DATA MANAGEMENT FOR LEP CONTROL

J.-P. Koutchouk and J. Poole
LEP Division, CERN,
Geneva, Switzerland

The LEP data include information relevant to the construction, installation, operation and maintenance of the accelerator. This paper describes their proposed organisation into a unified semi-integrated system. To provide the required efficiency and flexibility, the data will be grouped into various structures, optimised with respect to their functionality. Interactive databases are the primary source of information; these will also allow relational access to the archive of the control system. Live data structures, initialized from the interactive data, will allow simultaneous manipulation of three independant representations of the accelerator (reference, current and target machines). A user-friendly access to the control system entities will be provided by a Data Dictionary.

1. Introduction

The size of the LEP project has put particular emphasis on the organisational aspects of design, testing, installation and maintenance. In order to cope with the complexity, modern information management techniques have been introduced (automatic testing and data collection, etc...). The resulting wealth of information has been organised into databases pertaining to the various LEP subsystems, and must somehow be made available to the LEP control system.

The successful LEP operation requires efficiency, reliability and flexibility; with respect to the data structures, these requirements translate into fast and simple access, selective protection, adaptability to a changing environment and provisions to cope with the internal logic of accelerator operation and studies. This paper describes the logical model of a semi-integrated data organisation that would fulfil these aims.

2. The LEP data organisation

The LEP data do not share the same source and present differing access requirements; to simplify this problem, we have categorised the data according to their source (figure 1).

2.1 The interactive data

They are produced either manually or automatically, but off-line with respect to accelerator operation. Their volume is not constrained and the information stored may be only partly relevant to accelerator control. The access time not being critical, they can benefit from a database management system that relieves the user from programming, allows sophisticated access and easy maintenance. The commercial RDBMS Oracle [1,2] has been chosen for that purpose. The interactive data encompass :

- The equipment databases, which contain a description of each LEP equipment (already partly implemented).
- The system databases, which contain information relevant to the integrated system.
- The accelerator database which contain the inputs to the accelerator model.
- The archives, which permit the accumulation and cross-referencing of the real-time data.

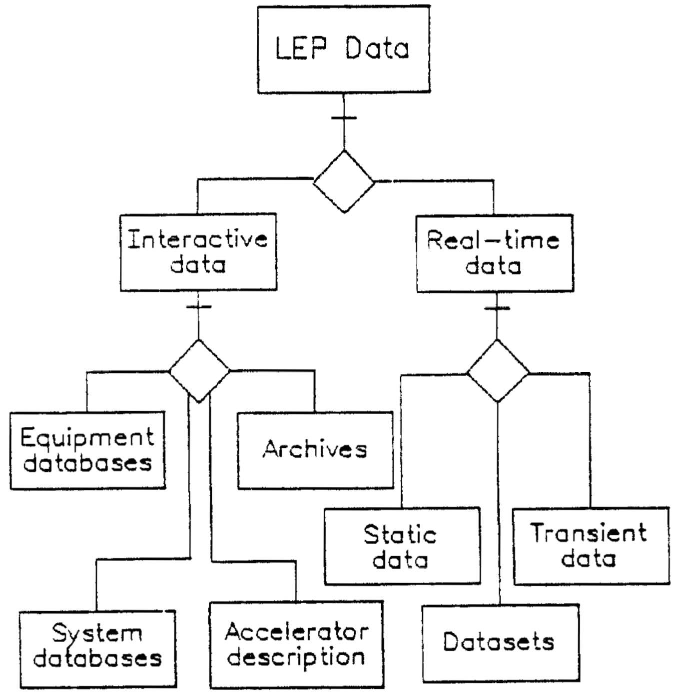

Fig.1 - General description of LEP data following SASD rules (see Annex).

2.2 The real-time data

The real-time data are either produced or consumed within the control system. Their volume must be limited and the access time minimized; they are organised into three subcategories in descending order of complexity :

- The Datasets which map the operational requirements into the dynamic data structures.
- The Static Data which are copied from the interactive databases.
- The Transient Data which are simply stored in the filing scheme.

3. Interactive Databases

The important characteristics of the data kept in Oracle databases is that they are and should remain the primary source of all data for LEP. Responsibility for their maintenance will reside with the owner of the equipment or software which they define;they are however considered operational, i.e. accessible by many authorised users (application programs).

3.1 Equipment Databases

Grouped under this title are all of the databases maintained by the LEP equipment groups. The tables will describe the equipment in terms of its composition, physical location, hardware address, calibration constants, allowable range for their parameters, status (up, down, suspect), etc.

The databases will undoubtedly also contain much information (such as ordering and manufacturing details, history of faults, ...) that will not be of direct use in machine operation and therefore only the relevant data will be mapped into the control system.

