

# Database System Implementation

## Project 2B - Implementing a Sorted File

Developed by –

Ghodkari Chowdary, Raghunatha Rao - UFID: 6218-1051

Mullapudi, Aseesh - UFID: 9175-1971

Aim of this project is to extend the DBFile to implement Sorted file capabilities.

We have changed the structure of existing code, we have written a virtual base class GenericDBFile which contains the virtual functions. There DBFile now contains an internal variable myInternalVar. New classes Heap and SortedFile are written which extend the virtual base class GenericDBFile. The DBFile now creates either a heap or sortedfile object depending on the type of the file required and corresponding functions in their respective classes are invoked.

We have created new SortedFile class by writing the required functions. The job of the SortedFile within the system is to store and retrieve records over the sorted order specified.

### **Method Definitions:**

- ***SortedFile::SortedFile () , SortedFile::SortedFile (OrderMaker \*sortOrder, int runlen)***

Constructor used to create instances fields required by thre sortedFile and also to initialize the sortOrder and the runlength.

- ***int SortedFile::Create (const char \*f\_path, fType f\_type, void \*startup)***

This method takes input f\_path, f\_type and creates a sorted bin file. It also creates a metadata file which stores the required the file type and sorted Order required by the file.

- ***void SortedFile::Load (Schema &f\_schema, const char \*loadpath)***

This method takes schema, loadpath as input and loads the records from the dataset into the input pipe which prepare it for the BigQ class.

- ***int SortedFile::Open (const char \*f\_path)***

This method takes the f\_path as input and points to the sorted bin file

- ***void SortedFile::MoveFirst ()***

This method points to the first record in the SortedDB File using readBuffer

- ***void SortedFile::Add (Record &rec)***

This method is used to add the records to the sorted DBFile. A new bigq instance is created if the mode is read, else record is added to existing BigQ.

- ***int SortedFile::GetNext (Record &fetchme)***

This method returns the record from the Sorted DBFile instance using readBuffer same as the Heap DBFile.

- ***int SortedFile::GetNext (Record &fetchme, CNF &cnf, Record &literal)***

This method returns the record from the Sorted DBFile if it matches the filter CNF using binary search. This function is little complex prepares the query OrderMaker from the CNF checks compatible with sorted order of file and fetches records.

- ***int SortedFile::Close ()***

If mode is writing it changes mode to reading and call the merge function to merge the sorted file with existing records from the BigQ output pipe.

- ***void SortedFile::Merge()***

Function merges the sorted records coming from the output pipe of the BigQ with existing sortedDB File.

File **gtest.cc** written to unit test the Create, Open and Close function of the Sorted DBFile. Also checking if the bin File is sorted or not.

### **Instructions to Run the Code:**

#### **Compile and run test.cc:**

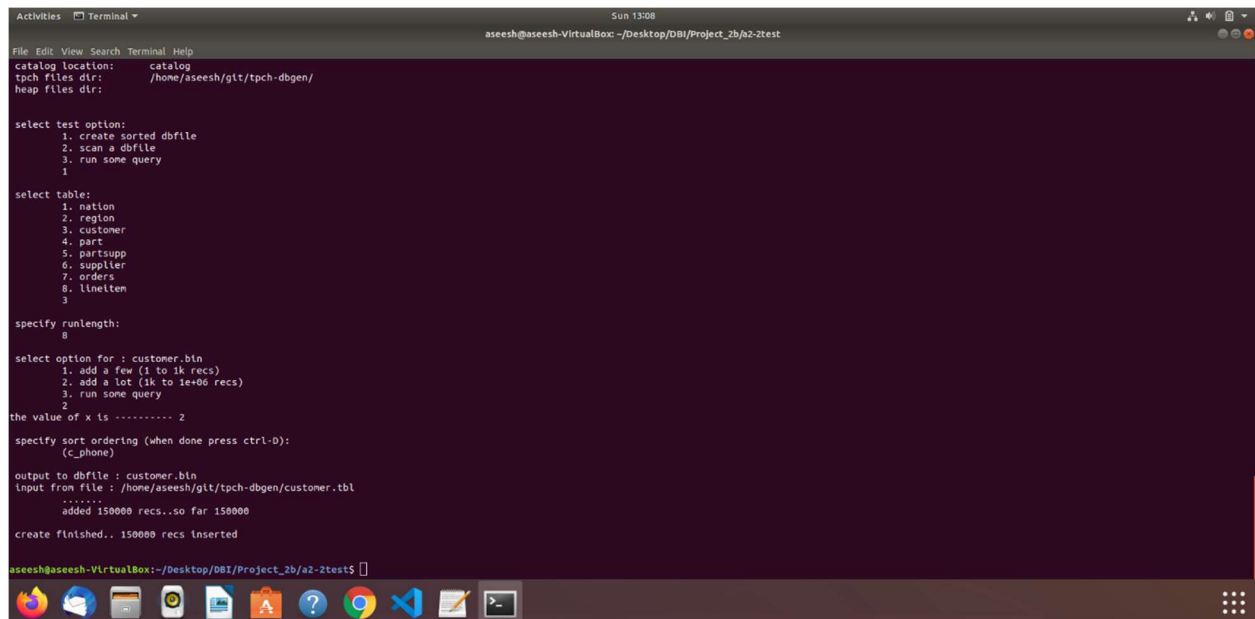
Ensure the directory contains .tbl files generated from the tpch-dbgen so that code has access to datasets. Ensure there is Catalog file in the root folder as well

