

Minutes of the HDRI-Meeting held at DESY on May 16, 2011

-- Draft (R. Gehrke, June 10, 2011) --

The meeting was attended by 25 participants who are engaged in the activities of the HDRI project at DESY, KIT, FZJ, HZB, HZG, and GSI. The aim of the meeting was to report on the present status of the various work packages, to review the milestones set in the project proposal and to discuss the status as well as actions to be taken on the short to medium term time scale.

1. Common Data Policy (WP1)

F. Schlünzen (DESY) reported on guidelines for a common data policy on the basis of a draft proposal published by the PaNdate Europe Strategic Working Group (http://www.pan-data.eu/PaN-Data_Scientific_data_Policy_Draft) in November 2010. In principal these guidelines could be adopted as rules for archiving and accessing data produced at the institutes involved in HDRI but there is probably demand for a discussion in the various centres about specific details, touching issues like public open access to data and also certain legal issues. This discussion should involve the management of the centres. A common advisory opinion of all centres should be established by the end of August 2011 and shall then be implemented in the archiving facilities and data access portals.

2. Data Storage, Archiving (WP1)

P. Millar (DESY) reported on the status of dCache at DESY and R. Stotzka (KIT) gave an overview on large scale data Management (LSDF) at KIT which involves storage facilities and a data browser for structured data archives. Both systems, LSDF and dCache, will be kept in parallel. At ANKA the work to design a data life cycle management for its beamlines with a transparent connection to LSDF is in progress. The definition and implementation of the corresponding data format and API will be done in close collaboration with the responsible person at DESY (E. Wintersberger). The contact has been established.

3. Data Management Portal (WP1)

I. Agapov (DESY) reported on the status of the web based data management portal at DESY and showed a demonstration of it. It can be considered as the portal prototype mentioned in the HDRI proposal. However, it is to some extent DESY specific as it utilizes dCache and DOOR. It has been discussed and agreed upon that the portal solutions of all centres within HDRI should provide similar functionality as well as similar "look and feel". It is also highly desirable, that from each portal the data at each centre can be accessed transparently.

4. Standard Data Format (WP1)

J.-U. Hoffmann (HZB) reported on the use of NeXus at the neutron facilities in Berlin where the existing implementation of NeXus into the data acquisition systems CARESS is in progress, extending to all beamlines on BER II and BESSY. Tests are already finished for three neutron diffraction instruments. The HZB has experience since 2006 with NeXus on the Flat-Cone Diffractometer E2. The Nexus/HDF5 format is used by different facilities world wide. The NeXus files were tested up to now by reading with TVnexus (HZB), MatLab, Lamp (ILL) and HDFview.

After this presentation E. Wintersberger (DESY) talked about the status of the creation of a standard data format. He proposes an implementation based on the NeXus data model and on an application definition which defines the location and content of the data elements that have been identified to be essential for a given experimental method (standard data set). In this approach the remaining data beside the standard data set could be organized without the mandatory use of objects defined in the NeXus base classes. This would provide greater

flexibility in the organization of the data outside the scope of the standard data set and could allow better adoption of frequently changing and often quite special setups encountered in synchrotron radiation experiments. The data structures have to be self-documenting. In the following discussion this approach was disapproved by J.-U. Hoffmann and the partners from HZB and DESY agreed upon an additional meeting in Berlin on June 8 in order to discuss the issues related to the PNI standard data format and NeXus in more detail (Note: In the meantime this meeting has taken place, for its results see the amendment of these minutes).

5. Dedicated Hardware for Real Time Data Processing and Real Time Data Assessment with Parallel Computing (WP2)

M. Weber (KIT) gave an overview of the activities in work package 2 which are partially in accordance or even far ahead of the milestones (review of GPU platforms, creation of fast GPU hardware, GPU based reconstruction of tomography data as the first application). First fast FPGA based hardware is available and has been used in a demonstrator. Work on the second application implementing a PX data flow is starting right now. First contacts among the interested partner have been established. Participation of others is encouraged.

Further presentations gave more detailed reports on the different achievements:

M. Balzer (KIT) reported on the development of a high speed camera demonstrator for synchrotron radiation based tomography applications. It features a 2.2 Mpixel sensor with 10 or 12 bit pixel depth and up to 340 frames/s on an exchangeable sensor carrier connected to readout electronics made up by an FPGA evaluation board (ML605) for readout, data preprocessing and DMA data transfer. First performance tests of the system are shown.

