

# Database System Implementation Project Report

## Assignment 2: Implementing a Sorted File

### **Team Members:**

Divyanjali Narkuti                      UFID: 9086-0043

Yashaswini Kondakindi              UFID: 7061-1190

### **BigQ Class:**

Creation and deletion of BigQ class objects is done by constructor and destructor methods of BigQ class.

Classes SortRunRecords and MergeSortedRuns are defined.

### **SortRunRecords:**

Records of each run is sorted by SortRunRecords. Run length is specified by the user in the input and each run would have pages equal to the Run length. From the input pipe, the input records are read and then the records are sorted for reading a single run. Built-in sort function sorts the records. vectorPage is used for storing the sorted pages and vectorRecord is used for storing the sorted record. All the pages in the run are written into the output pipe, once a single run is sorted.

### **MergeSortedRuns:**

Records from the sorted runs are merged by using MergeSortedRuns. Records from the first page of each run are taken and placed in the priority queue for sorting. Phase 2 of the built-in

algorithm is used to sort all the records. Records are placed back onto the output pipe after merging all the records from all the runs.

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

First, the input run length is checked and output pipe is shut down and returned if the run length is less than or equal to zero. Therefore, no records are written. If run len is greater than 0, then the next records are fetched from the input pipe until the run length is reached. Built-in algorithm is used to sort the records in each run. Then, the sorted runs are written into a temporary file. Then the total number of runs are calculated and for each run, the first page is added to the priorityQueue. All the pushed records are merged and placed back onto the output pipe.

**Steps to compile and run the code:**

- 1) Open the project file directory path “a2project” in command prompt and type
- 2) make clean
- 3) make
- 4) make test.out
- 5) ./test.out
- 6) Check if Tpch files, catalog location, and heap files directories are correctly mentioned in the test.cc file.
- 7) Load all the required database tables and place it in the bin file location

```

1  h_nationkey: [7], n_name: [GERMANY], n_regionkey: [3], n_comment: [l platelets. regular accounts x-ray: unusual, regular
2  n_nationkey: [19], n_name: [ROMANIA], n_regionkey: [3], n_comment: [ular asymptotes are about the furious multipliers. e
3  n_nationkey: [22], n_name: [RUSSIA], n_regionkey: [3], n_comment: [ requests against the platelets use never according t
4  n_nationkey: [23], n_name: [UNITED KINGDOM], n_regionkey: [3], n_comment: [eans boost carefully special requests. accoun
5  n_nationkey: [4], n_name: [EGYPT], n_regionkey: [4], n_comment: [y above the carefully unusual theodolites. final dugout
6  n_nationkey: [10], n_name: [IRAN], n_regionkey: [4], n_comment: [efully alongside of the slyly final dependencies. l]
7  n_nationkey: [11], n_name: [IRAQ], n_regionkey: [4], n_comment: [nic deposits boost atop the quickly final requests? qui
8  n_nationkey: [13], n_name: [JORDAN], n_regionkey: [4], n_comment: [ic deposits are blithely about the carefully regular
9  n_nationkey: [20], n_name: [SAUDI ARABIA], n_regionkey: [4], n_comment: [ts. silent requests haggle. closely express pac
10 consumer: removed 25 recs from the pipe
11 *****
12 producer: opened DBFile bin/customer.bin
13 producer: inserted 150000 recs into the pipe
14 consumer: removed 150000 recs from the pipe
15 *****

```

```

consumer: removed 1500000 recs from the pipe
*****
o_orderkey: [4515876], o_custkey: [100685], o_orderstatus: [F], o_totalprice: [510062], o_orderdate: [1993-11-02], o_ord
o_orderkey: [2185667], o_custkey: [51796], o_orderstatus: [F], o_totalprice: [511360], o_orderdate: [1992-10-08], o_orde
o_orderkey: [2199712], o_custkey: [66790], o_orderstatus: [0], o_totalprice: [515532], o_orderdate: [1996-09-30], o_orde
o_orderkey: [3586919], o_custkey: [24049], o_orderstatus: [F], o_totalprice: [522644], o_orderdate: [1992-11-07], o_orde
o_orderkey: [2232932], o_custkey: [13940], o_orderstatus: [0], o_totalprice: [522721], o_orderdate: [1997-04-13], o_orde
o_orderkey: [4576548], o_custkey: [108931], o_orderstatus: [0], o_totalprice: [525591], o_orderdate: [1997-12-26], o_ord
o_orderkey: [3043270], o_custkey: [144617], o_orderstatus: [0], o_totalprice: [530604], o_orderdate: [1997-02-12], o_ord
o_orderkey: [4722021], o_custkey: [128120], o_orderstatus: [F], o_totalprice: [544089], o_orderdate: [1994-04-07], o_ord
o_orderkey: [1750466], o_custkey: [21433], o_orderstatus: [F], o_totalprice: [555285], o_orderdate: [1992-11-30], o_orde
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
n_nationkey: [19], n_name: [ROMANIA], n_regionkey: [3], n_comment: [ular asymptotes are about the furious multipliers. e
n_nationkey: [22], n_name: [RUSSIA], n_regionkey: [3], n_comment: [ requests against the platelets use never according t
n_nationkey: [23], n_name: [UNITED KINGDOM], n_regionkey: [3], n_comment: [eans boost carefully special requests. accoun
n_nationkey: [4], n_name: [EGYPT], n_regionkey: [4], n_comment: [y above the carefully unusual theodolites. final dugout
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? qui
n_nationkey: [13], n_name: [JORDAN], n_regionkey: [4], n_comment: [ic deposits are blithely about the carefully regular
n_nationkey: [20], n_name: [SAUDI ARABIA], n_regionkey: [4], n_comment: [ts. silent requests haggle. closely express pac
consumer: removed 25 recs from the pipe

