Implementation of the vgosDb format

S. Bolotin, K. Baver, J. Gipson, D. Gordon and D. MacMillan

NVI, Inc., Greenbelt MD, 20770

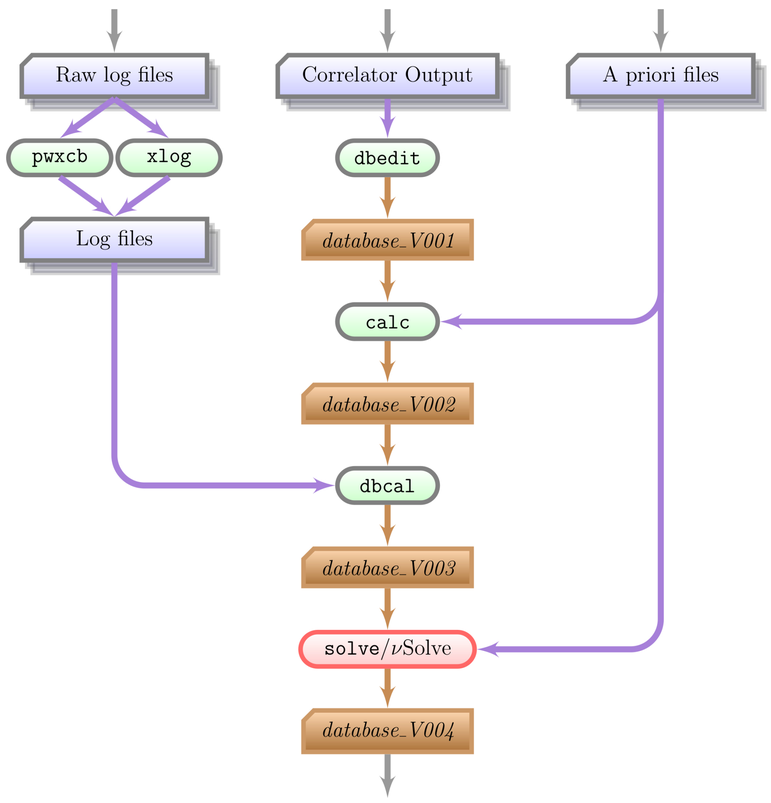
Introduction

The IVS Working Group 4 developed a new format to store and exchange data obtained from geodetic VLBI observations [1]. The new data format, vgosDb, will replace existing Mk4 databases this year.

At GSFC we are working on software that will implement the vgosDb format and will be used routinely to convert correlator output to the new data storage format.

Legacy geodetic VLBI data flow

Data produced at a correlator are subject to various changes before it becomes available to an end user. Historically, results of correlation of a VLBI session are stored in a special self-descriptive file called a *database*. Each modification or introduction of new information leads to a new version of the database. The figure bellow shows the traditional data flow of the geodetic VLBI observations.



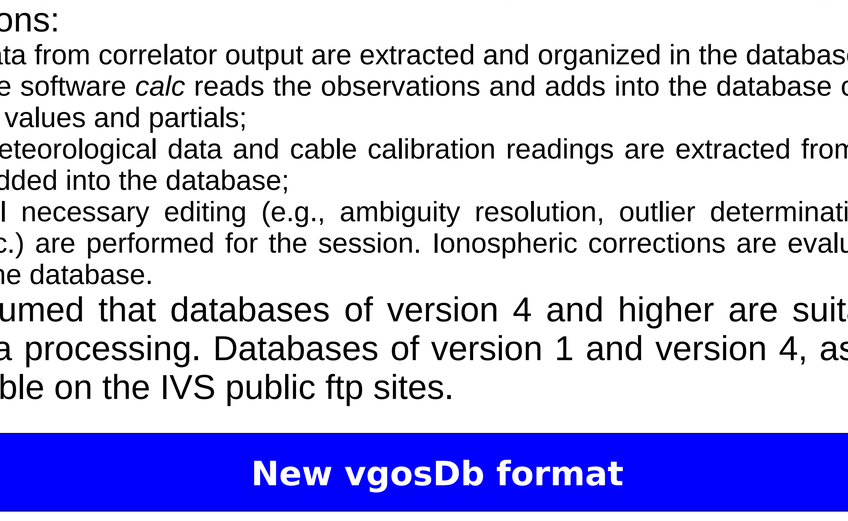
Traditionally, the version numbers correspond to the following Mk3 DBH modifications:

- *V001*: data from correlator output are extracted and organized in the database format;
- *V002*: the software *calc* reads the observations and adds into the database calculated theoretical values and partials;
- *V003*: meteorological data and cable calibration readings are extracted from station log files and added into the database;
- *V004*: all necessary editing (e.g., ambiguity resolution, outlier determination, clock breaks, etc.) are performed for the session. Ionospheric corrections are evaluated and stored in the database.

