

NIAC Meeting 2018

Brookhaven National Laboratory, October 25-26, 2018

Thursday Morning

We started by a group activity of arranging tables.

Round table introduction

Mark Basham's membership had expired. A letter of support from Diamond light source was provided. We voted on confirming Mark Basham as representative for Diamond: All YES

Finding an Agenda

Suggested Topics:

- NeXus as logbook format, #37, 3
- NXptycho, #36, 5
- Attribute mask, #35, 2
- Attribute Normalisation, #34, 2
- McStas-NAPI, #33, 2
- HDF5 XMP, #32, 5
- Messy, #37, 5
- NXquadric, #30, 3
- NXdata plotype, #29, 2
- Extended default, #28, 2
- Uncertainties, #27, 4
- NXdiffractometer, NXgoniometer, #26, 6
- NXdata, aux signal, #25, 2
- Voting, #22, 2
- 2014 Attributes, #4, 4
- Interfaces/Features/App Def, example data, #3, 8
- NXslit, #585,
- NXData, NXdet, #583, 5
- NXenvironment, #471, 1
- Rarely used NXDL #331, 0
- NXsample nature, #327, 0

- Branding, 4,

The `#numbers` point to issues on either the NIAC github or the NeXus definitions. The other numbers are votes on importance in the first pass.

We decided to put a time limit on discussions. Per default this is 20 minutes.

Interfaces, Features, Application Definitions, Example Data

Clarification of what Interfaces and Features are. The Interfaces are a better way to organize base classes, Features are more fine grained application definitions and are set to replace them.

In this first round we did not come to a conclusion. The issues at stake were unclear to most attendees. We decided to defer this discussion to after lunch. By then presentations on both topics will be presented by Mark Koennecke and Mark Basham.

NXdiffractometer, NXgoniometer

This is about having a special group in NXinstrument which describes a diffractometer. There is a lot of overlap with the content of NXtransformation.

We voted on: we add a group to NXinstrument named diffractometer which is of type NXtransformation. Voting result: All YES, no NO, no abstain.

Break

NXdiffractometer, continuation

We reopened NXdiffractometer because some attendees did not understand what they voted for.

For vote: a group of no specified name in NXinstrument of type NXtransformation which describes a diffractometer. Result: 7 for, 4 abstain

Uncertainties

Uncertainties is the name the NIAC agreed upon for discussing the storage of experimental or statistical errors associated with a dataset in a NeXus file.

This is more a discussion about a clarification of what the actual way of describing uncertainties is. There was a proposal to annotate uncertainties in 2014 but we did not arrive at a conclusion.

Pete Jemian will research this topic and bring up a proposal in the afternoon.

NXdata, NXdetector

In NXdata NeXus collects data to be plotted. In order to do perform more thorough data processing it is useful to link back from NXdata to the NXdetector group which describes the detector and data more fully. Up to now, NeXus uses the target attribute on the dataset for this function.

Mark Basham pointed out that the target attribute pointing to the original data in NXdetector from NXdata is broken when the data pointed to is an linked external file. We may not even have the right to open the external file and add the target attribute to the dataset.

Proposal from Mark B: The attribute datasetname_target can be used in the same way as the attribute target to a dataset is used but it is attached to the parent group.

Extending the target to annotate the target pointing to an external file does not work because the attribute may be attached to a dataset in an external file which we do not have control over.

There is a general problem relating data in different files with each other.

Armando suggests to have a dataset_target field which is a soft link to the group containing the dataset.

No conclusion was reached. We closed by defining an action item as follows: Armando and Mark Basham work out together how to do this in a proper way. Both taking into account NeXus ways of doing things and the technical possibilities of HDF5. They come back at a later stage with a documented proposal with examples.

NXptycho

This is coming from the CXI community who actual took NeXus, liked it but decided that reading and writing attributes is impossible. Mark Basham put the attributes back again. The resulting files are compatible with both XCI and NeXus.

A prerequisite for this application definition is an additional filed in the NXbeam base class which describes the extent of the beam. This change was proposed and voted for with:

For 7, 3 abstain, 0 against

NXptycho renamed to NXcxi and accepted. This is basically an application definition coming from an interested community.

Results of the vote: All in favor.

Lunch

Thursday Afternoon

NXptycho again

Mark Basham pointed out that XCI describes more methods than ptychography. Thus we need to be more specific with the name of the application definition.

The proposal then is to rename NXcxi to NXxcipycho. Result: All in favor

HDF5XMP

The NIAC agreed in an earlier meeting that it would be nice if file managers could show a sensible thumbnail when encountering NeXus files. A suitable approach to achieve this was implemented by Ben Watts with the help of some funds from PSI.

Ben Watts informed the NISC about what has been done. It is possible to have a user block in a HDF5 file, before the actual HDF5 content starts. The approach is to place meta data and thumbnails in there. There will be our own magic string which describes the content of our meta data in XMP format. This is a few keywords and a thumbnail. There are already plugins for most major file browsers on different OS which can work with this extra information. These plugins will be fed upstream into OS distributions. There is python code to add thumbnails and XMP data to the HDF5 file. Sidecar files with the xmp extension are also supported. The size of that user block has to be defined when opening the file.

