An Update on the CDDIS IGS Global Data Center

Abstract: The Crustal Dynamics Data Information System (CDDIS) supports data archiving and distribution activities for the space geodesy and geodynamics community. The archive consists of GNSS, SLR, VLBI, and DORIS data sets and products derived from these data. The CDDIS data system is a key component in several of the operational services within the IAG and GGOS, including the IGS, ILRS, IVS, IDS, and IERS. Over the last decade, CDDIS has seen its ingest volume explode to over 30 million files per year from over hundreds of simultaneous data providers. In order to accommodate this increase and to streamline operations, CDDIS has recently performed a significant computer system upgrade requiring updates to the data upload and distribution architecture. The poster provides background information about the system and its user communities, archive contents, and information about these updates and enhancements to the CDDIS. Information about the GNSS data holdings supporting the IGS will also be included.

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

Presented at: 2017 IGS Workshop
Paris France, July 03-07, 2017

Introduction:

- The Crustal Dynamics Data Information System (CDDIS) is NASA's active archive of space geodesy data, products, and information (Global Navigation Satellite System/GNSS, Satellite Laser Ranging/SLR, Very Long Baseline Interferometry/VLBI, and Doppler Orbitography and Radio-positioning Integrated by Satellite/DORIS).
- CDDIS is the principle data center for the geometric supporting services created under the umbrella of the International Association of Geodesy (IAG):
  - International GNSS Service (IGS)
  - International Laser Ranging Service (ILRS)
  - International VLBI Service for Geodesy and Astrometry (IVS)
  - International DORIS Service (IDS)

Archive Contents:

- Point data from permanent stations in the GNSS, SLR/LLR, VLBI, and DORIS networks generated on a multi-day, daily, hourly, and/or sub-hourly basis
  - GNSS: 585+ sites tracking GPS, GLONASS, and new GNSS (Galileo, QZSS, Beidou, IRNSS)
  - Laser Ranging (SLR and LLR): ~40 sites tracking 90+ satellites (including the Moon)
  - VLBI: 45 sites
  - DORIS: 58 sites tracking 6 satellites
- Products derived from these data including precise positions and velocities, satellite orbits, atmospheric parameters

Archive Statistics:

- Archive size: ~19.1TB/203M files
- Ingest rate: ~10GB (100K+ files)/day
- Distribution rate: ~575GB (~5M files)/day

Archive Usage:

- The CDDIS contains data and derived products from over 1500 observing sites located at about 1000 locations around the world, going back in time as far as 1975.
- In 2016, the CDDIS distributed nearly 1.5B files totaling 170TB in volume from over 320K distinct hosts (average of 35K unique hosts/month); we have seen at least a 20% increase in these figures thus far in 2017

IGS Global Data Center @ CDDIS:

GNSS Data @ CDDIS:

- GNSS data archive at CDDIS consists of daily, hourly, and 15-minute high-rate files
- General method for archive is for data providers to push files to CDDIS
- This policy puts the responsibility on the provider to ensure the latest data are archived; CDDIS does not have to poll the providers to make sure latest data are available
- The IGS Infrastructure Committee developed a RINEX 3 transition plan specifying steps to integrate V3 stations and their data into IGS operations
- Since 2016, data in RINEX V3 format are integrated in operational daily, hourly, and high-rate archive structure
- Therefore, multi-GNSS data in RINEX V3 format, and using the RINEX V3 filename specification, are available in the same directory structure as data in RINEX V2 format using the 8.3.Z filename format:
  ftp://cddis.nasa.gov/gnss/data/daily
  ftp://cddis.nasa.gov/gnss/data/hourly
  ftp://cddis.nasa.gov/gnss/data/high-rate
- CDDIS caster providing 270 real-time streams including NASA GDGPS as well as 37 IGS product streams
  - See poster "Real-time data and product performance metrics at NASA GSFC CDDIS" (Blevins, Michael, and Noll) for more information on real-time activities at CDDIS
- MGEX (campaign) data directories continue to contain files in RINEX V3 format using 8.3.Z naming convention

IGS Products @ CDDIS

- Standard final/rapid/ultra-rapid IGS products (station positions, satellite orbits, station and satellites clocks, ERP) by AC and combined
- Troposphere ZPD, ionosphere TEC (final, rapid, predicted)
- Other WG/pilot project products: DCB for multi-GNSS, MGEX campaign, and real-time combination products

