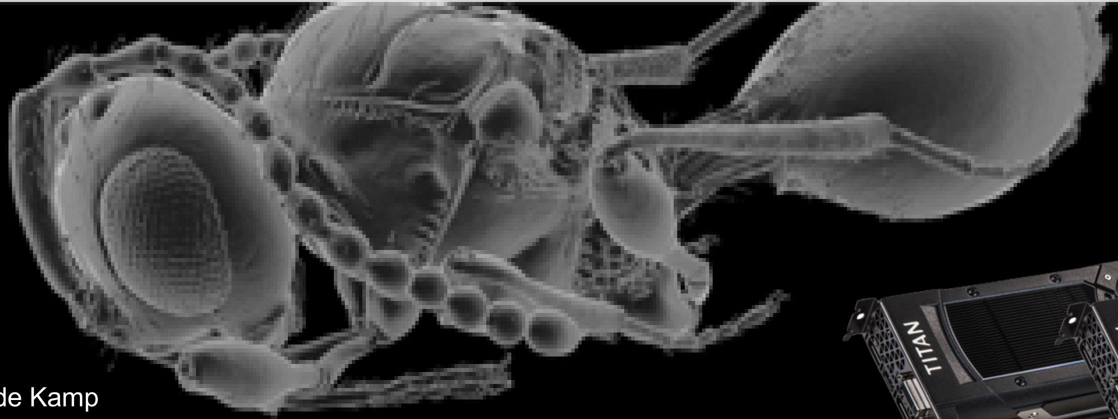
Karlsruhe Institute of Technology

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES

Fachgruppe “Prozessdatenverarbeitung”

Andreas Kopmann

Institute for Data Processing and Electronics



T. van de Kamp

„Prozessdatenverarbeitung“

- lat. procedere = voranschreiten
- „Als Prozessdatenverarbeitung […] bezeichnet man den **Teil der Informatik**, der sich mit den Prozessen zur **Überwachung und Steuerung** externer Geräte befasst.
- Gegenstand der Prozessdatenverarbeitung PDV ist die informationstechnische Verarbeitung der **Daten von technischen Prozessen**, […] im Gegensatz zur herkömmlichen Datenverarbeitung EDV, in der meist **"Bürodaten"** verarbeitet werden.“ (Wikipedia)

Our mission:

> *„Develop and advance instrumentation to operate large-scale scientific experiments. This includes (slow) control, detector readout, online data processing and data management.“*

Karlsruhe Institute of Technology

Milestones in IT history

3rd industrial revolution

1970
Micro-processors

1979
68000

1983-95
Transputers
Occam

1990
FPGAs

1995
GPS (civil use)

1985
GUIs

1990
Internet / www

1950-60
Control theory

1961-72
Apollo progr.
Analog comp.

1974
Automatic Control with µC

1985
Real-time OS
RTOS-UH

1988
Computer algebra
Mathematica

1985
Relational DBs

1995
mysql

New data/digital age?

2005
Multi-core

2007
GPU
CUDA

2008
Science 2.0
"Collaboration"

2011
Industry 4.0

2000
NoSQL

2002
Digital age

2007
Data-intensive *science*

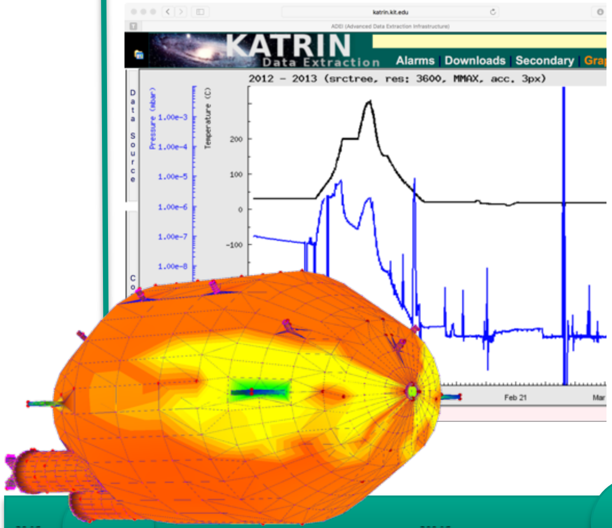
1960 — **1970** — **1980** — **1990** — **2000** — **2010**

