## IN31B-0794

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

AGU Fall Meeting 2019 San Francisco, CA; December 09-13, 2019

# SUPPORTING GLOBAL GEODETIC AND GEOPHYSICAL RESEARCH AND APPLICATIONS

ABSTRACT: The main objective of NASA's Crustal Dynamics Data Information System (CDDIS) is to provide first-class, robust archive, distribution, and real-time services (the IGS, ILRS, IVS, and IDS) and GGOS, the CDDIS receives and distributes a global set of data, derived products, and information to a broad internation to a broad international user community, and provides tools for the exploration and use of the archive. Over the last few years, the CDDIS has seen its download volume increase significantly, highlighting the need to improve and enhance the system to support new activities and to handle the increased load. Therefore, to meet these requirements, a new data ingest system was developed to improve automation and provide a broad range services, from data upload, quality control, metadata extraction, and documentation. In addition, as agencies become increasingly concerned about systems using legacy protocols for data access in order to ensure system security and data integrity for the user community, the CDDIS has implemented alternative methods for access to its archive contents. This poster will include background information about the CDDIS has implemented alternative methods for access to its archive for continued easy and automated download of its contents. This poster will include background information about the CDDIS has implemented alternative methods for access to its archive contents. This poster will include background information about the CDDIS has implemented alternative methods for access to its archive contents. overview of the improvements made to the system to address increase usage and the need to enable data discovery. Future plans to further enhance user support will be presented.

### CRUSTAL DYNAMICS DATA INFORMATION SYSTEM - CDDIS

- 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).
- The contents of the CDDIS archive are utilized for geodetic studies, e.g., plate tectonics, earthquake displacements, Earth orientation, Earth's surface deformation, Earth's gravity field, etc.
- The CDDIS archive also plays an interdisciplinary role in supporting the derivation of a Terrestrial Reference Frame (the foundation for virtually all airborne, space-based and ground-based Earth observations), precise orbit determination (POD) for NASA/international missions, atmospheric studies, etc.
- CDDIS is one of 12 Distributed Active Archive Centers (DAACs) within NASA's Earth Observation System Data and Information System (EOSDIS).
- CDDIS became a regular member of the World Data System (WDS) in March 2013 and a member of Earth Science Information Partners (ESIP) in August 2017.

#### The CDDIS and the IAG

- The largest CDDIS user community comes from the services within the International Association of Geodesy (IAG) where CDDIS serves as one of the primary data centers for the geometric services and its observing system, GGOS (Global Geodetic Observing System):
- + International GNSS Service (IGS)
- + International Laser Ranging Service (ILRS)
- + International VLBI Service for Geodesy and Astrometry (IVS)
- + International DORIS Service (IDS)
- These services function as cooperating federations dedicated to a particular type of data (e.g., GNSS, SLR, VLBI, or DORIS), providing data and products on an operational basis to geodesy analysts as well as a broader scientific community and are examples of a successful model of community management
- Analysis centers in these services produce derived products (e.g., positions of observing stations, Earth orientation parameters, precise satellite orbits, etc.) for use by a broader scientific community

Current space geodesy site locations (archived at CDDIS)

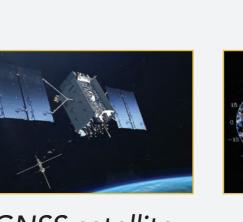
- Point data from permanent stations in the global GNSS, SLR/LLR, VLBI, and DORIS networks
- Data and derived products from over 1500 observing sites located at about 1000 locations around the world, going back in time as far as 1975.
- + Stations in the GNSS, SLR/LLR, VLBI, and DORIS networks generate point data on a multi-day, daily, hourly, and/or sub-hourly basis
- + GNSS: 600+ sites tracking GNSS (GPS, GLONASS, Galileo, QZSS, Beidou, IRNSS)
- + Real-time GNSS: 600+ streams from 400+ sites
- + Laser Ranging: ~40 sites tracking 115+ satellites (including reflectors on the Moon)
- + VLBI: ~50 sites
- + DORIS: 55 sites tracking 7 satellites
- Products derived from these data (some since 1976) including: + Precise network station positions (for ITRF)
- + Satellite orbits (for POD)
- + Station and satellite clocks (for timing)
- + Earth rotation parameters
- + Positions of celestial objects (for CRF)
- + Atmospheric parameters (ionosphere TEC, troposphere ZPD) Metadata information extracted from incoming files for archive management, data discovery