Current Status, Recent Developments, and Issues

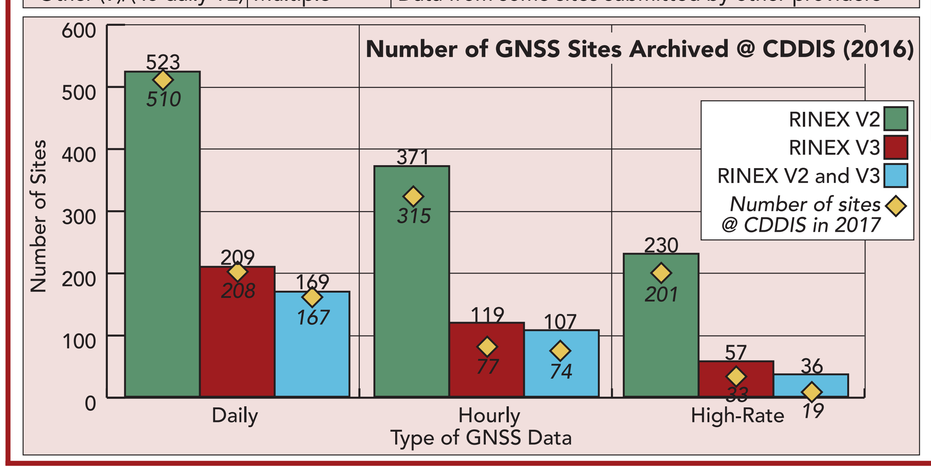
- Integration of RINEX V3 into the RINEX V2 archive has worked well
- Not all data providers have transitioned to the new CDDIS upload system
  - Working with providers to solve problems
  - Has forced CDDIS to script ftp GETs to keep archive current; need to catch up going back to Nov. 2016
- New GNSS processing system found problems with incoming data
  - Operations staff working on an individual basis with providers on any issues
  - See poster "GNSS Quality Control Improvements and Provider Performance Tracking at the Crustal Dynamics Data Information System (CDDIS)" (Woo, Limbacher, Noll, and Michael) for more information about improvements in CDDIS GNSS data operations
- New daily combined mixed GNSS broadcast ephemeris file for RINEX V3 includes all GNSS (thanks N. Romero!)
- Daily status files and other routine reports now available for data in both RINEX V2 (8.3.Z filenames) and V3 (RINEX V3 filenames)
- **Many sites (over 340!) continue to supply data only in RINEX V2 format**
- CDDIS does not yet incorporate a GNSS QC utility for RINEX V3 data similar to teqc used for RINEX V2; continue to investigate options as time permits

GNSS Data Provider Status (for 2017 as compared to 2016)

Provider	Country	RINEX Version 2 Daily	RINEX Version 2 Hourly	RINEX Version 2 High-Rate	RINEX Version 3 Daily	RINEX Version 3 Hourly	RINEX Version 3 High-Rate
Currently Active							
BKG	Germany	80	67	69	35	15	15
IGN (GDC)	France	73	(81)*	(31)*	39	(40)*	
JPL	USA	70	68	60	1	1	1
NRCan	Canada	42	32	3	31	31	3*
GA	Australia	40	39	37	31		
UNAVCO	USA	28	2				
GFZ	Germany	21		17	19		
NGA	USA	18					
PGC	Canada	17	17	13			
RDAAC	Russia	16	14				
KASI (GDC)	South Korea	16	6	1	1		1
ESA	Germany	10	10	10	10	10	10
Others (23)	Multiple	48	30	3	45	22	9
CDDIS retrieval (70+)	USA	33					
Currently Inactive							
IGN (GDC)*	France	No routine submission of hourly and high-rate data					
NGS	USA	Data submitted by other providers					
Other (9)/(45 daily V2)	Multiple	Data from some sites submitted by other providers					

Notes on provider status table:

- Number of sites fluctuate as sites are added/removed from IGS Network
- * Indicates no routine delivery in 2017
- (#) Indicates number of sites in 2016 with no routine delivery in 2017



These statistics show:

