TITLE: "WHY FOUR DATA BASE MODELS?"

by Tim Bryce
Managing Director
M. Bryce & Associates (MBA)
P.O. Box 1637
Palm Harbor, FL 34682-1637
United States
Tel: 727/786-4567
E-Mail: timb001@attglobal.net
WWW: http://www.phmainstreet.com/mba/
Since 1971: *"Software for the finest computer - the Mind"*

"A Data Base should naturally evolve over time and synchronize with all Information Systems."
- Bryce's Law

INTRODUCTION

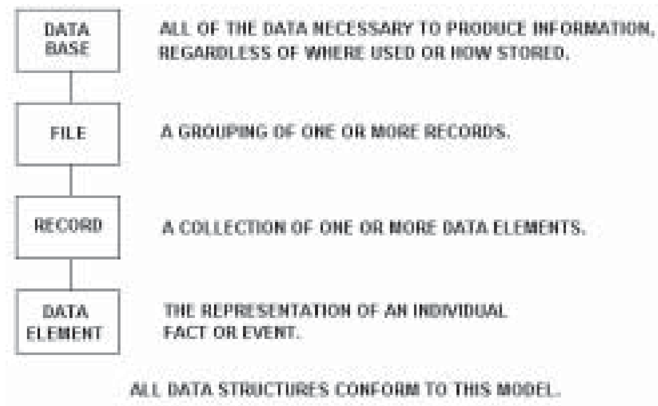
Managing the corporate data base has been an inherent part of "PRIDE" since it was first introduced in the early 1970's. However, by the early 1980's supplemental data base design techniques were the rage, such as Ted Codd's "normalization" and Entity/Relationship Diagrams. Such techniques, when applied with a little common sense, were helpful for developing data models representing the facts and events of the business. Some consultants were so enamored by these techniques they advised companies to suspend all development activity and build a global corporate data base model (I'm not going to mention any names, but you know who you are). Of course, this is not practical to do for Fortune 1000 companies (as many quickly discovered).

Our contribution to this area came in the form of the "PRIDE"-Data Base Engineering Methodology (DBEM) in 1987 which introduced the concepts of "data taxonomy," "objects" and four data base models. As to the latter, one might ask, "Why four data base models?" Quite simply, it is a convenient "divide and conquer" strategy for developing the corporate data base model without having to suspend development activities. Frankly, it is a simple and natural approach for development.

DATA BASE CONSTRUCTS

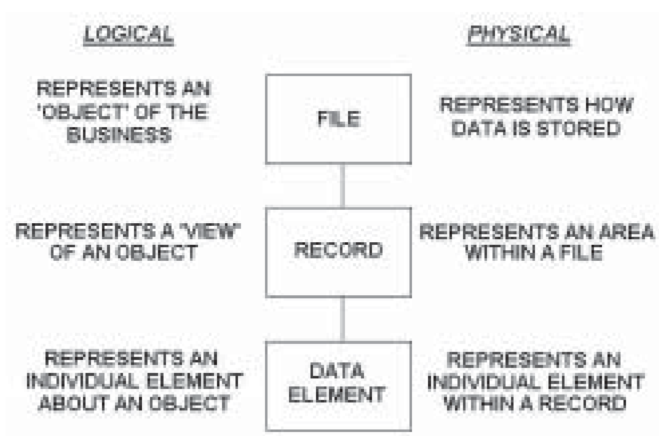
Before we can answer the question, *"Why Four Data Base Models?"* let's first review some basic concepts as used in "PRIDE"-DBEM.

Data resources can be organized into a generic and universal structure. The basic building block is the data element itself, the representation of an individual fact or an event. A collection of one or more data elements is a record and one or more records make up a file. And a data base represents all of the data used to produce information, regardless of where used or how stored.



All data resources are structured in this generic manner. Terms such as "schema," "sub-schema," "segments," "tracks," "cylinders," "sectors," "tables," "arrays," "tuples," "data stores," etc., (all of which deal with particular computer techniques and tools), can all be translated into the basic constructs mentioned above.

