Astronomical Data Analysis Software and Systems VI
ASP Conference Series, Vol. 125, 1997
Gareth Hunt and H. E. Payne, eds.

Case Study of RDBMS Use on The EUVE Mission

E. C. Olson

Center for EUV Astrophysics, 2150 Kittredge Street, University of California, Berkeley, CA 94720-5030

Abstract. The *Extreme Ultraviolet Explorer* satellite (*EUVE*) mission has employed a Relational Database Management System (RDBMS) in several ways. Shortly after the launch of *EUVE* in 1992, an indexing scheme for the *EUVE* telemetry was implemented to facilitate access to the calibration, in-flight, and supplemental telemetry data. The RDBMS has been used for tracking *EUVE* observations and other administrative information. Over time, it has evolved into the central data repository of the distributed *EUVE* data reduction system. A number of experimental systems also used the RDBMS. In 1992 *EUVE* began participating in the Astrophysics Data System, using the RDBMS as the repository for the *EUVE* bright source list and additional, ancillary information. With the advent of the World Wide Web, *EUVE* used its RDBMS to publish dynamic information for use within the project and across the Internet. More recently, *EUVE* has created prototype client programs to facilitate use of relational databases for scientific applications.

1. Introduction

The *Extreme Ultraviolet Explorer* (*EUVE*) mission (Bowyer et al. 1988; Haisch et al. 1993) employs relational database management systems for several mission-critical activities. Additionally, *EUVE* uses databases for a number of non-scientific, administrative, and other non-mission-critical areas. This paper will focus on the use of database technology for the mission critical aspects.

The Center for EUV Astrophysics (CEA), Space Sciences Laboratory, at the University of California, Berkeley, operates the scientific telescopes and instrumentation on board the *EUVE* satellite. CEA developed ground systems for acquiring, monitoring, archiving, and distributing the *EUVE* science data (Olson & Christian 1992; Christian & Olson 1993; Olson 1995; Olson et al. 1996). The primary mission-critical use of database systems has been in the archiving and distributing of *EUVE* telemetry.

2. Telemetry Archive

A terabyte optical disk jukebox serves as the primary storage facility for the Telemetry Archive (Girouard & Hopkins 1993). The jukebox also supports other hardware devices for various operational reasons beyond the scope of this paper. The jukebox supports the Network File System (NFS) protocol, which greatly simplifies access and manipulation of the archive. The Telemetry

Archive database itself is maintained in the Sybase RDBMS. The Telemetry Archive comprises two components: a general purpose archive database and a telemetry-specific archive database.

The general purpose archive is used for a wide range of purposes. Its primary purpose is to provide a mechanism for the archiving of telemetry. However, a number of supplementary files related to the telemetry and to the general operations of *EUVE* are also archived. This archive behaves like a simple file system that supports several different media. It maintains indexes, based on file names, tracking the location, size, and type of the file, including its media and media-specific information (such as file position and ID on the tape carousel device). This database resolves queries based on a unique ID, returning a specific file. The general-purpose archive also synchronizes processing. Handling telemetry is essentially an event-oriented processing problem. At any time the system must be prepared to receive telemetry. After the initial reception, the telemetry is processed in a variety of ways depending on its type. The features of RDBMS used are transaction locking, consistency checks, reliability, and robustness (particularly in light of machine crashes). Interestingly, the archive does not use the RDBMS relational features, and the system logic is coded in the application software. Essentially, the RDBMS is used as a reliable data store. Routinely, autonomous application software interacts with this database with little direct user interaction. The application software anticipates most routine exceptions and processes them accordingly. Highly unusual exceptions require the manual intervention of the system administrator. Since the system uses unique hardware, it is subject to single point failures. Additionally, the unique hardware and the 24 hour per day operation make testing of the database system difficult.

The telemetry-specific archive is used for maintaining information specifically required for accessing and manipulating the *EUVE* telemetry. The specific information includes time-stamps, exposure durations, and data quality information. This database is fairly straightforward since its organizing feature is time. Typical queries that the telemetry database can answer are related to time-oriented information. For example, the database maintains information about telemetry intervals and gaps in the telemetry stream. As with the archive database, the telemetry database synchronizes asynchronous processes where transaction-locking and reliability are key features. Again, the application software handles the logic, and the database is used as a reliable data store. More significant and less predictable errors can occur with this database. Although routine processing is handled autonomously, deleting badly formed telemetry from the index is problematic.

