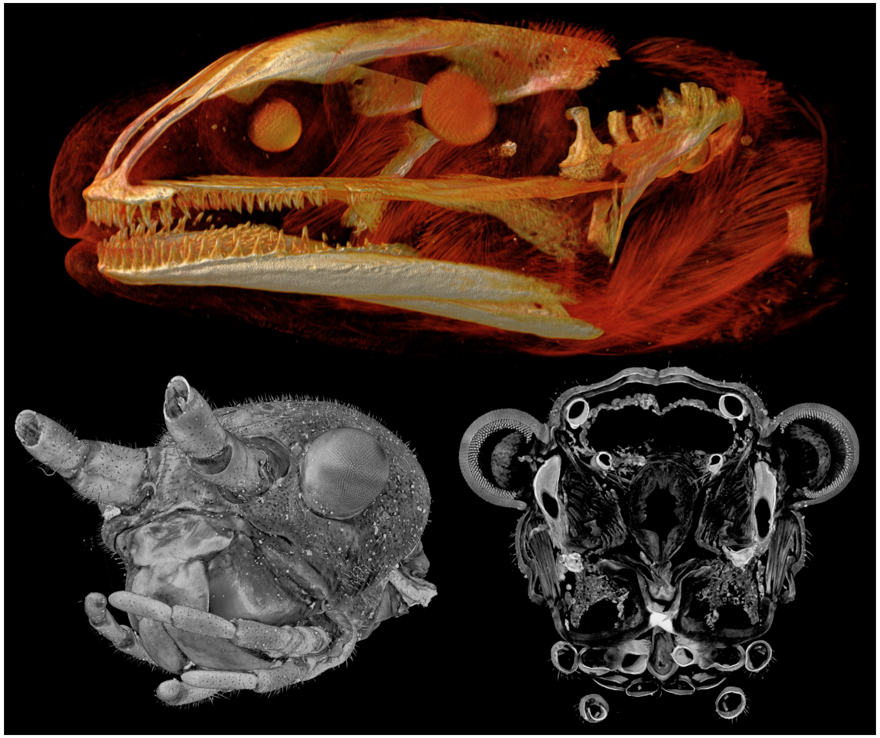
PCO.edge

Streaming camera
Resolution: 2560 x 2160
Dynamic Range: 16 bit
Frame Rate: 100 fps

Tomographic Reconstruction
Goal: 3D image w 2000^3 voxels
Projections: 2000
Acquisition time: 20 seconds

FBP Complexity: 144 Tflops
Xeon Performance: ~ 100 Gflops
Minimum time: ~ 15 minute (w DP)
Actually: ~ 1 hour

Data set: 21 GByte
20 seconds acquisition
CPU: 1 hour reconstruction



Heads of a newt larva showing bone formation and muscle insertions (top) and a stick insect (bottom), acquisition time 2s.

Recontruction with CPU Cluster

Karlsruhe Institute of Technology

REVIEW OF SCIENTIFIC INSTRUMENTS VOLUME 72, NUMBER 4 APRIL 2001

APS

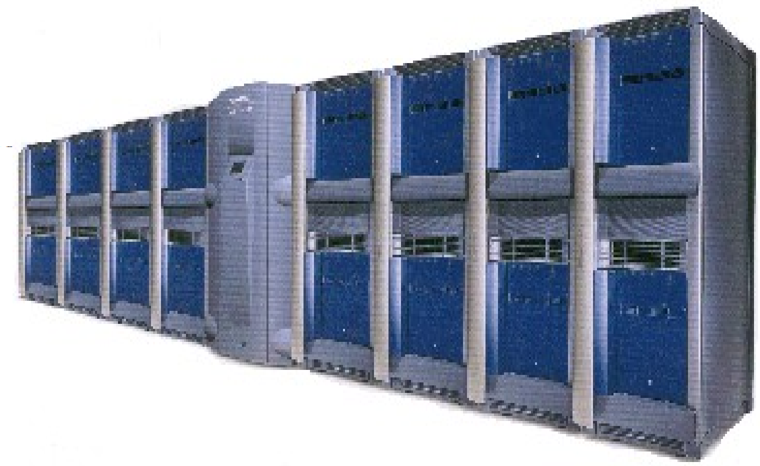
A high-throughput x-ray microtomography system at the Advanced Photon Source

Yuxin Wang,[a] Francesco De Carlo, Derrick C. Mancini, Ian McNulty, and Brian Tieman
Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois

John Bresnahan, Ian Foster, Joseph Insley, Peter Lane, and Gregor von Laszewski
Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois

TABLE II. Reconstruction time using the filtered backprojection method for various size data sets. In each case, 80 of the 128 processors of the parallel computer were used.

No. of projections	Projection size (pixels)	Calculation time (min)
721	1024×1024	17
721	512×512	2
361	512×512	1



SGI Origin 2000, 1996-2002
128 CPUs

Data set: 1.6 GByte

Cluster: 17 min → 1,4 MB/sec

First GPU Application

Towards Real-Time Tomography: Fast Reconstruction Algorithms and GPU Implementation
F. Marone et al., Swiss light source
2008 IEEE Nuclear Science Symposium

TABLE II
RECONSTRUCTION TIMES, INCLUDING FILE I/O

Code version	Runtime (s)
Original CPU code	20.9
Unoptimized GPU code	31.6
GPU code with optimized data access	9.7
GPU code with optimized data access and hardware linear interpolation	2.2