The organization of data serves two purposes: one is to logically describe the "objects" used to manage and operate the business, and; to express how data will be physically stored. The differences between logical and physical are substantial; there will not necessarily be a direct relationship between the two.



There is not necessarily a one-to-one relationship between logical and physical

Physical files may differ considerably from logical files. Here, the file represents a particular way of physically storing data. Data may be physically stored in a variety

(continued of page 2)

(continued from page 1)

of files, such as an indexed file, a "flat" file, a DBMS file, etc. Even manual files follow this model with the exception they also store inputs and/or outputs (both of which consist of records and data elements).

Unlike the logical file that is organized according to a unique data identifier ("primary basic grouping"), the physical file does not require any specific organization and can use any sort/access key desired. Ultimately, it depends on the file management technique or tool being used.

The logical view of data is the basis for all physical data base design, regardless of the file management technique or tool selected. The physical files must ultimately carry out the intentions of the logical files in terms of what data must be stored, the dependencies between data, and volume. As a matter of fact, all DBMS packages can implement these logical views, regardless of whether they have a hierarchical, network, relational, or object-oriented structure.

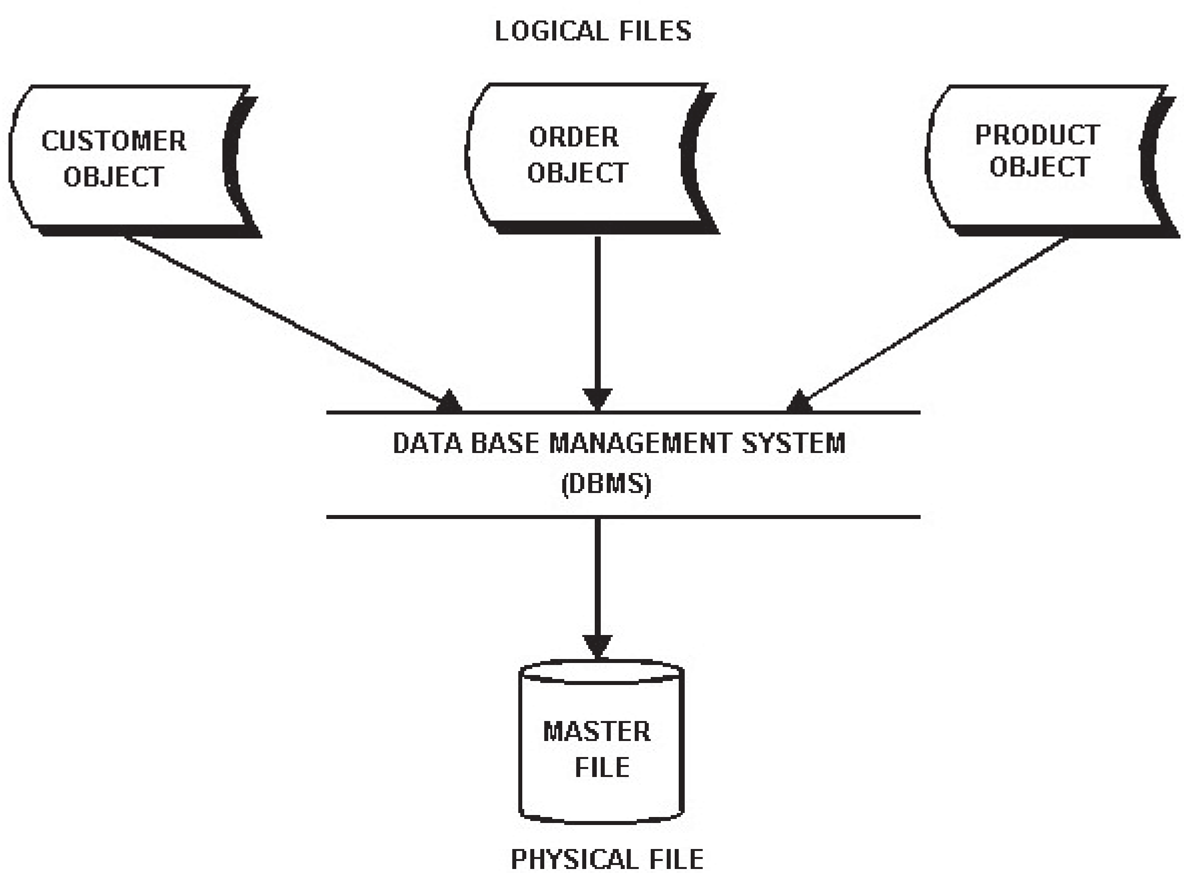
Again, there are substantial differences between logical and physical files. Perhaps the most noticeable difference is the logical file will remain relatively static while the physical file will change dynamically, based on advances in technology. One of the most important reasons for defining data resources logically is to seek data independence from the physical environment, thus allowing any physical implementation without disrupting systems.

