Check for updates

AN INTEGRATED CORPORATE DATA BASE CONCEPT AND ITS APPLICATION

BY: Bernard Plagman

Federal Reserve Bank of New York

Co-AUTHOR: Gene Altshuler

Diebold Europe, S.A.

PREPARED FOR PRESENTATION AT:

SIGFIDET WORKSHOP 1972

Denver, Colorado

ABSTRACT

AN INTEGRATED CORPORATE DATA BASE CONCEPT AND ITS APPLICATION

Authors: Bernard K Plagman and Gene Altshuler

The evolution of computer commercial applications has gone from simple, repetitive accounting to extremely complex information systems. The ability to support the multiplicity of applications and information needs of a business entity with a corporate wide information system has been discussed at length over the last few years. Data redundancy and the inability to transfer data on an intersystem basis within the corporation are among the chief problems to be overcome. The implementation of the Integrated Corporate Data Base (ICDB) in support of a total information system is a potential solution to these problems.

The ICDB concept can be defined as the logical centralization and control of all data necessary to manage the business enterprise. It entails five basic functional elements, which are defined and discussed in detail.

° The Data Bank

° The Data Dictionary/Directory

° The Data Base Management System

° The Data Base Administrator

° The User/System Interface

The ICDB concept is applied to the information system of the Federal Reserve Bank of New York. The information system is outlined and the ICDB concept applied in its enviroment.

I. INTRODUCTION

A. Purpose

There has been much said and written about data bases and their management and while advances are being made at the conceptual, (see Reference 5) technical (see Reference 3) and application (see Reference 10) levels, much of the attention has been focused on the "large applications" as opposed to the "corporate applications".

The purpose of this paper is to introduce the concept of the Integrated Corporate Data Base (ICDB) and relate it to the information system environonment of a specific application. The emphasis of the paper will be on explaining the five elements of the concept and demonstrating how the concept can be utilized in the information system environonment of the Federal Reserve Bank of New York.

B. The Evolution of Information Systems

The foundations of computer based information systems have gradually evolved over a period of approximately fifteen years. Initial computer applications were directed at the repetitive operations of the clerk and bookkeeper. Typically, these early applications were batch-oriented, not related to each other from a systems standpoint and did not share common data files or data input sources. These may be referred to as Accounting or Operational Level Systems. The introduction of second and third generation hardware and software brought a second phase in the evolution of information. It now became possible to collect relevant Accounting Level Systems under a functional umbrella and to relate these applications in some manner or means with the objective of providing management with a slightly broader scope of information about the business operation. Information systems supporting this level of information can be termed Functional Level Systems.

The third phase of information system evolution began approximately five years ago and was prompted by technological advances in computer hardware, communications and data base technology. These advances together with improvements made in systems design concepts, brought closer to reality a company-wide information system. The ability to provide management with information that is necessary for decision making on a corporate wide basis, depends on the ability to integrate the various Functional Level Systems. Data relating to, income, cost, expense and profits which cross functional lines may be termed Corporate Level Information. (see Figure 1.)

It has been this hierarchy of informational needs, the Accounting, Functional and Corporate Levels, that information systems have evolved to support. (The term evolution is emphasized because of the decentralized nature of development.) It is through no fault of the system designers, however, that when the need for Corporate Level Information had arisen, the hierarchy failed. Indeed, some have experienced problems at the Functional Level before attempting to develop Corporate Level Information. The