Nvidia GTS 8800, 2007

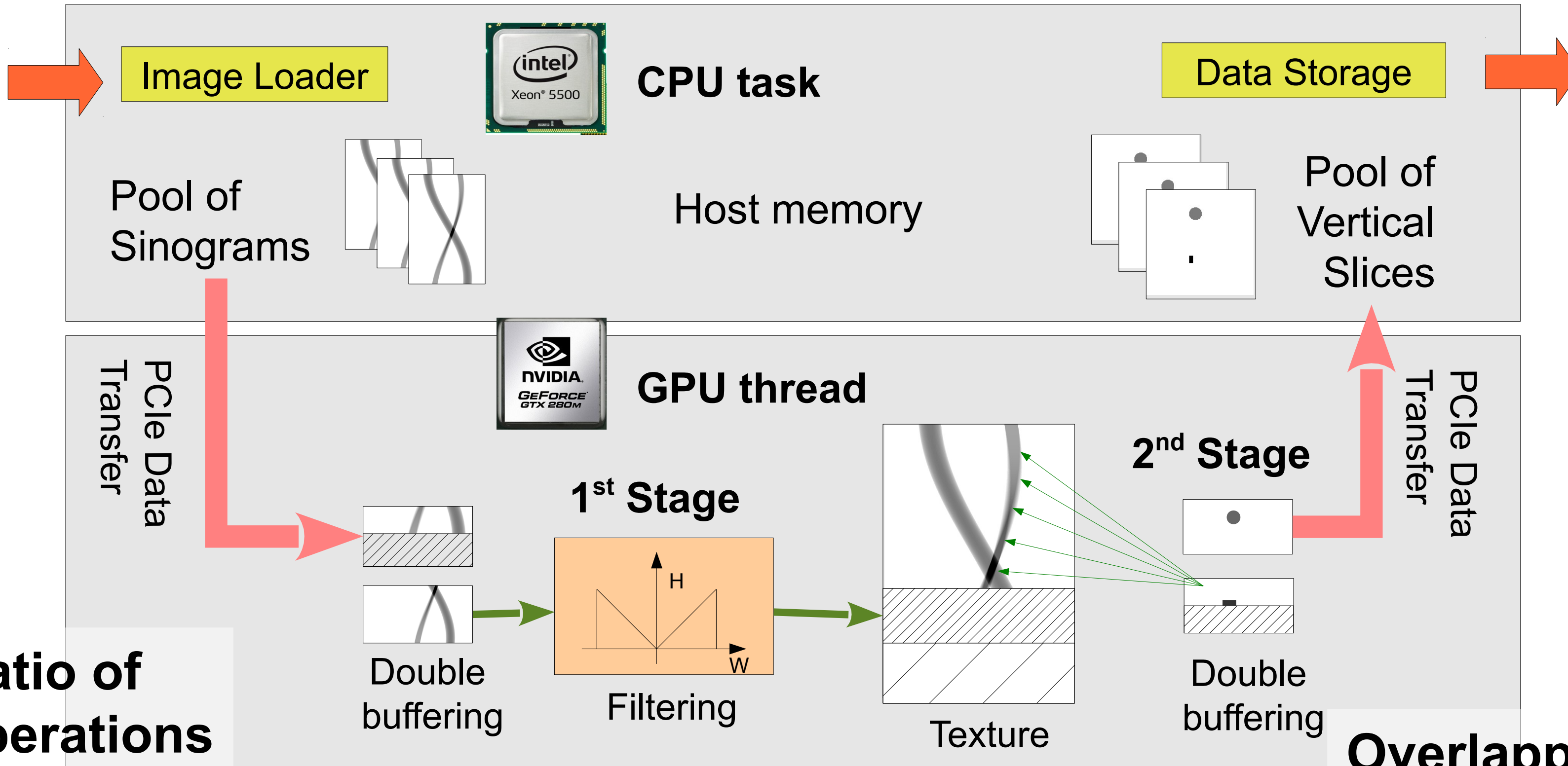
Data set: 3MB (only one slice)

CPU: 20.9 sec → 0,14 MB/s
GPU: 2.2 sec → 1,40 MB/s !

