IGS Data Centers: Status and Progress

Carey Noll, NASA GSFC/CDDIS
Fran Boler, UNAVCO
Heinz Habrich, BKG

IGS 20th Anniversary Workshop
Pasadena, CA
June 25, 2014
• Then and now
• Current status
  – MGEX support
  – Real-time activities
  – Metadata efforts
Using CDDIS as an example:

Data
- 1994:
  - ~85 daily 30-second sites
- 2014:
  - 500+ daily 30-second sites
  - 330+ hourly 30-second sites
  - 180+ high-rate 1-second sites
  - 140+ real-time sites
  - 110+ MGEX sites (daily, hourly, high-rate)

Products
- 1994:
  - Orbit and ERP from 7 ACs + combination
  - Weekly and daily rapid
- 2014:
  - Orbit, ERP, reference frame, clock solutions from 9 ACs
  - Reference frame solutions from 5 AACs
  - Weekly, daily, rapid, ultra-rapid
  - Orbit, clock, and reference frame combination
  - ~600 files/week
  - Ionosphere
  - Troposphere
  - MGEX
  - RTS
  - repro1 and repro2
20 Years of IGS Data Archiving

Year: 1994 to 2014

- Number of Sites
- Number of Days
- Number of High-Rate Sites
- Number of Hourly Sites
- Number of Daily Sites

*Estimated values for full year
CDDIS archive is currently nearly 10 Tbytes in size; over 9 Tbytes of this archive, or 90%, is devoted to the storage of GNSS data and derived products.
• IGS data centers currently supporting MGEX: CDDIS and IGN (Global DCs), BKG (regional DC)
• All data supplied in RINEX V3 format
• Separated from operational IGS data (archived in RINEX V3 format)
  – Under “campaign” directory structure at GDCs, e.g.,
    ftp://cddis.gsfc.nasa.gov/gnss/data/campaign/mgex
• Files utilize same filenaming convention, e.g.,
  ssssddd#.yyt.Z
  – Same filenames have caused confusion and errors in archive
• Current “separation” not conducive for integration in IGS operations
• Need to plan for integration of RINEX V3 into operational archive structure while ensuring clear “separation”
RINEX V3: Filenames

• Move to utilization of new filenaming convention as specified in RINEX V3.02 documentation
  – `mmmmsscccc_s_yyyydddhhm_m_fff_tt.rnx/crx.gz`
  – `ALGO00CAN_R_20121600000_01D_GN.rnx.gz`
  – `KIRU00SWE_R_20141600000_01D_30S_MO.crx.gz`

• New filenaming convention will require significant work for DCs (and ACs)

• Will remove some difficulties DCs have experienced in handling V2 and V3 data from same site

• Test:
  – CDDIS has started providing RINEX V3 data (from selected sites) in both old and new filename format
  – Purpose: Promote use of new filename format
  – Utilizes gzip for compression

• Need a “teqc-like” utility for data QC and metadata extraction
• Nearly 140 sites supply real-time streams to IGS RTS
• Some high-rate files created at the receiver, some from real-time streams
• DCs must ensure that files generated from these streams are sufficiently reliable; some suggestions:
  – tool for comparison of RINEX files from various construction approaches
  – minimum requirements for acceptance of an accumulated data stream of observations as a RINEX file in IGS data archives
  – mandatory/optional observation types to be included
  – procedures to fill the gaps in the case data streams have been interrupted
  – separate directories for distinguishing between files created from streams and by receivers?
• Past recommendation: IGS DCs investigate implementation of improved metadata standards and search and data discovery tools to improve access to the diverse archive structures available through the IGS and to aid in the integration of these archives into GGOS and other global data systems.
• Metadata is essential for data discovery
• DCs extract metadata from incoming files for managing their archives
• Tools, e.g., teqc, also useful for metadata extraction (DCs need similar tool for RINEX V3!)
• DCs provide statistics on availability of data
• Information about usability of the data would also be helpful
Statistics on Data Availability

- Daily status files available at CDDIS summarize data holdings and include key metadata
- Available for operational and MGEX daily data
- File location: ftp://cddis.gsfc.nasa.gov/gnss/data/daily
- Software available to other DCs

