CONSTRUCTION AND APPLICATION OF DATABASE FOR CSNS

Q. Gan, C.H. Wang, IHEP, P.O. Box 918, 100049, Beijing, China
P. Chu, SLAC, Menlo Park, CA 94025, USA

Abstract

The database of the China Spallation Neutron Source (CSNS) Accelerator is designed to store machine parameters, magnet measurement data, survey and alignment data, control system configuration data, equipment history data, electronic logbook, and so on. It will also provide project management quality assurance, error impact analysis, and assembly assistance including sorting. This paper introduces the construction and application of the database for CSNS. Details such as nameing convention rules, database model and schema, interface of import and export data, and database maintenance will be presented.

INTRODUCTION

The accelerator database plays an important role in accelerator control system. Nowadays, the database is not only used for storing the equipments parameters and control information, but also used for staff management, resource allocation and a wide range of other functions. It is an integral part of day to day running of accelerator operation.

Generally, the database is created by database group or control group. But the database tables and relevant applications are created and maintained by other groups and divisions. So it's necessary to make uniform database rules and provide advice on database construction.

DATABASE OVERVIEW

Today, many larger scale laboratories have built the global relational database, and being committed to combining the database with accelerator software and hardware. The global database can keep data consistent.

At present, the large database management system (LDBMS) product includes Oracle, Sybase, Informix, DB2, MS-SQL server and so on. Among them, the Oracle database whose technology is the most mature, and applications are the most extensive, is preferred solution for enterprise-class applications. And most of the high-energy physics laboratories chose the Oracle database product. Therefore, CSNS also will also use it.

The accelerator data can be divided into two types: static data and dynamic data. The static data includes equipments parameters, installation and survey information, design values of the accelerator physics, staff information and so on. While the dynamic data includes real-time data of the control system, MPS (Machine Protection System) and alarm information, operating log and so on. Some laboratories divide the database into multiple parts, or use two or more database products for the different type data. However, in order to manage and the maintain data efficiently, some other laboratories put these two types of data together in one global database. The CSNS will prefer the latter approach.

In addition, many laboratories which use EPICS software packages are working on importing EPICS database into Oracle. For example, the KEKB facility developed the accelerator database and configured it especially for EPICS IOC database. The SNS laboratory developed a global database for both accelerator data and staff information. They also developed relevant applications (JERI) for configuring EPICS IOC database [1]. The DESY laboratory and some collaborators are developing a new software package called CSS (Control System Studio), which as a future version of EPICS can better combine with the Oracle database.

DATABASE DESIGN

The development of CSNS accelerator database has six stages.

- Requirements analysis

Accurately understanding and analysis of the accelerator requirements (including data and processing) is the basis of the whole design process, and the most difficult, the most time-consuming step.

- Conceptual design

According to the requirements, abstract the results of the first stage. And design the E-R (Entity-Relationship) diagrams for every domain of accelerator. Then, integrate the part E-R diagrams into a global E-R diagram. As shown in figure 1, each block stands for a domain of the accelerator, and the lines between them show the entity relationship.
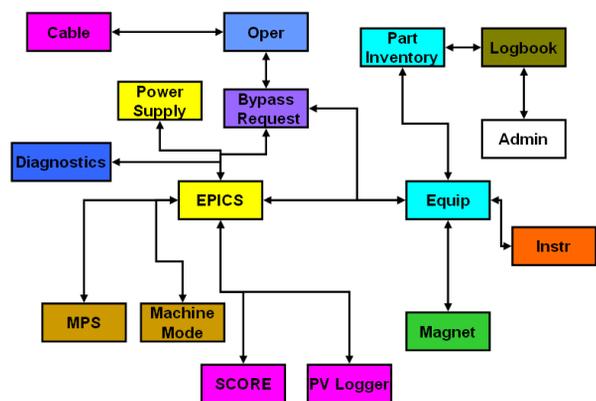

Figure 1. Layout of CSNS database model.

- Logic design

First, convert the E-R diagram to relation model. Then, adjust the model structure properly, and optimize the data model.

- Physical database design

Considering the time efficiency, space efficiency, maintenance costs and user requirements, design a suitable physical structure including storage structures and access methods.

