# Project Report - Implementing a Heap File

Arpan Banerjee UFID: 9359-9083 arpanbanerjee@ufl.edu

Krutantak Patil UFID: 5615-6343 Krutantakb.patil@ufl.edu

# 1) Steps to compile and run - code, tests and gtests

- i. **make main** Command to build the main program.
- ii. ./main Command to run the main executable.
- iii. make test.out Command to build the tests.
- iv. ./test.out Command to run the tests.
- v. **make DBFileTest.out** Command to build the gtests.
- vi. ./ DBFileTest.out Command to run the gtests.

# 2) DBFile implementation

#### **Data members**

- i. char \*filename Used to copy filename in Create and Open functions.
- ii. File \*heapfile Used to store pointer to the heapfile object used to store records.
- iii. Page \*deltapage Pointer to page holding recent record additions to the file.
- iv. Page \*readpage Pointer to page for reading records from the file.
- v. ComparisonEngine comp Used to apply and check CNF with a record
- vi. int readPageNumber Stores page number of current page being read.

#### Member functions

- i. void MoveFirst () This function resets the *readPageNumber* variable and empties the *readPage*.
- ii. [private function] void AddDeltaPageToFile() This function adds the *deltapage* to the *heapfile* and clears it out for further additions.
- iii. void Add (Record &record) This function adds a record to the *deltapage*. If the *deltapage* is full, it calls *AddDeltaPageToFile()* and adds the record to the fresh *deltapage*.
- iv. int GetNext (Record &fetchme) ) This function gets the next record from the file by reading from the *readpage* or incrementing the *readPage* if it runs out. If we reach the end of the file and *deltaPage* contains records, then it is added to the file and read.
- v. int GetNext (Record &fetchme, CNF &cnf, Record &literal)) This function repeatedly calls *GetNext* and checks the record against the CNF.
- vi. int Create (const\_char \* f\_path, fType f\_Type, void \*startup) ) This function creates a DBFile of the type heap. Returns 1 on success.

- vii. int Open (const char \*f\_path) ) This function opens an existing DBFile.

  Returns 1 on success and 0 on failure or when File not found.
- viii. int Close ()) This function calls AddDeltaPageToFile() and then closes the DBFile.
- ix. void Load (Schema &file\_schema, const char \*recPath) ) This function loads data from a given file path. It repeatedly calls *SuckNextRecord()* and then *Add()* to add it to the DBFile.

# 3) Test.out results for 1GB data

. **Query 1 :** (r\_name = 'EUROPE')

```
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-projl$_./test.out
* IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location: catalog
                      datagen/tpch-dbgen/
tpch files dir:
heap files dir:
select test:es
       1. load file
        3. scan & filter
select table:
        5. partsupp
        6. orders
Filter with CNF for : region
Enter CNF predicate (when done press ctrl-D):
      (r name = 'EUROPE')r regionkey: [3], r name: [EUROPE], r comment: [ly final courts cajole furiously final excuse]
selected 1 recs
rpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$
```

## ii. Query 2: (r\_name < 'middle east') AND (r\_regionkey > 1)

```
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-projl$ ./test.out
 * IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
 catalog location: catalog
 tpch files dir:
                     datagen/tpch-dbgen/
                     dbfiles/ 44 The expected output for these CNF's can be found in the file "output lon"
 heap files dir:
 select test:esto
       1. load file
       2. scan
       3. scan & filter
select table:
       1. nation
       2. region
       3. customer
       4. part
       5. partsupp
Filter with CNF for : region
Enter CNF predicate (when done press ctrl-D):
      (r name < 'middle east') AND
 r_regionkey > 1)r_regionkey: [2], r_name: [ASIA], r_comment: [ges. thinly even pinto beans ca]
 regionkey: [3], r name: [EUROPE], r comment: [ly final courts cajole furiously final excuse]
 regionkey: [4], r name: [MIDDLE EAST], r comment: [uickly special accounts cajole carefully blithely close requests. carefully final asymptotes haggle furiousl]
 selected 3 recs
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-projl$
```

#### iii. Query 3:

```
(n_regionkey = 3) AND
(n_nationkey > 10) AND
(n_name > 'japan')
```

### iv. Query 11:

```
(I_shipdate > '1994-01-01') AND
(I_shipdate < '1994-01-07') AND
(I_discount > 0.05) AND
(I_discount < 0.06) AND
(I_quantity = 4)
```

```
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-projl$ ./test.out
** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location: catalog
tpch files dir: c datagen/tpch-dbgen/ (o totalprice < 105 dbfiles/
 select test:est.cc
        1. load file
        2. scan
        3. scan & filter
select table:
        1. nation
      2. region
        customer
        4. part
        partsupp
        6. orders
        7. lineitem
 Filter with CNF for : lineitem
Enter CNF predicate (when done press ctrl-D):

(1 shindate > '1994-01-01') AND
        (lmshipdate > '1994-01-01') AND
(l shipdate < '1994-01-07') AND
(l discount > 0.05) AND
(l discount < 0.06) AND
(l quantity = 4.0) selected 0 recs
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$
```

