CDDIS Global Data Center Technical Report 2019

C. Noll (Carey.Noll@nasa.gov)
P. Michael (Patrick.Michael@nasa.gov)

NASA Goddard Space Flight Center, Code 61A
Greenbelt, MD 20771 USA
Carey.Noll@nasa.gov

1 Introduction

The Crustal Dynamics Data Information System (CDDIS) is NASA's active archive supporting the international space geodesy community. For over 35 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 scientific studies, 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 Active Archive Centers (DAACs) (see *https://earthdata.nasa.gov*); EOSDIS data centers serve a diverse user community and are tasked to provide facilities to search and access science data and products. The CDDIS is also a regular member of the International Council for Science (ICSU) World Data System (WDS, *https://www.icsu-wds.org*) and the Earth Science Information Partners (ESIP, *https://www.esipfed.org*).

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), in particular, the system has supported the International GNSS Service (IGS) as a global data center since 1992. The CDDIS activities within the IGS during 2019 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 globally accessible through anonymous ftp (address: *cddis.nasa.gov*) and through web-based archive access (*https://cddis.nasa.gov/archive*). 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.

2.1 Hardware Configuration

The CDDIS computer facility is fully redundant with primary and secondary/failover systems utilizing a virtual machine (VM) based system, configured with 100 TBytes of unified storage operating within the EOSDIS computer facility and network infrastructure. This system configuration provides reliable environment and network connectivity; a disaster recovery system is installed in a different location on the GSFC campus for rapid failover if required. Multiple, redundant 40G network switches are available to take full advantage of a high-performance network infrastructure by utilizing fully redundant network paths for all outgoing and incoming files. The use of the virtual machine technology provides multiple instance services for a load balancing configuration and allows for VM instances to be increased or decreased due to demand.

Furthermore, the VM technology allows for system maintenance (patching, upgrades, etc.) to proceed without any downtime or interruption to user access. The large, unified storage system will facilitate near real-time replication between its production and disaster recovery sites. The entire archive is also mirrored to traditional storage arrays for additional complete copies of the archive.

As shown in Figure 1, the providers of files for the CDDIS archive push their files (data, derived products, etc.) to the CDDIS ingest server, utilizing the Earthdata Login system for validating access. Incoming files are then handled by the processing system which performs file/content validation, quality control, and metrics extraction. Metadata and metrics (ingest/archive and distribution) information is pushed to the EOSDIS Common Metadata Repository (CMR) system. Content metadata, describing collections and granules, are available for access by a broad user community through the CMR. Valid files are then moved to the CDDIS archive for public access through the CDDIS ftp and web servers.

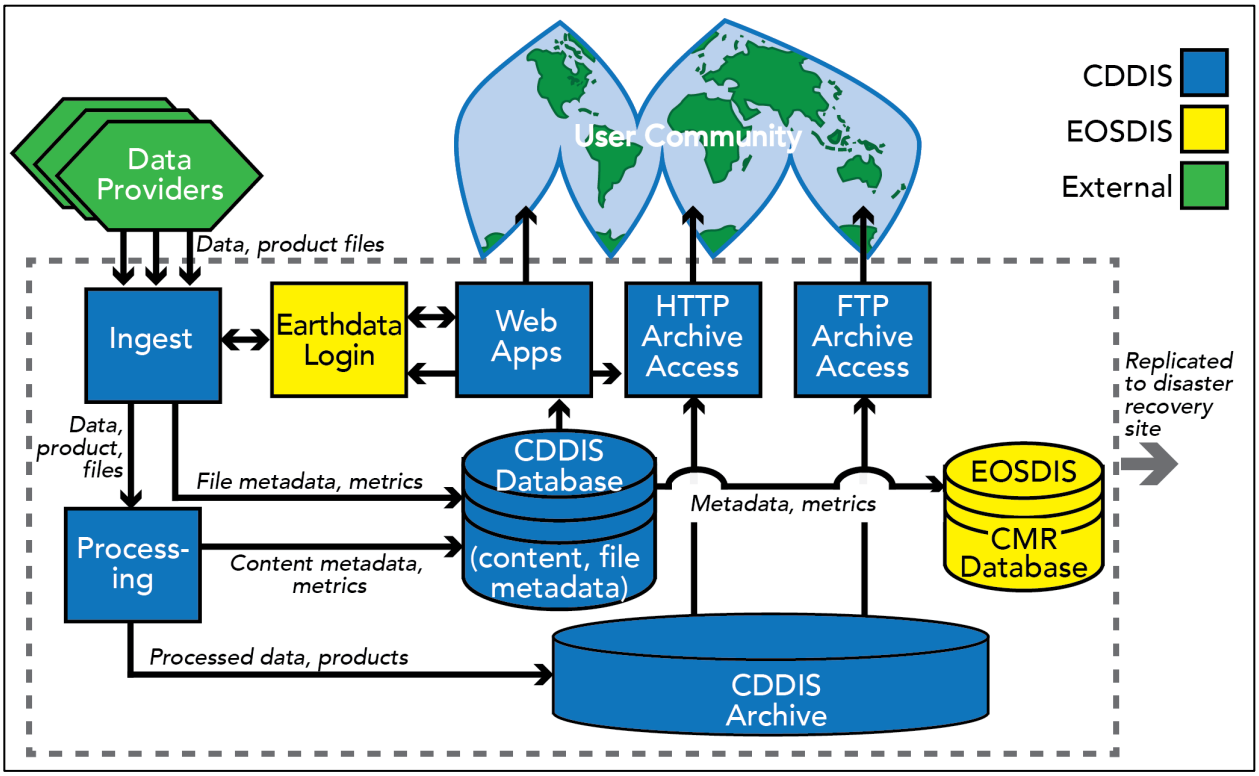

Figure 1: System architecture overview diagram for the CDDIS facility installation within the EOSDIS infrastructure.