PDV focus topics

Karlsruhe Institute of Technology

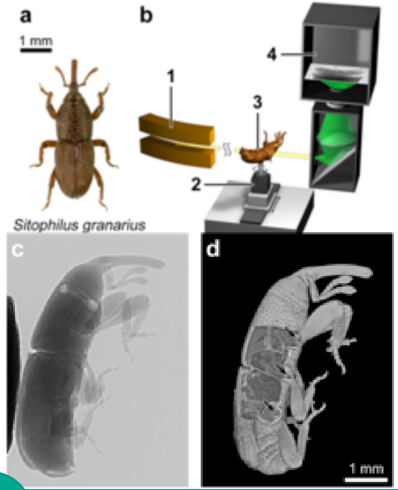
1 Online data management

Web technologies
Databases, Data aggregation



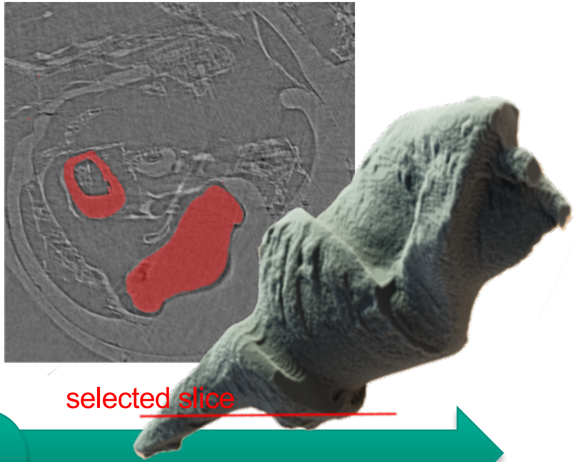
2 Scientific computing with GPUs

Hardware-awareness
Data streaming



3 Analysis environments for large datasets

Virtualization, remote access
3D web visualization

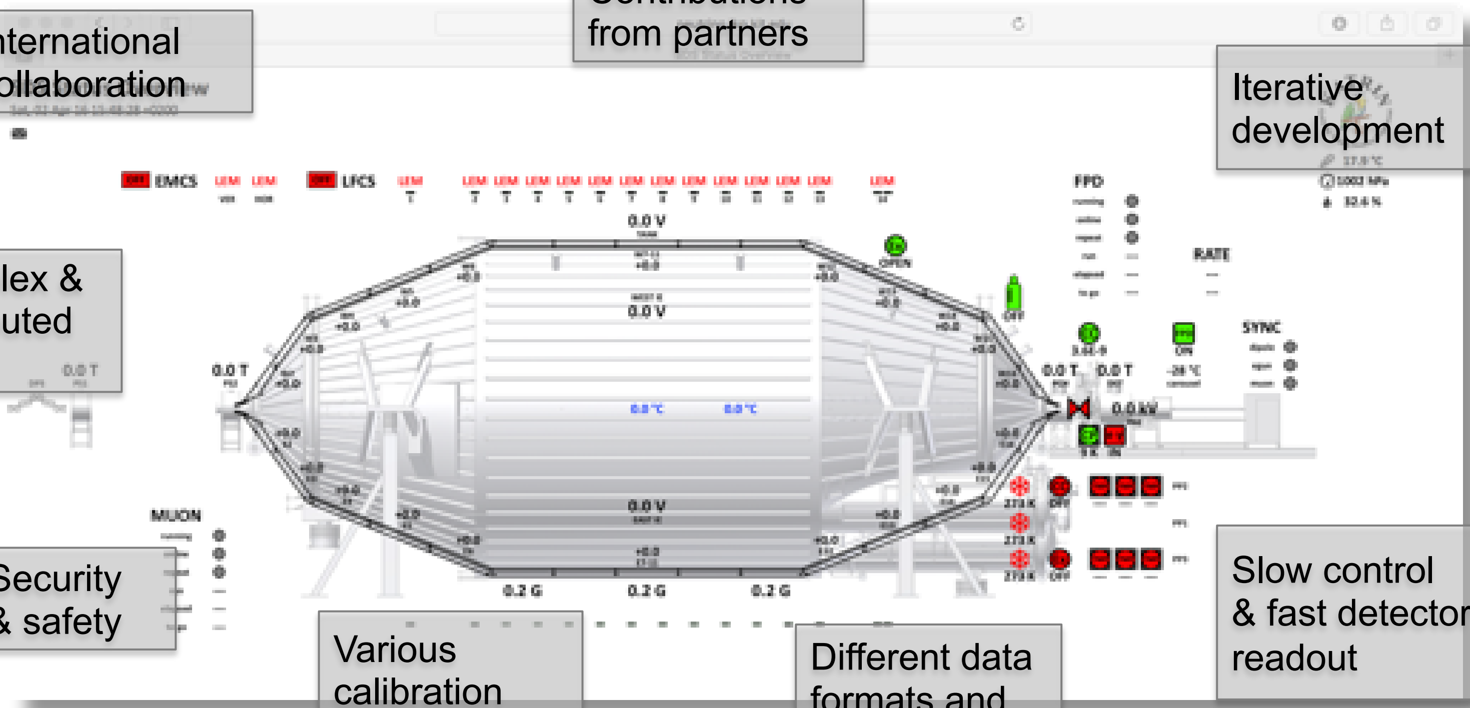


2007

2010

2013

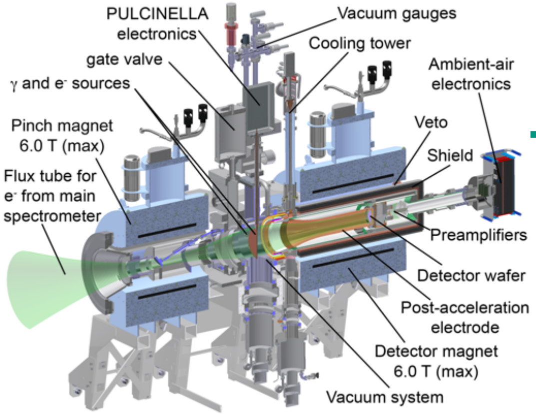
1 Manage data in large-scale experiments



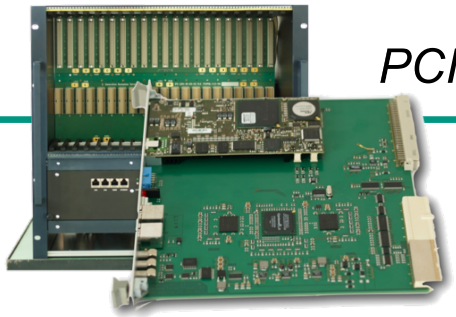
- Contributions from partners
- International collaboration
- Iterative development
- Complex & distributed setup
- Security & safety
- Various calibration devices
- Different data formats and databases
- Slow control & fast detector readout

Detector readout

KATRIN detector

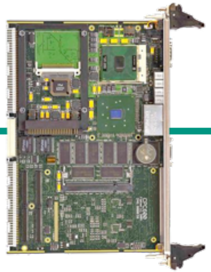


IPE DAQ electronics



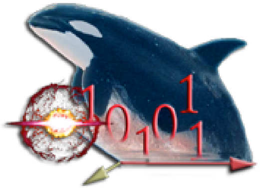
PCIe

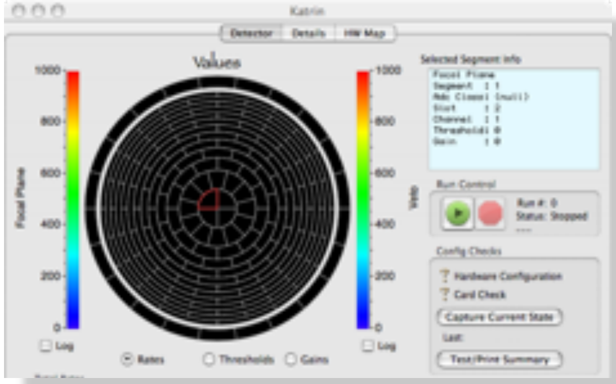
Crate PC



10GE

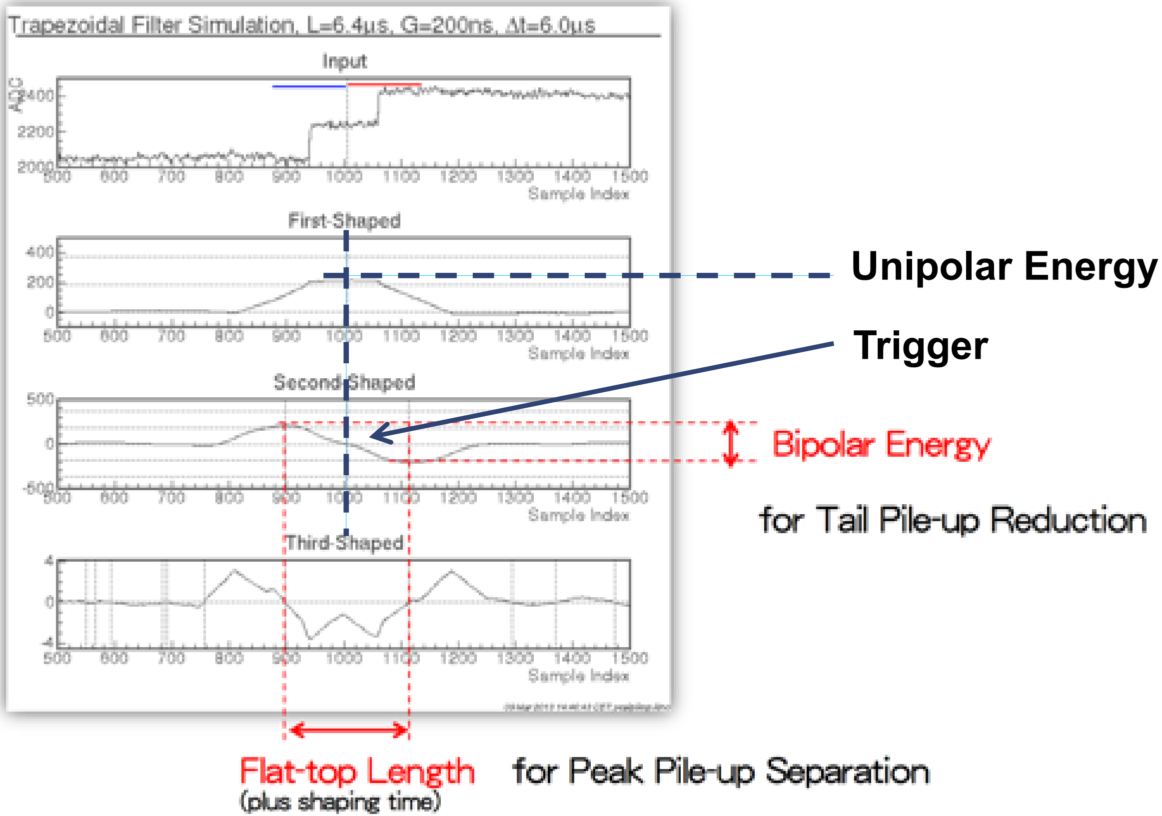
Orca





- Data rate 125 MB/s eq. 5 Mev/s
- Graphical operation by Orca
- Scriptable; integrated with slow control
- Flexible

