

**Progress Report – Year 3
April 2010–November 2010**

“Provide HDF Support for the ESDIS Project and the EOSDIS Standard Data Format”

Mike Folk and Kent Yang

December 2010

Principal Investigator: Mike Folk

Project Manager: Kent Yang

Institution: The HDF Group, 1800 So. Oak St., Suite 203, Champaign, IL 61820

Period covered: April 2010–November 2010

NASA Award Number: NNX08AO77A

Highlights

User support

- We received 438 requests in total; 92% of all ESDIS-related requests were addressed within two days.

DAAC and NASA HDF user community

- We maintained the HDF-EOS website and continued adding new tools and other tool examples on accessing NASA HDF4 and HDF-EOS2 data products.
- We continued adding new features for the tool that makes HDF-EOS5 files accessible by netCDF-4.

EOS support

- We provided comprehensive IDL/MATLAB/NCL examples for many NASA HDF data products.
- We helped the HDF-EOS development team add a new feature to the HDF-EOS5 library.

HDF4 and HDF5 libraries

- We released HDF5 1.8.5 and HDF5 1.8.5-patch1.
- Substantial work was done on integrating HDF products with CMake and on enhancing regression testing for the HDF4 and HDF5 libraries.

Applications and vendors

- We greatly enhanced the HDF4 OPeNDAP handler to serve many NASA HDF4 and HDF-EOS2 products.
- We communicated frequently with NASA GES DISC and JPL developers to help them access HDF-EOS data via HDF4 and HDF5 OPeNDAP handlers, netCDF Java tools and other tools.

Others

- We co-hosted the HDF and HDF-EOS Workshop XIV.
- We provided consultation to decadal mission teams.
- We enhanced the HDF4-OPeNDAP handler to access many NASA HDF4 and HDF-EOS2 data products.
- We released an alpha version of the software for the project to support independent mapping of HDF4 data objects to improve the long-term preservation of HDF4 EOS data.
- We provided HDF5 risk reduction support for NPOESS VIIRS, OMPS, and other sensor and environmental data products

1 Introduction

This document has three parts:

- (1) Project goals
- (2) Progress report for April 2010 through November 2010
- (3) Related activities supported by other funding sources

2 Project goals

The primary goals of this cooperative agreement for the period between July 15, 2008 and July 14, 2011 are to provide the following:

- (1) Support activities for data providers, for data users and for the EOSDIS Core System (ECS)
- (2) Maintenance and quality assurance for HDF4 and HDF5 libraries and utilities
- (3) Evolution of the HDF5 library and utilities to meet new EOS requirements
- (4) Integration with complementary technologies and application domains
- (5) Support for the long-term preservation of HDF-EOS and the transition to NPOESS

3 Progress report for April 2010 through November 2010

This progress report is organized to correspond with the project goals listed in Section 2.

3.1 Support activities for data providers, for data users and for the EOSDIS Core System (ECS)

Helpdesk support

Out of 438 requests received between April 16, 2010 and November 15, 2010, 52 requests were from ESDIS-related users. More than 90 percent of ESDIS-related requests were addressed within two days.

Support for NASA data centers and NASA HDF user communities

We continued adding new information to the HDF-EOS Tools and Information Center website (<http://hdfeos.org>) to help HDF and HDF-EOS users handle NASA HDF data. The main new contribution is a web page (<http://hdfeos.org/zoo>) that includes comprehensive example codes and plots for the popular Earth Science tools IDL, MATLAB and NCL. We have collected sample NASA HDF4, HDF-EOS2 and HDF-EOS5 files from all NASA data centers that distribute data in HDF.

Members of The HDF Group interacted with NASA HDF users as well as giving presentations and posters at earth science conferences hosted by NASA. At the 2010 Summer ESIP Federation meeting, we hosted a session and gave four talks on ways in which The HDF Group Supports easy access to NASA HDF data, and presented two posters. At the 9th ESDSWG conference we gave invited presentations on the HDF5 library and tools, and on interoperability between HDF5 and netCDF-4. At both meetings we provided help to NASA HDF users and collected valuable suggestions for improving HDF support. For example, during the 9th ESDSWG Conference, we consulted on HDF5 to decadal survey SMAP and ICESat-2 mission teams. We also gathered information from a MEaSURES science PI about a new HDF-EOS5 MEaSURES data product and provided IDL/NCL/MATLAB example codes and plots at the HDF-EOS website shortly after the conference. We also worked closely with the ESDIS Outreach manager to introduce the new features on the HDF-EOS website. For example, an article "The New HDF-EOS Tools and Information Center" was published at The Earth Observer.

