distributed data bases

edited by
h.-j. schneider

north-holland

DISTRIBUTED DATA BASES

Proceedings of the Second International Symposium on
Distributed Data Bases
Berlin, F.R.G., September 1-3, 1982

edited by

H.-J. SCHNEIDER
Technische Universität Berlin
F.R.G.

1982

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • NEW YORK • OXFORD

ISBN: 0 444 86474 1

Published by:

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • NEW YORK • OXFORD

Sole distributors for the U.S.A. and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 Vanderbilt Avenue
New York, N.Y. 10017

PRINTED IN THE NETHERLANDS

DISTRIBUTED DATA BASES

Second International Symposium on
Distributed Data Bases
September 1-3, 1982

Organized by
Gesellschaft für Informatik (GI)

in cooperation with
AFCET, AICA, IBM, IEEE, IFIP-TC2,
IGDD, INRIA, Nixdorf, Siemens,
TU Berlin, Universität Stuttgart

Symposium Chairman
H.-J. Schneider

Program Commitee
E. J. Neuhold, Chairman (FRG),
Ph. Bernstein (USA), H. Biller (FRG), G. Bracchi (I),
J. Bubenko (S), S. M. Deen (GB), G. Gardarin (F),
G. Le Lann (F), W. Litwin (F), R. Munz (FRG),
F. A. Schreiber (I), P. Selinger (USA),
R. P. van de Riet (NL), H. Weber (FRG)

Tutorial Session Chairman
G. Gardarin (F)

Organizer
S. Florek (FRG)

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • NEW YORK • OXFORD

PREFACE

Distributed data base systems and their components are becoming of ever increasing interest to our society. The data processing community has evolved towards a recognition of the importance for handling large masses of data in distributed and loosely coupled organizations. Users are now realizing a need for systems which support such distributed systems in industry, civil service and public administration. A new theory of handling distributed control has to be developed. Many indicators show that the development of distributed systems and, in particular, of distributed data base systems as the kernel part of such systems will be one of the major directions in computer science in the forthcoming decade.

In 1980, the first international symposium on distributed data bases was organized by INRIA (Institut National de Recherche en Informatique et en Automatique) in cooperation with many international scientific organizations. The results were edited by C. Delobel and W. Litwin and published by North-Holland.

Now the second international symposium on distributed data bases took place at the "Technische Universität Berlin" organized by the "Gesellschaft für Informatik" again in cooperation with many outstanding associations and institutions.

The conference concerned itself with the design, realization and evaluation of data base management systems on computer networks. Especially the following themes were deeply discussed :

- standardization
- computer networks
- multidatabase systems
- reliability
- query processing
- concurrency
- models and architecture

Three invited speakers gave overviews in the field of standardization issues, computer networks and multidatabase systems. The tutotial part of the conference served as the introduction and discussion of the four well-known distributed data base system proto-types POREL, R*, SIRIUS and VDN. Corresponding outlines are also included in this book.

The 30 papers submitted were reviewed by the Program Committee members (as listed opposite the title page), who together with the Tutorial Session Chairman, G. Gardarin, and the Organizing Chairman, S. Florek, helped in the preparation of the conference. We are greatly indebted for the hospitality of the Technische Universität Berlin.

The help of a number of very well-known specialists as referees provided the basis for the success of this conference. I would like to thank all of them very much for their contributions.

I hope that this book provides a challenge to the information system and data base system community and is a stimulus to researchers, implementers and users.

Hans-Jochen Schneider
Technische Universität Berlin
September 1982

TABLE OF CONTENTS

CONCURRENCY CONTROL

COMPUTER NETWORKS

RELIABILITY II

MULTIDATABASE SYSTEMS

MODELS AND ARCHITECTURES

TUTORIAL

DISTRIBUTED DATA BASES
H.-J. Schneider (editor)
North-Holland Publishing Company

INTERNATIONAL STANDARDIZATION AND DISTRIBUTED DATA BASES

T. B. Steel, Jr.

American Telephone and Telgraph Company
Basking Ridge, NJ, USA

The topic upon which I am asked to report in this paper is a curious one; at the moment it does not exist. There is no active work on standardization relative to distributed data bases currnetly under weigh in the International Organization for Standardization (ISO) and it seems likely that it will be several years before any substantial change in this situation will obtain. Given this fact, I could quite properly say "thank you very much", close the paper here and request an invitation to be issued some years hence.

There is, however, an enormous amount of activity that might be properly described as foreplay to distributed data base standardization. Several contemporary ISO efforts are laying the groundwork for eventual distributed data base standards. I propose to explore these efforts in some detail and attempt a prediction on the nature of their eventual marriage.