Event trigger by trapezidoial filters



by S Enomoto and
D Tschernikhovski

Slow control

Experimental monitoring
(mostly) with LabView

A Beglarian, H Bouquet

Magnet monitor

3D Temperatures

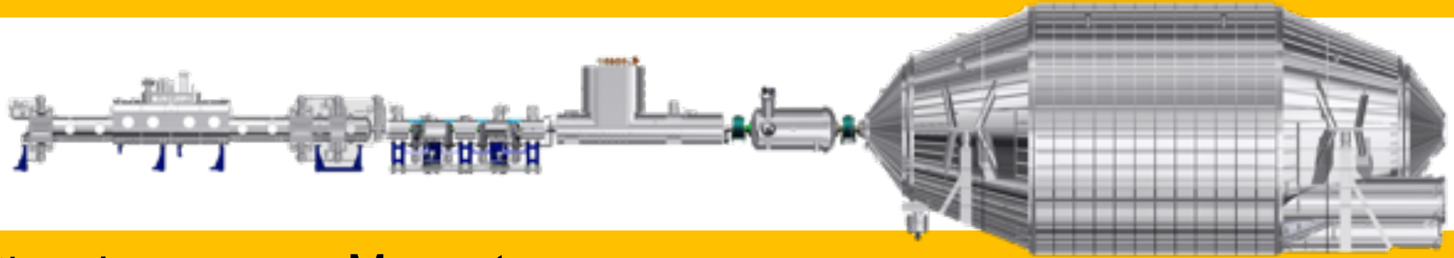
Air Coils

HV

Detector

E-gun

FBM



Tritium loops

Magnets

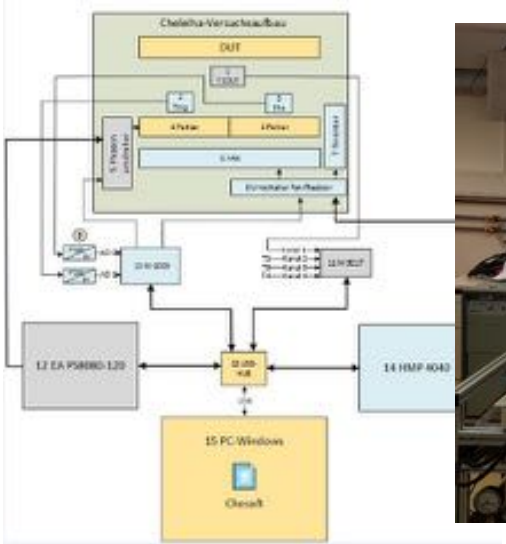
Control system / PLC

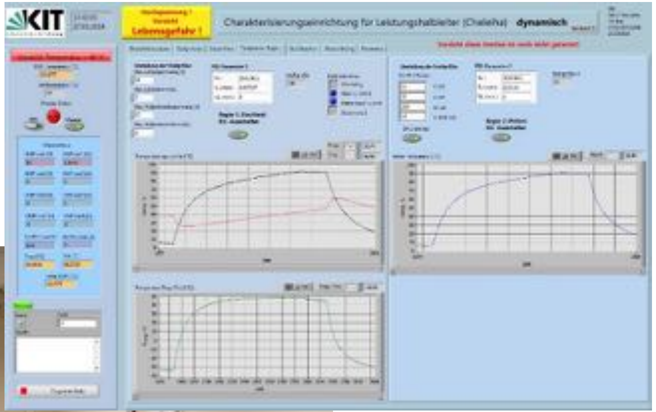
Vacuum

Heating / Cooling

Lab automation

- Test setup for characterization of power semiconductors (Cheleiha)

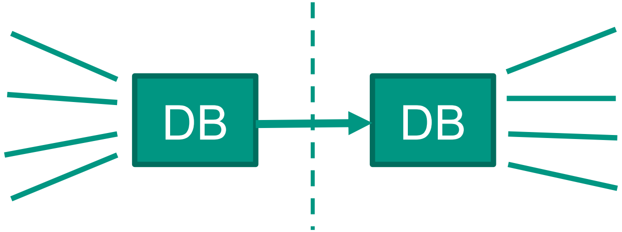




H. Demattio

Archival

- All monitoring data is stored in MS SQL databases
  - Access via ZEUS (LabView), C# libraries, files, etc.
- OPC Server interface for control system
- Conversion of time formats
- Isolation of databases by replication
- Management of run phases
- Calibration of raw data in the database
- Export of standard reports



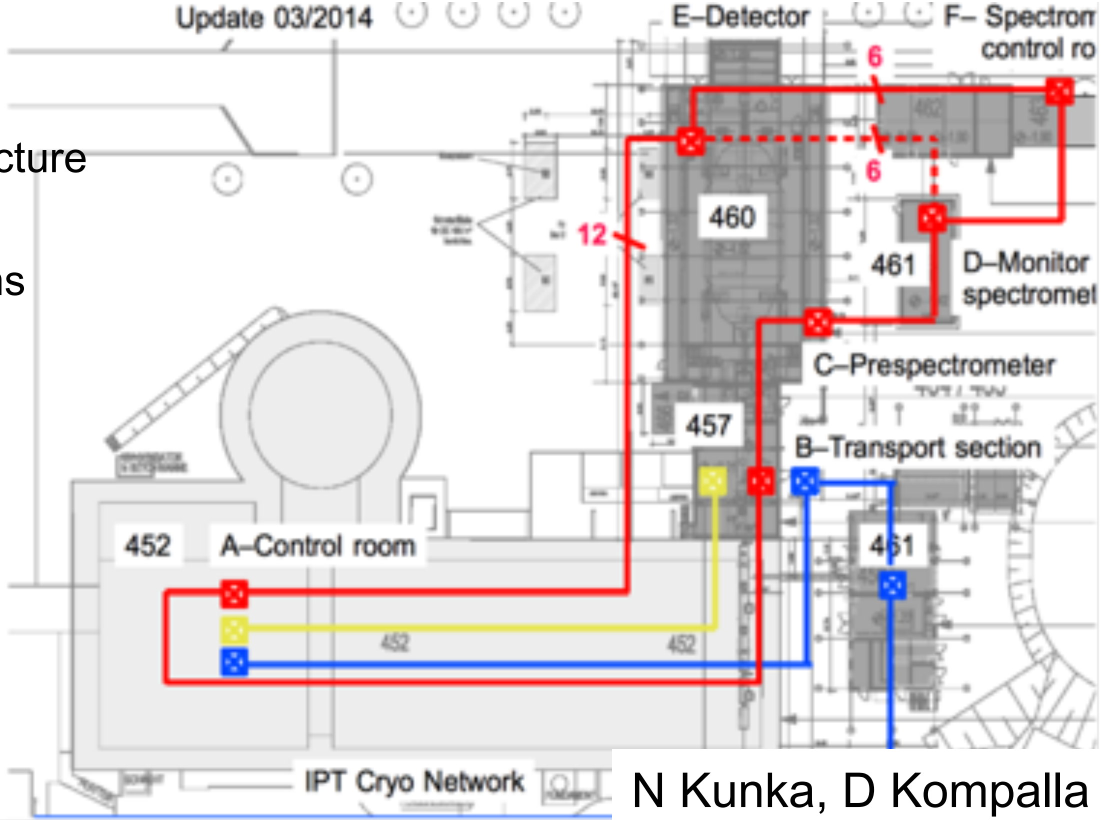
Column	Type
TS	Time
Sensor 1	Double
Sensor 2	Double
Sensor 3	Double

+ config information

P Rohr, H Frankrone

LAN Infrastructure

- Hierarchical network layer for isolation of services
- High(er) availability by ring structure
- Automated file transfer
- Virtualization of archival systems



N Kunka, D Kompalla

Advancing DAQ 1:
How to provide data to international collaborations?

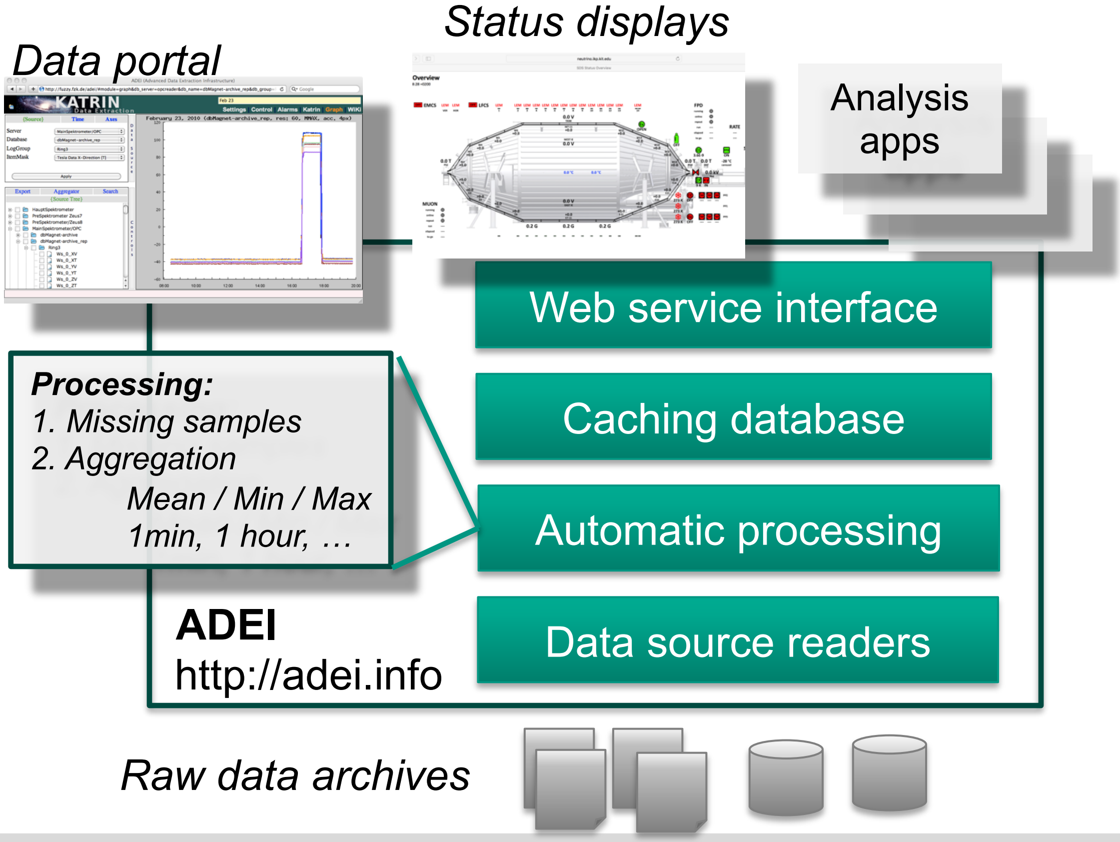
Karlsruhe Institute of Technology

- Many subsystems with various data formats, some proprietary
- Graphical near-time display required, fast browsing comparison of different devices
- Web-based; no installation required; for many devices
- Simplify communication on data by unique links to views
- Standard interfaces for applications; common naming scheme

Advanced data extraction infrastructure ADEI

S Chilingaryan

- **Integration + interpretation** of all data (slow control)
- Extensible **early** data **quality checks**
- **Uniform interface** for data analysis apps
- **Web data portal** for the collaboration
  - Aggregation techniques for fast access
  - Export in standard formats
- Experiment specific **status displays**



ADEI roadmap

Karlsruhe Institute of Technology

2007
TOSKA
W7X coil test

KA 2 RIN
Neutrino-Exp.

