

# **NPOESS Augmentation Tool Project Goals, Deliverables and Future Work Plan**

## **The HDF Groups**

### **Abstract**

This document summarizes goals, deliverables and plans for the “NPOESS Augmentation Tool” project.

### **Project goals, timeline and deliverables**

*Goal:* To provide access to the NPP/NPOESS/JPSS data and metadata via netCDF classical library and tools.

*Timeline:* Provide an initial solution acceptable to the NOAA scientists by June 1, 2011. Provide a final solution by November 1, 2011.

*Deliverable:* A documented and tested solution that can be a tool or a process (based on the feedback that will be solicited by the HDF developers).

### **Section 1: Prerequisites or accessing NPOESS data**

The HDF Group developers studied the problem and identified several possible solutions for accessing NPOESS data and NPOESS metadata stored in the XML control files.

All solutions require the following changes to the netCDF-4 library:

1. Allow opening an HDF5 file that contains datasets with the non-supported datatypes such as object and region references (present in the NPOESS HDF5 files)  
and/or
2. Add region and object references to the standard set of the netCDF datatypes  
*Status:* netCDF developers promised to add feature 1 and then 2; timeline will be known before November 1, 2010
3. Allow netCDF APIs to access a variable using file and variable identifiers. This feature is critical for accessing NPOESS HDF5 files by applications that use “classic” APIs only and cannot be modified to use new netCDF-4 APIs.  
*Status:* netCDF developers promised to add the feature; timeline will be known before November 1, 2010

*Note: the HDF developers developed a good understanding of the netCDF-4 library internals and can work with the netCDF developers to expedite the availability of the required features.*

If both features are added, any netCDF classic application using the netCDF-4 library will be able to access data stored in the NPOESS files.

To access NPOESS metadata additional solutions are required. They are discussed in the next section.

## **Section 2: Accessing NPOESS metadata available in the XML control files**

The HDF Group developers identified three possible solutions. Each of the solutions has pluses and minuses, and complements each other. This discussion is out of scope of this document and will be available in a separate document (see Section 3).

The first two solutions are based on already existing tools and techniques, or on tools under development. The third one is a new tool that will be specifically designed to achieve the goal.

### **1. Leveraging NcML – no file modification required**

Unidata provides netCDF Markup Language (NcML). A NetCDF-4 file or set of files can be modified with the metadata and data stored in a file written in NcML. After looking at NcML the HDF developers concluded that NcML is a viable solution to achieve the project goal especially when file modification is not desired.

This solution cannot be used alone. A tool should be developed to extract NPOESS metadata from the XML control files and possibly HDF5 files, and to write to the NcML file to facilitate access to the NPOESS files with the netCDF library and tools.

Access to metadata in NcML via C netCDF library is critical for this solution, but it is under development and its timeline is unknown.

For more detailed discussion see

<http://www.hdfgroup.uiuc.edu/RFC/HDF5/netCDF4-HDF5/NcML-study-note.pdf>

### **2. Using h5edit – file modifications are required**

The HDF Group developers have been working on the h5edit tool that will allow one to add/delete data and user-defined metadata (e.g., attributes) to/from the HDF5 files. The current design doesn't include dimensions and dimension scales desired by NPOESS.

The h5edit tool can be used to add some of the required attributes to the NPOESS file. It has to be used along with two tools:

- a. A tool to extract NPOESS metadata from the XML control files and to write a control file used by h5edit.
- b. A tool to extract NPOESS metadata from the XML control files and to add dimensions and dimension scales to the NPOESS file.

The HDF Group has a better control of this solution since it doesn't depend on any Unidata development activities except the ones described in the Section 1. H5edit may be extended to create high-level objects such as dimension scales, images, tables, etc.; then tool described in "b" will not be necessary.

For information on h5edit see <http://www.hdfgroup.org/pubs/rfc> , [h5edit - An HDF5 File Editing Tool](#) .

### **3. Creating augmentation tool – file modifications are required**

The HDF Group developers will design and create a specialized tool that will extract metadata from the XML control files and modify an NPOESS file according to the specification (TBD). This tool will be unique for the project.

### **Section 3: Development plans (for November-December 2010).**

1. Work with Unidata to resolve the issues described in the Section 1. The HDF Group developers will meet with the netCDF developers on October 27-29 and will discuss timelines and collaboration efforts. NcML support with the C netCDF library will be discussed at the same time.
2. Continue with the h5edit development effort and come up with the design of the tools described in sections 2.a and b. Prototype of the h5edit tool will be available in December-January 2011.
3. Create a formal RFC document that describes all three approaches in detail and circulate among NOAA scientists for comments.