The NIAC thanks Ben Watts and PSI for implementing this.

Uncertainties, continuation

Pete Jemian demonstrates his proposal. He proposes a uncertainties attribute to a dataset which as a value holds the name of another dataset containing the uncertainties. With the same shape as the original dataset.

Comment by Mark Basham: this does not work for linked datasets as you cannot add the attribute to the dataset in the external file. Datasetname_errors will work because in itself it can be a link.

Proposals:

Ben: uncertainties go into a dataset called datasetname_error. We remove the error and uncertainties field from NXdata.

Armando: datasetname_errors to be used as a general pattern when an error field is required.

Armandos proposal accepted with all YES

The error and uncertainties fields in NXdata to be marked deprecated, 8 YES

Attribute Mask and Normalisation

We started the discussion by clarifying that NXdata is not only for plotting. It is used as a container for a data object. Especially in processed data.

At times, there are invalid pixels in an array holding data from a detector. A means is needed to identify such pixels. This is commonly done with a mask field.

Proposal: datasetname_mask as a general way to specify a mask. We use the conventions as described for NXdetector/pixel_mask with the option to use less than 32 bit.

1 against, all others(10) YES

It occurs that there is a need to divide a dataset with another one to normalize it.

Rays proposal: datasetname_weights anywhere in a NeXus file. If it exists and has the same shape as the dataset, you are supposed to divide the data by the weights.

9 for, abstain 1,

Armando proposal: weights as an attribute to NXdata which denotes a dataset which the signal dataset has to be divided by. This dataset can have one value or as many values as items of the signal to be normalized. The signal can specify its type by the interpretation attribute.

1 in favor, 5 against, rest abstain, NOT accepted

Review of the agenda

- Base vs Interfaces, 4 – app def versus interface, 5

- example data, 2
- NXdata aux signals, 5
- 2014 attributes, close the ticket
- NXslit,
- NXenv, 2
- Rarely used NXDL, close the ticket,
- NXsample nature, not at this meeting
- NeXus as a logbook, 4
- McStas-NAPI, 4
- Voting, 1
- Messy specifications, 6
- NXquadric, 2
- Plottype, 1
- Extended default, 2
- Branding, 5

The numbers after the comma are the votes of importance.

Branding

Some NIAC members felt that the NeXus logo looks outdated. Stuart Cambell suggested to drop the NeXus graphics to the BNL design team for reevaluation. They will create a set of suggestions which the NIAC will discuss at a later stage.

Messy Specifications

Over time, strings have become a messy affair in NeXus:

- We started with arrays of bytes. Which later became NX_CHAR.
- HDF5 introduced variable length strings. Unfortunately, the popular h5py API for writing HDF5 files stores strings per default as such variable length strings. NeXus decided to support this, though the HDF-5 API for reading/writing such strings is of low quality
- After some pressure from the community, arrays of strings were introduced. Now there is a choice to store a string as a string or as an array of strings of length 1.

After some discussion we arrived at the following proposal:

Array of strings are not allowed in situations where a single string is expected. Cnxvalidate should flag that. Result of the vote: All YES

NXdata auxiliary signals

This is a recovery of the signal=1, signal=2, signal=n feature which we used to have in the old way of handling axes.

Proposal: additional group level attribute auxiliary_signals which is an array of strings holding the names of additional signals to be plotted with the signal. They all need to be of the same shape.

All in favor.

Extended Use of the Defaults Attribute

Currently we have a default attribute at root level which points to the default data. Currently this is restricted to NXroot, NXentry, NXsubentry.

Proposal: allow the default attribute for any NeXus group which contains a plotable NXdata.

All in favor

Friday Morning

Building #741, Room 156

Election of Officers

Stuart Campbell to be the technical release manager: All in favour

Pete Jemian as documentation release manager: All in favour

Ben Watts as NIAC Executive secretary: All in favour

Ben Watts steps back as NIAC executive secretary

Mark Basham as executive NIAC secretary: All in favour

Ben Watts as NIAC Chairman, All in favour

Voting

The NIAC constitution forces NIAC votes to be done by email. It is unclear however what the acceptance criteria for such email voting is. It happened often that not enough NIAC members actually voted. Moreover, there is a danger that the email

vote is badly understood and the NIAC ends up with bad decisions. On the other hand, the NIAC meets only every two years. A means to vote on issues when there is a dire need is good to have. For the discussion resulted the following proposal:

Email voting as used up to 2018 only for membership renewals

For issues other than membership renewals the following rule applies: we vote by teleconference where the proposal can be discussed. At least 2/3 of the NIAC have to be present at the teleconference. Then normal majority vote of the attendees applies. Teleconference decisions need to be published on the NIAC mailing list. NIAC Members not attending the teleconference have a one week period in which they can add their vote to the tally. Those not responding in that week are counted as abstained.

