Mission Critical Database for SPS Accelerator Measurements

R. Billen, E. Hatziangeli, I. Laugier, I. Reguero, N. Segura
CERN - European Laboratory for Particle Physics, Geneva, Switzerland

ABSTRACT

In order to maintain efficient control over the hadron and lepton beams in CERN's SPS accelerator, measurements are of vital importance. Beam parameters such as intensities, positions and losses need to be rapidly available in the SPS control room to allow the operators to monitor, judge and act on beam physics conditions.

For the 1994 SPS startup, a completely new and redesigned measurement system based on client and server C-programs running on UNIX-workstations was introduced. The kernel of this new measurement system is an on-line ORACLE database.

The NIAM method was used for the database design as well as a technique to tag synchronized data with timeslots instead of timestamps. A great attention was paid to proper storage allocation for tables and indices since this has a major impact on the efficiency of the database, due to its time-critical nature. Many new features of Oracle7 were exploited to reduce the surrounding software.

During the 1994 SPS physics run, this new measurement system was commissioned successfully and the infrastructure proved to be acceptably reliable. Hence, for the 1995 startup, the size of the measurement system was increased drastically to fulfill a variety of measurement needs. This proliferation of measurements beyond the initial scope showed the correct design of the system, as well as the performance limitations within the actual hardware configuration.

This paper describes the overall design and discusses performance issues of this critical system.

1 INTRODUCTION

The functionality of CERN's Super Proton Synchrotron (SPS) has changed drastically since it's commissioning in 1976. Originally designed as a 1×10^{13} particles per pulse (ppp), 300 GeV proton accelerator, nowadays 4×10^{13} ppp proton beams at 450 GeV are used for neutrino and other fixed target physics, lead ions reach 33 TeV per ion and 20 GeV lepton beams are injected into the Large Electron Positron Collider (LEP).

These different particle beams are guided and controlled by a team of operators in the SPS control room. Their observations, decisions and actions depend on a large number of measurements. For the high intensity, high energy proton beams, an especially correct and rapid knowledge of the beam parameters is vital to the effective operation of the accelerator. At any time the measured beam intensity, transmission, position, loss, etc. have to be available in the SPS control room.

Since 1976, equipment control and measurements were handled by a network of geographically distributed minicomputers. Since then, the SPS has evolved enormously and so has computer technology. For the 1994 SPS startup, the original control system was replaced completely by a modern network based on workstations and local microcomputers. With this transition, the original measurement system became obsolete. A project was created to reproduce and improve the functionality of the SPS measurement system [1]. Because of the need to progress, the urge to standardize and the availability of in-house knowledge, an on-line ORACLE database was chosen as the kernel of this new measurement system.

2 DATA DESCRIPTION

One can distinguish particle *beam data* and hardware *equipment data*. Both can be essential, but the beam data is used for tuning and optimizing the SPS performance. Data on equipment is important as well for the operators, but in most cases malfunctioning equipment will generate warnings or alarms and can even inhibit beam requests by means of hardwired or software interlock conditions and therefore it does not need as much attention.

Beam data

The first set of data concerns the most essential beam information related to the operation of the SPS:

- intensities in the injection lines
- intensities throughout beam acceleration
- intensities and positions in extraction lines and on the targets
- losses around the SPS
- losses in the extraction zones

The rate of availability of this data should be equal to the repetition time of the SPS supercycle (i.e. 14.4 or 19.2 seconds, depending on the physics program). The term “*supercycle*” is used, since it embodies a cyclic sequence of hadron and lepton cycles. Thus all beam data can be considered synchronous with respect to a specific SPS supercycle.

Equipment data

For 1995, an additional set of hardware information was added to the measurement system, mainly concerning detailed information for electrostatic (ZS) and magnetic (MS) extraction equipment [2]. The most important items are:

- ZS spark rate and occurrence within the supercycle
- ZS generator high voltage and current
- ZS anode currents of individual tanks
- ZS and MS anode, cathode and girder positions

The rate of storage for this equipment data is not necessarily equal to the SPS supercycle rate. Spark data is driven by the event and for fairly static data such as voltages, currents and positions, change-threshold values can be defined to avoid storing constant data.

