

DATA HANDLING WITHIN THE INTERNATIONAL VLBI SERVICE

Dirk Behrend

NVI, Inc./NASA Goddard Space Flight Center, Code 698.2, Greenbelt, MD 20771, USA

Email: dirk.behrend@nasa.gov

ABSTRACT

The International VLBI Service for Geodesy and Astrometry (IVS) is a globally operating service that coordinates and performs Very Long Baseline Interferometry (VLBI) activities through its constituent components. The VLBI activities are associated with the creation, provision, dissemination, and archiving of relevant VLBI data and products. The data and products are stored in dedicated IVS components called 'Data Centers.' The three Primary Data Centers provide identical data holdings. We give a brief overview of the organizational structure of the IVS and describe the general data flow among the various IVS components from preparing observational plans to creating the final products.

Keywords: Very Long Baseline Interferometry, VLBI, International VLBI Service, IVS, Space geodesy, Data center, Data, Products

1 INTRODUCTION

Very Long Baseline Interferometry (VLBI) is one of the most accurate methods used to measure the Earth and its orientation in space. It is one of four space-geodetic techniques, the others being SLR (Satellite Laser Ranging), GNSS (Global Navigation Satellite Systems), and DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite), which are used to determine the celestial and terrestrial reference frames, the Earth orientation parameters (EOP), atmospheric parameters as well as other ancillary parameters. The EOP parameters are precession/nutation, Earth rotation (UT1), and polar motion. Each space-geodetic technique has its own strengths and unique capabilities. VLBI is unique in its ability to measure precession/nutation and UT1. VLBI employs large radio telescopes to observe compact radio sources, usually quasars, in order to estimate the vector between the telescopes. The VLBI observable is the difference in arrival time of a radio signal at two (or more) telescopes; hence, VLBI requires at least two radio telescopes to furnish useful observations. The VLBI technique dates back to the late 1960s; high-precision data, however, have been collected from the mid-1980s onward.

In this paper we concentrate on the data aspect of the VLBI technique. After a brief overview of the international organizational structure, we describe the general data flow among the various VLBI components and then take a closer look at the VLBI data repositories. The technical and scientific aspects of VLBI are covered elsewhere; the interested reader is referred, for instance, to Sovers, Fanselow, and Jacobs (1998) and Schuh and Behrend (2012) as well as references therein.

2 INTERNATIONAL VLBI SERVICE FOR GEODESY AND ASTROMETRY

Geodetic/astrometric VLBI activities on global and regional scales are organized through the International VLBI Service for Geodesy and Astrometry (IVS), which is an international collaboration of institutions that operate or support VLBI components. The IVS was established in 1999 as a service of the International Association of Geodesy (IAG), recognizing the need to move away from the *ad hoc* basis of the VLBI operational activities, which were mostly organized through national or bi-lateral agreements until then (see, e.g., Schlüter & Behrend, 2007). In 2000, IVS became a service of the International Astronomical Union (IAU) and the Federation of Astronomical and Geophysical Data Analysis Services (FAGS). Following the dissolution of the FAGS federation at the end of 2008, IVS applied for membership in the newly established World Data System (WDS) and was approved by the end of 2011. IVS interacts closely with the International Earth Rotation and Reference Systems Service (IERS), which is tasked by IAU and IUGG (International Union of Geodesy and Geophysics) with maintaining the international celestial and terrestrial reference frames (ICRF and ITRF). The IVS currently

consists of around 80 permanent components, which are supported by about 40 organizations in 20 countries. The IVS permanent component types are: Network Stations, Operation Centers, Correlators, Data Centers, Analysis Centers, Technology Development Centers, and the Coordinating Center (see Figure 1). Up-to-date information about the service and its activities can be found online under the URL <http://ivscc.gsfc.nasa.gov>.

