SIGMOD 2001 Industrial Sessions

Guy Lohman (IBM)
lohman@almaden.ibm.com

César Galindo-Legaria (Microsoft)
cesarg@microsoft.com

Michael Franklin (U. California-Berkeley)
franklin@cs.berkeley.edu

Len Seligman (MITRE)
seligman@mitre.org

This report summarizes the Industry Track sessions at this year's International Conference on the Management of Data in Santa Barbara, California. There were three industrial sessions:

- Data Management in a Highly-Connected World,
- Applications and Lessons Learned, and
- Database Internals.

The papers are available in the SIGMOD 2001 conference proceedings and at http://www.acm.org/sigs/sigmod/sigmod01/eproceedings/.

Data Management in a Highly-Connected World

This session had three papers that exploit today's wide range of storage devices and the power of the Internet to deliver data transparently. The first described a product for managing a storage hierarchy for extremely large, multi-media databases (petabytes or even exabytes of data). The latter two pertained to obtaining data over the Internet.

The first paper, given by Felipe Carino of File Tek, Inc., described the StorHouse/Relational Manager (RM) product for managing an active hierarchy of devices that include RAID, disk, tape, and optical devices, for storing extremely large databases. While tape and optical disk have higher latencies, they are very cost effective – approximately 7% of the cost for disk – especially for large objects (LOBs). StorHouse RM is different than Hierarchical Storage Management (HSM), in which some of the data is archived and accessible only through an object identifier. StorHouse RM permits either horizontal partitioning of the data (e.g., some rows are on tape or optical disk) or vertical partitioning of the data (e.g., LOBs for each row are on tape or optical disk). The paper describes how StorHouse RM is implemented as a wrapper (i.e., a data source) to DB2 DataJoiner or DB2 UDB V7.1. A 0.05 scale factor TPC-H benchmark has been run on StorHouse RM, but was never published.

The second paper, presented by K. Selcuk Candan of NEC USA, Inc., was aimed at increasing the number of concurrent users that can connect to a database. Replicating servers may not scale well, and caching results or entire pages suffers from invalidation when updates occur. A technique used by IBM for the Olympics requires that the application be changed. The authors' propose instead to have a two-level approach, in which (1) a sniffer listens to HTTP requests from queries, and builds a map from the query to the appropriate URL; and (2) more selectively invalidating web content caches depending upon queries that are affected by updates (A comment from the audience noted, however, that a single URL usually maps to many queries.). This approach does not affect the applications or the application servers, as it runs on a separate machine and extracts the necessary information from three logs: (1) the HTTP request/delivery log, (2) the query instance

request/delivery log, and (3) the database update logs.

The third paper presented Netabase, a data management system developed by Siemens for a highly-distributed postal application with a large number (around 50,000) of sensors. Netabase exploits the Internet routing protocols to route queries to the proper data source, in order to ensure scalability and availability. The approach marries database and networking technology. It exploits a technique called "characteristic routing", which is implemented on top of IP and routes messages to nodes having certain "characteristics" (character strings that effectively serve as index keys). Netabase maps parts of an SQL query, such as the tables referenced and predicates, to form such characteristics. As the characteristics get more detailed, fewer nodes are involved in answering a query, but routing tables get larger and there is more administrative traffic to keep the advertised characteristics of nodes current. Intermediate routers aggregate the characteristics of nearby routers, permitting each query to be routed based upon increasingly precise information as it nears its destination and avoiding broadcasting to all nodes. Netabase adds network interfaces at the sensors to reply to queries, and at the querying points to assemble the data returned and perform joins. A prototype of this system for a parcel delivery application, called ParcelCall, has been completed in one fifth the time it took to implement a previous version, yet with more features.

Papers:

- Felipe Carino Jr., Pekka Kostamaa, Art Kaufmann, John Burgess, "StorHouse Metanoia - New Applications for Database, Storage & Data Warehousing"
- K. Selcuk Candan, Wen-Syan Li, Qiong Luo, Wang-Pin Hsiung, Divyakant Agrawal, "Enabling Dynamic Content Caching for Database-Driven Web Sites"
- Julio C. Navas, Michael Wynblatt, "The Network is the Database: Data Management for Highly Distributed Systems"

Applications and Lessons Learned

Overall, this session discussed some of the problems encountered when taking technical solutions and using them in real applications, as well as the commercial pressures and opportunities for data processing technology. Nazeri, Bloedorn, and Ostwald point out that there are interesting data mining problems that do not involve huge data volumes; the challenge is instead to find and exploit similarities on semi-structured data, where encoding domain knowledge can play a major role. Draper, Halevy, and Weld indicate that XML and distributed query processors can help make the old vision of federated databases a reality; that it is very much a real commercial opportunity; and that performance, manageability, and scalability cannot be ignored. Finally, Meseck shows that data processing technology got caught in the middle of the expectations of both customers and businesses about what Internet commerce *could* do. He details how the Internet environment changes the rules for what can be expected, and where current architectures and techniques have fallen short.

