A Gateway to Support Interoperability of OPeNDAP and OGC Protocols

Kenneth McDonald1, Yoseosok Enlou2, Leping Di2, Daniel Holloway3
1NASA Goddard Space Flight Center, Code 423, Greenbelt, MD 20771
2SGT Inc., 7600 Greenbelt Road, Suite 400, Greenbelt, MD 20770
3Center for Geospatial Information Science and Systems (CGISS), George Mason University, 6301 Ivy Lane, Suite 630, Greenbelt, MD 20770
4OPeNDAP Inc., 165 Dean Krauss Drive, Narragansett, RI 02882-1124

Abstract: Data access and analysis tools that are developed within specific disciplines and the protocols that they are built upon provide valuable services to their respective users but can actually be a barrier to the integration of data from a broad set of data sources. An example of this is data supported by OPeNDAP that is widely used in the ocean and atmospheric sciences, and data provided through the interface specifications of the Open Geospatial Consortium (OGC) that typically serve the land science community. This paper describes a project that is developing a gateway to bridge these two data system infrastructures, in response to a specific need expressed by CEOP, an international science program.

I. INTRODUCTION

The Coordinated Enhanced Observation Period (CEOP) is an element of the World Climate Research Program (WCRP) and was initiated by the Global Energy and Water Cycle Experiment (GEWEX) to focus on the measurement, understanding, and modeling of water and energy cycles within the climate system [1]. It has identified a network of thirty-six reference sites and several enhanced observing periods (EOPs) for which it collects and assembles data from satellite and in situ observations and satellite data and also model and data assimilation output products. To meet its goals, the CEOP science community requires data integration services that allow it to access and inter-compare these diverse data types from multiple sources.

The CEOP Program initiated a discussion with the Committee on Earth Observation Satellites (CEOS) Working Group on Information System and Services (WGISS) to determine if WGISS could assist in the development of advanced tools to access the various data collections with the data service needed to support data integration. Two of the WGISS agencies, the Japanese Aerospace Exploration Agency (JAXA) and NASA volunteered to work with CEOP in addressing these services. The NASA team chose to focus on the enhancement of access to its satellite data resources.

II. PROJECT DEFINITION

From early interactions with a number of CEOP scientists, NASA learned that they were currently using client/server applications based on the Open-source Project for a Networked Data Access Protocol (OPeNDAP) [2, 3] to access and analyze the field data and the model/assimilation products. The NASA team had been developing a set of data system components called the NASA Web-based GIS Services (NWGIS) [4] that could apply a number of data services to satellite data products, including geocoding, reformatting, resampling and reprojection, and makes the resulting products accessible via standard interfaces specified by OGC [5]. These data services are the same services needed to support data integration. The scientists saw the value of the NWGIS capabilities, but they needed the products to be served via OPeNDAP to be truly useful for data inter-comparisons of the diverse data types.

To address this issue, the NASA team proposed the development of a gateway that would allow an OPeNDAP client to access the data and services provided by an OGC-compliant web coverage service (WCS). With assistance from OPeNDAP Inc., a proof-of-concept demonstration was developed to allow a user of an OPeNDAP client (e.g., GRADS) to access a satellite data product from NWGIS and to overlay a model output from an OPeNDAP server [6]. While the demonstration was far short of a fully functional, generic capability, it did begin to address some of the challenges of translating between the two environments.

Building on the experience of the demonstration, a team of NASA, GMU, SGT Inc. and OPeNDAP Inc. wrote a successful proposal to NASA’s Advancing Collaborative Connections for Earth-System Science (ACCESS) Program. The proposed effort will integrate catalog services with the gateway and provide additional functionality to allow a user to fully specify the desired services to be applied to the satellite data. This effort will also merged with a second proposal to enable data held by OPeNDAP, or what was referred to as geoscientists, to be accessed by OGC clients, thus providing interoperability in the opposite direction.

III. THE GATEWAY DESIGN

While this ACCESS project is conceptually addressing two-way interoperability between OPeNDAP and OGC clients and servers, doing this requires the development of two gateways. The first one is called the CEOP Satellite Data Server (SDS) and the second gateway the OGC-Geoscience Gateway.

A. CEOP Satellite Data Server (SDS)

The CEOP SDS leverages the geospatial processing capabilities of the WCS with the OPeNDAP data access protocol or DAP, to expose a single, standardized data representation of a satellite data product consistent with other CEOP data sources. The project is using GRADS, a widely-used analysis and visualisation application to demonstrate access to satellite data products, in particular data from NASA’s EOS misions, for comparison with field measurements and model output. GRADS is an excellent example of a pre-existing application that has been OPeNDAP-enabled, meaning it has the additional capability to make requests to, and retrieve data from OPeNDAP servers located anywhere on the Internet.

There are two fundamental elements required to integrate these two data access technologies such that existing applications, like GRADS, can readily access the data. The first is a standardized representation or data model for satellite data that is consistent with other CEOP data sources and into which the server will transform the native satellite data product. The second is a set of software components that will transform elements of a DAP request into a WCS request, issue that request to a WCS instance and transform the WCS response into a DAP data object consistent with the standardized representation, and return that DAP object to the requesting client. Figure 1 shows the design of the server.

