

Discovery and Delivery of Space Geodetic Data Products from Distributed Archives

Annual Report

**Submitted by Frances Boler, PI
UNAVCO, Inc.
6350 Nautilus Dr.
Boulder, CO 80305**

Performance Period: April 1, 2011 through March 31, 2012

NASA Grant Number NNX10AF07A

Submitted to

**Stephen Berrick, NASA Headquarters, Technical Officer
Martha Maiden, NASA Headquarters, Technical Officer
NASA/NSSC, Grants Officer**

The major components of the GSAC-WS system are shown in Figure 1. The ability to utilize the GSL as a mechanism to federate queries is achieved by implementing the Federated Repository component, shown in Figure 2.

The GSL technology required refinements during the second year of the project as each of the partners worked to achieve implementations that provided robust access to their metadata, data and products. While the component software went through several releases during the year, the bulk of the work at the partner archives was in implementation, customization, and readying the systems for production and metrics reporting.

The project efforts for year for each partner are summarized below.

UNAVCO

In addition to adding to the overall system capabilities and working with the other partners to improve their implementations, UNAVCO improved the real-time query capabilities of the GSAC system. UNAVCO experimented with a separate repository for the Plate Boundary Observatory real-time holdings because that system employs an independent database within UNAVCO. Ultimately, the PBO real-time metadata delivery was converted to an in-house web service built on the PBO real-time database. This system provides connection information and statistics about each real-time stations data completeness and latency.

UNAVCO also developed code for logging GSAC-WS system queries and URLs returned (URLs returned are a measure of products delivered via the GSAC-WS system). These logging capabilities were included in a system release for dissemination to project partners. Separate scripts for parsing the GSAC-WS logs were developed and released to the archive partners. In addition, a system for reporting partner metrics to UNAVCO utilizing Google Docs spreadsheet was implemented.

SOPAC's Site Log XML with extensions for real-time station metadata was incorporated as an output encoding for GSAC-WS metadata.

The UNAVCO Repository web GUI can be accessed at:

<http://facility.unavco.org/gsacws/gsacpi/site/form> (site search)

<http://facility.unavco.org/gsacws/gsacapi/file/form> (file search)

The UNAVCO Federated Repository web GUI can be accessed at:

<http://facility.unavco.org/gsacfederated/gsacpi/site/form> (site search)

<http://facility.unavco.org/gsacfederated/gsacapi/file/form> (file search)

UNAVCO has had a personnel change in that Jeff McWhirter, lead developer for the project, has left full-time employment with UNAVCO. He remains on the project in an on-call capacity. Matt Beldyk has taken over GSAC-WS software support and system implementation at UNAVCO.

SOPAC

SOPAC staff continues to maintain the SOPAC repository component of GSAC. SOPAC added to the XML encodings of real-time metadata for augmenting their Site Log XML. SOPAC contributed documentation to the project which provides a comprehensive description of the GSAC-WS system in comparison to the legacy GSAC flat file exchange system.

The SOPAC Repository web GUI can be accessed at:

<http://geoappdev02.ucsd.edu:8080/gsacws/gsacapi/site/form> (site search)

<http://geoappdev02.ucsd.edu:8080/gsacws/gsacapi/file/form> (file search)

SOPAC is in the process of gathering metrics and moving the application to a virtual server from its current location on a development server. SOPAC is also incorporating additional links to their GSAC-WS API and GUI from their website, including access to real-time metadata.

CDDIS

The CDDIS staff continues development of a custom interface to the GSAC Service Layer (GSL) for use on the CDDIS website. This application leverages the capabilities of the GSL while maintaining a consistent user experience for those exploring the CDDIS website for data holdings. This interface allows for more customization and will allow users to search for sites and files by specifying spatial, temporal and other parameters in a way that is unique to CDDIS data holdings. The more generic parts of this work can be rolled into the GSL for use by other partners.

The new interface is being developed using an open source application framework and interacts with the GSL for both site and file searches. The custom user interface allows collection of fields related to both site and file searches. The new client requests raw data from the GSL and displays the results in JavaServer Pages, combining the results with a customized openlayers map.