3.2 Accelerator Description

Many application programs will need to know the structure of LEP, either optically or physically. The source of its description will be found in Oracle tables. At the heart of the machine description will be the MAD [3] input data detailing the constituent electromagnetic elements and their sequence (accelerator structure), together with a set of excitation tables (configurations). These tables will be joined with others describing the location of non-active machine components so that ultimately a full description of the machine and its components locations can be built.

In conjunction with the equipment databases, the machine description will allow powerful relational enquiries about the machine structure.

3.3 LEP Dictionary

Amongst the information which are relevant to the entire LEP complex, the method to access (address) hardware and software objects has been given attention. Given the very large number of individual LEP items, one needs long names to designate them. To help the mnemonics, the names presently used strongly reflect their significance to a specific group of people : installation uses 10-character names which specify the geographical location; designers use other names of the same length which specify their functional location within the LEP complex. Accelerator physicists and operators require yet another naming convention, closer to their natural language; moreover, the relation between accelerator parameters and hardware parameters is often not one to one and requires tree structures. The LEP Data Dictionary is meant to store the relationships between the various naming conventions and the corresponding network addresses. The Oracle RDBMS offers the necessary facilities to create, maintain and modify the reference copy of the dictionary; the definition of tree structures (e.g. the sextupole power converters; composed of F and D sextupoles; composed of SF1, SF2 etc.) is a normal Oracle construct and the control of uniqueness is similarly a standard Oracle tool.

3.4 Archives

To allow a better understanding of the development of machine performance, it is planned to archive beam and machine parameters. An index to these data will be constructed to allow queries in accelerator physics terms. This index will contain a summary of the accelerator performance and pointers to the archived data.

The bulk of the logged data is extracted from the dynamic data structures and archived within the control system itself. The archival software includes a number of steps to facilitate subsequent analysis of the data : it sets up pointers to the archived data files; it checks that the data is stamped with sufficient information to reconstruct the accelerator status at the time that the data were collected; it archives the complementary information and initiates a remote transaction with Oracle which creates an entry in the Archive Index database. This entry carries the name of what was saved and is stamped with data which concisely but precisely define the accelerator status, including pointers to larger sets of data. There is also provision for the entry to be labelled with a comment.

The Oracle database will thus cross-reference beam and machine measurements and the environment in which they were made. Use of the database will be through an Oracle interactive application program to retrieve pointers to files stored in the control system or to retrieve information stored in the Oracle database itself.

As an example of retrieving pointers one might have a query like "Find all orbits and tunes for which the luminosity exceeded a certain value", or "Find over the last 100 runs the collimator settings which gave a low background in a specified detector".

4. Real-time Data Structures

4.1 Datasets

A particular structure has been imposed on the dynamic data, following our experience of running the ISR :

- Data gain significance when related to other data : for example, a closed orbit measurement ought to be related to a machine configuration, a set of orbit corrector excitations and a tune value. To clarify these relations, the dynamic LEP data will be arranged in Datasets. A Dataset may be viewed as a sequence of composite data structures corresponding to the identifiable sequence of states of a LEP cycle (run); each of these structures contains related information (machine status information, equipment settings and readings, calculated and measured beam and machine parameters).

- Experience shows that reliable and flexible storage ring operation requires the simultaneous manipulation of three machine representations : the Reference Machine, not necessarily the best machine but more a proven starting point, the Current Machine, possibly different from the reference, and the Target Machine, resulting from calculations (for instance corrections) and possibly requiring some checking before any action is undertaken. Three Datasets are associated with these three representations of the machine (figure 2).

