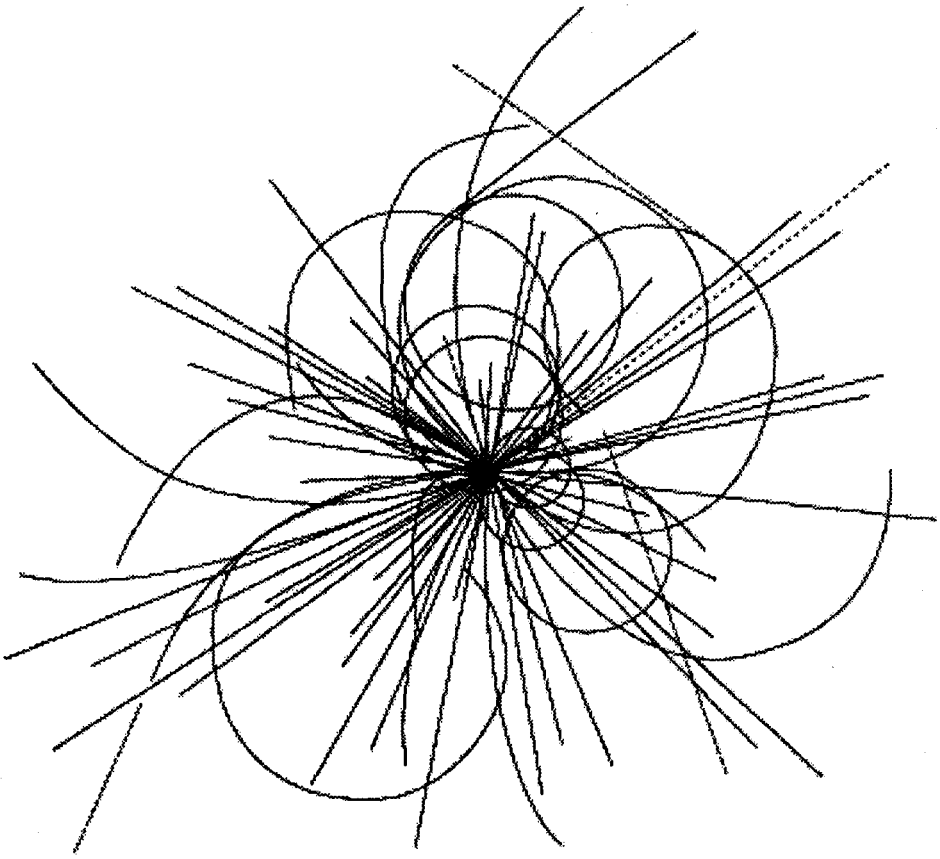
SSCL-Preprint-541

SSCL-Preprint-541
December 1993
Distribution Category: 414

U. Nixdorf

Experience with a Run File Archive Using Database Technology



Superconducting Super Collider Laboratory

SSCL-Preprint-541

Experience with a Run File Archive Using Database Technology*

U. Nixdorf

Superconducting Super Collider Laboratory†
2550 Beckleymeade Ave.
Dallas, TX 75237, USA

December 1993

---

*Presented at the Third International Workshop on Software Engineering Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics.

†Operated by the Universities Research Association, Inc., for the U.S. Department of Energy under Contract No. DE-AC35-89ER40486.

EXPERIENCE WITH A RUN FILE ARCHIVE USING DATABASE TECHNOLOGY

Ute Nixdorf
Superconducting Super Collider
Physics Research Division
Computing Department
2550 Beckleymeade Avenue
Dallas, Texas 75237-3997

ABSTRACT

High Energy Physics experiments are known for their production of large amounts of data. Even small projects may have to manage several Giga Byte of event information. One possible solution for the management of this data is to use today's technology to archive the raw data files in tertiary storage and build on-line catalogs which reference interesting data. This approach has been taken by the Gammas, Electrons and Muons (GEM) Collaboration for their evaluation of muon chamber technologies at the Superconducting Super Collider Laboratory (SSCL). Several technologies were installed and tested during a 6 month period. Events produced were first recorded in the UNIX filesystem of the data acquisition system and then migrated to the Physics Detector Simulation Facility (PDSF) for long term storage. The software system makes use of a commercial relational database management system (SYBASE) and the Data Management System (DMS), a tape archival system developed at the SSCL. The components are distributed among several machines inside and outside PDSF. A Motif-based graphical user interface (GUI) enables physicists to retrieve interesting runs from the archive using the on-line database catalog.