2.2 Data Upload System

The CDDIS utilizes an https-based protocol method for delivery of files from suppliers of data and products. As stated previously, the CDDIS is one of NASA's EOSDIS DAACs and through EOSDIS, has access to a world-class user registration process, the EOSDIS Earthdata Login, with over 650K users in its system. The validation for file upload to CDDIS is performed through this Earthdata Login system. The file uploads can be performed through a webpage interface or a command line application that can perform an http "post" operation, which is more commonly used for scripting. This process allows data suppliers to authenticate through the Earthdata Login system and provide their files through https to CDDIS for ingest into the archive. More information

on the CDDIS file upload system, including an FAQ, is available at URL: *https://cddis.nasa.gov/About/CDDIS_File_Upload_Documentation.html.*

2.3 Ingest Software

The CDDIS file ingest processing system allows staff to check for errors in a more consistent fashion, regardless of data type or file provider; the automated system allows the staff to identify several error types, such as problems with file naming, compression, and content. Any errors are further categorized as fatal or warning errors and are tracked in the CDDIS database allowing staff to more easily monitor data processing. Fatal errors include logic errors (e.g., data with a future date), an empty file, or an unknown file name/structure. Files with fatal errors are not moved to the archive; they are placed in a "quarantine" location for further examination by operations staff. Warning errors are generally auto-corrected/handled and the file is then archived; these errors include a significantly older file, invalid compression, etc. The ingest software also performs routine checksums of and anti-virus scanning on all incoming files, extracts uniform file-level and content-level metadata, and consistently tracks file and content errors. The number of errors detected in incoming files have been reduced significantly due to staff's outreach efforts with data suppliers to correct a large majority of errors. These efforts have resulted in an improved, more reliable CDDIS archive. Since GNSS data accounts for a majority of the incoming files to CDDIS, the staff has developed a guidelines document for data providers (*https://cddis.nasa.gov/docs/2017/GNSSDataStandards.pdf*).

3 Archive Contents

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. The CDDIS archive is approximately 27 TBytes in size (over 260 million files) of which over 95% is devoted to GNSS data (30 TBytes) and GNSS products (1.7 TBytes). All these GNSS data and products are accessible through subdirectories of *ftp://cddis.nasa.gov/gnss* and *https://cddis.nasa.gov/archive/gnss*.

3.1 GNSS Data

3.1.1 Main Data Archive

The user community has access to GNSS data available through the on-line global data center archives of the IGS. Nearly 50 operational and regional IGS data centers and station operators make data available in RINEX format to the CDDIS from receivers on a daily, hourly, and sub-hourly basis. The CDDIS also accesses the archives of other IGS global data centers (GDCs) 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 IGS GNSS data sets available in the CDDIS in the operational, non-campaign directories of the GNSS archive.

The main GNSS data archive (*https://cddis.nasa.gov/archive/gnss/data*) at the CDDIS contains GPS and GPS+GLONASS data in RINEX V2 format and multi-GNSS data in RINEX V3 format. Since January 2016, RINEX V3 data, using the V3 "long" filename specification, have been made available here along with the RINEX V2 data. The availability of RINEX V3 data into the operational, main archives at the IGS GDCs (and detailed in the "RINEX V3 Transition Plan") addressed a key recommendation from the IGS 2014 Workshop: "one network one archive" and provided for the better integration of multi-GNSS data into the entire IGS infrastructure. Starting

in 2015, stations began submitting RINEX V3 data using the format's "long" filename specification. The transition plan specified that RINEX V3 data from IGS network sites using the V3 filename structure should be archived in the same directories as the RINEX V2 data. Therefore, starting on January 01, 2016, all daily, hourly, and high-rate data submitted to the CDDIS in RINEX V3 format and using the long, V3 filename specification have been archived in the same directories as the RINEX V2 data (which use the 8.3.Z filename for daily and hourly files and the 10.3.Z filename format for high-rate files). In addition, these RINEX V3 files are compressed in gzip (.gz) format; files in RINEX V2 format continue to use UNIX compression (*.Z*). These data in RINEX V3 format include all available multi-GNSS signals (e.g., Galileo, QZSS, SBAS, BeiDou, and IRNSS) in addition to GPS and GLONASS. Figure 2 shows the network of IGS sites providing daily data in RINEX V2 and/or V3 formats.

Table 1: GNSS Data Type Summary.

Data Type	Sample Rate	Data Format	Available On-line
Daily GNSS	30 sec.	RINEX V2	Since 1992
Daily GNSS	30 sec.	RINEX V3	Since 2016
Hourly GNSS	30 sec.	RINEX V2	Since 2005
Hourly GNSS	30 sec.	RINEX V3	Since 2016
High-rate GNSS	1 sec.	RINEX V2	Since 2001
High-rate GNSS	1 sec.	RINEX V3	Since 2016
Satellite GPS	10 sec.	RINEX V2	2002-2012

The CDDIS archives three major types/formats of GNSS data, daily, hourly, and high-rate sub-hourly, all in RINEX format, as described in Table 1; the network distribution of submitted files is shown in Figure 3. Over 287K daily station days from 602 distinct GNSS receivers were archived at the CDDIS during 2019; of these sites, 294 sites supplied both RINEX V2 and V3 data (see Table 2). A complete list of daily, hourly, and high-rate sites archived in the CDDIS can be found in the yearly summary reports at URL *https://cddis.nasa.gov/reports/gnss/*. All incoming files for the CDDIS archive are now checked for conformance to basic rules, such as valid file type, non-empty file, uses correct compression, consistency between filename and contents, uses correct file naming conventions, and other logic checks. After incoming files pass these initial checks, content metadata are extracted and the files undergo further processing based on data type and format.

Table 2: GNSS Data Archive Summary for 2019.

Data Type	Number of Sites				Vol.	# files	Directory
	V2	V3	V2+V3	Unique			
Daily	552	344	294	602	892GB	1.3M	*/gnss/data/daily*
Hourly	378	255	222	411	948GB	15.1M	*/gnss/data/hourly*
High-rate	270	101	64	307	3,900GB	20.0M	*/gnss/data/highrate*

Daily RINEX V2 data are quality-checked, summarized (using UNAVCO's teqc software), 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. However, this data quality information, generated for data holdings in RINEX V2 format, is not available through the software used by CDDIS to summarize data in RINEX V3 format. CDDIS continues to investigate and evaluate software capable of providing data summary/QC information for RINEX V3 data.

