CONSTRUCTION AND APPLICATION OF DATABASE FOR CSNS

October, 2007
http://gcsns.ihep.ac.cn/

Q.Gan, C.H.Wang
GanQ@mail.ihep.ac.cn
Institute of High Energy Physics
P.O.Box 918, 100049, Beijing, China

P.Chu
Stanford Linear Accelerator Center
Menlo Park, CA 94025, USA

China Spallation Neutron Source (CSNS)

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.

Datase Overview

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 and alarm information, operating log and so on. Some laboratories divide the database into multiple parts, or create more databases 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.

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 part 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.

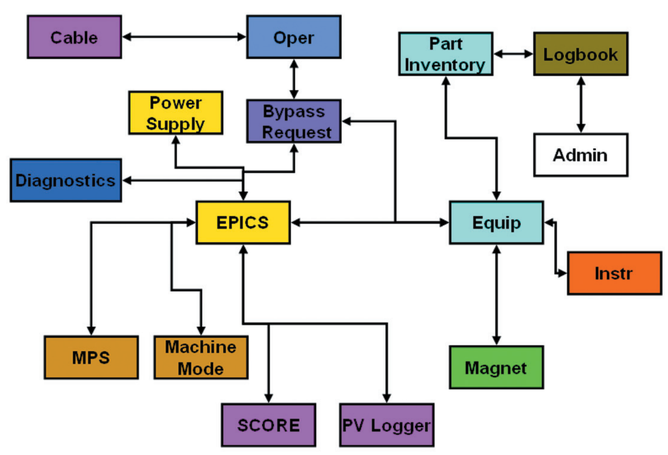

Figure 1. Layout of CSNS database model

- 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
  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)>.

In order to distinguish these different type entities, add the prefix for the entity name. The prefix of the table field is the domain name initials. The entity name and the field name are composed by 1-3 words. Every entity has an acronym 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.

Table and Key

The third normal form (3NF) is a good choice for the accelerator database. However, according to the experience of other laboratories, it's unnecessary to use normal form sometimes.

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. 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 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 tool which is offered by Oracle. The control files of it 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.

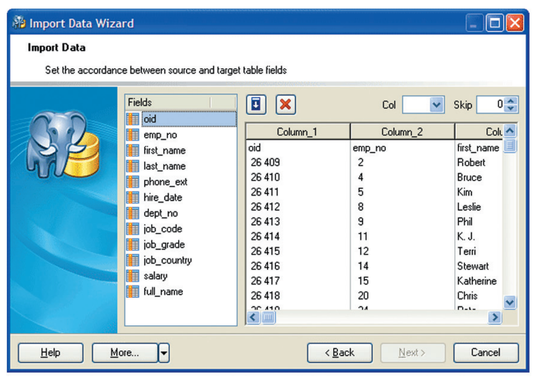

Figure 2. A CSV tool which can import CSV files to Oracle database

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. 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.

Database Application

Equipment Management

The most important purposes of the database is the management of accelerator devices. Except the data tables of devices established in Oracle, the database server net which based on web will be developed in CSNS. Further, management and analysis system of local devices will also be developed.

Project Management

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.

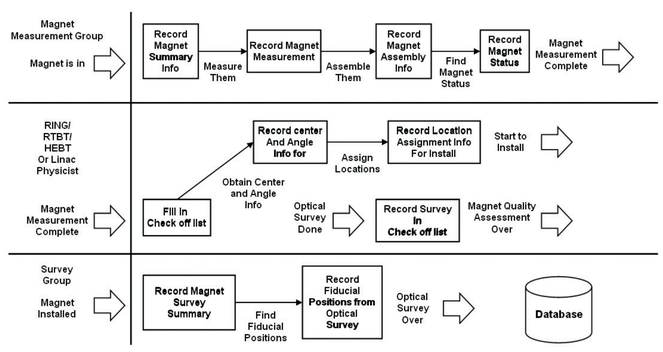

Figure 3. An example of database application for magnetic measurement process

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. The IRMIS 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.

High Level Application Software

The XAL is written by Java which provides many ways to connect Oracle. CSNS will mostly choose the JDBC+DATABASE way. In addition, the Hibernate, which applied most widely in ORM, will also be taken into account.

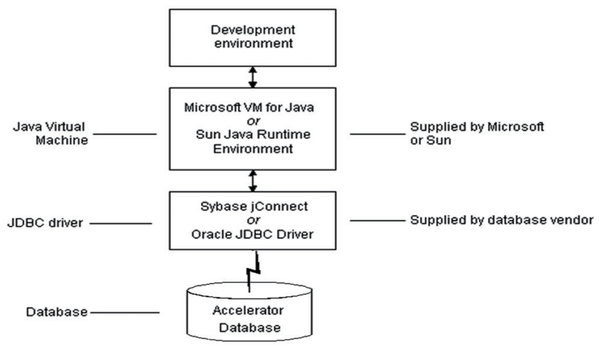

Figure 4. Connect to the database via the JDBC

Conclusion

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

http://neutrons.ornl.gov/conf/icalepcs07