APPLICATION VERSUS ENTERPRISE

It would be easy to say there are just two types of data base models, logical and physical. However, there is another perspective that adds another dimension to this, and that is how data is viewed from an "enterprise" versus an "application" perspective.

An "application" view refers to the data used in a specific system. In terms of the logical model, it represents the "local" data used to describe objects for a particular Information System. It represents only those data elements required to satisfy the information needs for a particular application. Obviously, this will not necessarily be the "global" view of the object, which is the intent of the "enterprise" view. In other words, the "application" view will usually be a subset of the "enterprise" view of data.

(continued on page 3)



(continued from page 2)

"APPLICATION VERSUS ENTERPRISE"

APPLICATION VIEW OF A CUSTOMER	ENTERPRISE VIEW OF A CUSTOMER
Customer Number	Customer Number
Name	Name
Credit Rating	Credit Rating
	Contact Number
	Title
	Telephone
	Address Code
	Address
	City
	State/Province
	Zip/Postal Code

The "enterprise" view represents a complete picture of the object, with all of the data required to satisfy all applications, not just one. Under this arrangement, there may be multiple "application" views of objects, but only one "enterprise" view of an object. In fact, it is quite common to have many different "application" views of an object. One system may require certain data elements about a customer object while another requires a totally different set of data elements to describe a customer. These legitimately separate views of the customer, as defined by Systems Engineering during design, are coordinated through the enterprise view of the customer as controlled by the Data Engineering function.

When a system is designed into sub-systems with logical files, the "enterprise" data base is adjusted to accommodate the "application" data base. If the objects encountered in the system are new to the enterprise, then new enterprise views must be defined. Initially, the application and enterprise views of an object are identical. As new applications are introduced with different views of the same object, then the enterprise view is modified accordingly by Data Engineering.

