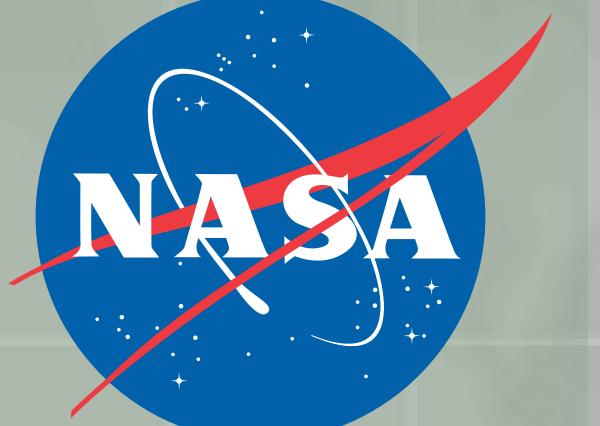
# G06P-298 IUGG19-1523

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

27th IUGG General Assembly Montreal Canada, July 08-18, 2019

# THE CRUSTAL DYNAMICS DATA INFORMATION SYSTEM: NASA'S ACTIVE ARCHIVE OF GEODETIC OBSERVATIONS SUPPORTING RESEARCH IN UNDERSTANDING OUR DYNAMIC EARTH

ABSTRACT: The main objective of the CDDIS system, as a key component in several of the IAG Services, the IGS, ILRS, IVS, and IDS, as well as GGOS, receives and distributes a global set of data and derived products in a central archive, maintains information and use of the archive. Over the last decade, CDDIS has seen its on-line holdings grow to over 26 Tbytes (250 million files) and its ingest volume explode by over 38 million additional files per year from hundreds of providers of files for archive. As the data volume increased load. Therefore, to meet these requirements, a new data ingest system was developed to increase automation. In addition, continued easy and automated download of its contents. This poster will include background information about the CDDIS and its user community. CDDIS has had to implement alternative methods for access to its archive to allow users continued easy and automated download of its contents. This poster will include background information about the CDDIS and its user some increasingly concerned about the CDDIS and its user some increasing the concerned about the CDDIS and its user some increasing the concerned about the CDDIS and its user some increasing the concerned about the CDDIS and its user some increasing the concerned about the CDDIS and its user some increasing the concerned about the CDDIS and its user some increasing the concerned about the CDDIS and its user some increasing th