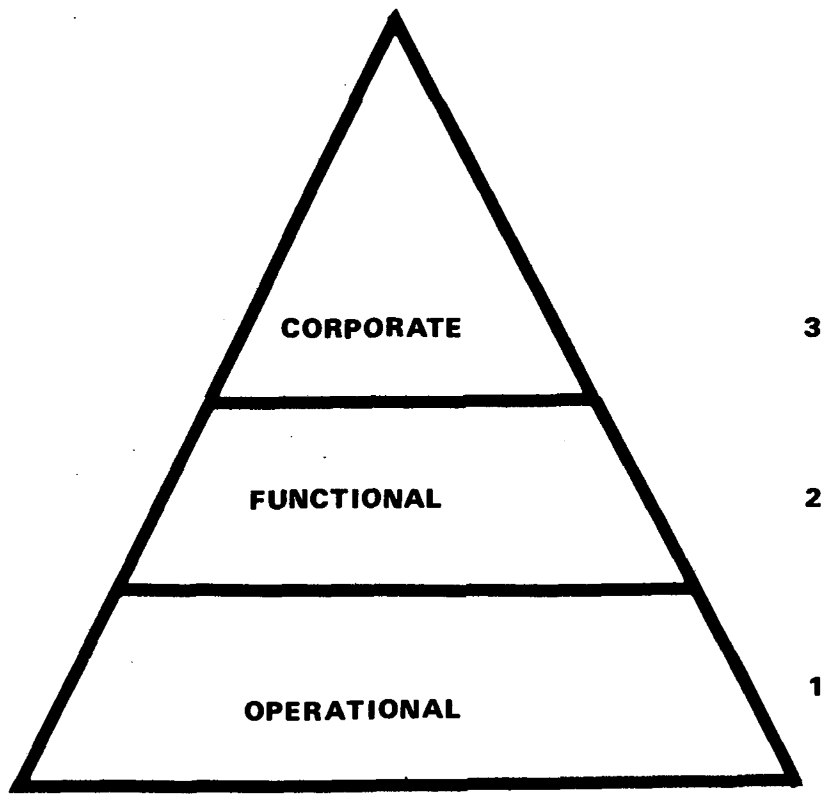

fig.1 LEVELS OF INFORMATION

problems of the implementation of these systems under a decentralized approach are characterized by the following difficulties:

1. Redundant Data

As each functional level system evolved in the corporation, the function created and maintained its own data files. Using an example in the banking industry, if a Deposit system and a Loan system were dealing with the same customer, typically the customer's identification data (e.g. name address, social security number) were stored redundantly in each system. Furthermore, if the bank were supporting a Customer Information File, typically the data would be stored all over again. This is but one type of redundancy: Duplication of data because it is actually needed in two places and the current technological trade-offs dictate redundancy. However, the problem has been compounded in large organizations where communication among staff has failed and inadvertent redundancies have been introduced in data systems.

2. Inconsistency/Incompatibility of Data

As data files were created and maintained by each Functional Level System, large corporations found themselves in a position where each system was dealing with data which was either inconsistent or incompatible with the data of other systems. At one high level meeting of a large computer manufacturer, attendees had brought what were purported to be comparable financial reports from their respective operating divisions, only to realize that the reports were unrelatable. The definition of terms and therefore the data collection in the various systems were so diverse as to render the comparison of information useless.

3. Software Data Dependence

As each Functional Level System was designed and implemented utilizing its own data files, the supporting software became bound to the data which it manipulated. This in turn created high maintenance costs in instances where the dynamic aspects of data manifested themselves causing large reprogramming and recompiling efforts. Witness, for example, the situations in most large financial institutions just a few years back when the Standard Industrial Codes were introduced into data files.

C. The Integrated Corporate Data Base

These problems and others not mentioned have one point in common. They all stem from and are fostered by decentralized control over data. The concept of allowing each application to "own" and control the data which it manipulates is the underlying cause of difficulty. One potential solution to these problems is the implementation of the Integrated Corporate Data Base concept (ICDB). Basically this concept treats data as a corporate resource, just as machines, personnel and money. An ICDB can be formally defined as:

> The consideration of the collection, storage and dissemination of data as a logical, centrally controlled and standardized utility function.

By creating an ICDB with central control over data it would be possible to minimize redundancy and maintain consistent and compatible data. Furthermore, by implementing centralized utility software to handle the storage, retrieval and maintenance of data it should be possible to attain a degree of data independence.

