The NASA Goddard Group's Source Monitoring Database and Program

John Gipson[1], Karine Le Bail[1], Chopo Ma[2]

Abstract Beginning in 2003, the Goddard VLBI group developed a program to purposefully monitor when sources were observed and to increase the observations of "under-observed" sources. The heart of the program consists of a MySQL database that keeps track of, on a session-by-session basis: the number of observations that are scheduled for a source, the number of observations that are successfully correlated, and the number of observations that are used in a session. In addition, there is a table that contains the target number of successful sessions over the last twelve months. Initially this table just contained two categories. Sources in the geodetic catalog had a target of 12 sessions/year; the remaining ICRF-1 defining sources had a target of two sessions/year. All other sources did not have a specific target. As the program evolved, different kinds of sources with different observing targets were added. During the scheduling process, the scheduler has the option of automatically selecting N sources which have not met their target. We discuss the history and present some results of this successful program.

Keywords Scheduling, sources

1 Problem: Under-observed Sources

In June 2003 we noticed that some sources were being observed a lot, while other sources, including many of the ICRF1 defining sources, were observed scarcely at all. Because there are literally thousands of sources that have been observed in VLBI, we decided to restrict our attention to two subsets. 'Geodetic' sources are in the geodetic catalog and are used for most IVS geodetic schedules. An ideal geodetic source is strong and compact. The other category, which we called 'Astrometric', consisted of the 212 defining sources of ICRF1 plus an additional 90 sources suggested by Martine Feissel. About a third of these sources were also Geodetic. After internal discussions, we decided that a reasonable observing target was that all 'Geodetic' sources should be scheduled and succesfully observed in 12 sessions/year, and the 'Astrometric' sources in two sessions/year. The remaining sources have no set observing target. Figure 1 plots the observing frequency of the Astrometric sources beginning in 2003 and continuing through 2004. In the

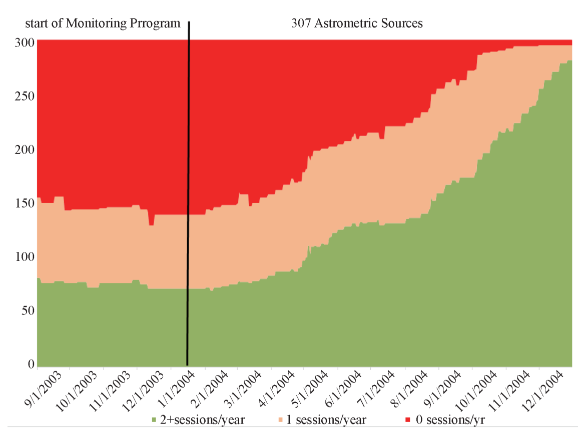

Fig. 1 Prior to the start of the Source Monitoring Program almost half of the Astrometric sources had not been observed during the prior year. Once the program starts the number of sources that do not meet their target [in red and orange (the top two layers)] decreases sharply.

1. NVI, Inc.
2. NASA Goddard Space Flight Center

remainder of this paper, we describe our approach to meeting these observing goals.

2 Source Monitoring Program

In principle, scheduling under-observed sources could be done by hand by the scheduler. The scheduler would need to do the following: 1.) find the under-observed sources; 2.) find out which ones are visible with the current network, and 3.) choose a set of these sources to observe. However, we wanted a solution that was more or less automatic and that required minimal human intervention. All of the above steps would still have to be done–but they would be done by the computer.

2.1 Overview of Database

Our approach to this was to design a database that kept track of when sources were observed. This database, called VDB, could be queried by humans or programs to extract the relevant information. We also decided to include additional information that would be useful outside of the needs for source monitoring. Table 1 summarizes the important tables in the database.

Table 1 Tables in the source monitoring database.

Name	Description
Master	Summary Information
SrcExp	Session specific source information
StatExp	Session specific station information
SrcObsDens	Source observing target
History	Database history

Master contains summary information about each VLBI session. This includes all of the information in the IVS Master files as well as additional information, for example, when the session is scheduled and when it is analyzed, and also individual session fits. This table is updated 1.) whenever the IVS Master files are updated; 2.) when the session is posted to CDDIS, and 3.) when the session is analyzed at GSFC to produce a Version 4 database.

SrcExp contains session-by-session information about each source. For each session, this table includes: the number of scheduled observations per source; the number of correlated observations per source, and the number of observations per source used in the solution. This table is updated: 1.) when a session is posted to CDDIS and 2.) when a session is analyzed at GSFC to produce a Version 4 database.