The mission objectives of the IVS include the provision of support for geodetic, geophysical, and astrometric research and operational activities as well as the integration of VLBI into a global Earth observing system. To meet these objectives, IVS coordinates VLBI observing programs, sets performance standards for VLBI stations, establishes conventions for VLBI data formats and data products, issues recommendations for VLBI data analysis software, sets standards for VLBI analysis documentation, and institutes appropriate VLBI product delivery methods to ensure suitable product quality and timeliness. The VLBI products currently available include the five EOP parameters, the TRF, the CRF, and tropospheric parameters. All VLBI data and products are archived in IVS Data Centers and are publicly available. The IVS data set extends from 1979.

3 GENERAL DATA FLOW

The general flow of VLBI data within the IVS is centered around a “data feedback loop” as depicted in Figure 1. Raw VLBI data are recorded by the IVS Network Stations. The IVS observational network currently consists of 40–45 radio telescopes worldwide. Subsets of these telescopes participate in 24-hour observing sessions (8–10 stations) that are run several times per week and in 1-hour intensive sessions (2–3 stations) for UT1 determination every day. The individual observing networks are planned ahead for an entire calendar year by the Coordinating Center in the so-called Master Schedule. About one to two weeks prior to the actual observation date, an Operation Center prepares the individual recording schedule of the session and uploads it to a Data Center. This schedule is subsequently downloaded by the stations and the Correlator that processes the data.

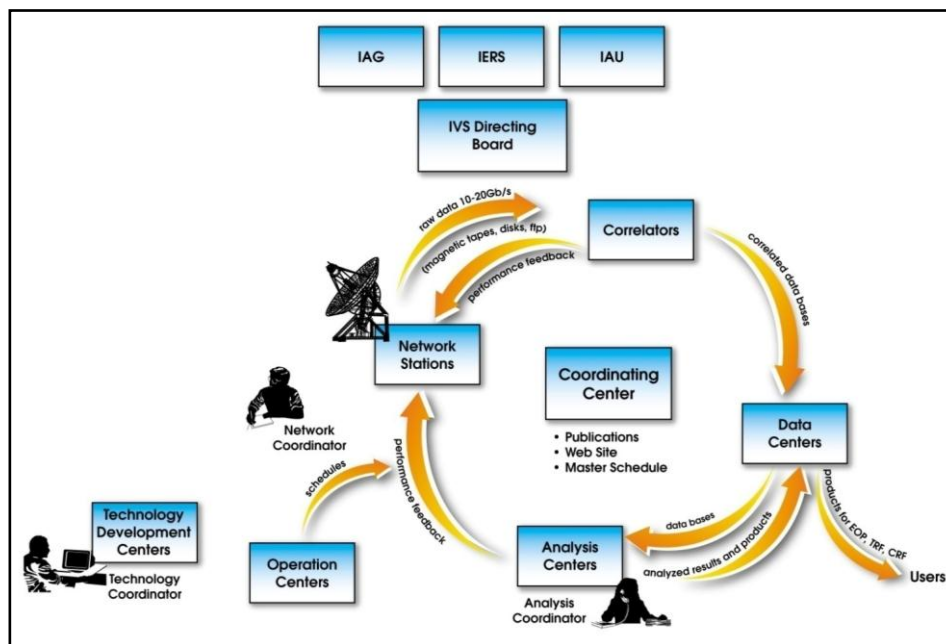


Figure 1. Organization of the IVS and general data flow between the various component types of the service

The raw VLBI data are recorded on storage media and then shipped to the pre-determined Correlator. With about 1–2 TB of raw VLBI data per station per 24-hour session, the entire session occupies up to 15–20 TB of storage space. In the correlation process the Correlator reduces the raw data to the ‘VLBI observables plus metadata’ in the so-called database (*db*) format, a binary format of about 1–2 MB size per 24-hour session. The database is uploaded to the Data Center; the raw VLBI data are erased and the storage media are prepared to record the next VLBI session. The *db* format was created in the mid-1970s for use with a specific VLBI analysis software. It has been the *de facto* standard for archiving and distributing geodetic/astrometric VLBI sessions until now. The IVS is currently working on the design and implementation of a new VLBI data structure, which is based on the NetCDF data storage format and uses modularization and wrapping techniques (Gipson, 2008; Gipson 2010).