1. make
2. ./test.out

### ***Output of the required Queries on 1GB tpch-dbgen files***

#### **(Q1)**

- create Sorted DBFile: 1
- table Customer: 3
- runlength: 8
- option to add records: 2
- sort Order: (c\_phone)



```
Activities Terminal Sun 13:08
aseesh@aseesh-VirtualBox: ~/Desktop/DBI/Project_2b/a2-2test

File Edit View Search Terminal Help
catalog location: catalog
tpch files dir: /home/aseesh/git/tpch-dbgen/
heap files dir:

select test option:
1. create sorted dbfile
2. scan a dbfile
3. run some query
1

select table:
1. nation
2. region
3. customer
4. part
5. partsupp
6. supplier
7. orders
8. lineitem
3

specify runlength:
8

select option for : customer.bin
1. add a few (1 to 1k recs)
2. add a lot (1k to 1e+06 recs)
3. run some query
2

the value of x is ..... 2

specify sort ordering (when done press ctrl-D):
(c_phone)

output to dbfile : customer.bin
input from file : /home/aseesh/git/tpch-dbgen/customer.tbl
.....
added 150000 recs..so far 150000

create finished.. 150000 recs inserted

aseesh@aseesh-VirtualBox:~/Desktop/DBI/Project_2b/a2-2test$
```

(Q2)