StatExp contains information about the stations. Its format and use is entirely analogous to SrcExp. This table is updated 1.) when a session is posted to CDDIS and 2.) when a session is analyzed at GSFC to produce a Version 4 database.

SrcObsDens contains source-specific information about the sources. This includes what kind of source it is (ICRF2, Defining, Special Handling, etc.) and the observing targets for each source. The only sources that are required to be in this table are sources that have some specific observing target. This table is only updated when we add new sources to the monitoring program or we change the observing targets of the source.

History contains a record of when the database is updated and is used primarily for debugging purposes.

2.2 Source Monitoring Dataflow

The Source Monitoring database is updated whenever 1.) the Master File is updated; 2.) a schedule is put on CDDIS, and 3.) a database is analyzed at GSFC. A schematic diagram of this is presented in Figure 2. The stages in this are described below.

Update_vdb_master keeps the Master Table current. This is run to create new entries in the Master Table when a new Master Schedule is created, for example, prior to the start of a new observing year. The program is also run whenever the Master Schedule is changed, which happens many times during the course of a year. When this is done, existing entries are updated to reflect the changes. This process happens automatically as part of posting the Master Schedule to the IVS Web pages.

Update_vdb_out is run whenever a new VLBI schedule appears on CDDIS. This program runs *sked* to extract information about the schedule. It checks to make sure that the schedule is in the Master Table, and if so, updates the Master Table indicating when the session was scheduled. In addition, it also updates SrcExp and StatExp indicating how many times each Source and Station in the schedule were scheduled.

