PNI Bonus Award

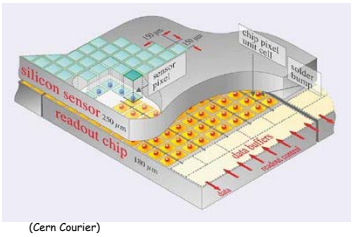
High Data Rate Processing and Analysis Initiative:

A proposal to develop and implement a common framework for
- Data acquisition and reduction
- Online data evaluation
- Simulation, modeling
- Analysis
- Visualization
- Data management

as a collaborative effort of the PNI centers.

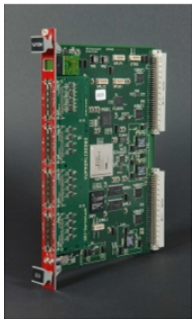
→ "Computing Initiative"

Data Acquisition and Reduction



(Cern Courier)

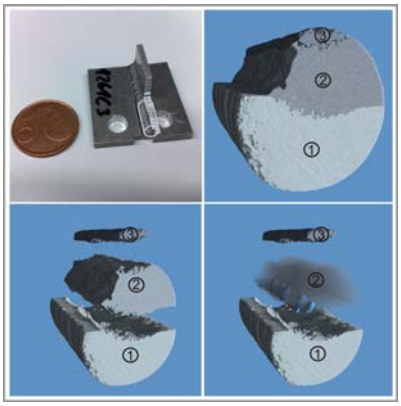
2D detectors generate ~100MB/s,
~ 300 TB/year



Data reduction is of general interest
- Lossless compression, reconstruction
- Digital signal processing
  - FPGAs (field programmable gate arrays), DSPs (digital signal processors): online feature extraction
  - Aim: enhance the flexibility

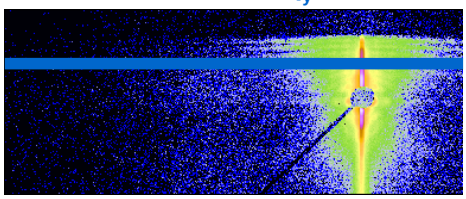
Tomography

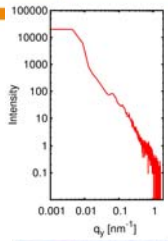
- Common to photon and neutron sources
- Data rate: 500 TB/year (> 2010)
- Online data evaluation
  - Monitoring
  - Quality of the sample
  - Parameter adjustment, statistics
- Reconstruction
  - Parallel computing
  - Offline: high quality reconstruction
- Data management
  - Remote access
  - Standard data format
    - Including description

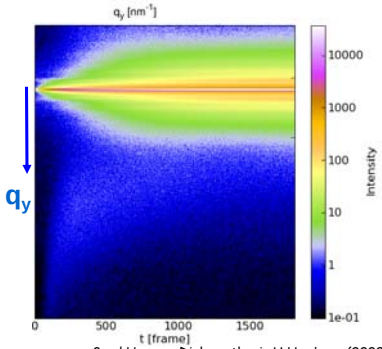


Laser-beam welding in civil aircraft production
(6 mm diameter, 40 keV, HARWI II)

GISAXS/GISANS





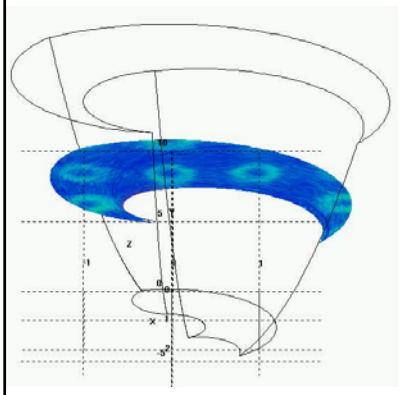


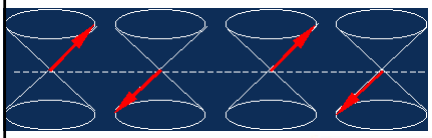
Online data evaluation
- Optimize measurement strategy
- Vital for in situ experiments
- Online modeling

Gerd Herzog, Diploma thesis U Hamburg (2009)
collaboration DESY, TU Munich, MPI Golm

