

For more information:
<https://cddis.nasa.gov>

Processing at NASA CDDIS

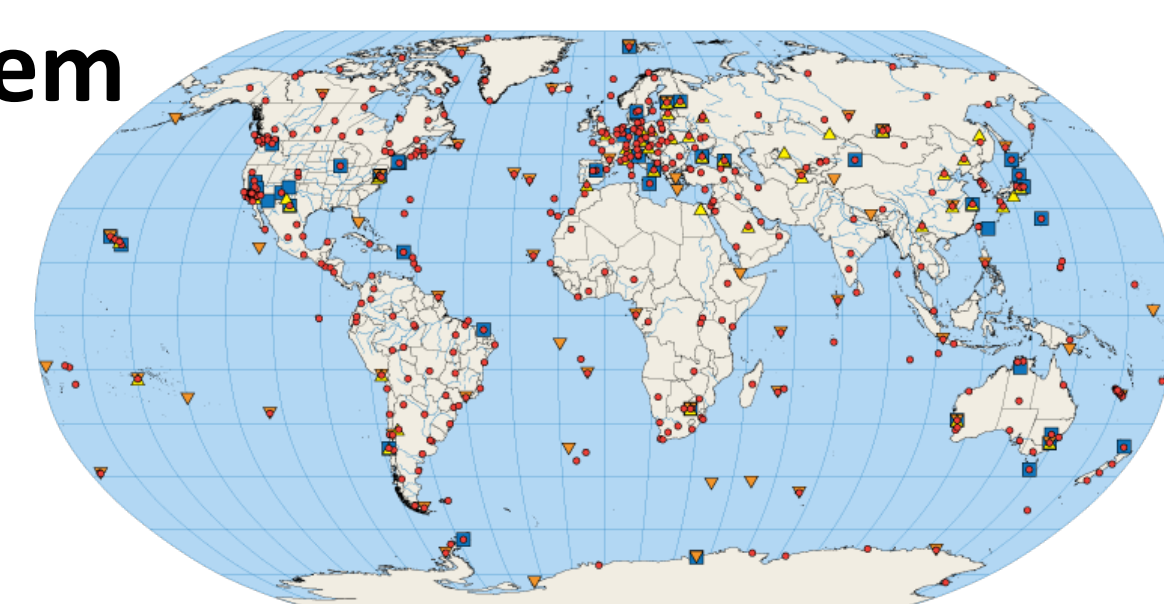
ABSTRACT: The Crustal Dynamics Data Information System (CDDIS) has been providing access to space geodesy and related data sets since 1982, and in particular, Global Navigation Satellite Systems (GNSS) data and derived products since 1992. The CDDIS became one of the Earth Observing System Data and Information System (EOSDIS) archive centers in 2007. As such, CDDIS has evolved to offer a broad range of data ingest services, from data upload, quality control, documentation, metadata extraction, and ancillary information. With a growing understanding of the needs and goals of its science users CDDIS continues to improve these services. Due to the importance of GNSS data and derived products in scientific studies over the last decade, CDDIS has seen its ingest volume explode to over 30 million files per year or more than one file per second from over hundreds of simultaneous data providers. In order to accommodate this increase and to streamline operations and fully automate the workflow, CDDIS has recently updated the data submission process and GNSS processing. This poster will cover this new ingest infrastructure, workflow, and current operations.

¹Patrick Michael (Patrick.Michael@nasa.gov)
¹Carey Noll (Carey.Noll@nasa.gov)
²Justine Woo (Justine.Y.Woo@nasa.gov)
³Rebecca Limbacher (Rebecca.L.Limbacher@nasa.gov)

Crustal Dynamics Data Information System

Background:

- Is NASA's active archive of space geodesy data, products, gathering information from:
 - Global Navigation Satellite System (GNSS)
 - Satellite Laser Ranging (SLR)
 - Very Long Baseline Interferometry (VLBI)
 - Doppler Orbitography and Radio-positioning Integrated by Satellite (DORIS)
- Collecting data from over 1000 sites worldwide.
- Providing data archiving & distribution support for over thirty years.
- The largest CDDIS user community comes from the services within the International Association of Geodesy (IAG).
- CDDIS has been an ESDIS DAAC since 2007



