The Databases of Our Lives

by Craig S. Mullins

When last we looked in on our favorite database, it was in the throes of a new release. Old man Plan was being packaged and versioned by the Collection agency for failure to perform. Mr. and Mrs. Join (Merge Scan and Nested Loop) had an offspring - little Hybrid Join. And all sorts of other interesting things were occurring . . . but what of the future?

Well, just like a soap opera, some of us in this business just love gossiping about the next release of DB2. Some of us can never get enough of DB2. Soooo, let me introduce you to the DB2 of the future! Well, of course I can't really see into the future, but I can bring you information "straight from the horse's mouth." IBM has been giving presentations at various user groups around the country discussing basic trends and directions for DB2's future development. In this article, I will consolidate information presented by IBM at four major industry conferences: the Midwest DB/DC User Conference (September 1992), the IBM DB2 Technical Conference (October 1992), Database World (December 1992), and GUIDE (March 1993).

Three DB2 Development Cycles

According to IBM, there have been two cycles for DB2 development so far, and a third cycle is just beginning. Each of these cycles has corresponded to a different version of DB2. Cycle one, lasting from 1984 to 1987, was for Version 1 of DB2. DB2 V1.1 was widely regarded as only an information center tool. In making modifications to DB2 during cycle one, IBM focused on introducing relational technology and making an RDBMS work in a production environment. During the second cycle, corresponding to DB2 Version 2, IBM concentrated mainly upon performance and on the advent of client/server and distributed database technology. This cycle lasted from 1988 through 1992.

In late 1992 the third cycle began. This development effort is multi-faceted, with the major emphasis being on performance, availability, distributed data, and functionality. The current strategy for this cycle concentrates development effort roughly as shown below:

Cycle 3 DB2 Development Effort	
Objective	% of Effort
Performance	25%
Availability	25%
Everything Else	50%

The "Everything Else" category includes data distribution, model extensions, system integration, and other miscellaneous objectives. Let's examine each of these items in turn.

Performance Objectives

Enhanced DB2 performance will continue to be an on-going objective for future releases of DB2. IBM's performance strategy will include the development of more advanced optimizer techniques, increased parallelism, ESA and S/390 exploitation, as well as new and refined algorithms throughout the DB2 code.

The major news within the world of DB2 performance is that the focus has changed. Future efforts to enhance performance will favor reduced response time over reduced CPU usage. In the past there were indeed cases where CPU costs rose but elapsed time decreased. In the future, this will most likely be the rule rather than the exception.

Historically, however, DB2 performance has benefited more from hardware enhancements than from software enhancements. In 1984, DB2 could process 30 transactions per second (tps), but in 1992 this rate was over 600 tps, a 20-fold increase in tps in eight years. The majority of the increase, however, was a result of hardware improvements. This trend should continue for some time to come.

Look for these future hardware enhancements to benefit DB2 performance:

- Utilizing multiple processors with shared memory and shared DASD. This hardware enhancement is projected to provide 1.5 BIPS of processing power by the middle 1990's.

- Further migration of specific software to hardware. Look for data compression to be available in a specialized chip very shortly. The goal is to support data compression much like DB2 V2.3 provides a hardware-assisted sort.

- Support for RAID technology. RAID stands for Redundant Array of Inexpensive Disk. RAID enables many disk drives to be used in tandem to support parallel data access.

Availability Objectives

In cycle three, improvements to the availability of DB2 data will be equal in importance to performance improvements. IBM's first cut at increased availability was package support in DB2 V2.3. When packages are utilized, less downtime is required to prepare applications for execution.

IBM's first item of business for improving availability is providing increased parallelism. Look for increased parallelism in the near future of DB2 through support for independent partitions. IBM's goal is to enable application logic, utilities, and DB2 commands to operate independently on each partition of a partitioned tablespace. Making each partition independent of the others provides the following benefits:

- One partition can be stopped without affecting the whole tablespace.

- Parallel utility jobs can be run concurrently, reducing application down time.

- Even non-partitioning indexes will eventually be independent by partition.

IBM has stated that DB2 will provide enhanced availability in the following three phases:

- Providing parallel I/O
- Providing a parallel processing across multiple CPUs
- Enabling multiple Central Electronic Complexes (CECs). Many small CPU clusters promote more parallelism in processing: access, joins, subselects, unions, distributed requests, etc.