II. THE ELEMENTS OF THE INTEGRATED CORPORATE DATE BASE

It should be emphasized at this point that an ICDB is not a system. It is a concept under which information systems should be implemented. The utilization of this concept implies the necessity for five sub-systems or elements of an ICDB:

- The Data Bank (DB) - the logically centralized respository of all the data utilized in a corporation.
- The Data Dictionary/Directory System (DD/DS) - the repository of all definitive information about the Data Bank such as characteristics, relationships and authorities.
- The Data Base Administrator (DBA) - a Human function responsible for coordination of all data related activities.
- The Data Base Management System (DBMS) - a software function performing the storage, retrieval and maintenance of data.
- The User/System Interface (USI) - the necessary sub-

systems to permit multiple classes and types of users to direct the system to effectively structure the available data into information and thus communicate with and fully utilize the resources at their disposal.

A. The Data Bank

The Data Bank is comprised of a collection of data bases organized to maximize the performance of the total system. These data bases, as implied by using the term "logically", may be physically stored in diverse locations, but are logically linked via communications and the DD/DS. The ICDB concept would therefore be capable of supporting an organizationly decentralized installation as well as a centralized one. Furthermore, data processing installations utilizing a multivendor hardware policy, may also take full advantage of this concept by logically centralizing the data bases of heterogeneous processors into a single Data Bank.

The storage facilities housing the DB should be designed to support a full range of storage and accessing requirements. This can be facilitated by utilizing hierarchical secondary storage. The utilization of different types of secondary storage media offering alternative access techniques and speeds at correspondingly adjusted costs would allow the DBA to design the optimum physical storage configuration for the DB. By providing just four levels of hierachy, (See Figure 2. for a listing of some of the trade-offs) an installation can capitalize on the unique requirements of individual users and save both time and money in the process of storing