3. Proposal Database

The Proposal Database of *EUVE* pointed observations is designed to provide not only information about the target, the observation, and the principal investigator for the data, but also information about the observation schedule, the location of the acquired data, and various types of historical data such as the software version used for processing. Over time, the design expanded to include the state of the data reduction processing. In general, information in the Proposal Database handles higher level information than the information found in the Archive Database. The Proposal database has two components: an observation database and a proposal database.

The observation database, closely related to the Archive Database, obtains input data from the telemetry database and from supplementary information about the observations. Using the telemetry interval information in the telemetry database, the observation database constructs higher level data abstractions about observation intervals. For example, these observation intervals tolerate data gaps and track contiguous exposure time intervals on a single target. Additionally, the observation database handles cases in which an observation may be interrupted for, say, calibration pointings. In this case, the database, not the application software, contains most of the logic. The observation database makes significant use of the database relational features as well as triggers for maintaining the integrity of this database. However, the more complex logic required is located in the application software as it was deemed too complex to maintain in SQL. The transaction and synchronization features of the database are used for coordinating access by asynchronous distributed processes. Much of the input for this database comes from operators, and is subject to the usual errors that occur from manual data entry. Therefore, this database also supports a WWW interface for correcting data, and includes a reporting mechanism that facilitates the detection of errors. In addition, the WWW interface allows easy handling of runtime exceptions.

The proposal component of the Proposal Database tracks high-level information about investigators, their proposals, the related observation requests, and completed observations. This database has the greatest amount of operator data entry. Several mechanisms have been used to check the consistency of the data in this database, since errors can directly effect the productivity of the *EUVE* mission. Moreover, the information changes frequently as adjustments are made to observing proposals. Runtime exceptions occur frequently and are handled by operators using a WWW interface. This interface is closely related to the schema structure of the database. Operators must be familiar with the schema structure in order to appropriately correct data in the database. However, this cost is small compared to the cost of identifying and anticipating the multitude of exception conditions that may occur in this database.

4. Discussion

The *EUVE* mission has used RDBMS technology in mission-critical processing. The features used are transactions and synchronization (particularly suited to distributed systems). In some cases the relational features of the RDBMS were also used. However, in many cases complex logical expressions were handled in the application software instead of the RDBMS, and the RDBMS was used as a reliable data store. Overall design, development, and testing of the databases was not difficult.

Providing operators with the capability to access and manipulate data has been an ongoing problem in working with databases on the *EUVE* mission. We experimented with several commercial products, as well as the Astronomical Data System, in an attempt to provide an operator front-end to the RDBMS. None of these systems provided an appropriate solution. Typically such products are designed to provide complete solutions that require extensive development of interface specifications intimately connected to the schema of the database. This type of extensive development was never justified, nor did it appear maintainable since the database schema have continuously evolved. The evolution of the database schema is a natural result of the evolution of the scientific goals of the mission. Therefore, instead of providing a complete solution only applicable at a given moment, the *EUVE* mission has concentrated on providing general-purpose, partial solutions. Specifically, the WWW interface used in the Proposal Database is derived from the database schema. This interface requires operators be trained in the structure of the database, but dramatically reduces development and maintenance costs. More recently, the *EUVE* mission has developed a WWW server prototype (named xdb) that provides structured access to databases based solely on their meta data information.

Acknowledgments. This work was supported by NASA contract NAS5-29298. The Center for EUV Astrophysics is a division of the Space Sciences Laboratory at UC Berkeley.

References

Bowyer, S., Malina, R. F., & Marshall, H. L. 1988, JBIS, 41, 357

Christian, C. A., & Olson, E. C. 1993, in ASP Conf. Ser., Vol. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (San Francisco: ASP), 56

Girouard, F. R., & Hopkins, A. 1993, JBIS, 46(9), 342

Haisch, B., Bowyer, S., & Malina, R. F. 1993, JBIS, 46(9), 330

Olson, E. C. 1995, in ASP Conf. Ser., Vol. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes (San Francisco: ASP), 425

Olson, E. C., & Christian, C. A. 1992, in ASP Conf. Ser., Vol. 25, Astronomical Data Analysis Software and Systems I, ed. D. M. Worrall, C. Biemesderfer, & J. Barnes (San Francisco: ASP), 110

Olson, E. C., Girouard, F., & Hopkins, A. 1996, in ASP Conf. Ser., Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes (San Francisco: ASP), 532