Architecture of Database System

---- Joseph M. Heller stein, Michael Stonebraker and James Hamilton.

Summary

The Intentions of this paper are

- Why database Architecture is not broadly known by the people.
- The two main reasons behind this is The database community is very small such that there are few successful DBMS Implementations due to less support from market and Academically concentrating more on Algorithms and theory rather than on architectural issues.
- He discussed the main architectural aspects of modern Database Systems.
- He focused on architectural discussion of DBMS Design Principles Which Include: Process Models, Parallel Architecture, Storage System Design, Transaction Systems, Query Processor & Optimizer, Shared Components & Utilities.

It is important to know the database system architecture which helps in academic research and developing industrial strength for high-end customers.

The Main Ideas proposed in the paper are

- The Main components of RDBMS are:
  - Process Manager
  - Client Communication Manager
  - Relational Query Processor
  - Transcational Storage Manager
  - Shared Components & Utilities
- He explained the life of a query in database system in following steps

1. When client calls an API, It establishes connection with Client Communication Manager. The connection establishment can be done via Two Tier (Client-server) or Three Tier (Web Server) or Four Tier (Application Server). So, The DBMS needs to be compatible with all different connectivity protocols. The main responsibility of client communicator is to establish connection, Remember connection state, Respond to SQL commands from caller and to return data & control messages.
2. After receiving SQL commands, DBMS Process Manager assigns a thread of computation to the command and make sure that thread's data and control

outputs are connected to the client. Then the Admission control decides whether to begin processing immediately or wait till enough resources are available.

3. After admitting, the client's query can begin to execute by invoking the code in Relational Query Processor. Then It checks whether the user is authorized to run the query and compiles SQL query into Query plan. The Query plan is handled by Plan executor which contains a set of operations assembled by Query Optimization for executing queries.
4. The operators from Query plan request data from DBMS Transactional Storage Manger which manages all data access and manipulation calls. The storage system includes algorithms & data structures for organizing and accessing data on disk. It also includes a Buffer management module which tells when and what data needs to be transferred between disk & memory. To ensure correct execution, Locks are acquired from Lock manager. Log Manager ensures that the transaction is committed or roll back if undone.
5. Finally, the query is processed, executed, computed results and placed in client communication manager which send the results to the client. After the end of transaction, connection is closed, Transaction manager is cleaned and process manager frees the control structures.

- The Shared Components & Utilities are vital to the fully functional operation of DBMS.
- The catalog is used by Query processor for authentication, Parsing and Query Optimization. The Memory Manger is used to allocate or de allocate memory dynamically.
- The remaining utilities are used to keep database well-tuned and reliable.

Question on this Paper:

Does the DBMS base architecture remain same for all softwares like MySQl, Oracle, Teradata because the relational model & query remains the same and even the output remains same?

Submitted By

Lakshman Madhav Kollipara

SID- 932655504

Submitted To

Arash Termehchy

CS540- Database Management Systems