

COP 6726:DataBase System Implementation

Project 2(Part 2):Extended DBFile Class

Venkateswarlu Tanneru
vtanneru@ufl.edu

Rohit Yerramsetty
rohityerramsetty@ufl.edu

Wednesday 15th March, 2022

1 Introduction:

The DBFile Class is now extended into different class files like SortedDB file and HeapDB, Where the functions and functionalities differ in both the DB Class files with Heap Implementation of the database values and also Sorted database on the given .bin files.

2 Files Location:

- **MakeFile:** Make File runs the file structure in making the interpreted files in project.
- **tpch/:** Contains .tbl files which were generated using tpch data generator.
- **gtest/:** Google Test Unit testing library.
- **catalog:** Catalog file for the schema of .tbl files
- **test/:** Contains .bin files.

3 Commands To Run the Program:

- **make clean**
- **make**
- **make test.out**
- **./test.out**
- Above commands give the test cases results by selecting the database files and prompt query from the database.
- Fill in any CNF from the result section below for the scan and filter operations. After you've finished creating the CNF, press ENTER and then Ctrl+D.
- The unit tests should run and the test results should be shown.

4 DBFiles for Heap and Sorted:

4.1 Generic DBFile:

This improved version of DBFile is designed to do both duties. As a result, it can handle both sorted and unsorted heap files. Generic DBFile Class is a virtual class that acts as the parent class for both heap and sorted DBFile classes. Both of the child classes separately implement the virtual methods in this class. Because these virtual methods are not shared by all of the children's classes, their implementation differs in each of them. This class contains protected variables for the class objects File, Page, and Comparison Engine, as well as a preference file pointer. Page class objects hold the items that will be written to the disk, whereas file class objects contain the pages.

4.2 Heap DBFile:

This class implements the DBFile unsorted Heap feature of the project. This class contains all the methods required to implement unsorted heaps. Aside from that, there is a constructor and a destructor in this class. The MoveFirst (), Add (), Load (), GenNext (), and Close () virtual methods are implemented individually by this class.

4.3 Sorted DBFile

This Class contains all of the methods for implementing the new sorted heap file capability. Input and output pipes for pushing and retrieving entries from the BigQ class are among the characteristics of the SortedDBFile Class. This class also includes a BigQ class pointer and a File object that contains the pages. This class also contains a buffer property, which is used before writing the pages into the file object. A queryorder maker element is used to generate a DBFile sort order as well as a sort order using a given CNF.

5 Functions in SortedDB Class File:

5.1 void Load(Schema &myschema, char *loadpath):

The load function can bulk add entries to the input stream. The load function and the add function both perform the same duties; the main difference is that the load method contains a loop to add data in bulk, whereas the add function can only add one item at a time. The loop reads items from the raw.tbl file and runs until the load function is not provided with any records.

5.2 void MoveFirst():

The MoveFirst () method in DBFile is implemented in the same way that it is in heap files. This heap file function resets the reference to the file's first entry. Each DBfile contains a reference to the first entry in the file. In response to file record retrievals, this pointer may move. If the page buffer is full and the DBFile is in write mode. The non-empty page buffer is saved into dbfile before the first page is read and the first record from the page is received. In other cases, the cursor is positioned on the first page and the content is read.

5.3 void Add(Record &addme):

The input pipe attribute in the class allows records to be written into BigQ class objects. Use the Add function to insert records into the input stream. If the buffer is in read mode, it is switched to write mode before the records are written to the pipe. When the buffer is set to write mode, sorted records from the BigQ class objects' output stream are written into the input pipe.

5.4 int GetNext (Record &fetchme):

This strategy is employed. GetNext() iteratively retrieves records from the dbfiles. The permanent dbfile (.bin file) is separated into pages, each of which contains records. The GetNext() function reads the pages of the persistent file one by one, obtaining all of the records contained in these pages numerous times. When all of the records in the file have been read, this method returns 0.

5.5 int GetNext(Record &fetchme, CNF &cnf, Record &literal):

This version of the GetNext method takes three parameters. The CNF predicate is one of the parameters. This supplied CNF is turned to an OrderMaker object using a CNF class method such as GetSortOrders. The supplied CNF is then compared to the current sorting order of dbfile. If the ordermakers' characteristics match the DBfile's sort order, the BinarySearch function is used to locate any entries that fulfill that CNF. With each consecutive execution of this procedure, we get the same results.

5.6 int create(char *fpath, fType f_type, void *startup):

This Function Creates the output based on the database .tbl files and take these inputs to work over sorting ,Heap and functioning on database files.

5.7 int Open(char *fpath):

This Open Function gets the file path as a input which traverses through the files and opens the bin filw which contains the database.

