DB&4GL
5 QUESTIONS & ANSWERS

by Dr. George Schussel

Based on information from independent consultants and software vendors, some key questions and answers have emerged. Here are five important points about DB and 4GL software productivity.

QUESTION 1

Using a modern relational DBMS and 4GL is supposed to improve applications development productivity. For most companies will this have an impact on decisions to purchase application software?

ANSWER

The applications software and productivity tools software businesses are converging. This has been caused by several factors:

□ Most successful companies are run by businessmen, who are interested in answers to problems, not improved software technology per se. If an application package can be purchased and customized to the customer's requirements, then in many cases that would be the best business solution. If such a package, for example financial, can be integrated with the manufacturing and personnel applications then so much the better. The creation of integrated applications/developmental tool software packages is the driving principle behind the strategies of companies like Cincom and Cullinet.

□ The major application houses, such as McCormack & Dodge, MSA, Software International, and Walker Interactive, came to the same conclusion but by different reasoning. Through the 1970's and early 1980's many of these companies had built sets of stand-alone applications based upon older file system logic. The conversion of these older systems to more modern, flexible, integrated, and on-line capabilities proved to be very difficult using COBOL, Assembler, and third-generation indexed file systems. These application houses, as a result, proceeded to redevelop their software using internally developed modern tools with integrated database capabilities and 4th generation languages. Having developed these new tools, in many cases these software vendors have decided to offer the new development tools to their customer base.

So for different reasons, all of the major software vendors have arrived at the conclusion that integrated applications and productivity software represent the best complement of software products to be marketed.

Dr. Schussel is President and Principal Consultant of Digital Consulting Associates, Andover, MA. He is chairman of the National Database and 4th Generation Language Symposium, and is a fellow of the American Assoc. for the Advancement of Science.

QUESTION 2

Some vendors recommend one DBMS/4GL for operational systems and another different DBMS/4GL for information centre/DSS applications, while other vendors say that a single system is the best approach. Your opinion?

DB&4GL

ANSWER

This is an issue similar to the one vs. two database argument. Only in the answer to this question one has to take into consideration not only the potential difference between production and information centre DBMS', but also the difference between 4th generation languages designed for end users (FOCUS, RAMIS, NOMAD, etc.) and 4GL's for teleprocessing applications such as ADS/O, MANTIS, IDEAL.

The software offerings of the large hardware vendors, including IBM, Sperry, DEC, and Data General, support the two database approach, e.g. IBM's IMS/VS is its production centre database with DB2 for the information centre and DXT for integrating the two. DEC's offerings are DBMS-32 for production applications and Rdb for the information centre with the VAX Information Architecture providing a common integrating facility.

On the other side are the leading independent software vendors such as Cullinet with IDMS/R, Cincom with TIS, ADR with DATACOM DB, etc.

These vendors have added 4th-generation language facilities and relational facilities to DBMS architectures that were designed for production centre and transaction processing environments. This approach supposes that most customers prefer to buy one integrated set of software development tools with a common database and software environment for all (or most) applications.

The dual database vendors argue that fundamentally changing an older network DBMS so that it is friendly enough for information centre users is a difficult if not impossible process. The hardware vendors seem to be saying that it is time for a new technology in user friendly environments and that "Truly" relational systems are the way to go for these "Information Centre" scenarios.

I think that there is some truth in both approaches. For the very largest applications and information centers, it's unlikely that one set of tools will satisfactorily perform for all of the applications. In this case, two or more DBMS/4GL combinations are required.

For smaller companies and/or in departmental computing situations we advise customers to look at relational DBMS with an appropriate 4GL as a single software development environment which can handle both production and information centre requirements.

QUESTION 3

Some vendors have described their DBMS as relational when they're hierarchical or network with some limited tabular extensions. Do these "Born Again" systems deserve to be called relational?

ANSWER

This issue has emerged as a hot debate topic. There is sharp disagreement between many software vendors and relational theoreticians on this point. For the purposes of this discussion let's define "Truly" relational systems as products which have evolved as attempted implementations of the relational theory first espoused by IBM's Dr. Edgar Codd and that did not exist before or without that theory. On the other hand, we will use the term " 'Born Again' " relational implementations to describe such products as Cullinet's IDMS/R, ADR's DATACOM DB and Cincom's TIS, which were built around DBMS models other than relational and which have been extended to encompass (some) relational theory.

Furthermore we need to understand what a relational database is:

□ A structural component consisting of groups of tables with named columns and unordered rows.

□ A data manipulation language consisting of set oriented operators (relational algebra) supporting SELECT, PROJECT, and JOIN functionality.

□ An integrity component for maintaining the consistency of data both within and between tables consisting of entity integrity and referential integrity.

Beyond considering the usual services expected from a DBMS when considering a Relational DBMS the buyer should be especially cognizant of the following points:

□ Reasonable performance from a relational database is going to require sophisticated statistically-based optimization techniques. If your vendor provides a lower level (than relational) access mode and suggests that approach for performance oriented applications then it is likely that that system's relational processing mode will never require the sophistication that is necessary (and achievable) for performance processing in relational mode.

This issue becomes particularly important when the database is distributed over a number of geographically distributed computers rather than being located in one facility.