Within minutes of receipt (typically less than 30 seconds), the hourly GNSS files are archived to subdirectories by year, day, and hour. Although these data are retained on-line, 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. As seen in Table 2, a total of 411 unique hourly sites (over

15 million files) were archived during 2019; 222 hourly sites provided data in both RINEX V2 and V3 formats.

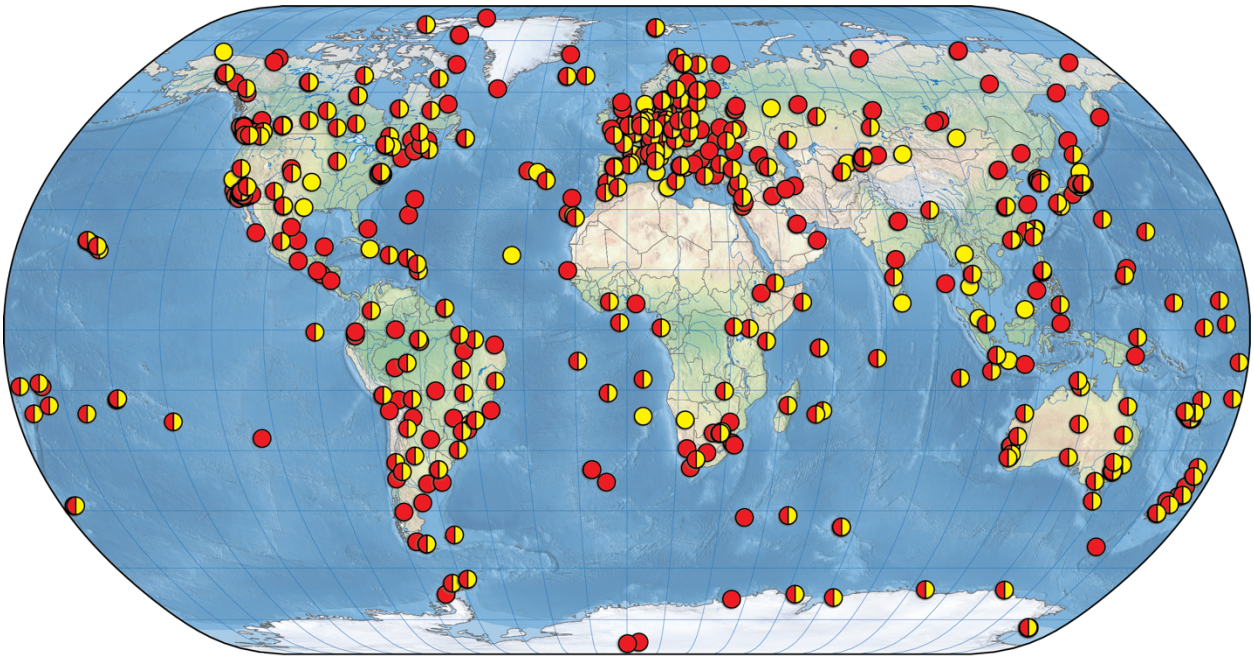

Figure 2: The main, operational archive at CDDIS now includes data in RINEX V2 format using the 8.3.Z filename specification (red) and RINEX V3 format using the V3 filename specification (yellow); sites providing both RINEX V2 and V3 formatted data are shown with the red+yellow icon.z

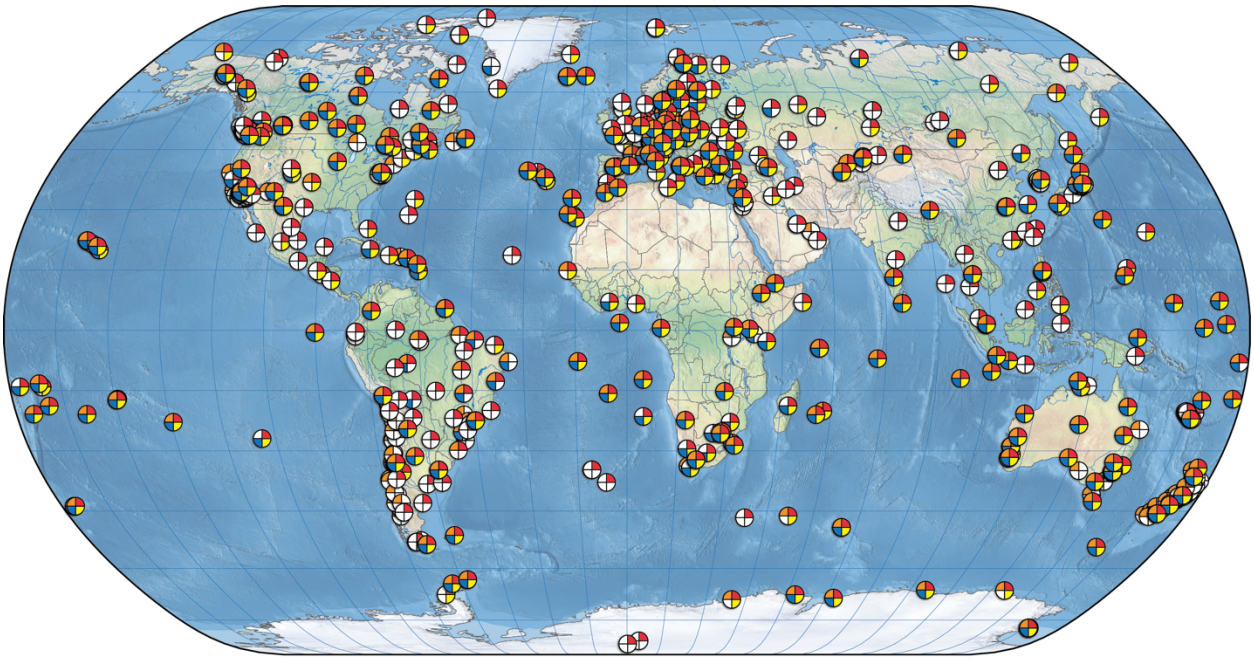

Figure 3: CDDIS GNSS archive includes data in daily (red), hourly (yellow), sub-hourly (blue), and/or real-time (orange) increments. Hourly, sub-hourly, and real-time data allow analysts to generate products for applications needing more frequent updates.

