An Update on the CDDIS

Carey Noll/NASA GSFC (Carey.Noll@nasa.gov)
Patrick Michael/Catholic University of America (Patrick.Michael@nasa.gov)
Maurice Dube/SSAI (Maurice.P.Dube@nasa.gov)
Nathan Pollack/SSAI (nathan.pollack@ssaihq.com)

NASA Goddard Space Flight Center
Code 690, Greenbelt, MD 20771
USA

Abstract: The Crustal Dynamics Data Information System (CDDIS) supports data archiving and distribution activities for the space geodesy and geodynamics community. The main objectives of the system are to store space geodesy and geodynamics related data products in a central data bank, to maintain information about the archival of these data, and to disseminate these data and information in a timely manner to a global scientific research community. The archive consists of GNSS, laser ranging, VLBI, and DORIS data sets and products derived from these data. The CDDIS is one of NASA's Earth Observing System Data and Information System (EOSDIS) distributed data centers; EOSDIS data centers serve a diverse user community and are tasked to provide facilities to search and access science data and products.

The CDDIS data system and its archive have become increasingly important to many national and international science communities, in particular several of the operational services within the International Association of Geodesy (IAG) and its project the Global Geodetic Observing System (GGOS), including the International DORIS Service (IDS), the International GNSS Service (IGS), the International Laser Ranging Service (ILRS), the International VLBI Service for Geodesy and Astrometry (IVS), and the International Earth Rotation Service (IERS).

The CDDIS has recently expanded its archive to support the IGS Multi-GNSS Experiment (MGEX). The archive now contains daily and hourly 30-second and sub-hourly 1-second data from an additional 35+ stations in RINEX V3 format. The CDDIS will soon install an Ntrip broadcast relay to support the activities of the IGS Real-Time Pilot Project (RTPP) and the future Real-Time IGS Service. The CDDIS has also developed a new web-based application to aid users in data discovery, both within the current community and beyond. To enable this data discovery application, the CDDIS is currently implementing modifications to the metadata extracted from incoming data and product files pushed to its archive.

This poster will include background information about the system and its user communities, archive contents and updates, enhancements for data discovery, new system architecture, and future plans.

Data Discovery Developments

Current access methods:

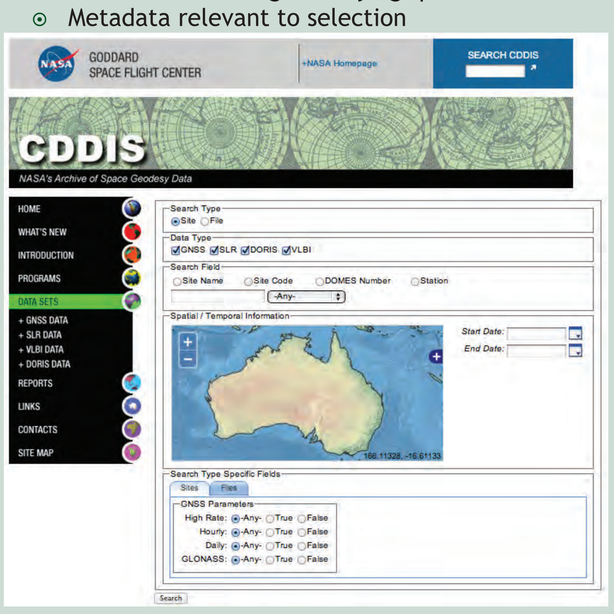
- The CDDIS contains data and derived products from over 1500 observing sites located at about 1000 locations around the world, going back in time as far as 1975.
- The archive is updated with new data/product files on varying time scales, dependent on the data type, from a sub-daily basis to monthly basis.
- The majority of CDDIS user community are analysts supporting the services within the IAG that produce derived products (e.g., positions of observing stations, Earth orientation parameters, precise satellite orbits, etc.) for use by a broader scientific community.
- 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 them within the online structure.