Analysis Centers download the databases from the Data Centers, analyze the VLBI data, and upload the analyzed results and products (e.g., EOP, CRF, TRF) to the Data Centers. Special Analysis Centers called Combination Centers use the results of the individual Analysis Centers to prepare the final IVS products. The combined solution is more accurate and robust as compared to the individual solutions of the Analysis Centers. Users can download the final products (either the combination solution or the individual Analysis Center solutions) from the Data Centers. Finally, in order to close the loop, the Correlators and Analysis Centers provide feedback to the Network Stations on their performance and suggest possible improvements.

4 DATA CENTERS

The Data Centers play an essential role within the IVS as the repositories of all geodetic/astrometric VLBI data and products. In addition to the data and products mentioned in Section 3, the Data Centers also archive auxiliary information such as station log files, correlator reports, and documentation files. As the radio telescopes constitute a very large financial investment, it is important that none of the collected data gets lost or unusable. Furthermore, users should be able to access the data and products efficiently and reliably. For that, the IVS is supported by three Primary Data Centers: CDDIS (Crustal Dynamics Data Information System), BKG Leipzig (Bundesamt für Kartographie und Geodäsie), and OPAR (Observatoire de Paris). The three Primary Data Centers mirror their data holdings several times per day in a predetermined scheme in pairs of data centers (Figure 2).

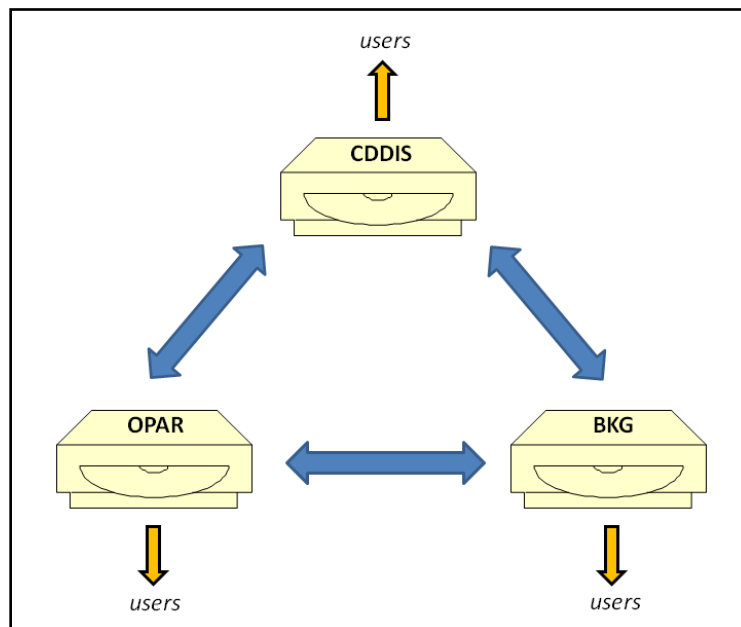


Figure 2. Mirroring scheme between the three IVS Primary Data Centers at CDDIS, OPAR, and BKG.

There are basically two ways the Primary Data Centers update their holdings. First, data and product files can be uploaded to a special incoming area of an individual Primary Data Center by an IVS component (e.g., Correlator or Analysis Center) using authenticated FTP from a registered IP address. Automated scripts check whether the names of the incoming files are either registered with the Coordinating Center in special code files or can be constructed from the information contained in the Master Schedule and, to a limited extent, whether the files comply with the expected data structure. If the tests are successful, the scripts relocate the incoming files to their appropriate archive directories. Secondly, new data and products are added in the mirroring process. Users are only allowed to retrieve files from the Data Centers using anonymous FTP; they are not allowed to upload files. The basic directory structure is identical for all Primary Data Centers. A detailed description of this structure is, for instance, given in Noll (2010); although this reference describes CDDIS only, the VLBI part is directly transferable to BKG and OPAR.