- Scan Sorted DBFile: 2
- table Customer: 3

```
Activities Terminal
Sun 13:09
aseesh@aseesh-VirtualBox: ~/Desktop/DBI/Project_2b/a2-2test

File Edit View Search Terminal Help
packages. furiously regular courts are furto]
.c_custkey: [86364], c_name: [Customer#000086364], c_address: [vWktn EzZa75U3vknT2nnp0[C9], c_nationkey: [24], c_phone: [34-997-251-4196], c_acctbal: [2691.67], c_mktsegment: [AUTOMOBILE], c_comment: [ even fox
es haggle quickly blithely regular pinto beans. slyly regular ideas sleep slyly along the f]
.c_custkey: [35286], c_name: [Customer#000035286], c_address: [P Hn6QXtMlBpJ5Xps], c_nationkey: [24], c_phone: [34-997-517-4660], c_acctbal: [120.81], c_mktsegment: [AUTOMOBILE], c_comment: [und the permanent p
ackages sleep blithely according to the special, even requests. express foxes ]
.c_custkey: [48253], c_name: [Customer#000048253], c_address: [pNusA89PvMERzgoor 510c9wK079YX1LRLKc], c_nationkey: [24], c_phone: [34-997-697-4316], c_acctbal: [4429.09], c_mktsegment: [AUTOMOBILE], c_comment:
[express forges use after the carefully ironic ideas. slyly express foxes sleep quickly, carefully special]
.c_custkey: [31537], c_name: [Customer#000031537], c_address: [82Fr5r7Gmxt2a5], c_nationkey: [24], c_phone: [34-998-190-8709], c_acctbal: [-252.7], c_mktsegment: [HOUSEHOLD], c_comment: [ructions cajole ideas. s
lyly express depoes]
.c_custkey: [317513], c_name: [Customer#00017513], c_address: [cuBZ TEEOOnPFMOodv6EIhXU43AV], c_nationkey: [24], c_phone: [34-998-315-4817], c_acctbal: [2334.66], c_mktsegment: [AUTOMOBILE], c_comment: [c ide
as are. accounts haggle agains]
.c_custkey: [53730], c_name: [Customer#000053730], c_address: [6fv7IW9Pr91qntCZzw51e2GX0xJ9A ], c_nationkey: [24], c_phone: [34-998-382-6518], c_acctbal: [4923.4], c_mktsegment: [MACHINERY], c_comment: [ silen
t requests haggle carefully regular, regular packages. enticing, even deposits wake slyly among the]
.c_custkey: [12022], c_name: [Customer#000012022], c_address: [PwVOMHrqztXWoo], c_nationkey: [24], c_phone: [34-998-408-4382], c_acctbal: [2877.36], c_mktsegment: [HOUSEHOLD], c_comment: [layers. regular pinto be
ans about the quickly regular ideas cajole blithely]
.c_custkey: [125210], c_name: [Customer#000125210], c_address: [B3L3Ag79TV5TdxgPeb63UqKA0Vv], c_nationkey: [24], c_phone: [34-998-711-2597], c_acctbal: [6678.32], c_mktsegment: [FURNITURE], c_comment: [ly ironi
c, ironic accounts. theodolites among the slyly even patterns nag slyly fluffly]
.c_custkey: [58607], c_name: [Customer#000058607], c_address: [9aGwCYLcGoy2eA], c_nationkey: [24], c_phone: [34-999-115-2195], c_acctbal: [8764.52], c_mktsegment: [FURNITURE], c_comment: [mptotes around the sly
ly bold instructions boost fur]
.c_custkey: [7288], c_name: [Customer#00007288], c_address: [dZCq3e945qXjAd7KxRHYyFVKGD4pFfAshk], c_nationkey: [24], c_phone: [34-999-121-7718], c_acctbal: [82.07], c_mktsegment: [HOUSEHOLD], c_comment: [ns
sublate furiously. Carefully ironic platelets grow furiously. ironic, bold theodolites hang above th]
.c_custkey: [89462], c_name: [Customer#000089462], c_address: [H3erpr0NUVRK], c_nationkey: [24], c_phone: [34-999-195-7029], c_acctbal: [6912.63], c_mktsegment: [BUILDING], c_comment: [totes above the pending, r
egular hockey players sleep blithely after the blithely even deposits. carefully regula]
.c_custkey: [148511], c_name: [Customer#000148511], c_address: [HoNKAELdZvMYGQWp3Z2g5NaDs7fODIH64], c_nationkey: [24], c_phone: [34-999-283-6448], c_acctbal: [24.12], c_mktsegment: [HOUSEHOLD], c_comment: [re
quests above the blithely pending asymptotes sleep carefully about the slyly even excu]
.c_custkey: [102689], c_name: [Customer#000102689], c_address: [uPFXngVSOaDQIL], c_nationkey: [24], c_phone: [34-999-363-7145], c_acctbal: [6256.97], c_mktsegment: [AUTOMOBILE], c_comment: [ly regular asymptotes
haggle fluffly according to the furiously bold courts. daring ]
.c_custkey: [3478], c_name: [Customer#00003478], c_address: [OtSzcyltLhZXVx4Y2FeToQih,lov6zyA], c_nationkey: [24], c_phone: [34-999-379-6016], c_acctbal: [6097.18], c_mktsegment: [MACHINERY], c_comment: [lt
hout the fluffly ironic packages. daring, even de]
.c_custkey: [103505], c_name: [Customer#000103505], c_address: [XPNXpdcD7HRqcVBA8pJf3Lx0c9A6X ZB2 ], c_nationkey: [24], c_phone: [34-999-404-9770], c_acctbal: [582.88], c_mktsegment: [FURNITURE], c_comment: [ful
ly even deposits boost. quickly ironic requests could have to are according to the final accounts. final]
.c_custkey: [62707], c_name: [Customer#000062707], c_address: [BRt5TIXhh], c_nationkey: [24], c_phone: [34-999-473-9448], c_acctbal: [7429.32], c_mktsegment: [FURNITURE], c_comment: [zzle slyly after the slyly l
ronic pinto beans. bold, final]
.c_custkey: [41774], c_name: [Customer#000041774], c_address: [DpHfJFKvkr4LQe1f9T5SSKZDUa4G3asKeyaSH7dz], c_nationkey: [24], c_phone: [34-999-535-1203], c_acctbal: [7118.29], c_mktsegment: [FURNITURE], c_comment
:[c requests sleep. quickly silent requests wake fluffly. furiously close packages nag furiously across the caref]
.c_custkey: [94388], c_name: [Customer#000094388], c_address: [5aE9PLelngelvgvKcGoyR,NhDT5uC4], c_nationkey: [24], c_phone: [34-999-618-6653], c_acctbal: [8636.13], c_mktsegment: [FURNITURE], c_comment: [g, eve
n deposits. carefully unus]
.c_custkey: [102414], c_name: [Customer#000102414], c_address: [Xzg9e1wbsiLwC,LUcYUACp0ENDLlUqt], c_nationkey: [24], c_phone: [34-999-618-6881], c_acctbal: [9614.62], c_mktsegment: [MACHINERY], c_comment: [usly
final packages. carefully even packages by the permanently special gifts ne]

scanned 150000 recs

aseesh@aseesh-VirtualBox:~/Desktop/DBI/Project_2b/a2-2test$
```