2009
KITcube
Meteorology

2010
SEVAN Space weather network

2013
BESS Renew. energies

2017
Updates

kaas.kit.edu

KATRIN-as-a-service

New features:

- 2008: Status display, detector data
- 2009: Download manager
- 2010: Tango Reader
- 2013: BORA configurable status display
- 2017: Cloud infrastructure, scalable services

Student work:

- 2008: Mobile devices
- 2009: Html5 Canvas
- 2013: BORA displays for KATRIN & BESS

2007 — **2008** — **2009** — **2010** — **2013** — **2017**

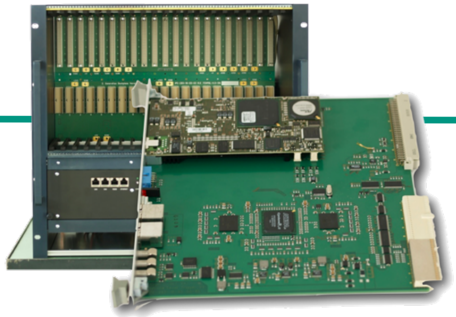
Advancing DAQ 2: How to compute GB/s in real-time?

Detector application

DAQ electronics
Hardware trigger

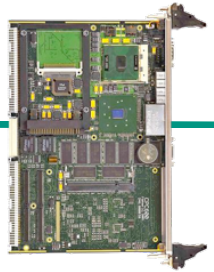
Computing

KATRIN
148 pixel @ 20MHz
Data rate 6 GB/s



PCIe
125 MB/s

red. 50x



Photon science
1-20 Mpixel @ kfps
1-10 GB/s

No trigger

Complex reconstruction

Computing platforms for online applications

Karlsruhe Institute of Technology

- Single core age ended in 2005 -> *parallel programming is required!*
- GPUs are fast, cheap and scalable (up to 4 in a PC)
- GOUs snable TeraFLOP applications: tomography 2TFLOP/GB

Moore's law and CPU

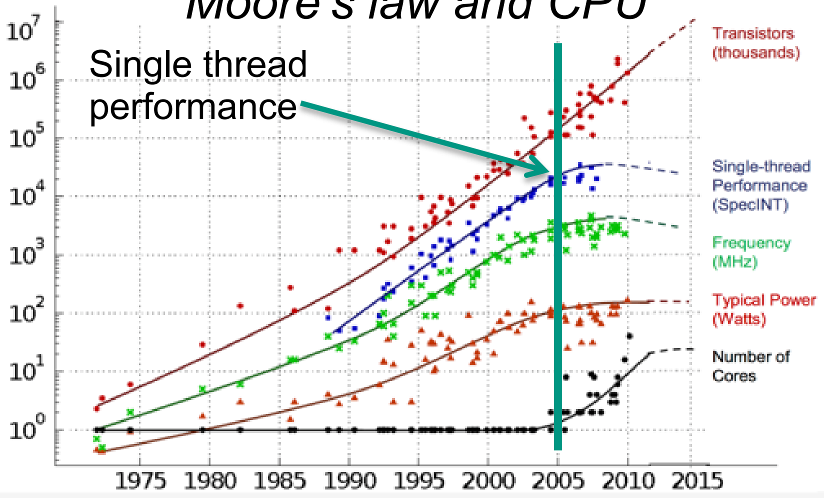
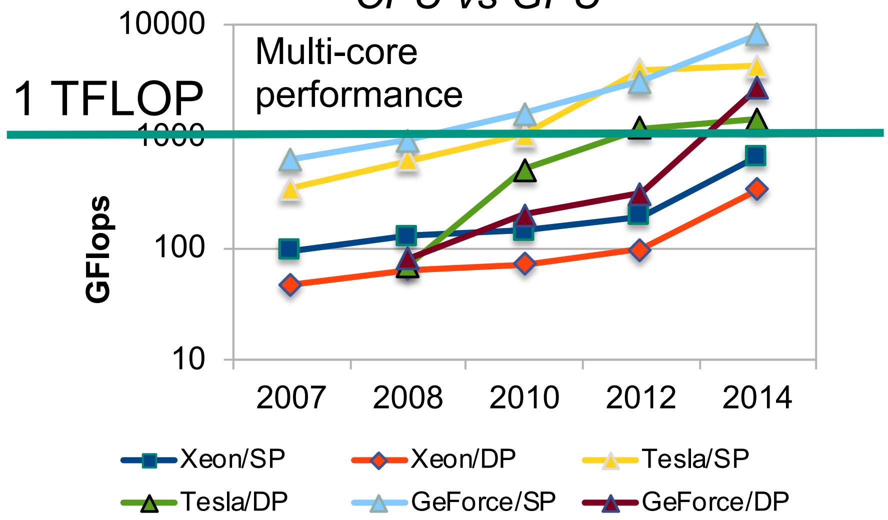


Figure: Chuck Moore, AMD Technology Group CTO

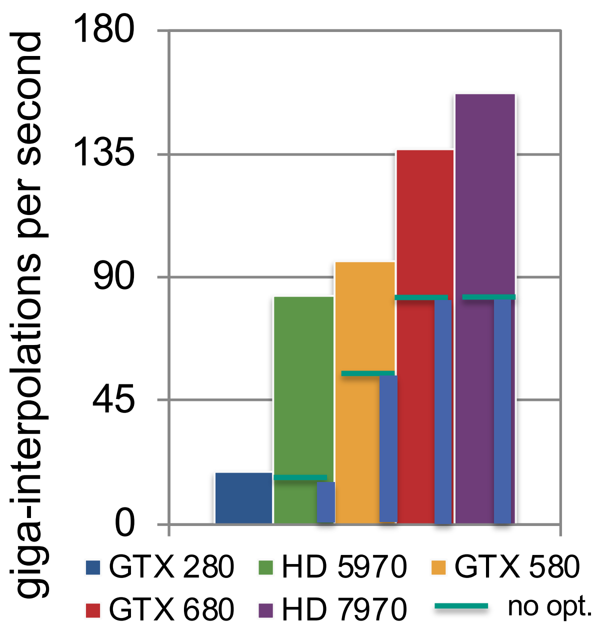
CPU vs GPU



Hardware-aware optimizations

- Example FBP / linear interpolation for different hardware architectures

Reference:
GTX 280, GT200
Uses texture engine



HD 5970, VLIW (+530%)
Multiple independent operations per thread

GTX 580, Fermi (+100%)
Higher kernel performance,
but under-performed texture unit

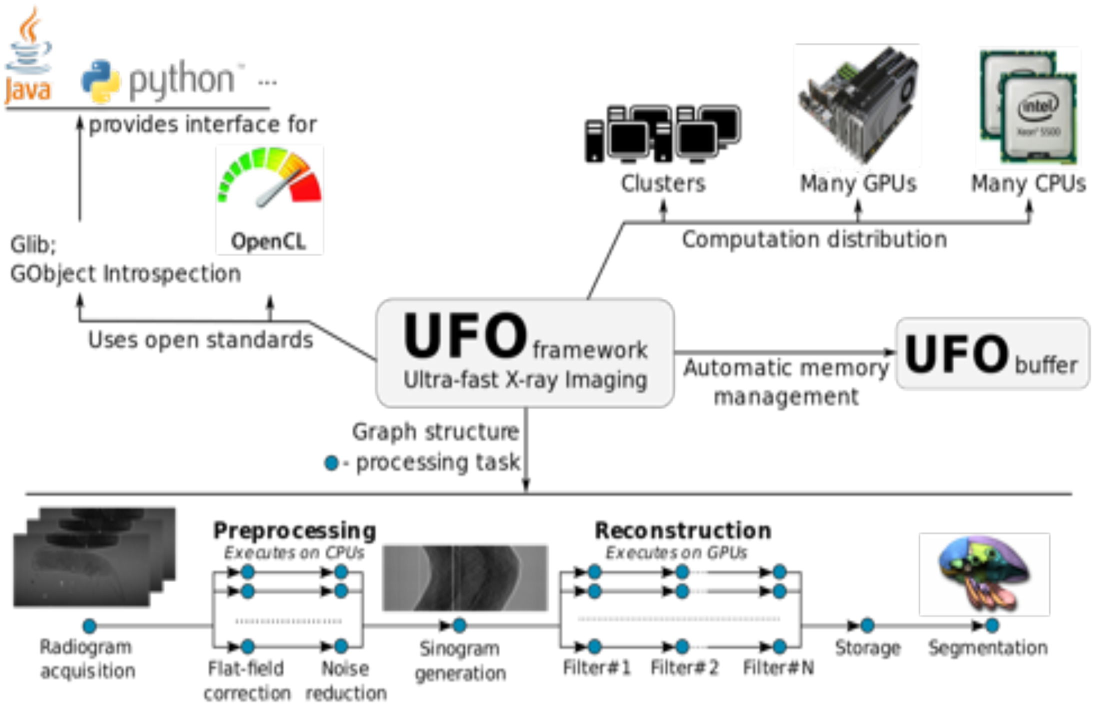
GTX 680, Kepler (+75%)
Low bandwidth of integer instructions,
but high register count

