JOMR: Multi-join Optimizer Technique to Enhance Map-Reduce
Job

أسماء المشتركين فى البحث :

Mina Samir Shenouda- Mohamed Helmy Khafagy- Samah Ahmed Senbel

ملخص البحث باللغة الانجليزية

Execution an environment developed by Google to process very large amounts of data. Query optimizer is needed to find more efficient plans for declarative SQL query. In classic database: join algorithms are optimized to execute the entire query result, but they ignore the importance of tables order especially in multi-join query. But we can see that the orders for tables are an important factor to get the best performance of a query plan and will be very effective in performance when join tables content huge number of rows in addition to more than one join operation. In this paper we proposed a new technique called JOMR (Join Order in Map-Reduce) that optimizes and enhances Map-Reduce job. This technique uses enhanced parallel Travel Salesman Problem (TSP) using Map-Reduce for improving the performance of query plans according to change the order for join tables. Also we build a cost model that supports our algorithm to find best join order. We will focus on Hive especially multi-join query and our experiments result for JOMR algorithm proving the effectiveness of our query optimizer and this performance is improved more when increasing the number of join and size of data.