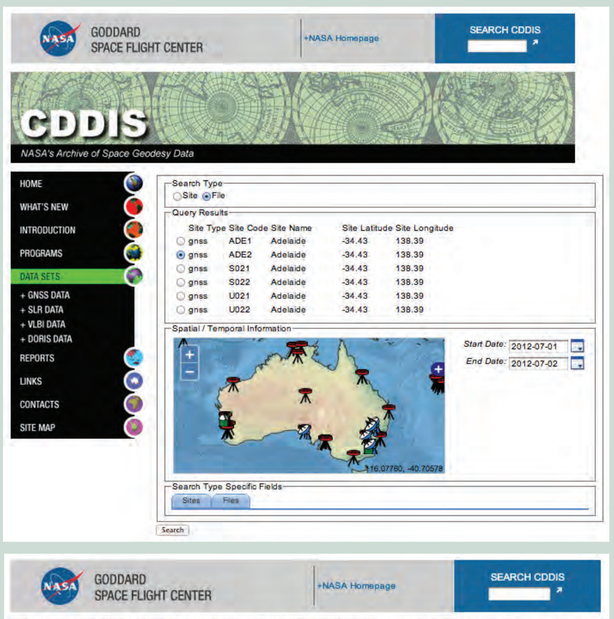
Development:

- SIO/SOPAC, UNAVCO, and CDDIS data centers are participating in an effort to modernize the GSAC (now GSAC-WS, the Geodesy Seamless Archive Center - Web Services) through a NASA ACCESS (Advancing Collaborative Connections for Earth System Science) proposal "Discovery and Delivery of Space Geodetic Data Products from Distributed Archives"
- The ACCESS team has developed the GSAC Service Layer (GSL), a Java-based application that is used at each repository (CDDIS, UNAVCO, SOPAC) to interface to the data center-developed GSAC Repository Implementation. The GSL is a middleware framework that provides different types of output (HTML, XML, wget scripts, etc.) from the data repository metadata databases. The GSL handles the incoming web service requests and routes it to the repository. A federated Repository Implementation has been developed that allows for an aggregate search across all repositories.
- The CDDIS has developed a custom interface, tailored to CDDIS user requirements, through its own GSAC Repository client based on an open source application framework. This application currently interacts with the GSL for both Site and File searches. This custom interface allows CDDIS to leverage the capabilities of the GSL while providing an interface tailored to CDDIS users. Parts of this interface could eventually be made more generic and rolled into the GSL.

Data Discovery Enhancements:

- New users of the CDDIS, both those familiar with space geodesy techniques as well as new research communities, would prefer a browsing interface to the archive contents.
- Furthermore, users also need to browse the archive for new or historic data sets.
- Therefore, the CDDIS has designed a web interface based search tool that queries the CDDIS metadata.
- Users have the ability to specify search criteria based on temporal, spatial, target, site designation, and/or observation parameters in order to identify data and products of interest for download.
- Results of these queries will include a listing of sites (or other metadata) or data holdings satisfying the user input specifications.
- Develop a search/metadata interface tool for CDDIS to:
  - Aid users in discovery of CDDIS data, products, and information
  - Aid staff in archive management
  - Promote CDDIS data holdings to a larger community (e.g., through metadata standards)
- Specify (any/all):
  - Temporal: Year, date/time, range
  - Spatial: Region, lat/lon, range
  - Target: Satellite (SLR, DORIS)
  - Designation: Station name/number/code
  - Parameter: Receiver type (GNSS), event timer (SLR), antenna type (GNSS, VLBI), …
- Results:
  - List of sites satisfying specifications
  - List of data holdings satisfying specifications
  - Metadata relevant to selection