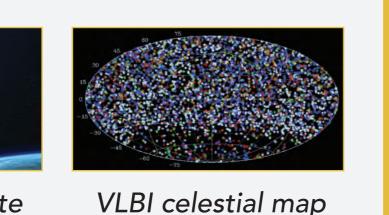
#### Archive statistics

- File size typically <2MB/data granule, <10MB/product granule
- Archive size: ~31TB/290M files Ingest rate: ~16.5GB (110K files)/day
- Distribution rate: ~785GB (~4.5M files)/day
- Multi-day, daily, hourly, sub-hourly
- Varying latencies (minutes, hours, days)
- Data (L1, L1B), products (L2) derived from these data, and information about data and products

- Archive updated with new data/product files on varying time scales, dependent on the data type, from a sub-daily basis to weekly basis
- Users require continuous access to data for generation of products on pre-determined schedules Average user of the CDDIS accesses the contents of the archive through anonymous ftp and
- https by means of automated scripts executed on predefined schedules (typically sub-daily) • Analysts can use this method for data transfer because they are familiar with the structure of the online archive and thus know what files they require, their availability schedule, and where to
- find them within the online structure • So far in 2019, the CDDIS distributed over 1.5B files totaling 270TB in volume from over 240K+

### Space geodesy techniques





**GNSS (95%)** 

SLR (1%)

**VLBI** (2%)

archive (nearly 31TB/290M files)

GNSS data and derived products

The majority of the CDDIS

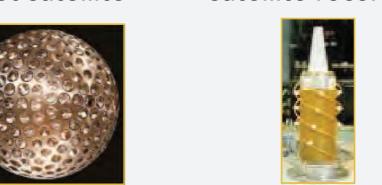
is devoted to the archive of

and antenna

All these space geodesy systems are located at NASA Goddard Space Flight Center, Greenbelt MD.