High-rate (one-second sampling rate) GNSS data are made available 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. As shown in Table 2, data from 307 unique high-rate sites (over 20 million files) were archived in the CDDIS in 2019; 64 high-rate sites provided data in both RINEX V2 and V3 formats.

3.1.2 Broadcast Navigation Files

The CDDIS generates global RINEX V2 broadcast ephemeris files (for both GPS and GLONASS) on a daily and hourly basis. The hourly concatenated broadcast ephemeris files are derived from the site-specific ephemeris data files for each hour and are appended to a single file that contains the orbit information for all GPS and GLONASS satellites for the day up through that hour. The merged ephemeris data files, named *hourDDD0.YYn.Z*, are 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 files (both GPS and GLONASS), using daily site-specific navigation files as input. These daily RINEX V2 broadcast ephemeris files, named *brdcDDD0.YYn.Z* and *brdcDDD0.YYg.Z*, are then copied to the corresponding year/day nav file subdirectory as well as the yearly *brdc* subdirectory (*/gnss/data/daily/YYYY/brdc*).

The CDDIS also generates daily RINEX V3 concatenated broadcast ephemeris files. The files are archived in the yearly *brdc* subdirectory (*https://cddis.nasa.gov/archive/gnss/data/daily/YYYY/brdc*) with a filename of the form *BRDC00IGS_R_yyyydddhhmm_01D_MN.rnx.gz*. The procedure for generating these files is similar to the V2 procedure in that site-specific, mixed V3 ephemeris data files are merged into to a single file that contains the orbit information for all GNSS satellites for the day. The chair of the IGS Infrastructure Committee provided the software that CDDIS staff uses to create these files. Users can thus download these single, daily (or hourly) files (in both RINEX V2 and V3 formats) to obtain the unique navigation messages rather than downloading multiple broadcast ephemeris files from the individual stations.

The CDDIS also archives a merged, multi-GNSS broadcast ephemeris file containing GPS, GLONASS, Galileo, BeiDou, QZSS, and SBAS ephemerides. This file, generate by colleagues at the Technical University in Munich (TUM) and Deutsches Zentrum f˙ur Luft- und Raumfahrt (DLR) from real-time streams, contains all the unique broadcast navigation messages for the day. The file, named *BRDM00DLR_S_YYYYDDD0000_01D_MN.rnx.gz*, is stored in daily subdirectories within the archive (*/gnss/data//daily/YYYY/DDD/YYp*) and in a yearly top level subdirectory (*/gnss/data/daily/YYYY/brdc)*. In addition, the TUM/DLR team provides a merged GPS/QZSS LNAV and CNAV navigation file generated from real-time streams; these files use the naming convention *BRDX00DLR_S_YYYYDDD0000_01D_MN.rnx.gz*.

For the near term, the CDDIS continues to archive a daily merged multi-GNSS broadcast ephemeris file and GPS/QZSS CNAV file using the RINEX V2 naming convention and archived in the MGEX campaign directories: (*/gnss/data/campaign/mgex/daily/rinex3/YYYY/DDD/YYp /brdmDDD0.YYp.Z* and */gnss/data/campaign/mgex/daily/rinex3/YYYY/cnav/brdxDDD0.YYx.Z* respectively. The archive of these files will discontinue in the near future as the IGS moves to complete the integration of the campaign directory contents into the main GNSS data archive.

3.1.3 Supporting Information

The CDDIS generates and updates "status" files, (*/gnss/data/daily/YYYY/DDD/YYDDD.status* for RINEX V2 data and *YYDDD.V3status* for RINEX V3 data) that summarize the holdings of daily

GNSS data. These status files of CDDIS GNSS data holdings reflect timeliness of the data delivered as well as statistics on number of data points, cycle slips, and multipath (for RINEX V2 data). The user community can thus view a snapshot of data availability and quality by checking the contents of such a summary file.

3.1.2 RINEX V3 (MGEX) Campaign Archive

During 2019, very little data in RINEX V3 format using the 8.3.Z filename specification were archived in the Multi-GNSS Experiment (MGEX) campaign directory structure at CDDIS *(/gnss/campaign/mgex/data*). The majority of data in RINEX V3 format utilize the "long" RINEX V3 naming convention with gzip compression and are integrated in the operational directory structure (*/gnss/data/daily, /gnss/data/hourly, /gnss/data/highrate*).

3.2 IGS Products

The CDDIS routinely archives IGS operational products (daily, rapid, and ultra-rapid orbits and clocks, ERP, and station positions) as well as products generated by IGS working groups and pilot projects (ionosphere, troposphere, real-time, MGEX). Table 3 below summarizes the GNSS products available through the CDDIS. The CDDIS currently provides on-line access 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 continue to be archived at the CDDIS in a directory structure within the file system */glonass/products*.

Table 3: GNSS Product Summary for 2019.

Product Type	Number of ACs/AACs	Volume	Directory
Orbits, clocks, ERP, positions	14+Combinations	3.5 GB/week	*/gnss/products/WWWW (GPS, GPS+GLONASS)* */glonass/products/WWWW (GLONASS only)*
Troposphere	Combination	3.4 MB/day, 1.2 GB/year	*/gnss/products/troposphere/YYYY*
Ionosphere	7+Combination	5 MB/day, 1.8 GB/year	*/gnss/products/ionosphere/YYYY*
Real-time	Combinations	28 MB/week	*/gnss/products/rtpp/WWWW*
MGEX	7	500 MB/week	*/gnss/products/mgex/WWWW*

Note: *WWWW*=4-digit GPS week number; *YYYY*=4-digit year

The CDDIS also continues to archive combined troposphere estimates in directories by year and day of year. Global ionosphere maps of total electron content (TEC) from the IONEX AACs are also archived in subdirectories by year and day of year. Real-time clock comparison products have been archived at the CDDIS in support of the IGS Real-Time Pilot Project, and current IGS Real-Time Service, since 2009.