3 DATABASE DESIGN

Since the design of the LEP logging database [3,4] has proven its correctness, a similar approach was taken for the database design. The difference between a measurement system and a logging system is mainly the data storage rate and the length of time for which a history is kept. Therefore this has only an influence on the implementation.

Design method and tool

For data modeling the Natural Information Analysis Method (NIAM) was used [5]. In a NIAM schema one defines objects, facts between objects, constraints on facts and between objects. This data modeling method was preferred above the Entity Relation Diagram method (ERD) because of the binary relationship approach, completeness and clarity of the method.

The RIDL* tool [6], based on the NIAM method allows a rapid creation of Oracle6 tables. Unfortunately, since the company has ceased to trade, the RIDL* tool has not been supported for Oracle7 compatibility, which implies editing of the files generated in order to make them compliant with the new syntax.

Tagging data with timeslots

In principle, one can build a measurement system in such a way that only the most recent measurement is stored and periodically overwritten with fresh data. To facilitate periodic statistics and off-line analysis, it was agreed to store some *measurement history* data for at least one week. One week can be considered as a sequence of 42000 14.4 second supercycles, so each individual supercycle can be identified with a unique integer, called “*timeslot*”. Since each measurement can be related to a specific supercycle within one week, one can refer to this measurement with the corresponding timeslot. A single table contains the correspondence between the timeslots and the actual supercycle start time called “*timestamp*”, while all tables containing the data have the timeslot as primary key.

Using timeslots has major advantages: it minimizes data storage and optimizes speed and ease of data retrieval.

Row packing

For measurements returning a large number of similar data (e.g. 216 beam losses from ring loss monitors), a logical grouping has been applied (e.g. 6 sextants of 36 monitors) rather than having one row per individual measurement. With this row packing technique, the column names are general (e.g. “loss_1”) and a group identifier (e.g. sextant number) is part of the primary key.

This technique mainly reduces index storage space, without penalizing access speed.

Object definition

Data volumes for tables and indexes have been calculated for one week's worth of data, taking into account the actual size (precision) of each individual data item. To avoid tablespace fragmentation, tables and indexes have been created with well defined storage parameters. At creation time, all tables have been filled with a sequence of timeslots and dummy, real size data. This is the only moment where data is *inserted* in the tables. Any other process writing to these tables may only *update* the existing data and timestamps.

An automatic roll-around is implicit: data and timestamp for the first timeslot will be overwritten when the maximum timeslot has been reached.

Primary keys on timeslots and unique keys on timestamps force index creation and imply fast data access.

All objects have a select grant to a user account without resources. All users interested in the measurement data only have access through this account with appropriate synonym definitions and therefore see the data as read only.

4 DATA GATHERING

The SPS measurement system uses the SL Measurement and Logging System infrastructure [7], and it comprises the measurement server (*msrv*) and the measurement black boxes (*mbb*s). The main aim of the measurement system is to have a single process to acquire measurement data for any given hardware system and to provide the data to one or more application programs known as *measurement clients*. TCP/IP sockets are used for inter-process communication between the measurement server and the measurement black boxes.

The measurement server

The measurement server is a single process which acts as a multiplexer between *mbbs* acquiring data from the hardware and measurement clients demanding this data. Once a client has asked for a measurement, the *msrv* starts off the appropriate *mbb*, if it is not already running. The *msrv* monitors the state of all the *mbb*s and guarantees that they are running.

The measurement black box

There is a single measurement back box (*mbb*) for each measured system, for example Beam Current Transformer (BCT), Secondary Emission Monitor (SEMs), couplers. The *mbb* performs single or periodic measurements, synchronized with the start of the SPS supercycle. The measured data is written to the database where it is available for a period of 7 days, after which it is overwritten by the roll-around mechanism.

The measurement client

Any program requiring data from the measurement system is a called a measurement client. A measurement client can specify the frequency of the measurement, how recent the data should be and whether the data should be written to the database or whether to receive it directly from the *mbb*. For specific clients such as the Fixed Displays (see section 6), data availability has been improved by using the latter, bypassing database accesses and avoiding synchronization problems.