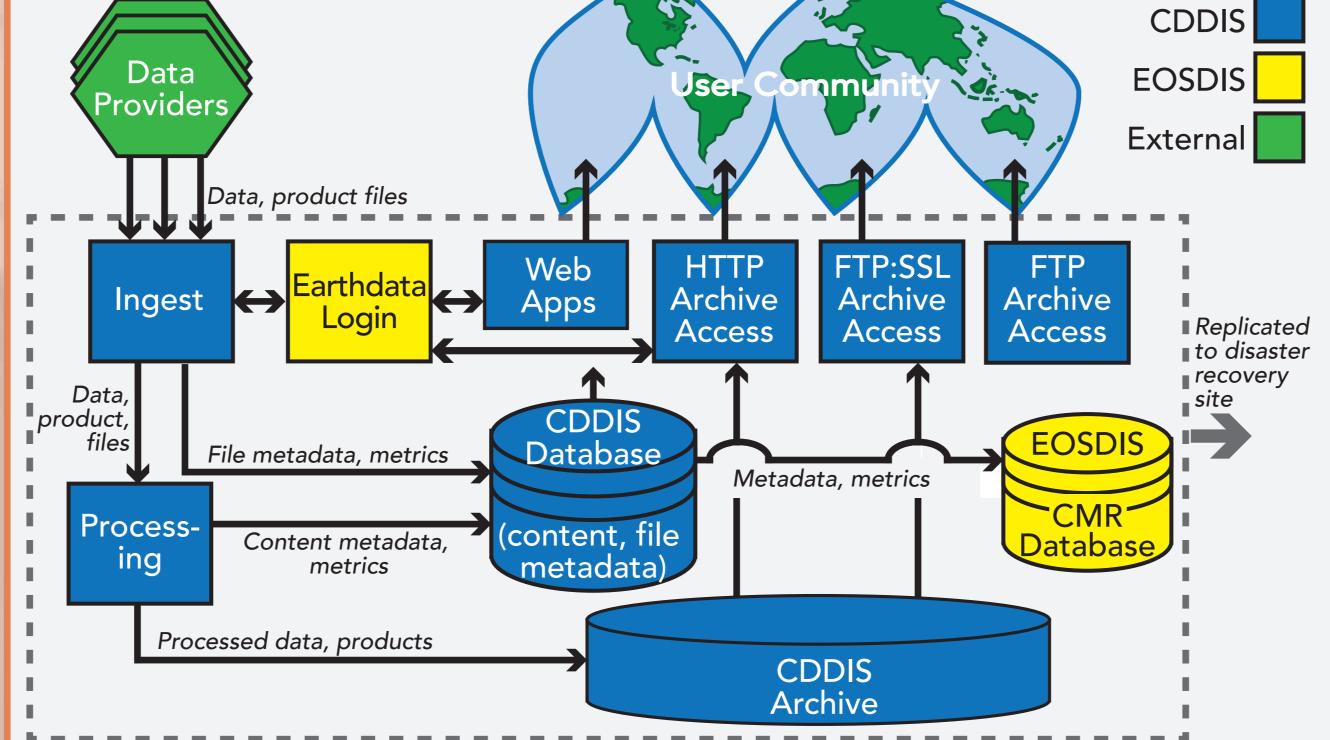




DORIS beacon and

# **Overall CDDIS system architecture**

CDDIS SYSTEM ARCHITECTURE

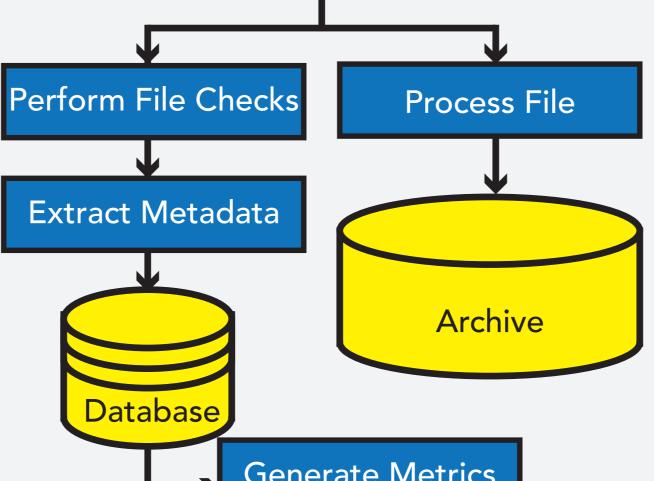


CDDIS system utilizes a virtual machine architecture for reliability and expandability. Both production and disaster recovery (DR) systems utilized and located at different buildings at GSFC. A hardware refresh is planned for installation/operations by end of 2019.

#### CDDIS file upload system

- CDDIS uses an https-based procedure for file upload
- Upload system uses a NASA (EOSDIS) single sign-on system
- + Web interface for simple, interactive uploads
- + Command interface for bulk uploads and scripting
- + cURL is the supported program for command line access but any program that can do HTTP GET and POST is usable

### Overview of CDDIS processing system



Processing of all incoming data files starts with driver software that calls file check and metadata extraction modules. The incoming file is then processed with data type-specific software due to the need to create ancillary files for the archive (e.g., metrics, summary reports, etc.).

- implementation of an upgraded ingest processing
- Goals of the new ingest system
- DORIS) processing levels (data, derived products, and ancillary information), and data providers
- checking, content consistency, etc.)
- + Improved timeliness of data availability
- + Increased automation
- + Improved metadata extraction for data discovery, archive management, and
- Improved data ingest software allowd CDDIS to re-process data and product files archived prior to 2017, ensuring data quality and allowing collection of valuable

### Improvements to CDDIS metadata

- CDDIS currently updating collection and granule level metadata to improve data discovery and satisfy EOSDIS requirements
- + EOSDIS implemented a "Common Metadata Repository" (CMR) + Single source of unified, high-quality, high-performance, and reliable
- Earth Science metadata + Metadata can be discovered and accessed through programmatic
- interfaces leveraging standard protocols and APIs + CDDIS modifying existing metadata to develop relationships with
- **EOSDIS-required metadata** + Initial focus on completion of collection-level metadata records describing CDDIS holdings of GNSS, SLR, DORIS, and VLBI data and derived products
- + Once complete, CDDIS will review and update granule-level metadata as

### CDDIS ARCHIVE ACCESS

#### Background

Earthdata Login and the CDDIS upload/download processes

**User DOWNLOAD Process** 

Archive

Processing

Data Provider

**UPLOAD Process** 

Use of https interface to Earthdata

protocol) and real-time GNS'S stream

downloads. Earthdata Login is also

required for file uploads using https.

Login for file downloads (https

Real-Time

- CDDIS has a large international user community: 320K unique hosts accessed the system in 2018
- + Users currently access the CDDIS archive through anonymous ftp + ftp allows users to easily automate file downloads but has problem
- from a system/security standpoint
- + Archives and users continue to move away from using ftp
- + CDDIS must begin to move users away from reliance on anonymous ftp due to U.S. Government mandates for improved security + CDDIS is committed to ensuring continued, easy, open access to its
- + Outreach materials aim to educate and assist users in transition to
- new access protocols

