Transition to the vgosDb Format

Sergei Bolotin, Karen Baver, John Gipson, David Gordon, Daniel MacMillan

Abstract The IVS Working Group 4 developed a new format to store and exchange data obtained from geodetic VLBI observations. The new data format, vgosDb, will replace existing Mark III databases this year. At GSFC we developed utilities that implement the vgosDb format and will be used routinely to convert correlator output to the new data storage format.

Keywords VLBI data analysis software, vgosDb

1 Introduction

Data produced at a correlator are subject to various changes before they become available to an end user. Historically, the results of correlation and fringe fitting of VLBI observations are stored in a binary self-descriptive file called a *database*. The format of the database file and implementation of input/output operations were developed in the early 1970s. Since then, the databases have been used as a standard for data exchange in the geodetic VLBI community.

The database format has disadvantages, mostly caused by hardware and software limitations that existed in the period when the format was developed. In addition, the format was not well documented. Several attempts to replace the database format with an alternative were made in the last few decades, but none of these were successful. The anticipated vast increase in the number of VLBI observations and the emergence of VLBI Global Observing System (VGOS) technology prompted the IVS Directing Board to establish the IVS Working Group on Data Structures. Efforts undertaken by the group were eventually realized in the creation of the new VLBI data format, *vgosDb* (see [4]).

In accordance with the vgosDb format, the VLBI data of one session are stored in various files in the form of $\{key \Rightarrow value\}$. Each file represents an atomic piece of data, e.g., observed values with their standard deviations, or station coordinates. An additional feature is that it is possible to keep alternative models or approaches to the editing of observations in the same session data tree. A set of data files that is available to the user is specified in a special file called a *wrapper file*. It is possible to have more than one wrapper file for one VLBI session.

2 vgosDb-compatible VLBI Data Analysis Software

First results of the implementation of the vgosDb format by the GSFC VLBI group were shown in 2013. The legacy VLBI data analysis software, *global solve*, is ready to use data in the vgosDb format [5]. A part of the *solve* distribution package is the utility db2vgosDB, which converts data of a VLBI session from the database format into the vgosDb format. The next generation VLBI data analysis software, *νSolve*, is capable of working with the new format [2].

In addition, vgosDb-compatible utilities were developed to support the transition to the vgosDb format. These utilities replace the legacy utilities dbedit and pwxcb/dbcal. The first utility, dbedit, creates database files from correlator output and fringe files. Routinely,

NVI, Inc.

it is executed at a correlator, and the database file for each band is then sent to one of the VLBI data centers. The purpose of the second pair of utilities, pwxcb and dbcal, is to extract, validate, edit (if necessary), and put into a database the information that is contained in the Field System log files produced by each station that participated in a VLBI session. Usually, two types of data are extracted: cable calibration measurements and meteorological parameters.

The new vgosDb-compatible utilities, vgosDbMake (which replaces dbedit) and vgosDbProcLog (which replaces pwxcb/dbcal), are part of the new VLBI data analysis software developed at NASA GSFC [1] and currently distributed in one package under the common name "nusolve". It is available on the FTP site:

`ftp://gemini.gsfc.nasa.gov/pub/misc/slb/nusolve-latest.tar.gz`

These utilities have the same design as other *v*Solve software as well as the same software development environment.

The utilities are designed to operate on any POSIX compatible operating system. We use C++ as the programming language due to its power, flexibility, and portability. The GNU Build System is used to make the software distribution portable. The software consists of two parts:

- Space geodesy library: a library where data structures and algorithms are implemented (about 90% of the total source code).
- Executables vgosDbMake and vgosDbProcLog: drivers that call library functions and organize work with an end user (about 10% of the total source code).

Such organization of the software allows us to share the source code between applications and reuse it in other projects.

The software has a modular structure that makes it flexible and scalable. A module is a logical block of code that is loosely tied with other parts of the software.

Obviously, not all the modules will be used by vgosDbMake and vgosDbProcLog. On the other hand, the modular design of the software allows us to easily add the functionality of the utilities to the interactive VLBI data editor, *v*Solve.

Modification of the program *calc* to be compatible with vgosDb is done in a different way. A library that mimics the Mark III database handler programming interface has been created. The library replaces database functions with vgosDb input/output operations. In this case we do not need to modify the *calc* source code at all, but just need to link the software with the new library. This approach was realized by Julia Ringsby, an NVI intern in 2013 [6], in the program vgosDbCalc. It was written in the C++ programming language, but its design is different from the design of the *vSolve* software. We use vgosDbCalc as a prototype for demonstration of the vgosDb format usage, but it will be replaced with a utility based on our common design for the new VLBI data analysis software.