5.8 int Close():

This Function Close () just closes the database file. when the application has finished writing the dbfile's state to the a.pref file. The order maker object is also saved in the form of a sorted DB File. It is done so that the file's sort order may be detected when it is read.

6 Results:

The Results are from different Test cases where we executed the following

- Test Case - 1 Created a sorted DB file based on Order database
- Test Case - 2 Scanned the sorted Order database
- Test Case - 3 Query Implementation to check output accuracy

6.1 Test Case 1

```
venkytanneru@ubuntu:/mnt/hgfs/D8SI/Projects/Project 2 part 2/a2-2test$ ./test.out
** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location:      catalog
tpch files dir:        tpch/10M/
heap files dir:        test/

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
7

specify runlength:
8
input from file : tpch/10M/orders.tbl

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

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

output to dbfile : test/orders.bin
meta data in test/orders.bin.meta
begin c Create
end Create
NumAtts =      1
0:      0 Int
begin writeMode
end writeMode
Work_func Starts
.....1
meta data in tempDiffFile.bin.meta
end Work_func
meta data in tempMergedFile.bin.meta
.....
added 235753 recs..so far 235753
```

Figure 1: Output Test case 1

6.2 Test Case 2

```
.o_orderkey: [41895], o_cuskey: [98973], o_orderstatus: [F], o_totalprice: [181581], o_orderdate: [1995-01-11], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000303], o_shippriority: [0], o_comment: [Ie quickly against the quickly ironic plants beans! quickly fin]
.o_orderkey: [41898], o_cuskey: [28632], o_orderstatus: [0], o_totalprice: [114834], o_orderdate: [1998-07-16], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000286], o_shippriority: [0], o_comment: [I have to wake blithely. furiously final foves alongside of t he fluffily final]
.o_orderkey: [41911], o_cuskey: [106838], o_orderstatus: [F], o_totalprice: [128134], o_orderdate: [1992-02-20], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000386], o_shippriority: [0], o_comment: [ts. platelets sleep. sl]
.o_orderkey: [41925], o_cuskey: [133998], o_orderstatus: [0], o_totalprice: [182464], o_orderdate: [1993-10-01], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000234], o_shippriority: [0], o_comment: [otes. ironic deposits wake regular deposits. blithely fin al packages ar]
.o_orderkey: [41923], o_cuskey: [96188], o_orderstatus: [0], o_totalprice: [236188], o_orderdate: [1996-07-22], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000751], o_shippriority: [0], o_comment: [Iola quickly about the ironic package]
.o_orderkey: [41924], o_cuskey: [60487], o_orderstatus: [0], o_totalprice: [62055.4], o_orderdate: [1995-05-26], o_orderpriority: [5-LOW], o_clerk: [Clerk#000000449], o_shippriority: [0], o_comment: [ly furiously final request]
.o_orderkey: [41925], o_cuskey: [133998], o_orderstatus: [0], o_totalprice: [17295.5], o_orderdate: [1993-02-24], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000293], o_shippriority: [0], o_comment: [ deposits. furiously special pac]
.o_orderkey: [41926], o_cuskey: [6618], o_orderstatus: [F], o_totalprice: [189864], o_orderdate: [1995-02-06], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000672], o_shippriority: [0], o_comment: [I nap carefully after the furiously express asymptotes. furiou sl]
.o_orderkey: [41927], o_cuskey: [23894], o_orderstatus: [F], o_totalprice: [39718.1], o_orderdate: [1992-07-14], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000790], o_shippriority: [0], o_comment: [ages. silyly express accounts are careful]
.o_orderkey: [41932], o_cuskey: [136864], o_orderstatus: [F], o_totalprice: [203298], o_orderdate: [1993-07-30], o_orderpriority: [5-LOW], o_clerk: [Clerk#000000228], o_shippriority: [0], o_comment: [lithely silyly even theodolites. quickly ironic asymptotes ha ]
.o_orderkey: [41933], o_cuskey: [71053], o_orderstatus: [0], o_totalprice: [43163.8], o_orderdate: [1997-09-06], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000068], o_shippriority: [0], o_comment: [e fluffily regular instructions haggle about the q uiet]
.o_orderkey: [41954], o_cuskey: [28684], o_orderstatus: [0], o_totalprice: [120643], o_orderdate: [1996-05-06], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000623], o_shippriority: [0], o_comment: [ally. silest, regular packages boost silyly final]
.o_orderkey: [41955], o_cuskey: [11977], o_orderstatus: [F], o_totalprice: [336349], o_orderdate: [1995-04-02], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000744], o_shippriority: [0], o_comment: [lilers use silyly far]
.o_orderkey: [41956], o_cuskey: [133888], o_orderstatus: [0], o_totalprice: [144728], o_orderdate: [1995-12-13], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000778], o_shippriority: [0], o_comment: [tes doubt furiously a]
.o_orderkey: [41957], o_cuskey: [87739], o_orderstatus: [0], o_totalprice: [348341], o_orderdate: [1997-12-28], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000995], o_shippriority: [0], o_comment: [otes cajole furiously! even courts are busily. carefully e ven]