1. Introduction

The GEM Collaboration designed a test stand to evaluate various muon chamber technologies for their future detector. A test rig was built, the chambers installed and the electronic data acquisition system (DAQ) set up to register cosmic muons passing through the detector prototype. Data was initially registered during a period of 6 months. This produced several Giga Byte of event data, which was first recorded on the disk and tape of the DAQ system. The data was then migrated to PDSF for long term storage. The purpose of this setup was the comparison of the quality of different types of muon chambers. The amount of data anticipated drove the requirement to be able to select interesting runs without having to retrieve and scan through all stored run files. Retrieval of run files from tape and sequential analysis of all run files would consume a lot of resources. A list of criteria was defined by which the collaboration members are able to decide, up front, which files they need to access. The criteria catalog is implemented using the relational database management system (RDBMS) SYBASE. The data itself resides in its original form in the tape archival system, DMS, of the PDSF. The project used commercial and in-house products and integrated them into an archival system that provides a fast and efficient way to access the run file archive. In order to ensure continuous operations even for several consecutive days of data recording, the software is designed to need as little attention as possible and is able to restart itself, in most cases.

2. System Architecture

The Texas Test Rig (TTR) archival system manages the run file archive. It is implemented in a client server environment. The archival system reads and analyzes the run files from the disk of

the DAQ system to build the criteria catalog. The run files are then moved from the DAQ disk to the DMS of PDSF. Users select files, using the Motif-based graphical user interface to the database, and request associated runs to be restored from tape. Figure 1 gives an overview of the hardware and software components of the archival system. Catalog data is transferred, from the DAQ System to the DBMS, and retrieved by a user on the physics workstation. The files containing the run data are moved, from the DAQ system to DMS, and loaded to a physics workstation, on request.

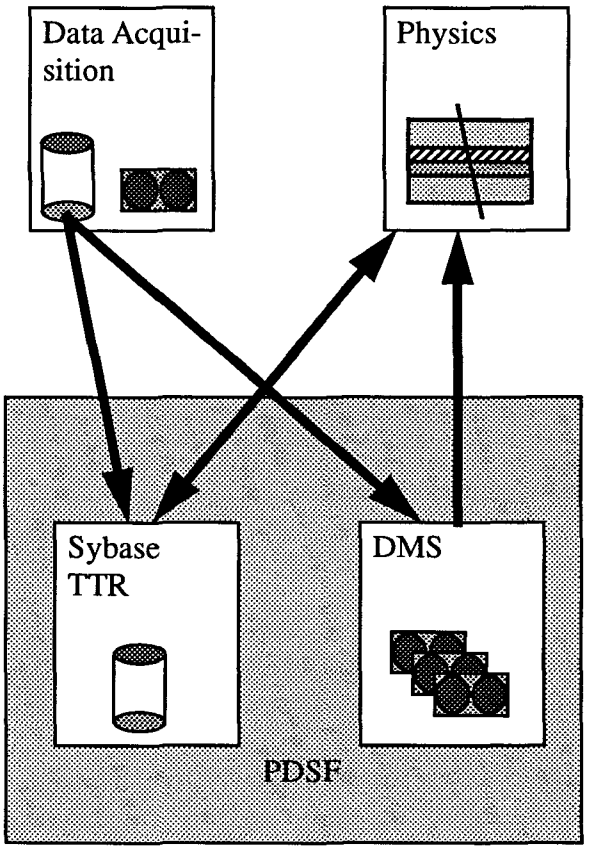


Figure 1: System Overview of the TTR Archival and Catalog System

The TTR software system is implemented in a client-server fashion. Different machines execute parts of the overall task. Authorized users can access the catalog and the DMS from any PDSF workstation. The database management system is installed on one of the workstations of the PDSF. DMS uses the batch ranches of PDSF and drives the two tape robots. Each robot contains a carousel which can hold up to 54 8mm ANSI-labelled tapes. Part of the DMS software is a SYBASE database application.

The TTR System is a complex network of machines and software which will usually fail if one of the components does not function properly. The SYBASE software for the TTR is installed on participating physics workstations. The database server, the database application programs and query tools are installed on the TTR database server machine, as well as on the physics workstations. It has to be ensured that the DAQ system is not affected by these off-line activities. It will continue to store recorded data as long as sufficient space is available. This guarantees that the events can be recorded, even if the long term archival is not available.