Seven AACs (CODE, GFZ, GRGS, JAXA, TUM, SHAO, and Wuhan) generated weekly products (orbits, ERP, clocks, and others) in support of MGEX; these AACs now utilize the "long" filename convention for their products. These files are archived at the CDDIS in the MGEX campaign subdirectory by GPS week (*/gnss/products/mgex/WWWW*).

Colleagues at DLR and the Chinese Academy of Sciences (CAS) provide a differential code bias (DCB) products for the MGEX campaign. This product is derived from GPS, GLONASS, Galileo, and BeiDou ionosphere-corrected pseudorange differences and is available in the bias SINEX format. DLR has provided quarterly DCB files containing daily and weekly satellite and station biases since 2013 in CDDIS directory */gnss/products/biases*; CAS provides files on a daily basis. Additional details on the DCB product are available in IGSMail message 6868 sent in February

2015 and message 7173 sent in October 2015. Both products use the RINEX V3 file naming convention.

3.3 Real-Time Activities

The CDDIS real-time caster has been operational since early 2015 in support of the IGS Real-Time Service (IGS RTS). By the end of 2019, the CDDIS caster broadcasts 39 product and more than 600 data streams in real-time. The caster runs the NTRIP (Network Transport of RTCM via Internet Protocol) format. Figure 4 shows the distribution of stations providing real-time streams to the CDDIS caster by source. The CDDIS caster accesses streams from several regional casters as shown in Table 4.

Table 4: CDDIS Caster Stream Availability.

Acronym	Agency/Country	Approximate Number of Streams*
Data		
ASI	Italian Space Agency (Italy)	9
BKG	Bundesamt für Kartographie und Geodäsie (Germany)	106
CNES REGINA	Centre National d'Etudes Spatiales Reseau GNSS pour l'IGS et la Navigation (France)	27
CNS	Centro Sismológico Nacional, University of Chile (CNS, Chile)	62
FinnRef	National Land Survey of Finland (Finland)	1
GA	Geoscience Australia (Australia)	88
GDGPS	Global Differential GPS, Jet Propulsion Laboratory (USA)	98
GFZ	GeoForschungsZentrum (Germany)	21
GSI	Geospatial Information Authority of Japan (Japan)	6
IBGE	Instituto Brasileiro de Geografia e Estatística (Brazil)	30
ICGC	Institut Cartográfic i Geològic de Catalunya (Spain)	2
IGN	Institut Geographique National (France)	20
LINZ	Land Information New Zealand (New Zealand)	98
NRCan	Natural Resources Canada (Canada)	21
ROB	Royal Observatory Belgium (Belgium)	2
TrigNet	TrigNet (South Africa)	4
UNAVCO	UNAVCO (USA)	16
Total Data:		611
Product	Multiple	39
Total Streams		650

Note: *Data stream count includes streams using both 5 and 10 character mount point naming convention in most cases.

The CDDIS caster serves as the third primary caster for the IGS RTS, thus providing a more robust topology with redundancy and increased reliability for the service. User registration, however, for all three casters is unique; therefore, current users of the casters located at the IGS/UCAR and BKG are required to register through the CDDIS registration process in order to use the CDDIS caster. By the end of 2019, over 330 users from 47 countries have registered to use the CDDIS caster; approximately 65 users were added in 2019. More information about the CDDIS caster is available at *https://cddis.nasa.gov/Data_and_Derived_Products/Data_caster_description.html*.

The CDDIS staff updated the caster to provide new 10 character mount point names as per direction of the IGS Real-Time Working Group (RTWG). The expanded mount point names align with the RINEX V3 naming convention utilized within the IGS to accommodate multi-constellation data.

As stated previously, the CDDIS utilizes the EOSDIS Earthdata Login, for authenticating file uploads to its incoming file server. Since the NTRIP-native registration/access software was not compatible with NASA policies, the CDDIS developed software to interface the caster and the Earthdata Login within a generic Lightweight Directory Access Protocol (LDAP) framework.

Access to the CDDIS caster requires that new users complete two actions: 1) an Earthdata Login registration and 2) a CDDIS caster information form, providing the user's email, institution, and details on their planned use of the real-time data. Following completion, the information is submitted to CDDIS staff for the final steps to authorize access to the CDDIS caster; this access is typically available to the user within 24 hours.

