

The Prototypical NASA HDF-EOS Web GIS Software Suite (NWGISS)

Liping Di, Wenli Yang, Meixia Deng, and Donna Y. Deng
Center for Earth Observing and Space Research, School of Computational Sciences, George Mason
University, Fairfax, Virginia 22030, USA
lpd@rattler.gsfc.nasa.gov

Ken McDonald
NASA Goddard Space Flight Center, Code 423, Greenbelt, MD 20771, USA
mcdonald@rattler.gsfc.nasa.gov

Abstract: NASA EOS program is producing huge amounts of data in HDF-EOS format. Those data are valuable not only for global change research but also for resource and environmental management. GIS is one of the important tools for analyzing NASA's EOS data. Making HDF-EOS accessible to GIS through Internet will greatly enhance the interoperability and public use of EOS data. This paper describes a suite of web GIS software that makes HDF-EOS data available to GIS users based on Open GIS Consortium's (OGC) interoperability protocols. The software is called NASA HDF-EOS Web GIS Software Suite (NWGISS). NWGISS consists of the following components: a map server, a coverage server, a catalog server, a coverage client, and a toolbox. Those components can work both independently or collaboratively. NWGISS can work with all three types of HDF-EOS data, namely swath, grid, and point. Currently, the map server, compliant with OGC web mapping specification 1.0, has been developed and tested. The coverage server, which serves real data to GIS clients, has been developed and tested as a part of OGC WMT II activities. The catalog server provides catalog search capabilities to GIS clients. The state-full catalog server has been implemented and the http-based stateless catalog server is being implemented. The coverage client, which enables GIS users to access multi-source data served by OGC compliant web coverage servers, is being developed now. The toolbox consists of two-way translators between HDF-EOS and major GIS formats, as well as the CreateCapabilities tool that automatically creates the XML capabilities descriptions from the metadata in HDF-EOS files. Both tools are available now. NWGISS map and coverage servers have been used by NASA and other space agencies. NWGISS is free to data providers who want to serve HDF-EOS data to GIS clients.

I. INTRODUCTION

NASA EOS program is generating huge volume of data in HDF-EOS format. Those data are valuable not only for global change research but also for resource and environmental management. GIS is one of the important tools for analyzing NASA's EOS data. However,

current GIS systems use different internal formats and are very difficult to interoperate with each other and with EOSDIS directly. Most of GIS software packages are also incapable of ingesting data in HDF-EOS. Therefore, the development of capability for delivering GIS-ready EOS data directly to user's GIS systems through Internet for analysis will greatly enhance the interoperability and public use of EOS data.

II. THE OGC WEB MAPPING TESTBED AND INTERFACE STANDARDS

The Open GIS Consortium, Inc. (OGC) is a not-for-profit membership organization founded in 1994 to address the lack of interoperability among systems that process georeferenced data, and between these systems and mainstream computing systems. OGC's mission is to give the world's information systems a new connection to physical reality by making georeferenced data behave like just another standard data type in systems of all kinds. To achieve that mission, OGC has successfully engaged key user organizations and technology providers in a consensus process to develop technology standards and business process innovations that support widespread adoption and use of georeferenced data and services.

In the past two years, OGC successfully implemented two Web Mapping Testbeds, WMT I and WMT II, and produced a set of web-based interoperability specifications. The WMT I was finished in 1999 and produced an OGC Web Mapping Specification (WMS) version 1. WMS allows interactively assembling maps from multiple servers. The WMT II was finished in December 2000 and produced a set of new interoperability specifications. One of the most important specifications for NASA from WMT II is the Web Coverage Specification (WCS). It allows a WCS client to access real multi-dimensional, multi-temporal data from coverage servers. WCS provides an interoperable way of

accessing geospatial data, especially those from remote sensing. In addition to WMS and WCS, OGC also developed other interoperability specifications. OGC specifications mentioned in this paper can be found in OGC website [1].

This paper describes a prototypical software suite that demonstrates the interoperability and access of NASA EOSDIS data by using OGC interoperability technologies. The software developed in the project is called NASA HDF-EOS Web GIS Software Suite (NWGISS).

III. OBJECTIVE

The objective of this project is to develop a prototypical NASA HDF-EOS Web GIS Software Suite (NWGISS) that makes HDF-EOS data available to GIS applications according to OGC Web GIS Specifications. Through the development of the prototype, we will evaluate the suitability of OGC interface standards for making NASA data accessible to GIS community. The NWGISS should be able to work with all three types of HDF-EOS data files, and is easy-to-use and easy-to-setup.