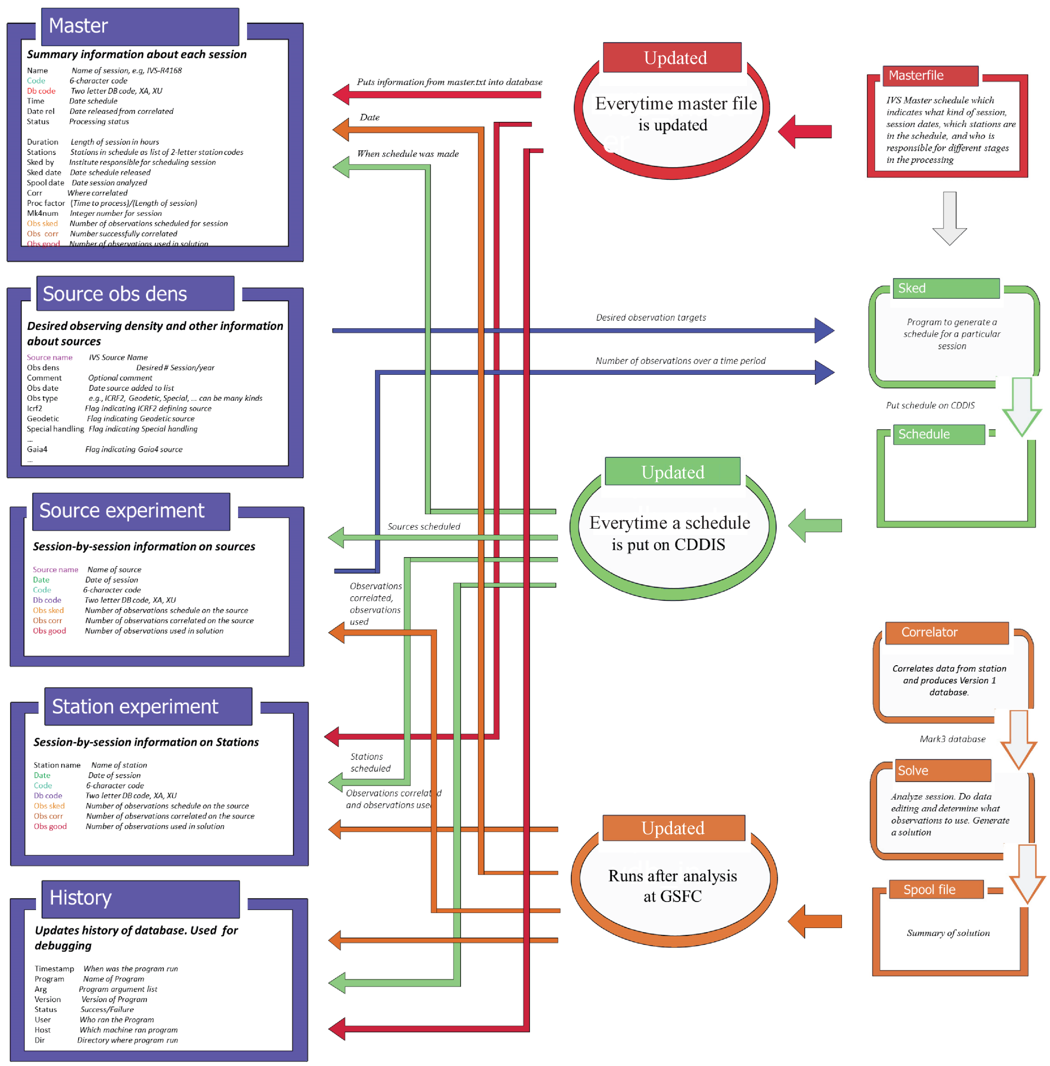

Fig. 2 Events that update the Source Monitoring Database.

This program is run automatically as part of posting the schedule the IVS data centers.

Update_vdb_in is run whenever GSFC analyzes an IVS session as part of its regular data analysis. The program reads the session spool-file (which contains the results of the analysis). It then updates the Master Table indicating when the analysis was done, as well as the session fit of the solution. It then updates SrcExp and StatExp with the number of observations that were correlated and used in the solution for each source and station. This program is run automatically as part of posting the Goddard solution to the IVS Web site.

The above describes how we keep the database updated to reflect the state of processing and the analysis of VLBI data. An equally important part of the source monitoring program is the ability to query the database. We built an interface between *sked* and the source monitoring database. If the user types “monitor N”, *sked* will attempt to select N under-observed sources to include in the current schedule. These schedules are put in the $ASTROMETIC section with observing targets of 1% to 1.5%, and *sked* will devote up to $1.5 \times N\%$ of the observations in the schedule to these sources. We do not have space to completely describe the process, but here is a summary. Using SQL commands, *sked* queries the database to find out how many successful sessions the sources in the monitoring program have been observed during the last year. Here $successful$

means that they were scheduled and at least three observations were used in the solution. It then compares this against the target for the sources and gives each score a rank by how much it has been under-observed. *Sked* will then select up to *N* sources from this list (starting at the top) that are observable with the current network. For a more complete description of *sked* and its commands, see [1].

3 History and Applications

The monitoring program started on February 1, 2004 and is now entering its tenth year. Initially only NASA was involved in this effort, and only R1s and RDV sessions participated in the program. In each of these sessions we scheduled up to ten monitoring sources and used up to 15% (although the actual number is typically less). We chose this number because we felt that this would not significantly degrade the measurement of EOP, which was the primary purpose of the R1s.

3.1 Starting the Monitoring Program

Figure 1 shows the observing history of the sources initially in the Astrometric source list before and after the Source Monitoring program. Initially only 25% of the sources in the list met their observing target of two sessions/year. We began to see results almost immediately, and by the end of the first year, 95% of the Astrometric sources met their target. The sources that did not meet their targets tend to be southern sources that are not visible by most stations in the IVS.

3.2 Enlarging the Geodetic Catalog

In 2008, GSFC decided to significantly enlarge the geodetic catalog. A series of four R&D experiments scheduled in April (RD0803), May (RD0804), September (RD0807), and October (RD0808) was devoted to observing these sources and establishing a baseline flux model. Once the sources were successfully observed, they were added to the catalog. Figure 3 shows the observing history of the new enlarged geodetic catalog.

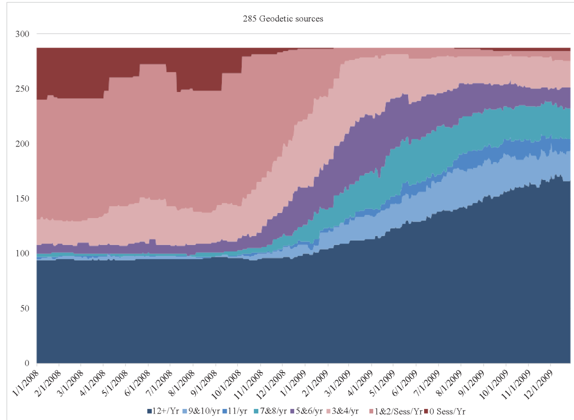

Fig. 3 Updating the geodetic catalog.