```

## Gtests for BigQ Class Methods:

Gtests.cc class is added which has the google tests to run the tests on the BigQ Class Methods **Steps to compile and run the code:**

- 1) make Gtests.out
- 2) ./Gtests.out
- 3) Select the table name for which gtests needs to be run
- 4) Enter the run length and the sort ordering CNF value

## Results:

## Output for Lineitem table:

```
ubuntu@primary:~/Home/desktop/DivyanjaliNarkuti_YashaswiniKondakindi/a2project$ ./Gtests.out
```

```
/** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
```

```
catalog location:      catalog
tpch files dir:        tpch-dbgen/
heap files dir:        bin/
```

```
SELECT table Name:
```

```
1)Nation
2)Region
3)Customer
4)Part
5)Partsupp
6)Orders
7)Lineitem
7
```

```
Enter the Run Length:
```

```
2
```

```
[=====] Running 7 tests from 2 test suites.
[-----] Global test environment set-up.
[-----] 4 tests from BigQClassTest
[ RUN    ] BigQClassTest.TestRunLength
[ OK     ] BigQClassTest.TestRunLength (0 ms)
[ RUN    ] BigQClassTest.CheckNumberOfRecords
[ OK     ] BigQClassTest.CheckNumberOfRecords (0 ms)
[ RUN    ] BigQClassTest.InsertAndRemoveFromPipe
[ OK     ] BigQClassTest.InsertAndRemoveFromPipe (26 ms)
[ RUN    ] BigQClassTest.CheckIfBinFileEmpty
[ OK     ] BigQClassTest.CheckIfBinFileEmpty (10 ms)
[-----] 4 tests from BigQClassTest (36 ms total)

[-----] 3 tests from gtest
[ RUN    ] gtest.testDBFileCreate
[ OK     ] gtest.testDBFileCreate (7 ms)
[ RUN    ] gtest.testDBFileOpen
[ OK     ] gtest.testDBFileOpen (5 ms)
[ RUN    ] gtest.testDBFileClose
[ OK     ] gtest.testDBFileClose (3 ms)
[-----] 3 tests from gtest (15 ms total)

[-----] Global test environment tear-down
[=====] 7 tests from 2 test suites ran. (53 ms total)
[ PASSED ] 7 tests.
```

## Output for Region table:

```
ubuntu@primary:~/Home/desktop/DivyanjaliNarkuti_YashaswiniKondakindi/a2project$ ./Gtests.out
```

```
/** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
```

```
catalog location:      catalog
tpch files dir:        tpch-dbgen/
heap files dir:        bin/
```

```
SELECT table Name:
```

```
1)Nation
2)Region
3)Customer
4)Part
5)Partsupp
6)Orders
7)Lineitem
2
```

```
Enter the Run Length:
```

```
2
```

```
[=====] Running 7 tests from 2 test suites.
[-----] Global test environment set-up.
[-----] 4 tests from BigQClassTest
[ RUN    ] BigQClassTest.TestRunLength
[ RUN     OK ] BigQClassTest.TestRunLength (0 ms)
[ RUN     OK ] BigQClassTest.CheckNumberOfRecords
[ RUN     OK ] BigQClassTest.CheckNumberOfRecords (0 ms)
[ RUN     OK ] BigQClassTest.InsertAndRemoveFromPipe
[ RUN     OK ] BigQClassTest.InsertAndRemoveFromPipe (21 ms)
[ RUN     OK ] BigQClassTest.CheckIfBinFileEmpty
[ RUN     OK ] BigQClassTest.CheckIfBinFileEmpty (11 ms)
[-----] 4 tests from BigQClassTest (34 ms total)

[-----] 3 tests from gtest
[ RUN     OK ] gtest.testDBFileCreate
[ RUN     OK ] gtest.testDBFileCreate (4 ms)
[ RUN     OK ] gtest.testDBFileOpen
[ RUN     OK ] gtest.testDBFileOpen (5 ms)
[ RUN     OK ] gtest.testDBFileClose
[ RUN     OK ] gtest.testDBFileClose (5 ms)
[-----] 3 tests from gtest (14 ms total)

[-----] Global test environment tear-down
[=====] 7 tests from 2 test suites ran. (48 ms total)
[ PASSED ] 7 tests.
ubuntu@primary:~/Home/desktop/DivyanjaliNarkuti_YashaswiniKondakindi/a2project$
```