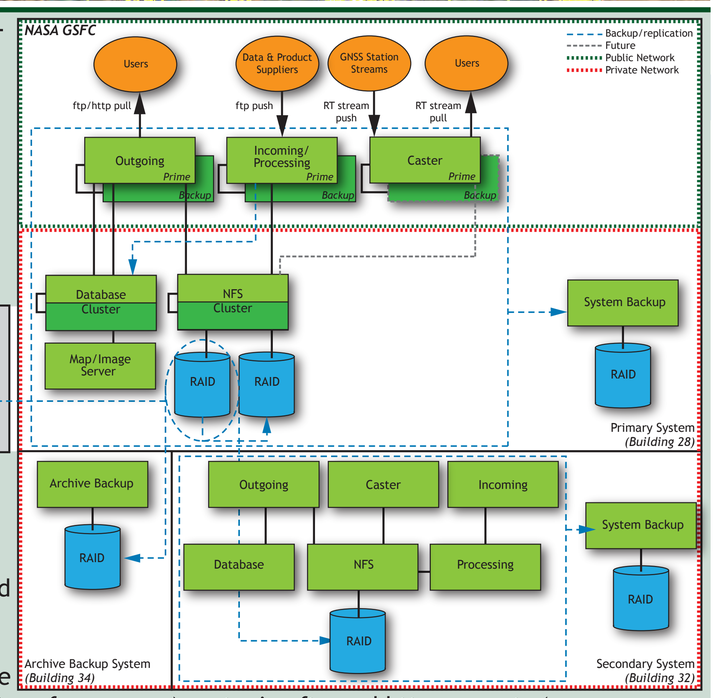




- *Above left:*
  User selects "Site" as the search type and queries for all GNSS, SLR, VLBI, and DORIS sites in the Australian area.
- *Above right:*
  The query yields a map and list of valid sites that can be identified for retrieval. The user selects "Files" as the search type, picks a single site of interest, and specifies a time range in July 2012.
- *Right:*
  A list of valid URLs are presented for user download.

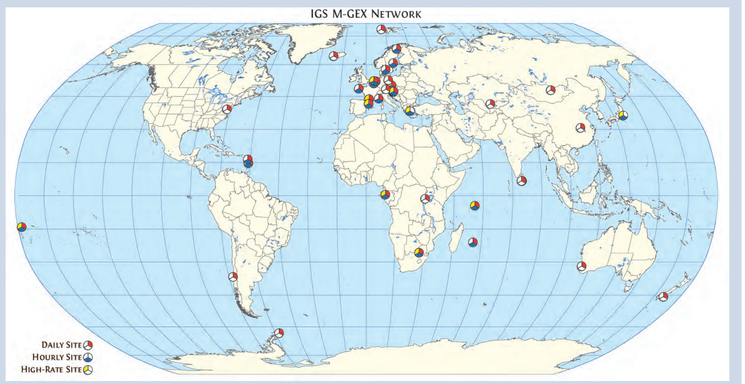
New CDDIS System Architecture

- On May 09, 2012 CDDIS operations were transferred to new server environment. Although the new servers continue to utilize a distributed configuration for the CDDIS, other functions were further distributed for more efficient operations:
  - outgoing servers: handle all web and ftp traffic
  - incoming servers: accept transfers from authorized data suppliers
  - database cluster: handles database for metadata management (internal), data transfer logs (internal), data discovery (external)
  - processing/NFS server: archive processing and NFS access to RAID array
  - caster server: provide real-time streams to authorized clients
- This new configuration allows for efficient and timely processing of incoming files as well as enhanced system reliability by separating user/archive functions. Distinct servers handle incoming data and product files (server *cddisin.gsfc.nasa.gov*), outgoing ftp and http requests (server *cddis.gsfc.nasa.gov*), real-time caster (server *cddis-caster.gsfc.nasa.gov*), and archive operations to the RAID storage.
- The new system is fully redundant with backup servers/clusters which can take over operations should a failure occur with the prime server. A secondary system installed in a second location at GSFC can accommodate CDDIS operations should the prime server be offline for an extended period of time. Both primary and secondary systems have backup systems.
- The archive is equipped with a multi-Tbyte RAID storage system and is scaled to accommodate future growth. The CDDIS archive increases in size by approximately 12 million files/1 Tbyte per year; the existing storage will accommodate the archive requirements for the near future.
- Additional archive backups are made to both local and off-site disk arrays.
- The cut over to the new system was accomplished with minimal downtime (3 hours for data download, 10 hours for data receipt)



MGEX Support

- The CDDIS is one of three IGS data centers supporting the IGS Multi-GNSS Experiment (MGEX). Following a successful proposal to the IGS MGEX CfP, the CDDIS expanded its data archive and distribution service to include data from participating multi-GNSS receivers, products derived from the analysis of these data, and any required metadata for the experiment. The archive now contains daily and hourly 30-second and sub-hourly 1-second data from an additional 40+ stations in RINEX V3 format. The CDDIS developed software to extract metadata from the RINEX headers for use in reporting (see figure) and archive management. The software is also used to generate daily status files (see figure) similar to those provided in the CDDIS IGS operational directories. However, data quality information, as produced by teqc, is not available through this software.