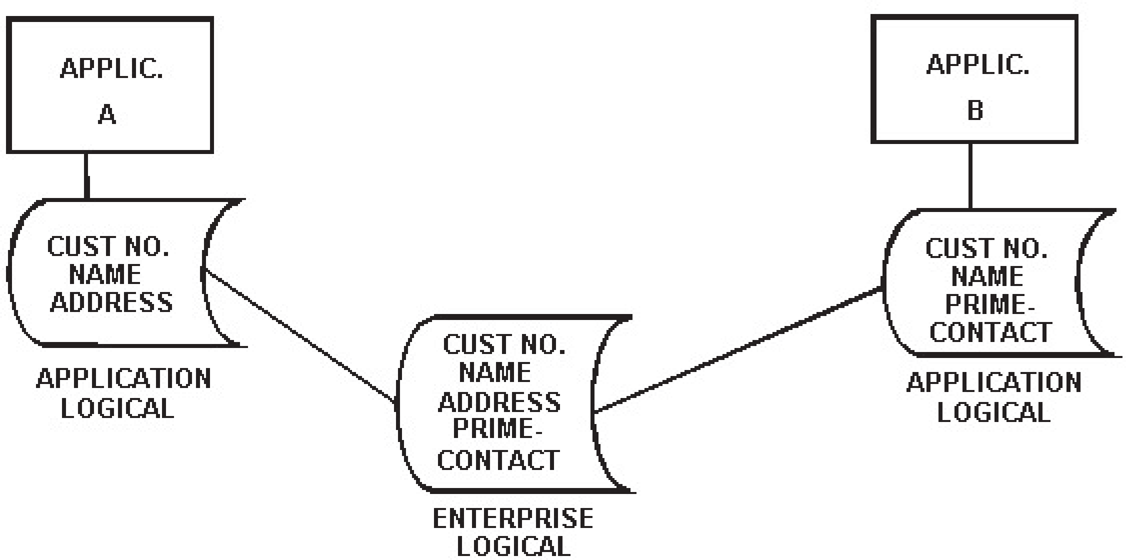
THE FOUR DATA BASE MODELS

The variables of logical versus physical and application versus enterprise results in four data base models:

- The **Application Logical Data Base Model (ALDBM)** represents all of the primary data elements needed to satisfy the information requirements of a single application. In other words, all of the data needed to describe the objects pertinent to a given information system. The ALDBM defines the logical files used in a single system. It also represents a subset of the Enterprise Logical Data Base Model.

- The **Enterprise Logical Data Base Model (ELDBM)** represents the primary data elements used to describe all objects in an enterprise, not just that data used in a single system. It represents all logical files in the corporate data base.

(continued on page 4)



(continued from page 3)

• The **Enterprise Physical Data Base Model (EPDBM)** represents how the data in the ELDBM is physically stored in files. The corporate data base can be either centralized or distributed. A variety of file management techniques can be used to store the data, e.g., computer files, manual files, etc. The EPDBM, therefore, defines all of the physical files in the corporate data base.

• The **Application Physical Data Base Model (APDBM)** represents subsets of the EPDBM used to fulfill a specific application. It satisfies the data requirements of the ALDBM and denotes the physical files used in the system.

These four models represents the rationale for Phases 2 - 5 in "PRIDE"-DBEM.

Phase 2 - Application Logical Data Base Design - define the ALDBM for a given system.

Phase 3 - Enterprise Logical Data Base Design - merge the ALDBM into the ELDBM.

Phase 4 - Enterprise Physical Data Base Design - implement the ELDBM into the EPDBM.