Future availability goals include improved concurrency, reduced index locking, row-level locking support, and read-through locks.

Distribution Objectives

Enhanced DB2 support for distributed data will continue to be of

> "Enhanced DB2 support for distributed data will continue to be of paramount importance during cycle three."

paramount importance during cycle three. Of course, IBM will stress DRDA support. In terms of functionality, look for two-phase commit across sites, distributed request support, referential integrity across sites, and multi-site join capability. Another future goal is to provide multi-site access and update across multiple database managers (both IBM and non-IBM). However, IBM hinted that this would probably be handled by a gateway product, rather than being built internally in the DBMS.

Another of IBM's current distributed directions is embodied within their DataHub framework. DataHub provides distributed DBA functionality from a workstation with a graphical user interface (GUI). It supports

The Databases of Our Lives

IBM and ISV tools to provide a front-end to all relational database environments. DataHub will enable a DBA to control utilities, authorization, copying objects and data, and other administrative efforts, all from a single workstation.

Object-Orientation Objectives

One of the hottest topics within the DP industry these days is object-orientation (OO). The trend is to incorporate object-oriented technology into our business databases, and IBM plans to do just that with DB2! A division within IBM, the Database Technology Institute, has been working on a new database model which they have coined *Object Relational (OR)*.

The Object Relational model provides a marriage of the best relational and object-oriented concepts. In the long run, DB2 will be enhanced to incorporate object-oriented features. So, rather than introducing a new ODBMS, for example, a DB3, IBM will be promoting the extension of DB2 to the Object Relational model, making, for example, DB2++.

In an object-oriented paradigm, the data and the processes that act upon that data are stored together. This concept is called *encapsulation*. (See the illustration at the right.) IBM has indicated that they intend to support object orientation by incorporating a rules manager into the database. The level of abstraction will be also increased by hiding the software within the data.

Because data and processes are coupled, code can be easily reused and the cost of maintenance reduced. However, the purpose of this article is not to teach object-oriented technology, so let's stop before we get side-tracked. For those of you unfamiliar with object-oriented technology, refer to the sidebar of reference books at the end of this article.

Migration To An Object-Oriented Paradigm

The first OO features that will probably be seen in DB2 are abstract data types and user-defined functions. Abstract data types will enable different types of data to be stored in DB2 tables. Columns will not have to be limited to the small scope of data types currently available within DB2. The DBA will be able to define his own data types. And, user-defined functions and/or triggers will enable changes in the state of data to execute software further affecting data. This enables IBM to achieve a limited amount of encapsulation and information hiding.

By merging object-oriented techniques into DB2, IBM can support the best of both the OO and relational worlds. It remains to be seen whether this is completely possible. IBM did not indicate whether OO support will be included in the third cycle or a future cycle. It is likely, however, that limited OO enhancements will be seen during cycle three.

TODAY

Logic

Rules

Data

FUTURE

Logic

Rules Manager

Data
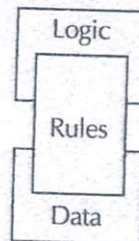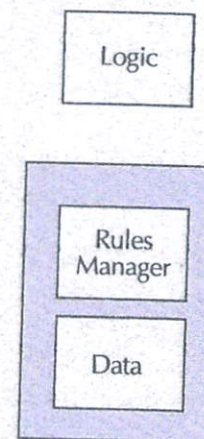
Other Objectives

IBM has announced numerous future directions for DB2. I have grouped these into the category of Other Objectives because they either didn't fit into any of the above categories or they fit into more than one of them. What follows is a list of items that IBM indicated will be in the future of DB2. When a time frame was given by IBM, I have included that as well.

Optical Disk Support. Although optical disk is slower than DASD, it is also much faster than tape and much cheaper than DASD. This makes it ideal for large historical databases.

Domain Support. Domains provide a means of grouping data elements into logical families of data. The domain concept has been around as long as the relational model (since 1970) but few, if any, relational database implementations support domains. This is a near-term goal (1994-1996 time frame).

Support for ANSI Standards. Primarily supporting the SQL2 standard.

SQL Extensions. IBM plans to enable updates for cursors that employ joins and ORDER BY; provide the fabled Outer Join; and offer support for UNION in views. This is also a near-term goal, but no dates were mentioned by IBM.

