

BKG Data Center

Anastasiia Walenta, Markus Goltz

Abstract In 2021–2022 we supported as-identical-as-possible data handling in the triad of the primary IVS Data Centers. We employ the same validation procedures and data definition files (DDFs) to accept new data as well as to keep the old data. As the storage demands of new VLBI data types, i.e. vgosDb and corresponding VGOS station log files and Level 1 data files in Swinburne format (also referred to as SWIN files), grow quickly, we address this issue by allocating disk space. Due to security reasons, the access to our server has been restricted. The File Transfer Protocol (FTP) over Secure Socket Layer (SSL) and the Hypertext Transfer Protocol Secure (HTTPS) were set up for all our users in the middle of 2022. In addition to the service for IVS, the BKG Data Center provides an exchange server for projects of relatively small disk space demands: PWLO and EU-VGOS. These projects are powered by the same ingest procedures which we have in place for the IVS. Thus, an IVS-like structure is arranged. As follows, the projects can be integrated into the official IVS structure, if it will be desired in the future. The projects are fully independent at the same time, which lets us test different data handling procedures.

1 General Information

The BKG Data Center (BKG DC) is hosted by the German Federal Agency for Cartography and Geodesy

Federal Agency for Cartography and Geodesy (BKG)

BKG Data Center

IVS 2021+2022 Biennial Report

(BKG) as one of the three primary IVS Data Centers on terms as defined by the International VLBI Service for Geodesy and Astrometry (IVS).

According to the IVS regulations we are responsible for accepting the recognized VLBI data from all IVS components—other Data Centers, Operation Centers, Network Stations, Correlators, Analysis Centers, and Combination Centers—and provide them in open access. At this time, access to our server is restricted to File Transfer Protocol (FTP) over Secure Socket Layer (SSL) and Hypertext Transfer Protocol Secure (HTTPS). The principle of the open access is empowered by anonymous downloads over both protocols. The IVS users are offered to receive an account as described at our web page (<https://ivs.bkg.bund.de/>) to upload data. To achieve a data set which is identical with the data sets of the other official IVS DCs, the mirror procedures are set up.

2 Data Center at BKG

The BKG Data Center works together with CDDIS and OPAR representatives to support the proper data ingest among the official IVS DCs [2]. In this report we highlight the major items in the operations of the BKG DC in the last two years.

- **Establishing common validation procedures, before August 2021**

The internal BKG DC structure was re-organized significantly: the development and production servers are arranged on the new machines instead of the old server. The internal testing of the ingest script was carried out on the development server for the entire data set, so that many issues could

be addressed without disturbing the Data Center routines. The work on the ingest script was finalized by establishing the script as the main routine to accept data. The common validation procedures are shared via GitHub. The DDFs are maintained by the IVS Coordinator.

- **After the switch, August 2021**

The received data has been recognized according to the Name Convention since August 2, 2021. As follows, many routinely uploaded data were moved to the designated directory ‘unknown’ for the files rejected by these procedures at the BKG DC. In some cases the validation routines had to be adjusted. In the other cases, our team was in personal contact with the users who submit the data and products as a part of an operational service to fix the issues.

- **Level 1 data, 2022**

The raw VLBI observations are referred to as Level 0 data; then these data are correlated, producing visibilities called Level 1 data. In the IVS community these data are distributed with the so-called SWIN files, short for Swinburne format to acknowledge Swinburne University of Technology, where the DiFX correlator was developed and SWIN files were created initially. The applications of the incomparably more descriptive than the VLBI group delays (vgosDb) SWIN files are expected to be exploited far beyond the geodetic IVS community. Notwithstanding the large size of the files, which exceeds extensively any other data type at the IVS DCs, the SWIN files are validated with the ingest script as any other data type. The only difference with respect to other data types is that the SWIN files are received in a separate incoming area, where a separate anti-virus check is performed because of the file size. We have established the contact to receive SWIN files from the correlator in Shanghai. SWIN files are also mirrored from CDDIS, which completes the list of all available Level 1 data at the BKG DC.