The development of NWGISS will make NASA EOS data immediately available and accessible to GISs developed by major GIS vendors. Hence, this project will significantly increase the accessibility, interoperability, and inter-use of HDF-EOS data, improve the data analysis and visualization, promote the use of HDF-EOS data not only in the global change research but also in the public concerned issues such as environment and resource management, education, and community planning.

IV. ARCHITECTURE AND COMPONENTS

Currently, NWGISS consists of following components: a map server, a coverage server, a catalog server, a coverage client, and a toolbox. The detailed description of each component is given in the following subsections. All NWGISS components can work both independently and collaboratively. With the development of new OGC interface specifications, additional components will be added into the suite.

The OGC interface specification to be implemented in NWGISS includes the Web Mapping Specification (WMS), Web Coverage Specification (WCS), and Catalog Interoperability Specification (CIS). Figure 1 shows the overall architecture of NWGISS.

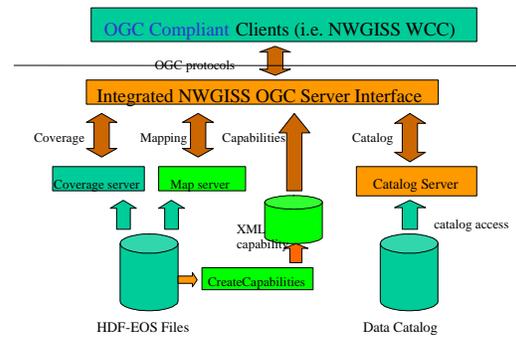


Figure 1. The high-level architecture of NWGISS

A. The Map Server

The map server enables GIS clients to access HDF-EOS data as maps. The OGC WMS defines three interfaces, namely GetCapabilities, GetMap, and GetFeatureInfo. The GetCapabilities interface allows an OGC map client to obtain information about the server capabilities as well as the availability of map layers. Once the client gets those information, the client can formulate a GetMap request based on user requirements as well as the information obtained in the GetCapabilities request. The map server makes a map based on the GetMap request, and sends the image map to client. The GetFeatureInfo interface allows the client to query geographic features associated with the map.

In order for the NWGISS map server to comply with OGC WMS, all three-interface protocols have to be implemented. Because we want our server to be generic, the server has to be able to work with all three types of HDF-EOS files.

The OGC WMS requires the map object returned from a server to be a georeferenced object with exact geographic coverage and map size specified by the client. This requirement implies that the server has the capability of geographic-coordinate based accurate subsetting and subsampling on georeferenced data. In the three HDF-EOS data models, both grid and point data are already georeferenced. However, the swath data are not georeferenced although georeferencing information is available in HDF-EOS swath files. Therefore, in order for NWGISS also to serve the swath data, NWGISS has to provide on-the-fly georeferencing and accurate subsetting services to the swath data. The HDF-EOS library provides subsetting on swath data, based on geographic coordinates. But

the data obtained from the subsetting are not georeferenced and the data of a whole scanline that touches the subsetting box defined by geographic coordinates are returned. Therefore, the HDF-EOS swath subsetting service is not useful to NWGISS. We need to implement on-the-fly georectification and resampling based on the geolocation information associated with the data.

Two georectification methods are implemented in the NWGISS map server. The first one is polynomial fitting and the second one is piece-wise linear interpolation. The later one provides more accurate georectification than the former one but requires more geolocation information (e.g., Ground Control Points). Therefore, the selection of a method to be used is based on the amount of geolocation information available in the file. If several pixels in every scanline have the geolocation information (e.g., the lat/lon locations), the piece-wise linear interpolation method is used. Otherwise, the polynomial fitting method is used. The detailed description of the piece-wise linear interpolation method can be found in [2]. The polynomial fitting is one of standard methods for georeferencing the swath data [3]. In current implementation, a third order polynomial is applied to whole scene of swath data for georectification. For both methods, the nearest-neighboring resampling is used. We also plan to implement the bi-linear and cubic convolution methods for resampling.

B. Coverage Server

OGC Web Coverage Specification (WCS) is designed for enabling GIS clients to access real multi-dimensional, multi-temporal data. WCS defines two interface protocols, namely GetCapabilities and GetCoverage. Similar to one in the WMS, the GetCapabilities in WCS allows a client to obtain the capabilities as well as the available data in the server. Same as in WMS, the capability description is encoded in XML. The OGC specification provides a DTD for the description. The GetCoverage allows a client to request multidimensional data by specifying extents in the spatial, temporal, elevation, band, and other dimensions, and the encoding format for the returning coverage. The server has to extract the data from archives based on client's specifications, package the data in the format specified by the client, and return the data to client.

