



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

P. Fox, S. Zednik, and P. West

Rensselaer Polytechnic Institute, World Tetherless Constellation, Troy, United States (pfox@cs.rpi.edu, +15182764464)

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 dataset 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 also can 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 operations. We describe our current approach, implementation and benefits obtained from this approach and indicate how it aids data fusion.