HD 7970, GCN (+95%)
Balance between high performance texture engine and computing kernels

Chilingaryan S et al, IEEE TNS (2011)
Chilingaryan S et al, SBAC-PAD (2019)

Ultra-fast X-ray imaging – "UFO Framework"

- Processing of **data streams**
- (Re-)use of **optimized algorithms**
- Automatic scheduling
  - **CPUs, GPUs**, …
- **Integration** in
  - Control system
  - Analysis tools
- **Easy to use** for
  - Users, Admins, Developers



http://ufo.kit.edu

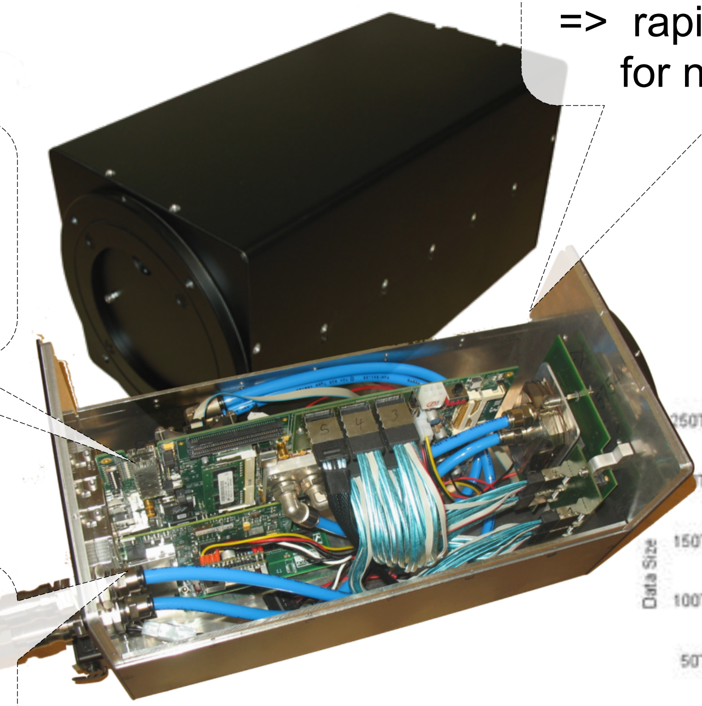
Vogelgesang M et al, Proc HPCC-ICESS (2012)
Vogelgesang M et al, Proc ICALEPS (2013)

Integration with electronics: Smart scientific camera

Modular sensor interface
=> rapid development for new sensors

Embedded FPGA for online data analysis
=> open design

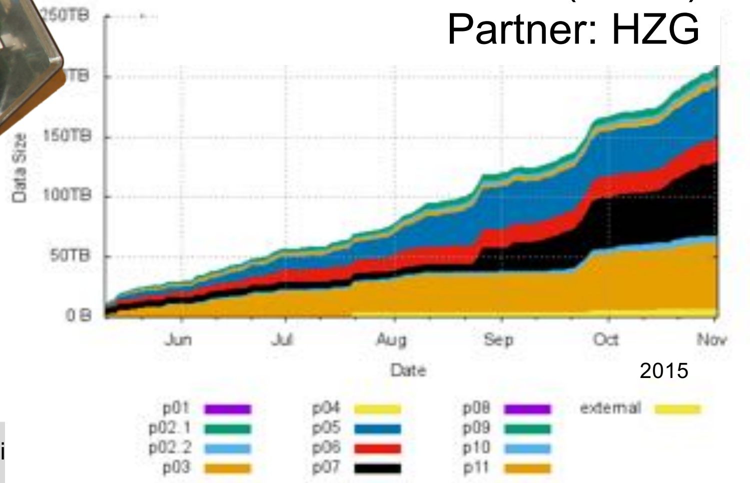
High-throughput link with PCIe
=> streaming to GPU for processing



up to 6 GB/s

Av. image sensors:
CMOS 1-20MPixel, 30-5000fps

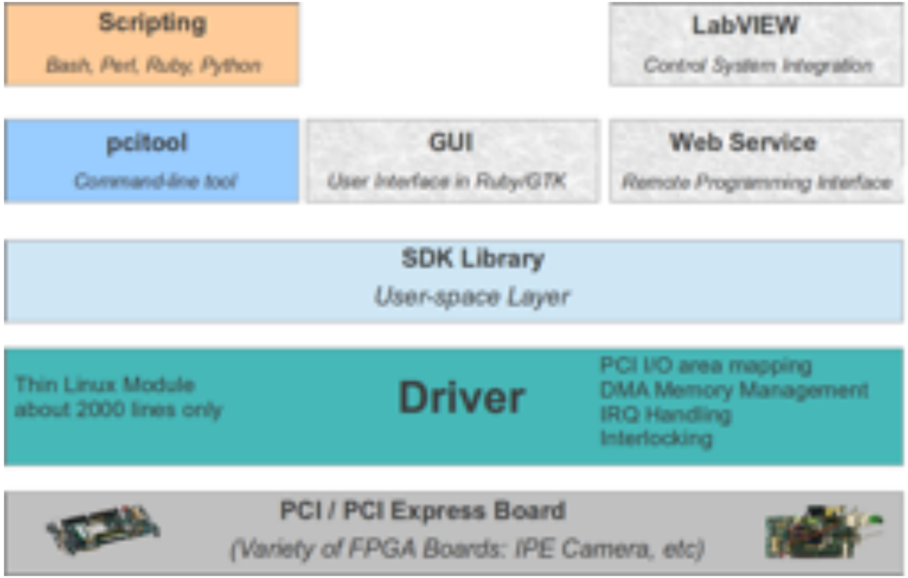
Phase contrast tomography at P07/PETRA III (**black**)
Partner: HZG