- Database implementation

According to the relation model and physical structure, create the database by the Oracle language and relevant tools. Then, import the data and develop the database applications.

- Database operation and maintenance

After the above-mentioned work, the database will formally run. During the database system running, the database should be adjusted and improved continuously. The database availability should be 100% except for routine maintenance.

Database Model

First, according to CSNS accelerator requirements, lists out the accelerator data. Generally, this work requires all groups' cooperation. Then, divide these data into several domains, such as magnet, cable, survey, optics, and so on. Analyze the links between these domains, and draw the ER diagram. Here, a specialized database modelling tools (such as PowerDesigner) is recommended. It can not only greatly simplify the creation of the database, but also directly generate the SQL code for various database products.

Naming Convention

CSNS accelerator database use a unified naming rule. The entity name format is <entity type> _ <entity full name> _ <added notes (optional)>, and the field name format is <domain acronym> _ <field full name> – <added notes (optional)>.

The database entities include tables, query and report. In order to distinguish these different type entities, add the prefix for the entity name (such as tbl, qry and rpt). The prefix of the table field is the domain name initials (such as MAG stand for magnet domain).

The entity full name and the field full name are composed by 1-3 words. Each entity also has an acronym name which is composed by 1-4 capital letters. But the field name should not use acronym to avoid confusion field meaning.

The suffix is used as the supplementary note of entity and the field (such as tmp and bkp). Particularly, table which is used for application software should use "work" as prefix, and the application's name as suffix.

Table and Key

Database normalization can help to reduce data redundancy. The third normal form (3NF) is a good choice for the accelerator database. It is easy to insert of data, and give better chance to change the database design compared to NON 3NF design.

However, according to the experience of other laboratories, it's unnecessary to use normal form sometimes. For example, some magnet tables are designed for magnet measurement and survey. Although

these tables contain a lot of redundant information, they can be easily filled and checked by relevant groups.

The selection of table key should follow 4 rules: create for foreign key for the correlative fields; all keys must unique; avoid using composite key; the foreign key must reference a unique key. [2] When necessary, the auto-generate key should be used as the main keys. For frequently insert and delete data, some small tables can not set key.

In addition, each table should have three fields: ModifiedUser - The user who last changed the table record; ModifyDate - date and time when the table was last modified; RecordIndicator – a free form text field relating to the reasons about loss of data.

Data Interface

The construction of accelerator is a collaborative work by various departments. Each department may use different software to generate different data table with different format. It's a great trouble for the data entry. Therefore, a uniform data interface specification is needed.

The CSV (Comma Separated Values) format is chosen as the uniform data form. The CSV has a compact definition and is compatible with most software.

Several methods can be used to import CSV files to Oracle database such as SQL*Loader which is offered by Oracle. The control files of SQL*Loader are required to edit manually. The command line mode is used to import data of it. This method is used in BESSY and SNS. Besides, the ODBC (Open Database Connectivity) drive program can also be used to exchange data with Oracle.

Data Query

Many data selection and management tools are offered by Oracle product. There are subsets of Oracle Server Manager which supplies inspection function and SQL input function. The user can write a C/C++ program, complete data query with program interface OCI, then export the needed data to program file. The user of Microsoft can use ODBC to import data from Oracle to MS Excel spreadsheets. In addition, a web browser window can be established to access Oracle with tools supplied by Oracle or C language. Data can be imported with Java GUI (Graphical User Interface). In addition, there are two-dimensional or three-dimensional image display tool offered by a third party. It does a great contribution to the offline analysis of the data, as well as the control of accelerator operation and quality of the control system.

APPLICATION OF DATABASE

Equipment Management

One of the most important purposes of the database is the management of accelerator devices. Except the data tables of various devices established in Oracle, the database server net which based on web will be developed in CSNS. Through the network, the user can inquire

device data directly. Further, management and analysis system of local devices will also be developed to facilitate the analysis of equipment in each system.

Project Management

In the traditional construction of an accelerator, database is used as a subsidiary data storage tool. Datasheets from various departments are added to the accelerator database with the progress of the project. The disadvantage is that the datasheet is made by different engineering staffs who have different bias and no unified data monitoring mechanism is provided. Thus it is hard to share data. Furthermore, data tidying and submission will be delay due to the tight schedule.