3. TTR Catalog Database

For the TTR, we used SYBASE a relational, client-server based database management system. The TTR database management system provides concurrent access to the stored data through a standard interface. The TTR software was implemented using the C Library interface of SYBASE. The DBMS allows storage of data in a central place, so applications can access it from any workstation that has a network connection to the database machine and the appropriate permissions. Consistency and safety of the stored data are basic features of database management systems. To be able to build an efficient relational database, certain rules have to be applied. These rules determine the way data is stored during the design of the database. The purpose of the system is to select run files based on the technology used during the run, environmental conditions, gas mixture used for the chambers and other technical information. The graphical database design tool ERdraw[1] and the associated schema translation tool SDT[2] were used for database design and implementation. ERdraw is based on an extended entity relationship model and generates SYBASE SQL code for database creation and consistency control. It proved to be a very useful tool for SYBASE database development. Figure 2 shows an overview of the TTR database design using ERdraw notation.

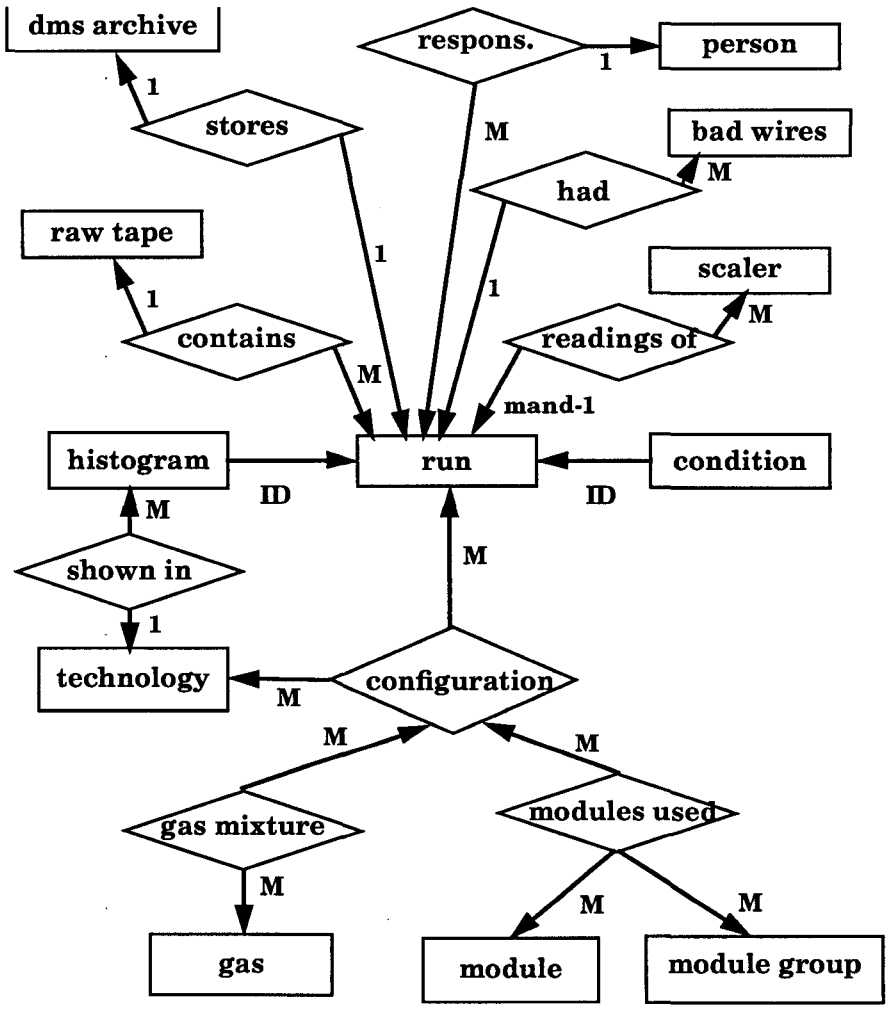


Figure 2: TTR Database Design (ERDraw notation) [1]

A graphical user interface was built for database access. It ties predefined database queries, programmed using SQL, to buttons and menus in the window. Physicists can build queries by the "point and click" method and review the results on-line. The query results can be saved in files for later reference. Additionally selected run file identifications are used directly to invoke other tools supported by the TTR window environment. Generated histograms are also managed and stored on PDSF disks by the TTR software. This allows physicists to view results immediately after selecting the interesting runs without having to restore any file from the archive.

