

# **Database System Implementation**

## **Project 2 - Implementing a Sorted File**

**Developed by –**

Ghodkari Chowdary, Raghunatha Rao - UFID: 6218-1051

Mullapudi, Aseesh - UFID: 9175-1971

Aim of this project is to implement a Sorted file.

We have completed the BigQ class by writing the required functions. The job of the BigQFile within the system is to sort the records based on the input predicate, write to console and store the sorted file with .bigq extension.

### **Implementation Details:**

We have created the following private fields in the BigQ.h

**Pipe \*inputPipe;** // input Pipe to get the records

**Pipe \*outputPipe;** // output Pipe to push records

**OrderMaker \*sortedOrder;** // sorted Order required for sorting

**int \*runlength;** // the run length input from console

**File \*runsFile;** // File pointer for Runs File which stores the Runs

**vector <int> runPointers;** // list of pointers to all the runs

### **Method Definitions:**

- ***BigQ::BigQ(Pipe &in, Pipe &out, OrderMaker &sortorder, int runlen)***

Constructor used to initialize the fields of the BigQ class, A new runs File is created to store the runs

- ***static void\* invoke\_tpmmsAlgo(void\* args);***

Function which is the start of the worker thread which performs the sorting, we are typecasting the void arguments to bigQ and invoking the worker function

- ***void BigQ::worker()***

worker function which does the TPMMS Algorithm, Algorithm is written in two phases where first phase breaks the File into runs and sorts each run. In second phase all the runs are merged into a single sorted file using priority Queue

- ***void BigQ::sortRun( vector <Record\*> & )***

function which sorts the vector of records using the input predicate

- ***int BigQ::addRunToFile( vector <Record\*> & )***

function take input of sorted records into vector and writes them to File and return the current size of file which act as pointer to the next run

File **gtest.cc** written to unit test the TPMMS Algorithm of the BigQ on different tables

## Instructions to Run the Code:

### Compile and run test.cc:

Ensure the directory contains the bin files generated from the Project1 so that code has access to Heap files