Making use of new Oracle7 features

In order to reduce the work of the measurement software, a part of the logical work can be performed inside the database. This results in a reduction of network traffic, database access and software maintenance, while database security and integrity is maintained.

These new Oracle7 features (see section 5) have been applied to the 1995 equipment data. The software only updates a measurement table containing a single measurement. A *trigger* can decide on the relevance of the data (e.g. freshness, threshold), if the measurement data needs to be cascaded into a long term history table by means of a *procedure*. A simple *function* can calculate the right timeslot for the history table concerned.

5 DATABASE IMPLEMENTATION AND ADMINISTRATION

Hardware and Operating System configuration

The platform chosen for the SPS Measurement DB server was the one on which the DB support section has standardized: a SUN running the Solaris 2 operating system. The main reasons for choosing this platform are: a stable UNIX OS, support of multi-CPU servers, the excellent ORACLE support and a good price-performance ratio.

The model chosen was SPARC 10/51 server with one CPU module and 64 Mbytes memory, one internal 1 Gbyte system disk and one external enclosure with four 1 Gbyte disks for the data. The disks on the external enclosure were mirrored using the Sun Online:DiskSuite 2.1 software in order to be protected against single disk failures. The network interface was FDDI, using a 100Mbit token ring protocol running on fiber optics, directly connected to the accelerator control network. The system software running on this configuration was Solaris 2.3.

Reaching the system limitations

The original configuration was sized for the SPS measurements application and ran happily for the whole of 1994, offering the real time performance required and proving to be very stable. At the beginning of 1995 a second SCSI bus to support a second external enclosure with two 2.1 Gbyte disks was added to host LEP measurement data.

Last minute changes in the LEP physics program required urgent expansion of the LEP measurement system, which brought the system capacity required to be above that of the original design. The system came under high pressure, degrading performance for all users. At first, 64 Mbytes of main memory and afterwards a second CPU module was plugged in, giving the equivalent of a SPARCstation 10/512 (i.e. the two CPU model).

With this configuration Oracle started being unstable, freezing frequently. The combination of performance pressure, Oracle level, Operating system level and multi-CPU hardware seems to trigger bugs in the multi-threading and asynchronous I/O support of the Oracle DBMS, but no clear answer on these problems has been given by ORACLE or SUN.

The mission critical nature of the service for SPS/LEP operations prevented more debugging or upgrading the software and an urgent solution had to be set up using a four CPU SPARCcenter 2000 with 256 Mbytes of main memory. The situation has stabilized since then.

Oracle installation

Since the ORACLE RDBMS is the software chosen for the management of databases at CERN, it is being used for the Measurement Database, running currently release 7.1.6, recently upgraded from 7.1.4. The upgrade was done in order to cope with performance problems in the old hardware. The database uses some of the new functionality provided by the RDBMS, such as:

- *Stored procedures* : the possibility of keeping compiled PL/SQL code in the database which can be easily shared between different applications.
- *Multi Threaded Server* : this feature available with SQL*Net V2 offers the possibility of handling a large number of connections the database with a reduced number of “shared” processes. Previous versions of SQL*Net (V1) always created a dedicated server for each connection request.
- *Alarms* : it is possible to send ORACLE alarms between sessions connected to the database, allowing a more intelligent way of handling communications between the applications. This interchange of information between sessions is done entirely inside ORACLE.

The database was configured according to the tight constraints imposed by the application needs. Since the database has a very high I/O activity, the main effort was put onto a careful distribution of the I/O system following ORACLE standard recommendations in this field (e.g. separate index/data in different disks). Moreover, other parameters that may have an impact on the performance of the database (e.g. amount of memory dedicated to ORACLE buffers) are also constantly monitored and tuned.

6 DATA RETRIEVAL

The data retrieval facilities are an extremely important item in the measurement system, since the data flow out of the database is much larger than the data flow into the database. This is inherent to the desired concept where only one process reads out the equipment, stores it and thus makes it available to all potential users. *Fixed displays* in the SPS control room are permanent users, while an interactive *graphical interface* serves occasional users.