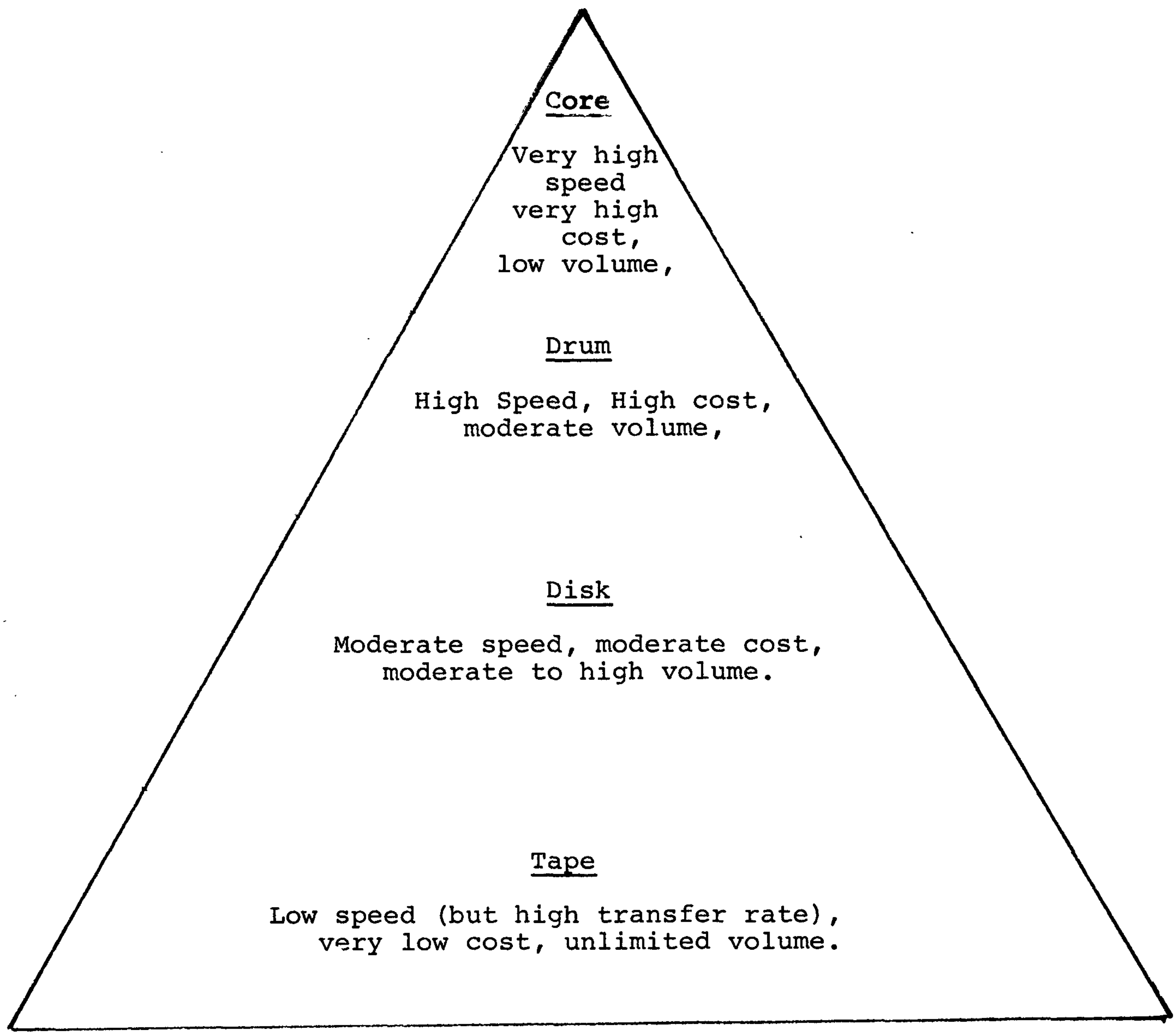

Figure 2. Physical storage hierarchy.

and transferring data.

The efficiency of the total system will depend to a large degree on the specific organization of the DB. There can be only one physical representation of the Data, and the DBA's choice as to exactly which representation to employ is an important decision if not a critical one. The applications' view of the data (sometimes referred to as the logical representation) are equally as important, in as much as the applications modules of the user systems will be designed to utilize these representations. The final section of this paper discussess the basis for a logical representation of the data bases of the user systems of the Federal Reserve Bank of New York.

The critical areas of the DB are those identified as being common to more than one application. Typically, a corporation will have one facit of its business which is central to its operations and which will require a large data base. In turn, various other applications in the corporation will require access to this data base in one form or another (i.e. update, query, reference). It is these data concentrations which provide the focal point for the organization of the DB. The final section of this paper provides an example of this concept in the Federal Reserve Bank environment.

B. The Data Dictionary/Directory System

The design of a DD/DS will generally address two objectives, but in varying degrees of emphasis.

- Collection and Dissemination of Data - This entails the function of supplying the users of data with meta-data, (i.e. date about data) and providing the DBMS with the information it needs to operate.
- Establishment of Standards - This addresses the need of establishing standards for such things as data naming, usage and coding conventions.

The amount of emphasis placed on each of these objectives will have a profound effect on specific aspects of the DD/DS design. If standardization is emphasized then the scope of what is allowable in the system can be narrowed down to standardized elements. If collection and dissemination is the primary objective, then a wider range of possibilities must be provided for, since "what is," must be collected as opposed to "what should be". The degree of emphasis to be placed on either of these objectives will usually be indicated by the individual characteristics of the systems environment. For example, a highly decentralized systems structure would probably require emphasis on collection and dissemination, while a centralized environment with stricter controls would more aptly lean towards emphasis on standardization. Regardless of emphasis, however, both objectives must be addressed and met to the degree necessary.

