Exploring the Sun and its effects on the Earth's atmosphere and physical environment...

HIGH ALTITUDE OBSERVATORY

Earth System Grid

The OPeNDAP and Remote NetCDF Invocation (RNI) middleware platform for Scientific Data Fusion.

Stephan Zednik[1] (zednik@ucar.edu), Jose Garcia[2] (jgarcia@ucar.edu), Peter Fox[3] (pfox@cs.rpi.edu), Patrick West[1] (pwest@ucar.edu)

[1]High Altitude Observatory/ESSL/NCAR P.O. Box 3000, Boulder, CO, 80307 United States

[2]Computational and Information Systems Laboratory/NCAR P.O. Box 3000, Boulder, CO, 80307 United States

[3]Tetherless World/Rensselaer Polytechnic Institute 110 8th St., Troy, NY, 12180 United States

Abstract

Across geosciences, there are large data holdings being made available via the DAP protocol by means of OPeNDAP software. A lot of the underlying data are in the NetCDF format. Often, each individual dataset is a combination of hundreds of individual NetCDF files. Requesting such datasets for analysis is an expensive data fusion transaction, especially as the number and size of datasets increase. We present a set of solutions that instead request needed portions of the data fused just-in-time. The fusion includes both subsetting and aggregation operations as well as analysis and data manipulation steps. We have modified the NetCDF C library for Remote NetCDF Invocation (RNI) that is, to operate on remote dataset, over HTTP, HTTPS or GSIFTP (or any) protocols, individual NetCDF Application Programming Interface (API) calls as if they were local. This Invocation model can also be applied to OPeNDAP data streams, and local files. This mechanism resembles the well known Remote Procedure Call (RPC) yet it radically differs on the binding between local and remote remote operations. We describe our current approach, implementation and benefits obtained from this approach and indicate how it aids data fusion.

The Problem

It is impractical to fetch dataset files in their entirety from the remote repository for client analysis. Modern datasets can be multi-GB in size and will be tens of TB in the near future. These datasets are comprised of potentially thousands of distributed files putting an undue burden of knowledge of dataset representation on the client. Data acquisition should be streamlined and efficient, should support common security mechanisms, and only the data requested should be retrieved. Data access representations should be complementary to analysis operations and not require client knowledge of the file-level dataset representation.

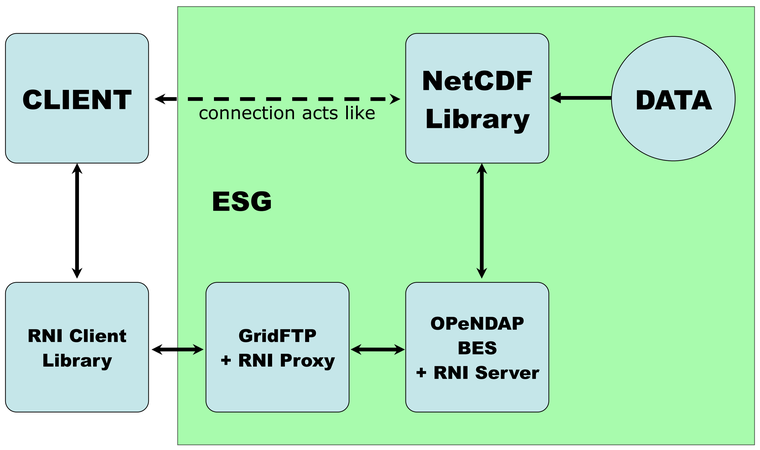
The **Earth System Grid (ESG)** is an example of a distributed data system confronting this problem

- **TB+ datasets distributed across thousands of files**
- **Federation of distributed data nodes with a common security infrastructure**
- **Data requests must support subsetting and aggregation**

Remote NetCDF Invocation

Remote NetCDF Invocation (RNI) is an Application Programming Interface (API) based system for secure data access designed to meet ESG needs for a heavy-weight data access client. The RNI client is dynamic library that preserves Application Binary Interface (ABI) compatibility with the NetCDF C dynamic library. This means that applications built against NetCDF need only re-link against the RNI client to support RNI data access.

To support the ESG security infrastructure RNI uses a 3-tier architecture. The client communicates with a ESG GridFTP server which authenticates the user and forwards the client request to a data node OPeNDAP server which authorizes the data request and executes it on behalf of the client. Data operation results are packaged as an OPeNDAP data object and passed back to the client.



RNI linked client access to NetCDF data on the ESG using a GridFTP server as grid proxy

GLOSSARY

OPeNDAP: Open-source Project for a Network Data Access Protocol

OLFS: OPeNDAP Light-weight Frontend Servlet. HTTP front-end to the OPeNDAP Hyrax architecture

BES: OPeNDAP Back End Server. Extendable High-performance server to the OPeNDAP Hyrax architecture

GridFTP: a high-performance, secure, reliable data transfer protocol, based on FTP, optimized for high-bandwidth wide-area networks. Provided by the Globus Alliance.

GSIFTP: a subset of the GridFTP protocol using GSI, provided by the Globus Alliance.