New access protocols: webpage documentation

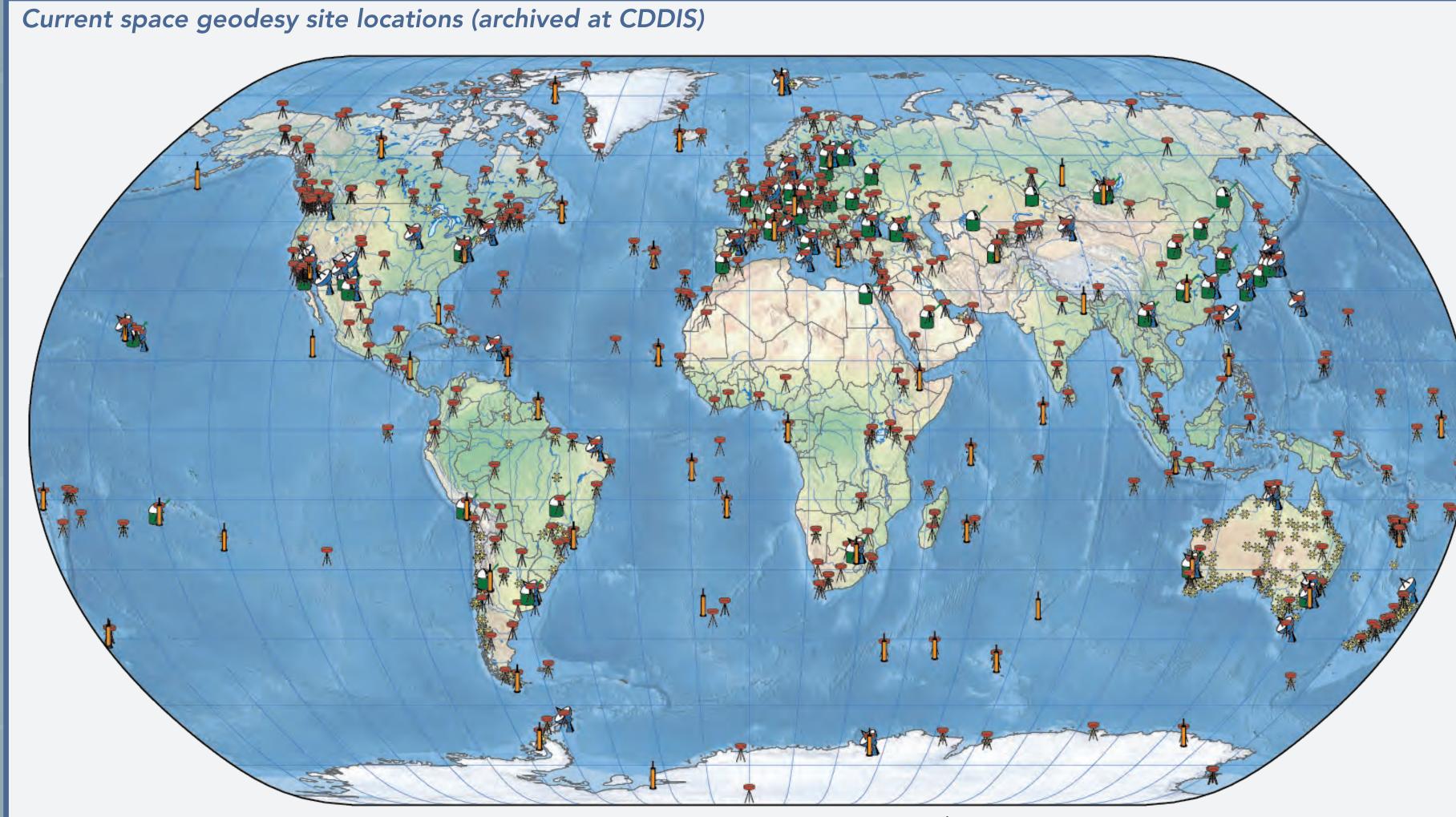
# CRUSTAL DYNAMICS DATA INFORMATION SYSTEM - CDDIS

• 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,
- 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, atmosph
- 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
- CDDIS has extensive partnerships through the IAG, serving as one of the primary data centers for the geometric services and its observing system, GGOS (Global Geodetic Observing System).
- CDDIS is the principle data center for the geometric supporting services created under the
- + 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) • The services provide 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
- Successful operation through cooperation of many international organizations who leverage their respective limited resources to all levels of service functionality The CDDIS user community primarily consists of analysts supporting the services within the
- International Association of Geodesy. These groups produce derived products (e.g., positions of observing stations, Earth orientation
- parameters, precise satellite orbits, etc.) for use by a broader scientific community.



## 

# 



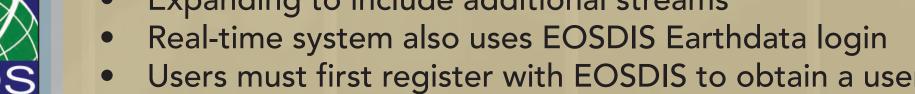












# CDDIS ARCHIVE DESCRIPTION

communities, and recent updates to its archive contents and system architecture, enhancements for data discovery, as well as some future plans.

SLR system and target satellite

VLBI antenna and

All these space geodesy

systems located at NASA

Goddard Space Flight

Center, Greenbelt MD.

