

Parallel database management

Parallel database management refers to the management of data in a multiprocessor computer and is done by a parallel database system, i.e., a full-fledge DBMS implemented on a multiprocessor computer. ... This enables supporting very large databases with very high query or transaction loads.

- The main advantage to parallel databases is speed. The server breaks up a user database request into parts and dispatches each part to a separate computer. They work on the parts simultaneously and merge the results, passing them back to the user.
- Parallel databases can be roughly divided into two groups :
 - The first group of architecture is the multiprocessor architecture.
 - the other are group includes

Shared memory architecture:- Where multiple processors share the main memory (RAM) space but each processor has its own disk (HDD). If many processes run simultaneously, the speed is reduced, the same as a computer when many parallel tasks run and the computer slows down.

Shared disk architecture:- Where each node has its own main memory, but all nodes share mass storage, usually a storage area network. In practice, each node usually also has multiple processors.

Shared nothing architecture:- Where each node has its own mass storage as well as main memory.

- The other architecture group is called hybrid architecture, which includes:
 - Non-Uniform Memory Architecture (NUMA), which involves the non-uniform memory access.
 - Cluster (shared nothing + shared disk: SAN/NAS), which is formed by a group of connected computers.

Parallel DBMS Techniques

- Data placement
 - ➔ Physical placement of the DB onto multiple nodes
 - ➔ Static vs. Dynamic
- Parallel data processing
 - ➔ Select is easy
 - ➔ Join (and all other non-select operations) is more difficult
- Parallel query optimization
 - ➔ Choice of the best parallel execution plans
 - ➔ Automatic parallelization of the queries and load balancing
- Transaction management
 - ➔ Similar to distributed transaction management

Parallel DBMS Functions

❖ Session manager

- ◆ Host interface
- ◆ Transaction monitoring

❖ Request manager

- ◆ Compilation and optimization
- ◆ Data directory management
- ◆ Semantic data control
- ◆ Execution control

❖ Data manager

- ◆ Execution of DB operations
- ◆ Transaction management support
- ◆ Data management