#### Changes to CDDIS access protocols

- As archives and users move away from using ftp, CDDIS has implemented new protocols for accessing it's archive: https and ftp-ssl
- + Current directory structure of CDDIS archive does not change
- + https (browser and command line)
- URL: https://cddis.nasa.gov/archive
- + ftp-ssl (command line); most similar to standard "anonymous" ftp; Host name: gdc.cddis.eosdis.nasa.gov
- + Using these access protocols requires users to update existing scripts
- used for accessing CDDIS archive + EOSDIS-developed Earthdata Login system used for archive access through https
- > Earthdata login (https://urs.earthdata.nasa.gov/)
- > Single mechanism for user registration and profile management for all EOSDIS components
- Anonymous ftp access to CDDIS archive will continue until OCTOBER
- + Users are STRONGLY encouraged to explore https and ftp-ssl capabilities now in order to transition their scripts by this deadline + CDDIS system will implement bandwidth-limiting on anonymous ftp
- access in the June 2020 time frame to encourage late adopters to switch to using new protocols

#### **Future plans**

- Staff testing implementation of WebDAV (Web Distributed Authoring) and Versioning) interface to provide another method for accessing **CDDIS** archive
- + If feasible for CDDIS, interface would allow users to securely connect to CDDIS archive as if it were a local drive on their computer
- Additional examples (cURL, Wget, Python commands/scripts; ftp:ssl protocol) are available on CDDIS website
- See documentation on CDDIS website:

https://cddis.nasa.gov/About/CDDIS\_File\_Download\_Documentation.html

https://cddis.nasa.gov/Data\_and\_Derived\_Products/CDDIS\_Archive\_Access.html

### Conclusion

- CDDIS https and ftp-ssl access now available!
- Users are STRONGLY encouraged to start testing
- Changes are coming to CDDIS that are beyond our control and you MUST implement scripts to use https or ftp-ssl protocols

### New access protocols: webpage documentation

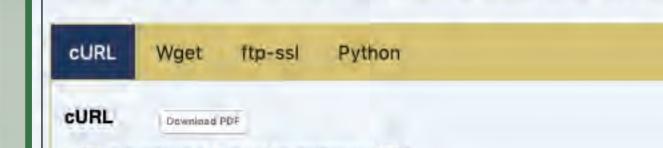
global user community. Since its inception, users have relied on anonymous ftp for accessing and downloading files from the system. Although this protocol allows users to easily automate file downloads, many organization

Therefore, the CDDIS will discontinue anonymous ftp access to its archive in October 2020. Users must now begin

and provide alternative methods for access to its archive for continued easy and automated download of its

their transition to utilize more secure access protocols. In order to access the CDDIS archive, users will need to use a client that supports Transport Security Layer (TLS When using the https protocol (for example, cURL, Wget, or Python commands/scripts) users will need to have a Earthdata Login account. In addition, when utilizing the cURL and Wget, users will need to create a .netrc file:

Examples of accessing the archive via TLS using various methods can be found below. Click on the name of a method to view the instructions; click on a link within a method to see expected results of that command. In the examples, <email address> means type in your actual email address without the < and > symbols.

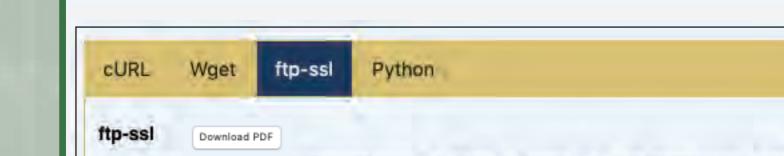


For all of these examples: [file] = file to hold cookie

instructions for creating this file are also available

cURL is a command line tool for accessing URLs

- -n = use .netrc file -L = follow redirects O = write to output file
- -c = file to save cookies to Download Single File
- > curl -c [file] -n -L -0
- 2. Get a list of all files in a directory
- 3. Get a list of files in a directory matching a pattern: (\*\_2018\*.log)
- 5. Download a specific group of files from a list of files: - curl -c .urs cookles -n -L -O
- 6. Download a set of files matching a pattern and output those to a single .tar file.
- https://cddis.nasa.gov/archive/doris/data/cs2/2017/cs2rx1700\*.001.Z



Iftp -d -u anonymous,<email address> -e 'set ftp:ssl-force true' -e 'get

Download an entire directory using lftp: -d -u anonymous,<email address> -e 'set ftp:ssl-force true' -e 'mget vlbi/ivscontrol/";exit

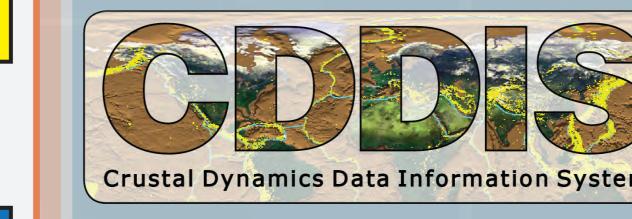
> lftp -d -u anonymous, <email address> -e 'set ftp:ssl-force true' -e 'mget vlbi/ivscontrol/\*;exit' gdc.cddis.eosdis.nasa.gov

NOTE: In this technique, the path to the archive is gdc.cddis.eosdis.nasa.gov. This is different than for the

Logs you into the archive enabling you to use any FTP commands.