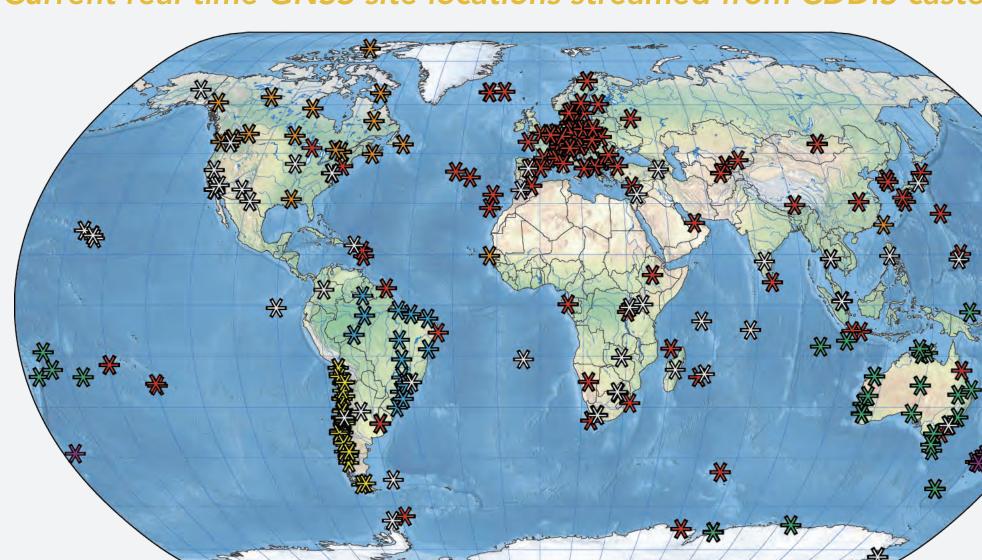
- Point data from permanent stations in the global GNSS, SLR/LLR, VLBI, and DORIS networks
- The CDDIS contains data and derived products from over 1500 observing sites located at about 1000 locations around the world, going back in time as far as 1975.
- + Stations in the GNSS, SLR/LLR, VLBI, and DORIS networks generate point data on a multi-day, daily, hourly, and/or sub-hourly basis
- + GNSS: 585+ sites tracking GPS, GLONASS, and new GNSS (Galileo, QZSS, Beidou, IRNSS)
- + Real-time GNSS: >500 sites
- + Laser Ranging: ~40 sites tracking 100+ 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: + Non-standard metadata, data type specific
- + Extracted from incoming files
- + Internal access to metadata database

### Archive statistics

- File size is typically <2MB/data "granule", <10MB/derived product "granule"
- Archive size: ~27TB/270M files
- Ingest rate: ~15GB (100K files)/day Distribution rate: ~650GB (~5.75M files)/day
- Data (L1, L1B), products (L2) derived from these data, and information about data and products
- Multi-day, daily, hourly, sub-hourly Varying latencies (minutes, hours, days)

them within the online structure

- The archive is 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. • The average user of the CDDIS accesses the contents of the archive through anonymous ftp 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
- In 2018, the CDDIS distributed over 2B files totaling 235TB in volume from over 320K distinct hosts.



CDDIS contributes to the IGS Real-Time Service, a clock correction service enabling monitoring, etc.) at global scale. for scientific and hazard detection applications. This map shows global sources for the real-time streams available at CDDIS. **★**BKG (Germany)

(nearly 27TB) is devoted to the

y data type for 2018)