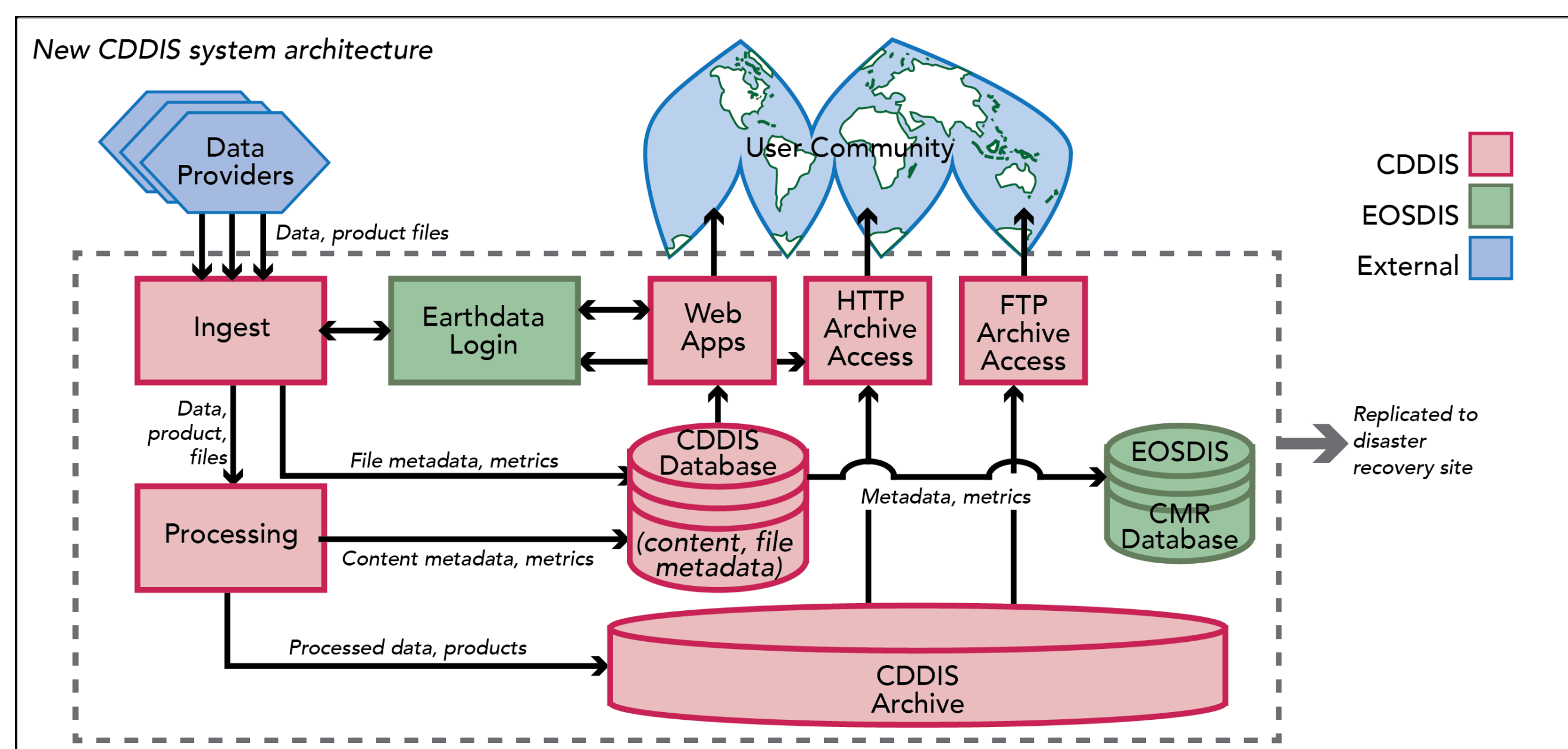
Space Geodesy Site Locations

Archive Statistics:

- Archive size: ~17.5TB
- Ingest rate: ~9GB (100K files)/day
- Distribution rate: ~500GB (~4.7M files)/day
- Data (L1, L1B), products (L2) and information about data and products
- Latencies ranging from: multi-day, daily, hourly, sub-hourly, and sub-minute
- Expected distribution in 2016 is 1.5B files, 180TB and over 140,000 users served

Operations Overview:

- CDDIS's primary operations can be categorized as:
 - Provide systems to collect data & derived products from worldwide suppliers
 - Perform quality control and make data & products available for distribution
 - Collect & store metadata & provide data discovery both internally & externally
 - Develop valued-added reports from archive contents for science community
- CDDIS's archive and distribution have been growing at over 30% annually
- CDDIS transitioned to all new infrastructure and implemented an entirely new GNSS processing system on 1 December 2016.

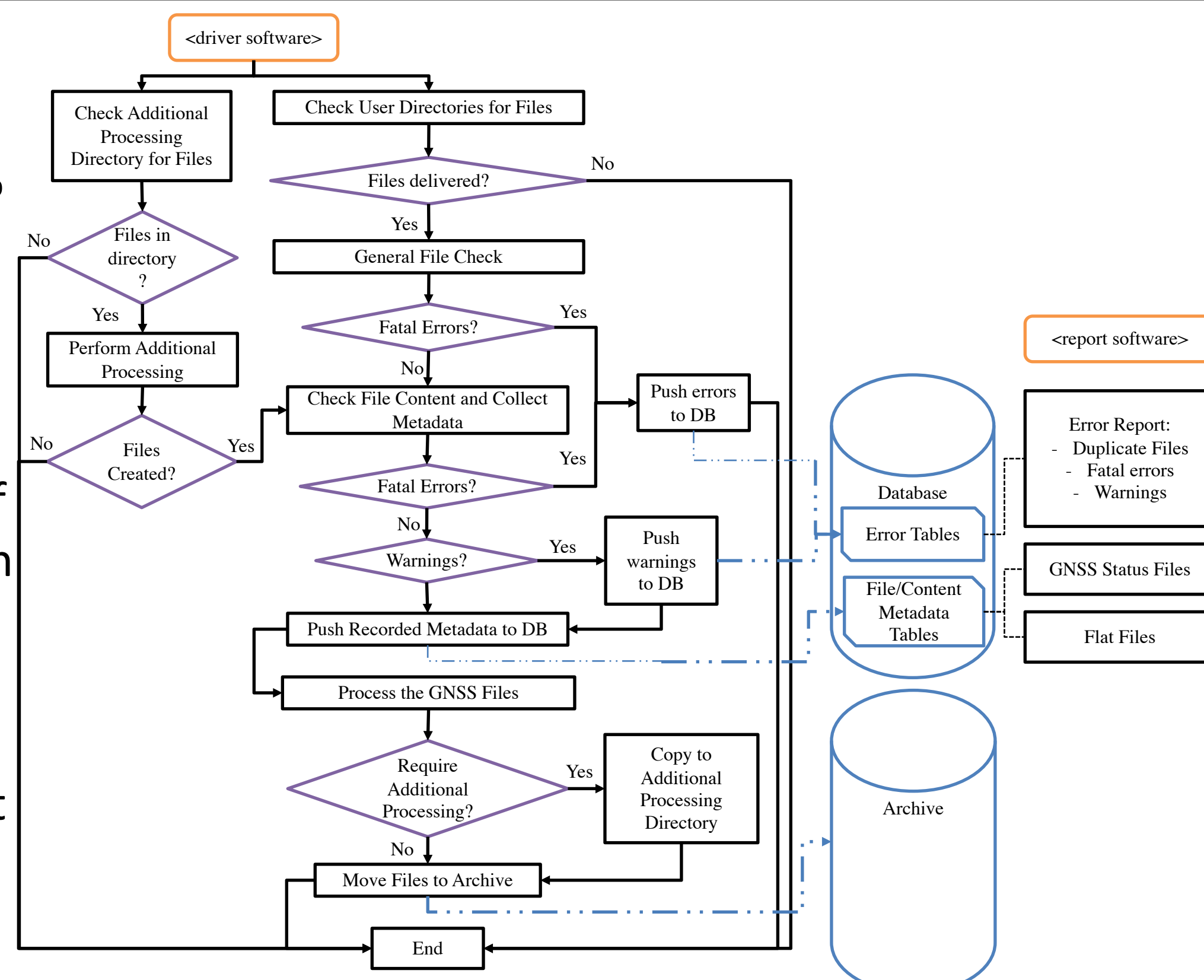


CDDIS Operational Overview

GNSS Processing:

Overview:

- GNSS accounts for 95% of all incoming files
- ~30 million files/year
- Delivery rates up to 100 files/second
- Latency from receipt of data files to availability in the archive is commonly less than 30 seconds
- Files cannot just be copied to the archive but MUST be quality tested before making available



GNSS Workflow Diagram

Ingest Operations:

- All new file upload process. See poster [G41B-1013](#) and https://cddis.nasa.gov/About/CDDIS_File_Upload_Documentation.html
- All files undergo a battery of security checks before any processing takes place.
- 24 total quality control checks
- 5 warning checks – some can be fixed in real-time

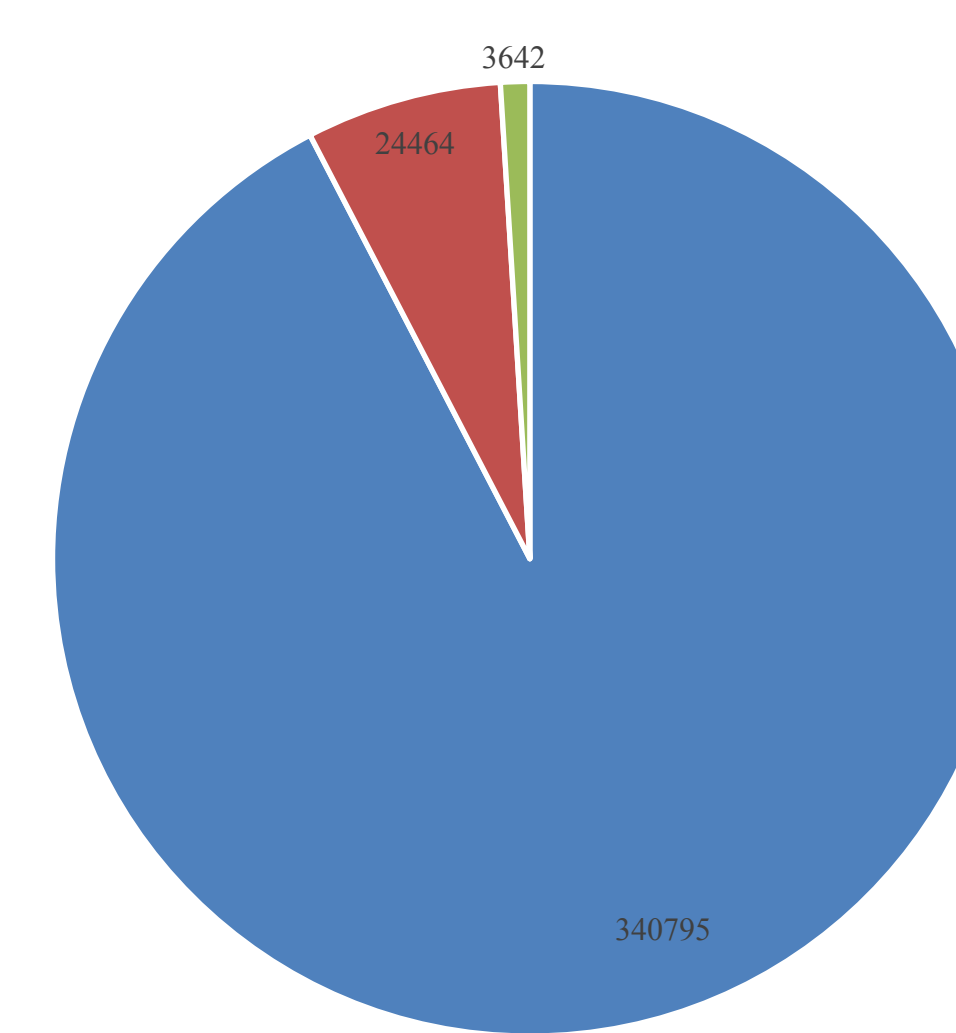
Examples include:

- Incorrect compression scheme
- Incorrect file type submitted
- Incorrect naming scheme

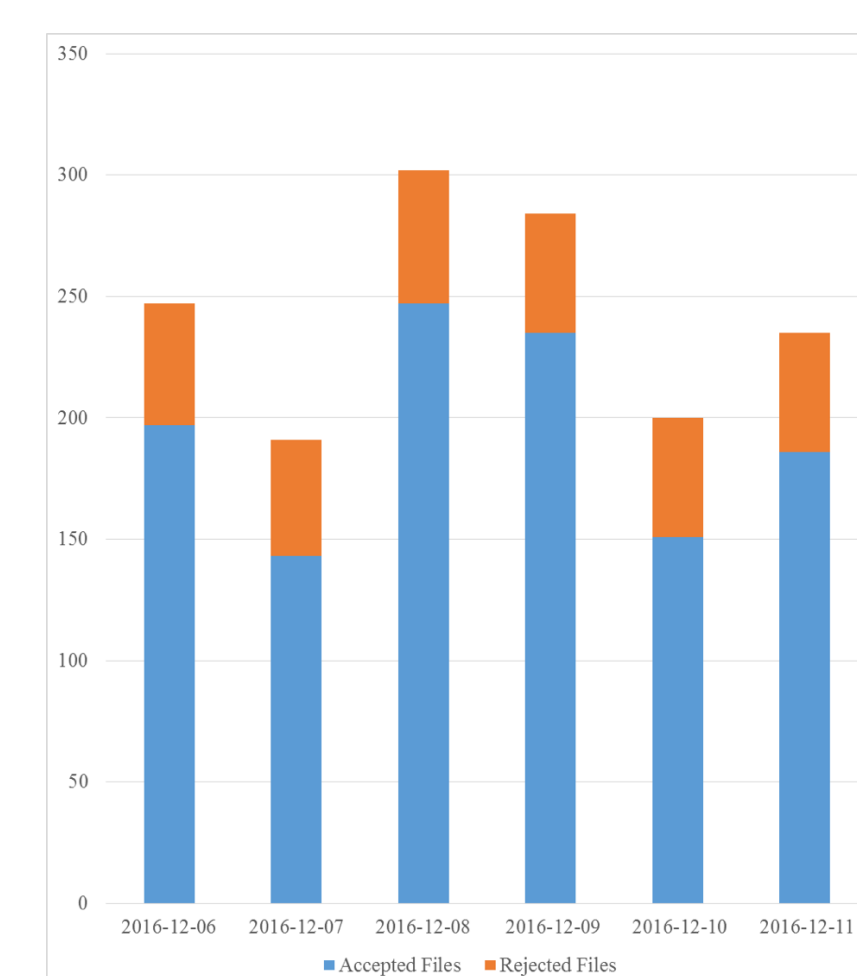
• 19 "fatal" checks – errors that stop file processing

Example includes:

- Unknown file type
- Filename does not match file contents
- Unknown RINEX version
- Unknown monument marker
- Virus infected file



Total Files Processed Report Over 6 Day Period
 Pass = Files Successfully Processed
 Warnings = Files Corrected During Processing
 Fatal = Files Not Able to Be Processed Due to Errors