- The MGEX data are available within the campaign subdirectory structure in the CDDIS archive:
  ftp://cddis.gsfc.nasa.gov/gnss/data/campaign
  - Daily 30-second files:
    /gnss/data/campaign/mgex/daily/rinex3/YYYY/DDD/YYT
    /gnss/data/campaign/mgex/daily/rinex2/YYYY/DDD/YYT
    /gnss/data/campaign/mgex/daily/raw/YYYY/DDD/YYT
    40 sites
  - Hourly 30-second files:
    /gnss/data/campaign/mgex/hourly/rinex3/YYYY/DDD/HH
    /gnss/data/campaign/mgex/hourly/rinex2/YYYY/DDD/HH
    /gnss/data/campaign/mgex/hourly/raw/YYYY/DDD/HH
    21 sites
  - 15-minute 1-second files:
    /gnss/data/campaign/mgex/highrate/rinex3/YYYY/DDD/HH/YYT
    /gnss/data/campaign/mgex/highrate/rinex2/YYYY/DDD/HH/YYT
    /gnss/data/campaign/mgex/highrate/raw/YYYY/DDD/HH/YYT
    11 sites

IGS MGEX GNSS Data Holdings for 2012 (as of 13-Jul-2012 13:20:48)

Mon. Name	Site Name	Lat.	Lon.	Receiver	Start Date	Start Day	End Date	End Day	No. Sites
ABMF	Abymes	16 16	-61 32	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	132
BRST	Brest	48 23	355 30	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	142
BRUX	Brussels	50 48	4 22	SEPT POLARX4TR	08-Mar-12	12068	12-Jul-12	12194	127
CONZ	Concepcion	-36 51	286 58	LEICA GRX1200+G	01-Feb-12	12032	12-Jul-12	12194	155
CUT0	Bentley	-32 00	115 32	TRIMBLE NETR9	01-Feb-12	12032	11-Jul-12	12193	151
DLF1	Delft	51 59	4 23	TRIMBLE NETR9	08-May-12	12129	12-Jul-12	12194	61
DYNG	Dionysos	38 05	23 56	TRIMBLE NETR9	14-May-12	12135	12-Jul-12	12194	47
FTNA	Futuna	-14 18	-178 07	TRIMBLE NETR9	02-Jul-12	12184	09-Jul-12	12191	8
GOP6	Ondrejov	49 55	14 47	LEICA GRX1200+G	01-Feb-12	12032	12-Jul-12	12194	162
GRA2	Graz	47 04	15 30	LEICA GRX1200+G	16-Feb-12	12045	08-Jul-12	12190	66
GRAB	Graz	47 04	15 30	IFEN SX_NSR_RT_	08-Feb-12	12039	29-Mar-12	12089	51
GRAB	Graz	47 04	15 30	IFEN SX_NSR_RT_	30-Mar-12	12090	12-Jul-12	12194	86
GRAC	Grasse	43 45	6 55	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	83
HARB	Hartebeesthoek	-25 53	27 41	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	67
KIR8	Kiruna	67 51	20 58	TRIMBLE NETR9	01-Mar-12	12061	12-Jul-12	12194	113
LMMF	Lamentin	14 36	-61 00	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	128
MOSE	Rome	41 32	01 14	LEICA GR25	27-Apr-12	12118	12-Jul-12	12194	75
MAR7	Maartsbo	60 35	17 16	TRIMBLE NETR9	01-Mar-12	12061	12-Jul-12	12194	126
MYVA	Reykjavik	64 08	338 03	LEICA GR10	14-Feb-12	12045	12-Jul-12	12194	148
NKLG	Libreville	0 21	9 40	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	118
NURK	Kigali	-1 34	30 03	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
NYA2	Ny Alesund	78 56	11 52	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
OBE3	Oberpfaffenhofen	48 56	11 18	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
OHIX	O'Higgins	-63 19	302 06	LEICA GRX1200+G	01-Feb-12	12032	12-Jul-12	12194	158
ONS1	Onsala	57 24	11 56	TRIMBLE NETR9	01-Mar-12	12061	12-Jul-12	12194	127
OUS2	Dunedin	-45 52	170 31	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
POTS	Potsdam	52 23	13 04	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	74
REUN	La Reunion	-21 12	55 34	TRIMBLE NETR9	16-Feb-12	12047	08-Jun-12	12160	107
SEYG	Seychelles	-4 40	55 29	TRIMBLE NETR9	03-Jul-12	12185	09-Jul-12	12191	7
SGOC	Narahenpita	06 32	79 31	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
TASH	Tashkent	41 18	69 17	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
TLSE	Toulouse	43 34	1 29	TRIMBLE NETR9	16-Feb-12	12047	12-Jul-12	12194	112
ULAB	Ulaanbaatar	47 40	107 03	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75
UNB3	Fredericton	45 57	293 21	TRIMBLE NETR9	15-Jun-12	12167	12-Jul-12	12194	28
USN4	USNO, Washington	38 55	282 56	SEPT POLARX4TR	01-Feb-12	12032	12-Jul-12	12194	149
USN5	USNO, Washington	38 55	282 56	NOV OEM6	07-May-12	12128	12-Jul-12	12194	67
WARN	Warnemuende	54 10	12 06	JPS LEGACY	30-May-12	12151	12-Jul-12	12194	40
WIND	Windhoek	-22 34	17 05	JAVAD TRE_G3TH	08-Jun-12	12160	28-Jun-12	12180	20
WTZR	Wettzell	49 09	12 53	LEICA GRX1200+G	04-Jun-12	12156	10-Jun-12	12162	7
WTZZ	Wettzell	49 09	12 53	JAVAD TRE_G3TH	01-Feb-12	12032	08-Jul-12	12190	127
WUH2	Wuhan	30 32	114 21	JAVAD TRE_G3TH	29-Apr-12	12120	12-Jul-12	12194	75

