

ISDC's DAL and DAL3 libraries

B. O'Neel¹, M. Beck, L. Lerusse², R. Walter

*Integral Science Data Centre, Chemin d'Ecogia 16, 1290 Versoix
Switzerland [bruce.oneel,mathias.beck,roland.walter]@obs.unige.ch*

J. Peachey

*NASA/GSFC, Laboratory for High Energy Astrophysics, Code 660,
Greenbelt, MD 20771, Email: peachey@lheamail.gsfc.nasa.gov*

D. Jennings

Celera Inc, Rockville MD, Email: Donald.Jennings@celera.com

Stéphane Paltani

*Observatoire Astronomique de Marseille-Provence, Marseille France,
Email: stephane.paltani@oamp.fr*

Abstract. This paper discuss the ISDC's experiences with our Data Access Layers (DAL). Now that Integral has launched and we have a year of operational experience we can more fully evaluate DAL's benefits for Integral as well as potential benefits to other missions. ISDC's DAL was designed to solve problems anticipated by the combination of Integral's 4 instruments, wide fields of view, many instrument modes, and new pointings approximately every half hour. This would mean that any sensible scientific data analysis would process hundreds of individual files. Selecting and managing these files was seen as a large and unnecessary overhead to push off to the scientist and DAL and DAL3 have generally allowed scientists to work while remembering only one file name.

1. Introduction

Integral Science Data Centre (ISDC) has written 3 layers of libraries on top of CFITSIO (Pence 1999) for data access. DAL (O'Neel, Jennings, Rohlf, & Paltani 2000) is the data access layer designed to encapsulate CFITSIO and is not ISDC specific. The top layers are called DAL3 and they consist of ISDC specific functions which are both instrument independent and instrument dependent.

¹NASA/GSFC, Laboratory for High Energy Astrophysics and SP Systems, Seabrook MD

²Now at Max-Planck-Institut für extraterrestrische Physik, Garching, Germany, Email: Laurent.lerusse@mpe.mpg.de

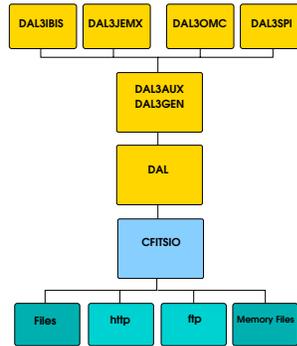


Figure 1. The Hierarchy of DAL libraries.

All of the software described in this paper can be downloaded from the ISDC Download Page¹ and the structure can be seen in Figure 1.

2. DAL Libraries

2.1. CFITSIO

The lowest layer in our system is CFITSIO. CFITSIO is a fast efficient, easy to use library for reading and writing FITS files. CFITSIO has the ability create FITS files from ASCII templates and this feature is used heavily by ISDC in order to produce a consistent set of products. Programs only fill in existing keywords and columns in precreated files.

2.2. DAL

Built on top of CFITSIO is ISDC's Data Access Layer or DAL. DAL is a mission independent library written in ANSI standard C that can be called from C, C++ and Fortran. It allows one to build up structures of FITS extensions in a way which is independent of the disk layout of your data but which matches the scientific processing needs for that data. Once this is done you can easily find and select different data structures by name and not just by position, independent of the file location on disk.

In order to build these structures DAL uses FITS binary tables called DAL Groups (Jennings, D., Pence, W., Folk, M., & Schlesinger, B. 1997). These groups, which are analogous to disk file directories, consist of FITS binary tables with 6 predefined columns. Entries in these 6 columns point to other FITS extensions either in the same file, in other files, or to remote systems using FTP or HTTP. Since these groups are nothing more than FITS extensions it is possible for groups to point to other groups allowing one to build hierarchies which match the science needs of the mission. An example showing DAL grouping is shown in Figure 2.

¹<http://isdc.unige.ch/index.cgi?Software+downloads>.