A.1 Data Model

To facilitate interoperability with CEOP model data, the server utilizes a CF-based (Climate and Forecast) metadata multi-dimensional grid structure for representing the requested satellite data. The CF standards use the dimension ordering of (x,y,z,t) to regularize the shape of multi-dimensional data sources. Additionally, CF implements and utilizes extended naming convention with regard to physical quantities, their units of measure and associated attributes to facilitate data access and use within analysis applications. Initially, the server implementation will focus on the two dimensions of the CF grid structure (x,y), and not attempt to interoperate with the remaining dimensions (t).

A.2 Software Components

The server comprises components that evaluate the OPeNDAP client request, generate valid OGC WCS Coverage Service (WCS) requests, and return the response from those services into the structural Grid representation prescribed by the common data model. Each component can be envisioned as a package consisting of classes operating together to provide a functional element for the server. The following paragraphs provide a brief introduction of each component.

- DAP Satellite Data Model Container: This component instantiates a DAP data object conforming to the CF-based representation of the satellite product, relative to the default grid parameters specified for the product in the external configuration document. This component is the primary DAP element when the server operates in the non-function-based, or default DAP mode.
- DAP Function Evaluator: This component evaluates any server-side functions that are contained in the client request and creates a DAP data object consistent with the results of that functional request. As described earlier, the functional interface will consist of a predefined, extensible, set of function names with well-defined argument structures. The set of functions will consist of functions to specify geospatial selection with latitude-longitude boundaries or center latitude-longitude coordinates, grid cell resolution and interpolation methods. As a result of function evaluation a DAP Satellite Data Model Container will be instantiated with the parameterization necessary to represent the result of the functional request. That container will then operate to create the WCS request and process the WCS response into an appropriate DAP response to the request.
- WCS Request Generator: This component converts the WCS request elements generated by the DAP Satellite Data Model Container and/or DAP Function Evaluator, with additional WCS request elements from the Client Request to generate a valid WCS request for processing the satellite data product.
- WCS Response Format Handler Plug-in: The WCS Implementation Specification allows a WCS to return, minimally, any one of five well-known binary file formats, and optionally any other file format it chooses for a "Layer". It advertises the multiple response formats from a WCS, the server utilizes a plug-in framework to instantiate the proper format handler for the

Figure 1: Architecture diagram for CEOP Satellite Data Server Design

https://ntrs.nasa.gov/search.jsp?R=20070014931 2020-01-13T11:38:55+00:00Z
resulting response format from the WCS. The plug-in can be viewed as a specialization of a base class that operates to read WCS well-known binary formats, creating a standardized data structure that can be used by the OpenDAP interface module to form the response back to the requesting client. This component defines the behavior of these base classes, thus enabling multiple implementations to provide support for different WCS data sources.

**Scheduler:** This component is the primary engine of the server. It coordinates the activities of the worker processes comprising the server. Active management is provided to support multiple threads for handling simultaneous WCS requests, and the potential asynchronous nature of their completion. The component determines the required format Handler from the Configurator component and instantiate the appropriate plug-in module. Additionally, this component will support caching to provide enhanced throughput for particular DAP usage patterns.

**Configurator:** Initially this component will utilize an external XML configuration document to identify the satellite data products available to the server from a WCS. The configuration parameters include WCS layer information for each satellite product available for transformation by the server. Included in the configuration is the default grid resolution and interpolation parameters to use for transforming the satellite product from its native representation to the server's DAP representation. The configuration document also stores which WCS response format to request, and the format handler to plug-in for processing the WCS response for this layer. Further development is envisioned to support integration with OGC CSW services for maintaining the configuration of new products and services on the server.

**THREDDS Catalog Response Generator:** The OpenDAP servers and clients generally support the THREDDS catalog interface. To expose the available satellite data products from the server to the OpenDAP community, this will be a "simple" THREDDS catalog response. This component translates the relevant elements of the external XML configuration document to a THREDDS catalog. THREDDS catalogs are themselves XML documents that can describe both directory and inventory level information for a data interface for those base classes. THREDDS catalogs will include the specific information OpenDAP URLs to the satellite products available from the server, as well as the spatial and temporal extent of the individual satellite products, and information available from the WCS describing the satellite data. THREDDS catalogs are becoming generally available from data providers using OpenDAP servers and the suite of OpenDAP-enabled client implementations are migrating toward this catalog system. The CEOP SDS will undergo extensive testing to prepare for the capability to provide this catalog available to other interested CEOS agencies (e.g., JAXA, ISRA) willing to host the software with Web Coverage Service access or OpenDAP server access. After this initial deployment of this gateway, the CEOP science community will test the SDS and provide feedback on ideas for iterative enhancements.