#### v. Query 12:

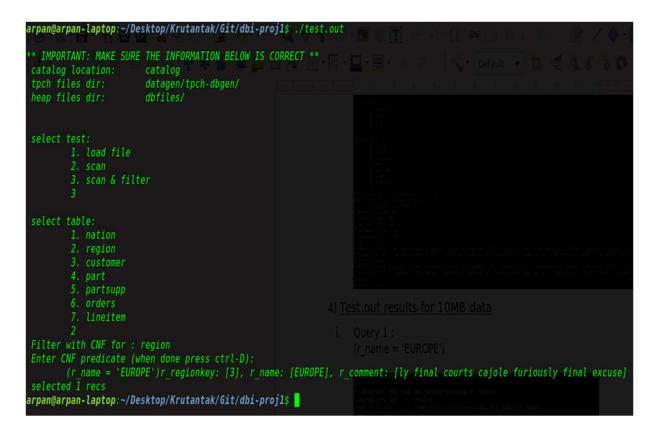
```
(I_orderkey > 100) AND
(I_orderkey < 1000) AND
(I_partkey > 100) AND
(I_partkey < 5000) AND
(I_shipmode = 'AIR') AND
(I_linestatus = 'F') AND
(I_tax < 0.07)
```

```
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$ ./test.out
   IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
                                     datagen/tpch-dbgen/ (o totalprice < 1851.89)
  tpch files dir:
  heap files dir:
  select test:
             1. load file
  select table:
            3. customer
             5. partsupp
             6. orders
  Filter with CNF for : lineitem
  Enter CNF predicate (when done press ctrl-D):
           (l orderkey > 100) AND
   orderkey < 1000) AND
   partkey > 100) AND
   partkey < 5000) AND
   shipmode = 'AIR') AND
  Lax < 0.07)Lorderkey: [130], Loartkey: [1739], Lsuppkey: [4240], Llinenumber: [2], Lquantity: [48], Lextendedprice: [78755], Ldiscount: [0.03], Ltax: [0.02], Lreturnflag: [R], Linestatus: [F], Lshipdate: [1992-07-01], Lcommitdate: [1992-07-12], Lreceiptdate: [1992-07-24], Lshipinstruct: [NONE], Lshipmode: [AIR], Lcomment: [Lithely alongside of the regu] orderkey: [194], Lpartkey: [2594], Lsuppkey: [5095], Llinenumber: [1], Lquantity: [17], Lextendedprice: [25442], Ldiscount: [0.05], Ltax: [0.04], Lreturnflag: [R], Llinestatus: [F], Lshipdate: [1992-05-24], Lcommitdate: [1992-05-22], Lreceiptdate: [1992-05-30], Lshipinstruct: [COLLECT COD], Lshipmode: [AIR], Lcomment: [regular deposi]
  selected 2 recs
 rpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$
```

# 4) Test.out results for 10MB data

#### i. Query 1:

(r\_name = 'EUROPE')



## ii. Query 2:

(r\_name < 'middle east') AND
(r\_regionkey > 1)

#### iii. Query 3:

```
(n_regionkey = 3) AND
(n_nationkey > 10) AND
(n_name > 'japan')
```

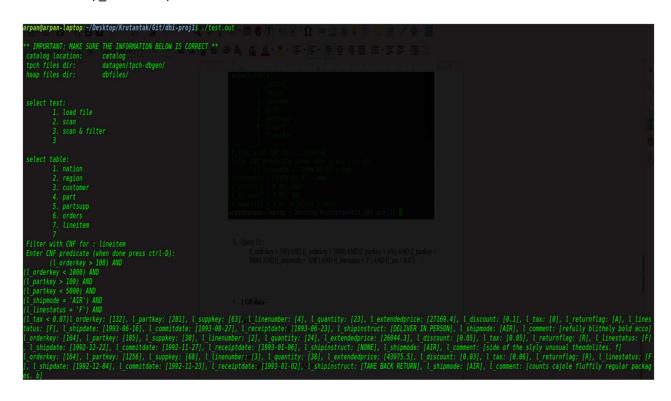
```
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$ ./test.out
** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location:
                       catalog
 tpch files dir:
                       datagen/tpch-dbgen/
 heap files dir:
                       dbfiles/
 select test:
        1. load file
        2. scan
        3. scan & filter
 select table:
        1. nation
        2. region
        3. customer
        4. part
        5. partsupp
        6. orders
        7. lineitem
 Filter with CNF for : nation
 Enter CNF predicate (when done press ctrl-D)://metafileName.c strl
       (n regionkey = 3) AND (n nationkey > 10) AND (n name > 'japan') selected 0 recs
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$
```

#### iv. Query 11:

```
(I_shipdate > '1994-01-01') AND
(I_shipdate < '1994-01-07') AND
(I_discount > 0.05) AND
(I_discount < 0.06) AND
(I_quantity = 4.0)
```