**GNSS (95%)** 

The majority of downloads (over

were for GNSS data and derived

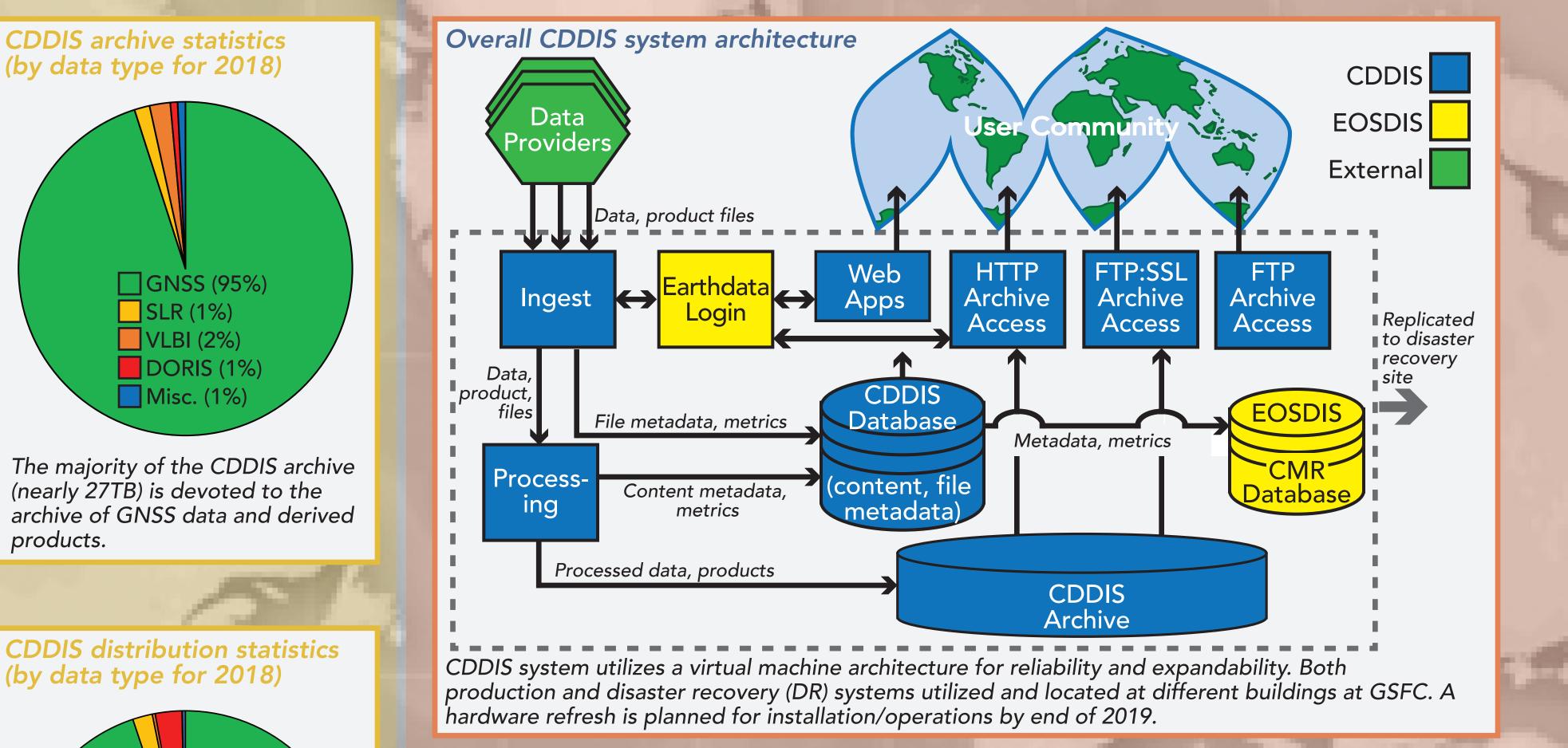
212TB, 1.3M files) from the CDDIS

**X** LINZ (New Zealand) ★NRCan (Canada) ₩GA (Australia)

# CDDIS support of real-time data and products

- CDDIS streams real-time GNSS data and derived products in support of the IGS Real-Time Service (RTS) One of three real-time "casters"
- Data from over 500 global sites and 40 derived product steams Expanding to include additional streams
- Users must first register with EOSDIS to obtain a user ID for access to real-time GNSS caster

# CDDIS SYSTEM ARCHITECTURE



## CDDIS file upload system

- CDDIS uses an https-based procedure for file upload Upload system uses a NASA (EOSDIS) single sign-on system
- Users can upload files through webpage interface or through command line using cURL, etc.
- + 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

gest/processing improvements

checking, content consistency, etc.)