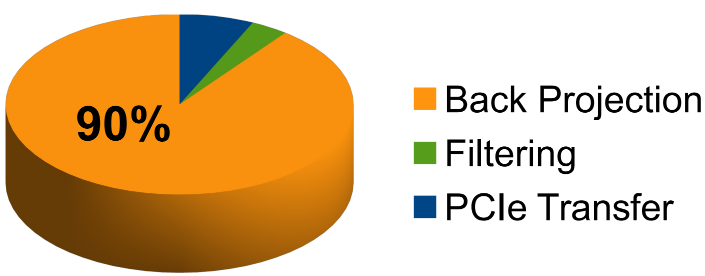
BUT: Full reconstruction > 1h

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

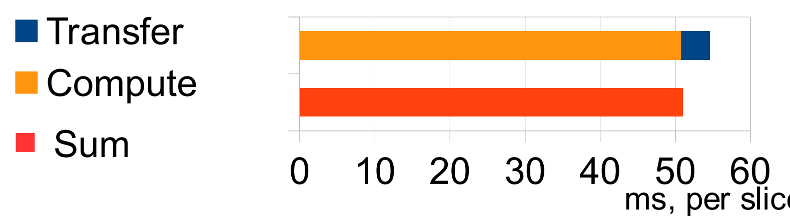
Filtered Back Projection on GPU



Ratio of operations

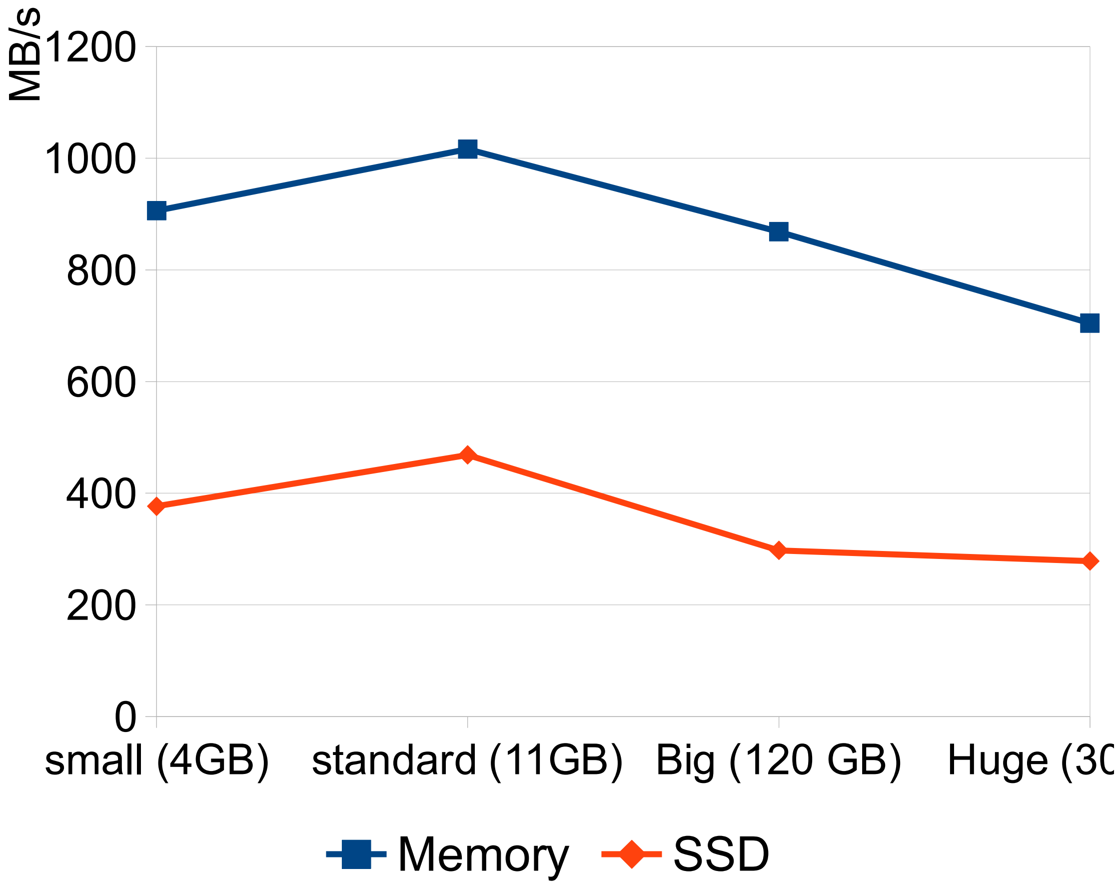


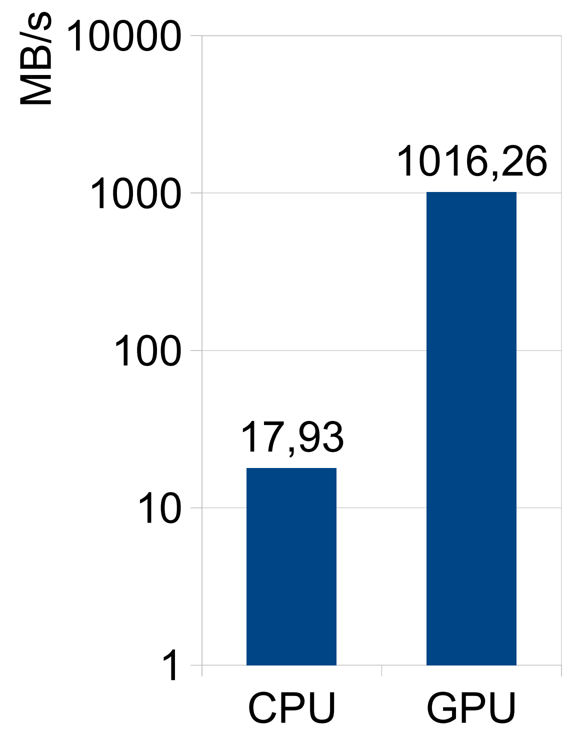
Overlapping efficiency



A. Kopmann et. all, GPUs for synchrotron applications
Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Filtered Back Projection Performance



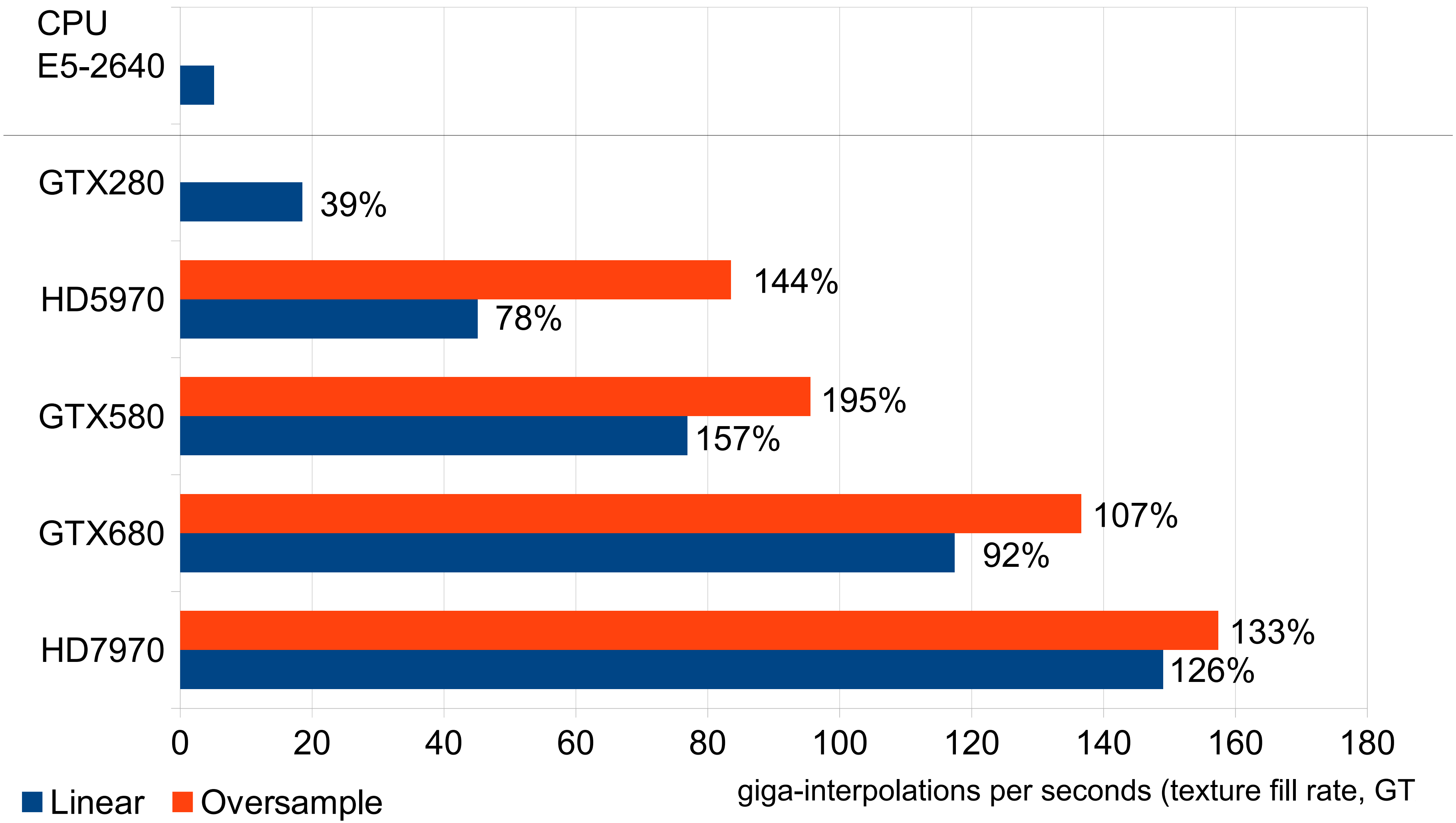


11 GB data set

GPU: 4 x GTX590 , 8 cores
CPU: 2 x Xeon X5650, 12 cores
(Both from 2011)

Computing for FBP is solved!
BUT fast imaging requires more effort

Back Projection: Evolution of GPUs



Optimization for GPU architectures

Karlsruhe Institute of Technology

NVIDIA

GT200
Base version with texture engine

Fermi + 100%
High computation power, but low speed of texture unit
Reduce load on texture engine: use shared memory to cache the fetched data and, then, perform linear interpolation using computation units.