(Q3)

- run Query over Sorted DBFile: 3
- table Customer: 3
- CNF literal: (c\_phone > '34-999-195-7029') AND (c\_mktsegment = 'FURNITURE')

```
Activities Terminal
Sun 13:10
aseesh@aseesh-VirtualBox: ~/Desktop/DBI/Project_2b/a2-2test

File Edit View Search Terminal Help
aseesh@aseesh-VirtualBox:~/Desktop/DBI/Project_2b/a2-2test$ ./test.out

** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location: catalog
tpch files dir: /home/aseesh/glt/tpch-dbg/en/
heap files dir:

select test option:
1. create sorted dbfile
2. scan a dbfile
3. run some query
3

select table:
1. nation
2. region
3. customer
4. part
5. partsupp
6. supplier
7. orders
8. linetem
3

enter CNF predicate (when done press ctrl-D):
(c_phone > '34-999-195-7029') AND (c_mktsegment = 'FURNITURE')
.c_custkey: [103505], c_name: [Customer#000103505], c_address: [XPNXpdcD7HRqcVBA8pJf3Lx0c9A6X ZB2 ], c_nationkey: [24], c_phone: [34-999-404-9770], c_acctbal: [582.88], c_mktsegment: [FURNITURE], c_commen
t: [fully even deposits boost. quickly ironic requests could have to are according to the final accounts. final]
.c_custkey: [62707], c_name: [Customer#000062707], c_address: [BRt5TIXhh], c_nationkey: [24], c_phone: [34-999-473-9448], c_acctbal: [7429.32], c_mktsegment: [FURNITURE], c_comment: [zzle slyly after the slyly l
ronic pinto beans. bold, final]
.c_custkey: [41774], c_name: [Customer#000041774], c_address: [DpHfJFKvkr4LQe1f9T5SSKZDUa4G3asKeyaSH7dz], c_nationkey: [24], c_phone: [34-999-535-1203], c_acctbal: [7118.29], c_mktsegment: [FURNITURE], c_comment
:[c requests sleep. quickly silent requests wake fluffly. furiously close packages nag furiously across the caref]
.c_custkey: [94388], c_name: [Customer#000094388], c_address: [5aE9PLelngelvgvKcGoyR,NhDT5uC4], c_nationkey: [24], c_phone: [34-999-618-6653], c_acctbal: [8636.13], c_mktsegment: [FURNITURE], c_comment: [g, eve
n deposits. carefully unus]

query over customer.bin returned 4 recs

aseesh@aseesh-VirtualBox:~/Desktop/DBI/Project_2b/a2-2test$
```