Results of the vote: 11 in favor, 1 abstain, 0 against

Review of the agenda

- NeXus as logbook, 7
- OFF-geometry, 4
- NXquadric,3
- NXdata plotype, 2
- NXslit,
- NXenv,1
- NXpdb, 3
- Base CL v Interfaces, 6
- App Def vs Features, 6
- Example data, 3
- NeXus @ Accelerators, 2
- Target: NXdetector/NXdata, 2

The numbers after the comma are the voted importance values.

NeXus as Logbook

Presentation by Mark Koennecke about how to store logbook information in a NeXus file. The added value is that a suitable tool can interact with the logbook data.

Results of some discussion: Nice idea, but not in the NeXus scope. Nothing is stopping us to use NeXus in this way, however. No new features in NeXus are required.

Interfaces

Presentation about using interfaces for structuring the base classes better by Mark Koennecke.

Result: We bury the idea. It adds more complexity to NeXus than it is worth. We look at the github branch for interfaces for ideas on structuring the documentation of the base classes better. We add some documentation on inheritance and interfaces into the design section of the manual to the tune: we thought about it and stepped back from it because of the additional complexity introduced by this.

Voting results: 1 against, 2 abstain, rest in favor

Application Definitions vs. Features

Mark Basham presents on features. Features can also test logic. Features can also be used to select software to process a file.

Clarification: Application definitions stay.

Features provide an example of reading and writing files and also do some validation.

Proposal:

We will make features available through the NeXus-WWW site as technical preview and encourage contributions from NIAC members. We accept the feature attribute as official.

9 in favor, 2 abstain, rest against

OFF geometry

This is about the mesh geometry as described in NXoff and discussed multiple times in the NIAC.

Proposal: accept NXoff into NeXus

3 abstains, rest in favor, accepted

NXquadric

A quadric is a 10 parameter functions which defines sheets in space. This would go into infinity: CSG needed to limit it. This is more precise than the mesh representation in NXoff. This is also the motivation for introducing it.

Result of the discussion: This stays in the contributed definitions for the time being.

Friday afternoon

NXpdb

Herbert Bernstein presents the proposal. There are tools to convert pdb data to NeXus and back again. This code is in CBFlib.

Proposal: we merge this when the tool is available as a standalone program. We also want example data. At least a pdb and a converted file.

All in favor

Example data

There are two types of files in the NeXus example data repository: files suitable as a reference and NeXus files as seen in the wild.

Discussion if github is the right place to store the possibly large NeXus files. No good alternative was named.

Closed with: reorganize this into reference and nexus-in-the-wild examples. We encourage facilities to add more example files.

No vote required

NXdata Plottype Attribute

This is about describing in more detail how the data is to be plotted. Color maps, linear or logarithmic axis etc. This could be achieved by having a plottype attribute at group level which contains keywords hinting at the intended presentation style.

After some discussion the NIAC came up with this proposal: NeXus is about how to find and annotate the data to be plotted but not to describe how the data is to be plotted.

Voting results: All in favor

This is to be placed into the documentation.

NeXus Target Attribute

The target attribute denoting the origin of a linked data item is broken as it only works for internal links.

We discussed this a little but could not come to an agreement. The technical issues raised by this problems need to be looked at in more detail.

Thus this Proposal: While the NIAC is working at the problem of the broken target attribute for external links, we document the problem with it.

7 in favor, 0 abstains, 0 no

NeXus at Accelerators

Feedback from the controls groups at PSI about storing data from accelerators in NeXus:

- They like NeXus
- They stumbled over the NXsample because they do not have a sample. And NXsample defines the origin of the NeXus coordinate system.

The situation is that they are using NeXus look alike things to describe an accelerator. They share some with other facilities. NeXus is not built for accelerators, thus there is a case for a new tree of groups including one which describes the origin. The suggestion is to continue working on this and when they are ready to standardize, to come back to the NIAC.

NXenvironment

Proposal: In NXsample mark magnetic_field_log and temperature_log as deprecated. To be replaced with NXlog classes with the log dropped of the name.

1 abstain, rest in favor

NXslit

After some discussion we concluded that the NXslit definition is sufficient. The slit can always be transformed to wherever it should go.

End of NIAC

Attendees

- Mark Basham, Diamond Light Source
- Stuart Campbell, NSLS-2
- Ben Watts, PSI, SLS
- Jiro Suzuki, JPARC
- Michele Brambilla, PSI, (Guest)
- Mark Koennecke, PSI, SING
- Takahiro Matsumoto, Spring-9, Japan, (Guest)
- Pete Jemian, APS
- Armando Sole, ESRF
- Herbert Weinstein, CIF,
- Andy Moesch, DECTRIS, (Guest)
- Matt Clarke, DMSC, (Guest)
- Ray Osborn, Member at large, (Thursday)
- Freddie Akeroyd, ISIS, (remotely)
- Jens Hoffman, HZB, (remotely)
- Marie XingXing Yao, APS, (Friday morning, guest)