The software vgosDbCalc is distributed in a separate package and can be downloaded from the FTP site

`ftp://gemini.gsfc.nasa.gov/pub/misc/slb/vgosDbCalc.tar.gz`

The distribution contains instructions on how to compile the source code and to run the utility. A collection of necessary files with *a priori* data is also in the distribution. It should be noted that one of these files, the table of the Earth rotation parameters, needs to be updated on a regular basis.

3 Use of the VLBI Data in the vgosDb Format

The vgosDb format is used in routine data analysis utilizing the set of the three vgosDb compatible software packages: solve, vgosDbCalc, and *v*Solve.

Before switching to the new vgosDb format we made the VLBI observations in this format available for public access in two groups of files. The first is the data that are distributed by the official IVS ftp sites. The vgosDb files for this set can be downloaded from

`ftp://gemini.gsfc.nasa.gov/pub/vgosDB_IVS/`

The second collection of the vgosDb files corresponds to the GSFC-analyzed VLBI observations. It can be obtained from

`ftp://gemini.gsfc.nasa.gov/pub/vgosDB_GSFC/`

The main difference between the two data sets is in the data editing options. These publicly accessible data will help users to transition to the new format.

The first practical use of the VLBI observations in the vgosDb format was reported by MIT Haystack Observatory [7]. All VGOS-related observations since January 2016 were analyzed with vgosDb data flow. These observations include broadband VGOS sessions with up to three stations and S/X sessions using mixed Mark IV and broadband stations.

4 Conclusions

Our group will switch to the new VLBI data format in mid-2016. At the time of writing this paper we are performing extensive testing of the software, the legacy solve package, and new utilities from the nusolve distribution. In July 2016, we will perform tests of the entire VLBI data flow using the vgosDb format and then start to release VLBI sessions in the new format routinely. We strongly encourage all Analysis Centers to switch to the new VLBI data format.

Acknowledgements

We are grateful to Tim Morin, Mike Titus, and Arthur Niell of the MIT Haystack Observatory for installing, testing, and helping debug the vgosDb format.

References

1. S. Bolotin, J.M. Gipson and D. MacMillan. *Development of a New VLBI Data Analysis Software*. In D. Behrend, and K.D. Baver, editors, *IVS 2010 General Meeting Proceedings*, NASA/CP-2010-215864, NASA GSFC, Maryland, pages 197–201, 2010.
2. S. Bolotin, K. Baver, J.M. Gipson, D. Gordon and D. MacMillan. *The VLBI data analysis software νSolve: development progress and plans for the future*. In D. Behrend, K.D. Baver, and K. Armstrong, editors, *IVS 2014 General Meeting Proceedings*, ISBN 978-7-03-042974-2, Science Press, Beijing, China, pages 253–257, 2014.
3. J.M. Gipson. *IVS Working Group 4: VLBI Data Structures*. In D. Behrend, and K.D. Baver, editors, *IVS 2012 General Meeting Proceedings*, NASA/CP-2012-217504, NASA GSFC, Maryland, pages 212–221, 2012.
4. J. Gipson, J. Böhm, S. Bolotin, D. Gordon, T. Hobiger, C. Jacobs, S. Kurdubov, A. Nothnagel, O. Sovers, O. Titov, H. Takiguchi. *Final Report of IVS Working Group 4 (WG4) on Data Structures*. In K.D. Baver, D. Behrend, and K. Armstrong, editors, *International VLBI Service for Geodesy and Astrometry 2013 Annual Report*, NASA/TP-2014-217522, NASA GSFC, Maryland, pages 11–25, 2014.
5. J. Gipson. *Practical Uses of VGOSDB Format*. In R. Haas and F. Colomer, editors, *Proceedings of the 22nd European VLBI Group for Geodesy and Astrometry Working Meeting*. ISBN 978-989-20-6191-7, pages 97–101.
6. D. Gordon, C. Ma, D. MacMillan, J. Gipson, S. Bolotin, K. Le Bail, K. Baver. *GSFC VLBI Analysis Center Report*. In K.D. Baver, D. Behrend, and K. Armstrong, editors, *International VLBI Service for Geodesy and Astrometry 2013 Annual Report*, NASA/TP-2014-217522, NASA GSFC, Maryland, pages 272–275, 2014.
7. A. Niell, R. Cappallo, B. Corey, C. Eckert, P. Elosegui, R. McWhirter, G. Rajagopalan, C. Ruszczyk, M. Titus. *VGOS Observations with Westford, GGAO, and the new station at Kokee, Hawaii*. This volume.