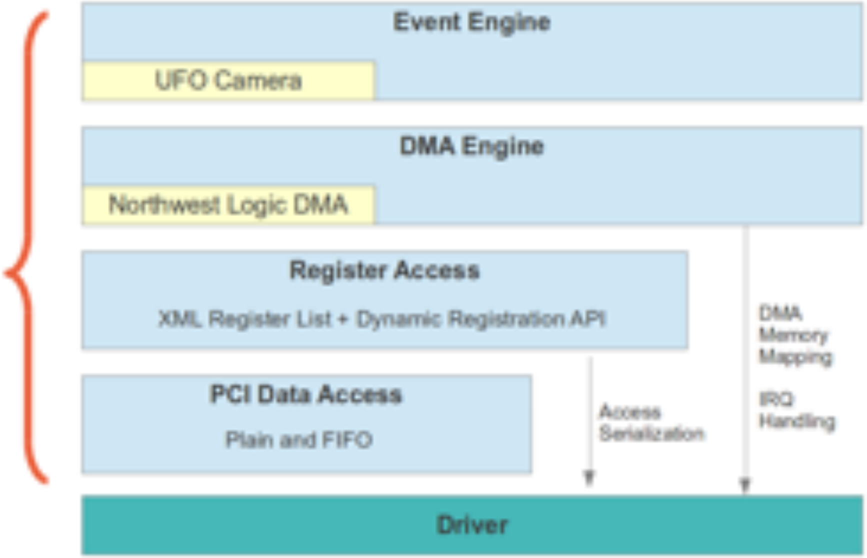


High performance data transfer

Karlsruhe Institute of Technology

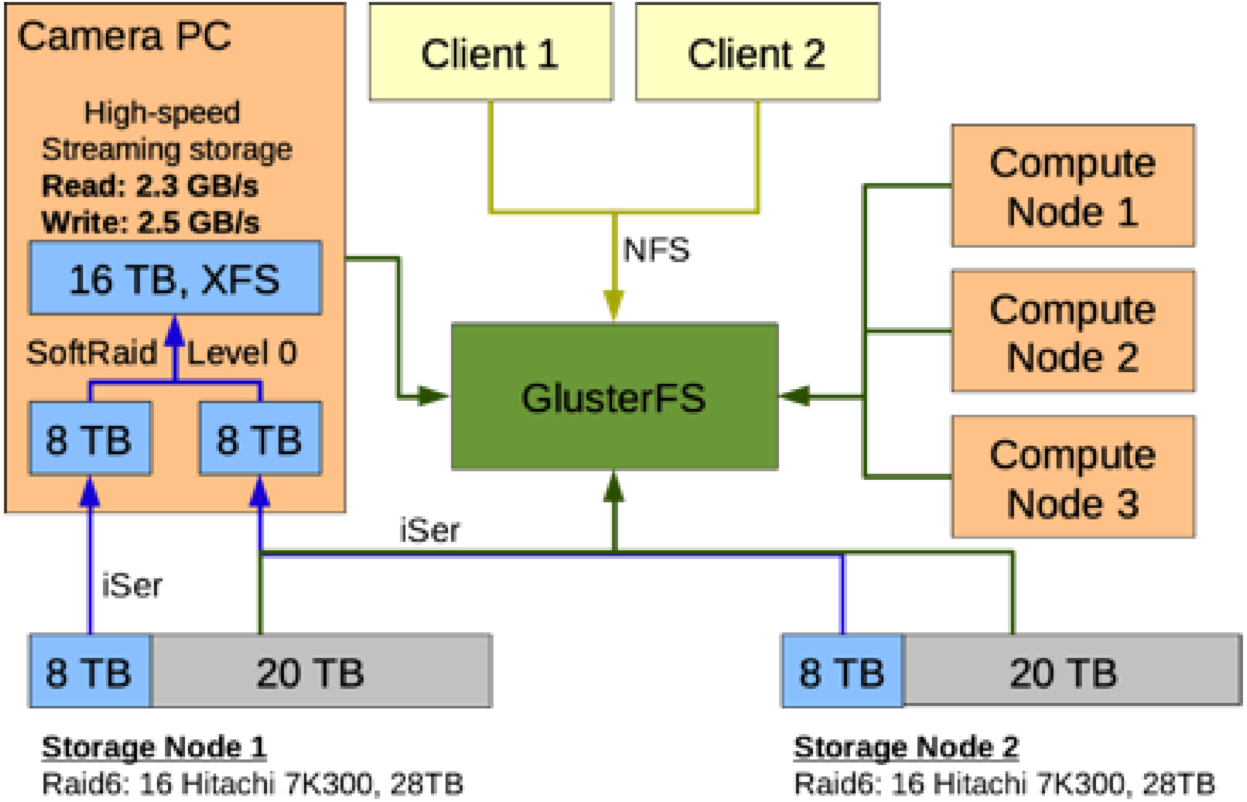
- ALPS - Advanced Linux PCI Services:





S Chilingaryan, IEEE RT2012

UFO real-time storage system



Enabling GPU computing for DAQ systems

funded by UFO, UFO2
coordinated by IPE

Actions:

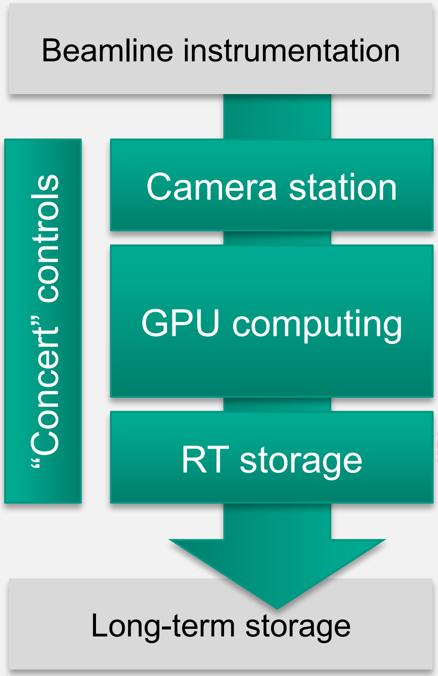
- Hardware-aware programming; orders-of-magnitude speed-up
- Algorithm management with UFO framework; broad camera support
- Computing integrated in experimental environment;

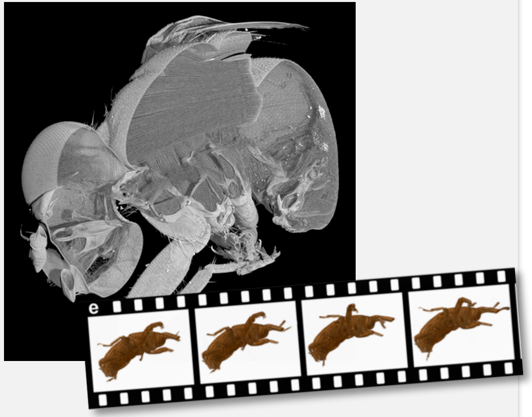
Results for X-ray imaging:

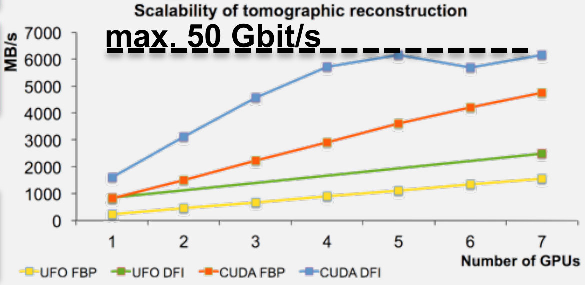
- **Reconstruction in seconds**
- New methods: e.g. **Cine-4D-tomography**
- UFO is in Ubuntu and interoperable
- In use at KIT, HZG, HZDR, Soleil, …

GPU computing is game changer for high-rate DAQ applications

Ultrafast X-ray tomography enabled by GPU computing

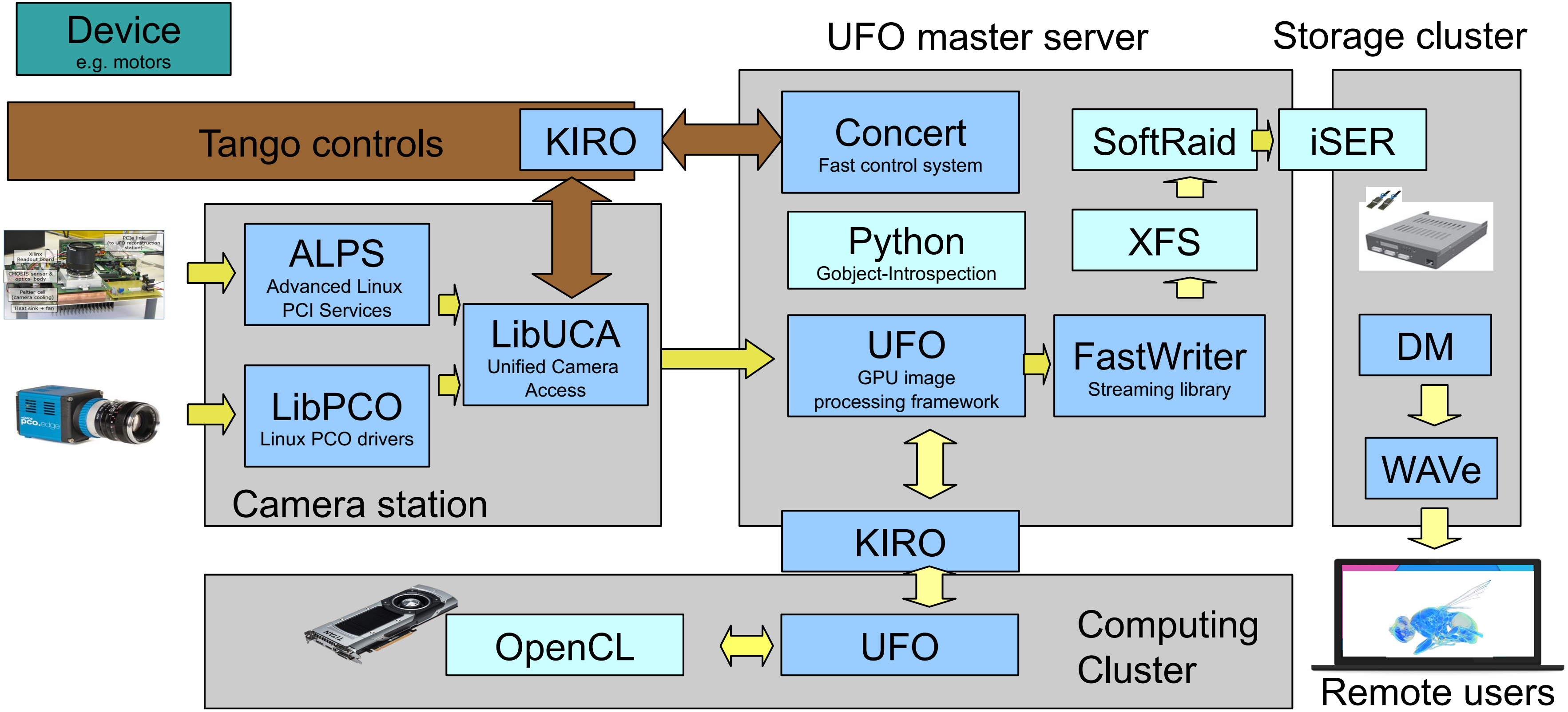






Vogelgesang et al., Proc. SPIE 9967 (2016)

