CDDIS Global Data Center Technical Report 2011

C. Noll

NASA Goddard Space Flight Center, Code 690.1
Greenbelt, MD 20771 USA
Carey.Noll@nasa.gov

1   Introduction

The CDDIS is NASA's data archive and information service supporting the international space geodesy community. For over 30 years, the CDDIS has provided continuous, long term, public access to the data (mainly GNSS-Global Navigation Satellite System, SLR-Satellite Laser Ranging, VLBI-Very Long Baseline Interferometry, and DORIS-Doppler Orbitography and Radiopositioning Integrated by Satellite) and products derived from these data required for a variety of science observations, including the determination of a global terrestrial reference frame and geodetic studies in plate tectonics, earthquake displacements, volcano monitoring, Earth orientation, and atmospheric angular momentum, among others. The specialized nature of the CDDIS lends itself well to enhancement to accommodate diverse data sets and user requirements. The CDDIS is one of NASA's Earth Observing System Data and Information System (EOSDIS) distributed data centers; EOSDIS data centers serve a diverse user community and are tasked to provide facilities to search and access science data and products.

The CDDIS serves as one of the primary data centers and core components for the geometric services established under the International Association of Geodesy (IAG), an organization that promotes scientific cooperation and research in geodesy on a global scale. The system has supported the International GNSS Service (IGS) as a global data center since 1992. The CDDIS activities within the IGS during 2011 are summarized below; this report also includes any recent changes or enhancements made to the CDDIS.

2   System Description

The CDDIS archive of IGS data and products are accessible worldwide through anonymous ftp. The CDDIS is located at NASA's Goddard Space Flight Center (GSFC) and is available to users 24 hours per day, seven days per week.

The CDDIS computer system consists of incoming, outgoing, and processing servers. All ftp and web access is performed on the outgoing server, which is equipped with a hot spare. Data centers, stations, and analysis centers push files to the CDDIS incoming server, which is also configured with a hot spare. Processing of incoming files for the on-line archive is performed in a separate environment that also includes a database server for managing metadata extracted from incoming data.

3   Archive Content

As a global data center for the IGS, the CDDIS is responsible for archiving and providing access to GNSS data from the global IGS network as well as the products derived from the analyses of these data in support of both operational and working group/pilot project activities. Approximately 6 Tbytes of the CDDIS archive are devoted to GNSS data (5.5 Tbytes), products (250 Gbytes), and ancillary information. All data and products are accessible through subdirectories of *ftp://cddis.gsfc.nasa.gov/gnss* (a symbolic link to *ftp://cddis.gsfc.nasa.gov/gps*).

3.1   GNSS Tracking Data

The user community has access to the on-line archive of GNSS data available through the global data center archives of the IGS. Over 50 operational and regional IGS data centers and station operators make data (observation, navigation, and meteorological) available in RINEX format to the CDDIS from selected receivers on a daily, hourly, and sub-hourly basis. The CDDIS also accesses the archives of the other three IGS global data centers, Scripps Institution of Oceanography (SIO) in California, the Institut Géographique National (IGN) in France, and the Korea Astronomy and Space Science Institute (KASI) to retrieve (or receive) data holdings not routinely transmitted to the CDDIS by an operational or regional data center. Table 1 below summarizes the types of GNSS data archived at the CDDIS.

Table 1a: GNSS Data Type Summary.

Data Type	Sample Rate	Data Format	Available On-line
Daily GNSS	30 sec.	RINEX and compact RINEX	Since 1992
Hourly GNSS	30 sec.	Compact RINEX	2+ years
High-rate GNSS	1 sec.	Compact RINEX	Since May 2001
Satellite GPS	10 sec.	Compact RINEX	Since 2002

Table 1b: GNSS Data Archive Summary for 2011.

Data Type	Avg. No. Sites/Day	Avg. Volume/Day	Total Volume/Year	Directory Location	Latency of Majority of Data
Daily GNSS	425	850 Mb	285 Gb	/gnss/data/daily	Within 1 hour
Hourly GNSS	260	300 Mb	95 Gb	/gnss/data/hourly	Within 10 minutes
High-rate GNSS	135	1500 Mb	460 Gb	/gnss/data/highrate	Within 10 minutes
LEO GPS	1	0.5 Mb	200 Mb	/gnss/data/satellite	Within 10 days

Data, in RINEX V2.10 or V2.11 format, from GPS and GPS+GLONASS receivers are archived within the GNSS directory structure */gnss/data*.

The CDDIS archives four major types/formats of GNSS data, all in RINEX format, as described in Table 1a. Daily RINEX data are quality-checked, summarized, and archived to public disk areas in subdirectories by year, day, and file type; the summary and inventory information are also loaded into an on-line database. Nearly 150K station days from 490 distinct GNSS receivers were archived at the CDDIS during 2011; a complete list of these sites can be found in the yearly summary reports at URL *ftp://cddis.gsfc.nasa.gov/reports/gnss/*.