We continued adding new features for the tool to augment HDF-EOS5 files so that HDF-EOS5 data can be read by netCDF-4 applications. The tool can now support all NASA Aura and MEaSUREs HDF-EOS5 data products.

EOS support

The HDF Group added many IDL/NCL/MATLAB example codes and plots at the HDF-EOS website to help users access NASA HDF-EOS2 and HDF-EOS5 data easily.

We implemented a utility that can dump the latitude and longitude values for HDF-EOS2 data. This information is necessary for tools such as IDL and MATLAB to display the HDF-EOS2 data correctly. We are also in the process of implementing an HDF-CF conversion library aiming to make NASA HDF-EOS2 and some HDF4 products to be CF-compliant. This library will be very useful for users who want to access HDF4 and HDF-EOS2 data via some popular open source tools that follow CF conventions.

The HDF Group continued providing automatic configuration support to the HDF-EOS2 and HDF-EOS5 libraries. Developers at The HDF Group continued performing automatic daily tests of HDF-EOS2 and HDF-EOS5 with HDF4 and HDF5 development snapshots.

The highest priority was assigned to HDF libraries and tools bugs that were reported by the HDF-EOS development team; developers fixed the reported bugs in a timely manner and provided patches if necessary.

We continued working very closely with vendors, such as The MathWorks and ITTVis, whose support of HDF is important to the EOS community.

We co-hosted HDF and HDF-EOS Workshop XIV on Sep 28–30, 2010. We gave 15 tutorials and presentations.

3.2 Maintenance and quality assurance for HDF4 and HDF5 libraries and utilities

HDF4 and HDF5 libraries

During this period, The HDF Group released HDF5 1.8.5, HDF5 1.8.5-patch1 and Java HDF Products 2.7-beta.

New releases of HDF5 included several important enhancements:

- The source code was modified to handle all optimization levels for the latest GNU C compilers.
- All known memory leaks were fixed.
- A defect that could cause data corruption was removed. To our knowledge the problem was addressed before any corrupted data was generated.
- A CMake build option for Windows and Linux was officially released.
- HDF5 libraries and command line tools were ported to use the latest GNU, PGI and Intel compilers.

During this period The HDF Group's developers continued their efforts to improve the quality of the HDF5 libraries and command-line utilities by using the "Coverity" tool to analyze source code and to eliminate software problems. The HDF Group also expanded and improved regression testing of HDF4 and HDF5, adding a test configuration to catch memory leaks in HDF5, and expanding the pool of the systems used in daily testing by adding PPC64 (big-endian) and Mac OS X Intel 64 machines. Every day more than 250 different configurations are tested on Linux, UNIX, Mac OS X, and Windows platforms. The configurations include testing HDF-EOS and netCDF-4 with HDF software under development and testing HDF5 backward/forward compatibility with Aura and NPP files downloaded from NASA and NOAA FTP servers to ensure the robustness and quality of HDF5 and its interoperability with software critical to ESDIS users.

Currently The HDF Group developers are working on the HDF5 1.8.6 and HDF 4.2.6 releases. The HDF5 1.8.6 release will include, among other enhancements, performance improvements for parallel I/O and support for new Sun OS 5.10 C, C++ and Fortran compilers. HDF 4.2.6 will include support for CMake and new functions that were created as part of the HDF4 mapping project reported in Section 4.2.

3.3 Evolution of the HDF libraries and utilities to meet new EOS requirements

HDF4 and HDF5 libraries

We implemented a CMake build option for all HDF products. Usage of CMake has improved testing and maintainability on Windows, the most popular ESDIS platform. The usage of the CDash interface to monitor daily testing drastically reduced the time needed to review test results. Currently we are looking into CMake for building and testing stackable software; for example, how to conveniently build and test HDF4 software that depends on third party compression libraries. This work has also helped us develop expertise that we will share with HDF-EOS developers to improve maintainability of HDF-EOS products.

Interoperability of different software packages used in the Earth Science community is critical when working with the data stored in different formats. The HDF Group and Unidata have been working together to enhance HDF and netCDF libraries and documentation to achieve better interoperability as requested by our users.

Tools

The HDF Group fixed bugs and improved testing for tools critical to the ESDIS users, such as h5dump, h5diff, and h5repack, and HDF Java browser HDFView.