That *some* consummation is inevitable cannot be doubted. The issue is too vital to be dismissed; it will not be. Even though the title of this meeting is "Symposium on Distributed Data Bases", I doubt that most participants appreciate the economic and social impact of a successful development resulting from their reseaches. In a pithy and not too inaccurate a phrase, what the confluence of standards in data base and data communications — the essence of distributed data bases — will cause to happen is the establishment of a single, world-wide information processing system to which everyone has direct access. This is the stuff of science fiction, but I predict that by the turn of the millennium, such an artifact will be recognizably in place.

Consider what this entails. Taking an abstract and non-limiting view of the matter, all data — that is, all *recorded* information generated by mankind — will be available to anyone, anywhere with the proper authorization. This is not an idle dream for some back of the envelope calculations suggest that a system accomplishing this could be put into place world-wide with about ten years effort and an investment of some $150 billion. Now, it is not going to happen that way; no massive, single effort using just comtemporary technology. The point is that it *could* be done; no breakthroughs are required. Much work needs to be done to be able to do the job in an efficient and cost-effective manner, but that is engineering.

Let me digress from the main theme of this paper for a moment to substantiate the above observation. How much data *is* there? The answer to this question clearly depends on a precise definition of "data", but reasonable upper limits can be put on the number. [Senders 1963] estimates the non-redundant content of the world's libraries and its growth rate as yielding about $4 \cdot 10^{14}$ bytes in 1985. Of course, not all relevant data is contained in libraries. There are perhaps 10^8 white collar workers in the developed world, including managers. If we assume that they all spend half their time (surely an overestimate) keying data that is permanently captured (also an overestimate), then about 10^{15} bytes per year are added to the world's data base. So, by the year 2000 there will be no more than 10^{16} bytes of data. This ignores data derived from sensors, but examination suggests that this data is a couple of orders of magnitude smaller and much of it is unlikely to be desired as part of a permanent archive.

There is already enough magnetic storage to contain that number of bytes, or at least close to it, and I am confident that the media manufacturers would be glad to oblige by producing more if necessary.

What about access? The public telephone network, the burgeoning cable television systems and the bandwidth that will be engendered by optical fibres should make it clear that a solution to this problem is also at hand. There is no question that it will all be very expensive, but I submit that the demand will be so high that economies of scale will make the cost sufficiently low that the kind of system I am contemplating will be essentially in place by the end of this century.

The important point in the above discussion is that there is no fundamental barrier to a world-wide distributed data base and that the using community will exert sufficient demand to bring such a thing into being in the next two decades. Two subsidiary questions emerge: Can we — the information processing and information handling community — organize this data effectively and can we make all the necessary pieces fit together? I believe the answers to those questions are "Yes" and that it will be the international standards effort that makes it so.

Having said above that there is not at present any standardization work *directly* addressing distributed data bases, I must qualify that in two directions. First, there is considerable work concerned with data bases *per se*, and much of that is applicable to the distributed situation. Second, there is a major effort in the development of standards for arbitrary interconnection of information processing systems. It is the ultimate coupling of these efforts that is relevant to the distributed data base problem.

Consider first the data base managment system situation. The basic paradigm here remains the ANSI/X3/SPARC/DBMS Framework [SPARC 1975] and its three schema model of a DBMS. It should be clearly understood that this Framework is a *model*, not an implementation plan. Among other things this Framework identifies a number of interfaces, some of which are appropriate for standardization. Identification of interfaces was the original objective of the Study Group that developed the model and in that it is fair to say the Study Group was successful. While this model was developed as a domestic activity in the United States, it has guided much of the research community activity in the data base area since its publication. For recapitulations of the Framework, see [Tsichritzis and Klug 1977], [Jardine 1977] and [Yormark 1977].

In addition to its impact on research directions, the Framework was adopted as a basis for its work by ISO/TC97/SC5/WG3, the ISO Working Group on Data Base Management Systems established in 1976. It rapidly became clear to WG3 that the priority item in its work program should be the development of the concepts and terminology for conceptual schemas, the seminal idea in the Framework. It should be noted in passing that among the other items considered but not explicitly included in the WG3 program of work is the study of distributed data bases.

