

PANEL

Are Standards the Panacea for Heterogeneous Distributed DBMSs?

Panel Chair:

Glenn Thompson, Amoco Production Company, Research

Panelists:

Hector Garcia-Molina, Princeton University

James Gray, Digital Equipment Corporation

Avi Silberschatz, University of Texas, Austin

Michael Stonebraker, University of California, Berkeley

Applications exist today in a wide variety of industries (for example, the oil industry, automobile manufacturing, and the banking industry) that require the integration of multiple existing distributed databases. Current database management systems (DBMSs) do not satisfy the requirements for these complex applications. Therefore, it is important for the database research community to provide leadership in this rapidly growing area so that future database systems better serve the needs of industry.

The importance of solutions to heterogeneous distributed database management system (HDDDBMS) problems has been extensively discussed. Discussions focusing on the potential problem areas can be found in the *Report of the Workshop on Heterogeneous Database Systems* (held at Northwestern University in December, 1989), in *Database Systems: Achievements and Opportunities* (Report on the NSF Invitational Workshop on Future Directions in DBMS Research), and in the *Proceedings of the 1990 ACM SIGMOD Conference*.

While HDDDBMS research has substantially benefited from homogeneous distributed DBMS research, there are a significant number of open problems unique to

HDDDBMS systems that must be addressed. For example, how can issues such as performance, extensibility, and security be satisfactorily resolved? What type of unified cost model is appropriate for HDDDBMS query optimization? Which DBMS standards are most critical to future HDDDBMS development? For example, if all local DBMSs support strict two-phase locking and two-phase commit, can the recovery processes used for homogeneous distributed database management systems be used for HDDDBMS systems? How does two-phase commit affect the autonomy of the local DBMSs? How does supporting non-standard DBMSs complicate matters?

How can databases that are incomplete or inconsistent be merged? How can effective data browsing be provided so that the compatibility of distributed databases be tested and verified? What types of HDDDBMS models would support scaling up to 1,000 sites (or even 10,000 sites) and how can data replication be handled effectively in these large systems?

These problems and others must be addressed before HDDDBMS systems become a reality.