**Output for Nation table:**



```

ubuntu@primary:~/Home/desktop/DivyanjaliNarkuti_YashaswiniKondakindi/a2project$ ./Gtests.out

** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location:      catalog
tpch files dir:        tpch-dbggen/
heap files dir:        bin/

SELECT table Name:
1)Nation
2)Region
3)Customer
4)Part
5)Partsupp
6)Orders
7)Lineitem
1

Enter the Run Length:
2
[=====] Running 7 tests from 2 test suites.
[-----] Global test environment set-up.
[-----] 4 tests from BigQClassTest
[ RUN    ] BigQClassTest.TestRunLength
[ OK     ] BigQClassTest.TestRunLength (0 ms)
[ RUN    ] BigQClassTest.CheckNumberOfRecords
[ OK     ] BigQClassTest.CheckNumberOfRecords (0 ms)
[ RUN    ] BigQClassTest.InsertAndRemoveFromPipe
[ OK     ] BigQClassTest.InsertAndRemoveFromPipe (19 ms)
[ RUN    ] BigQClassTest.CheckIfBinFileEmpty
[ OK     ] BigQClassTest.CheckIfBinFileEmpty (10 ms)
[-----] 4 tests from BigQClassTest (29 ms total)

[-----] 3 tests from gtest
[ RUN    ] gtest.testDBFileCreate
[ OK     ] gtest.testDBFileCreate (4 ms)
[ RUN    ] gtest.testDBFileOpen
[ OK     ] gtest.testDBFileOpen (3 ms)
[ RUN    ] gtest.testDBFileClose
[ OK     ] gtest.testDBFileClose (4 ms)
[-----] 3 tests from gtest (13 ms total)

[-----] Global test environment tear-down
[=====] 7 tests from 2 test suites ran. (44 ms total)
[ PASSED ] 7 tests.
ubuntu@primary:~/Home/desktop/DivyanjaliNarkuti_YashaswiniKondakindi/a2project$

```

## Methods in Gtests.cc Class:

**TEST (BigQTest, TestRunLength):**

**EXPECT EQ(outputLength,0);**

In this gtest test run length, the input run length is given as less than zero(-20). Therefore, the output is not written into the output pipe and the length of the output pipe would be zero. Therefore, nothing will be written into the output pipe as input length will not be equal to output length.

**TEST (BigQTest, CheckSorting):**

**EXPECT EQ(0,val);**

In this gtest check sorting, as input we pass the input pipe, output pipe, sorting order CNF, run length and 2 threads are created and we check if the given records are placed on the output pipe in the sorted order.

**TEST (BigQTest, CheckNumberOfRecords):**

**EXPECT EQ(inputLength,outputLength);**

In this gtest-check number of records, we check if the input number of records is equal to output records after successfully sorting of the pages in the runs and merging the runs to form the sorted output.

**TEST(BigQClassTest, InsertAndRemoveFromPipe):**

**EXPECT TRUE(hasRecords == 1);**



Insertion and removal from the pipe is being checked in this testcase. Read the given file path and then records are read from the file into the input pipe and then data is removed from the input pipe, and we are checking if the output has the records.

**TEST( BigQClassTest, CheckIfBinFileEmpty):**

**EXPECT TRUE(isNotEmpty == 1);**

In this test case, we are checking if the bin file is empty. If the bin file is not empty, then 1 is returned and the expected statement is satisfied.

**TEST (gtest, testDBFileCreate) :**

**EXPECT EQ(1,newFile.Create (rel->path(), heap, NULL));** In this gtest create, creation of new file is tested and 1 is returned on successful creation using create method of DBFile class.

**TEST (gtest, testDBFileOpen) :**

**EXPECT EQ(1,newFile.Open (rel->path()));**

In this gtest open, Using the above statement a created file is opened. This will returns 1 if success and 0 if failure.

**TEST (gtest, testDBFileClose) :**

**EXPECT EQ(1,newFile.Close ());** In this gtest close, closing an opened file of DBFile instance is tested using the above statement. Returns 1 if success and 0 on failure.