Volume 8, No. 5, May – June 2017
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

Parallel Database for Student Counselling through Single Window System for Admission in Engineering Colleges

Dr. D. Asir Antony Gnana Singh
Department of Computer Science and Engineering
Anna University, BIT-Campus,
Tiruchirappalli, India

Dr. E. Jebamalar Leavline
Department of Electronics and Communication Engineering
Anna University, BIT-Campus,
Tiruchirappalli, India

B. Preethi
Department of Computer Science and Engineering
Anna University, BIT-Campus,
Tiruchirappalli, India

Abstract: Nowadays, the growth of the educational sector generates massive data such as details of students, staff and their academic performance, administrative and research details, etc. These data are stored in distributed databases. The parallel database performs parallel query processing hence, the query processing time is reduced and throughput for the transaction processing is improved. The parallel database system also improves the performance by employing parallelism through various database management operations such as loading the data, building indexes, and evaluating queries. The data is partitioned and placed across multiple disks for parallel input and output (I/O) operation to achieve parallelism on queries using pipeline parallelism. The individual relational operations such as sort, join, aggregation, etc. are executed in parallel database by each processor that works independently with its own partition. Queries are expressed in high level language such as structured query language (SQL) that are translated into relational algebra for query processing. Thus, the parallel database increases the throughput in database management and reduces the query processing time. This paper presents a parallel database for student counselling for engineering university admission using MySQL relational database management system.

Keywords: Distributed database; Parallel Database; Structured query language; Database management system; MySQL;

I. INTRODUCTION

In recent past, thanks to the massive growth of the engineering colleges, huge numbers of students are admitted to pursue engineering courses. The counselling for admission in various engineering colleges is carried out in a single window system in a centralized manner by the universities. The increasing number of engineering colleges and students widen the data required for process during the counselling time. Hence, processing the data with traditional database is quite complex. The parallel database system seeks to improve the performance through parallelization of various operations such as loading data, building indexes and evaluating queries. Parallel database improves processing and the speed of input and output operation using multiple central processing units (CPU) that run in parallel on the partitions of data that reside over in multiple disks in parallel.

The concurrency control is used to avoid conflict among the different queries which are processed in parallel manner by different CPU. Thus, the database increases the parallelism and reduces the time required to retrieve relations from disk by partitioning the relations as multiple disks. Each processor Pi sorts it's partitioned of the relation locally redistribute the relation using range partition all tuple that lie in the range are sent to processor Pi. Parallel systems usually give results which fall somewhere between pure plurality or majority, purity and pure system. Parallel processing system is extensively developed and used in numerous commercial servers for large-scale data analysis. Parallel computing is a fast growing area since cost of the microprocessor, memory and the disks is dropped down and the growth of the data is massively increasing day-by-day through various transactions that are carried out in various sectors such as banking, trading, shopping, etc. The parallel database can be categorized into different multiprocessor architectures such as shared memory, shared disk, and shared nothing architecture. In multiprocessor architecture, more than one processor is participating for computation. In some multiprocessor system, all the processors are treated equally and these processors are known as symmetric multiprocessing (SMP).

In some multiprocessor systems, all processors cannot be treated equally, some of the processors can be used for special operations. Such processors are called as asymmetric multi processing (AMP). The multiprocessor system can be divided into two types namely tightly and loosely coupled. In tightly coupled system, the process can share the central shared memory. In loosely coupled system, multiple stand-alone systems are inter-connected through a network. In shared memory architecture, multiple processors share the common memory but each processor has its own disk. In shared disk architecture, multiple processors share the common disk space but each processor has its own memory. In shared nothing architecture, each processor has its own disk and its own memory.

This paper presents a parallel database application for student counselling for admission through single window system for engineering college. The vacancies of the departments of the colleges are maintained in three sites. Seats are allotted with these three sites simultaneously.

The rest of this paper is organized as follows: Section 2 reviews the literature. Section 3 presents the parallel database for the students counseling. Section 4 gives the implementation details. Section 5 discusses the outputs that are obtained. Section 6 concludes this paper.

