NeXus Recapitulation and Developments

Mark Könnecke

NeXus International Advisory Committee

October 1, 2012

Table Of Content

- Brief Introduction to NeXus
- Decisions from NIAC 2010
- Developments since NIAC 2010
- Topics for this NIAC Meeting

PAUL SCHERRER INSTITUT

NeXus Mission

- Definition of a standard data format
  - Rules
  - Validation tools
- Promotion of NeXus
  - Documentation
  - NeXus API
  - Outreach to the scientific community

PAUL SCHERRER INSTITUT

NeXus Design

- Complete data for typical use
- Extendable, add additional data as you please
- Self describing
- Easy automatic plotting
- Store a full beamline description (FBD)
- Platform independent, public domain, efficient
- Suitable for a wild variety of applications

PAUL SCHERRER INSTITUT

NeXus Levels

1. Physical file format and API for accessing files
2. Rules for storing data in files
3. Component and application definitions
4. NeXus Utilities

PAUL SCHERRER INSTITUT

Decisions from NIAC 2010

- NXsubentry
- NXcollection
- Support for CIF style coordinate systems
- Non C-storage order arrays: offset, stride attributes
- Python tree API becomes part of NAPI
- Look into NAPI thread safety and PHDF

PAUL SCHERRER INSTITUT

PSI

entry:NXentry
- sample:NXsample
- instrument:NXinstrument
- .....
- sas:NXsubentry
  - sample:NXsample
  - instrument:NXinstrument
    - source:NXsource
    - velocity_selector:NXvelocity_selector
    - detector:NXdetector
      - data[xsize,ysize], signal=1 (1)
  - control:NXmonitor
    - data
  - data:NXdata
    - link to (1)

PAUL SCHERRER INSTITUT PSI

entry,NXentry
- measurement:NXcollection
  - positions:NXcollection
    - om
    - two_theta
  - scalars:NXcollection
    - title
    - wavelength
  - data:NXdata
    - detector1
    - mca5

PAUL SCHERRER INSTITUT

Developments since 2010

- HDRI Meeting Hamburg
- PanData floundered
- Code Camp 2011
- Collaboration with DECTRIS
- Code Camp 2012
- NAPI release 4.3, NXDL 3.2 == 1.0

PAUL SCHERRER INSTITUT

HDRI: High Data Rate Initiative

- German synchrotron and neutron sources
- Collaboration with little money
- Invented something NeXus alike
- Later convinced to go NeXus
- Additional and revised synchrotron base classes
  - NXcapillary
  - NXbending_magnet
  - NXinsertion_device
  - NXxray_lens
- Current state: Eugen?

PAUL SCHERRER INSTITUT

PSI

PanData

- European project to standardize logins, data catalogs and data
- NeXus at first well received
- Then plans to develop method specific formats between facilities
- Dormant ever since

PAUL SCHERRER INSTITUT

- At APS
- NX_UNLIMITED for all dimensions
- 64 bit dimensions
- HDF-5 1.8
- Documentation updates
- WWW-site from manual
- NXimpatient dodument: NeXus in 8 pages
- Test for python-API 30% complete
- Python Tree API cleaned up
- PHDF not useful
- New C++ Tree API (Eugen)