There is other data base work in ISO, specifically ISO/TC97/SC5/WG5, chartered to coordinate data base proposals submitted to SC5 for consideration as possible candidates for standardization. The establishment of WG5 was a result of a recommendation by WG3. Originally WG3 had been given that assignment as well as its other charges, but WG3 realized that it did not have the resources, nor did the active participants have the interest, to evaluate various proposals for data description languages, external schema languages in the terminology of the Framework. It was anticipated that many such proposals might appear. To date only one such proposal has been taken seriously by SC5. In 1981 an experts group on the CODASYL data description language was established. It seems likely that as time goes on additional DDL proposals based on data models other than the network model will surface.

From the point of view of distributed data base management it doesn't really make a gread deal of difference *which* external schema languages become standards. It does matter that there *be* standards for the popular and widely used such langauges. As will be seen below, however, such standards do not have to have the status of International Standards developed by the ISO. It appears likely that conventional market forces will result in the development of a suitable suite of external schema language standards by the time they are relevant to distributed data base applications.

Crucial to distributed data base applications, however, is the development of an International Standard for a conceptual schema language. This is the work whose basis is being laid currently by SC5/WG3. WG3, thus far, has been developing a definition of exactly what a conceptual schema entails [WG3 1982].

It is important to appreciate that a conceptual schema is not a model of data, it is a model of a relevant portion of *reality!* As such, it deals with the semantics of the data, not the syntax. The entities to which terms in a conceptual schema refer are objects in the real world, not bit or character strings. This point has been made repeatedly but is of sufficient importance that it must be repeated over and over again. A conceptual schema is *not* a description of data.

There are a variety of proposals for conceptual schema languages in the literature [Nijssen 1976], [WG3 1982] and the reader is referred there for details. It makes little difference to the distributed data base problem exactly what language is chosen for standardization, provided it is adequate to the job. Adequacy in this context really means being able to model arbitrary aspects of reality. *All* of the semantic content of information that is in the distributed data base must be expressible in the standard conceptual schema language. This places rather severe requirements on such a language.

It is also important to the distributed data base application that the conceptual schema language be formal and so interpretable by a digital data processing machine. Too many proposals for such languages involve attractive but imprecise diagrams which, even if made machine readable, would be uninterpretable by an algorithm and, so, unsatisfactory.

There are, however, several approaches, doucmented in [WG3 1982], that show some promise of being satisfactory. WG3 is currently developing criteria to judge such proposals and it is reasonable to assume that detailed development work on a standard for a conceptual schema language can begin within a year or two under the auspices of SC5.

To summarize the data base situation in ISO, current work is underway on one external schema language (CODASYL DDL) and work is close to commencing on a conceptual schema language.

The question of communications standards is somewhat more involved. ISO/TC97/SC6 has been in being for many years and is charged with the development of standards for transparent transmission of data across a digital data network. While it is not to be denied that this work is essential to the emplacement of distributed data base applications, it became clear a few years ago that there was a great deal more to communication between cooperating application programs in different systems that simply getting bits across the wires accurately.

In 1978 a new Subcommitte was established under TC97, ISO/TC97/SC16, on Open Systems Interconnection. In emulation of the data base management activity described above, the first task SC16 set for itself was the development of a model, now known as the Reference Model of Open System Interconnection. There is now a Draft International Standard being circulated for vote by TC97 Member Bodies and it is documented as ISO/TC97/SC16 N890.

This model consists of a specification of seven layers: Physical, Data Link, Network, Transport, Session, Presentation and Application. It details the services provided by each layer to the next higher layer and from those services it is possible to derive the semantics of the protocols that must be exchanged between systems to provide implementation of the services. These protocols are objects that are subject to standardization and there is considerable activity in SC6 (Physical, Data Link, Network) and in SC16 (higher layers) on the development of appropriate standards. Indeed, there are either in place or near to the status of draft proposals, standards for protocols for all layers up to and including the Session layer. While the distributed data base application will depend on the presence of such standards, so will all other communications and there is nothing in this area peculiar to the data base situation. The same is not true at the Presentation and Application layers.

Before commenting on the data base specific aspects of open systems interconnection, it is necessary to detail the proposed structure of the Application and Presentation layers of the OSI Reference Model. More detail than can be given here is found in ISO/TC97/SC16 N918 and N919, respectively.

To begin with, the Session layer provides a connection between two (or more) application processes that wish to communicate. All of the detailed aspects of interconnection across networks that are independent of the semantics and the syntax of the data being transferred are cared for by the Session or lower layers. The Presentation layer resolves any syntax differences and the Application layer deals with the semantics of communication. Now, while that is very easy to say, it probably doesn't convey much comprehension to anyone not already immersed in the open system interconnection work. Space available prevents a precise account of what is actually going on here. All I can do is paint with a broad brush and hope some of the flavor comes through.