As an important element of the ICDB the DD/DS is the central source of control over data specification. The DBA will control all data specification using the DD/DS and its facilities. Furthermore, the DD/DS is the central source of control over the flow of data within the ICDB. By supplying meta-data to the DBMS to execute user requests, the DD/DS is the final source of control with respect to DBMS accesses to the DB. (See Reference 1)

The characteristics or attributes of data items stored in the Data Bank make up the contents of the dictionary/directory. In order to allow for maximum flexibility and minimum redundancy in organizing the DD/DS its contents may be divided into two parts, variable and non-variable.

- Non-variable meta-data - This refers to attributes which cannot change from one use of the data element to another. A data element name and description, for example, would usually be non-variable.
- Variable meta-data - This refers to attributes which can change from one use of the data element to another. The representation (binary string as opposed to character string) of a data element would be classified as variable.

By utilizing this dichotomy of meta-data, the DD/DS provides data specifiers with the ability to define similar data elements by specifing the attributes that are similar only once, and then defining the variable characteristics. This referencing is then extended within the DD/DS and minimizes redundancy by storing non-variable attributes only once.

C. The Data Base Administrator

The functional responsibilities of the DBA include technical aspects, user interface and control functions. The following is a partial list of his major responsibilities:

1) Definition of the content and structure of the DB.
2) Control of data access and modification rights to the DB.
3) Advising DB users on efficient techniques for extracting data
4) Establishing data entry, edit and validation standards.
5) Maintenance of the DD/DS.
6) Maintenence of the DBMS.
7) Keeping track of available physical storage.

The primary tool of the DBA in carrying out his functions will be the DD/DS and a special type of data maniputation language using certain features not common to other users.

Data base administration, being a human function, will be affected by oranizational placement. Much debate has been going on with respect to exactly where the DBA should reside organizationally. What has become evident, however is that there is no optimal placement from a generalistic point of view. Each systems environment operates with distinct organizational characteristics and deciding where to place the DBA will invariably involve trade offs.

There are three organizational entities, each with their own unique set of interests which the DBA must contend with. Computer operations is concerned with performance; computer system design is concerned with ease of use and complexity, and the user is concerned with ease of access. Each of these concerned parties have their own interests to consider, thus making the DBA placement in these organizational units a questionable policy. The solution most often offered for this problem is placing the DBA in a staff position, reporting directly to the highest data processing officer in the organization.

D. The Data Base Management System

The DBMS is a software system and does not contain any meta-data. The DBMS must consult the DD/DS for information about data to provide the parameters necessary for the generalized software/hardware to execute user requests.

The DBMS is the element in which the major future technological breakthroughs in data management will take place. Faster, more efficient and different types of software will emerge to facitilate the carrying out of the DBMS functions. Due to this expected change it is important that the implementation of the ICDB maintain a degree of data management system independance. One degree of independence, that of data with respect to the DBMS, is relatively easy to attain by installing an interface mechanism between the DD/DS and the DBMS. A second degree, and more difficult to implement, is program independence with respect the DBMS. In order to accomplish this a standard instruction set for user requests would have to be developed to interface with the DBMS.

The DBMS is the element which makes possible the fullfillment of many of the advantages to be accrued by implementing the ICDB concept. One of the most important aspect of the DBMS must be the high degree of flexibility and range it permits in data base organization.

E. The User/System Interface

The USI is a critical element of the ICDB for without it the system will not serve its users. Without the appropriate interfaces for various levels of users (see Reference 2) the facility will fall into disuse. Among the necessary areas that require USI support are:

1) Language Capability - The utilization of natural language for communicating with the system will enhance the ability of users to formulate requests and problem definitions.In addition a wide variety of special purpose languages will serve to satisfy the requirements of certain users.

2) Inter-active Capability - The ability to browse in search of solutions will support the user by providing the means to develop new alternatives. Through an inter-active capability it should be possible to complement the human thought process and aid management in its decision making activities.

3) Auxiliary Subsystems - The USI will provide subsystems to assist users in massaging system output into the most humanly comprehensible form. This category includes graphic techniques, algorithmetic processes and modeling and/or simulation tools.