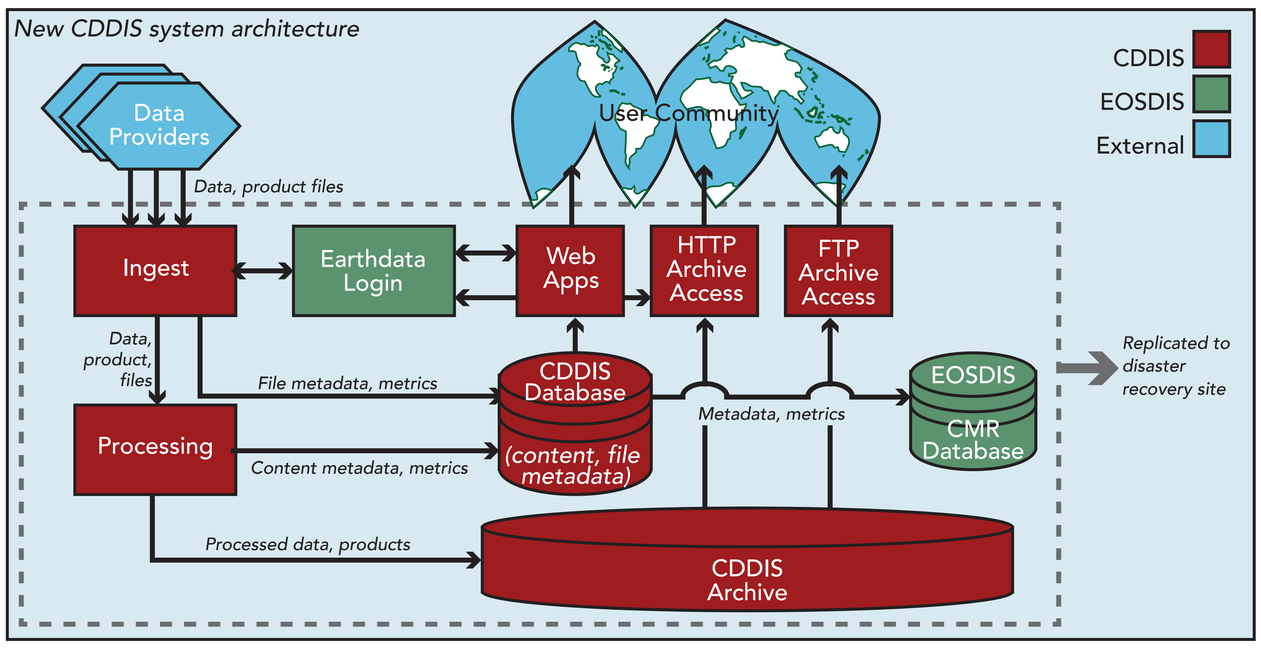
- **New upload procedure is working for providers**
- **However, CDDIS is lacking GNSS data, daily, hourly, and high-rate, in both RINEX V2 and V3 formats, from too many sites.**
- **CDDIS continues to work with providers on their issues**

Number of GNSS Providers (2017)

Data Type	Submitted RINEX Version V2	V3	V2 w/o V3	V3 w/o V2
Daily data	32	23	16	7
Hourly data	20	12	9	1
High-rate data	11	10	4	3

System Facilities/Architecture Improvements:

- Over the past 7 years CDDIS has experienced double-digit growth culminating in over 1.5B downloads and over 170TB of data distributed in 2016
- On path to distribute over 210 TB from 1.75B files in 2017
- Upgraded CDDIS system installed in new location providing better infrastructure (power, network connectivity, etc.)
  - IT infrastructure designed for 4 "nines" uptime
  - Multiple redundant 40Gb networks directly connected to the Internet
  - New system implemented with virtual machine architecture for reliability and expandability
  - Both production and disaster recovery (DR) systems located at different buildings at GSFC
  - Unified storage across both production and DR systems
  - File processing software re-designed for more efficient operations and additional QC
- New system has been operational since December 01, 2016
- Improvements to CDDIS operations
  - Streamlining archive operations across data types
  - Improved metadata for archive operations and data discovery

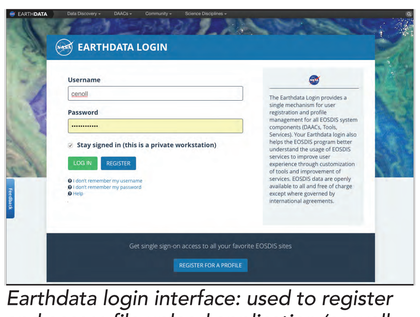


File Upload Procedure @ CDDIS:

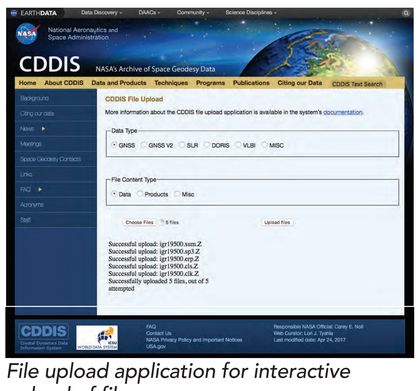
- Because of NASA security restrictions, CDDIS can no longer use non-secure FTP for file uploads from data providers
- New, upgraded system was designed to use HTTPS protocol for file upload
- Implemented both web and command line interfaces
  - Web interface for simple, interactive uploads
  - Command interface for bulk uploads and scripting
  - Users can make simple modifications to existing scripts for uploads to the new system
- cURL is the supported program for command line access but any program that can do HTTP GET and POST is usable
- Sample code (Java, bash) provided for bulk uploading and scripting
- New system uses EOSDIS Earthdata login
- Users must first register with EOSDIS to obtain an Earthdata login ID for access to CDDIS upload system
- For more information: *http://cddis.nasa.gov/Data_and_Derived_Products/CDDIS_File_Upload_Documentation.html*

```
[/home/user]$ curl -c .urs_cookies -n -L
https://depot.cddis.eosdis.nasa.gov/CDDIS_FileUpload/login/

Welcome to CDDIS File Upload

[/home/user]$ curl -X POST -b .urs_cookies -F "fileType=GNSS" -F "fileContentType=PRODUCTS" -F
"file[]=@igr19500.clk.Z" -F "file[]=@igr19500.cls.Z" -F "file[]=@ igr19500.erp.Z" -F
"file[]=@igr19500.sp3.Z" -F "file[]=@igr19500.sum.Z"
https://depot.cddis.eosdis.nasa.gov/CDDIS_FileUpload/upload/
Successful upload: igr19500.clk.Z
Successful upload: igr19500.cls.Z
Successful upload: igr19500.erp.Z
Successful upload: igr19500.sp3.Z
Successful upload: igr19500.sum.Z
Successfully uploaded 5 files, out of 5 attempted
```

Command line program example using cURL for scripted uploads



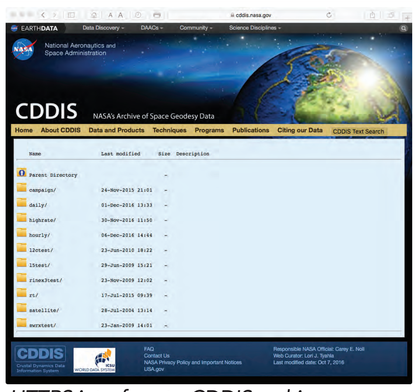
Earthdata login interface: used to register and access file upload application (as well as CDDIS real-time caster)



File upload application for interactive upload of files

What's Up Next?:

- All web activities at NASA, including CDDIS, transitioned to HTTPS in the Fall of 2016 per U.S. government policy
- CDDIS currently developing enhanced HTTPS access to CDDIS archive
  - Archives and users continue to move away from using FTP
  - Therefore, CDDIS will implement access to it's full archive through HTTPS
  - HTTPS access is as efficient as FTP transfer without the firewall/router issues of FTP
    - FTP is a two-port protocol; users can have connectivity problems (e.g., with firewall, DNS, etc.)
    - HTTP is a one-port protocol, fewer issues with downloads
  - HTTPS access will continue to use same structure as provided through FTP
  - Earthdata login will also be used for access through HTTPS
  - FTP access to CDDIS archive will continue but users are encouraged to explore HTTPS capabilities
  - Early results from user testing are promising
- Other plans
  - Develop real-time metrics capability (completeness, latency, etc.)
  - Capture real-time streams to files for archive (coordinate with RTWG and IC)
  - Reprocess historic archive of GNSS data (pre-2016) using new file ingest software to fully represent archive contents with metadata and metrics
  - Rename RINEX V3 data (8.3.Z files) in MGEX campaign directory structure (prior to 2016) to RINEX V3 filenames and move to operational archive



HTTPS interface to CDDIS archive
https://cddis.nasa.gov/archive

More Information/Feedback:

- Data and products are acquired as part of NASA's Earth Science Data Systems and archived and distributed by the Crustal Dynamics Data Information System (CDDIS):
  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.
- The staff welcomes feedback on the CDDIS and in particular the ideas expressed in this poster; contact Carey Noll (Carey.Noll@nasa.gov)

CDDIS
Crustal Dynamics Data Information System