Multi-Table Indexes. Instead of using multiple indexes to accomplish a join, a multi-table index is based upon a common key and supports RIDs from more than one table. All of this is housed within a single index. This will reduce I/O and enhance join performance.

External Security Exit. To enable RACF to control DB2 security, IBM intends to provide a standard DB2 exit point to interface with other security packages.

Multi-Media Support. Multimedia capability enables a database to store pictures, music, sound, etc. This is usually accomplished via Binary Large Objects, or BLOBs. BLOBs will probably not be stored in the database, though. The DBMS will not want to log BLOBs because of their large size. They will probably be supported by pointers within the DBMS pointing to other types of storage.

Temporary Tables. A temporary table provides an application with the ability to create a table that exists only during a program's execution. This could be used to enable SQL access to code tables in a program's Working-Storage, to involve a result set in relational operations, for example.

The Next Release

But, the bottom-line is: what can I expect to see in the next release of DB2? And when will I see that release? Well, DB2 V3.1 should be available 18 to 24 months after the last release of DB2. For those interested in details, DB2 V2.3 was released for general availability in March 1992. IBM has just announced DB2 V3.1.

Look for the following functionality in the next release:

- Enhanced support for DB2 ob-

"By merging object-oriented technology into DB2, IBM can support both the object-oriented and the relational worlds."

---

ject statistics, including separate statistics tables, timestamping of RUNSTATS values, and the removal of the non-uniform distribution statistics from SYSFIELDS

- Enhanced bufferpool support (more bufferpools, more flexible modification, and bufferpools in hiperspace)

- Increased, but not full, partition independence

- Package information recorded in all the accounting trace records

In closing, as we once again look in on our favorite database, the evil Doctor Da (DRDA) is infiltrating our lovely Miss DB2. Slyly operating from multiple sites, will his two-phased plan to conquer her be suc-

The Databases of Our Lives

New Product Developments: PLATINUM Expands Its Line of QMF Tools

by Michelle Edelman

With the release of the P303 tape, PLATINUM announces the production version of PLATINUM Object Administrator for QMF™. This product is the latest addition to PLATINUM's comprehensive tool kit for managing the QMF environment.

All the QMF management products are fully integrated with each other and with PLATINUM's other tools for DB2. And, they will reflect PLATINUM's unique development expertise and match our DB2 products in design elegance.

PLATINUM Object Administrator provides an easy-to-use online interface that simplifies and streamlines tiresome QMF object management tasks. By working from a selection list or specifying global selection criteria, you can delete, copy, or move single or multiple objects, increasing your level of control over the QMF environment. QMF object management is key to effective QMF usage, as it reduces the amount of DASD used and improves performance.

PLATINUM Object Administrator joins PLATINUM Compile/QQF and PLATINUM Governor Facility as part of a total solution for managing objects and resources in the QMF environment.

PLATINUM's next QMF product will be PLATINUM Object Tracker, which will enable you to monitor the usage of QMF objects so that you can detect resource intensive queries and develop alternative methods of controlling the query, the end-user, or the DB2 resources involved. Deleting superfluous objects releases critical DASD space and improves QMF performance by reducing the number of QMF object table rows that must be searched.

PLATINUM will also introduce a Query Analyzer Facility to determine exactly how much a query will cost in terms of I/O and CPU time. You may then easily cancel resource-intensive queries before execution begins.

We are committed to providing the most complete set of QMF management tools on the market. For all the details on these products, contact a PLATINUM sales representative. 凰

cessful? Well, we'll just have to keep tuning in . . . as the database turns!

Caveat: All of this information was obtained from IBM's stated trends and directions at industry conferences. Some of these anticipated directions could change in the next few months. 凰

Further Reading on Object-Oriented Technology

- Cantrell, R.G.G., *Object Data Management*, Addison-Wesley Publishing Company, Inc., Reading, MA, 1991.

- Kim, W., *Introduction to Object-Oriented Databases*, MIT Press, Cambridge, MA, 1990.

- Taylor, D.A., *Object-Oriented Technology: A Manager's Guide*, Addison-Wesley Publishing Company, Inc., Reading, MA, 1990.

- Zdonik, S.B., Maier, D., Eds., *Readings in Object-Oriented Database Systems*, Morgan-Kaufman, San Mateo, CA, 1989.