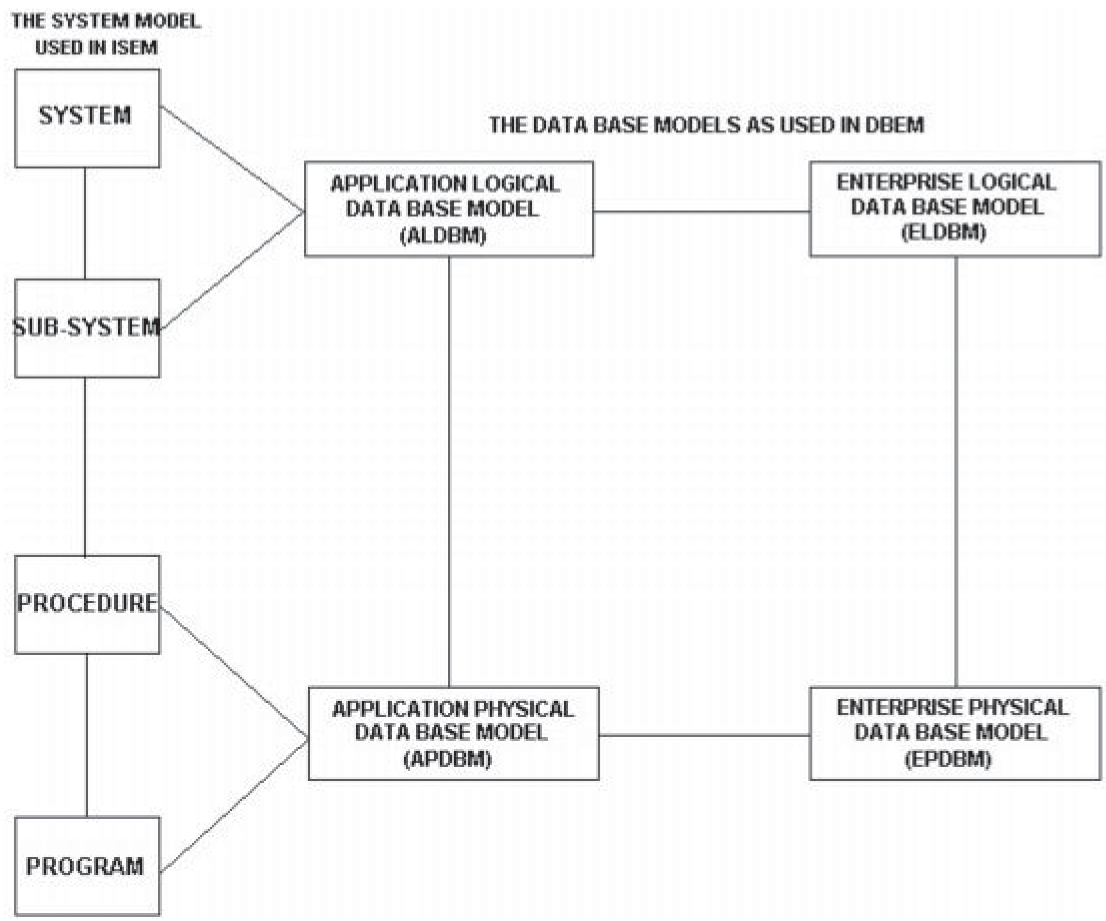
Phase 5 - Application Physical Data Base Design - create a subset of the EPDBM as used by a given system.

CONCLUSION

This application/enterprise relationship highlights the fact that data base design is an evolutionary process. Other data base design techniques typically take a "revolutionary" approach by trying to identify all of the data requirements for the entire company at one time. Obviously, the problem with this approach is that it becomes an enormous and unmanageable data base design project with questionable results. Whereas the evolutionary approach naturally synchronizes the data base with all of the applications, the revolutionary approach develops a data base that will not necessarily match the applications.

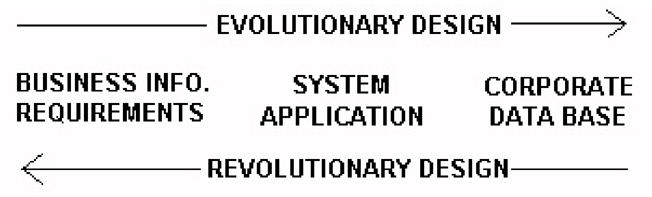
Under the evolutionary approach, the corporate data base will expand and contract naturally as the business and

(continued on page 5)



(continued from page 4)

applications change. Consequently, excessive or unnecessary data definitions will be avoided.



For more information on the four data base models, see:

http://www.phmainstreet.com/mba/pride/dbmeth.htm#4dbmodels

END

"PRIDE" Special Subject Bulletins can be found at:

http://www.phmainstreet.com/mba/mbass.htm

They are also available through the "PRIDE Methodologies for IRM Discussion Group" at:

http://groups.yahoo.com/group/mbapride/

You are welcome to join this group if you are so inclined.

The "Management Visions" Internet audio broadcast is available at:

http://www.phmainstreet.com/mba/mv.htm

"PRIDE" is the registered trademark of M. Bryce & Associates (MBA) and can be found on the Internet at:

http://www.phmainstreet.com/mba/pride/