PAUL SCHERRER INSTITUT

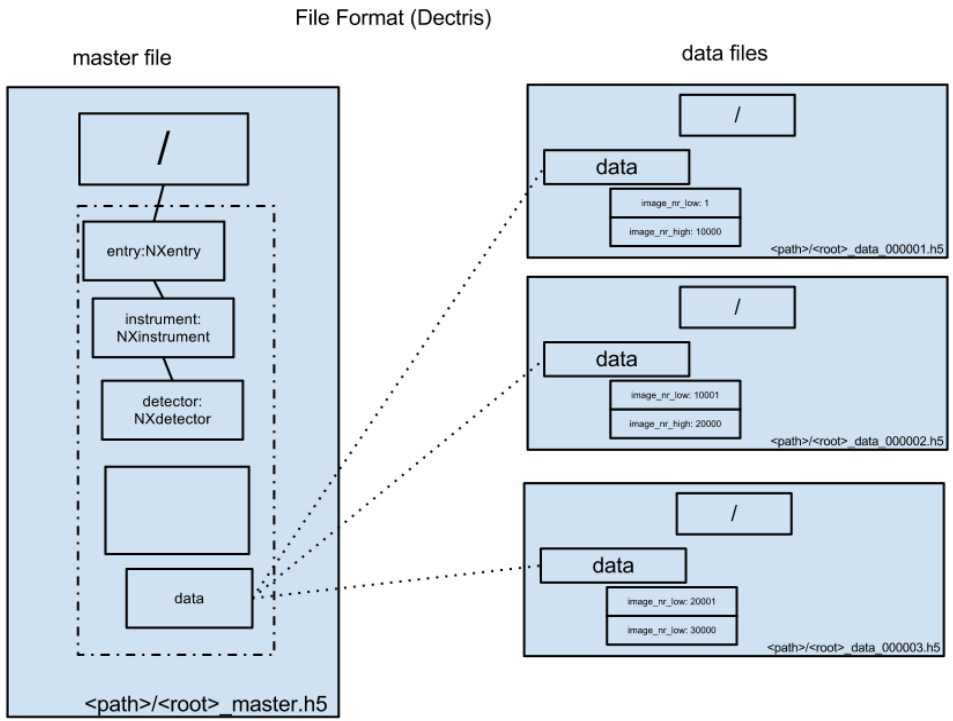
PSI

Collaboration with DECTRIS

- Manufacturer of Pilatus and Eiger detectors
- Going for NeXus/HDF-5 for Eiger
- Detector Software writes NXdetector group
- More fields added to NXdetector to accomodate pixel detectors
- This is well on its way

PAUL SCHERRER INSTITUT

DECTRIS File Structure



Rationale

- DECTRIS has a problem:
  - Detector outputs 5-10 GB/sec
  - The deliver the detector and the computer going with it
  - They cannot ask their customers to provide the appropriate hardware for such a detector: parallel file system etc.
  - Must compress and write the file on one computer
  - Compression has to be parallel as CPU intensive
- File structure a workaround for HDF-5 not allowing sections of datsets in different files
- Sidenote: LZ4 or snappy compression; up to 450MB/sec on write

PAUL SCHERRER INSTITUT

PSI

Upcoming DECTRIS Meeting with Community

- DECTRIS aims at meeting customers in october
- How far can we compromise?
- Anyone from the NeXus community who wishes to join?
- Comment from code camp: ask for tool to convert to HDF-5 standard compression

PAUL SCHERRER INSTITUT

- May 2012
- Organised with DECTRIS to address performance issues
- HDF people gave overview of new developments in HDF-5
- DECTRIS (and DESY) pays for:
  - Writing pre compressed chunks
  - Dynamically loadable filters

PAUL SCHERRER INSTITUT

New Feautures for HDF-5 1.10

- Asynchronous I/O
- Journaling
- Single Writer, Multiple Reader semantics
- Better fault tolerance
- In memory HDF-5 files
- Shared object headers

PAUL SCHERRER INSTITUT

PSI

Other Things the HDF people work on

- Better multi threading support
- Virtual Object Layer, completely replace storage layer
  - Use HDF-5 data model but not file format
  - Opens path to more storage models
  - Metadata server for better parallel support
  - Mirroring, stacking
- Better parallel processing support: meta data server

PAUL SCHERRER INSTITUT

PSI

NAPI Release 4.3, Application Definitions 3.1

- 64 bit dimensions
- HDF-1.8
- NX_UNLIMITED everywhere
- Alpha python tree API included
- First release of applicaton definitions
- Updated manual

PAUL SCHERRER INSTITUT

PSI

- Moved documentation to sphinx
  - Wiki like syntax allows for easier editing then docbook
  - URL:
- Cleaned up trac tickets
- Decided to drop autoconf in favour of CMake
- Resolved CIF coordinate issue
- Devised a good suggestion for handling axes at multidimensional datasets
- Cleanup of NeXus applications: nx2dtd, NXDump, nxtraverse, NXformat_dfn dropped
- Got NAPI 4.3 release ready

PAUL SCHERRER INSTITUT

Topics for NIAC 2012

- Review of NeXus: where are we headed?
- Roadmap OO-NeXus
- CIF coordinates
- Process for changing base classes
- Review synchrotron beamline classes
- Review additions to NXdetector
- Materials definition
- Multi dimensional array axes encoding
- What to do about expired NIAC members?
- Electing new officers

PAUL SCHERRER INSTITUT

Questions from Code Camp

- Is anyone using NXcharacterization? Remove?
- Who is using F77 NAPI? How much effort to put into this?
- Do we go into timed data?