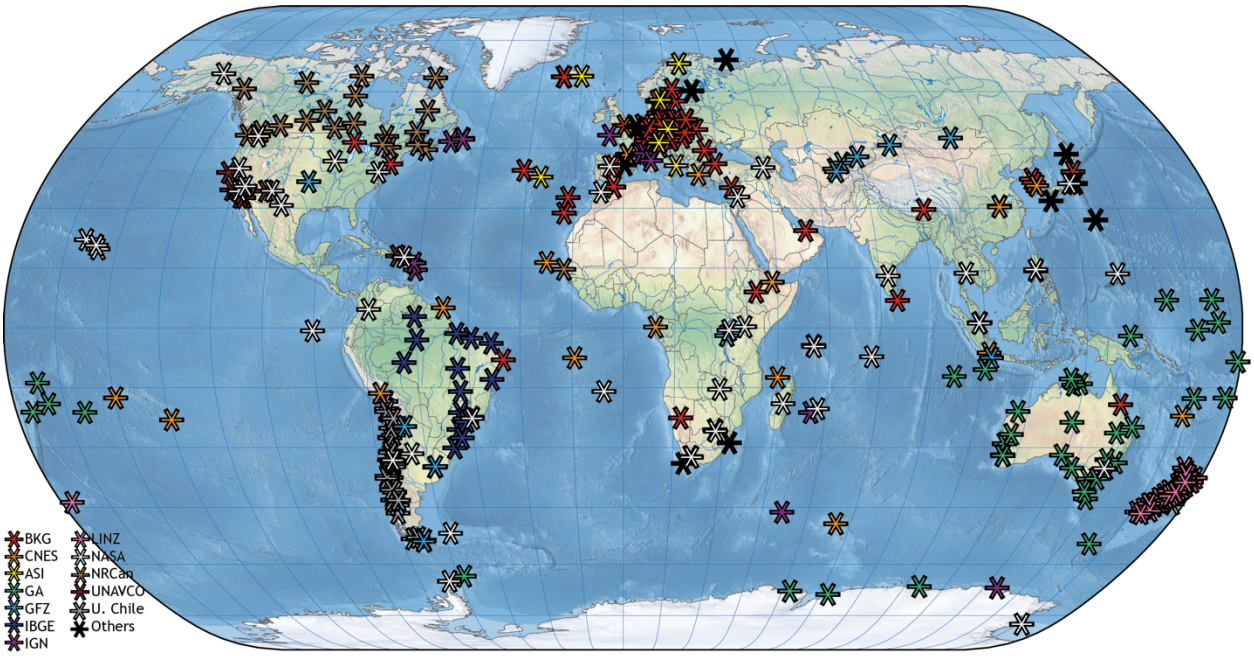

Figure 4: CDDIS is operationally supporting the dissemination of data from over 650 real-time GNSS sites as well as near real-time products derived from these data.

3.4 Supporting Information

Daily status files of GNSS data holdings, show timeliness of data receipt and statistics on number of data points, cycle slips, and multipath, continue to be generated by the CDDIS for RINEX V2 data; status files, with limited information, summarizing RINEX V3 data holdings are also available. These files are archived in the daily GNSS data directories and available through at URL *https://cddis.nasa.gov/reports/gnss/status*..

Other available ancillary information at CDDIS include daily, weekly, and yearly summaries of IGS tracking data (daily, hourly, and high-rate, in both RINEX V2 and V2 formats) archived at the CDDIS are generated on a routine basis. These summaries are accessible through the web at URL *https://cddis.nasa.gov/reports/gnss.* The CDDIS also maintains an archive of and indices to IGS Mail, Report, Station, and other IGS-related messages.

4 System Usage

Figure 5 shows the usage of the CDDIS, summarizing the retrieval of GNSS data and products from the online archive in 2019. This figure illustrates the number and volume of GNSS files retrieved by the user community during the past year, categorized by type (daily, hourly, high-rate, products). Over 1.5 billion files (nearly 286 TBytes) were transferred in 2019.

As for real-time system usage, an average of 15 users consistently accessed the CDDIS real-time caster on a daily basis in 2019, with on average 9,000 stream connections to over 500 streams

through a day. Figure 6 summarizes the primary applications the community uses from CDDIS caster streams; this information is provided by users during the caster registration process.

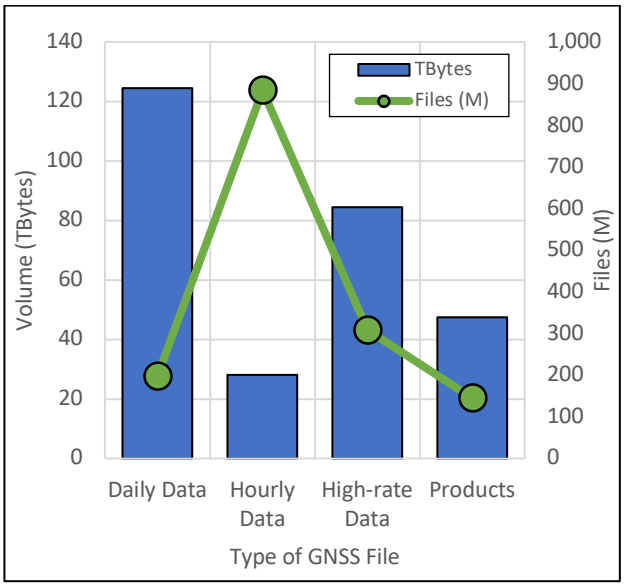

Figure 5: Number and volume of GNSS files downloaded from the CDDIS in 2019.

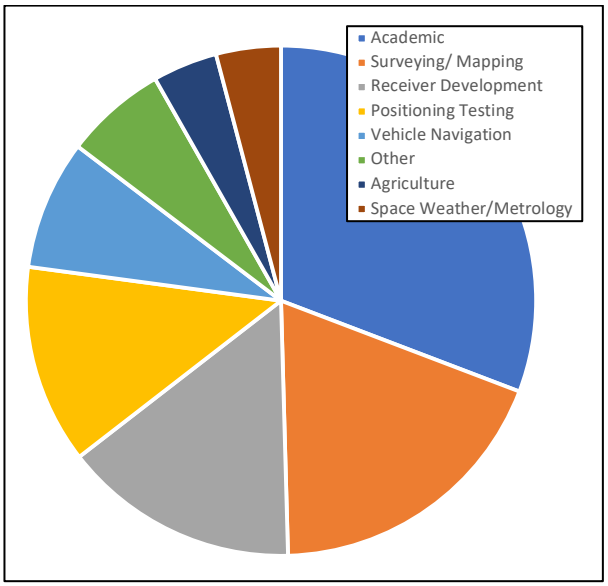

Figure 6: Primary applications supported by CDDIS real-time caster streams.

5 Recent Developments

5.1 Updates to Archive Access

The CDDIS has a large international user community; nearly 300K unique hosts accessed the system in 2019. Today, users access the CDDIS archive through anonymous ftp and https. The ftp protocol allows users to easily automate file downloads but has problems from a system/security standpoint. Starting in 2018, as per U.S. Government and NASA directives, the CDDIS began to move users away from reliance on anonymous ftp. Despite this requirement, the CDDIS staff is committed to ensuring continued, easy, open access to its archive. Access to data in the CDDIS archive using anonymous ftp will continue until October 2020; users are strongly encouraged to implement procedures to use the https (address: *https://cddis.nasa.gov/archive*) and ftp-ssl (address: *gdc.cddis.eosdis.nasa.gov*) capabilities as soon as possible.