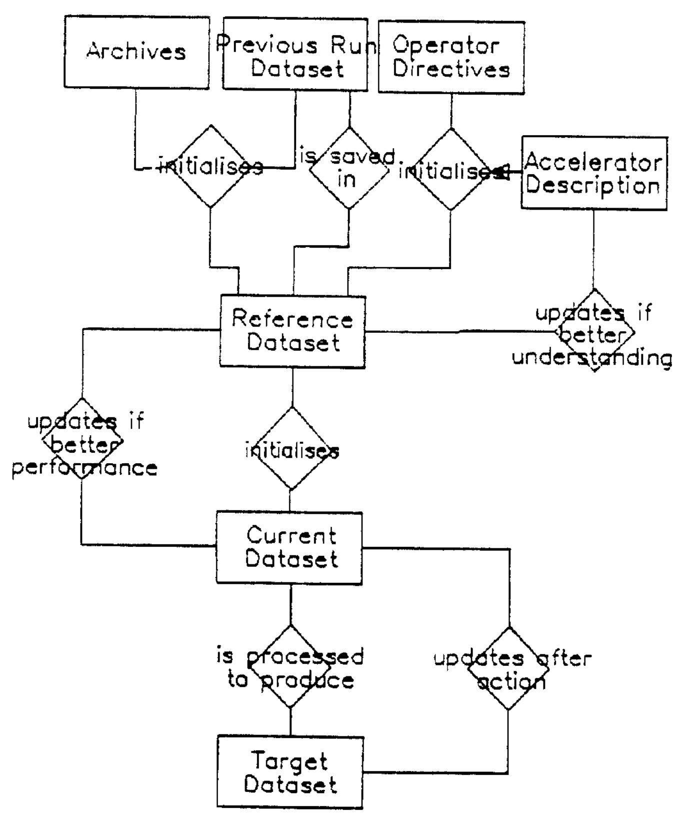

Fig.2 - Relationships between the LEP Datasets (ERD notation : refer to Annex)

A fourth Dataset keeps available the Reference Dataset of the previous run to allow a future automation of the storage ring operation.

The datasets are also mirrored in temporary files to make a 'virtual machine' so that off-line simulations can run with real machine processes and data without interfering with operation.

4.2 Static Data

All the information related to individual pieces of equipment or software which are unlikely to change during one run (e.g. calibration constants) are simply extracted from the interactive databases and distributed within the control network. This operation is foreseen to be done before each run. Once distributed, this information is undistinguishible from private data and does not require any further data management.

The LEP Dictionary is an important member of this class.

4.3 Transient Data

The transient data are only relevant to one process and do not require more than the standard file system.

5. Operation of the data structures

It is planned to initialize the real-time data at the beginning of each run by remote procedure calls to the Oracle RDBMS. Data coming from Oracle will be mapped into control system files [4] and will become the source of all information for the duration of a run. The rate of change of these data is expected to be very low so that data consistency should not be a problem. Critical data, if any, may be supervised by a process which would inform the operator of a change in the database.

Once the control system is initialized with the interactive data, a Reference Dataset is created; it may equally be copied from the last run allowing autonomous production running; alternatively an access to the archives would allow a reloading of the corresponding dataset from any previous run. The Reference Dataset is meant to be gradually improved following operational experience. In case of difficulty, it may always be recalculated using the accelerator model and description.

The Current Dataset is initially loaded from the Reference Dataset and then follows the actual machine status. It also records the history of changes and collects the logged information.

In case of a modification to the accelerator settings, the calculated or estimated modifications together with the predicted new characteristics are stored in the Target Dataset for checking and approval. For instance, the calculated increments of the orbit correctors together with the new predicted orbit would be stored there prior to any action. The last changes made would thus be always available.

It is foreseen to automatically archive the logged information and the history of changes at three break-points of the run : after accumulation, after ramping and at the end of the run.It is estimated that the volume of data to be archived each run will be ~ 0.5 Mbytes and therefore a long period can be kept on-line. The volume of data sent to the Oracle Archive database only amounts to ~ 1 kbytes per entry in the index.

The life of an accelerator is made of runs and shutdowns. The activity on accelerator components follows almost exactly a complementary cycle. It seems thus judicious to separate the interactive data from the control system, because they are mainly concerned with the equipment. In this way a non-availability of the control system during shut-downs would not impair the work carried out on LEP components; conversely a non-availability of the interactive data during running would only degrade the flexibility of operation and possibly the data consistency.

6. Conclusion

The proposed LEP data organization implements the well-known computer science concept of separating the data from the processes. We have tried to avoid an abstract approach by systematically taking into account the meaning and use of the data; their logical organisation aims at fulfilling the expected requirements of the various parties (design, maintenance, physics production and machine studies). The requirements in the accelerator field are fluctuating; we have tried to identify the invariants and provide sufficient flexibility to cope with the accelerator evolution.

ANNEX

The Entity Relationship Diagram

The Entity Relationship Diagram (ERD) is a rather straightforward but powerful data representation; it is used within the context of Structured Analysis (SASD) which has been adopted for the LEP software. The boxes represent data entities, the diamonds and the lines express the relationships between the data; diamonds with a crossed line are used to represent data hierarchies and should be read " is a subtype of"

REFERENCES

[1] M.C. Crowley-Milling, The Control System for LEP, IEEE Trans. Nucl. Sci. Vol. NS-30, No 4 (1983).

[2] Oracle, product of Oracle Corporation, Belmont, California, USA.

[3] E. Keil, F.C. Iselin, The MAD Program (Methodical Accelerator Design) Reference Manual, CERN/LEP-TH/85-15.

[4] L.A. Tausch, The Distributed Database for the LEP Control System and the new SPS Control System, LEP Controls Note 66 (Revised) 1986 (unpublished).