Ensure there is Catalog file in the root folder as well

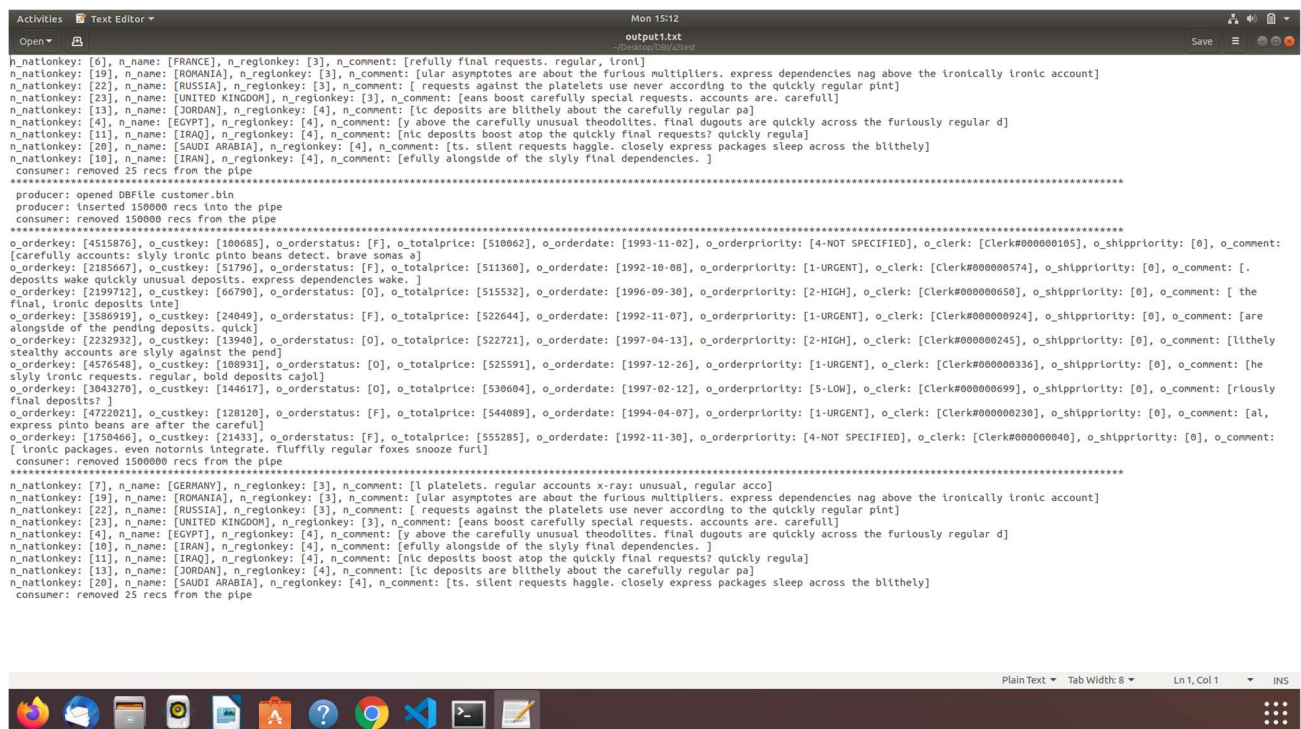
1. make
2. ./test.out

### To run the test cases

1. make
2. ./runsTestCases.sh

File output1.txt is generated

Screen Shot of the output1.txt is below



```
h_nationkey: [6], n_name: [FRANCE], n_regionkey: [3], n_comment: [refully final requests. regular, iron]
n_nationkey: [19], n_name: [ROMANIA], n_regionkey: [3], n_comment: [ular asymptotes are about the furious multipliers. express dependencies nag above the ironically ironic account]
n_nationkey: [22], n_name: [RUSSIA], n_regionkey: [3], n_comment: [ requests against the platelets use never according to the quickly regular pint]
n_nationkey: [23], n_name: [UNITED KINGDOM], n_regionkey: [3], n_comment: [eans boost carefully special requests. accounts are. careful]
n_nationkey: [13], n_name: [JORDAN], n_regionkey: [4], n_comment: [lc deposits are blithely about the carefully regular pa]
n_nationkey: [4], n_name: [EGYPT], n_regionkey: [4], n_comment: [y above the carefully unusual theodolites. final dugouts are quickly across the furiously regular d]
n_nationkey: [11], n_name: [IRAQ], n_regionkey: [4], n_comment: [nic deposits boost atop the quickly final requests; quickly regula]
n_nationkey: [20], n_name: [SAUDI ARABIA], n_regionkey: [4], n_comment: [ts. silent requests haggle. closely express packages sleep across the blithely]
n_nationkey: [10], n_name: [IRAN], n_regionkey: [4], n_comment: [efully alongside of the slyly final dependencies. ]
consumer: removed 25 recs from the pipe
*****
producer: opened DBFile customer.bin
producer: inserted 150000 recs into the pipe
consumer: removed 150000 recs from the pipe
*****
o_orderkey: [4515876], o_custkey: [100605], o_orderstatus: [F], o_totalprice: [510062], o_orderdate: [1993-11-02], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000105], o_shippriority: [0], o_comment: [carefully accounts: slyly ironic pinto beans detect. brave sonas a]
o_orderkey: [2185667], o_custkey: [51790], o_orderstatus: [F], o_totalprice: [511360], o_orderdate: [1992-10-08], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000574], o_shippriority: [0], o_comment: [. deposits wake quickly unusual deposits. express dependencies wake. ]
o_orderkey: [2199712], o_custkey: [66790], o_orderstatus: [0], o_totalprice: [515532], o_orderdate: [1996-09-30], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000650], o_shippriority: [0], o_comment: [ the final, ironic deposits lnte]
o_orderkey: [3586919], o_custkey: [24049], o_orderstatus: [F], o_totalprice: [522644], o_orderdate: [1992-11-07], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000924], o_shippriority: [0], o_comment: [are alongside of the pending deposits. quick]
o_orderkey: [2232932], o_custkey: [13940], o_orderstatus: [0], o_totalprice: [522721], o_orderdate: [1997-04-13], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000245], o_shippriority: [0], o_comment: [lithely stealthy accounts are slyly against the pend]
o_orderkey: [4576548], o_custkey: [108931], o_orderstatus: [0], o_totalprice: [525591], o_orderdate: [1997-12-26], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000336], o_shippriority: [0], o_comment: [he slyly ironic requests. regular, bold deposits cajol]
o_orderkey: [3043270], o_custkey: [144617], o_orderstatus: [0], o_totalprice: [530604], o_orderdate: [1997-02-12], o_orderpriority: [5-LOW], o_clerk: [Clerk#000000699], o_shippriority: [0], o_comment: [rlously final deposits? ]
o_orderkey: [4722021], o_custkey: [128120], o_orderstatus: [F], o_totalprice: [544089], o_orderdate: [1994-04-07], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000230], o_shippriority: [0], o_comment: [al, express pinto beans are after the careful]
o_orderkey: [1750466], o_custkey: [21433], o_orderstatus: [F], o_totalprice: [555285], o_orderdate: [1992-11-30], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000040], o_shippriority: [0], o_comment: [ ironic packages. even notornis integrate. fluffily regular foxes snooze furl]
consumer: removed 1500000 recs from the pipe
*****
n_nationkey: [7], n_name: [GERMANY], n_regionkey: [3], n_comment: [l platelets. regular accounts x-ray: unusual, regular acco]
n_nationkey: [19], n_name: [ROMANIA], n_regionkey: [3], n_comment: [ular asymptotes are about the furious multipliers. express dependencies nag above the ironically ironic account]
n_nationkey: [22], n_name: [RUSSIA], n_regionkey: [3], n_comment: [ requests against the platelets use never according to the quickly regular pint]
n_nationkey: [23], n_name: [UNITED KINGDOM], n_regionkey: [3], n_comment: [eans boost carefully special requests. accounts are. careful]
n_nationkey: [4], n_name: [EGYPT], n_regionkey: [4], n_comment: [y above the carefully unusual theodolites. final dugouts are quickly across the furiously regular d]
n_nationkey: [10], n_name: [IRAN], n_regionkey: [4], n_comment: [efully alongside of the slyly final dependencies. ]
n_nationkey: [11], n_name: [IRAQ], n_regionkey: [4], n_comment: [nic deposits boost atop the quickly final requests; quickly regula]
n_nationkey: [13], n_name: [JORDAN], n_regionkey: [4], n_comment: [lc deposits are blithely about the carefully regular pa]
n_nationkey: [20], n_name: [SAUDI ARABIA], n_regionkey: [4], n_comment: [ts. silent requests haggle. closely express packages sleep across the blithely]
consumer: removed 25 recs from the pipe
```