User Report on Duplicate Files

Processing Reports, Metrics and Graphics

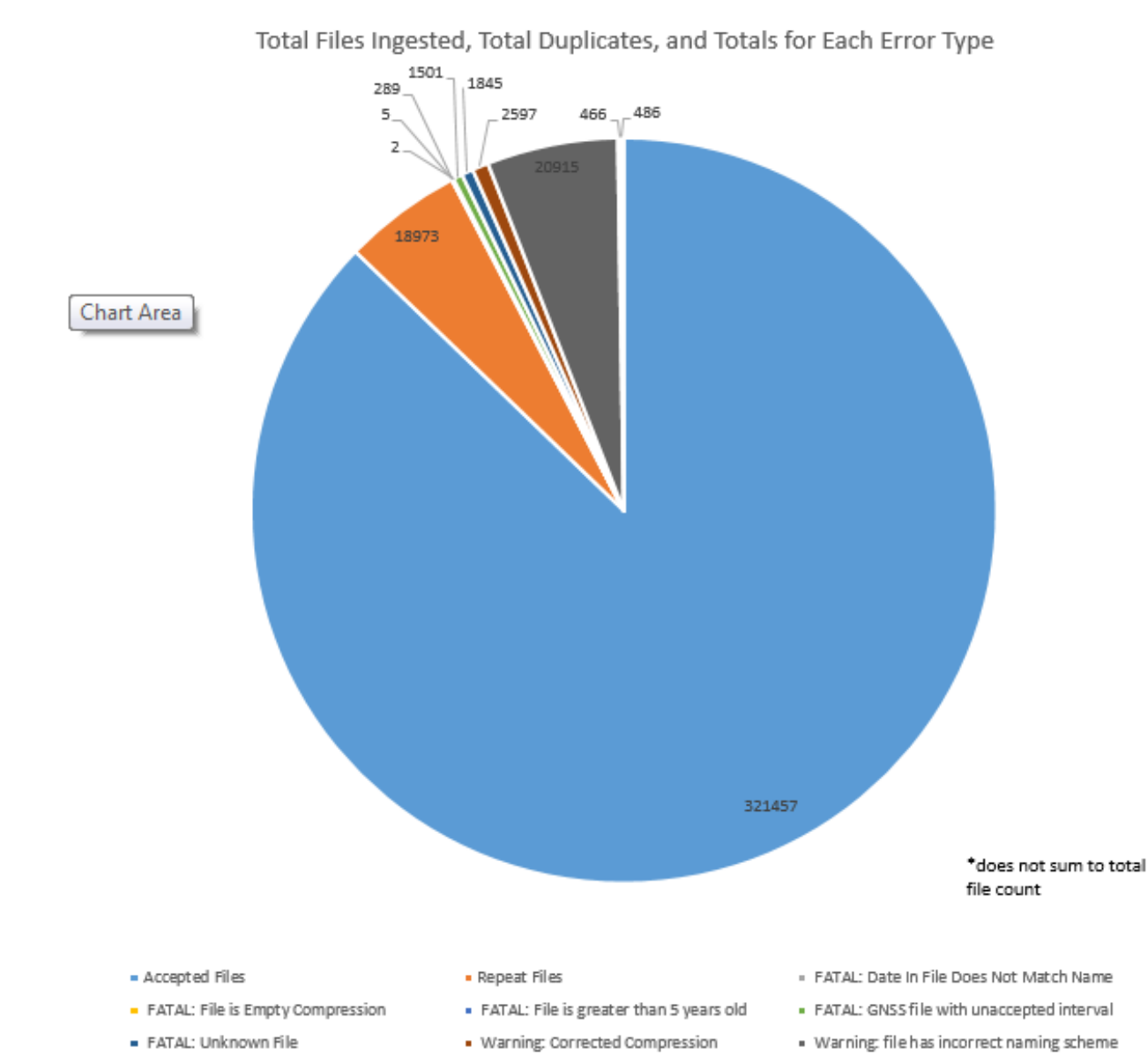
- Metrics on individual data providers
- System wide statistics
- Metrics on file quality
 - By data provider
 - By file type
- Metrics on number/size of incoming files