The current version of WCS defines three coverage-encoding formats, namely HDF-EOS, NITF, and GeoTIFF. The later two can only handle georectified data (i.e., NASA level-3 data). For NASA level 1 and 2 data, they are not very useful because they are not capable of handle geolocation information in swath data. For handling NASA EOSDIS swath data, HDF-EOS is the only choice of the returning format.

The NWGISS coverage server has implemented all three WCS coverage-encoding formats. In order for our coverage server to work with all three HDF-EOS data models, on-the-fly data subsetting, assembly, and format conversion for all three data models are implemented. Because WCS requires the geographic coordinate based subsetting, for the swath data the on-the-fly georectification and re-sampling as well as geographic bounding-box subsetting are provided. The bounding-box subsetting will return only the data within the bounding box but retaining the original scanline structure by filling 0s in the area outside of the bounding box to maintain the same numbers of pixels for each subset scanline. This subsetting is different from the one implemented in HDF-EOS library that returns whole scanlines that touch the bounding box. Therefore, for the swath data, the client can requested data either in georectified form or in the swath form. The georectification methods implemented in the WCS are the same as those in WMS.

C. Catalog Server

Both WCS and WMS have the GetCapabilities protocol for clients to find geographic data/map and services available at servers. This protocol works nicely when a server has a small data archive. If the server has a lot of data, the capabilities description, which basically is a data catalog, is become very large. A better way to handle large data archives is to enable a client to search the data catalog in a server to find the needed data. The catalog server in NWGISS is designed for this purpose. The catalog server allows GIS clients to search and find available geographic data and services in a NWGISS site based on OGC catalog interoperability specification (CIS).

The CIS has two operational models, the state-full and state-less. We reuse the Data and Information Access Link (DIAL) [4] catalog server for NWGISS. DIAL catalog server has implemented Catalog Interoperability Protocol

(CIP) of Committee on Earth Observation Satellites (CEOS), and also has its native http-based catalog protocol. CIP is a profile of OGC state-full catalog protocol. For implementing OGC state-less catalog interoperability protocol, we plan to put an OGC wrapper on top of the DIAL http-based native catalog server.

D. Web Coverage Client

The combination of WMS and WCS allows GIS users to browse data on servers and obtain the data from servers in the form that exactly matches their needs. Currently various agencies have developed several WMS clients. However, there is no good WCS client available because WCS is very new and there are not many coverage servers existing yet. In fact, the NWGISS coverage server was the first one in the world that implemented the OGC WCS.

The NWGISS coverage client is a comprehensive OGC WCS client. It is able to interactively communicate with all OGC-compliant coverage servers (not only with NWGISS) for accessing multi-dimensional geospatial data and handling all three coverage-encoding formats. Besides performing basic WCS client-server communication, coverage access, visualization, and user interaction, the client will also provide georectification, reprojection, and reformatting functions. Those functions can also be applied to local HDF-EOS files. The execution of those functions will be automatically arranged based on the user's data requirement and the information about the data in servers. With the client, users can obtain data from multiple coverage servers in the format, projection, and spatial and temporal coverage that exactly match user's specifications. Those data are ready to be integrated and analyzed in most of GIS and image processing packages.

We also plan to make the client multiple protocols compliant (e.g., map, catalog, and feature) in the future.

E. Toolbox

In order for data providers to easily set-up and manage NWGISS servers, a toolbox is provided. The toolbox contains tools for automated data ingestion and catalog creation. Currently, two types of tools are provided: the format conversion tools and XML capabilities creation tools. A third type of tools, the catalog creation tools, will be provided later.

Because NWGISS servers only work with data in HDF-EOS format and data providers may have data in different formats, it is nice to have data conversion tools so that NWGISS can also serve data in formats other than HDF-EOS. Currently we have tools for converting GIS formats to HDF-EOS available [5].

Another tool provided in the toolbox is the CreateCapabilities program. It automatically creates the XML-based capabilities description based on the NWGISS capabilities as well as data in HDF-EOS files. CreateCapabilities is a generic tool that can work with all HDF-EOS files.