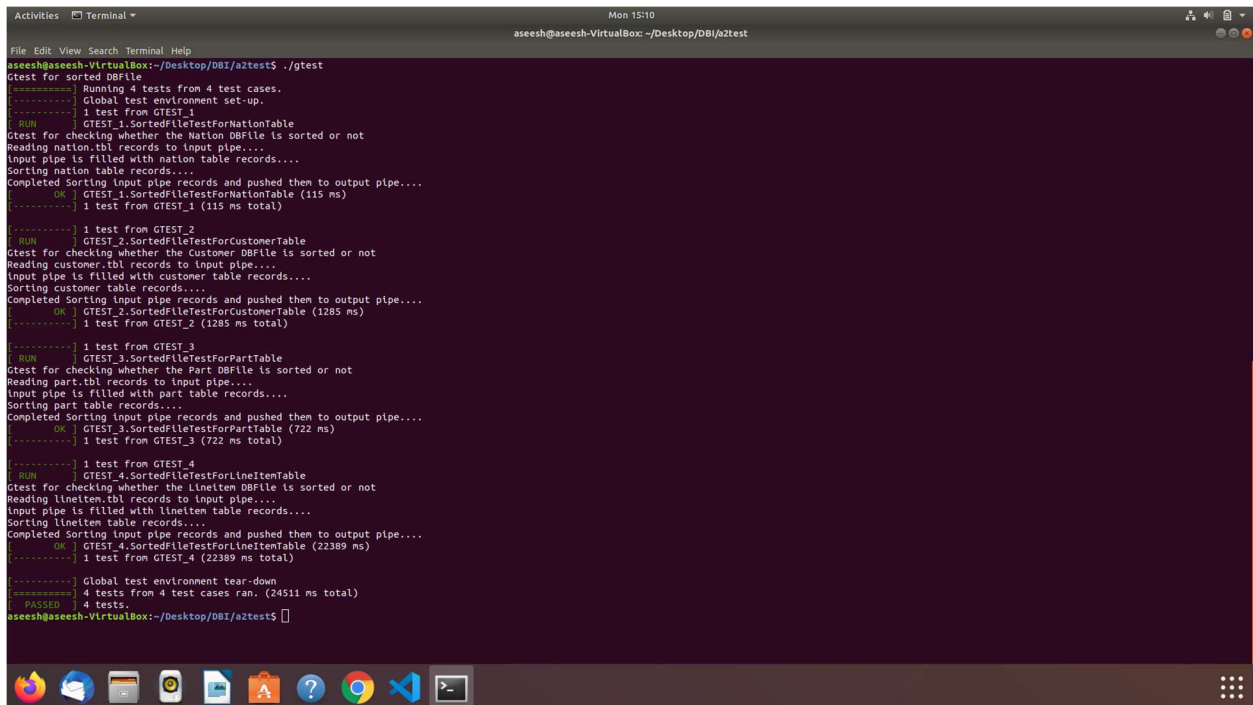
## Instructions to run gtests:

### Compile and run code gtest.cc:

Ensure the directory contains the .tbl files generated from the tpch-dbgen so that code has access to datasets

1. make gtest
2. ./gtest

The screenshot of ./gtest is present on the next page



```
aseesh@aseesh-VirtualBox:~/Desktop/DBI/aztest$ ./gtest
Gtest for sorted DBFile
[=====] Running 4 tests from 4 test cases.
[-----] Global test environment set-up.
[-----] 1 test from GTEST_1
[ RUN      ] GTEST_1.SortedFileTestForNationTable
Gtest for checking whether the Nation DBFile is sorted or not
Reading nation.tbl records to input pipe....
Input pipe is filled with nation table records....
Sorting nation table records....
Completed Sorting input pipe records and pushed them to output pipe....
[ OK      ] GTEST_1.SortedFileTestForNationTable (115 ms)
[-----] 1 test from GTEST_1 (115 ms total)

[-----] 1 test from GTEST_2
[ RUN      ] GTEST_2.SortedFileTestForCustomerTable
Gtest for checking whether the Customer DBFile is sorted or not
Reading customer.tbl records to input pipe....
Input pipe is filled with customer table records....
Sorting customer table records....
Completed Sorting input pipe records and pushed them to output pipe....
[ OK      ] GTEST_2.SortedFileTestForCustomerTable (1285 ms)
[-----] 1 test from GTEST_2 (1285 ms total)

[-----] 1 test from GTEST_3
[ RUN      ] GTEST_3.SortedFileTestForPartTable
Gtest for checking whether the Part DBFile is sorted or not
Reading part.tbl records to input pipe....
Input pipe is filled with part table records....
Sorting part table records....
Completed Sorting input pipe records and pushed them to output pipe....
[ OK      ] GTEST_3.SortedFileTestForPartTable (722 ms)
[-----] 1 test from GTEST_3 (722 ms total)

[-----] 1 test from GTEST_4
[ RUN      ] GTEST_4.SortedFileTestForLinetItemTable
Gtest for checking whether the Linetitem DBFile is sorted or not
Reading linetitem.tbl records to input pipe....
Input pipe is filled with linetitem table records....
Sorting linetitem table records....
Completed Sorting input pipe records and pushed them to output pipe....
[ OK      ] GTEST_4.SortedFileTestForLinetItemTable (22389 ms)
[-----] 1 test from GTEST_4 (22389 ms total)

[-----] Global test environment tear-down
[=====] 4 tests from 4 test cases ran. (24511 ms total)
[ PASSED  ] 4 tests.
aseesh@aseesh-VirtualBox:~/Desktop/DBI/aztest$
```