Kepler + 75%
Low bandwidth of integer inst-ructions, but high register count
Uses texture engine, but processes 16 projections at once and 16 points per thread to enhance cache hit rate

AMD

VLIW + 530%
Executes 5 independent operations per thread
Computes 16 points per thread in order to provide sufficient flow of independent instructions to VLIW engine

GCN + 95%
High performance of texture engine and computation nodes
Balance usage of texture engine and computation nodes to get highest performance

Parallel Computing Framework

Karlsruhe Institute of Technology

How to support code development for GPUs?

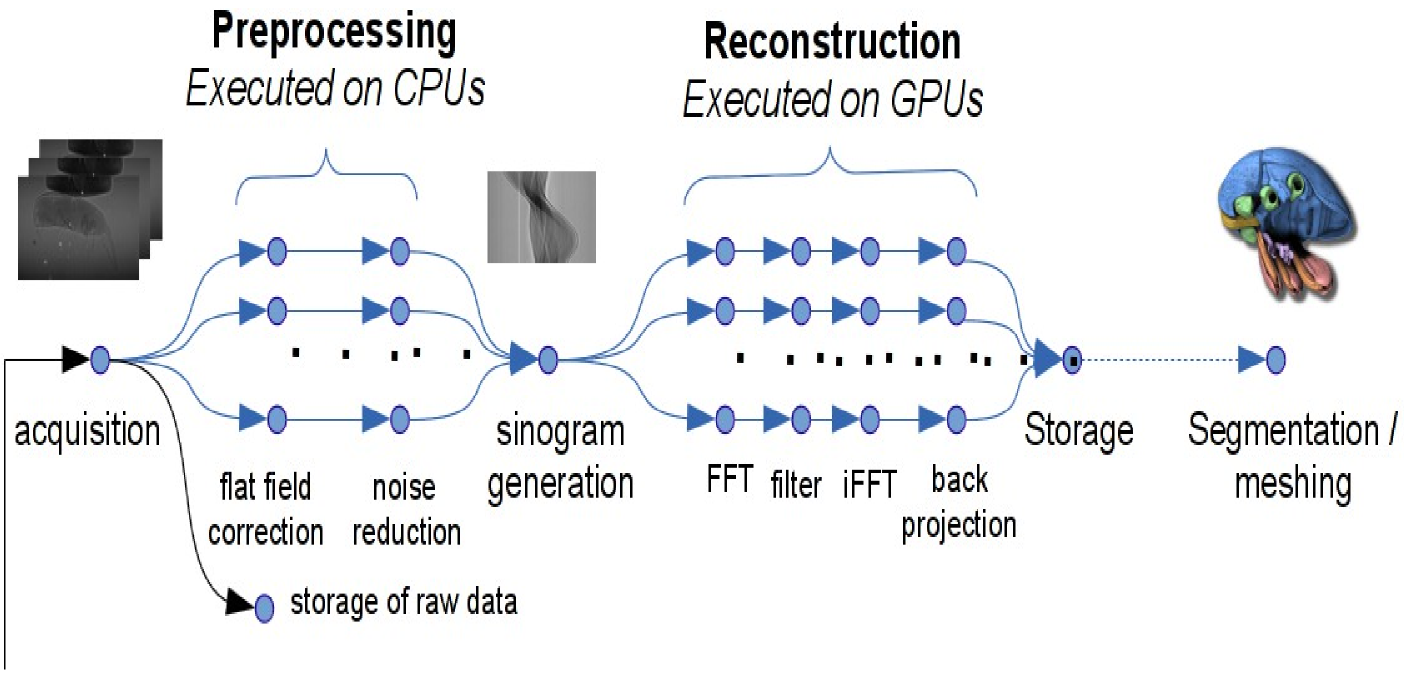
Requirements:
- Processes data streams (usually 1 to 4 dimensional floating point data)
- Detect and use all hardware resources

Developer
- Hides parallelization and concurrency details
  - Management of memory transfers
  - Multiple implementations (e.g. for CPU + GPU)

User
- Simple end-user interface
  - GUI + Scripting
- Modular algorithm design

Realization

- Define algorithms as self-contained tasks
- Specify data flow as edges in a graph



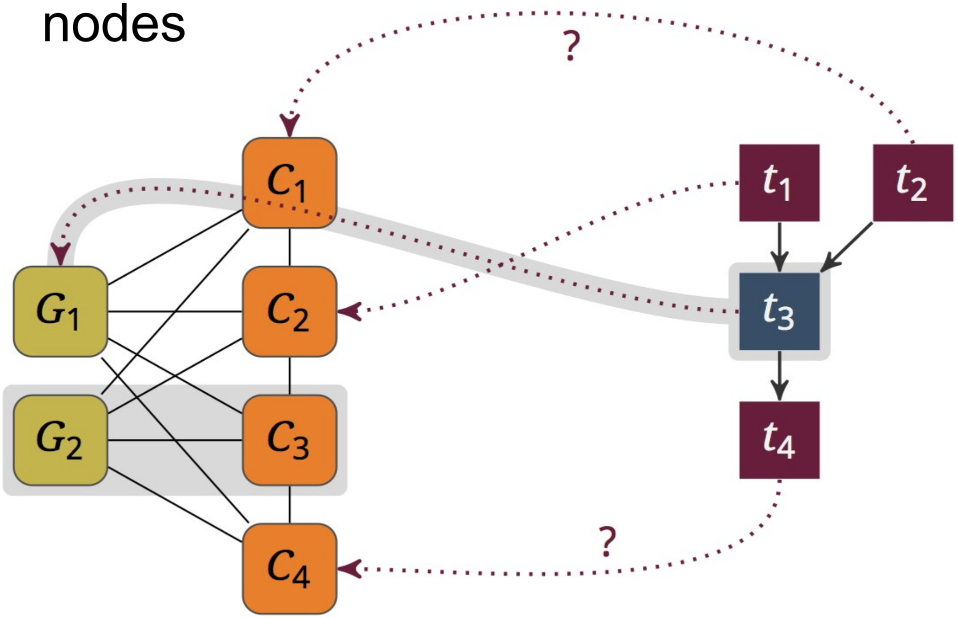
Institute for Data Processing and Electronics
Karlsruhe Institute of Technology

