Advanced Database System Lab Assignment 6 Parallel Databases

Introduction to Parallel Databases

Companies need to handle huge amounts of data with high data transfer rates. The client server and centralized system is not very efficient. The need to improve the efficiency gave birth to the concept of Parallel Databases.

Parallel database system improves performance of data processing using multiple resources in parallel, like multiple CPU and disks are used parallely.

It also performs many parallelization operations like, data loading and query processing.

Goals of Parallel Databases

The concept of Parallel Database was built with a goal to:

Improve performance:

The performance of the system can be improved by connecting multiple CPU and disks in parallel. Many small processors can also be connected in parallel.

Improve availability of data:

Data can be copied to multiple locations to improve the availability of data.

For example: if a module contains a relation (table in database) which is unavailable then it is important to make it available from another module.

Improve reliability:

Reliability of the system is improved with completeness, accuracy and availability of data.

Provide distributed access of data:

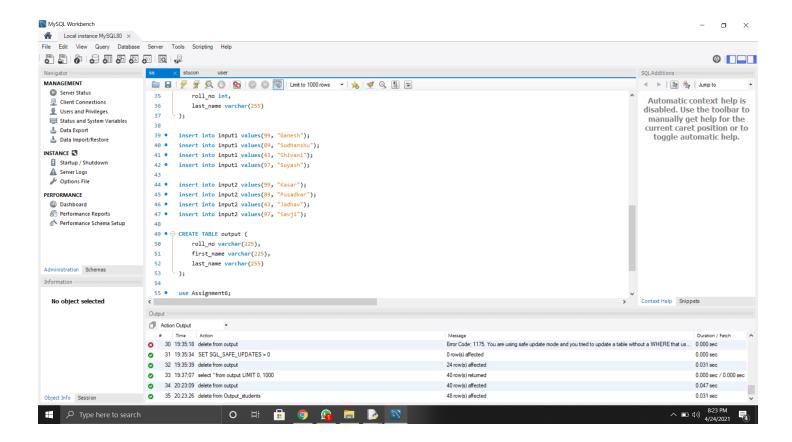
Companies having many branches in multiple cities can access data with the help of a parallel database system.

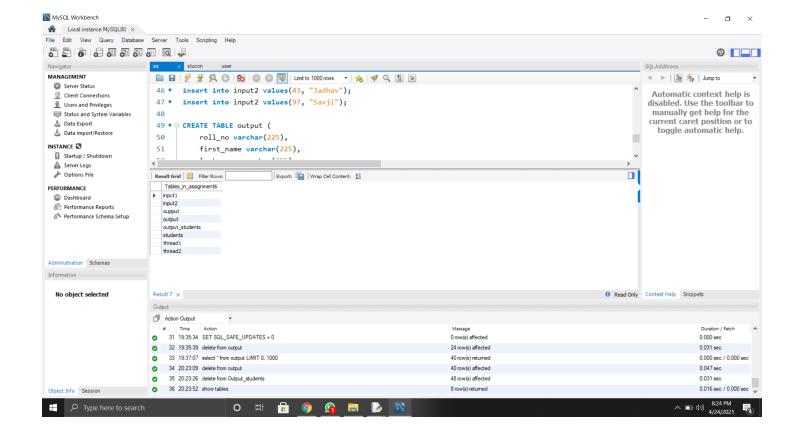
Steps: 1. Created the Tables ParallelSort():

- 1. Students
- 2. Output_students

ParallelJoin():