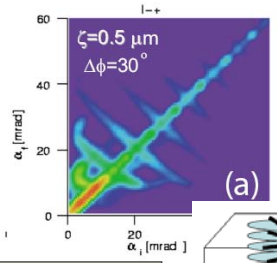
Visualization

Single X-tal TOF Spectroscopy

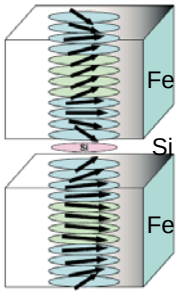




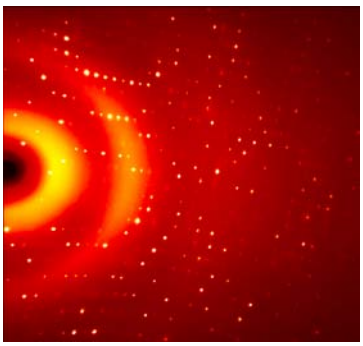
GISANS/GISAXS



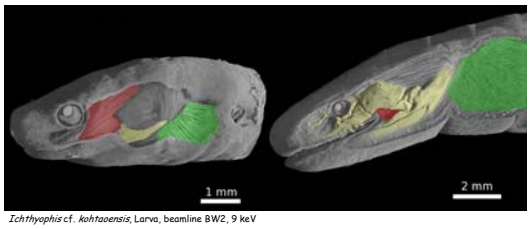
Reciprocal to real space
- Raw data → Q, E
- Visualize Q,E in 4D
- Modeling, comparisons



Visualization



Raw/reconstructed data display
Comparison: simulation vs. data



Ichthyophis cf. *kohtaoensis*, Larva, beamline BW2, 9 keV

Caecilians, comparison of species

In cooperation with T. Kleinteich and A. Haas, University of Hamburg

Computing Initiative
- Survey of existing applications
- Repository of graphics/utility libraries
- Widget sets for GUIs

Standard Data Format

Self describing
- Data structure
- Instrument

NeXus is a potential candidate
- Converters to other formats

Automated documentation
- ELog book

Data Management

- File servers, archive
  - PNI centers provide resources
  - Data lifetime policy
  - Accounting
- Common access method
  - Option: Web-based
- Common authentication scheme

Compute Servers

Analysis, modeling, simulation, fitting

Online and offline
- Interactive
- Batch operation
- Fast data access
- Parallel computing

Computing Initiative
- Parallel computing systems
  - MPI (message passing interface)
  - OpenCL: CPU, GPUs (graphics processing unit)

Objectives

To enhance the scientific outcome of the experiments

- Data reduction: FPGA, DSP, GPU
- Online data evaluation is essential
  - Sample quality, parameters, statistics
- Parallel computing
- Data management
  - Organize disk space
  - Archive
  - Remote access
- Convenient user interfaces

→ Highly integrated solutions

Examples for Collaborative Work

Open source projects

- DANSE, SNS, computational neutron science, kick-off 2006
- CCP4 – Protein Crystallography, set up 1979, 18 developers/contributors
- CCP14 – Powder and Small Molecule Single Crystal Diffraction, since 1994

→ Computing Initiative contacts these projects

The Computing Initiative

- WP-1: Data management
- WP-2: Parallel computing
- WP-3: Data processing
- WP-4: Visualization
- WP-i: Applications
  - 2012: First releases
  - 2012 – 2014: Extend applications

Open source project

The Application Workpackages

- A steering committee selects 2-3 experimental techniques, installs working groups, e.g.:
  - GISAS, SAS
  - Tomography
  - Diffraction
  - TOF spectroscopy
  - X-ray micro/nano probe
  - Coherent x-ray diffraction
- Design a framework for online/offline analysis, visualization and simulation for the specific technique.
  - Independent of the instrument
- Identify generic components
  - Image processing, simulation, fitting …
  - Utilize existing code, write new procedures, if necessary
  - Create a public code repository
- Build the application
  - Release first versions early
- Organize training workshops

Resources

Task	2010	2011	2012	2013	2014
Data management	3	3	3	3	3
Parallel computing	3	3	3	3	3
Data processing	1	1	1	1	1
SW: analysis/simulation	4	9	9	9	9
Visualization	1	1	1	1	1
Code mgt., standards	1	1	1	1	1
Coordination	0.5	0.5	0.5	0.5	0.5

Sum: 87.5 FTE years → 6.1 MEuro (Bonus Award)

Summary

- The PNI centers strongly support the Computing Initiative on common data analysis, visualization, simulation, data management to enhance the scientific outcome of the experiments.
- Combining the expertise of the various institutes creates synergy.
- This project establishes a communication platform.
- The time frame of 5 years and the proposed budget should be sufficient to install systems for data management and parallel computing and to develop first integrated applications.
- We consider this initiative as a nucleus for long-term collaborative efforts of PNI centers.

Contributors

- Felix Beckmann, GKSS
- Dietmar Herrendörfer, HZB
- Jens-Uwe Hoffmann, HZB
- Andreas Kopmann, FZK
- Thorsten Kracht, DESY
- Stefan Mattauch, FZJ
- Wolfgang Mexner, FZK
- Michael Monkenbusch, FZJ
- Mathias Münch, GSI