PAUL SCHERRER INSTITUT

CIF Revisited

- Finalize
- Hope for endorsement of CIF community
- AIM: provide data necessary to derive positions of components from transformation matrics

PAUL SCHERRER INSTITUT

$$T = \begin{pmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

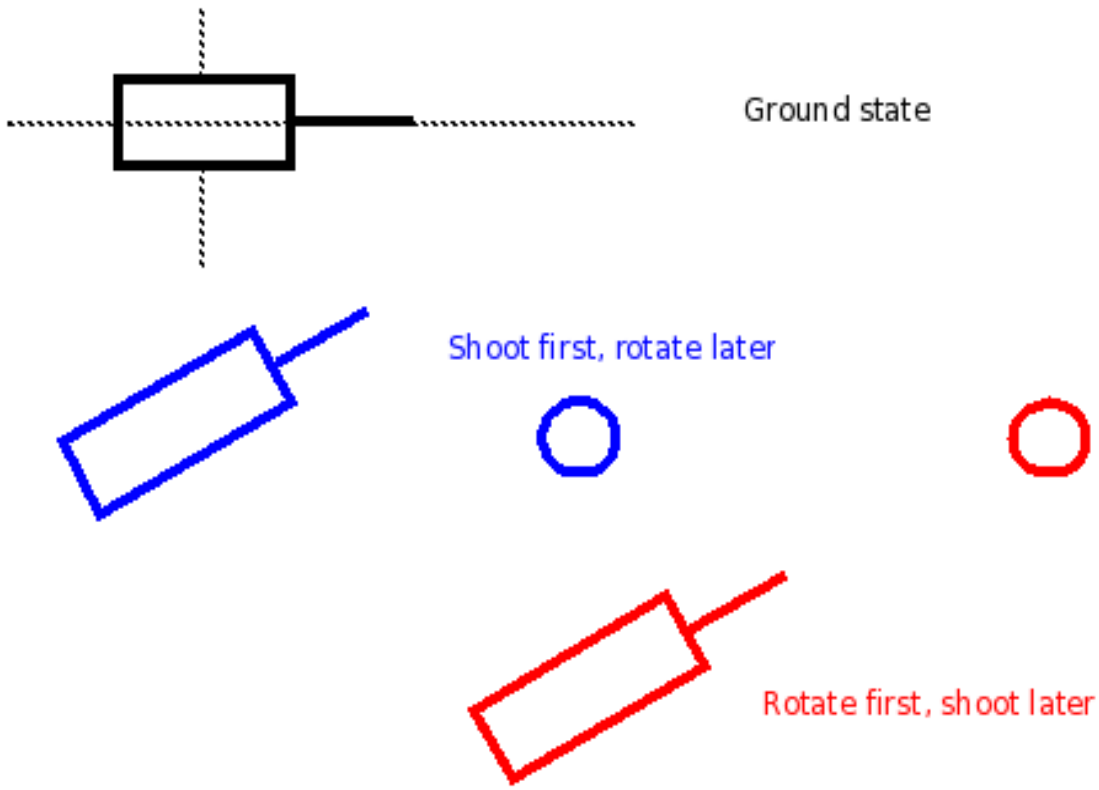
PAUL SCHERRER INSTITUT

$$T = \begin{pmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R = \begin{pmatrix} r11 & r12 & r13 & 0 \\ r21 & r22 & r23 & 0 \\ r31 & r32 & r33 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

PAUL SCHERRER INSTITUT

Combining Transformations



PAUL SCHERRER INSTITUT

Some Properties

- Transformations can be combined by matrix multiplications
- Individual matrices can be derived by looking at the situation when everything else is 0
- Absolute positions can be obtained by multiplying the resulting matrix with its transpose
- Defines new coordinate systems at components
- CIF contains a duplication: vector, offset scheme

PAUL SCHERRER INSTITUT

PSI

What Use Is This?

