Development of a spatial data infrastructure for cryosphere community based on open source software tools

Weibo Liu (KU Geography, CReSIS), Kyle Purdon (NSIDC), Trey Stafford (UNL Geography), John Paden (CReSIS), Xingong Li (KU Geography)

Abstract

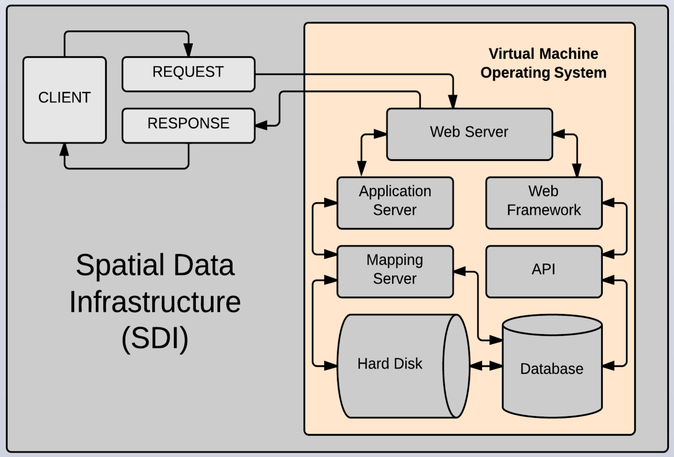
The Center for Remote Sensing of Ice Sheets (CReSIS) at The University of Kansas has collected approximately 700 TB of radar depth sounding data over the Arctic and Antarctic ice sheets since 1993 in an effort to map the thickness of the ice sheets and ultimately understand the impacts of climate change and sea level rise. In addition to data collection, the storage, management, and public distribution of the dataset are also one of the primary roles of CReSIS. The OpenPolarServer (OPS) project developed a free and open source spatial data infrastructure (SDI) to store, manage, analyze, and distribute the data collected by CReSIS. The OPS SDI includes a spatial database management system (DBMS), map and web server, JavaScript geoportal, and application programming interface (API) for the inclusion of data created by the cryosphere community. Open source software including GeoServer, PostgreSQL, PostGIS, OpenLayers, ExtJS, GeoEXT and others are used to build a system that modernizes the CReSIS SDI for the entire cryosphere community and creates a flexible platform for future development.

Objectives

- Develop and deploy a database management system.
- Develop and deploy a web-based data retrieval system (geoportal).
- Develop and deploy an API for interaction between MATLAB and OPS system.
- Deploy OPS system to the public and include community derived datasets.
- Share the source code following standard free and open source software (FOSS) guidelines.

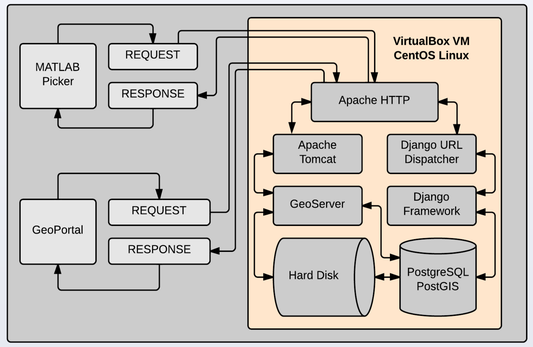
SDI System Structure

The OPS SDI is a collection of software and custom code used to store, update, analyze, manage, and distribute spatial data. The below figure shows the conceptual structure of the OPS SDI. The SDI is composed of seven basic components. 1) The OS serves as the platform to install and execute other SDI components on the server; 2) A web server allows and controls the flow of data from anyone connecting to the server via the Internet; 3) Web Application Server stores and manages one or more server applications; 4) Map Server is to generate georeferenced spatial data formats; 5) Web Framework connects server-side code to the Internet via Web Server; 6) RDBMS is used to store and organize the spatial data; 7) A client is any software or service that makes a request to the OPS Web Server.



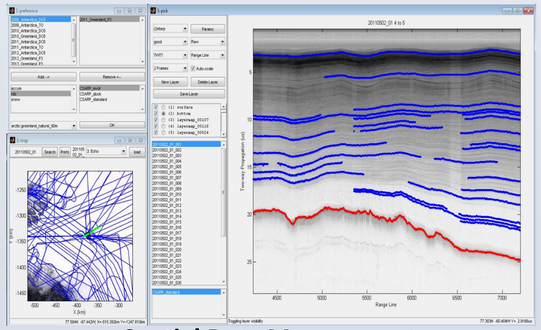
OPS SDI Implementation

The implementation of the OPS SDI follows the conceptual structure. The below figure demonstrates the actual software selected and its implementation on the OPS.



