

Topic 05

Parallel and Distributed Databases

Burkhard Freitag and Kader Hameurlain

Co-chairmen

Although parallel and distributed database have been known for several years and are an established technology, new application domains require a deeper understanding of the underlying principles and the development of new solutions.

Besides such important questions as parallel query processing, optimization techniques, parallel database architectures, load sharing and balancing to name but a few, transaction management, data allocation, and specialized algorithms for advanced applications are topics of high relevance.

The parallel transactional execution of operations is addressed by the following three papers.

Data checkpointing is essential in distributed transaction processing and thus in distributed database systems. Roberto Baldoni et al. address the problem of how to determine consistent checkpoints in distributed databases and present formal solutions.

Armin Fessler and Hans-Jörg Schek investigate the similarities of database transaction management, in particular the scheduling of nested transactions, and parallel programming.

The standard serialization approach to object locking causes a considerable performance loss on shared memory architectures. Christian Jacobi and Cedric Lichtenau describe a serialization free approach to object locking in parallel shared memory architectures.

Data allocation and redistribution is critical for the overall performance of a distributed or parallel database system.

The article by Holger Märtens explores ways of allocating intermediate results of large database queries across the disks of a parallel system. As a solution, declustering even of self-contained units of temporary data, e.g. hash buckets, is proposed. An analytical model is presented to show that the performance gain by improved parallel I/O justifies the costs of increased fragmentation.

Particular problem classes and/or architectures allow specialized algorithms and thus some extra gain in performance.

Basilis Mamalis et al. describe a parallel text retrieval algorithm designed for the fully connected hypercube. The parallel protocols presented are analyzed both analytically and experimentally.

The papers reflect the vivid discussion and ongoing research and development in the field of distributed and parallel databases which is not only intellectually challenging but also commercially most relevant.