F. Interaction of the ICDB Elements

The manner in which the five elements of the ICDB will interact is depicted in Figure 3. The sequenced events trace a typical request by a programmer for data in the Data Bank.

III. THE ICDB CONCEPT ALLIED

The application of the ICDB concept to the functions and operations of the Federal Reserve Bank of New York is comprised of the following two related aspects.

- Information System Design
- Installation of the ICDB elements

The first of the ICDB elements to be implemented is to be the DD/DS (see Reference 1). The emphasis in the initial implementation will be towards collection and dissemination of meta-data. The DBA function has been initiated in the establishment of a Data Base Task Force which will eventually transfer its responsibilities to a formal Data Base Administator. The first assigment of the task force is to develop the requirements for and select a DBMS. The USI requirements are being designed in

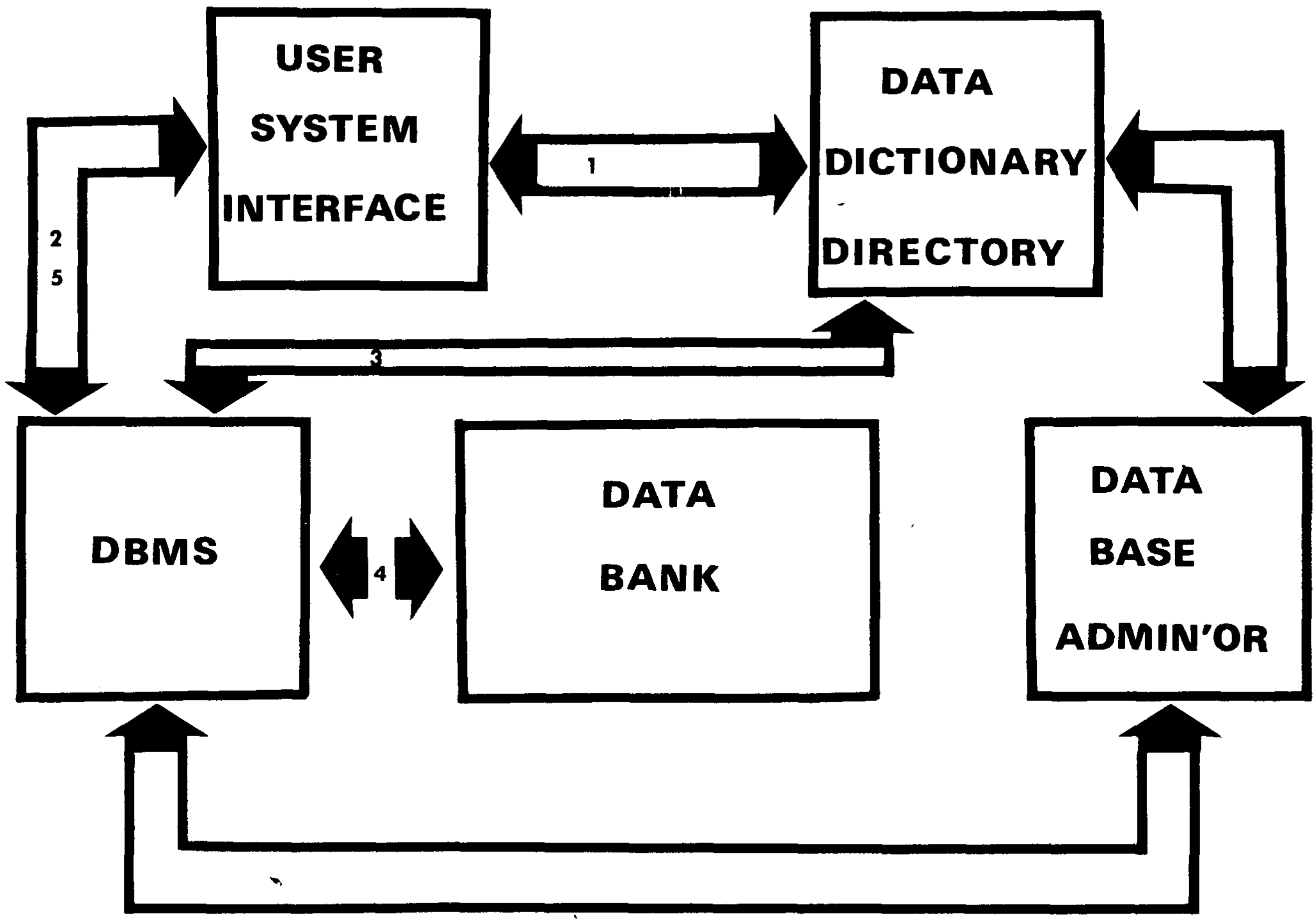

fig. 3 DD/DS driven ICDB

1. Programmer queries DD/DS for information relating to his request for data.
2. Request is coded and transmitted to DBMS.
3. DBMS obtains from the DD/DS meta-data parameters to be used in the physical access.
4. The physical access is performed.
5. Data is transferred to the application program.

conjunction with the information system development.

The remainder of this paper will focus on the information system design as it relates to the organization of the Data Bank.

A. The Federal Reserve Bank of New York

The Federal Reserve Bank of New York is the arm of the central bank in the Second District of Federal Reserve System. Each of the twelve Federal Reserve Banks perform four primary functions, as follows:

1. Acts as a bank for banks - Among the services performed for banks are, processing and clearing, wiring money, maintaining checking accounts. (i.e. reserve accounts), making loans, holding securities and providing information on business and banking.
2. Acts as a Fiscal Agent for the U.S. Treasury Department - the two primary functions performed are keeping the Treasury's checking accounts and helping handle the public debt.
3. Supervising banks - In fullfilling this responsibility the Bank issues operating rules and performs examinantions of state member banks.
4. Managing Money - This function is dedicated to fostering a flow of credit and money that will facilitate orderly economic growth, a stable dollar, long-run balance in our international payments and most importantly maintains a high level of employment. The principle tools to accomplish this are open market