It is important to recognize, as many have not, that the term "syntax" as used with respect to the Presentation layer means *all* aspects of syntax, not just the resolution of differences such as ASCII and EBCDIC. Thus, in the data base application, the Presentation layer will interpret the syntactical aspects of an external schema and take the appropriate action. Thus, the Presentation layer of the open system interconnection mechanism must know the details of all possible external schema languages, either by name or by description. Of course knowledge of an external schema language by name implies the ability to call up from somewhere a description of the named language. Here, then, is the reason for the need for a standard syntax description language. SC16/WG5 is currently studying this problem. It is known to be a soluble problem but the optimum solution — or even a workable solution — is far from obvious.

It is also known that universal translation algorithms do not exist between arbitrary syntaxes. How far this can be pushed is an excellent research topic and should be pursued vigorously. At this stage of our understanding it is a matter more appropriate to academic study than standards committee development.

In a very real sense the Presentation layer is nothing more than a function which maps transmitted data from one syntax to another without altering the semantics of the information. That makes it simple to describe but hardly simple to implement.

The Application layer deals with the semantic aspects of communication. As such it is the only direct connection visible to the application processes. *All* of the requests for service to the open system interconnection mechanism by an application process must proceed by requests for service to points, known as "service elements", in the Application layer. One of the major tasks for all relevant standardization bodies is the definition of the semantics of the service elements.

Four broad categories of service elements can be identified. First, there are the Common Application Service Elements, providing such services as the initiation of a connection with a named peer application process (such as a given data base), the transmission of data to or from a peer application process, the termination of a connection, and so forth. These services are those concerned with pure communication, independent of any semantic content of the data transferred. It is generally acknowledged within ISO that the responsibility for development of standards for the *semantics* of these service elements lies with SC16, while the responsibility for development of the *syntax* required to invoke them lies with whatever ISO group is concerned with the invoking structure (SC5 for the procedure languages such as FORTRAN and COBOL, other ISO Technical Committees for non-procedural techniques that are application dependent).

The second category of service elements are the (badly named) Application Specific Service Elements, providing such services as terminal presentation, file transfer, access and management, job transfer and manipulation, and store and forward messaging. Each of these is application independent (hence my objection to the chosen name), but are related to a particular information processing technique. The semantics of these service elements is in the provenance of SC16 as well, and the syntax resolution is as above. It should be obvious that the distributed data base application is an aspect of the file transfer, access and management service. I will return to that point briefly below.

The third category of service elements are the User Specific Service Elements (again badly named) which are those that are, in fact, application dependent. It will be the application areas of ISO that will do the entire job here; areas such as banking, trade data exchange, airline reservations, and the like. For the record, the fourth category, which has not been given a name, are those service elements defined by a particular using institution for its own purposes. The relevance of these last two categories to the dicussion herein is that they must adhere to conventions established by SC16 in order to fit smoothly into the scheme of open systems interconnection. They will *use* service elements defined in the other categories, invoking them just as a computer program invokes subroutines.

It is important to note that the description of the above categories has been cast in terms of which group develops the standards and, as such, is quite arbitrary. It is not possible, except possibly for the Common Application Service Elements, to give a formal definition of the distinctions. Nevertheless, the discussion is relevant, showing the complexity that will be involved in the total distributed data base application.

What has been said above is just a hint of the nature of the Application and Presentation layers. In a paper as limited as this, the only way to give some real hint of what will happen in the distributed data base environment is to provide an approximate scenario for a single user accessing a distributed data base, making a series of assumptions as the discussion unfolds.

Consider a user attached to an open system — in this context the user is an application program embedded in a computing system whose communications component obeys the protocol rules of the ISO defined OSI standards — who wishes to make an inquiry of *the* data base (from the user's point of view there is exactly *one* data base in the world). The questions here are: What does the user have to do and what standards are necessary to permit him to do it? Given answers to those questions, there are the subsidiary questions of who will develop the standards and when will they be in place.

Taking this step by step, the user must be able to formulate the query; this will require, at least, a *local* syntax for a query language and a standard semantics for inquiry languages. This requires knowing what a valid question is as well as what counts as an answer to the question. That these issues have a formal solution has been demonstrated in [Belnap and Steel 1976]. There is ample theoretical basis for the development of formal query languages. At the moment this task is assigned within ISO to TC97/SC18, largely because it is assumed that the query language problem is a simple matter of office automation — the "office of the future" run wild. Although this venue is not really the appropriate one to make this an issue, it is my belief that strong representation should be made to TC97 that the question of query languages should be in the provenance of TC97/SC5.