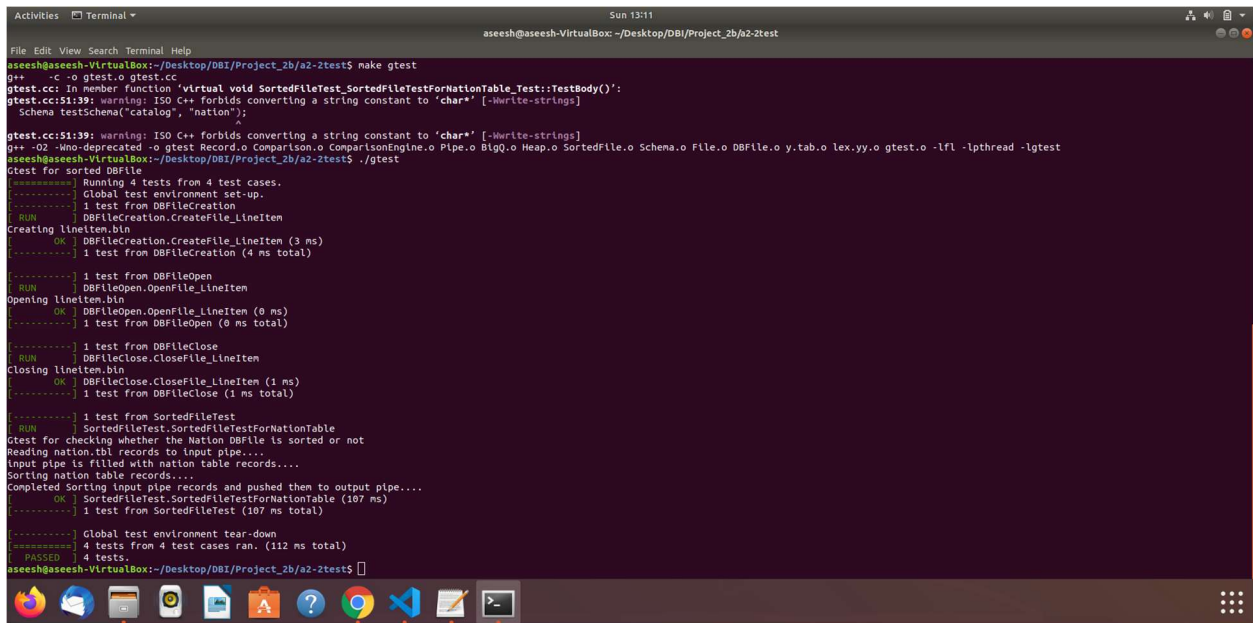
## Instructions to run gtests:

### 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/DB1/Project_2b/a2-2test$ make gtest
g++ -c -o gtest.o gtest.cc
gtest.cc: In member function 'virtual void SortedFileTest_SortedFileTestForNationTable_Test::TestBody()':
gtest.cc:51:39: warning: ISO C++ forbids converting a string constant to 'char*' [-Wwrite-strings]
   Schema testSchema("catalog", "nation");
                             ^
gtest.cc:51:39: warning: ISO C++ forbids converting a string constant to 'char*' [-Wwrite-strings]
g++ -O2 -Wno-deprecated -o gtest Record.o Comparison.o ComparisonEngine.o Pipe.o BigQ.o Heap.o SortedFile.o Schema.o File.o DBFile.o y.tab.o lex.yy.o gtest.o -lfl -lpthread -lgtest
aseesh@aseesh-VirtualBox: ~/Desktop/DB1/Project_2b/a2-2test$ ./gtest
Gtest for Sorted DBFile
[*****] Running 4 tests from 4 test cases.
[*****] Global test environment set-up.
[*****] 1 test from DBFileCreation
[ RUN ] DBFileCreation.CreateFile_LineItem
Creating linitem.bin
[ OK ] DBFileCreation.CreateFile_LineItem (3 ms)
[*****] 1 test from DBFileCreation (4 ms total)

[*****] 1 test from DBFileOpen
[ RUN ] DBFileOpen.OpenFile_LineItem
Opening linitem.bin
[ OK ] DBFileOpen.OpenFile_LineItem (0 ms)
[*****] 1 test from DBFileOpen (0 ms total)

[*****] 1 test from DBFileClose
[ RUN ] DBFileClose.CloseFile_LineItem
Closing linitem.bin
[ OK ] DBFileClose.CloseFile_LineItem (1 ms)
[*****] 1 test from DBFileClose (1 ms total)

[*****] 1 test from SortedFileTest
[ RUN ] SortedFileTest.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 ] SortedFileTest.SortedFileTestForNationTable (107 ms)
[*****] 1 test from SortedFileTest (107 ms total)

[*****] Global test environment tear-down
[*****] 4 tests from 4 test cases ran. (112 ms total)
[ PASSED ] 4 tests.
aseesh@aseesh-VirtualBox: ~/Desktop/DB1/Project_2b/a2-2test$
```