

# Very Large Data Bases

## FIFTH INTERNATIONAL CONFERENCE ON VERY LARGE DATA BASES

SPECIAL ISSUE ON  
WORK IN PROGRESS  
IN DEVELOPING COUNTRIES

Editors:  
Antonio L. Furtado  
and Howard L. Morgan

005.7406  
I61

October 3 – 5, 1979  
Rio de Janeiro – Brazil

An experience with the ANSI/SPARC

approach to database design

Claudio A. Teixeira - Marinha do Brasil

Rubens N. Melo - Pontifícia Universidade Católica do RJ

Brasil.

This paper reports an experience of designing a database system following the three-schema approach suggested in the ANSI/SPARC Report (1). Such system is an integrated budget-planning system (2), which is proposed as an alternative to the current budget-planning system used by different government agencies.

Many people understand from the ANSI/SPARC Report that the conceptual schema should be defined first. This schema should be a central control point, an unconstrained long-term view of the enterprise, a sort of common denominator between optimized storage descriptions (internal schema) and the multiple user views (external schemas). The external schema defining particular applications' (or application classes) views should be consistent with and mappable from the conceptual schema.

Some people, however, feel that it is very difficult and unreasonable to start by describing the conceptual schema model of a large enterprise. They think that the only way to gather information about the needs of an enterprise is to start by analysing its functions and therefore to consider first specific applications and then proceed to the amalgamation of applications of some area into external schemas which are used to derive the conceptual schema. We may even agree that this latter point of view is apparently the natural approach when integrating existing applications and this kind of "integration"

is indeed a very common trend in many enterprises. In fact we also started by describing the external schemas of our system, but we first realized the existence of three main flows of information basic to the visualization of functions, procedures and responsibilities in the enterprise. These flows are concerned with the three main entity types of the problem, namely, project, credit and cash. Based on these flows we defined twenty two external schemas to support the principal activities such as "credit planning", "financial planning", etc.

In part this policy was also influenced by the project structuring we adopted, which starts by system functional specifications. Thus we first defined the system functions, the users and their prerogatives, access constraints to data, security controls and other aspects related to the input and output documents. At the next phase we were guided by the three flows mentioned above and described the external schemas in terms of their (external) entities and relationships and the transactions and consistency constraints corresponding to the relevant applications.

The conceptual schema was then obtained by the integration of these external schemas. In fact some iteration was needed to resolve some conflicts before we arrived at a final description of both conceptual and external schemas.

At this point we felt that in reality neither the conceptual nor the external schema comes first. The realization of those three flows of information concerning projects, credits and cash provided indeed a sort of a "model of the enterprise" a sort of incipient conceptual schema from which we derived the main external schemas which in turn suggested the basic structure of the conceptual schema and so forth.

## Conclusion

This experience has shown us some points which sometimes are not perceived by database designers.

### First

We should be aware that the integration of existing functions (or systems) tends to induce us to produce a system which will reflect the enterprise as it is now and maybe not as it should be. This policy is a sort of bottom-up approach where we try to build integrated functions on top of others that would not necessarily prove adequate after a new global analysis of the enterprise.

### Second

We should not be misled by typical project structuring habits which are more process-oriented and in a sense give inadequate consideration to data analysis. This may result in a narrowly defined database system which may be unresponsive to the changing needs of the enterprise.

### Third

We should understand that neither the conceptual nor the external schema must be completed first. By now we agree that we should begin to analyse the problem in terms of a general model and try not to think only in terms of today applications. In a way we feel that we should go in parallel through both functional and data analysis in order to obtain an adequate and stable conceptual schema together with its proper external schemas.

The project herein mentioned has not yet been finished. The internal schema is currently under development and the implementation on a particular DBMS is also under investigation.

- (1) ANSI/X3/SPARC - "Interin Report ANSI/X3/SPARC Study Group on Data" - FDT Bulletin (1975).
- (2) C. A. Teixeira - "Um Sistema de Orçamento - Programa para Órgãos da Administração Direta, Apoiado em Banco de Dados" MSc Thesis - PUC-RJ 1979.