Scheduling

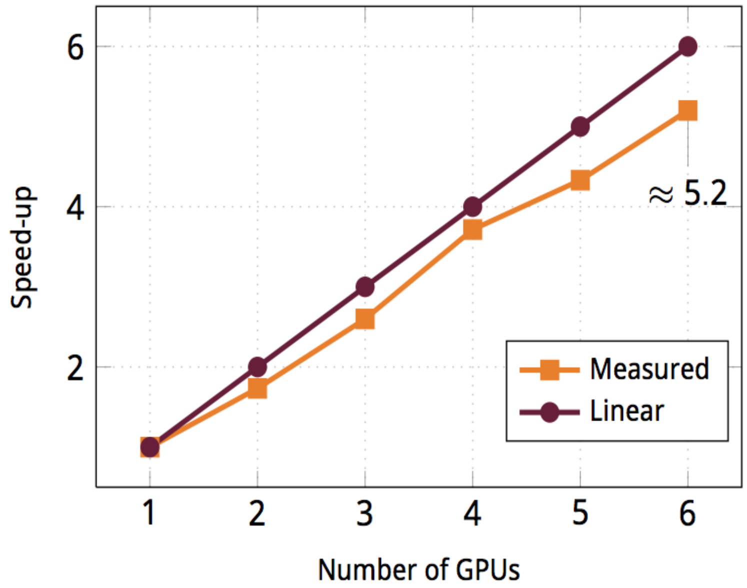
- Detect the architecture
  - Type of nodes
  - Interconnects
  - NUMA
- Map Tasks to nodes
  - Consider CPU + GPU implementations
- Split and merge data streams
- Optimization

Computing nodes

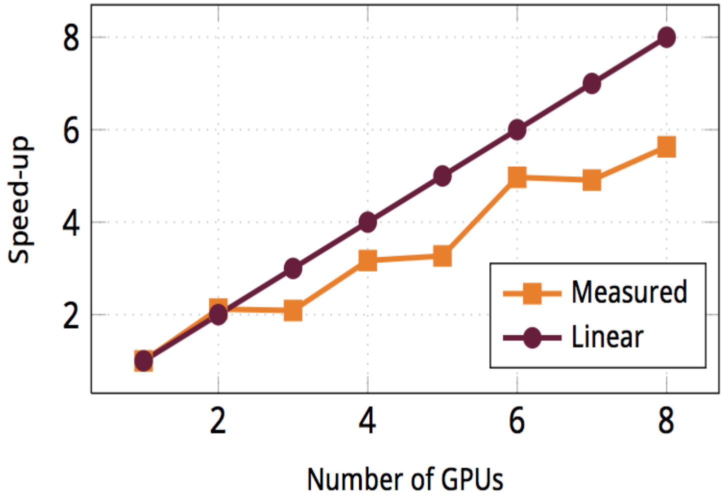
Task



Scheduling



GPU server: 6 x GTX 580



GPU cluster: 4 x (2 x GTX 580)

- Number of GPUs / node is limited
- GPU cluster scale quite well
  - Relevant for advanced algorithms

UFO-Algorithms are prepared for cluster usage

Cluster again?

- MASSIVE1, located at the Australian Synchrotron:
- 42 nodes with 12 cores per node running at 2.66GHz (504 CPU-cores total)
- 48 GB RAM per node (2,016 GB RAM total)
- 2 nVidia M2070 GPUs with 6GB GDDR5 per node (84 GPUs total)
- 58 TB of fast access parallel file system
- 4x QDR Infiniband Interconnect

Required for higher rates and complexer algorithms

Implementation Details

Karlsruhe Institute of Technology

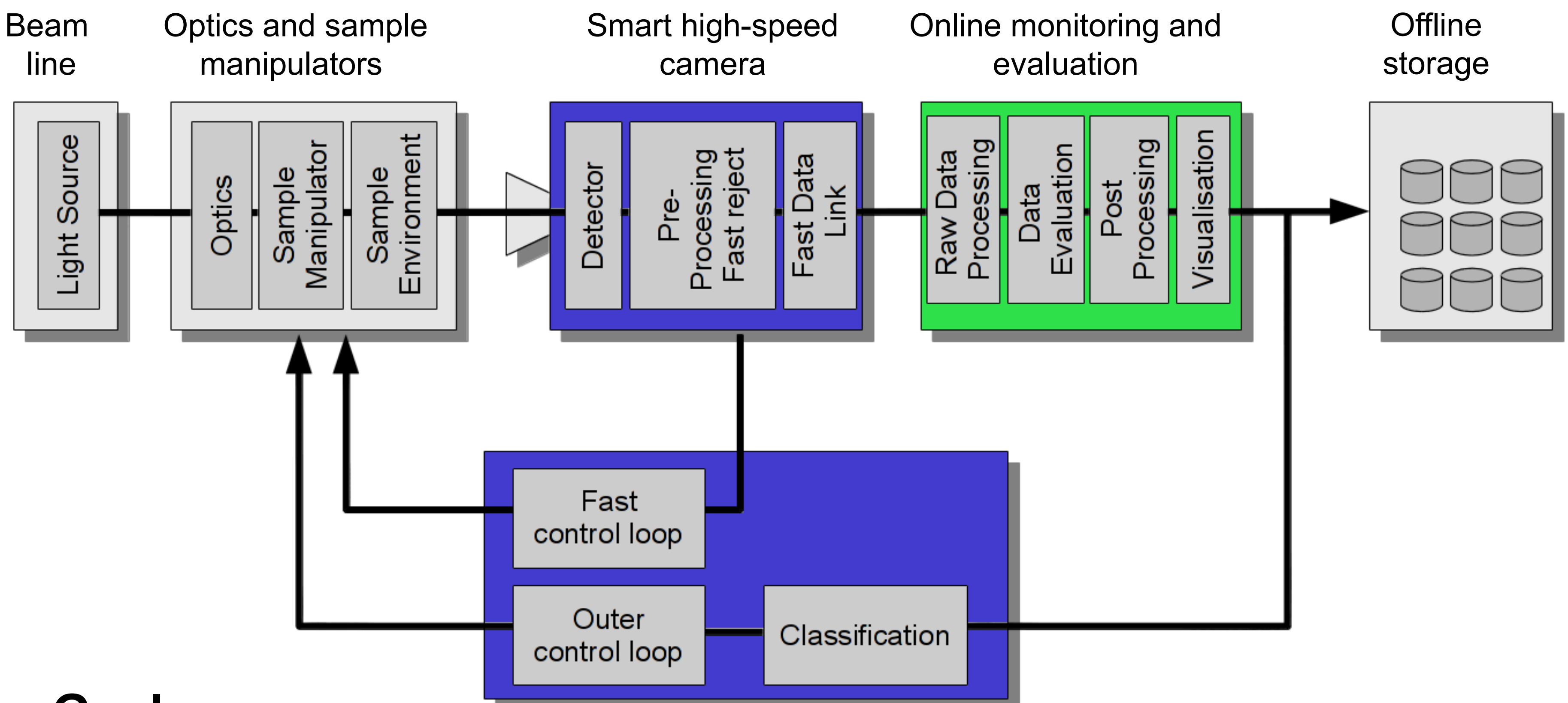
Dependencies & Tools

- Standard C99
- GLib/GObject + GObject introspection
- OpenCL 1.1 or 1.2
- ZeroMQ 3.2
- Nightly builds and unit test execution via Jenkins
- API documentation built with Gtk-Doc, manual with Sphinx