U. Spillmann (GSI) reported on the readout electronics demonstrator for a Si(Li)-Compton-Polarimeter to be used by the SPARC collaboration at GSI and FAIR. An already existing Si(Li) strip detector is utilized and the NIM and VME electronics forming the readout system have been replaced by the new development based on custom designed FPGA hardware.

A. Kopmann (KIT) presented a thorough review of presently available GPU platforms (mainly Nvidia products, which clearly provide the best performance) including benchmarks for the GPUs, OpenCL vs. CUDA, internal GPUs vs. GPU boxes via external PCIe, and storage systems (Raid). He describes the selected setup of a station for fast reconstruction of tomography data (12 CPU cores at 2.66 Ghz, 6 Nvidia GTX580 GPUs, 2 internal, 4 via external PCIe), the implementation of the data treatment, and demonstrates the performance of the system operating on a sample data set. 2000 projections of size 1776 x 1707 (24 GByte) are reconstructed into a 1691 x 1331 x 1311 3d image (11 GByte) within some ten seconds. New Nvidia chips announced for 2012 (Kepler) and 2013 (Maxwell) will be up to 40 times faster than the present state-of-the-art GPUs based on the Fermi architecture (GTX580). Thus the market has to be observed further.

6. Data Analysis, Modelling, and Simulation (WP3)

E. Kentzinger (FZJ) explained the outcome of the WP3 workshop held on January 24, 2011 to identify the requirements and approaches concerning scientific software for data analysis, modelling, and simulation. According to the proposal the work package concentrates on the development of software for simulation and modelling of off-specular scattering and for online visualization of standard SAS analysis. G. Benecke (DESY) presented the present status of the program DPDAK (Directly Programmable Data Analysis Kit for Online Analysis of 2D Scattering Data). This program is currently developed within the frame of a co-

operation between DESY and MPI Golm. The data workflow is realized via plugins written in Python which can be arranged in a sequential order according to the requirements of the analysis and are controlled by a GUI. The system is platform independent, easily extendible by adding new plugins, and open source. It can be used for offline and online analysis and the presently available plugins are tailored to the needs of GISAXS experiments at DESY beamlines BW4 and P03. Typical applications involve e.g. online visualization during fast time resolved experiments or offline fitting of data.

For modelling of GISAS data the program IsGisaxs is widely used. At present at DESY attempts are made to modify this program in different aspects: Extend the scope of possible model structures that can be simulated, make it executable in a Linux environment (the original program runs under windows), and make use of a multiprocessor system to speed up execution.

7. General Topics

The participants agreed upon having teleconferencing meetings on a regular basis in order to intensify communication among the partners of the project. The first meeting is scheduled for June 20 at 10 a.m.

A follow-up meeting to the present one should take place in November 2011. The precise date and location has still to be negotiated.

All documents produced in relation to the HDRI project (reports, publications, talks, etc.) shall be submitted to and collected by the project co-ordinator and shall be made available on the HDRI wiki.

8. Amendment: Minutes of the Meeting at HZB on NeXus Issues

The meeting took place on June 8, 2011 at HZB in Berlin-Wannsee and was attended by E. Wintersberger (DESY), R. Gehrke (DESY), J.-U. Hoffmann (HZB), and D. Herrendörfer (HZB). It was intended to discuss open questions of the meeting on May 16 related to the implementation of NeXus as the PNI standard data format.

Following the recommendations of PanData the partners agree on the implementation of NeXus according to its specification despite of the concerns of E. Wintersberger and R. Gehrke related to flexibility issues.

- The NeXus base classes shall be used to build up the data structures whenever it is possible. The partners agree that a fallback solution for items not covered by the provided base classes shall officially be implemented in NeXus by introducing a special class for this.

- A new NeXus API that fulfils the requirements of high data rates will be implemented. E. Wintersberger is responsible for this activity which shall be located at the corresponding NeXus development platform.

- The application definitions for the considered methods shall soon be consolidated. R. Gehrke will do this for the method "scattering with monochromatic radiation" by the end of August 2011 at latest, this includes reviewing of the application definitions by other HDRI partners involved in that method.

- In parallel, the implementation of NeXus in the data acquisition system of selected Petra III instruments shall start.

- J.-U. Hoffmann is ready to assist in all problems related to NeXus and its implementation.