purchases and sales, mainly of U.S. Government securities but also of foreign currencies; lending operations and changes in reserve requirements.

In addition to these four functions the Federal Reserve Bank of New York is unique among Reserve banks in that it manages the portfolio of government securities used to accomplish the goals of monetary policy. The Federal Reserve Bank of New York is also solely responsible for carrying out foreign exchange operations, handling gold transactions, and executing investment transactions for foreign and international accounts.

B. Federal Reserve Bank of New York Information System Structure

The underlying concept upon which the information system structure is based is a combination of communications and data base technology. The combination of a transaction driven system, using remote terminals, with the ICDB concept, allows the applications systems to be designed from the point of view of information usage. The total information system structure represents the functions and operations of the Bank as they relate to the use of information.

There are four information systems which correspond to the Functional Level Information described in the first section of this paper. Each one of these systems is comprised of Accounting Level Systems which feed or are an integral part of their respective Functional Level Systems.

The five information systems and a representive sample of their respective components are outlined in Figure 4. to indicate the nature of the information system structure.

C. The Data Bank

Each of the five information systems indicated utilize data bases designed specifically to support the functional purpose of that system. Integration is attained by identifying the data bases which contain concentrations of data, central to two or more of the information systems and centralizing the organization of that data base with respect to the overall system. The derivation of Corporate Level Information, such as comprehensive profit and loss analysis is insured by providing for the compatilibitily of the data bases comprising the Data Bank.

The largest central data concentration in the Federal Reserve Bank of New York information system structure lies in the data which contains information relative to state member banks. All of the information systems, require access to member bank related data. Thus the Member Bank Data Base is organized in such a manner so as to facilitate access on the basis of member bank ID for the Cash Deposit Information System and by securities held, for the Securities Information System.