High-level architecture

- Core framework manages OpenCL resources, graph and execution
- Shared library plugins implement further functionality
  - Reading, writing, filtering, …
  - Algorithms

UFO – Ultrafast X-Ray Imaging



Goals

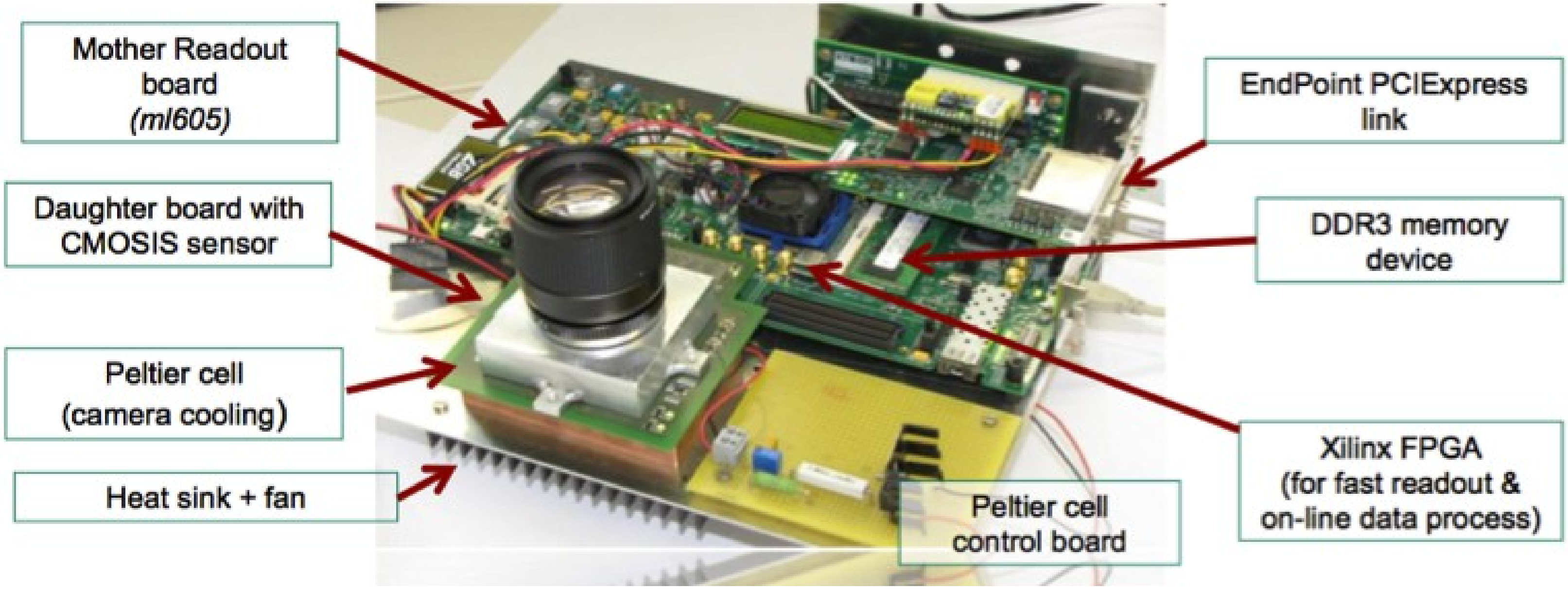
- High speed tomography
- Increase sample throughput
- Tomography of temporal processes
- Allow interactive quality assessment

- Enable data driven control
  - Auto-tunning optical system
  - Tracking dynamic processes
  - Finding area of interest

High-throughput camera platform

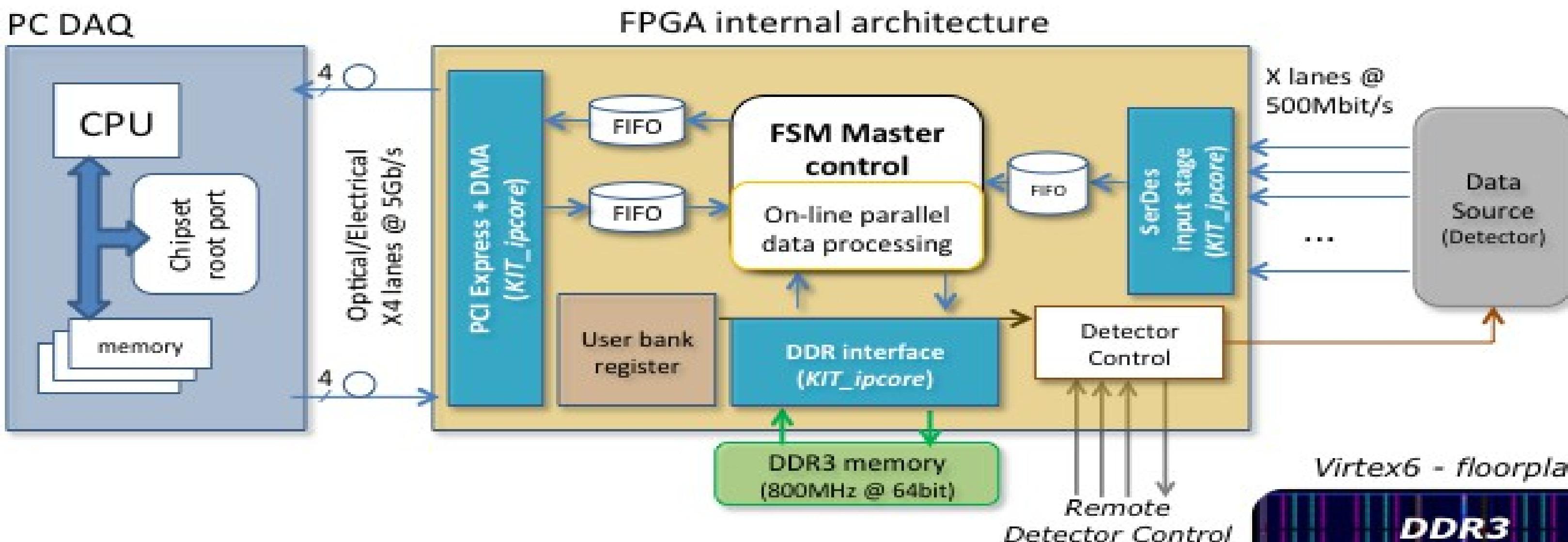
- Remove bottleneck between camera and GPU
  - High speed interface with PCI Express
  - Trigger logic, Compression
- Modular design

- Prototype:

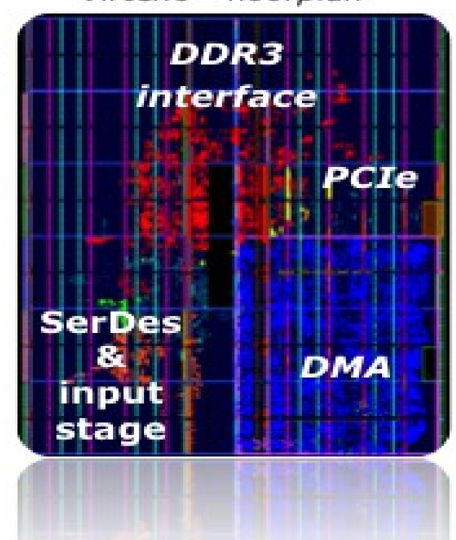


Flexible high-throughput FPGA platform

HELMHOLTZ ASSOCIATION



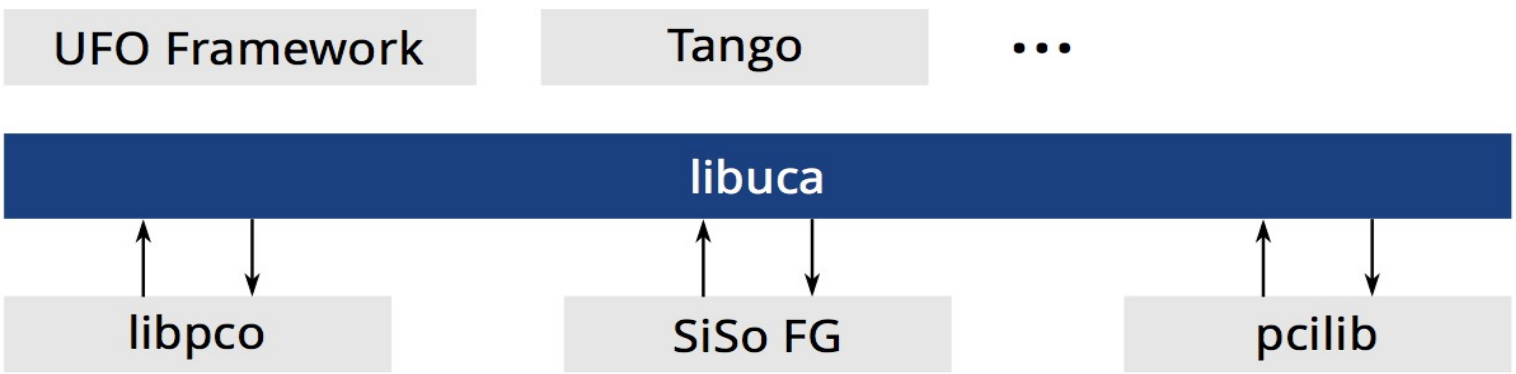
Virtex6 - floorplan



- ✓ Three logic cores have been developed for a flexible high-throughput platform
  - ✓ PCIe-Bus Master DMA readout architecture
  - ✓ Multi-port high speed DDR3 interface
  - ✓ Configurable 2..16 bits "SerDes" (Serializers /Deserializers) architecture
- ✓ PCI Express/DMA Linux 32-64 bits driver with ring buffer data management
- ✓ Integration in the parallel GPU/CPU computing framework
  - 64-bit Linux drivers

Drivers and Camera Abstraction

- Generalized access to streaming cameras (C-API)
- 64-bit linux support for PCO cameras
- Licensed under LGPL with permission from PCO
- TANGO driver



KIT high-throughput camera platform

UFO Computing Infrastructure

Camera

PCO.edge
PCO.dimax
PCO.4000

CameraLink
850MB/s

Ethernet
10 Gb/s

Storage

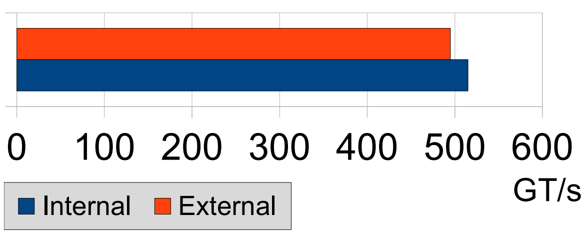
LSDF
Large Scale Data Facility

External PCIe x16 (8 GB/s)

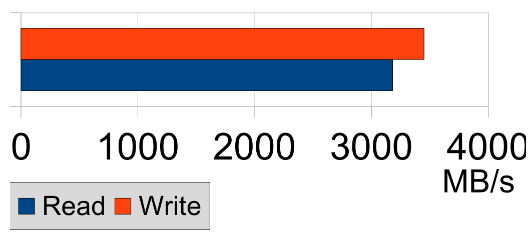
SFF8088 (2.4 GB/s)

<u>SuperMicro 7046GT-TRF </u>(Dual Intel 5520 Chipset)
CPU: 2 x Xeon X5650 (total 12 cores at 2.66 Ghz)
GPUs: 4 x GTX590 External
Memory: 96 GB / 12 DDR3 slots (192GB max)
Network: Intel 82598EB (10 Gb/s)
Camera Link Frame Grabber (850 MB/s)
Storage: Areca ARC-1880-ix-12 SAS Raid
  16 x Hitachi A7K200 (Raid6)
   8 x Samsung 840 Pro 510 (Raid0)