4. Integration with Data Acquisition

The database applications are integrated with the DAQ only through the data itself. At the end of a run, the file containing the events is written to a specific directory on one of the disks of the DAQ system. The archival software periodically checks this directory, analyzes the files for catalog data and sends it to DMS for long term storage. The file is then removed from the DAQ disk. The run files contain "special events" and "special comments" which, together with other information, is used to build the database catalog. This causes the archival software to be sensitive to changes in the run file structure. The mechanism used for periodic checking is a UNIX "cron" table invoking a "fresh" copy of the program every time. This enables the software to recover automatically from most system problems.

5. DMS

DMS, the Data Management System of the PDSF[3,4,5], provides tools to support users in the off-line tape storage of files. DMS shields users from having to know anything about the particular storage medium. The storage is addressed like a standard UNIX file. The actual transfer of the data to tape is done transparently by internal mechanisms of DMS. The DMS software uses the four batch ranches, currently part of PDSF. DMS also provides a catalog about files and storage locations via a SYBASE application. The tapes are managed by two tape carousels with the least recently used tapes moved to a shelf. This can sometimes lead to long response times. The TTR applications follow a policy of restoring the second file requested while the first one is being processed. The transfer of a file to/from DMS currently involves an intermediate staging of the file on one of the PDSF batch ranches.

6. Graphical User Interface

The TTR catalog provides on-line information about runs and associated technologies. Users select runs using a MOTIF-based graphical interface to determine interesting run files before they request them from the DMS. The graphical user interface provides the user with a set of predefined buttons and menus to compose a query by the "point and click" method. It retrieves the data from the database using SQL, the standard query language for RDBMS, and displays the results on the screen. Users can then detail the query or use the resulting file name to retrieve the run from the tape storage. The run file can be sent to the event display or other physics analysis programs by clicking on the appropriate buttons of the TTR interface. Query results can also be saved in files for future reference. The graphical user interface also allows changes to the database, provided the user has the necessary permissions.

The TTR GUI currently supports a predefined set of queries. The system is designed to extend this set of queries as required. Any query to the database can also be generated using the SYBASE toolset.

Figure 3 shows a snapshot of a session using the TTR GUI.