Totals: 3,669 days of data from 39 stations

MGEX Tracking Network Status for 08-Jul-12 120708 12190 | GPS Week 1696 Day 1 | As of date: Jul 11 2012 09:11:22

Site	Dly (H)	Receiver Type	Antenna Type	Ant. Height	Marker Name	Marker Number	Type	RINEX Version	Dly (M)
abmf	3	1TRIMBLE NETR9	TRM57971.00 NONE	0.0000	ABMF	97103M001	M	3.02	230
brst	19	1TRIMBLE NETR9	TRM57971.00 NONE	0.2431	BRST	10004M004	M	3.02	1190
brux	1	1SEPT POLARX4TR	JAVRINGANT_DM NONE	0.4689	BRUX	13101M010	M	3.01	50
conz	11	1LEICA GRX1200+GNSS	LEIAR25.R3 LEIT	0.0574	CONZ	41719M002	M	3.01	715
cut0	1	1TRIMBLE NETR9	TRM59800.00 SCIS	0.0000	CUT0	59945M001	M	3.02	47
dlf1	1	1TRIMBLE NETR9	LEIAR25.R3 LEIT	0.0000	DELFT-22	13502M009	M	3.02	1
dyng									
ftna	60	1TRIMBLE NETR9	TRM59800.00 NONE	0.0000	FTNA	92902S002	M	3.00	3631
gop6	1	1LEICA GRX1200+GNSS	LEIAR25.R4 LEIT	0.0653	GOP6	11502M006	M	3.01	15
gmsd									
gra2	24	1LEICA GRX1200+GNSS	LEIAR25.R4 LEIT	0.0600	GRA2	11001M004	M	3.01	1483
grab	1	1IFEN SX_NSR_RT_800	LEIAR25.R4 NONE	0.0600	GRAB	11001M004	M	3.00	42
grac	3	1TRIMBLE NETR9	TRM55971.00 NONE	0.0500	GRAC	10002M010	M	3.02	230
harb	1	1TRIMBLE NETR9	TRM59800.00 NONE	3.0520	HARB	30302M009	M	3.00	50
kir8	3	1TRIMBLE NETR9	LEIAR25.R3 LEIT	0.0000	KIR8		M	3.00	230
lmmf	3	1TRIMBLE NETR9	TRM55971.00 NONE	0.0000	LMMF	97205M001	M	3.02	230
mose	4	1LEICA GR25	LEIAR25.R4 LEIT	0.0000	MOSE	12772M001	M	3.01	253
mar7	3	1TRIMBLE NETR9	LEIAR25.R3 LEIT	0.0000	MAR7		M	3.00	230
matg									
myva	1	1LEICA GR10	LEIAR25.R4 LEIT	0.0000	MYVA	10205M001	M	3.01	110
nklg	12	1TRIMBLE NETR9	TRM59800.00 SCIS	3.0430	NKLG	32809M002	M	3.00	759
nurk	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1300	NURK	34001M001	M	3.02	3
nya2	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1206	NYA2	KOLD	M	3.02	3
obe3	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1206	OBE3	14208M005	M	3.02	3
ohix	11	1LEICA GRX1200+GNSS	LEIAR25.R3 LEIT	0.0668	OHIX	66008M006	M	3.01	715
ons1	1	1TRIMBLE NETR9	LEIAR25.R3 LEIT	0.0000	ONS1		M	3.00	230
ous2	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.0711	OUS2	50212M002	M	3.02	5
pots	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1206	POTS	14106M003	M	3.02	6
reun									
seyg	60	1TRIMBLE NETR9	TRM59800.00 NONE	0.0000	SEYG	39801M004	M	3.00	3631
sgoc	7	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1220	SGOC	SGOC	M	3.02	439