Within minutes of receipt, the hourly GNSS files are archived to subdirectories by year, day, and hour. These data are retained on-line indefinitely; the daily files delivered at the end of the UTC day contain all data from these hourly files and thus can be used in lieu of the individual hourly files.

High-rate (typically 1-second sampling) GNSS data are archived in files containing fifteen minutes of data and in subdirectories by year, day, file type, and hour. Many of these data files are created from real-time streams.

The CDDIS generates a global broadcast ephemeris file on an hourly basis. This file is derived from the site-specific ephemeris data files for each day/hour. These files are appended to a single file that contains the orbit information for all GNSS satellites for the day up through that hour. This merged ephemeris data file is then copied to the day's subdirectory within the hourly data file system. Within 1-2 hours after the end of the UTC day, after sufficient station-specific navigation files have been submitted, this concatenation procedure is repeated to create the daily broadcast ephemeris file, using daily site-specific navigation files as input. The daily file is copied to the corresponding subdirectory under the daily file system. Users can thus download this single, daily (or hourly) file to obtain the unique navigation messages rather than downloading multiple broadcast ephemeris files from the individual stations.

The CDDIS continues to archive data from space-borne GPS receiver data from selected missions (e.g., SAC-C). The staff hopes to add data from other satellites such as Jason, GRACE, and ICESat.

3.2   IGS Products

The CDDIS routinely archives IGS operational products (daily, rapid, and ultra-rapid orbits and clocks, and weekly ERP and station positions) as well as products generated by IGS working groups and pilot projects (ionosphere, troposphere, real-time clocks). The CDDIS currently provides on-line access through anonymous ftp or the web to all IGS products generated since the start of the IGS Test Campaign in June 1992 in the file system */gnss/products*; products from GPS+GLONASS products are available through this filesystem. Products derived from GLONASS data only continued to be archived at the CDDIS in a directory structure within the file system */glonass/products*.

The CDDIS also continued to archive combined troposphere estimates in directories by GPS week. Global ionosphere maps of total electron content (TEC) from the IONEX AACs were archived in subdirectories by year and day of year. New ionosphere products include hourly and sub-hourly rapid products and predicted products. Real-time clock comparison products have been archived at the CDDIS in support of the IGS Real-Time Pilot Project since 2009. Table 2 below summarizes the GNSS products available through the CDDIS.

Table 2: GNSS Product Summary.

Product Type	Number of ACs/AACs	Volume	Directory
Orbits, clocks, ERP, positions	13+Combinations	750Mb/week	*/gnss/products/WWWW (GPS, GPS+GLONASS)* */glonass/products/WWWW (GLONASS only)*
Troposphere	Combination	2.5 Mb/day, 930Mb/year	*/gnss/products/troposphere/YYYY*
Ionosphere	4+Combination	4Mb/day, 1.5Gb/year	*/gnss/products/ionex/YYYY*
Real-time clocks	Combination	5.5Mb/week	*/gnss/products/rtpp/YYYY*

In 2011, the archive of products for the first IGS reprocessing campaign (repro1) was completed. GNSS data collected by the IGS network from 1994 through 2007 (GPS weeks 0730 through 1459) were re-analyzed by the IGS ACs in a consistent way using the latest models and methodology. The reprocessed files were submitted to the data centers for archive in a "repro1" directory structure (*/gnss/products/WWWW/repro1*); to maintain consistent access, the original set of IGS products continue to be archived in the weekly directories (*/gnss/products/WWWW*).

3.3   Supporting Information

Daily status files of GNSS data holdings, reflecting timeliness of the data delivered as well as statistics on number of data points, cycle slips, and multipath, continue to be generated by the CDDIS. By accessing these files, the user community can receive a quick look at a day's data availability and quality by viewing a single file. The daily status files are available through the web at URL *ftp://cddis.gsfc.nasa.gov/reports/gps/status*. The daily status files are also archived in the daily GNSS data directories.

Ancillary information to aid in the use of GNSS data and products are also accessible through the CDDIS. Weekly and yearly summaries of IGS tracking data (daily, hourly, and high-rate) archived at the CDDIS are generated on a routine basis. These summaries are accessible through the web at URL *ftp://cddis.gsfc.nasa.gov/reports/gps*. The CDDIS also maintains an archive of and indices to IGS Mail, Report, Station, and other IGS-related messages.

4   System Usage

Figure 1 summarizes the usage of the CDDIS for the retrieval of GNSS data and products in 2011. This figure illustrates the number and volume of GNSS files retrieved by the user community during 2011, categorized by type (daily, hourly, high-rate, products). Over 500 million files (40 Tbytes) were transferred in 2011, with an average of over 40 million files per month. Figure 2 illustrates the profile of

users accessing the CDDIS IGS archive during 2011. The majority of CDDIS users are from hosts in North America and Europe.
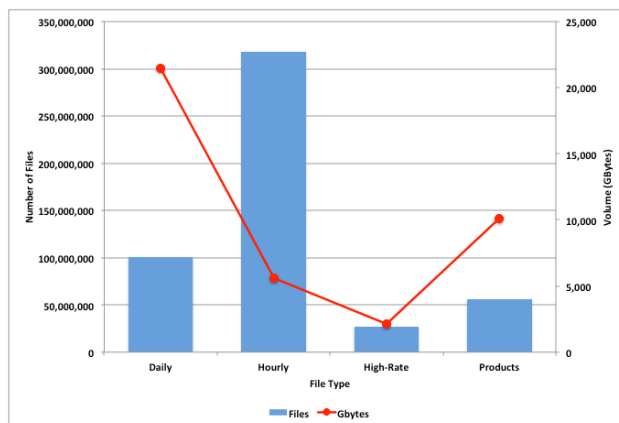