V. CURRENT STATUS

Although we have not finished the development of NWGISS, some components of NWGISS have been fully developed and are ready to use. The NWGISS map server is fully operational currently and has been installed at multiple sites. The map server was a part of WMT I Testbed and is compliant with OGC WMS version 1.0. Both Sun Solaris and Windows versions are available now.

The NWGISS coverage server was developed as part of WMT II activities. The OGC WCS was based on this implementation. Currently the coverage server is fully operational and has been installed at several sites. It can encode the coverage in all three standard encoding formats. It can work with HDF-EOS swath and grid data. The Sun Solaris version of coverage server is available now and Windows version will be available soon.

The NWGISS coverage client is being developed currently. It is expected that a fully functional client will be available by June 2002.

The toolbox has both the data conversion and CreateCapabilities tools in place. The data conversion tools include those for converting GeoTIFF, Arc/INFO grid, ArcView Shape, ERDAS LAN, and binary to HDF-EOS. The CreateCapabilities tool currently has a prototype working with HDF-EOS swath and grid data based on OGC WMS version 1.0. An upgrade of the tool to handle multi-dimensional data for latest WMS and WCS is needed.

The development of NWGISS catalog server for state-full model is finished. The server has been tested internally and externally. The state-less server is being implemented currently.

VI. CURRENT USE OF NWGISS

NWGISS map and coverage servers are the only OGC-compliant ones in the world that serves HDF-EOS data directly to GIS clients. Currently, NWGISS is used in CEOS WTF/GOFC [6], Digital Earth Testbed Alpha version [7], OGC WMT I & WMT II [1], NASA ESIP Digital Earth Cluster [8] as well as NASDA (Japan), ESA/ESRIN (EU), and GISTDA (Thailand).

We have set up a demo site for NWGISS map and coverage servers at NASA Goddard Space Flight Center. The web address for the servers are: <http://heineken.gsfc.nasa.gov:8910/cgi-bin/dib/WGISS> (map server) and <http://heineken.gsfc.nasa.gov:8080/cgi-bin/dib/wcshdfeos> (coverage server). The demo site contains MODIS level 1b, TOMS ozone, elevation, and other data. More data will be added once the disk space becomes available.

The map server in the demo site can be accessed by many map clients, including Cubwerx client (<http://www.cubwerx.com/wmt/cubeview/cubeview.cgi>) and NASA Digital Earth client (<http://viewer.digitalearth.gov>). There are many servers accessible through those clients. Our server is named NASA-EOSDIS in those clients.

VII. FUTURE WORK

Although some NWGISS components are fully functional now, we still need to develop many others. In the near term (within one-year time frame), we plan to

- Modify the map server to comply with OGC WMS version 1.1 from WMT II, which represents a big enhancement in mapping capabilities;
- Make CreateCapabilities component working for both map and coverage servers and compliant with OGC specifications from WMT II;
- Implement the state-less catalog server;

- Develop the NWGISS coverage client and evolve the client into a multi-protocol comprehensive one that handles requests for coverage, map, catalog, as well as feature.
- Integrate the components into a single system

ACKNOWLEDGEMENT

This project was supported mainly by grants from NASA Earth Science Technology Office (ESTO). Additional funding was provided by Open GIS Consortium (OGC) for the development of NWGISS coverage server as a part of OGC WMT II.

REFERENCES

- [1] OGC Homepage, <http://www.opengis.org>
- [2] Di, L., and Rundquist D., 1994. "A One-Step Algorithm for Geometric and Radiometric Correction and Calibration of AVHRR level 1b DATA," *Photogrammetric Engineering and Remote Sensing*. Vol. 60, No. 2, pp. 165-171.
- [3] FGDC, 1999. The FGDC Content Standard for Remote Sensing Swath Data. FGDC-STD-09-1999.
- [4] Di, L., R.Suresh, K. Doan, D. Ilg, and K. McDonald, 1999. "DIAL-an Interoperable Web-based Scientific Data Server". In M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman edit, *Interoperating Geographic Information Systems*, Section 4. System Experiences. Kluwer Academic Publishers, Norwell, MA. pp 301-312
- [5] HDF-EOS GIS translators, <http://heineken.gsfc.nasa.gov/eosgis/eghome.html>
- [6] CEOS WTF-GOFC, <http://lennier.gsfc.nasa.gov/wtf-gofc/contents.html>
- [7] Digital Earth Alpha version, <http://viewer.digitalearth.gov>
- [8] NASA ESIP, <http://www.esipfed.org>