Dataworks provides data management and distribution software subsystems as open source modules that can be employed by regional GPS/GNSS managers for small to medium scale networks (e.g. 10-100 stations).
Motivation

Recognizing that many organizations operate GNSS stations but do not have the expertise to write their own software systems for the fundamental tasks of GNSS data and metadata management, UNAVCO created Dataworks for GNSS. These software modules are intended to keep the tasks of handling incoming data, ingesting, metadata storage, and presentation to the users manageable for smaller institutions. UNAVCO’s development of Dataworks was supported by NSF funding to COCONet for Regional Data Centers.

Functions

- Handle Incoming Data
- Extract Metadata
- Maintain Database of Metadata
- Data Stored and Backed-Up
- Data and Metadata Distribution Web User Interface
- Web Services
- Mirroring and Federation Options
GSAC (Geodesy Seamless Archive Centers)

- built to enable simplified search across archives
- implemented in Java
- uses uniform web services, enables federation
- built by partners UNAVCO, SOPAC, CDDIS
- NASA ACCESS funding 2010-2012
- NSF funding 2013-2014
- adopted in Europe by EPOS
- installations in Europe revealed lack of software infrastructure
- GSAC prototype database initiated

GSAC has been installed at 8 centers in Europe
<table>
<thead>
<tr>
<th>Implementation</th>
<th>Website URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>INGV, Italy</td>
<td><a href="http://bancadati2.gm.ingv.it:8081/gsacring/gsacapi/site/form">http://bancadati2.gm.ingv.it:8081/gsacring/gsacapi/site/form</a></td>
</tr>
<tr>
<td>RENAG, France</td>
<td><a href="http://webrenag.unice.fr/gsac/">http://webrenag.unice.fr/gsac/</a></td>
</tr>
</tbody>
</table>
DATAWORKS MODULES

Required:
• Specified OS (CentOS Linux) and Services (e.g. Tomcat, FTP)
• Dataworks Database Schema running on MySQL
• GSAC (Geodesy Seamless Archive Centers Software)

Optional:
• Mirroring (from partner data centers)
• Receiver download and associated data management
• Metrics and Backup
Database schema
- purpose-built for Dataworks
- evolved from GSAC prototype database schema
- fundamental tables: station, file, equip_config
- lookup tables aid maintenance, clean metadata
Python Modules
- downloadManager
- downloadIngester
- downloadExporter

These modules download files from remote receivers and manage interaction with GSAC through Dataworks database (ingest metadata), storing in local ftp server for distribution.
REGIONAL DATA CENTERS

COCONet Regional Data Centers running Dataworks:
- Geological Survey of Colombia (SGC)
- CIMH Barbados
- INETER, Nicaragua

Selected from proposals submitted to RFP

TLALOCNet Regional Data Center running Dataworks
Universidad de Guadalajara / UNAM

- Server with Dataworks installed by UNAVCO delivered to each center
- Training in Dataworks at UNAVCO
DATA & METADATA MIRRORING

Facilitates data sharing across regional networks with individual data centers having responsibility for part of the network.

Used at COCONet and TLALOCNet Regional Data Centers to Mirror UNAVCO holdings form COCONet or TLALOCNet sites.
DATA & METADATA FEDERATION

Facilitates data search across regional networks with individual data centers having responsibility for part of the network.

Used among US Data Centers (SOPAC, UNAVCO, CDDIS)
Dataworks web site and documentation

http://www.unavco.org/software/datamangement/dataworks/dataworks.html
Development is needed in these areas:

• Improve scalability
• Add tools for operator metadata management
• Add functionality related to realtime
• Expand the receivers and formats of the Download Module
• Add QC and formatting functionality
• Test in a Cloud VM environment and distribute as a VM image
• Sustainable support

Additional development will depend on securing additional resources