II. LITERATURE REVIEW

This section presents the related research works that are carried out by many researchers. The parallel databases are used for various commercial applications and they are used in different servers to handle massive data. Achieving the quality attributes such as efficiency, reliability, scalability is a challenging task in these parallel database systems especially when performing the parallel operation for query processing on the massive data. Hence, the multilevel-parallel database system is used to handle the massive data. The multilevel parallel data base system can reduce the query processing time and data access time. It also reduces the network input and output operations. Furthermore, it reduces the disk input and output operations [1].

In parallel database environment, the data are placed in different sites and these sites are connected with the network. The parallel query processing is executed by accessing the data from various dataset sites through the network. Hence, the less network bandwidth makes the bottleneck for parallel databases and increases the data accessing time for execution. Xuan Ping presented a parallel database system to reduce the execution time using the better node section algorithm with intra-operation parallelism. Moreover, it improves the execution of parallel multi-join queries [2]. The large scale databases adopt some techniques to improve the performance of the parallel databases such as fault tolerance, distributed transactions, query optimization, and distributed query processing [3].

In general, load balancing is performed in the parallel databases in order to improve the effective utilization of parallel database. Load balancing can be carried out in different levels such as initiating the transactions, query processing and sub-query processing at individual sites, etc. In general, the intra and inter query parallelism is carried out in many distributed databases. The joining operation algorithms are used to improve the effectiveness of parallelism. This joining operation algorithm incorporates load balancing mechanisms to reduce the skew effect. The skew effect is a scenario of uneven distribution of data or queries or workload among the distributed systems [4]. Kien A et al. presented the load balancing scheme for parallel hash-joining algorithm and presented three parallel hash-joining scenarios to improve load balancing to reduce the effect of skew effect [5]. The heterogeneous sites can lead to problems in balancing the parallel database queries. In order to avoid these problems, the load balancing mechanisms are developed with different approaches. Anastasios Gounaris et al. developed a load balancing scheme with linear quadratic regulation (LQR) for balancing parallel query execution in parallel databases [6].

The computational grid is another approach to increase the performance of the system using parallel computing by sharing the resources [7]. Wu Weining et al. presented a performance analysis on the parallel database in the grid environment. The performance metrics such as scalability, speed up, execution time are used to evaluate the performance of the parallel processing systems [8]. D.A.A.G Singh et al. presented a distributed database for bookstore application [9]. From the literature review, it is observed that the parallel databases are used for various applications for managing data. The parallel databases can reduce the time taken to execute the query for different database management operations and improve the throughput. Thus, this paper develops a parallel database for student counselling through single window system for admission in engineering colleges.

III. PARALLEL DATABASE FOR STUDENT COUNSELLING

The architecture of the parallel database for the student counselling for admission in three engineering colleges is illustrated in Figure 1. This parallel database consists of three sites namely Site 1, Site 2, and Site 3. Each site is located in an engineering college and the main server is located in the University in order to access the data from each site. Each site has its own database. Each site is connected with the main server through the network connection. The main server contains different seat allocation details of three colleges. The main server computes the sorting and merging algorithm to prepare the student rank list by collecting the details from the three sites that are located in three different engineering colleges.

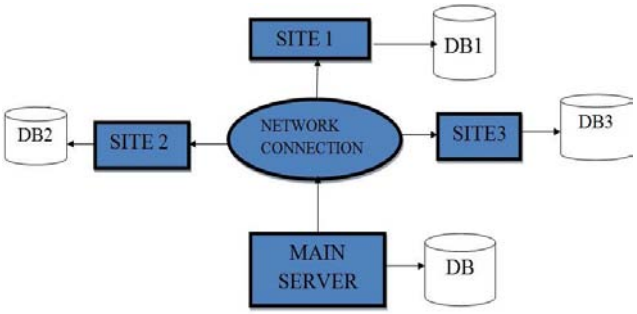

Figure 1. Architecture of parallel database for student counselling

IV. IMPLEMENTATION DETAILS AND PROCEDURE

This section provides the details of implementation and implementation procedure for developing the parallel data base for student counselling. In order to implement this parallel database, computer system is used with the configuration of Processor: Intel(R) Core (TM) i5-3470 CPU @ 3.20 GHz, RAM: 4 GB, Hard disk: 500GB and the operating system: Windows 8. This parallel database system is implemented using MySQL Server version 5.7. The following steps are carried out to implement the parallel database.