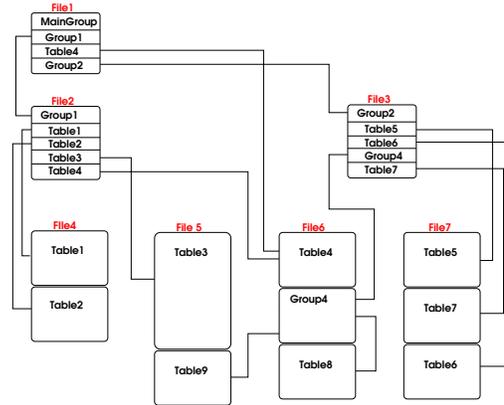


Figure 2. DAL Grouping Structures.

2.3. Instrument Independent DAL3 Libraries

Integral, like many other multi-instrument missions, has mission specific but instrument independent routines. These are stored in our two libraries DAL3AUX and DAL3GEN. DAL3AUX is the library which gives programs easy access to attitude, orbit, and time correlation data. DAL3GEN contains the indexing functions described later as well as support functions for Integral software.

2.4. Instrument Dependent DAL3 Libraries

In order to allow programmers to write more consistent programs and standardize certain operations which might be difficult, ISDC wrote an instrument specific library for each instrument. These libraries call routines in DAL3GEN, DAL3AUX, and DAL but give programmers an easier interface to the instrument data. Since Integral event data is stored in multiple tables, one of the primary purposes of these libraries is to make data selection with cfitsio selection strings easier across multiple tables for different instrument modes.

3. Indexing functions

Index groups expand upon the DAL groups concept above to allow for easy selection of scientifically interesting data. This is done by adding columns to the DAL group which contain the value of interesting keywords in the FITS tables pointed to by the DAL group. For example, if each of your files has a TSTART and TSTOP keyword listing the start and stop time, one would add columns named TSTART and TSTOP to the DAL index group. Then when you added this file to the index the TSTART and TSTOP columns would be filled with the values of the TSTART and TSTOP keywords from the header of your data extension. An example index table is shown in Figure 3.

Once the index is built one can easily select, using CFITSIO selection strings, data which matches your criteria for subsequent processing. In this way you don't have a large mass of data you have a simple table which allows queries for selection.

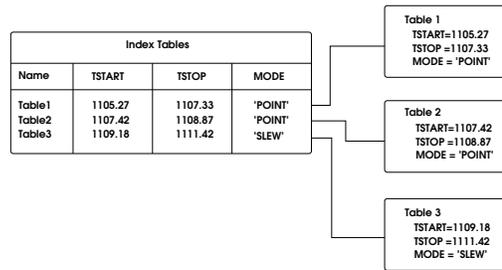


Figure 3. DAL Indexing.

4. Lessons Learned

4.1. Successes

- It is likely that without DAL and the DAL3 libraries the ISDC software would have been slower to write and less reliable while running. Most low level operations are quite standardized in both DAL and DAL3 and these routines have been heavily tested in daily operations.
- DAL has allowed us to make all of the tools very flexible from the run time perspective. Most of the time the scientist only has to specify one file, a group, and the rest of the data is found by searching that group or its subgroups.
- We have been pleased with the overall performance of DAL and CFITSIO even on our slow 300MHz operational processing systems.
- It is very useful to have the whole processing chain in FITS. Many tools such as ftools can easily examine any FITS file in our processing system allowing easier debugging and analysis.

4.2. Difficulties

- Scientists who wrote much of ISDC's software had some difficulties with the abstractions presented by DAL. This was solved by user education similar to the education required when one moved from flat ASCII or binary files to FITS files.
- Since DAL decides when to open and/or close physical files it was not always obvious when an operation was efficient and when it was slow.

References

- Jennings, D., Pence, W., Folk, M., & Schlesinger, B, A Hierarchical Grouping Convention for FITS, <http://fits.gsfc.nasa.gov/group.html>.
- O'Neel, B., Jennings, D., Rohlfs, R., & Paltani, S., 2000, in ASP Conf. Ser., Vol. 216, ADASS IX, ed. N. Manset, C. Veillet, & D. Crabtree (San Francisco: ASP), 187
- Pence, W. D., 1999, in ASP Conf. Ser., Vol. 172, ADASS VIII, ed. D. M. Mehringer, R. L. Plante, & D. A. Roberts (San Francisco: ASP), 487