- **EU-VGOS and PWLO projects, 2021–2022**

The requirements of both projects, i.e., EU-VGOS and PWLO ([1] and [5]), were met successfully as reported at the IVS 2022 General Meeting [4]. We have faced several technical problems, the solution of which helped us to support the IVS DC in the transition from FTP to the secure protocols. The work on the projects has helped us to understand better the validation procedures of the ingest script.

- **FTP over SSL and HTTPS, Summer 2022**

The IVS users were informed about our ultimate decision to discontinue the unsecured FTP access. The ivsincoming user was told to support the transitions, so that users in a first step would adjust their routines to access with FTP over SSL. And in a following step, users were offered to obtain an account at the BKG DC in order to download and upload data according to the new procedures. The data download is also possible with HTTPS (path to the BKG DC root directory: https://ivs.bkg.bund.de/data_dir/vlbi/).

- **Renewal of the website, Summer 2022**

The BKG DC has an old website with some general information about our Analysis and Data Center team and their main tasks. The website has since been updated significantly (<https://ivs.bkg.bund.de/>). The current team members and activities of our two IVS Components, i.e., Analysis and Data Centers, are presented for historical reasons, as they belong to the same section G1 at BKG and work closely together. A brief description of the Combination Center, the third IVS component operated within section G1 at BKG, is given for the time being while the corresponding website is under maintenance. The BKG DC exercises the ability to make public the regulations on how to access and upload the data at our Data Center. The graphical representation of the user access statistics (Figure 1) is posted on a daily basis. The effect of switching to the secure protocols can be seen for uploads as well as for downloads. The amount of downloads dropped at first, but after a while, an approximately previous level through the sum of FTPS and HTTPS downloads was reached.

- **Dealing with an upload of a new complete solution (SINEX files) to the Data Center, 2021–2022**

A complete VLBI solution of Intensives or 24-hour sessions includes about 7,600 and 6,000 files, respectively. If it is uploaded all at once, it might look like a DDoS attack. Then the ingest script will attempt to ingest the massive amount of files, which will put the operational uploads of the other users on “pause” due to the absence of a prioritized selection. Moreover, many files of the new solution might be rejected and moved to the unknown area. The files need to be managed manually in the unknown directory while the incorrect file names

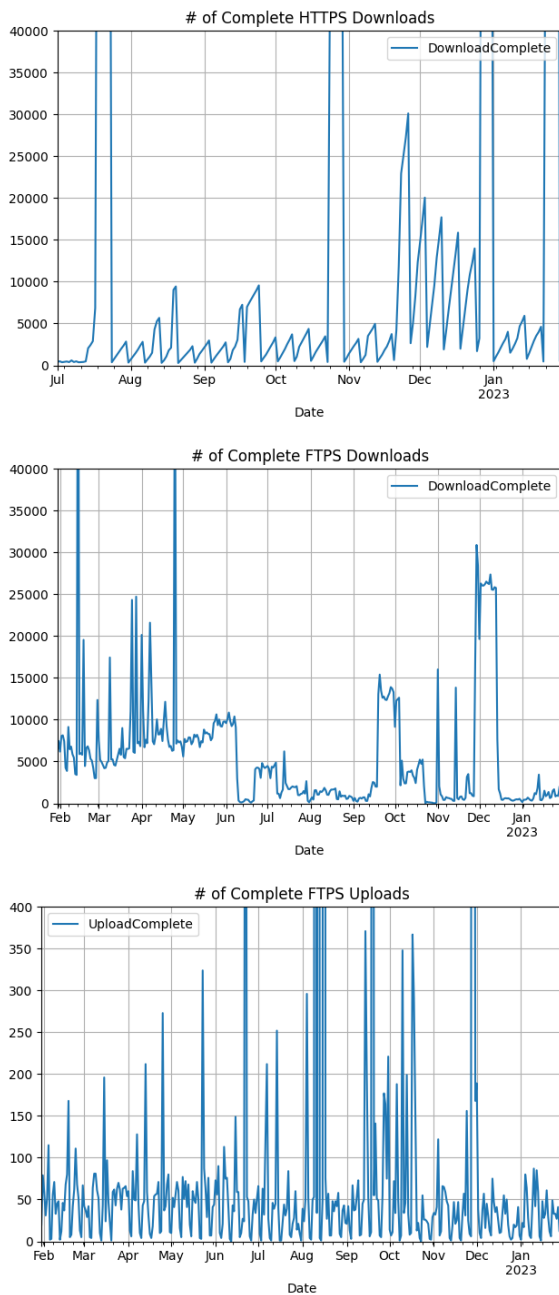


Fig. 1 User activity statistics (from top to bottom) represent the number (#) of successful downloads using HTTPS and FTPS respectively and the number of uploads via FTPS client.