In view of these shortcomings, data-driven workflow approach will be adopted in CSNS. The content and format of datasheet will be constituted in the prophase of the project. Then corresponding plan will be established according to the flow of collection, tidying, submissions and inspection. This project management can not only monitor the progress of works, but also greatly improve the quality of accelerator data to avoid data delays and data loss. There is an example of database application for magnetic measurement process as shown in figure 2 [3].
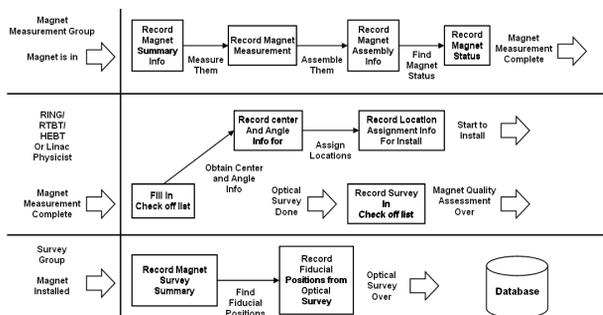

Figure 2. An example of project management.

It should be noted that, because there are many uncertain situations in projects, the establishment of data tables should make allowance to meet special events.

Real-time Data Storage

In EPICS system, the real-time data from IOCs is archived with Channel Archiver offered by EPICS or database software supplied by a third-party generally. While this approach can be realized for the great mass of data archiving function, in order to achieve a unified data management, the increasing laboratories choose to develop their own software to import EPICS real-time data into Oracle global database directly. The IRMIS [4] software developed by Argonne lab is relatively mature and is adopted by many laboratories. This real-time storage method may be taken into account in CSNS.

ELog

The elog of the CSNS accelerator data will track the commissioning path of the physicist. Transit the real-time

data in various components and sensors of the accelerator, the work log is recorded in a decided time with this software. Once the components found in abnormal, the alarm will be issued. E-mail or phone message will be sent to corresponding staff and the error log will be recorded.

High Level Application Software

There are plenty of interfaces supplied by Oracle database for different programming language. The XAL and its improvement version developed by SNS will be mainly used in CSNS accelerator applications. The XAL is written by Java which provides many manners to connect Oracle. [5] CSNS will mostly choose the JDBC+DATABASE connection manner as shown in figure 3.
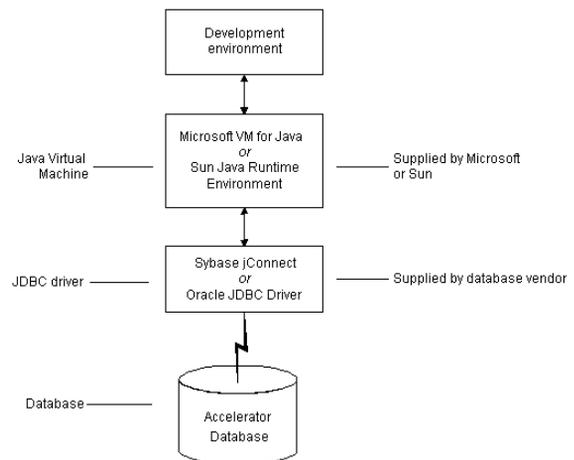

Figure 3. Connect to the database via the JDBC.

In addition, the object-relational mapping (ORM) is used prevalently in the rapid development environment. The open source tool which applied most widely in ORM is Hibernate. Therefore, Hibernate connection manner will also be taken into account. [6]

CONCLUSION

The CSNS will be constructed with the standard rule and interface. This will facilitate the database development and maintenance.

REFERENCES

[1] Jeff Patton, "Rdb systems at SNS", EPICS Meeting Talk, 2005.
[2] http://searchoracle.techtarget.com
[3] J. Wei, et al, Magnet & Survey Database for the SNS Accumulator Ring, SNS, 2000.
[4] http://www.aps.anl.gov/epics/irmis/index.php
[5] J. Galambos, et al, SNS Global Database use in Application, SNS, 2003.
[6] W. Clay Richardson, et al, Professional Java, JDK5 Edition, 2006.

Data and Information Management