tash	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1206	TASH	12327M001	M	3.02	6
tlse	1	1TRIMBLE NETR9	TRM59800.00 NONE	1.0530	TLSE	10003M009	M	3.00	50
ulab	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1206	ULAB	24201M001	M	3.02	7
unb3	1	1TRIMBLE NETR9	TRM57971.00 NONE	0.1245	UNB3	40146M002	M	3.02	13
usn4	1	1SEPT POLARX4TR	AOAD/M_T NONE	0.0000	usn4	40451S007	M	3.01	89
usn5	1	1NOV OEM6	AOAD/M_T NONE	0.0000	usn5	40451S007	M	3.01	90
warn	7	1JPS LEGACY	LEIAR25.R3 LEIT	0.0690	warn	14277M002	M	3.00	469
wind									
wtzr									
wtzz	58	1JAVAD TRE_G3TH DELTA	LEIAR25.R3 LEIT	0.0450	WTZZ	14201M014	M	3.01	3508
wuh2	1	1JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T NONE	.1206	WUH2	WUH2	M	3.02	7
zim2									

Program: QC 2009Mar23 by UNAVCO run with elevation angle cutoff of 10 degrees

Field	Size	Type	Explanation
Site	4	char	Site name
Dly (H)	3	number	Delivery delay in hours
Receiver Type	20	char	Type of GPS receiver from RINEX header
Antenna Type	20	char	Type of GPS antenna from RINEX header
Ant. Height	6	number	Height of antenna from RINEX header
Marker Name	20	char	Marker name from RINEX header
Marker Number	10	char	Marker DOMES number from RINEX header
Type	1	char	GPS type
RINEX Version	7	char	RINEX Version number
Dly (M)	3	number	Delivery delay in minutes

Above: Summary report of current M-GEX daily data holdings for 2012. Similar reports are available for hourly and high-rate data holdings. Right: Daily status file listing RINEX header information summary for day 12190.

Other Recent Developments

- The CDDIS has installed a dedicated server to serve as an NTRIP broadcast relay system to support the activities of the IGS Real-Time Pilot Project (RTPP) and the future Real-Time IGS Service. Software has been installed and is under testing. Applications have been submitted to GSFC network management board for activation of appropriate ports for exchange of streams using NTRIP. The CDDIS will also investigate process for capturing incoming streams and generating/comparing high-rate data files for archive.
- A design for a CDDIS website upgrade is currently underway. The new website will feature access to a map selection tool and data discovery application. An improved navigation design will also be incorporated.