3.3 USNO Joins the Effort

In August 2010, USNO joined the Source Monitoring program, scheduling up to ten sources in each of the R4s. Because of security reasons, they could not directly access the database over the Internet. Instead we used the following work-around. Every evening the mysql database is tarred and gzipped and made available via anonymous ftp. USNO picks up the database the next morning and makes a local copy. This local copy is queried by sked when USNO generates schedules.

3.4 Gaia Transfer Sources

In mid-2013, we began introducing the Gaia transfer sources to the source monitoring program. This is a set of 195 sources submitted by the Laboratoire d'Astrophysique de Bordeaux (LAB) that will be used to tie the ICRF and the Gaia Reference Frame. These 195 sources are divided into four categories. The first two contain sources already in the IVS monitoring program, with a yearly observation target of 12. Categories 3 (16 sources) and 4 (24 sources) contain sources that are not in the monitoring program, and, for category 4, have a poor position accuracy. These sources are now scheduled in R&D and RDV sessions. We can see preliminary results in Figures 4 and 5.

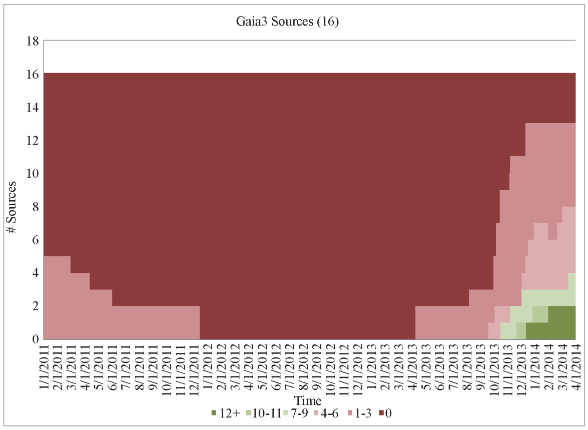

Fig. 4 Category 3 of the Gaia transfer sources.

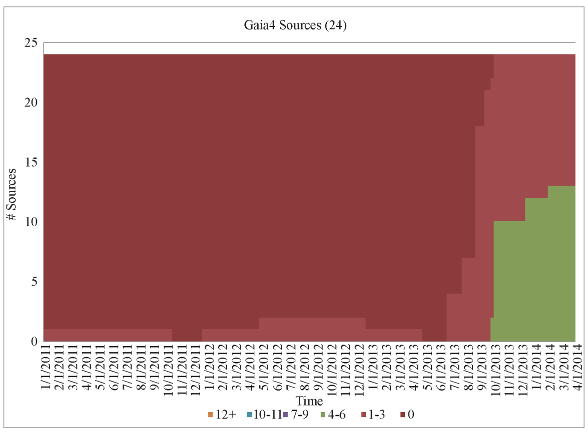

Fig. 5 Category 4 of the Gaia transfer sources.

4 Using SQL to Query the Database

The VDB database can be used for purposes other than to automate the observing of sources. The figures displayed previously were the results of querying the database about the frequency of observation of various source sets. Other queries are also possible. For example, in the query below (Figure 6) we inquire about all sources in the Intensives over a one-year period. More complex queries are possible, such as asking about: 1.) all sessions involving Kokee; 2.) all sessions involving Kokee and Ny-Ålesund which scheduled 3C84; 3.) observations of sources in different declination bands; and 4.) all observations of a source in the RDVs. Because of space limitations, we forgo further examples.

```
mysql> select src_name, count(*)
from srcexp where date
> "2007-8-31" and date <
"2008-9-1" and code like "i0%"
group by src_name;

+----------+----------+
| src_name | count(*) |
+----------+----------+
| 0014+813 |        2 |
| 0059+581 |      156 |
| 0119+115 |        9 |
| 0201+113 |        4 |
. . .
| 3C371    |      178 |
| 3C418    |       98 |
| 4C39.25  |       78 |
| DA426    |        2 |
| OJ287    |       24 |
+----------+----------+
31 rows in set  (0.07 sec)
```

Fig. 6 Sample SQL query and results.

5 Conclusion

The source monotiring mySQL database was originally designed as a tool to facilitate the automatic monitoring of sources. It has been very successful in this capacity. However, because it is a full relational database, you can use it to examine the data in different ways. In this note, we have presented some of the ways that we use database. A copy of the database is tarred and gzipped on a daily basis and made available via the Web at ftp://gemini/pub/vdb/gsfc_mysql_db.out.gz

References

1. J. Gipson, Sked User Manual, http://lupus.gsfc.nasa.gov/files_user_manuals/sked/sked.pdf