The success or failure of the application of the ICDB concept is a direct consequence of the adequecy of the information system design. The ICDB concept is a utility function

I. Cash Deposit Information System
   A. Federal Reserve Bank Deposit System
      1. Second District Deposit Accounting
      2. Check Processing
      3. Funds Transfer
      4. Cash Inventory
      5. Non-Cash Collections
      6. Credit and Discount
   B. Treasury Tax and Loan (TT&L) System
   C. Foreign Accounting System

II. Securities Information System
   A. Fiscal Agency System
      1. Securities Offerings
      2. Transfer Agent
      3. Public Debt Inventory and Reporting
      4. Savings Bond Consignment
   B. Securities and Acceptances Trading System
      1. Securities Trading
      2. Acceptances Trading
   C. Open Market Accounting System (SOMA)
   D. Custody Accounting System

III. Economic Analysis and Reporting Information System
   A. Bank Analysis and Supervision Information System (BASIS)
      1. BASIS Status
      2. BASIS Structure
      3. BASIS Financial Data
      4. BASIS Regulation Information
   B. Research Information System
      1. Balance of Payments
      2. Foreign Research
      3. Domestic Research
      4. Financial Statistics
      5. Market Statistics

IV. Finance and Resource Information System
   A. Financial System
      1. Financial Control
      2. Accounting Control
      3. Budget and Expense
      4. General Ledger
   B. Resource Control System
      1. Personnel
      2. Supplies Inventory

Figure 4. Federal Reserve Bank of New York Information System Structure.

providing the end user of the information system with data collection, storage, and dissemination services. Thus, if the information system design itself leaves something to be desired, the ICDB cannot fill the gap and the total system will fail.

IV. CONCLUSION

The advantages of applying the ICDB concept can be summarized by highlighting two important benefits expected from its application at the Federal Reserve Bank of New York.

- Adaptability to Change - Due to the very nature of the Bank, its information systems are subject to frequent and volitile change. The ICDB concept will provide a degree of insulation against the adverse effects that change has had in the past on automated systems.
- Central Control of Data Resources - While the Federal Reserve Bank of New York is the arm of a central bank its functions and operations are organized on a decentralized basis. The successful implementation of the information system requires centralized control of data resources. The ICDB concept will provide some of the necessary tools, in terms of the DD/DS and the DBA, which will facilitlate this centralized control.

NOTE

The references contained in this paper to the application of the ICDB concept to the Federal Reserve Bank of New York are not to be construed as the accepted policy of the Bank. All such references are, in fact, currently in the form of prosposals.

ACKNOWLEDGEMENTS

A number of people contributed towards clarifying thought and improving the content of this paper. Mr. Gerald Hayden, Mr. Alan Disman and the Public Information Department of the Federal Reserve Bank of New York provided invaluable and timely assistance.

REFERENCES

1. A Data Dictionary/Directory Within the Context of an ICDB, B.K. Plagman and G. Altshuler, Proceedings of FJCC 1972.

2. A User/System Interface Within the Context of an ICDB, G. Altshuler and B.K. Plagman, Conference Proceedings, Meeting XVIII, Diebold Research Program-Frankfurt, Europe, November 14-16, 1972, document No. EC 18.

3. Data Base Task Group Report, CODASYL Systems Committee, April, 1971.

4. Requirements for a Data Base Management System, GUIDE, SHARE, November 11, 1971.

5. Data Base Management System Requirements, Construction Management Systems Action Group, June 23, 1971.

6. Liu, Ho Nien, A File Management System for a Large Corporate Information System Data Bank FJCC Proceedings, 1968.

7. Organizing for Data Base Management, the Diebold Group Inc Doc. No. 516, December 1971.

8. Ackoff, Russel L. Management Misinformation Systems, Management Science, December, 1967.

9. The Large Data Base, Its Organization and User Interface, Olle, William, Chairman. Data Base, Fall, 1969.

10. Moehrke, D.P. Evolution of Data Base Management at A.O. Smith Proceedings of SIGFIDET Workshop, November, 1971.