Metadata Operations:

- All incoming files are processed for both file and content metadata
- 68 total metadata parameters are gathered for each GNSS file
 - File metadata – 16 parameters per file
 - Content metadata - 52 parameters per file
- Content & file metadata are publically searchable in multiple data discovery tools
- Metadata are provided to the Earth Observing System Data Information System (EOSDIS) through their Common Metadata Repository (CMR) for data discovery across all 12 Distributed Active Archive Centers (DAACs).
- Metadata will eventually be made available to the Global Geodetic Observing System (GGOS) for discovery by the geodetic community

Reporting Operations:

- Data provider reports
 - Error reports
 - Trending reports over time
- System level reporting
 - Metrics generation
 - Ingest rates
 - Data provider error rates
 - Trending reports
- Science community reports
 - GNSS status reports



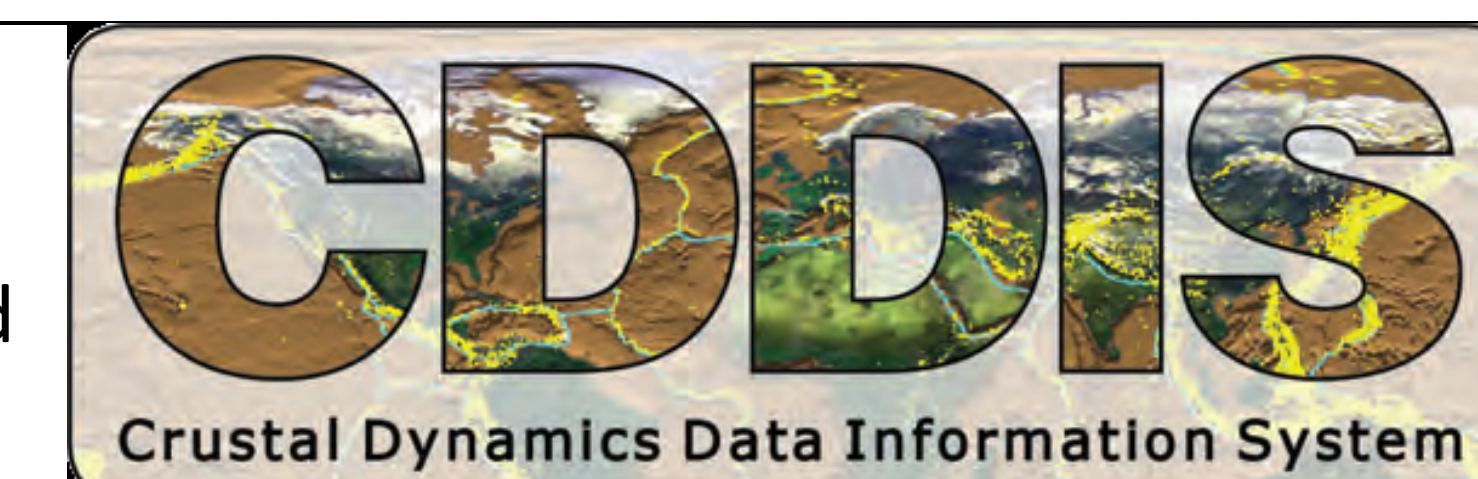
Total File Processing Over a Six Day Period

Future Capabilities:

- System will eventually incorporate processing of SLR, DORIS, and VLBI data files
- Additional content metadata collections
- Automated ingest reports by data provider
- Better analysis and handling of duplicate files
- Web interface for viewing real-time processing statistics
- Additional geodetic community directed reports/summaries
- Transition entire process to an agile software development path

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



- NASA Goddard Space Flight Center, Greenbelt, MD, USA
- Stinger-Ghaffarian Technologies, Lanham, MD, USA
- Science, Systems and Applications, Lanham, MD, USA