### B. OGC-Geoscience Gateway

The OGC-Geoscience Gateway uses a set of interface specifications developed by the Open Geospatial Consortium (OGC) to access geospatial data served by geospatial protocols such as OpenDAP and THREDDS. The major work of the geospatial gateway is to rapidly develop and deploy an OGC-geospatial gateway for facilitating the interoperability from the OGC catalog and data access protocols (CSW [7], WCS [8], and WFS [9]) to the catalog and data access protocols used in the geoscience communities (THREDDS [10], OpenDAP, and netCDF DAP). The gateway will enable OGC clients to search and access data served through these geospatial protocols. Figure 2 shows the overall architecture of the gateway.

![The Architecture of OGC-Geoscience Gateway](image)

**Figure 2. The Architecture of OGC-Geoscience Gateway**

The shaded components in the figure are those that the prototypes already exist. This project only needs to update, enhance, assemble, and test them in the OGC-Geoscience Gateway instead of developing them from scratch. As shown in Figure 2, many of the components already exist. This allows us reusing those components to rapidly develop and deploy the gateway.

The following paragraphs describe the major new developments for the three sub-gateways: CSW, WCS, and WFS.

**CSW to THREDDS sub-gateway:** This sub-gateway will provide CSW interface to OGC clients. The NCWIGS CSW server will be reused in this sub-gateway. The major development in the sub-gateway is to develop the catalog endpoint that reads the THREDDS XML data catalog document and ingests it into the CSW searchable MySQL database. Tools also need to be developed for facilitating the creation of semantic catalog. The CSW will be enhanced so that it can combine the information in the data catalog and the semantic catalog to generate valid CSW search results that can be used by OGC clients to formulate valid WCS or WFS data retrieval requests.

**WCS to geoscience-server sub-gateway:** The THREDDS prototypes already developed by UCSD will be used as the starting point for the sub-gateway. The major work for this sub-gateway is to enhance the prototype so that it becomes a fully compliant, fully functional OGC WCS server, from the viewpoint of OGC WCS clients. The enhancement includes providing services of map-coordination information, indexing, reformatting, etc. to coverage data before the data being sent to OGC clients. Those services will be exposed to the WCS clients through the capabilities description.

**WFS to geoscience-server sub-gateway:** Currently, there are no components available for this sub-gateway. Therefore, the sub-gateway will need to be developed from scratch. The similar architecture as the WCS sub-gateway will be used. Besides the protocol translation, the major work for this sub-gateway will be the development of constraint constraint and implementation of features data selection function.

To design the gateway for facilitating the protocol interoperability, the general pattern for data access in OGC client-server model was considered. A typical OGC session for accessing geospatial data starts with a client searching for a data catalog at a CSW server (data discovery step). Once the requested dataset is found, the client can obtain the dataset by issuing a getCoverage request to the WCS server or a getFeature request to the WFS server which hosts the dataset (data retrieval step). The design and implementation of the Geoscience gateway will require an understanding of the metadata used by the OGC CSW and the THREDDS protocols to determine what additional metadata is needed for CSW access to THREDDS catalogs and also what metadata is needed in the catalog search results to construct a valid WCS or WFS access request.

Under the current ACCESS award, the project will develop the CSW/THREDDS subgateway (the catalog subgateway) in the first and the second years. The OGC-GeoscienceGateway sub-gateway (the WCS subgateway) will be developed in the optional third year, if it is funded.

### IV. Conclusion

The successful implementation and deployment of the gateway will improve access to a large amount of NASA EOSDIS data from OpenDAP clients to large amounts of geoscience data from OGC clients. Currently, NASA EOSDIS data is available to CEOP scientists only via the pre-processed datasets for a limited CEOP reference site. The gateway will also greatly facilitate the utilization of satellite data in the CEOP modeling efforts. The gateway will provide many of the data services needed to support data integration and enable the quick and effective inter-comparison between model output and satellite observations and in situ reference data. The gateway will significantly improve the scientific productivity of CEOP scientists. The data services needed to support data integration include online visualization tools that facilitate inter-comparison of satellite data with model output data and in situ reference data are also needed by other scientists outside the CEOP community. The gateway could be offered to other science communities needing the capability to inter-compare between model output and in situ data. The gateway will also significantly reduce the time for obtaining data access and provide much better data services to CEOP scientists. The gateway will also facilitate the pre-processing of making CEOP datasets be automated and the CEOP scientists use this pre-processed, on-demand data products generated on their specifications.

The gateway will positively impact not only the CEOP program, but also other programs that use OpenDAP protocols for data access and want to use the satellite data in their efforts. The CEOP Satellite Data Server developed in this project will use the standard WCS protocols to communicate with any OGC WCS compliant server. Therefore, any WCS server can act as an OpenDAP server by deploying this server which acts as a gateway server. Currently, most space agencies around the world are developing or deploying WCS servers. The availability of such a gateway server will make all satellite data in those space centers immediately available to Earth science research and modeling community who commonly use OpenDAP protocols for data access. Finally, this project also facilitates the interoperability from geospatial community to the Earth science research and modeling community and makes the use of geospatial technology in the Earth science community.

### Acknowledgements

This research is supported by NASA ROSES 2005 ACCESS program (05-ACCESS-03-28).

### References