.o_orderkey: [41958], o_cuskey: [216612], o_orderstatus: [0], o_totalprice: [382795], o_orderdate: [1997-05-25], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000089], o_shippriority: [0], o_comment: [st carefully regular pinto beans. doggely even id nap use fur]
.o_orderkey: [41959], o_cuskey: [47838], o_orderstatus: [F], o_totalprice: [93888.8], o_orderdate: [1993-04-29], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000761], o_shippriority: [0], o_comment: [ickly pending theodolites. quickly fluffy a]
.o_orderkey: [41964], o_cuskey: [28126], o_orderstatus: [0], o_totalprice: [176342], o_orderdate: [1997-01-31], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000372], o_shippriority: [0], o_comment: [I hinder around the regular packages. carefully stealthy pack ages wake regulat]
.o_orderkey: [41983], o_cuskey: [7181], o_orderstatus: [0], o_totalprice: [244606], o_orderdate: [1996-07-12], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000854], o_shippriority: [0], o_comment: [lilyly pending theodolites. repu]
.o_orderkey: [41986], o_cuskey: [82490], o_orderstatus: [F], o_totalprice: [118893], o_orderdate: [1993-02-14], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000626], o_shippriority: [0], o_comment: [he carefully pending d]
.o_orderkey: [41987], o_cuskey: [58371], o_orderstatus: [0], o_totalprice: [173876], o_orderdate: [1995-10-22], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000883], o_shippriority: [0], o_comment: [uriously special excuses run alon]
.o_orderkey: [42016], o_cuskey: [119302], o_orderstatus: [F], o_totalprice: [195117], o_orderdate: [1993-07-19], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000996], o_shippriority: [0], o_comment: [tallers are furiously pending deposits. unusual, quick foves a]
.o_orderkey: [41989], o_cuskey: [15711], o_orderstatus: [F], o_totalprice: [77381.7], o_orderdate: [1993-04-10], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000991], o_shippriority: [0], o_comment: [Instructions x-ray furiously alongside of ]
.o_orderkey: [41990], o_cuskey: [29473], o_orderstatus: [0], o_totalprice: [243131], o_orderdate: [1997-06-08], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000333], o_shippriority: [0], o_comment: [tegrate among the quickly regular plat]
.o_orderkey: [41991], o_cuskey: [35948], o_orderstatus: [0], o_totalprice: [218986], o_orderdate: [1995-12-19], o_orderpriority: [5-LOW], o_clerk: [Clerk#000000637], o_shippriority: [0], o_comment: [I deposits cajole quickly final]
.o_orderkey: [42016], o_cuskey: [119302], o_orderstatus: [F], o_totalprice: [195117], o_orderdate: [1993-07-19], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000996], o_shippriority: [0], o_comment: [eans wake blithely above the carefully express braids. re gular theodo]
.o_orderkey: [42017], o_cuskey: [49318], o_orderstatus: [F], o_totalprice: [128207], o_orderdate: [1995-06-01], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000428], o_shippriority: [0], o_comment: [requests cajole even, pending pinto beans. doggelyly unusal]
.o_orderkey: [42018], o_cuskey: [136661], o_orderstatus: [0], o_totalprice: [139317], o_orderdate: [1996-09-25], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000674], o_shippriority: [0], o_comment: [ges are silyly final, quiet accounts. furiously final acco a]
.o_orderkey: [42019], o_cuskey: [44822], o_orderstatus: [0], o_totalprice: [48884], o_orderdate: [1997-09-29], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000190], o_shippriority: [0], o_comment: [uffy foves impress fluffily. furiously regular depende]
.o_orderkey: [42038], o_cuskey: [124838], o_orderstatus: [0], o_totalprice: [54388.7], o_orderdate: [1995-03-21], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000761], o_shippriority: [0], o_comment: [counts are requests. platelets boost blithely furiously regular pinto beans]
.o_orderkey: [42021], o_cuskey: [88138], o_orderstatus: [0], o_totalprice: [238468], o_orderdate: [1996-06-14], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000387], o_shippriority: [0], o_comment: [ect evenly furiously bold packages. final epitaphs are thinly. regular a]
.o_orderkey: [42022], o_cuskey: [97938], o_orderstatus: [F], o_totalprice: [65855.2], o_orderdate: [1994-09-25], o_orderpriority: [3-MEDIUM], o_clerk: [Clerk#000000141], o_shippriority: [0], o_comment: [ic packages hinder fur]
.o_orderkey: [42023], o_cuskey: [71807], o_orderstatus: [F], o_totalprice: [285188], o_orderdate: [1992-09-21], o_orderpriority: [2-HIGH], o_clerk: [Clerk#000000000], o_shippriority: [0], o_comment: [illy regular packages can wake expor]
.o_orderkey: [42048], o_cuskey: [102403], o_orderstatus: [0], o_totalprice: [282551], o_orderdate: [1995-10-18], o_orderpriority: [5-LOW], o_clerk: [Clerk#000000439], o_shippriority: [0], o_comment: [ blithely regular theodolites are blithely. requests detect blithely pend]
.o_orderkey: [42049], o_cuskey: [38272], o_orderstatus: [0], o_totalprice: [77973.2], o_orderdate: [1997-08-18], o_orderpriority: [4-NOT SPECIFIED], o_clerk: [Clerk#000000694], o_shippriority: [0], o_comment: [sly special instructions. carefully even ]
.o_orderkey: [42050], o_cuskey: [134558], o_orderstatus: [F], o_totalprice: [47873.4], o_orderdate: [1995-03-31], o_orderpriority: [5-LOW], o_clerk: [Clerk#000000029], o_shippriority: [0], o_comment: [lar dependencies wake life pinto beans. quickly bold packag
```