The major reason for changing the archive access methods at CDDIS is system security and data integrity; ftp with its clear text username and password and lack of encryption, is just not acceptable in the current internet environment. The ftp protocol also has the disadvantage of being a two-port protocol that can result in connectivity problems (e.g., with firewall, router/switches, etc.). Unfortunately, proper network configuration is too often not the case and, in most instances, outside the control of CDDIS or the data provider to fix.

The CDDIS has configured servers to utilize protocols that allow two new methods for system access: https (browser and command line) and ftp-ssl (command line). The https protocol is as efficient as ftp transfer without the firewall/router issues of ftp; unlike ftp, https is a one-port protocol with fewer issues with downloads. The access to the CDDIS archive through both methods continues to present the same structure as that provided through anonymous ftp.

Archive access through the https protocol utilizes the same NASA single sign-on system, the EOSDIS Earthdata Login utility, as is used for the file upload and real-time caster user

authentication. Before using the https protocol to access the CDDIS archive, new users must initially access the webpage, *https://cddis.nasa.gov/archive*, to establish an account and authorize access; this page will then redirect the user to the Earthdata Login page. Earthdata Login allows users to easily search and access the full breadth of all twelve EOSDIS DAAC archives. Earthdata Login also allows CDDIS staff to know our users better, which will then allow us to improve CDDIS capabilities.

Once an account is established, the user has all permissions required to access the CDDIS archive using the https protocol, via a web browser or via a command line interface (e.g., through cURL or Wget) to script and automate file retrieval.

In addition, ftp-ssl access, an extension of ftp using TLS (transport layer security), can be used for scripting downloads from the CDDIS archive. The ftp-ssl is the option most similar to standard anonymous ftp. As with https, ftp-ssl will satisfy U.S. Government/NASA requirements for encryption.

Examples on using these protocols, including help with the cURL and Wget commands, are available on the CDDIS website; users are encouraged to consult the available documentation at: *https://cddis.nasa.gov/About/CDDIS_File_Download_Documentation.html* and examples documentation at: *https://cddis.nasa.gov/Data_and_Derived_Products/CDDIS_Archive_Access.html*. Various presentations on these updates to the CDDIS archive access are also available (see Section 7 below and *https://cddis.nasa.gov/Publications/Presentations.html*).

5.2 Metadata Improvements

The CDDIS continues to make modifications to the metadata extracted from incoming data and product files pushed to its archive and implemented these changes in the new file ingest software system. These enhancements have facilitated cross discipline data discovery by providing information about CDDIS archive holdings to other data portals such as the EOSDIS Earthdata search client and future integration into the GGOS portal. The staff continues work on a metadata evolution effort, re-designing the metadata extracted from incoming data and adding information that will better support EOSDIS applications such as its search client and the metrics collection effort. The CDDIS is also participating in GGOS metadata efforts within the Bureau of Networks and Observations.

The CDDIS continues to implement Digital Object Identifiers (DOIs) to select IGS data sets (GNSS data and products). DOIs can provide easier access to CDDIS data holdings and allow researchers to cite these data holdings in publications. Landing pages are available for each of the DOIs created for CDDIS data products and linked to description pages on the CDDIS website; an example of a typical DOI description (or landing) page, for daily Hatanaka-compressed GNSS data files, can be viewed at: *https:/cddis.nasa.gov/Data_and_Derived_Products/GNSS/daily_gnss_d.html*. DOIs have now been assigned to the majority of GNSS data and product sets archived at CDDIS.

5.3 Real-time Caster Updates

By the end of 2019, the CDDIS real-time caster was configured to stream data from over 600 GNSS data mount points and 39 product streams. The caster added over 150 10-character mount point names as per recommendations from the IGS Real Time Working Group (RTWG). These

streams, along with new product streams using the 10-character naming, will eventually replace the corresponding 5-character mount point names by August 2020.

6 Future Plans

6.1 Archive Access

As discussed in Section 5 above, the CDDIS cannot continue to support of non-encrypted anonymous ftp access to its archive; access to the archive through https and ftp-ssl have already been implemented. The schedule for terminating anonymous ftp access is as follows:

- July 01, 2020 – all anonymous ftp service will be bandwidth limited to 1MB/s
- August 10, 2020 – all anonymous ftp service will be bandwidth limited to 250KB/s
- October 01, 2020 – all anonymous ftp service will be bandwidth limited to 50KB/s
- October 31, 2020 – all anonymous ftp service will be permanently discontinued at CDDIS

The staff is also testing providing a WebDAV (Web Distributed Authoring and Versioning) interface to provide another method for accessing CDDIS archive. If feasible for CDDIS, this interface method would allow users to securely connect to the CDDIS archive as if it were a local drive on their computer.

6.2 RINEX V3 Data and Reprocessing Older GNSS Data

The CDDIS will continue to coordinate with the Infrastructure Committee and other IGS data centers to implement steps outlined in the RINEX V3 transition plan to complete the incorporation of RINEX V3 data into the operational GNSS data directory structure. The CDDIS began this process with multi-GNSS, RINEX V3 data from January 2016 onwards; the CDDIS will continue these efforts by integrating RINEX V3 multi-GNSS data from years prior to 2016 into the IGS operational archives. MGEX campaign directories will continue to be maintained during this transition to the operational directory archive. Furthermore, the CDDIS staff will continue to test software to copy RINEX V3 data (using the older filename format) into files with RINEX V3 filenames as well as QC RINEX V3 data and files and incorporate the software into operational procedures.

In mid-2016 CDDIS installed a new ingest processing system (see section 2.3) providing more extensive quality control on and metadata extraction of incoming files. The CDDIS staff plans to use this new software to validate the older GNSS archive (daily starting in 1992, hourly starting with 2005, and high-rate starting in 2001); this process will ensure that these historic files are valid and accurately archived for the user community. The additional metadata will also help the staff to better manage the CDDIS GNSS data holdings, provide improved metrics on data availability, and extensive data search capability for the EOSDIS Earthdata Search utility.

