



Grid-enabled OPeNDAP Hyrax servers: parallel/multiple back-end services and remote netCDF invocation

Garcia, J., Zednik, S., West, P. and **Fox, P.**

(1) HAO/NCAR, Boulder CO USA

We present our developments for data-intensive server-side parallel processing in Grid environments in the context of two science applications: inversion of Stokes polarimeter parameters for magnetic field components on the surface of the Sun, and the analysis of very large climate model records via the Earth System Grid.

In the first application, we have constructed a dedicated Grid using the OPeNDAP Hyrax Back-End-Server (BES), as the middleware that facilitates all the interprocess communication required to find the complete cohesive solution for each dataset. We describe the parallel approach taken to speed up the computation of the solution, as well as the specifics of the Grid design. Each BES acts as a computational node, obtaining an OPeNDAP URL specifying the dataset on which to operate and the specific slice using OPeNDAP constraints. Each BES returns the solution to the controller as an OPeNDAP data object. The Controller collects all solutions and assembles a complete result as a netCDF data file.

In the second application the vast repositories of netCDF data in the Earth System Grid (ESG), make it impossible or impractical for users to download and replicate the complete database. Furthermore, each individual dataset, is a combination of hundreds of individual netCDF files. We present our efforts with Remote NetCDF Invocation (RNI), that is, to operate on a remote dataset, over HTTPS and gsiFTP protocols, individual netCDF Application Programming Interface (API) calls as if they were local. Our design is based on the extensibility mechanism provided by the OPeNDAP Hyrax BES middleware platform with Globus GridFTP and Apache modules acting as

the proxy transport mechanism (binding) between the local and remote transactions.