This development effort was presented at the 2011 Fall AGU: C. Noll, N. Pollack, P. Michael. "Improvements in Space Geodesy Data Discovery at the CDDIS", Abstract IN41B-1410 presented at 2011 Fall Meeting, AGU, San Francisco, Calif., 05-09 Dec.

UNR

The primary objective of the UNR team was to provide quality assessment statistics on the GPS data, which UNAVCO could then distribute to users. The statistics are based on post-analysis rather than pre-screening, as it more closely meets user needs. To meet this objective the UNR team has continued to download data from all continuous GPS stations that are in the UNAVCO/SOPAC/CDDIS archives. They process the data as part of their global analysis. During this project year they changed to using JPL's ITRF2010/IGS08 products and GIPSY 6.1 software with a new single-station ambiguity resolution method known as WLPB (Wide-Lane and Phase Bias), recently developed by JPL. The re-analysis of all data (which include more than those in

the aforementioned archives) took ~3 months. From this analysis, they now routinely extract various parameters concerning the quality of each rinex file. These parameters include number of phase biases, residual scatter of the observations, satellites tracked, number of outliers, formal errors, and level of multipath. These products have been provided to UNAVCO, thus closing the loop. The positive benefit of using the WLPB method of single-station ambiguity resolution (instead of the now-deprecated Ambizap network algorithm) is that data that arrives late can be included and quality-assessed without affecting all other stations. Next step during the no-cost extension period is to create a routine weekly update of the solutions and QA parameters, and make that automatically available for UNAVCO pickup.

The second major objective of the project was for the UNR partners to use the gsac client in their data ingestion software. Because that software is operating well, it was decided not to replace it with the gsac client and to use the gsac client only as a new feature in UNR's software that it currently lacks: to easily pick up data that has arrived late (i.e., after UNR's latest fetch) in the UNAVCO/SOPAC/CDDIS archives. So far the success of this approach has been quite limited, mostly because the gsac client doesn't respond well to queries that involve a lot of stations at once. This will be worked on with UNAVCO engineers during the no-cost extension of the project.

In parallel activity at UNR that closely relates to this project, UNR in September 2011 started a prototype public "next-day" service that provides time series and post-analysis statistics on the scatter of 5-minute precise point positions with ambiguity resolution. These QA statistics are currently provided for over 1,000 stations in the western US, and provide rapid feedback on low-latency data, such as the growing number of stations with real-time capability. As they are public, these data are also available for UNAVCO pickup, and could also be made part of the existing QA stream. Results are very encouraging, as the RMS scatter of 5-minute positions are strongly site-dependent, and they perhaps provide the most relevant statistic to assess the damaging effect of multipath on GPS positions.

ESDSWG Participation

The ESDSWG representative (Jeff McWhirter) participated in the Technology Infusion Working Group (TIWG), and attended the Software Reuse Working Group telecons and the Service Interoperability and Orchestration (SIO) group telecons. However, since Jeff has left the project his role on ESDSWG will be fulfilled by Fran Boler and/or Matt Beldyk for the remainder of the project.

Project Schedule

The software technology development for the project is largely completed.

For the no cost extension period, the following milestones and deliverables will be achieved:

- Implementation of weekly runs of the UNR QA/QC product generation is a remaining deliverable for UNR. Archiving and dissemination of these products in a production capacity is a remaining deliverable for UNAVCO.

- SOPAC's GSAC-WS implementation will be migrated to a production server.
- SOPAC will add to their real-time GSAC-WS API and GUI implementation by enabling queries that allow filtering of sites returned based on whether the site has real-time capability and will provide real-time metadata including connection information, latency and completeness statistics.
- SOPAC will incorporate links to their GSAC-WS API and GUI from their website.
- CDDIS will complete the GSAC-WS implementation and migrate to a production, publicly queryable system.
- Documentation aimed at software engineers exists on the UNAVCO website and within the SourceForge software package. Additional documentation aimed at project managers, archive managers, or science users is a remaining deliverable. Documentation of the QA/QC product is also a deliverable to be completed.