IGS Tracking Network Status for 09-Jun-14 140609 141601

<table>
<thead>
<tr>
<th>Slps</th>
<th>Receiver Type</th>
<th>Antenna Type</th>
<th>Height</th>
<th>Marker Name</th>
<th>Number</th>
<th>Type</th>
<th>Version</th>
<th>(M)</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td>aber</td>
<td>TRIMBLE NETR9</td>
<td>TRM599000.00</td>
<td>SCIS</td>
<td>3.9650</td>
<td>ABER</td>
<td>13231M001 M</td>
<td>2.11</td>
<td>9430 64de0c941419e4e7e47f2589c65a0119</td>
<td></td>
</tr>
<tr>
<td>abmf</td>
<td>TRIMBLE NETR5</td>
<td>TRM57971.00</td>
<td>NONE</td>
<td>0.0000</td>
<td>ABMF</td>
<td>97103M001 M</td>
<td>2.11</td>
<td>4242 ac7e0b95b0a08a5c0dca5e9a96e9f7</td>
<td></td>
</tr>
<tr>
<td>abpo</td>
<td>ASHTECH Z-XII-12</td>
<td>ASH701945M_M</td>
<td>SCIT</td>
<td>0.0083</td>
<td>ABPO</td>
<td>33028M001 G</td>
<td>2.11</td>
<td>73 18e06dd22589b0a1bce6537bc56ed6d</td>
<td></td>
</tr>
<tr>
<td>aisa</td>
<td>LEICA GR25</td>
<td>TRM579800.00</td>
<td>SCIS</td>
<td>0.0000</td>
<td>AISA</td>
<td>21742M001 G</td>
<td>2.11</td>
<td>100 018266f66edde28865a337e5f9a0e6a</td>
<td></td>
</tr>
<tr>
<td>airc</td>
<td>AOA BENCHMARK ACT</td>
<td>AOAD/M_T</td>
<td>SCIS 1.0000</td>
<td>AIRC</td>
<td>10077M005 G</td>
<td>2.11</td>
<td>69 39b21a6e96b6fdec786a1f02967f1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>albh</td>
<td>TRIMBLE NETR9</td>
<td>TRM59800.00</td>
<td>SCIS</td>
<td>3.9650</td>
<td>ALBH</td>
<td>13231M001 M</td>
<td>2.11</td>
<td>9430 64de0c941419e4e7e47f2589c65a0119</td>
<td></td>
</tr>
<tr>
<td>alh2</td>
<td>AOA BENCHMARK ACT</td>
<td>AOAD/M_T</td>
<td>SCIS 1.0000</td>
<td>ALC2</td>
<td>10077M005 G</td>
<td>2.11</td>
<td>69 39b21a6e96b6fdec786a1f02967f1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alr2</td>
<td>LEICA GRX1200GGPRO</td>
<td>LEIAR25.R3</td>
<td>NONE</td>
<td>0.0015</td>
<td>ALIC</td>
<td>50137M001 M</td>
<td>2.11</td>
<td>79 71e055d1a59afec636aa193936d20</td>
<td></td>
</tr>
<tr>
<td>alm2</td>
<td>LEICA GR25</td>
<td>TRM57971.00</td>
<td>NONE</td>
<td>0.0000</td>
<td>ALMA</td>
<td>50137M001 M</td>
<td>2.11</td>
<td>79 71e055d1a59afec636aa193936d20</td>
<td></td>
</tr>
<tr>
<td>am2</td>
<td>ASHTECH Z-XII-12</td>
<td>ASH701945G_M</td>
<td>SCIT</td>
<td>0.0000</td>
<td>AM2</td>
<td>60400M002 G</td>
<td>2.11</td>
<td>84 69f50126ca2cd9329ef36a5966ec3</td>
<td></td>
</tr>
<tr>
<td>amn2</td>
<td>AOA BENCHMARK ACT</td>
<td>AOAD/M_T</td>
<td>SCIS 1.0000</td>
<td>AMN2</td>
<td>10077M005 G</td>
<td>2.11</td>
<td>69 39b21a6e96b6fdec786a1f02967f1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>antc</td>
<td>TPS E_GGD</td>
<td>TPSCR3_GGD</td>
<td>CONE</td>
<td>0.0700</td>
<td>ANTC</td>
<td>20805M002 M</td>
<td>2.11</td>
<td>105 f571136cdd0d87504bbcdd634f4beadb</td>
<td></td>
</tr>
<tr>
<td>arqc</td>
<td>LEICA GRX1200GGPRO</td>
<td>LEIAR25.R3</td>
<td>NONE</td>
<td>0.0015</td>
<td>ARTQ</td>
<td>50137M001 M</td>
<td>2.11</td>
<td>79 71e055d1a59afec636aa193936d20</td>
<td></td>
</tr>
<tr>
<td>arqc</td>
<td>LEICA GRX1200GGPRO</td>
<td>LEIAR25.R3</td>
<td>NONE</td>
<td>0.0015</td>
<td>ARTQ</td>
<td>50137M001 M</td>
<td>2.11</td>
<td>79 71e055d1a59afec636aa193936d20</td>
<td></td>
</tr>
<tr>
<td>aten</td>
<td>TRIMBLE NETR9</td>
<td>TRM59800.00</td>
<td>SCIS</td>
<td>3.9650</td>
<td>ATEN</td>
<td>13231M001 M</td>
<td>2.11</td>
<td>9430 64de0c941419e4e7e47f2589c65a0119</td>
<td></td>
</tr>
<tr>
<td>artu</td>
<td>ASHTECH Z-XII-12</td>
<td>ASH701945C_M</td>
<td>SCIS</td>
<td>0.0792</td>
<td>ARTU</td>
<td>12321M002 G</td>
<td>2.11</td>
<td>73 6744aab22bfad3938e05bbf5b42cc9d</td>
<td></td>
</tr>
<tr>
<td>aspa</td>
<td>TRIMBLE NETR5</td>
<td>TRM55971.00</td>
<td>NONE</td>
<td>0.0000</td>
<td>ASPA</td>
<td>50503M006 M</td>
<td>2.11</td>
<td>88 128e5af0e7743045f52754d5fd3f3bc4</td>
<td></td>
</tr>
<tr>
<td>axc2</td>
<td>ASHTECH UZ-12</td>
<td>ASH701945B_M</td>
<td>SCIT</td>
<td>0.0000</td>
<td>AXC2</td>
<td>50299M001 M</td>
<td>2.11</td>
<td>457 5d231532b55212ba831e267634562b</td>
<td></td>
</tr>
<tr>
<td>axtc</td>
<td>LEICA GRX1200GGPRO</td>
<td>LEIAR25.R3</td>
<td>NONE</td>
<td>0.0015</td>
<td>AXTC</td>
<td>50137M001 M</td>
<td>2.11</td>
<td>79 71e055d1a59afec636aa193936d20</td>
<td></td>
</tr>
<tr>
<td>axtu</td>
<td>LEICA GRX1200GGPRO</td>
<td>LEIAR25.R3</td>
<td>NONE</td>
<td>0.0015</td>
<td>AXTU</td>
<td>50137M001 M</td>
<td>2.11</td>
<td>79 71e055d1a59afec636aa193936d20</td>
<td></td>
</tr>
</tbody>
</table>