□ Be careful if your relational DBMS provides both a combination of relational "set at a time" operators and lower level record-at-a-time access. In such systems these lower level access methods can subvert or override the higher level relational rules of constraints and impact the integrity of the database.

□ Check the content of the relational data manipulation language. Some are less full than others. While operators for the full relational algebra (union, intersection, etc.) can be constructed from the 3 basic operators (PROJECT, SELECT, JOIN), the relational DML will be more useful with a full complement of functionality.

The vendors of the "Born Again" relational systems argue that many of the points above are of academic interest only. In addition to supporting relational-like tabular views and verbs, their systems also allow lower level record-oriented processing. Especially in transaction processing environments which are by nature record oriented, the resulting performance potential means that these "Born Again" products can be used for both information centre and production centre applications. The proponents of these views argue that if a user acquires a "Truly" relational system such as IBM's DB2 or DEC's Rdb, then the vendor will sell two different database systems, a hierarchical or network system (IMS for IBM and DBMS-32 for DEC) and a relational system for infor-

Continued on page 55

DB&4GL

mation centre applications.

This issue is frequently enjoined in debates at the *National Database and 4th Generation Language Symposium.* In those debates the theoreticians, who generally support the approach of "pure" relational win. However, in the real world, we have spoken with many experienced customers of "born again" implementations who are convinced that "pure" relational systems can't solve their current real world needs while the "born again" implementations are truly the best of both worlds.

QUESTION 4

Are the differences between inverted list DBMSs such as System 1032, Model 204 and ADABAS and truly relational DBMS such as Oracle, INGRES, and DB2 important?

ANSWER

The technology in modern inverted list DBMS (SAS Institute's System 2000, Software AG's ADABAS, Computer Corp. of America's Model 204) was pioneered in the 1960s and 1970s. From a programmer point of view such a data model consists of looking at sets of two dimensional tables which are indexed on specified fields. The user views both the ordered database tables and the indexes. Much of the processing speed advantage of an indexed DBMS results from the fact that many queries can be answered simply by processing in the indexes without going to the DBMS. A few years ago, these systems owned the "user-friendly" segment of the DBMS market. Index systems were considered highly efficient for multiple key query and interactive searching, and not so efficient for high volume transaction processing, but definitely friendlier and easier to implement than CODASYL or hierarchical database structures.

In the 1980s relational databases came along which while also offering user views of two dimensional tables did away with visibility of the indexes. Relational DBMSs also offered high level set oriented retrieval operators as part of their basic data manipulation language. A single relational operator such as SELECT will produce a set of records where in an inverted DBMS the basic data manipulation language is record oriented. The handling of set oriented questions in an indexed system is through a separate query language.

The advent of relational DBMS with their enhanced data manipulation languages took the high ground from the inverted systems in the user friendly wars.

When we look at the products that are available in today's marketplace, we make the following conclusions:

- For most users there are more similarities than dissimilarities between using inverted and relational DBMSs.
- Implementations of inverted DBMSs are more mature, may perform better than today's relational systems, and because of their maturity are more likely to be trustworthy in serious strategic applications.
- The simplicity of the powerful set oriented data manipulation language of a relational DBMS offers the potential for more programmer productivity than does the typical indexed DBMS.
- Conclusion: The widespread implemtation of relational DBMS means that database design and implementation techniques for this model are evolving as standards.

QUESTION 5

Within the next five years, will artificial intelligence find its way into 4th generation languages?

ANSWER

Artificial Intelligence (AI) is one of the 5th Generation Technologies. By 1990 these 5th generation methods will pervade modern software development tools. In some products these techniques will be used to enhance existing 4th (and 3rd) generation tools by making them smarter and friendlier. In other products, we will see truly new 5th generation approaches. These new 5th generation approaches will use software to write applications software directly from input specifications.

In the first category there currently are a number of artificial intelligence languages: ENGLISH by Mathematica, THEMIS by Frey Associates and INTELLECT by Artificial Intelligence, are starting to generate some interest from DBMS users. These AI tools now provide interfaces to a number of popular DBMSs, including RAMIS II, DB2, IMS and VSAM.

One problem with the current crop of AI query languages is that it is not clear that the use of such languages offers end users a significant savings in time and effort over the use of more contemporary 4GL techniques. There's a fair amount of work to generate an AI language environment before it can be used. 4GL's require more training on the part of the end user, but have the potential for being implemented more quickly than the current AI languages.

One necessary and likely development from AI is the development of more friendly natural human interfaces for existing products. We expect to continue to see 4th generation languages and application develpoment environments becoming more menu-driven, more dialogue oriented and therefore operating in friendlier environments.□

Toronto meet to probe 4GL software productivity

The area of 4th Generation Language, DB management systems, productivity tools information centres and prototyping will get close attention in the Canadian National Database and 4th Generation Language Symposium, scheduled for Feb. 3-6, 1986 at the Ramada Renaissance Hotel in Toronto.

Organized by Digital Consulting Associates Inc., the event will feature product-oriented discussions and will show how the new generation of software can be profitably used to build systems both for the DP department and for the end user in the information centre. Included are 60 individual presentations by guest lecturers on the most popular DBMS and 4th Generation Languages.

Details are available from Digital Consulting Assoc., Andover, MA, (617) 470-3870. □