Figure 1: Number and volume of GNSS files transferred from the CDDIS in 2011.
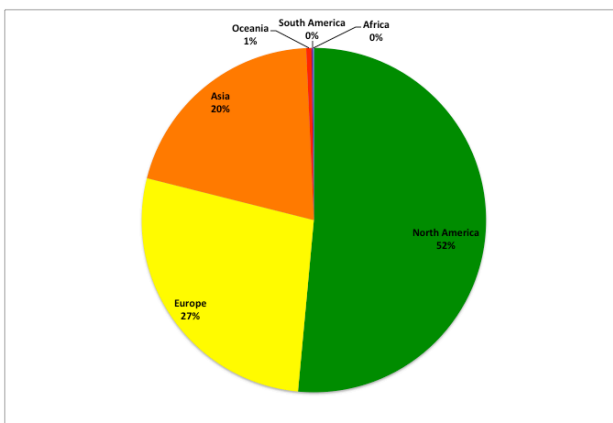

Figure 2: Geographic distribution of IGS users of the CDDIS in 2011.

5   Recent Developments

The CDDIS is cooperating in the development of Geodetic Seamless Archive Centers (GSAC) with colleagues at UNAVCO and SIO. The activity will provide web services to facilitate data discovery within and across participating archives. A prototype implementation of these GSAC web services at the CDDIS is under development and should be operational in mid-2012. In addition, the CDDIS is currently implementing modifications to the metadata extracted from incoming data and product files pushed to its archive. These enhancements will facilitate cross discipline data discovery by providing information about CDDIS archive holdings to other data portals such as Earth Observing System (EOS) Clearinghouse (ECHO) and integration into the Global Geodetic Observing System (GGOS) portal.

6   Publications

The CDDIS staff attended several conferences during 2011 and presented papers on or conducted demos of their activities within the IGS, including:

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.

Electronic versions of this poster and other publications can be accessed through the CDDIS on-line documentation page on the web at URL *http://cddis.gsfc.nasa.gov/reports.html*.

7   Future Plans

In 2011, the CDDIS staff procured new server hardware to further enhance the capabilities of the system and ensure a robust archive environment. The new system will be fully redundant with the primary and secondary/failover system located in different buildings on the GSFC campus. Each system will utilize a distributed functionality (incoming, outgoing, processing servers) and will be configured with a local backup system as well as a full backup system located in a third building at GSFC. The archive is equipped with a 32 Tbyte RAID storage system and is scaled to accommodate future growth. The new server environment will become operational in early 2012.

The CDDIS successfully submitted a proposal to the IGS Multi-GNSS Experiment (M-GEX) call for proposals for archive and distribution of data and products. During 2012 the CDDIS will expand its data archive and distribution service to include data from participating multi-GNSS receivers, products derived from the analysis of these data, and any required metadata for the experiment. The data will include

newly available signals (e.g., Galileo, QZS, and Compass). The CDDIS data ingest procedures will be modified to accommodate these new data sets, the majority of which will be archived in RINEX V3. This data format will require development of new software to extract metadata from incoming data files; the software package currently used for summarization and metadata extraction on RINEX V2 data, teqc, will not process data in RINEX V3 format.

The CDDIS is supporting the IGS Real-Time Pilot Project as a data center. During 2012, the CDDIS will implement an NTRIP Castor to transmit real-time data streams from stations to users. CDDIS will set up a dedicated server for this task. Possible future activities include capturing the streams for generation of 15-minute high-rate files for archive at the CDDIS.

8 Contact Information

To obtain more information about the CDDIS IGS archive of data and products, contact:

Ms. Carey E. Noll	Phone:	(301) 614-6542
Manager, CDDIS	Fax:	(301) 614-6015
Code 690.1	E-mail:	*Carey.Noll@nasa.gov*
NASA GSFC	WWW:	*http://cddis.gsfc.nasa.gov*
Greenbelt, MD 20771		*http://cddis.nasa.gov*

9 Acknowledgments

References

C. Noll, The Crustal Dynamics Data Information System: A resource to support scientific analysis using space geodesy, Advances in Space Research, Volume 45, Issue 12, 15 June 2010, Pages 1421-1440, ISSN 0273-1177, DOI: 10.1016/j.asr.2010.01.018.

C. Noll, Y. Bock, H. Habrich and A. Moore, "Development of data infrastructure to support scientific analysis for the International GNSS Service", Journal of Geodesy, Feb 2009, pages 309-325, DOI 10.1007/s00190-008-0245-6.