6.3 Real-Time Activities

The CDDIS will add real-time data and product streams to its operational caster in support of the IGS Real-Time Service. The CDDIS continues to review the implementation of software to capture real-time streams for generation of 15-minute high-rate files for archive. This capability requires further testing and coordination with the IGS Infrastructure Committee. The staff is also developing software to provide metrics on usage of the CDDIS caster.

CDDIS staff members continue to investigate the use of DLR's ntripchecker software for updating the caster source table in real-time, maintaining stream record consistency among the CDDIS and

regional casters. The staff is also working on developing scripts to monitor and report interruptions and outages in broadcast streams.

6.4 High-rate Archive Modifications

CDDIS staff put forward a recommendation at the 2018 IGS Workshop to consolidate the sub-hourly high-rate data files into a tar archive, one file per site per day. At this time, each site supplies up to 96 files per day; the bundling of the files into a single daily site-specific tar file would simplify downloads for the user as well as streamline the directory structure at the data centers. CDDIS plans to begin these modifications to the high-rate data archive starting with 2001 and work toward the present; the data from the current year will remain in the standard, submitted 15-minute file format. The CDDIS staff will coordinate with the IGS Infrastructure Committee, users, and data centers on moving forward with this recommendation.

6.5 System Upgrades

The CDDIS has procured new systems, storage, and network hardware for its next hardware refresh. Staff members have begun the integration of this next system; plans are to have the upgraded system installed by mid-2020. The server and network hardware will remain within the same physical infrastructure as today's system, thus providing a reliable hosting environment with fully redundant networking paths and backup sites.

6.6 Repro3 Support

The CDDIS provided support through the upload of files from the ACs and online archive of the IGS repro1 and repro2 campaigns (*/gnss/products/WWWW/repro[1,2]* and */gnss/products/repro[1,2]/WWWW*). As the ACs work on solutions for the next reprocessing campaign (repro3), CDDIS will try to provide archive support; initial size requirements, however, may not be sufficient to allow for a complete upload of files. The CDDIS staff continues to work with the ACC and ACs on possible solutions.

7 Publications

The CDDIS staff attended several conferences during 2019 and presented, or contributed to, papers on their activities within the IGS, including:

M. Pearlman et al. GGOS: Current Activities and Plans of the Bureau of Networks and Observations (poster), presented at the EGU General Assembly, Vienna, Austria, April 07-12, 2019.
https://ilrs.gsfc.nasa.gov/docs/2019/GGOSEGU201904.pdf

C. Noll, P. Michael. The Crustal Dynamics Data Information System: NASA's Active Archive of Geodetic Observations Supporting Research in Understanding our Dynamic Earth, presented at the 27th IUGG General Assembly, Montreal Canada, July 08-18, 2019.
https://cddis.nasa.gov/docs/2019/CDDISposter_IUGG2019_v3.pdf

C. Noll. Role and function of the SLR Data Centers, presented at the First SLR School, Stuttgart, Germany, October 20, 2019.
https://cddis.nasa.gov/2019_Technical_Workshop/docs/2019/presentations/SLRschool/Session2/SLRschool_session2_Noll_presentation.pdf

C. Noll, M. Pearlman. ILRS: Recent Developments, presented at the 2019 ILRS Technical Workshop, Stuttgart, Germany, October 21-25, 2019.

https://cddis.nasa.gov/2019_Technical_Workshop/docs/2019/presentations/Session1/session1_Noll_presentation.pdf

T. Otsubo, Y. Aoyama, A. Hattori, K. Doi, M. Pearlman, C. Noll. The Final Frontier for Satellite Laser Ranging: Antarctica, , presented at theTenth Symposium on Polar Science, Tokyo, Japan, December 03-05, 2019.
https://ilrs.gsfc.nasa.gov/docs/2019/PolarSci10-otsubo-finalfrontier-s.pdf

P. Michael, C. Noll. NASA CDDIS: Supporting Global Geodetic and Geophysical Research and Applications, presented at the 2019 Fall AGU meeting, San Francisco, CA, USA, December 09-13, 2019.
https://cddis.nasa.gov/docs/2019/CDDISposterAGU201912v0.pdf

S. Blevins, N. Pollack, P. Michael, C. Noll. Enhancements to the GNSS Real-time System at CDDIS, presented at the 2019 Fall AGU meeting, San Francisco, CA, USA, December 09-13, 2019.
https://cddis.nasa.gov/docs/2019/SMBlevins_AGU2019_final.pdf

Electronic versions of these and other publications can be accessed through the CDDIS on-line documentation page on the web at URL *https:/cddis.nasa.gov/Publications/Presentations.html*.

8 Contact Information

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

Patrick Michael	Phone:	(301) 614-5370
Manager, CDDIS	Fax:	(301) 614-6015
Code 61A	E-mail:	*Patrick.Michael@nasa.gov*
NASA GSFC	WWW:	*https://cddis.nasa.gov*
Greenbelt, MD 20771		

General questions on the CDDIS, archive contents, and/or help using the system, should be directed to the user support staff at: *support-cddis@earthdata.nasa.gov*.

9 Acknowledgments

Funding for the CDDIS, and its support of the IAG, IGS, and other services, is provided by NASA through the Earth Science Data and Information System (ESDIS) project, which manages the EOSDIS science systems and DAACs.

The authors would like to acknowledge the entire CDDIS staff. The success of the CDDIS and its recognition in the many international programs supported by the system can be directly attributed to the continued dedicated, consistent, professional, and timely support of its staff.

10 Additional Resources

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.

"Access NASA Earth Science Data", from Earthdata website, *https://earthdata.nasa.gov*.

"IGS RINEX 3 Transition Plan v3.0", from IGS website, *http://kb.igs.org/hc/en-us/article_attachments/202584007/Rinex_3_transition_plan_v3.0.pdf*.

"The Receiver Independent Exchange Format. Version 3.04", from IGS website, *ftp://ftp.igs.org/pub/data/format/rinex304.pdf*.