+ Increased automation

collected

quality control monitoring

+ Improved code maintenance

+ Improved timeliness of data availability

implementation of an upgraded ingest processing

+ Consolidation of key software components across data types

+ Consistent processing of incoming data across data types (GNSS, SLR, VLBI,

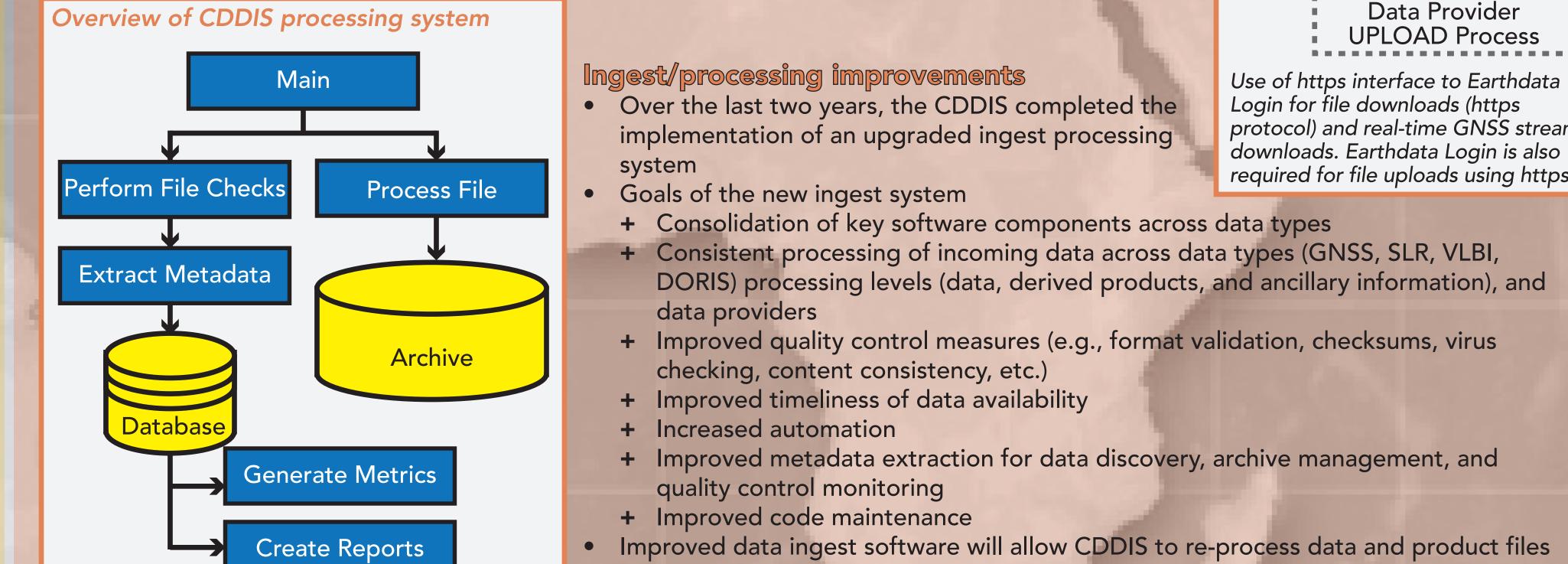
+ Improved quality control measures (e.g., format validation, checksums, virus

+ Improved metadata extraction for data discovery, archive management, and

archived prior to 2017 to ensure data quality and allow valuable metadata to be

Flow of CDDIS metadata

DORIS) processing levels (data, derived products, and ancillary information), and



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.).

# mprovements 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

Earthdata Login and the CDDIS upload/download processes

User DOWNLOAD Process

Processing

Ingest

Data Provider

**UPLOAD Process** 

Systems

Use of https interface to Earthdata

protocol) and real-time GNSS stream

downloads. Earthdata Login is also

required for file uploads using https.