are considered generally as exceptions. Multiple USNO submissions to the Data Centers have triggered CDDIS and BKG feedback, based on which the USNO group has made the exceptional effort

to facilitate the Data Centers' work. The key to the solution of a considerable data uploading is to split the files in chunks and upload them irregularly in time during a few hours. This issue might become more important, when, for instance, the switch to ITRF2020 will be effective: every AC will deliver new solutions and increase vastly the amount of incoming data.

- **Master file format version 2 and EOP 3.0 format, the very end of 2022**

The validation procedures of the ingest script as well as the script itself were adjusted to the requirements of the master file format 2 and EOP 3.0 format. In particular, the internal ingest database was modified. At this moment the EOP-I and EOP-S files in EOP 3.0 format are received by BKG DC for test purposes.

3 Personnel

- Markus Goltz:
[markus.goltz@bkg.bund.de]
- Anastasiia Walenta:
[anastasiia.walenta@bkg.bund.de]

Both team members are involved in the regular representative meetings of the official IVS Data Centers. Markus Goltz is engaged mainly in the technical Data Center support: maintenance of the server including close work with BKG internal IT support, documentation of the internal procedures, ingest support, graphical presentation of the data statistics, and web-site support. Anastasiia Walenta mainly covers the ingest support: check-ups, special cases, ingest of the new data formats, and internal documentation.

We support our users and every IVS member who wants to work with us. The BKG DC Team can be reached by writing to our mailbox: vlbi@bkg.bund.de.

4 Future Plans

For the next two years we intend to focus on covering some remaining flows in the Data Center ingest procedures. The main concern is aimed at the proper ingesting of vgosDb. Also, we will work on turning our services to being more user-oriented and user-friendly,

specifically: establishing an automatic feedback, a development of appropriate web services, and data uploading via HTTPS. We expect to increase gradually the Data Center disk space to meet VLBI Level 1 data requirements. As the Data Center efforts are significantly higher to support SWIN files, the number of users and the cadence of their access to these data will be considered, to advocate the further support of these data at the Data Center.

References

1. E. Albentosa, F. Jaron, et al., Current Status of the EU-VGOS Project, IVS 2022 General Meeting Proceedings, edited by Kyla L. Armstrong, Dirk Behrend, and Karen D. Baver, NASA/CP-20220018789, 2023.
2. D. Behrend, M. Bérubé, J. Gipson, A. Girdiuk, M. Goltz, T. Yates, P. Michael, C. Barache, Data Ingest at the IVS Data Centers, IVS 2022 General Meeting oral presentation, Finnish Cyberspace, 2022. https://space-geodesy.nasa.gov/docs/2022/GM2022_behrend_ingest.pdf
3. D. Behrend, Data Center file conventions, https://ivscc.bkg.bund.de/products-data/DataCenter_File_Conventions.pdf
4. A. Girdiuk, M. Goltz, D. Thaller, IVS Data Center at BKG, IVS 2022 General Meeting Proceedings, edited by Kyla L. Armstrong, Dirk Behrend, and Karen D. Baver, NASA/CP-20220018789, 2023.
5. A. Nothnagel, S. Böhm, R. Dach, M. Glomsda, H. Hellmers, A.-S. Kirkvik, T. Nilsson, A. Girdiuk, D. Thaller, First results of project on six-hourly EOP piecewise linear offset parameterization, IVS 2022 General Meeting Proceedings, edited by Kyla L. Armstrong, Dirk Behrend, and Karen D. Baver, NASA/CP-20220018789, 2023.