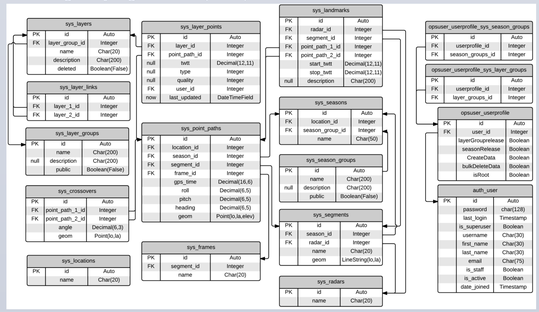
Spatial Data Generation

The Data Picker in the below figure is a MATLAB tool developed at CReSIS for the manual and automated digitization of ice layers. The Data Picker generates the ice surface, ice bottom, and other layers, and the OPS allows the inclusion of any layers with spatial information via the OPS MATLAB API.



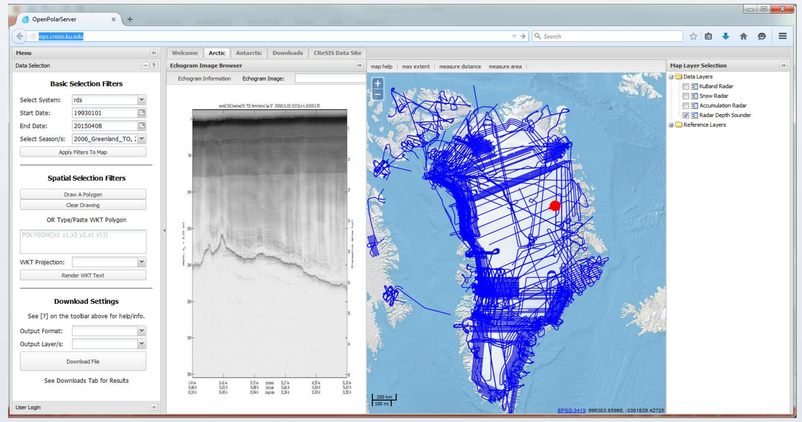
Spatial Data Management

One of the major contributions of the OPS SDI is the use of a relational database management system (RDBMS). The RDBMS chosen is PostgreSQL, and PostGIS was chosen as the spatial database extender for PostgreSQL and adds spatial queries in SQL. The complete OPS database schema is shown in the below entity-relationship (ER) diagram. Each table corresponds to a single Django model (class).



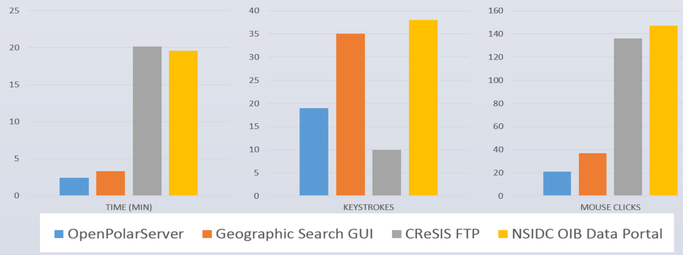
OPS SDI GeoPortal

The OPS GeoPortal is a JavaScript Web application that is designed with the discovery and download of geographic content (OPS data) as the primary function. The GeoPortal GUI provides the spatial and temporal data filtering options, download format options, map interface for Arctic and Antarctica, echogram browsing panel, layer selection tree with four major Radar systems (KuBand, Snow, Accumulation, and Radar Depth Sounder), and reference layer selection tree.



Usability Analysis

In order to verify that the OPS SDI has achieved the goal of providing an improved experience for end users. Four existing systems were compared with the same task, “download CReSIS L2 data in the csv-good format over a NE Greenland outlet glacier (79N) including only data from 2010 and 2011”. The below figure shows that the OPS had the fastest processing time, the least amount of mouse clicks, and the second least amount of keystrokes.



Conclusions & Future Work

The OPS has been developed using all FOSS and conforms to practical data and coding standards. The system is public at https://www.ops.cresis.ku.edu and many cryosphere community data providers have or are preparing to include their datasets in the OPS. The code is open source and provided free of charge or restriction on GitHub at https://github.com/CReSIS. As an open source project the OPS will only be successful if the community contributes to its growth. Some possibilities for future exploration of the OPS are:

- Cloud-based hosting of the SDI on systems such as RACKSPACE or Amazon EC2.
- An interactive (web based) data picking system build in JavaScript to replace the MATLAB data picker.
- An interactive JavaScript echogram browser that loads dynamic data instead of static images.
- A Web based data loading system so community members can load their own data into a centrally hosted system without CReSIS intervention.

Apache HTTP SERVER · Apache Tomcat · python · django · PostgreSQL · PostGIS · GeoServer · GeoExt · OpenLayers™ · CentOS · VAGRANT

NSF · NASA · The Kansas Board of Regents 1925

CReSIS Center for Remote Sensing of Ice Sheets · The University of Kansas