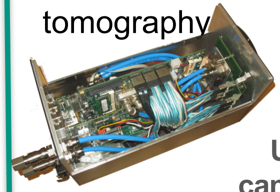
Overview: Software Stack



Computing roadmap

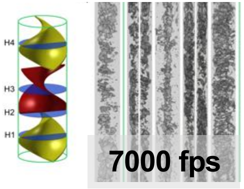
2010
ANKA,
X-ray cine-
tomography

2012
PETRA-3,
Phase contrast
tomography

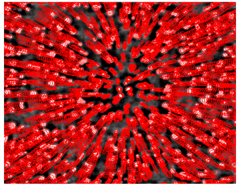


UFO
camera

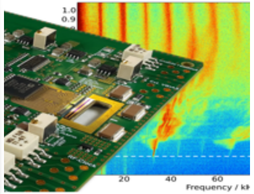
2013
ROFEX, HZDR
Fluid dynamics



2013
uPIV for
thin films



2013
Electron beam
monitoring



New features:

Camera
Driver
(e.g pco)

Control
system
"Concert"

Algebraic
reconstruction

Infiniband
communication

RDMA
FPGA-GPU

Student work:

Cluster
communication

Python to CUDA
Compiler

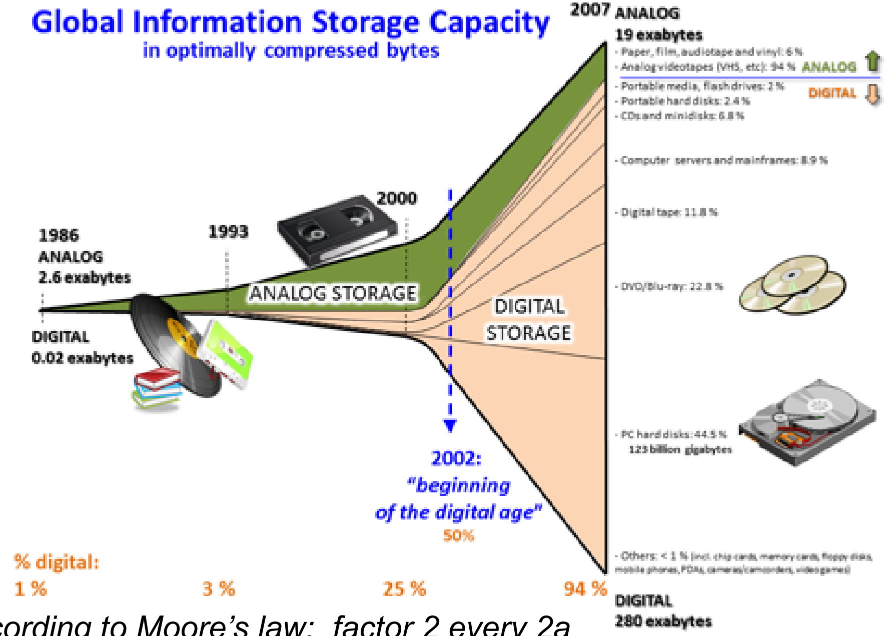
Cristallography
Laminography

GPU Trigger for
HEP experiments

2010 **2012** **2013**

3 Advancing DAQ 3: How to organize analysis for large datasets?

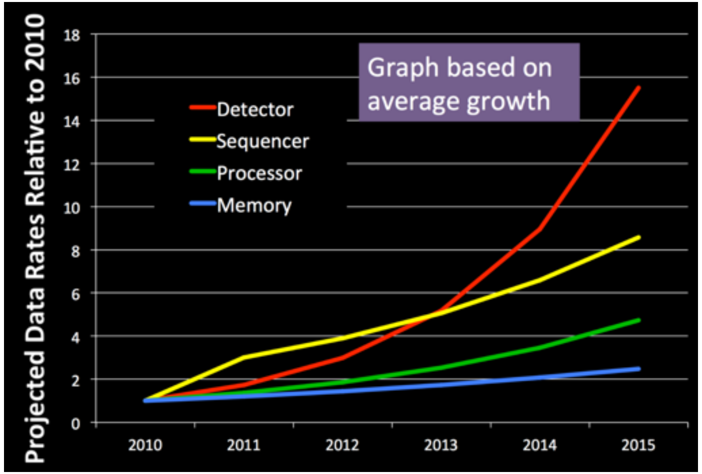
“*Big datasets are so large or complex that traditional data processing applications are inadequate*” (Wikipedia)



According to Moore’s law: factor 2 every 2a

Nersc: “Data growth is outpacing Computer growth”