IGS Workshop 2014

IGS Data Centers: Status and Progress

June 25, 2014 | 11
• Increasingly, data centers are storing site, file and status metadata in databases (example, IGS Central Bureau Site Log Manager)

• With metadata stored in a database, GSAC (Geodesy Seamless Archive Centers) software then allows access via web services to useful metadata and to data

• GSAC in federated mode allows additional useful queries such as tracking of status of file delivery across multiple data centers

• CDDIS, SOPAC, and UNAVCO are providing GSAC services.

• Several European geodetic data centers are also trying GSAC (ROB, INGV, SEGAL, and others)
20 years ago – text based site logs stored station metadata which was shared via file transfer (site log, SINEX, station.info, RINEX headers)

2014 – metadata is often stored and managed in databases but sharing is still largely by file transfer in same formats as 20 years ago

Management of metadata in files is highly inefficient and error prone

Migration to use of web services (GSAC or other), and away from metadata in files, should be considered as a way to efficiently transfer the best metadata available

The format of the metadata output by web services is flexible i.e., can simulate familiar file formats
• Address issues relevant to effective operation of all IGS data centers

• Meeting:
  – When: Splinter Session 4
    Thursday, June 26
    16:00-17:30 (concurrent with Antenna and TIGA WG meetings)
  – Where: Noyes 147