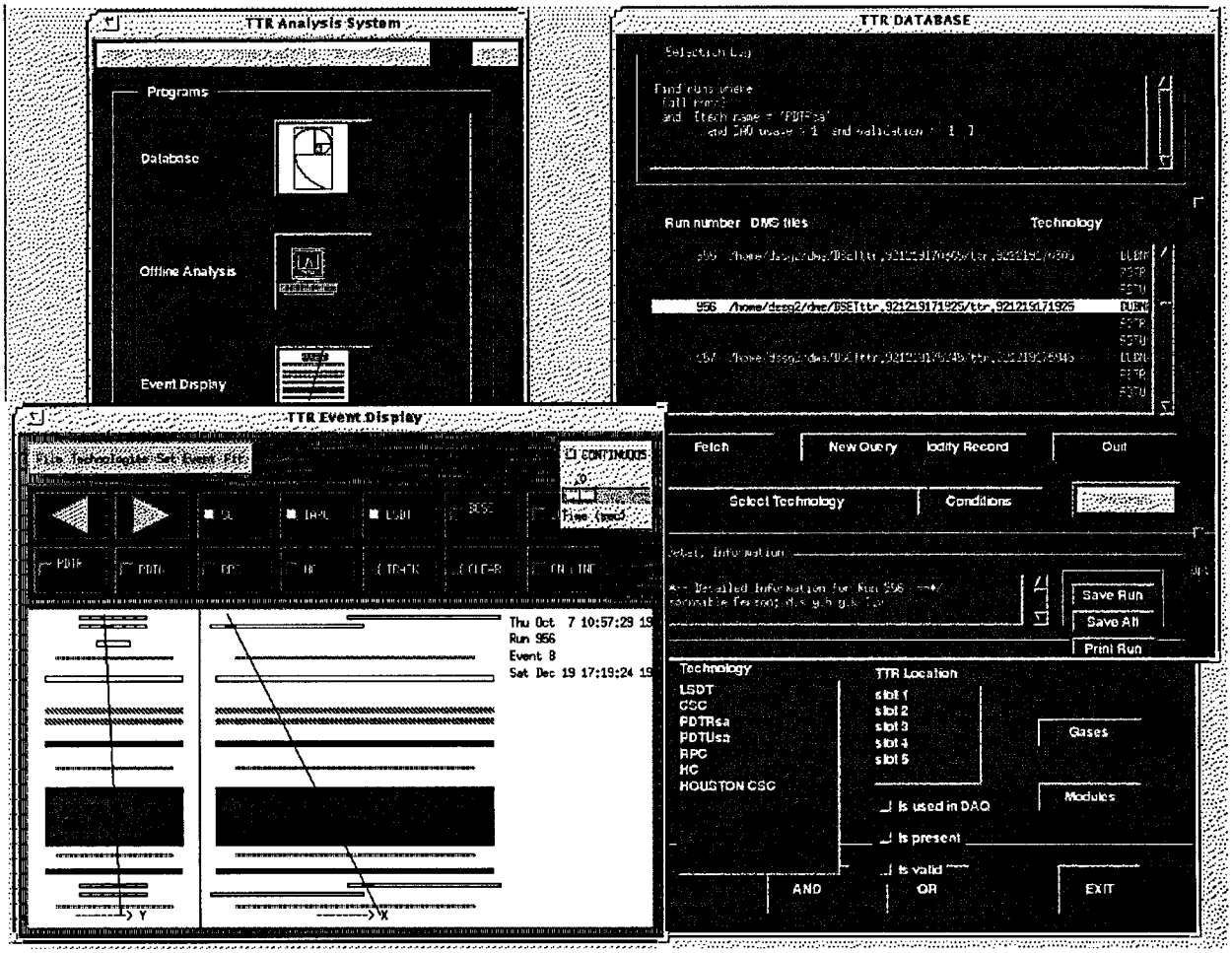


Figure 3: Graphical User Interface

7. Integration with Physics Analysis Programs

The graphical user interface integrates the database query application with physics analysis programs. It is possible to determine the name of a file, which contains interesting events, by applying the query mechanisms. The user can then immediately apply the result of the database query, - usually a run file name - to retrieve this file from DMS. It can then be input into other physics analysis programs or displayed directly with the on-line event display.

The results of run file analysis frequently are kept in histograms files, which are also archived in PDSF and entered into the TTR catalog for future reference. This allows users to reference the associated histograms immediately after selecting a run.

8. Status and Conclusions

Between October 1992 and March 1993 approximately 560 runs were recorded and stored in DMS using the TTR catalog software. DMS tapes currently store 7 Giga Byte of run files, more than 12 million events. Run file sizes vary between a few Kilo Byte and 50 MB. Sybase and DMS have proven, for the scope of this project, to be reliable and stable. The TTR software and database maintenance is fully automated and needs attention only occasionally. The TTR system, for long term storage and selective retrieval for run files provides tool to select interesting files without having to retrieve and analyze all of them.

The database catalog is generated automatically. This requires the presence of the special events and special comments in the run file. The Entry was encouraged but not enforced. This caused one third of run files not to have any catalog entry besides their name.The lesson learned here is that as much as possible has to be done automatically. If someone has to remember to add certain information to run files and to enter some keystrokes to start a catalog entry, at least part of the files will not use this feature.

The interface between the DAQ and the catalog application is the run file itself, which makes the catalog builder dependent on changes in the run file structures. These were frequent, especially in the beginning of the project. The dependencies between run file and catalog software are not obvious. The advantage of this approach is that the DAQ system can run completely independent from the database system. On the other hand, the integration of the DAQ software and the catalog builder would make things faster and less vulnerable to changes in the run file. Future projects may be able to solve this trade off.

9. Acknowledgments

The author would like to thank Ivan Chow for his development of the graphical user interface for the TTR software package, Luis Villasenor for the event display, Jeff Allen for his support as a DMS expert and George Yost for the many discussions and suggestions. She would also like to acknowledge the contributions from our students and the members of the GEM collaboration who participated in this project.

10. References

1. E. Szeto, V. Markowitz, *"ERDRAW A Graphical Schema Specification Tool"*, Lawrence Berkeley Laboratory 1990
2. V. Markowitz, Weiping Fang, *"SDT A Database Schema Design and Translation Tool"*, Lawrence Berkeley Laboratory 1990
3. Jeffrey Allen, *"SSCL-PDSF Data Management System"*, presented at the CHEP 92
4. Jeffrey Allen, *"SSCANSI, SSC Laboratory's Implementation of ANSI Labelled Tapes"*, 1992
5. Ch. Chang, Cathy Nelson, *"Computer System Operator's Manual for the Data Management System (DMS)"*, SSCL Documentation, September 1993