MATLAB: numerical computing environment and programming language specializing in matrix manipulation

MEXNC: NetCDF interface for MATLAB

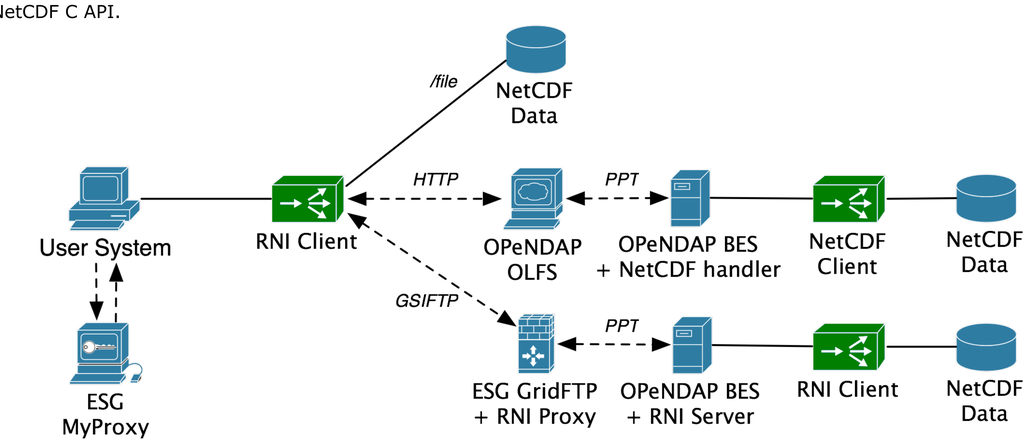
PPT: Point-to-Point Transfer protocol. OPeNDAP project protocol used primarily between the OLFS and BES in the OPeNDAP Hyrax architecture, a simple protocol designed for high-performance and not meant for public network use.

Leveraging the NetCDF API

RNI chose the NetCDF C interface because it is a well-known standard with an established community. A large portion of the ESG data holdings are in the NetCDF format; many analysis programs already exist for ESG data built upon the NetCDF C API.

- RNI is ABI compatible with the NetCDF C library, ESG users will be able to integrate RNI with their existing NetCDF-based data analysis programs with little effort. For example, linking MEXNC (NetCDF interface for MATLAB) against the RNI client provides secure access to ESG data holdings from MATLAB. No MEXNC modifications or rebuild required, just re-link and run.
- The NetCDF interface also provides a mechanism for user specified subsetting. NetCDF variable retrieval methods offer robust support of several subsetting techniques and fulfill ESG's requirement for data subsetting support.

RNI chose the NetCDF C interface because it is a well-known standard with an established community. A large portion of the ESG data holdings are in the NetCDF format; many analysis programs already exist for ESG data built upon the NetCDF C API.



RNI client access to NetCDF files over local file system, OPeNDAP streams over HTTP, and ESG data over GSIFTP. *Dotted lines represent network connections.*

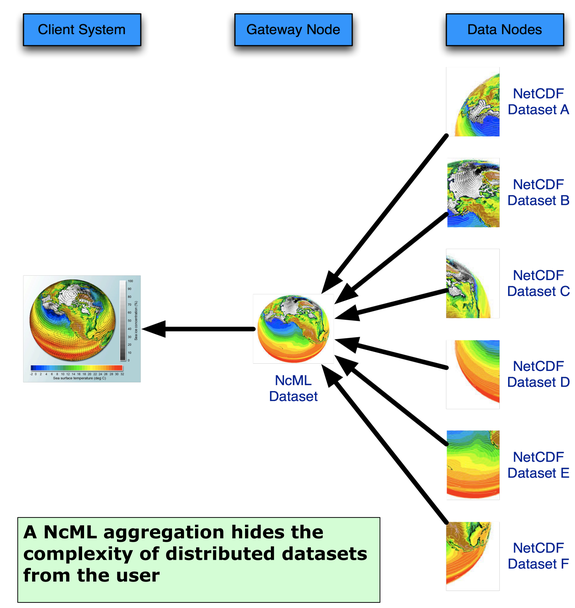
Fusing Datasets with NcML

The NetCDF Markup Language (NcML) is an XML representation of the metadata and structural content of a NetCDF dataset. NcML can be used to describe the structure of a 'virtual' dataset, that is, a dataset comprised of variables aggregated from other NetCDF datasets. NcML supports variable aggregation by

- **Union**: join data from different datasets
- **Tiling**: join data along existing dimension
- **Layering**: join data along new outer dimension

NcML is a community standard geared towards the NetCDF format that fulfills the aggregation requirements of ESG.

RNI leverages NcML by treating an NcML aggregation representation as if it were a NetCDF dataset. When a file open request is made on a local NcML file, the NcML engine within the RNI client builds an internal mapping of NcML dataset variables to the NetCDF dataset where they are stored, accessible by local file system, HTTP/S, PPT (RNI 2-tier), or GSIFTP (RNI 3-tier). When a data request is made the RNI NcML engine makes the request across the proper connection(s), builds an aggregated response (if needed) and returns data as if the NcML file had contained the aggregated data itself.



A NcML aggregation hides the complexity of distributed datasets from the user

Performance through Parallelization

The RNI architecture for ESG is parallelizable, and this parallelization improves performance for access to very large datasets. Distributing the data files of NcML defined aggregations across many data nodes allows the RNI NcML engine to parallelize data retrieval, spreading disk read penalties for many large files across several nodes.

RNI also improves performance by

- High-performance GridFTP parallel stream connection utilized between ESG Gateway GridFTP server and client.
- Data requests are evaluated lazily, so only the data explicitly requested is ever fetched.
- Metadata and structural information about the dataset is cached, so the client can answer non-data requests.

Conclusion

RNI provides fast and streamlined client access to both local and remote datasets over a variety of data protocols. RNI was designed as a NetCDF ABI compatible library so a single client library can power many heavy-weight clients designed for NetCDF. The client library is modular in design so new access protocols (HTTP/GSIFTP) and new dataset representations (NcML) can be added.

The RNI 3-tier architecture provides secure API-level access to remote ESG data holdings; supporting data subsetting and data aggregation.