It is assumed that databases of version 4 and higher are suitable for batch data processing. Databases of version 1 and version 4, as a rule, are available on the IVS public ftp sites.

New vgosDb format

The new VLBI data format removes unnecessary redundancy. It keeps data in netCDF binary files (which allows flexible, platform independent access) and is well documented. The interaction between various utilities and VLBI data is shown on the figure.



According to the vgosDb format, the VLBI data are stored in various files in a form *key –> value*. Each file represents an atomic piece of data, e.g., observables with their standard deviations, station coordinates and so on. Also, it is acceptable to have several representations of the same information. This makes it possible to keep alternative models or approaches to editing of observations in the same data tree. The set of data files a user should use is specified in a special file called a "*wrapper file*". It is possible to have more than one wrapper file for one VLBI session.

Such an approach to organization of the VLBI data departs from the traditional meaning of "version".

The work on implementation of the vgosDb format by the GSFC VLBI group started in mid 2013. Now, the νSolve software as well as *global solve* are ready to use VLBI data in vgosDb format. The traditional utilities *dbedit* and *dbcal* will be replaced by new software. The software *calc* will be extended to use vgosDb format.

Design of the new software

The new utilities, *vgosDbMake* and *vgosDbCal*, are parts of the new VLBI data analysis software developed at NASA GSFC [2] and currently distributed under the name "nuSolve". The utilities have the same design as νSolve software. The same software development environment is also used.

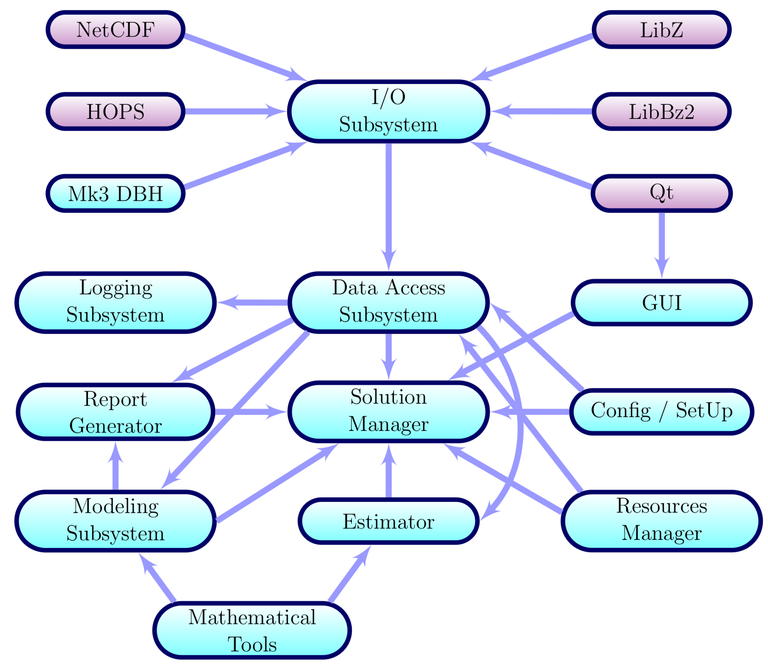
The utilities are designed to operate on any POSIX compatible OS. We use C++ as the programming language due to its power, flexibility and portability. *GNU Build System* is used to make the software distribution portable.

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 *vgosDbCal,* 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 reuse the source code in other applications.

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. The figure displays the general modular structure of the whole software package.



Several modules in the figure (e.g., HOPS, NetCDF, Qt) represent external libraries. The sizes of the modules could vary.

Obviously, not all the modules will be used by *vgosDbMake* and *vgosDbCal*. On the other hand, the design of the software makes it easy to add the functionality of the utilities to the interactive VLBI data editor, νSolve.

Modification of the software *calc* will be done in a different way. A library that mimics the *Mk3 DBH* programming interface will be created. The library will replace database functions with vgosDb I/O operations. In this case we do not need to modify *calc* source code at all, but just need to link the software with the new library.

Tentative timetable

Our group will switch to the new VLBI data format at the beginning of 2016. The work on *vgosDbMake* is at the final stage. In June we expect to finish tests of the utility. The utility *vgosDbCal* will be available in early Fall. Also, we expect to prepare *vgosDbCalc* in mid Fall. During the winter months, we will perform extensive tests of the whole VLBI data flow using the new vgosDb format and then start to release VLBI sessions in the new format routinely.

We strongly encourage all VLBI data analysis centers to switch to the new VLBI data format.

References

1. J. M. Gipson. **The Report of IVS-WG4**. *In IVS 2012 General Meeting Proc., NASA CP 2012-217504, pp. 212–221. 2012.*
2. S. Bolotin, J.M. Gipson and D. MacMillan. **Development of a New VLBI Data Analysis Software**. *In IVS 2010 General Meeting Proc., NASA CP 2010-215864, pp. 197–201. 2010.*

E-mail: sergei.bolotin@nasa.gov