Login for file downloads (https

- CDDIS has a large international user community: 320K unique hosts
- + Users currently access the CDDIS archive through anonymous ft
- + ftp allows users to easily automate file downloads but has proble
- + Archives and users continue to move away from using ftp
- ftp due to U.S. Government mandates for improved security + CDDIS is committed to ensuring continued, easy, open access 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
- + https (browser and command line)
- URL: https://cddis.nasa.gov/archive
- Host name: gdc.cddis.eosdis.nasa.gov
- used for accessing CDDIS archive
- through https
- for all EOSDIS components
- Anonymous ftp access to CDDIS archive will continue until OCTOBER
- + Users are STRONGLY encouraged to explore https and ftp-ssl
- access in the June 2020 time frame to encourage late adopters to switch to using new protocols

### Future plans

- and Versioning) interface to provide another method for accessing **CDDIS** archive
- to CDDIS archive as if it were a local drive on their computer
- Additional examples (cURL, Wget, Python commands/scripts; ftp:ssl
- 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

- MUST implement scripts to use https or ftp-ssl protocols

- accessed the system in 2018
- from a system/security standpoint
- + CDDIS must begin to move users away from reliance on anonymous

- + Current directory structure of CDDIS archive does not change
- + ftp-ssl (command line); most similar to standard "anonymous" ftp;
- + Using these access protocols requires users to update existing scripts
- + EOSDIS-developed Earthdata Login system used for archive access
- > Earthdata login (https://urs.earthdata.nasa.gov/)
- > Single mechanism for user registration and profile management
- capabilities now in order to transition their scripts by this deadline + CDDIS system will implement bandwidth-limiting on anonymous ftp

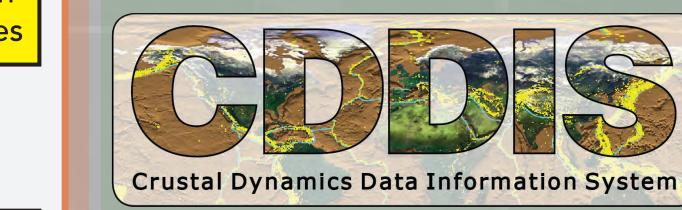
- Staff testing implementation of WebDAV (Web Distributed Authoring
- + If feasible for CDDIS, interface would allow users to securely connect
- protocol) are available on CDDIS website

### 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

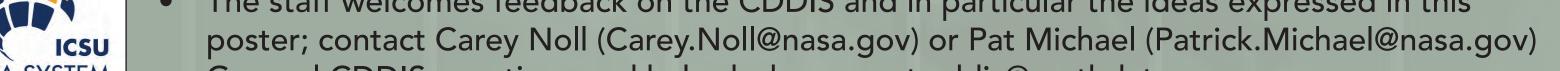
### gdc.cddis.eosdis.nasa.gov Logs you into the archive enabling you to use any FTP commands.

## 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.



General CDDIS questions and help desk: support-cddis@earthdata.nasa.gov

ftp-ssl is an extension to the commonly used File Transfer Protocol (FTP) that adds support for the Transport Layer Security (TLS). NOTE: In this technique, the path to the archive is gdc.cddis.eosdis.nasa.gov. This is different than for the Iftp -d -u anonymous,<email address> -e 'set ftp:ssl-force true' -e 'get

2. Download an entire directory using Iftp: o -d -u anonymous,<email address> -e 'set ftp:ssl-force true' -e 'mget vlbl/ivscontrol/";ex > lftp -d -u anonymous, <email address> -e 'set ftp:ssl-force true' -e 'mget vlbi/ivscontrol/\*;exit' gdc.cddis.eosdis.nasa.gov Downloads all files in the /vlbi/ivscontrol/ directory to your local file system.

-- curl -c [file] -n -L "https://cddis.nasa.gov/archive/gps/data/daily/2019/045/19o/"?list"

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"

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:

3. Log into the archive using ftp commands

URL is a command line tool for accessing URLs

> curl -c [file] -n -L -0

2. Get a list of all files in a directory

- curl -c .urs cookles -n -L -O

-- curl -c .urs\_cookles -n -L -o files.tar

cURL Wget ftp-ssl Python

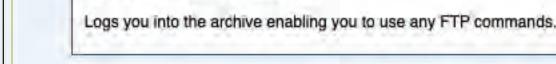
For all of these examples:

[file] = file to hold cookies

-c = file to save cookies to

-n = use .netrc file

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













• The staff welcomes feedback on the CDDIS and in particular the ideas expressed in this