- 1. Input1
- 2. Input2
- 3. Output



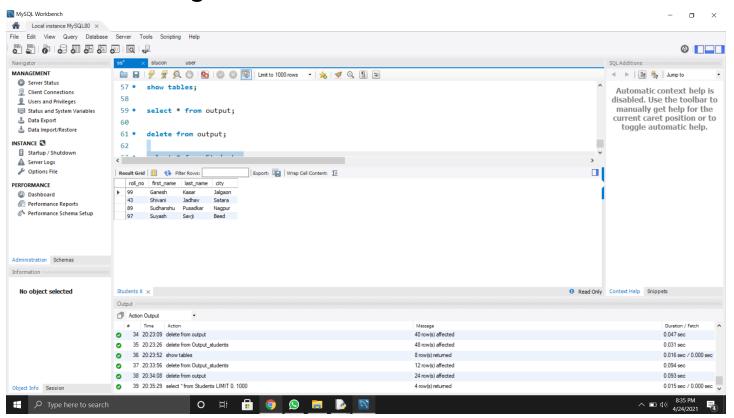


Step 2: Created Python Program

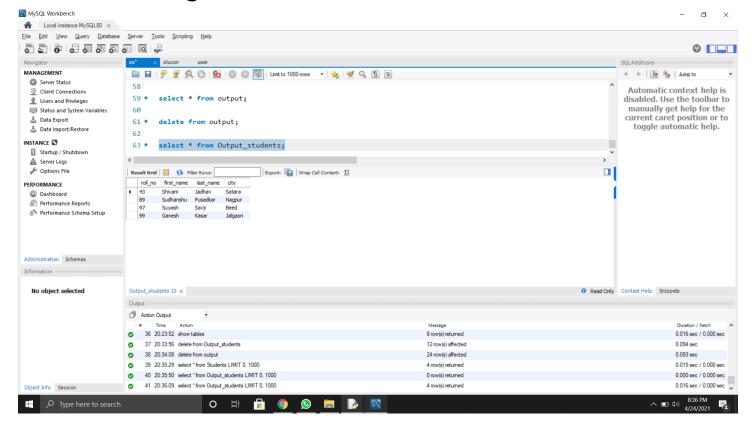
- 1. Created Connection to MySQL using MySQL connector
- 2. Sorted the Students table using threads
- 3. Joined the input1 & input2 stored in new table output

```
IDLE Shell 3.9.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - 🗆 X
                               mysql_python.py - D:\TY\ADS\Assignment6\mysql_python.py (3.9.2)
                                                                                                                                                                                                                                                                                                                                                                                                              File Edit Shell Debug Options Window Help
                             File Edit Format Run Options Window Help
                                                                                                                                                                                                                                                                                                                                                                                                             Python 3.9.2 (tags/v3.9.2:la79785, Feb 19 2021, 13:44:55) [MSC v.1928 64 bit (AM D64)] on win32
     o ·
                             Hile cost Format Kun Oppons Window Help
#Parallel Join
def ParallelJoin(input1, input2):
    threads = ["thread1", "thread2", "thread3", "thread4", "thread5"]
    temp threads = []
    for i in range(5):
                                                                                                                                                                                                                                                                                                                                                                                                                                  "help", "copyright", "credits" or "license()" for more information.
                                                                                                                                                                                                                                                                                                                                                                                                                         comparison | 
                                                           input1,
                                                                                                            input2
                                                                                                                                                                                                                                                                                                                                                                                                                           ad <Thread(Thread-2, started 11112)> is sorting [(99, 'Ganesh', 'Kasar', 'Jal n')]
3, 'Shivani', 'Jadhav', 'Satara'), (89, 'Sudhanshu', 'Pusadkar', 'Nagpur'), (
'Suyash', 'Sayji', 'Beed'), (99, 'Ganesh', 'Kasar', 'Jalgaon')]
Joining the inputl and input2
,'Ganesh', 'Kasar']
,'Sudhanshu', 'Pusadkar']
,'Shivani', 'Jadhav']
,'Suyash', 'Sayji']
,'Ganesh', 'Kasar']
,'Suhanshu', 'Pusadkar']
,'Shivani', 'Jadhav']
,'Suyash', 'Sayji']
,'Ganesh', 'Kasar']
,'Sudhanshu', 'Pusadkar']
,'Shivani', 'Jadhav']
,'Suyash', 'Sayji']
,'Ganesh', 'Kasar']
,'Suchanshu', 'Pusadkar']
,'Suyash', 'Sayji']
,'Ganesh', 'Kasar']
,'Suyash', 'Sayji']
                                                             temp_threads[i].start()
                                            for t in temp_threads:
                                            t.join()
res = []
                                            res = []
[res.append(x) for x in output if x not in res]
for item in output:
    sql = "INSERT INTO output VALUES (%s, %s, %s)"
    val = (str(item[0]), item[1], item[2])
    mycursor.execute(sql, val)
  7
                                                           mydb.commit()
                              Database Connection
                               Database Connection
ydb = mysql.connector.connect(
   host="localhost",
   user="2018BCGRP28",
   passwd="pass",
   database="Assignment6"
                             mycursor = mydb.cursor()
                             sql = "SELECT * FROM Students"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ln: 34 Col: 4
                                                                                                                                                                                                                                                                                                                                                      In: 75 Col: 34
                                  ganekasar.c...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ヘ ロ か) 8:24 PM 4/24/2021
Type here to search
                                                                                                                                                                                                                   O # 🔒 🧿 😘 📜 🕟
```

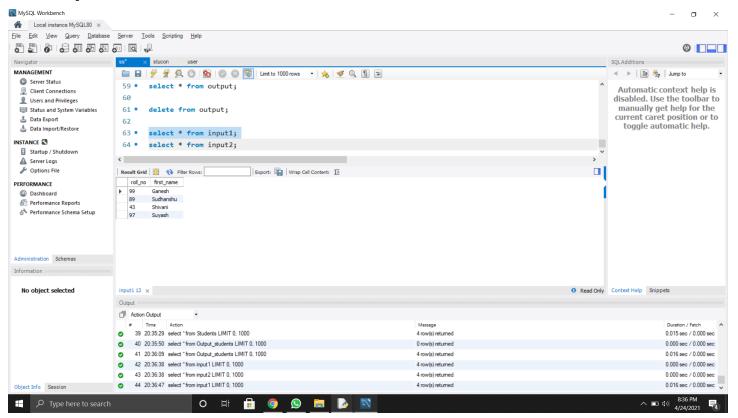
1. Before Sorting



2. After Sorting



1. Input Tables



2. Output tables