The first paper, presented by Zorah Nazeri of MITRE, described a system to track and extract information from aviation *incident reports*, which are standard forms with both text and structured data. It was a research project conducted for two European airlines, and it analyzed data collected from 1991 to 2000. One issue they faced was that reports are de-identified to protect pilots, before it is made available for processing, and this screening of detail adversely affected the efficacy of data mining. The other issue discussed was generation of many

uninteresting rules, which are difficult to filter out. A challenging part of the project was the need to consider the text part of the reports in the mining, so the authors developed similarity measures based on the text vocabulary.

The second paper, presented by Denise Draper of Nimble, described the architecture of the Nimble integration engine. Their goal is similar to that of federated databases -- to present a unified view of data from diverse sources. XML plays a central role in their system as the data interchange format. They use a hybrid data model
that allows mixing structured and unstructured data. The Integration Layer is a distributed query processor that maps queries from the unified model into the multiple sources, pushes work off to the sources, collects results and finishes off the processing. Updates are not supported in general on the global views (as that is unfeasible); however, they have found it is important for customers to expose primitive functionality of data sources, on the appropriate data fragments. Performance is critical for commercial success, as well as manageability and scaling to hundreds of data sources, many of them non-relational.

The third paper, presented by consultant Reed M. Meseck, described experiences of a typical Web business. Internet transactions open many possibilities, which on the technical side stretch traditional data processing architectures and techniques, and on the business side introduce costs and opportunities that are not fully understood yet. Customers expect immediate response times, any hour of the day, from any place in the world. Businesses expect they can track very detailed data that can later be exploited through things like data mining, and turned into further business opportunities. And data processing is caught in the middle of these various expectations. For Internet commerce to become a reality, data processing technologies need to evolve by better and more gracefully handling scalability, including on-line upgrade and expansion. And businesses need to structure their plans with a better understanding of cost - the Internet does make many things *possible*, but not all of them are *worthwhile* in terms of cost/benefit.

Papers:

- Z. Nazeri, E. Bloedorn, P. Ostwald, "Experiences in Mining Aviation Safety Data"
- D. Draper, A. Halevy, D. Weld, "The Nimble Integration Engine"
- R. Meseck, "Data Management: Lasting Impact of the Wild, Wild, Web Management"

Database Internals

This session included talks by representatives of the three major relational database vendors. The first two presentations, from Microsoft and IBM dealt with query processing issues, while the third talk, from Oracle, dealt with recovery issues.

The first paper, presented by César Galindo-Legaria described optimization techniques for subqueries and aggregations implemented in Microsoft SQL Server. The techniques are based on the observation that the processing of subqueries and aggregations are closely related, and that it is possible to define a small set of primitives that can be combined in various ways to implement these features. Use of these primitives permits cost-based optimization to be extended to classes of queries that previously had to depend on heuristic rewriting techniques.

The second paper, presented by Parke Godfrey of York University, described work being done in the context of the IBM DB2 system, to exploit so-called "soft constraints" for query optimization. Soft constraints are analogous to integrity constraints except they are not maintained by the database system, and thus, may become invalid over time. These constraints

can be used to perform various types of semantic query optimizations. Two types of soft constraints are investigated: "absolute" ones, which are consistent with respect to the current state of the database, and "statistical" ones, which may not be completely consistent with any state of the database. The former can be used during query rewrite and cost estimation, while the latter can only be used for cost estimation.

The final paper of the session, presented by Tirthankar Lahiri, described innovative techniques used in the Oracle database system to give users control over recovery time and to improve recovery performance. The main techniques presented was the "Fast-Start" roll-forward and roll-back mechanisms. Fast-Start roll-forward is implemented in a way that gives administrators some control over redo time by adjusting checkpointing frequency to specified recovery time requirements. Fast-Start roll-back exploits the Oracle locking mechanism to allow online access to the database while roll-back is in progress. Finally, the notion of "flash-back" queries, in which queries are executed against earlier states of the database were presented. These can be used as a form of disaster recovery.

Papers:

- César Galindo-Legaria and Milind Joshi (Microsoft), "Orthogonal Optimization of Subqueries and Aggregation"
- Parke Godfrey (York Univ.), Jarek Gryz (York Univ.), and Calisto Zuzarte (IBM Canada), "Exploiting Constraint-Like Data Characterizations in Query Optimization"
- Tirthankar Lahiri (Oracle), Amit Ganesh (Oracle), Ron Weiss (Oracle), and Ashok Joshi (NuGenesis Tech.), "Fast-Start: Quick Fault Recovery in Oracle"