In any event, our user knows how to make an inquiry, semantically standard and syntactically local, and to make it in an OSI standard manner. Now we must ask about what information the OSI mechanism requires to interpret this query. First, it must know the local syntax in which the query is couched. Using the technical terms of ISO/TC97/SC16 N890, the Application-entity associated with the using application process must have access to a syntactical description of the local syntax used by our user. The Application-entity must also have the capacity to access an appoporiate directory to determine *where* to direct the inquiry. This is the data dictionary problem; it is being addressed in the United States through ANSI/X3/H4 and there is a suggestion that such work become an experts' group in the ISO.

A key point here is that in order to properly interpret the query and identify the specific pieces of data necessary to construct a response to the query, it is likely in the general case that the conceptual schema needs to be consulted. This will be true regardless of whether the data base is distributed, but in the event that it *is* distributed, the data dictionary, or at least portions of it, may well be distributed itself. It is then imperative that there be a *standard* language for the conceptual schema.

Having located the sources of data items required through consultation of the data dictionary by name, the Application-entity associated with the inquiring application process must request, via the service elements it contains, access to the systems containing those data items in their local data bases. The definition of the semantics of the service elements to do this is an assigned responsibility of SC16/WG5 under its File Transfer, Access and Management Rapporteur's Group. This group has concentrated thus far on the pure file transfer problem, basically that of reading a tape remotely. It is not a major step from there to reading a part of a file, given that the proper part of the file can be located.

A complication in all this occurs as a result of differences in syntax between the source and target systems. As noted above, it will be the responsibility of the Presentation layer to resolve those differences. Given a standard syntax description language, the problem is to translate between the specified local syntaxes. This may involve a mapping between different data models, a not yet solved matter.

If what has been described above sounds all too easy, it shouldn't. Many of the tasks described are trivial to state and very complex to implement. However, the development of standards for open system interconnection, as currently being done under SC16, reduces the distributed data base problem to a reasonably close approximation of a conventional data base problem. If the OSI job is done properly, a data management system will hardly know the difference between a data base that is located in a single system and one that is spread about the world.

What is, of course, not known at the present time is whether the theoretical solutions being dicussed in the various standards bodies will result in potential implementations that arecost-effective and have sufficient performance to be usable. That they will *work* in principle is not to be doubted, but when one looks at the demands on processing capacity and bandwidth, there is reason to be, at the very least, skeptical. It is my view that the principal challenge facing us all in this area is the design of proper indices; that is, effective design of the data dictionary standards. Unless rather clever means of locating the sources of required data items are devised, any distributed data base access is going to spend most of its time sending messages about the world to which the response will be something like, "sorry, not here". Performance will be destroyed.

I am convinced that use of the semantics of the data will be an essential part of this task and that implies use of the conceptual schema. Thus, in addition to the development of the OSI protocol and service standards, active pursuit of the standardization of a conceptual schema language is imperative. Both SC5 and SC16 have a great deal of work before them.

Another essential aspect of this, derivable from the above observation, is the need for SC5 and SC16 to work closely together. This cooperation has just now begun. Indeed, at the very moment of the presentation of this paper, the two subcommittees are holding their first joint working meeting. If this cooperative spirit continues, the distributed data base problem is likely to be on the road to successful resolution. It will become simply a very difficult problem, not an intractible one.

REFERENCES:

Bartoli, P.D. and Steel, T.B., Jr., "Formal Communication", in preparation [198?].

Belnap, N.D., Jr. and Steel, T.B., Jr., "The Logic of Questions and Answers", Yale University Press (New Haven, 1976).

Jardine, D.A. (ed,), "The ANSI/SPARC DBMS Model", North-Holland Publishing Company (Amsterdam, 1977).

Nijssen, G.M., (ed.), "Modelling in Data Base Management Systems", North-Holland Publishing Company (Amsterdam, 1976).

Senders, J.W., "Information Storage Requirements for the Contents of the World's Libraries", *Science*, 141: 1067-1068.

SPARC, "Interim Report — ANSI/X3/SPARC Study Group on Data Base Management Systems", CBEMA (Washington, DC, 1975). Reprinted in *ACM SIGMOD Bulletin*, 7:2 (1976).

Tsichritzis, D. and Klug, A. (eds.), "The ANSI/X3/SPARC DBMS Framework", AFIPS Press (Montvale, NJ, 1977).

WG3, "Concepts and Terminology for the Conceptual Schema and the Information Base", ISO/TC97/SC5 N695, ANSI (New York, 1982).

Yormark, B., "The ANSI/X3/SPARC/SGDBMS Architecture", in [Jardine 1977], 1-21.