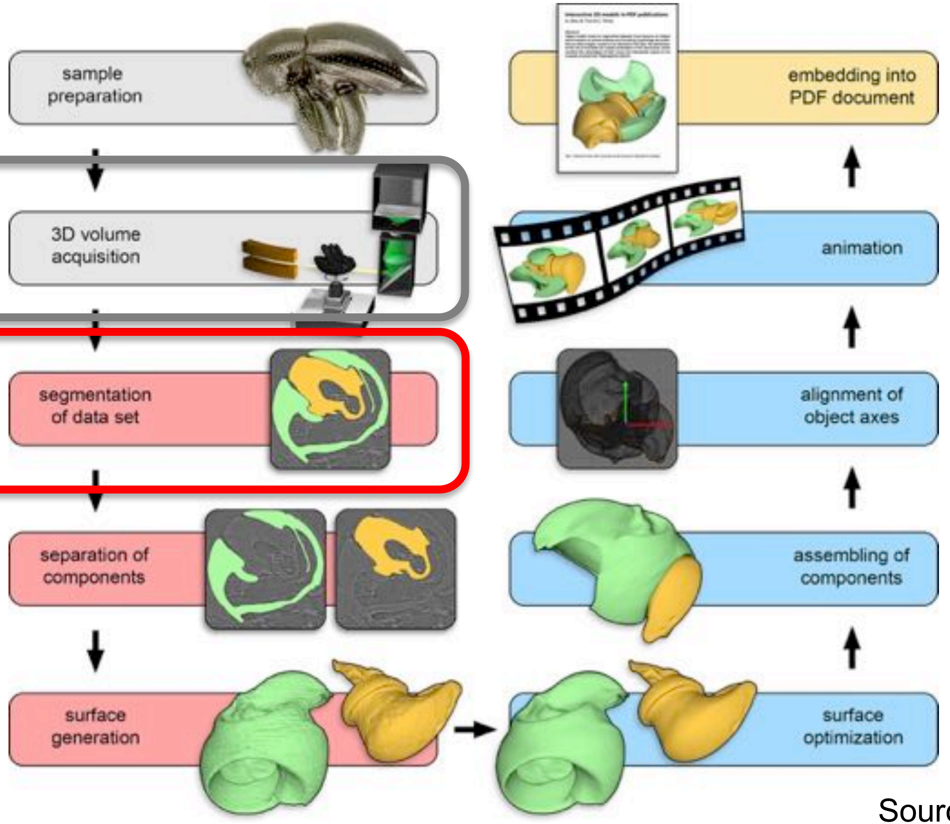
Source: Yelick K, LBNL/Nersc



Analysis workflow in life sciences

Reconstruction
Solved

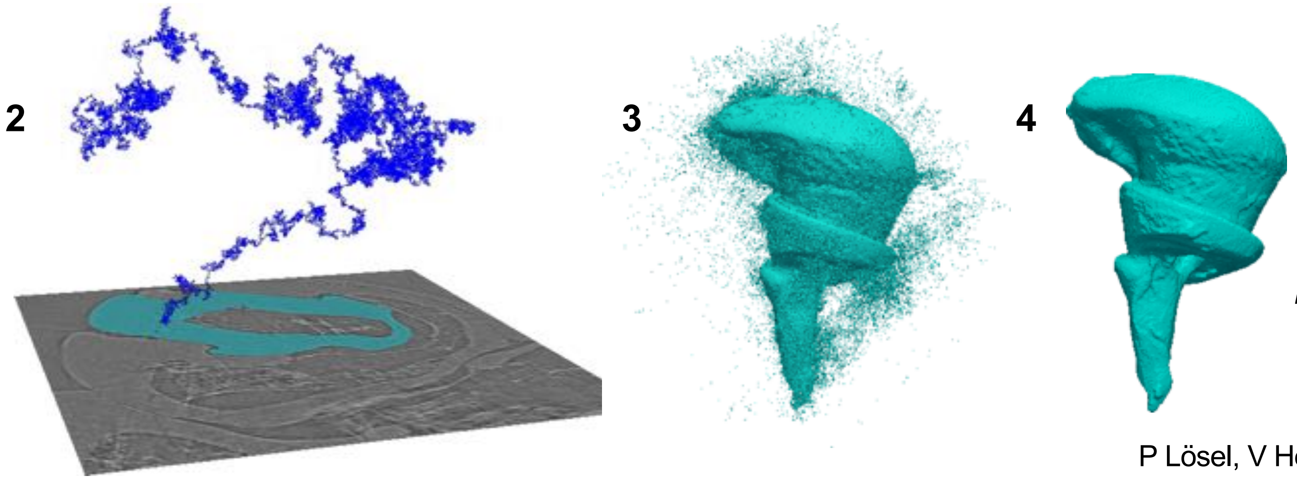
Time-critical
Weeks to months



Source: T van de Kamp, IPS

Segmentation of 3D data

1. Preprocess: Label the segments in some well chosen slices.
2. Start a great number of weighted random walks at each labeled pixel.
3. The number of hits by random walks which were started in the same segment leads to the probability that a voxel belongs to this segment.
4. Postprocessing using Active Contour Method in 3D.



like reco.
req. perform.
of GPUs

P Lösel, V Heuveline, SPIE 2016

casual interpolation

Biomedisa.de

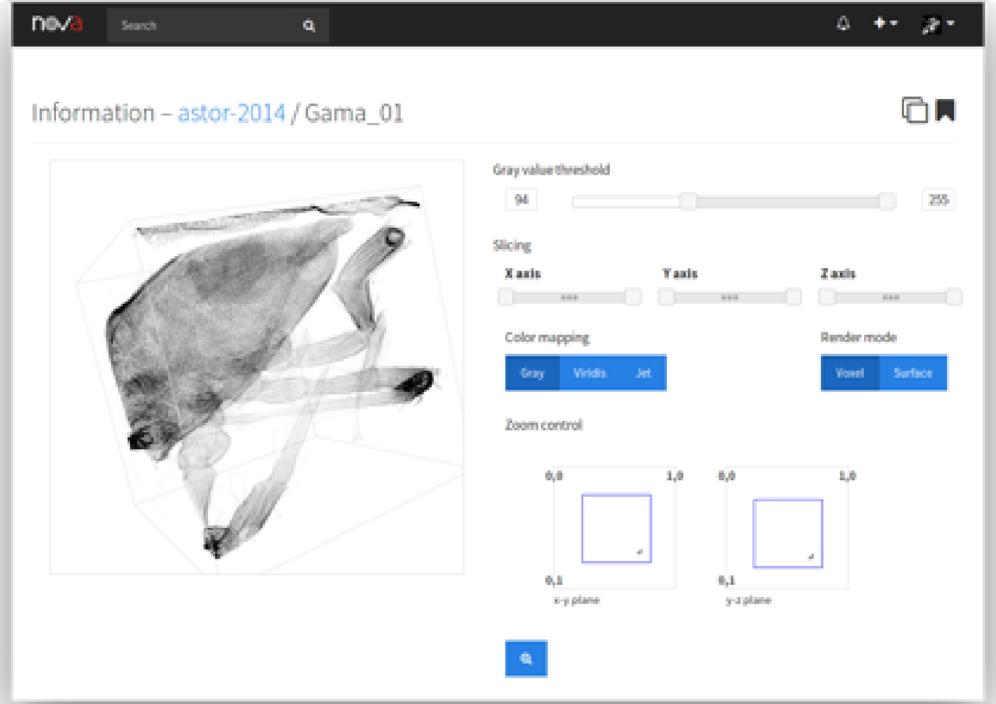
Visualization of 3D data

funded by nova astor

Actions:

- Fast 3D web visualization of raw data
- Pre-processing; aggregation, noise filtering, …
- Progressive loading
- Started to develop collaborative tools

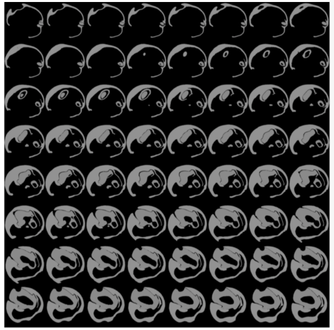
Data catalog with 3D visualization



Results:

- WAVe library
- NOVA data portal (demo)

Slicemap



Visualization of 3D data

funded by nova astor

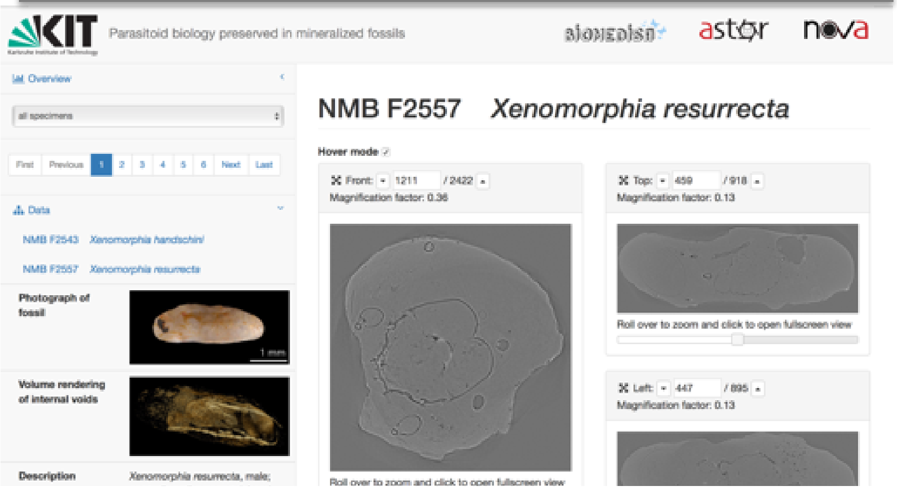
Actions:

- Fast 3D web visualization of raw data
- Pre-processing; aggregation, noise filtering, …
- Progressive loading
- Started to develop collaborative tools

Results:

- WAVe library
- NOVA data portal (demo)
- Fossil study of 1500 samples
  http://www.fossils.kit.edu

Data catalog for fossils project



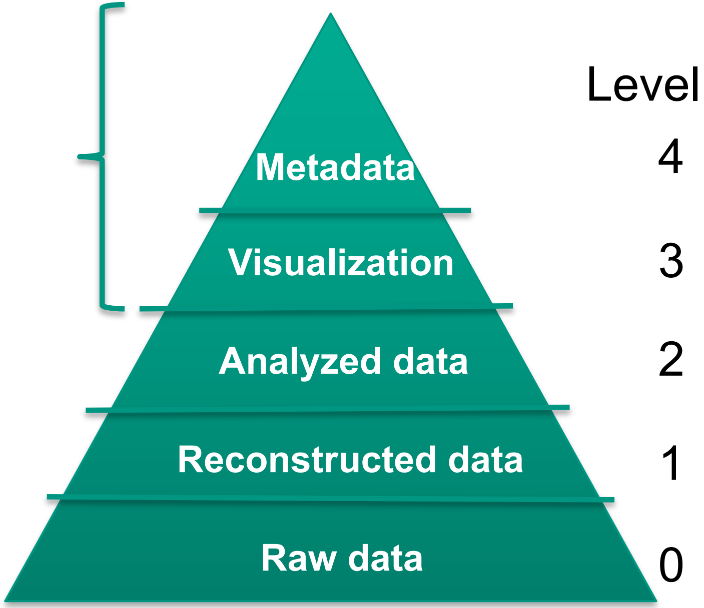
T. van de Kamp et al., Nature Comm (2018)
Review: „astounding, impressive, important"

Hierarchical organization of datasets

- **Long-term archive have long access times**
- Organize large dataset in levels with decreasing size is needed
- Only metadata is available in data catalogs

Solution:

- **Extend classical metadata by visualization data**
- **Develop 3D-visualization for data catalogs**



Conclusion and future directions

Karlsruhe Institute of Technology

1. Data needs to be available at any time
   - Challenge: intelligent metadata, aggregation techniques

Don't store data and forget

2. Computing needs to be parallel – *in best case on GPU*
   - GPUs are at same price or power 20x faster
   - TFLOP applications can be turned online

GPUs are 5-8a ahead of CPUs

3. Remote interactive analysis is possible; large datasets require
   - Better analysis
   - Efficient computing
   - Data management and visualization

DAQ systems be will changed by parallel computing and Big Data requirements