- Allows to calculate absolute positions of components in the laboratory coordinate systems
- Can directly convert from a detector coordinate system to vectors in Lab coordinate system
- Calculate things like impact of primary beam on detector, SAS
- Allows arbitray axis to be expressed
- Intuitively describe an instrument with angles and translations and still be able to recover absolute coordinates

PAUL SCHERRER INSTITUT

NeXus Axis Mapped

- rotation_angle, polar_angle, rotate 0 1 0
- azimuthal_angle, rotate 0 0 1
- distance, translate 0 0 1
- chi, rotate 0 0 1
- phi rotate, 0 1 0
- NeXus polar coordinate system: rotate azimuthal_angle, rotate polar_angle, translate by distance

PAUL SCHERRER INSTITUT

PSI

CIF Dependency Table

axis-id	type	equipment	dependson	vector	offset
gonio_phi	rotation	goniometer	.	1,0,0,	...
det_z	translation	detector	.	0,0,-1	0 0 0
det_y	translation	detector	det_z	0,1,0	0,0,0
det_x	translation	detector	det_y	1,0,0	0,0,0

PAUL SCHERRER INSTITUT

Expressing Axis Dependency in NeXus

- Implied: use existing NeXus coordinate system
- dependson attribute pointing to depending axis
- transform field in base classes which becomes a comma separated list of the path to the transformations required to position this component
- Create a special container to hold axis dependencies, NXdependency, to collect the dependencies in one place for easy access. This is what CIF does

PAUL SCHERRER INSTITUT

PSI

Dependons Option

sample,NXsample
    rotation_angle
    chi (dependson rotation_angle)
    phi (dependson phi)

PAUL SCHERRER INSTITUT

Transform Option

```
sample,NXsample
        rotation_angle
        chi
        phi
        transform = rotation_angle,chi,phi
```

PAUL SCHERRER INSTITUT

Separate Group Option

```
sample,NXsample
        rotation_angle
        chi
        phi
dependency,NXdependency
        sample/chi =
                sample/rotation_angle
        sample/phi =
                sample/chi
        instrument/detector/x_translation =
                instrument/detector/distance
        instrument/detector/distance =
                instrument/detector/polar_angle
```

PAUL SCHERRER INSTITUT

PSI

```
sample,NXsample
        rotation_angle (vector 0,1,0)
        chi (depends_on rotation_angle, vector 0,0,1)
        phi (depends_on chi, vector 0,1,0)
        depends_on
                phi
```

PAUL SCHERRER INSTITUT

Tech Committe Recommendation Continued

- Add offset attribute to fully cover CIF. This is an extra translation
- offset_unit to give units for offset
- The vector attribute becomes mandatory
- This gives us CIF endorsement!

PAUL SCHERRER INSTITUT

Process for Adding Fields and Classes

- Passing them through NIAC is slow (2 years!)
- Must be documented well enough, no duplicates
- Suggestion: leave to technical group
- Suggestion2: technical group publishes changes to NIAC for intervention

PAUL SCHERRER INSTITUT

Review of Synchrotron Beamline Components

- NXcapillary
- NXbending_magnet
- NXinsertion_device
- NXxraylens

PAUL SCHERRER INSTITUT

PSI

NXdetector Additions for DECTRIS

```
acquisition_mode:NX_CHAR
angular_calibration:NX_FLOAT[i,j]
angular_calibration_applied:NX_BOOLEAN
bit_depth_readout:NX_INT
countrate_correction__applied:NX_BOOLEAN
countrate_correction:NX_FLOAT[i,j]
detector_readout_time:NX_FLOAT
exposure_time_time:NX_FLOAT
flatfield:NX_FLOAT[i,j]
flatfield_applied:NX_BOOLEAN
flatfield_error:NX_FLOAT[i,j]
frame_start_number:NX_INT
frame_time:NX_FLOAT[NP]
```

PAUL SCHERRER INSTITUT

PSI

NXdetector Additions for DECTRIS 2

```
gain_setting:NX_CHAR
pixel_mask:NX_FLOAT[i,j]
pixel_mask_applied:NX_BOOLEAN
saturation_value:NX_INT
sensor_material:NX_CHAR
sensor_thickness:NX_FLOAT
threshold_energy:NX_FLOAT
trigger_dead_time:NX_FLOAT
trigger_delay_time:NX_FLOAT
```

PAUL SCHERRER INSTITUT

PSI

Materials Definition Recommendation

- This is about defining materials: samples, filters, multi layers etc.
- Bag of worms
- Recommendation
  - Use textual description
  - When decisive for DA: community should suggest enums

PAUL SCHERRER INSTITUT

Time Based Data

- Neutron event data to be correlated with other time based data
- Dynamic scans collecting detector frames and other data with possibly different sampling rates
- FELS will collect data only dynamically
- We have NXlog
- Questions to the NIAC:
  - Shall NeXus expand into this market?
  - Is the tech committee to be tasked to develop a recommendation how to store this better?

PAUL SCHERRER INSTITUT