Abstract

Database is defined as collection of files or table, whereas DBMS stands for Database Management System which is collection of unified programs used to manage overall activities of the database. The two dominant approaches used for storing and managing database are centralized database management system and distributed database management system in which data is placed at central location and distributed over several locations respectively. Independent of the database approach used, one of the foremost issues in the database is the retrieval of data by using multiple tables from central repository in centralized database and from number of sites in distributed database. Joins and semi-joins are primitive operations used to extract required information from one, two or multiple tables. In this paper the focus is given on computing and analyzing the performance of joins and semi-joins in distributed database system. The various metrics that will be considered while analyzing performance of join and semi-join in distributed database system are Query Cost, Memory used, CPU Cost, Input Output Cost, Sort Operations, Data Transmission, Total Time and Response Time. In short the intention of this study is analyze the performance and behavior of join and semi-join approach in distributed database system.
References

- Nilarun Mukherjee, Synthesis of Non Replicated Dynamic Fragment Allocation Algorithm in Distributed Database System, Published in Proceeding of international conference on advances in Computer Science, 2010.
- Deepak Shukla, Dr. Deepak Arora, An Efficient Approach of Block Nested Loop Algorithm based on Rate of Block Transfer, IJCA, Vol. 21, No. 3, May 2011.

Index Terms

Computer Science  Databases

Keywords

Distributed Database  Data Transmission  Response Time  Total Time  Join
Semi join