STEP 1: Install the MySQL installer and set the own password.
STEP 2: Open the MySQL.
STEP 3: Create a database (DB) with the name of "college" for each site located in the different colleges
STEP 4: Create three tables with the name of eng1, eng2, eng3, respectively for each database "college"
STEP 3.1: Insert the seat allocation details such as college code, college name, department name, total seats, available seats with the name of coll_code, coll_name, dept, total_seats, avail_seats, respectively for each table "college".
STEP 4: Calculate and display the available seats of different college databases.
STEP 4.1: Calculate and display the available seats for departments such as CSE, ECE in each engineering colleges.
STEP 5: Compute the merging algorithm to display

the overall college wise seat availability.

V. OUTPUT AND DISCUSSION

This section discusses the output obtained from the experiments that are conducted.

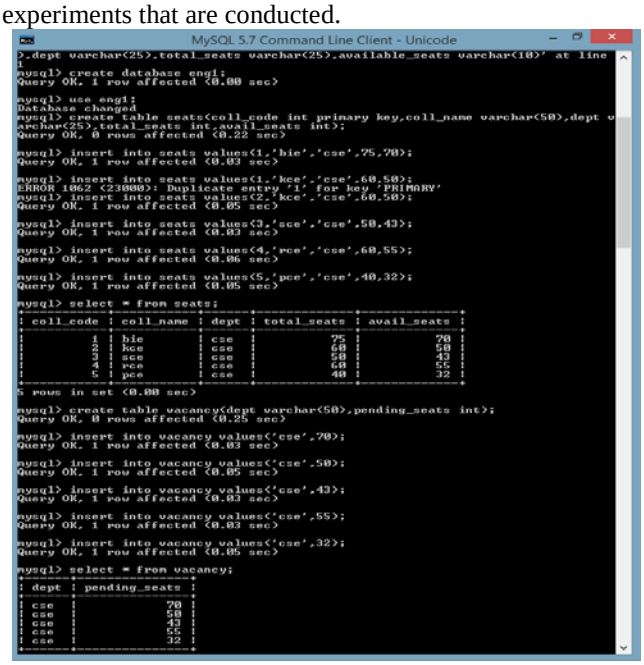

Figure 2. Creation of database namely 'college' that contains a table namely 'eng1' with the attributes namely coll_code, coll_name, dept, total_seats, avail_seats and displaying the department wise available seats of engineering college eng1.

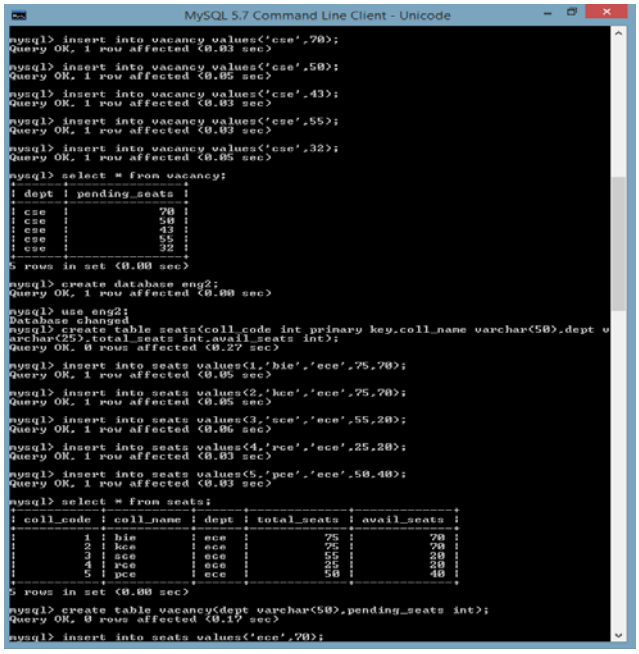

Figure 3. Creation of database namely 'college' that contain a table namely 'eng2' with the attributes namely coll_code, coll_name, dept, total_seats, avail_seats and displaying the department wise available seats of engineering college eng2.