3.4 Integration with complementary technologies and application domains

We continued communicating frequently with NASA GES DISC developers to help them use the HDF4-OPeNDAP and HDF5-OPeNDAP handlers to serve AIRS, OMI, MLS and HIRDLS data.

We continued working closely with Unidata netCDF-4 developers. We participated in a netCDF workshop, and had detailed discussions on improving interoperability between netCDF and HDF. We tested the daily netCDF-4 snapshot against the HDF5 daily snapshot to ensure early identification of potential bugs.

3.5 Support for the long-term preservation of HDF-EOS and the transition to NPOESS

The HDF Group's efforts in the area of long-term preservation mostly involved a project to create maps of HDF4 data. This work is described in section 4.2.

The HDF Group's efforts in support of the transition to NPOESS are described in section 4.3.

4 Related activities supported by other funding sources

Much of the work by The HDF Group during this reporting period was supported through other funding sources, as indicated in the following subsections.

4.1 HDF and OPeNDAP

Because of the success of HDF5-OPeNDAP work, NASA provided additional funding to enhance the HDF4-OPeNDAP handler so that OPeNDAP can access NASA HDF-EOS2 and HDF4 data. We greatly enhanced the HDF-EOS2 module of the HDF4-OPeNDAP handler to serve many NASA HDF-EOS2 products including MODIS, AIRS, AMSR-E and MISR. We also provided a special HDF4 module to handle some NASA HDF4 products including SeaWIFS, TRMM and CERES. The enhanced HDF4-OPeNDAP handler was released at the OPeNDAP website (http://www.opendap.org/download/hdf4_server.html).

4.2 Improving long-term preservation of HDF4 EOS data by independently mapping HDF4 data objects

The NASA EOS has more than a petabyte of critical earth science data stored in the HDF4 format, comprising a core component of the long-term climate record. It will be important to have access to these data long into the

future, but there is a risk in depending solely on the HDF libraries to access HDF-formatted data over the long term. To address this risk, the HDF4 Independent Mapping Project was created to find a way to retrieve the data *independently* from HDF4 software.

In Phase 1 of the HDF4 Independent Mapping Project, we teamed with the National Snow and Ice Data Center (NSIDC) and Goddard Earth Sciences Data and Information Services Center (GES DISC) to explore this concept. We produced XML-based maps of the layout of selected HDF4 files, a tool to create maps, and tools to read the maps without using the HDF4 library.

Phase 2 of the project, which will complete in July 2011, will create production quality mapping tools and XML schema, and begin deploying these at NASA data centers. The XML schema is nearly complete. A prototype of a tool to create XML maps was completed in November 2010, and a prototype of a tool to read data from HDF4 files using the maps was completed in summer 2010. Completion of these tools is expected by March 2011. Discussions have begun with NSIDC, GES DISC, and Raytheon to prepare for deployment at data centers. It is anticipated that deployment testing will begin in January 2011, and actual deployment at NSIDC and GES DISC in late spring. It is uncertain whether any other deployment will begin during this phase of the project.

4.3 HDF5 Support for NPOESS

This project provides HDF5 risk reduction support for the distribution of NPOESS VIIRS, OMPS, and other sensor and environmental data products. It has allowed The HDF Group to greatly improve its level of support for NPOESS, but also has freed up resources for important new ESDIS project activities.

The initial contract, which ended August 31, 2010, included implementing and maintaining NPOESS-specific software developed by The HDF Group, implementing testing aimed at NPP/NPOESS needs, providing NPOESS with rapid response user support, performing special projects as defined by the NPP/NPOESS communities, and performing research projects as requested.

A second contract spans September through January 14, 2011, giving special emphasis to finding ways to make NPOESS files accessible by the netCDF-4 library. Accomplishing this in a timely manner will help NPOESS science teams' calibrate instruments in anticipation of the NPP launch. It should also expand the accessibility of NPOESS data, since many applications of this data require data to conform to the netCDF data model.

4.4 Research on data management with HDF5 in support of military operations

This investigation for the U.S. Army Engineering Research and Development Center (ERDC) demonstrated the use of HDF5 as a data management platform for mission operations and showed how to adapt HDFView to present mission operational data. The project showed how to integrate a wide range of spatio-temporal data in HDF5. A paper describing the concept, and a demo can be found at the project website (<http://www.hdfgroup.org/projects/erdc/>).