The main geodetic VLBI data and products available at the Data Centers are summarized in Table 1. Auxiliary data files such as schedule files, station log files, or correlator reports are not included in the list, since they mostly support VLBI operations and are of marginal relevance for most scientific applications. They may be considered metadata. All data and products are freely available at no cost to the user. The earliest VLBI data and products date back to 1979.

Table 1. Summary of the major geodetic VLBI data and products available in the three Primary Data Centers (from Noll, 2010)

Data set	Processing level	Granule	Time span
Correlated experiment databases	Data	Daily	1979–date
Baselines	Derived product	Daily	1979–date
EOP (all)	Derived product	Daily	1979–date
Station positions and velocities (TRF)	Derived product	Daily	1979–date
Source positions (CRF)	Derived product	Daily	1979–date
Zenith tropospheric wet delay	Derived product	Weekly	2002–date

Noll (2010) reported the usage statistics of CDDIS for the year 2008. By volume, VLBI accounted for about 2% of archive downloads, similar to SLR (also 2%) and DORIS (1%), whereas GNSS accounted for the major bulk of the downloads (95%). About 400 user organizations accessed CDDIS on a regular basis in 2008 in order to retrieve VLBI-related files. They downloaded over 13.5 TB of data and products in about 640,000 files. The users recruited from government agencies (53%), educational institutions (29%), networks (15%), commercial companies (2%), and some miscellaneous groups (1%). The other two Primary Data Centers are expected to have similar usage numbers.

5 CONCLUSIONS

The IVS was formed at the end of the last century to serve the scientific community in the fields of geophysics and astrometry. The service brought under one roof the operational activities and standardized the data flow from capturing the raw VLBI data, to correlating and preparing databases, to creating the final VLBI products. Various component types specialize in given aspects of the VLBI technique. The task of archiving and distributing the VLBI data and products rests with the IVS Primary Data Centers.

The concept of having three Data Centers with identical data holdings by means of a daily mirroring process has proven to be very successful. It ensures continuous and reliable access to the data without possible disruptions through maintenance work or IT security issues. The growing size of the repositories necessitates that the computing facilities be upgraded in terms of data capacity and speed. A challenge for the future will be the integration of near real-time data generation into the overall data flow.

6 REFERENCES

- Gipson, J. (2008) IVS Working Group 4: VLBI Data Structures. In: Finkelstein, A., Behrend, D. (Eds.), *Measuring the Future*. Saint Petersburg, Russia, Nauka, ISBN 978-5-02-025332-2, pp 143-152.
- Gipson, J. (2010) IVS Working Group 4: VLBI Data Structures. In: Behrend, D., Baver, K. (Eds.), *IVS 2010 General Meeting Proceedings*, NASA/CP-2010-215864, Greenbelt, MD, USA, pp 187-191.
- Noll, C.E. (2010) The crustal dynamics data information system: A resource to support scientific analysis using space geodesy. *Adv. Space Res.* 45(12), pp 1421-1440.
- Schlüter, W. & Behrend, D. (2007) The International VLBI Service for Geodesy and Astrometry (IVS): current capabilities and future prospects. *J. Geod.* 81(6-8), pp 379-387.
- Schuh, H. & Behrend, D. (2012) VLBI: A Fascinating Technique for Geodesy and Astrometry. *J. Geodyn.* 61(2012), pp 68-80.
- Sovers, O.J., Fanselow, J.L., & Jacobs, C.S. (1998) Astrometry and geodesy with radio interferometry: experiments, models, results. *Rev. Mod. Phys.* 70(4), pp 1393-1454.

(Article history: Available online 17 February 2013)