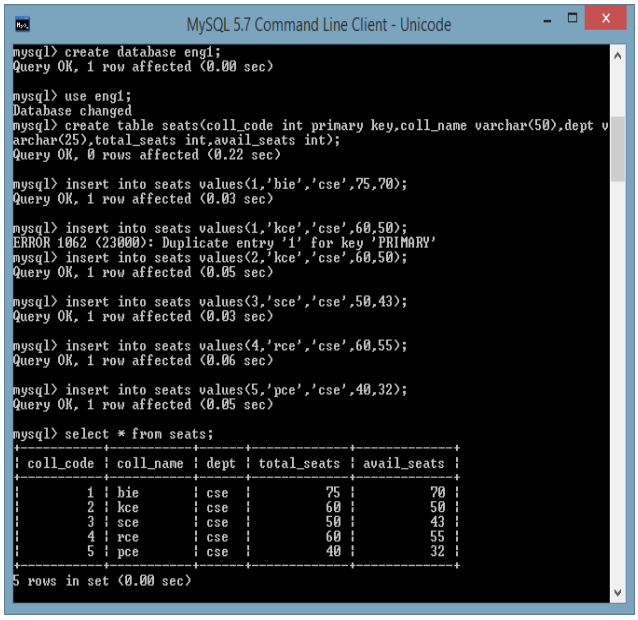

Figure 4. Creation of database namely college that contain a table namely eng 3 with the attributes namely coll_code, coll_name, dept, total_seats, avail_seats and displaying the the department wise available seats of engineering college eng3.

Thus, the parallel database is implemented for the student counselling as shown in Figure 2 to Figure 4. The implementation is carried out with the creation of database namely college that contain a table namely eng1, eng2 and eng 3 with the attributes namely coll_code, coll_name, dept, total_seats, avail_seats and available seats are displayed department wise with respective engineering colleges.

VI. CONCLUSION

This paper presented a parallel database for student counselling through single window system for admission in engineering colleges. The college details, department details and vacancy details are maintained in three different sites, by using the information from the three sites seats can be allocated to a student using parallel query processing. It can also provide information of different colleges for the students. In future, this work can be extended to implement this system with more sites for university counselling for engineering students using parallel database system.

VII. REFERENCE

[1] Lei Yu, Ge Fu, Yan Jin, Xiaojia Xiang Huaiyuan Tan, Hong Zhang, Xinran Liu, Xiaobo Zhu "MPDBS: A multi-level parallel database system based on B-Tree," 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Takamatsu, 2015, pp. 1-7.

[2] Xuan Ping, "Selection of Nodes for Distributing Relations in Parallel Database", 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics,

[3] M. W. Bright, A. Helal and A. R. Hurson, "A mini-track on parallel and distributed database systems: track coordinators' introduction," 1994 Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, Wailea, HI, USA, 1994, pp. 203-205.

[4] Märtens H. (2001) A Classification of Skew Effects in Parallel Database Systems. In: Sakellariou R., Gurd J.,

Freeman L., Keane J. (eds) Euro-Par 2001 Parallel Processing. Euro-Par 2001. Lecture Notes in Computer Science, vol 2150. Springer, Berlin, Heidelberg

[5] K. A. Hua, Chiang Lee and C. M. Hua, "Dynamic load balancing in multicomputer database systems using partition tuning," in IEEE Transactions on Knowledge and Data Engineering, vol. 7, no. 6, pp. 968-983, Dec 1995.

[6] A. Gounaris, C. A. Yfoulis and N. W. Paton, "An efficient load balancing LQR controller in parallel database queries under random perturbations," 2009 IEEE Control Applications, (CCA) & Intelligent Control, (ISIC), St. Petersburg, 2009, pp. 794-799.

[7] A. Barmouta and R. Buyya, "GridBank: a Grid Accounting Services Architecture (GASA) for distributed systems sharing and integration," Proceedings International Parallel and Distributed Processing Symposium, 2003, pp.

[8] W. Weining, S. Guozhi and L. Xingang, "Performance analysis of parallel database in grid," 2010 International Conference on Computer and Information Application, Tianjin, 2010, pp. 323-325.

[9] D. Asir Antony Gnana Singh, E. Jebamalar Leavline, and S. Malini “Distributed Database for Bookstore Application” International Journal of Advanced Research in Computer Science, Volume 8, No. 3, pp. 309-312, March – April 2017