Downloads all files in the /vlbi/ivscontrol/ directory to your local file system.

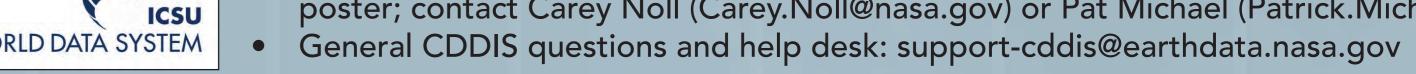
### FOR MORE INFORMATION



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) or Pat Michael (Patrick.Michael@nasa.gov)

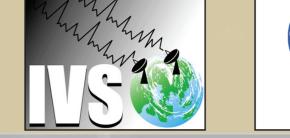












Real-time GNSS ■SLR VLBI DORIS



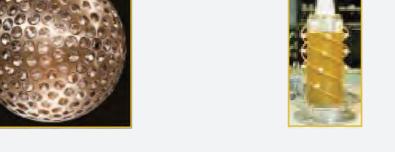


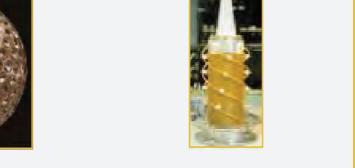


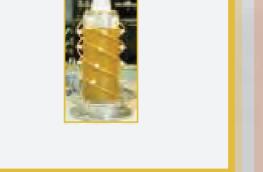


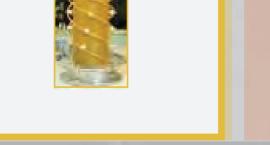


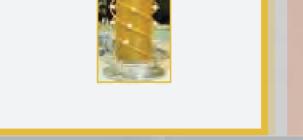


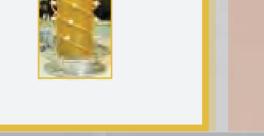


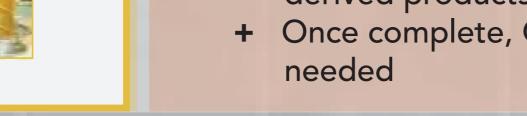


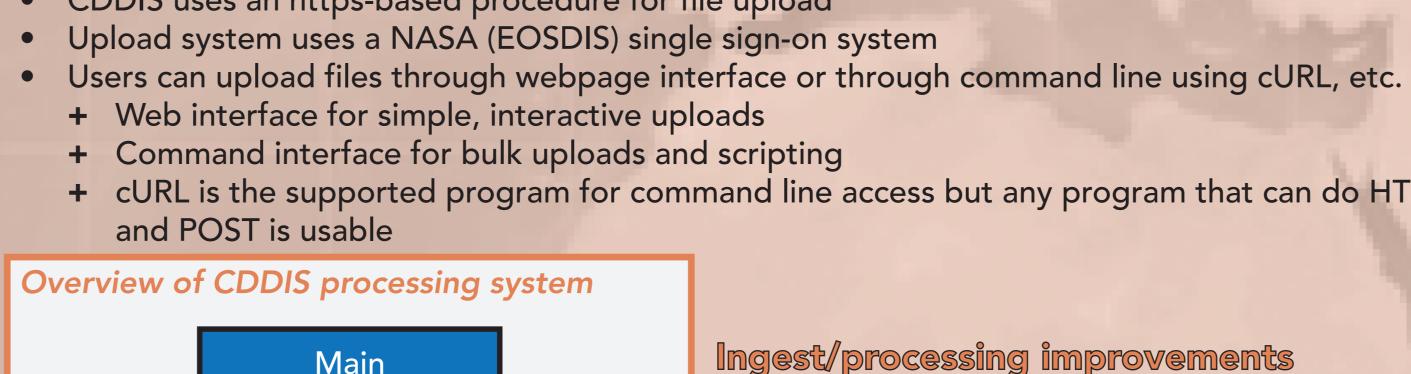












Ingest/processing improvements Over the last two years, the CDDIS completed th

+ Consolidation of key software components across data types + Consistent processing of incoming data across data types (GNSS, SLR, VLBI,

+ Improved quality control measures (e.g., format validation, checksums, virus

quality control monitoring + Improved code maintenance