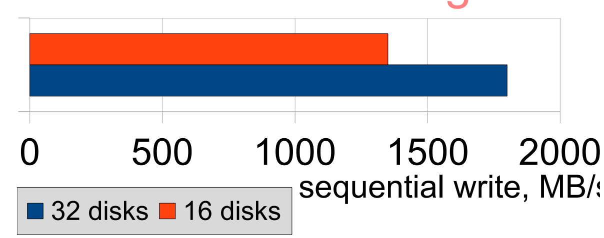
External GPU Box



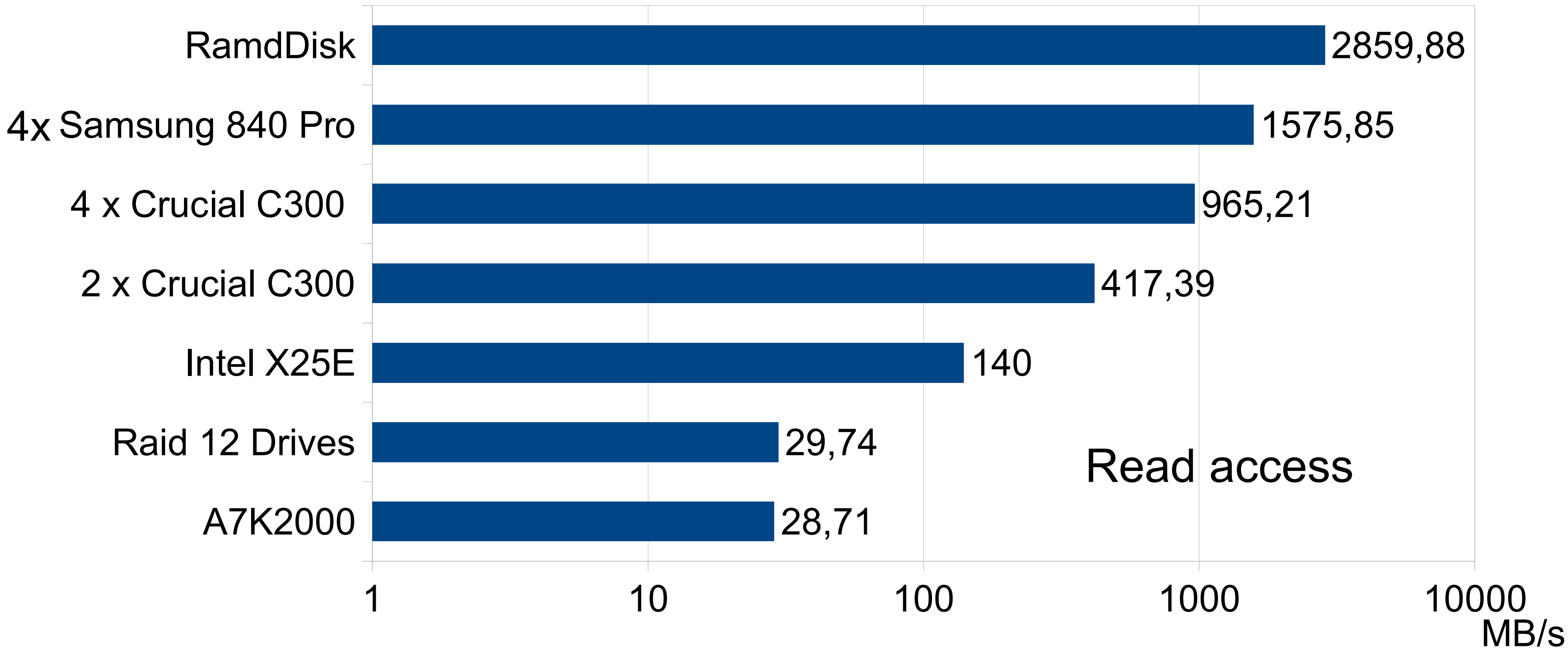
SSD Raid



SAS Attached Storage

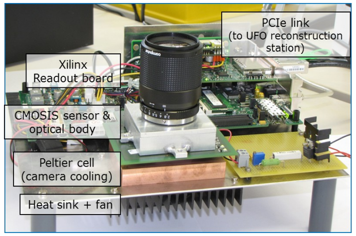


Handling large data sets

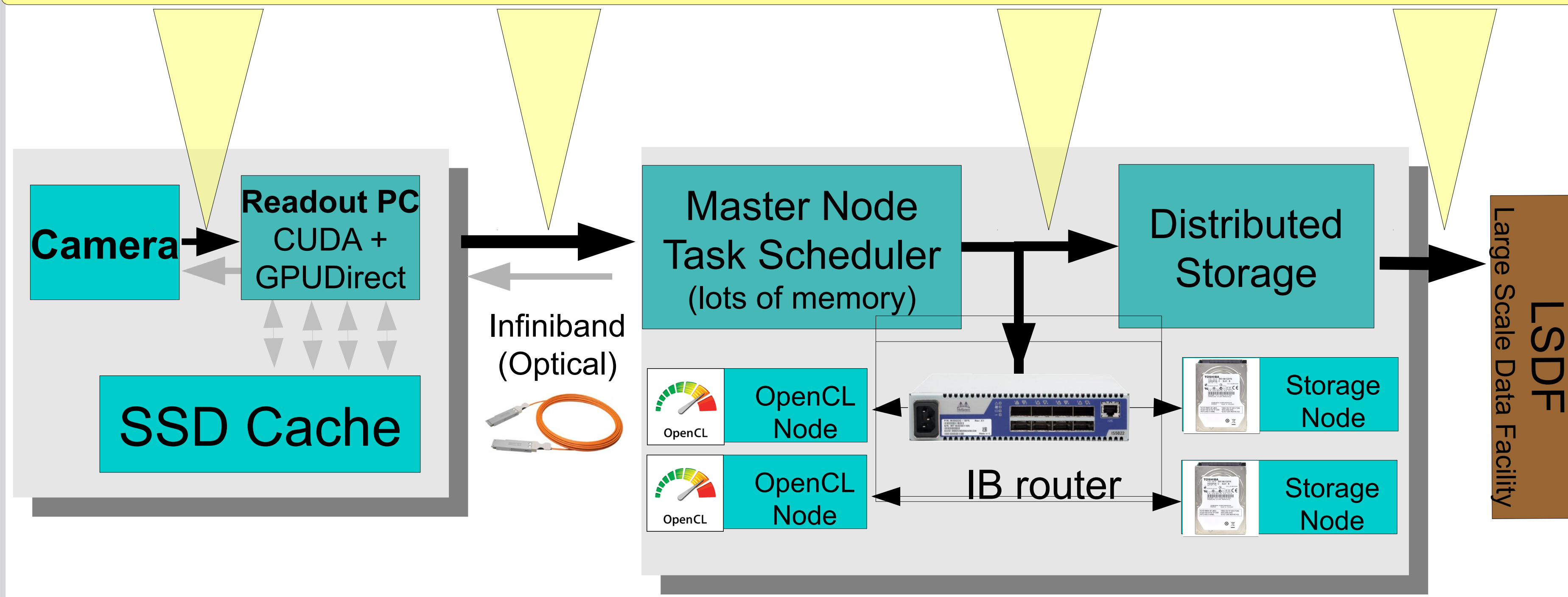


Using SSD drives may significantly increase random access performance to the data sets which are not fitting in memory completely. The big arrays of magnetic hard drives will not help unless multiple readers involved.

Scaling up to Cluster



Transfer rates up to 8 GB/s   4GB/s   <2 GB/s   0,25 GB/s



Summary

- GPUs are a powerful tool for synchrotron applications
- UFO framework support development and management of optimized code for parallel architectures
- FBP is now faster than DAQ
  - Throughput ~1 Gbyte/sec

Next generation of applications (> 1 GByte/sec) requires:

- Clustered DAQ systems for high-throughput
- GPU clusters (and software that supports these!)

- AND technologies are getting more complicated
  → More common activities required

- Getting involved – More information: http://ufo.kit.edu

Institute for Data Processing and Electronics
Karlsruhe Institute of Technology