STATEMENT OF WORK

Streaming IDX Server in Earth Systems Grid (ESG) Node

Required Tasks

The subcontractor shall perform the following tasks

Install ViSUS IDX server with ESG Node

Built on the hierarchical streaming IDX data format, the ViSUS server and client facilitate exploratory analysis of disparately located massive datasets. This functionality can be used as a front end that will enable climate scientists to first explore a variety of analyses, interactively varying inputs and operations until a desirable combination has been discovered, after which a traditional comprehensive analysis can be carried out across all time steps using existing climate analysis tools.

The subcontractor shall install the ViSUS server in association with an ESG node. The customized IDX server will accept requests for ESG datasets and either provide them from a cache or automatically convert them to IDX before serving. The listing and querying of available datasets shall continue to be performed using existing ESG interfaces.

Provide dynamic data conversion to streaming IDX

The subcontractor shall implement dynamic data conversion from HDF5 NetCDF files to streaming IDX. When a request for a given ESGF dataset is received, the IDX server should provide the data to the client, first considering the cache to see if the dataset has already been converted. If the data is not already available as IDX, the server should instantiate an automatic conversion request for the data and inform the user that a conversion is in progress. Once available, if the user is still connected a notification will be sent and the request can be fulfilled. If the user is no longer connected, the client itself can periodically request the data which will be delivered when available. The subcontractor should explore the viability of maintaining (secure?) connections with the server versus periodic re-requests from the client.

Rudimentary Cache Management

The amount of space to be reserved for dynamic exploratory analysis can vary from node to node. The subcontractor will implement simple least recently used (LRU) caching for converted data.

Additional work (outside scope of this SOW)

There are several additional tasks that might be useful, but fall outside the scope of this project:

- Integrate ViSUS client with UV-CDAT or other existing climate analysis tools
- List, search, or organize ESG datasets
- Provide a client interface to facilitate analysis
- Explicitly determine hierarchies of datasets to be combined, e.g. by combining datasets from server A and B using server C, then combining the result with data from server D using the client.
- Finalize (i.e. polish) integration of the ViSUS server with the ESG node installation