Figure 2: Output Test case 2

6.3 Test Case 3

```
root@ubuntu:/mnt/hgfs/DBS/Projects/Project 2 part 2/a2-2test$ ./test.out

** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location:      catalog
tpch files dir:        tpch/10M/
heap files dir:        test/

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

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

enter O/E predicate (when done press ctrl-D):
(o_orderpriority = '1-URGENT') AND (o_orderstatus = 'F') AND (o_totalprice > 470000.0)
NumWatts = 1
0: 0 Int
( Att 5 from left record = Att 8 from literal record (String)) AND
( Att 2 from left record = Att 1 from literal record (String)) AND
( Att 3 from left record > Att 2 from literal record (Double))
original lower bound: 0
updated lower bound by Binary Search: 284
updated higher bound by Binary Search: 284
.o_orderkey: [101708], o_cuskey: [147288], o_orderstatus: [F], o_totalprice: [472441], o_orderdate: [1994-05-18], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000621], o_shippriority: [0], o_comment: [ express Tiresias unwind s]
.o_orderkey: [809158], o_cuskey: [63994], o_orderstatus: [F], o_totalprice: [474183], o_orderdate: [1994-12-17], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000693], o_shippriority: [0], o_comment: [y even accounts use among the pending, even pinto ]
.o_orderkey: [869154], o_cuskey: [125666], o_orderstatus: [F], o_totalprice: [489837], o_orderdate: [1994-06-12], o_orderpriority: [1-URGENT], o_clerk: [Clerk#000000690], o_shippriority: [0], o_comment: [ are silyly. furiously ironic ]

query over test/orders.bin returned 3 rows
```

Figure 3: Output Test case 3

7 Gtest:

To begin running the gtest, go to the project's root directory, which includes the gtest.cpp file.

7.1 Commands To Run the Gtest Program:

- Type **make clean** and then **make gtest** to compile Google Test.
- Type **./gtest.out** to run the unit tests.
- The unit tests should run and the gtest results should be shown.

7.2 Gtest results:

we ran three tests on gtests. The testing is done on paper records. Create is being tested after successfully comparing records. On successfully comparing runs testing non-class method records, Open and Close. Check to see if the records are successfully placed in the Created file. Below is a snapshot of the gtest.

```
venkytanneru@ubuntu:/mnt/hgfs/DBSI/Projects/Project 2 part 2/a2-2test$ ./gtest.out
[=====] Running 3 tests from 1 test suite.
[-----] Global test environment set-up.
[-----] 3 tests from DBFileTest
[ RUN    ] DBFileTest.FileCreation
meta data in ./gtest.bin.meta
begin c Create
end Create
==1
[      OK ] DBFileTest.FileCreation (8 ms)
[ RUN    ] DBFileTest.FileClose
NumAtts =    1
  0:    4 String
[      OK ] DBFileTest.FileClose (15 ms)
[ RUN    ] DBFileTest.FileCreateCloseTest
meta data in ./gtest.bin.meta
begin c Create
end Create
[      OK ] DBFileTest.FileCreateCloseTest (11 ms)
[-----] 3 tests from DBFileTest (34 ms total)

[-----] Global test environment tear-down
[=====] 3 tests from 1 test suite ran. (35 ms total)
[ PASSED ] 3 tests.
```

Figure 4: Gtest Results