Fixed Displays

In the SPS-LEP (SL) division at CERN, a local teletext service is configured for up to 899 teletext pages, which can be distributed all over CERN and displayed on television monitors with teletext capabilities [8]. Four pages are used to display particle beam information in the SPS complex. Derived data such as beam transmission and loss percentages are shown for ease of interpretation. Teletext features such as colors and special fonts are used for clear visualization.

Making use of the same technique as used for the LEP Graphical Fixed Displays [9], currently 3 X-terminals display on-line graphical data. The data concerns particle losses around the accelerator and in the extraction zones, as well as sparks in the electrostatic extraction equipment. The displayed information is highly important for the health of the equipment.

The display layout, data source and destination is completely data driven and defined in dedicated database tables. A single process picks up the definition parameters and creates the dynamic SQL to fetch the desired data.

The Graphical Interface

The potential number of interactive users of measurement data is large, therefore it is a good idea to provide a standard way to access the data. In 1993, the Graphical User Interface to the Logging System (GUILS) was built to access the LEP Logging database [10]. GUILS is a C-program where data retrieval is performed by Pro*C routines and where the graphics part makes use of the CERN X-Window User Interface Management System [11]. The functionality of this interface was extended to query the measurement database which is physically different but logically similar.

With GUILS available on HP-UX workstations as well as on CERN's PC-DOS based Office-LAN, users can select a set of data of interest and choose a certain time period. The data retrieved is displayed in an externally spawned window. Standard features such as zooming, printing, input from a file and output to a file are available. Advanced features such as graph arithmetic and transformations are not pursued since commercial software packages can be invoked on data retrieved by GUILS.

7 CONCLUSION

This new SPS measurement system, based on the measurement database, has completely replaced the original. The availability and reliability of the beam measurements had reached a satisfactory level towards the end of 1994. However, the implementation of a multitude of additional measurements in 1995 , beyond the scope of the initial design, resulted in a noticeable performance degradation. The current ORACLE server configuration is unable to handle this extreme load and therefore a set of vital data is not rapidly available in the SPS control room. This makes the system to be unreliable from the operator's point of view and thus unacceptable.

A number of hardware and software modifications on the ORACLE server platform have been introduced in order to improve the overall performance. At the present time, the measurement system is still at the limits of the acceptable performance level. Since scalability is a key issue, a large effort continues in that direction to be able to handle future measurement requirements.

ACKNOWLEDGMENTS

The authors would like to thank A. Silverman for hardware support and management, G. Wood (ORACLE UK) for performance tuning on the database and R. Bailey and J. Poole for checking contents and form of this paper.

REFERENCES

[1] E. Hatziangeli, G. Morgan, “The New SPS Statistics and Fixed Displays User Requirements and Software Requirements Document”, SL/Note 95-92 (CO), 1994
[2] A. Ferrari et al., “Specifications for the extension of the surveillance and diagnostics of the SPS extraction channels”, SL/Note 95-12 (BT), 1995
[3] R. Billen et al., "LEP accelerator logging system using on-line database", Nuclear Instruments and Methods in Physics Research , Section A 352 , pp. 296-299 , 1994.
[4] R. Billen, “The evolution of the LEP Logging Database”, Paper to be presented at the IADBG Workshop on Databases in Accelerators, Chicago, November 1995
[5] J. Wintraeken, “The NIAM Information Analysis Method - Theory and Practice”, Kluwer, 1985
[6] RIDL* (V1.2) manual set, IntelliBase nv/sa, Antwerp, Belgium, 1991
[7] E. Hatziangeli, F. Roeber, "A Description of the SL Measurement and Logging System and the Application Interface Libraries", SL/Note 93-38 (CO)
[8] Ph. Nouchi, private communication
[9] M. Lamont, private communication
[10] F. Bordry, J. Fritze, P. Ninin, “Graphical User Interface for Logging Systems”, CERN Note, 1993
[11] M. Vanden Eynden, “The X-Window User Interface Management System at CERN”, paper presented at ICALEPCS, Chicago, November 1995