## v. Query 12:

```
(I_orderkey > 100) AND
(I_orderkey < 1000) AND
(I_partkey > 100) AND
(I_partkey < 5000) AND
(I_shipmode = 'AIR') AND
(I_linestatus = 'F') AND
(I_tax < 0.07)
```



```
orderkey: [164], l.partkey: [1089], l.suppkey: [90], l.linenumber: [6], l.quantity: [27], l.extendedprice: [26732.2], l.discount: [0.1], l.tax: [0.04], l.returnflag: [R], l.linestatus: [F]
l.shipdate: [1992-12-23], l.commitdate: [1993-01-16], l.receiptdate: [1993-01-10], l.shipinstruct: [DELIVER IN PERSON], l.shipmode: [AIR], l.comment: [ayers wake carefully a]
orderkey: [226], l.partkey: [407], l.suppkey: [95], l.linenumber: [4], l.quantity: [45], l.extendedprice: [58833], l.discount: [0.1], l.tax: [0.02], l.returnflag: [R], l.linestatus: [F], l.
shipdate: [1993-04-17], l.commitdate: [1993-05-27], l.receiptdate: [1993-05-11], l.shipinstruct: [DELIVER IN PERSON], l.shipmode: [AIR], l.comment: [carefully pending pi]
Conderkey: [261], L partkey: [1740], L suppkey: [25], L linenumber: [3], L quantity: [28], L extendedprice: [45968.7], L discount: [0.08], L tax: [0.03], L returnflag: [R], L linestatus: [F, L shipdate: [1993-07-24], L commitdate: [1993-08-20], L receiptdate: [1993-08-05], L shipinatruct: [COLLECT COD], L shipmode: [AIR], L comment: [ironic packages mag slyly, carefully fin] orderkey: [261], L partkey: [970], L suppkey: [5], L linenumber: [6], L quantity: [20], L extendedprice: [37419.4], L discount: [0.06], L tax: [0.06], L returnflag: [A], L linestatus: [F], L shipdate: [1993-10-15], L commitdate: [1993-09-05], L receiptdate: [1993-11-07], L shipinatruct: [NONE], L shipmode: [AIR], L comment: [ing to the special, ironic deposi] orderkey: [294], L partkey: [597], L suppkey: [88], L linenumber: [1], L quantity: [31], L extendedprice: [46425.3], L discount: [0], L tax: [0.01], L returnflag: [R], L linestatus: [F], L shipdate: [1993-08-06], L commitdate: [1993-08-19], L receiptdate: [1993-08-13], L shipinstruct: [TAKE BACK RETURN], L shipmode: [AIR], L comment: [le fluffily along the quick]
 orderkey: [295], l_partkey: [914], l_suppkey: [17], l_linenumber: [2], l_quantity: [26], l_extendedprice: [47187.7], l_discount: [0.04], l_tax: [0.03], l_returnflag: [R], l_linestatus: [F
  l shipdate: [1994-12-13], l commitdate: [1994-11-30], l receiptdate: [1995-01-06], l shipinstruct: [DELIVER IN PERSON], l shipmode: [AIR], l comment: [ts above the slyly regular requests a
 orderkey: [322], L partkey: [1833], L suppkey: [77], L linenumber: [4], L quantity: [10], L extendedprice: [17348.3], L discount: [0.06], L tax: [0.03], L returnflag: [R], L linestatus: [F L shipdate: [1992-04-12], L commitdate: [1992-05-13], L receiptdate: [1992-04-14], L shipinstruct: [DELIVER IN PERSON], L shipmode: [AIR], L comment: [ deposits grow slyly according to the
 orderkey: [322], l partkey: [334], l suppkey: [91], l linenumber: [6], l quantity: [3], l extendedprice: [3702.99], l discount: [0.08], l tax: [0.05], l returnflag: [A], l linestatus: [F]
  shipdate: [1992-07-03], l commitdate: [1992-05-10], l receiptdate: [1992-07-28], l shipinstruct: [NONE], l shipmode: [AIR], l comment: [ending, ironic deposits along the blith]
 orderkey: [327], l partkey: [418], l suppkey: [48], l linenumber: [2], l quantity: [9], l extendedprice: [11865.7], l discount: [0.09], l tax: [0.05], l returnflag: [A], l linestatus: [F]
  shipdate: [1995-05-24], l commitdate: [1995-07-11], l receiptdate: [1995-06-05], l shipinstruct: [NONE], l shipmode: [AIR], l comment: [ asymptotes are fu]
Conderkey: [353], L partkey: [1194], L suppkey: [67], L linenumber: [1], L guantity: [41], L extendedprice: [4490.8], L discount: [0], L tax: [0.06], L returnflag: [A], L linestatus: [F], shipdate: [1994-03-25], L commitdate: [1994-03-31], L receiptdate: [1994-03-30], L shipinstruct: [DELIVER IN PERSON], L shipmode: [AIR], L comment: [refully final theodoli] orderkey: [416], L partkey: [1742], L suppkey: [85], L linenumber: [3], L quantity: [25], L extendedprice: [4109.5], L discount: [0.07], L tax: [0.01], L returnflag: [R], L linestatus: [F], L shipdate: [1993-10-16], L commitdate: [1993-12-03], L receiptdate: [1993-10-29], L shipinstruct: [NONE], L shipmode: [AIR], L comment: [ses boost after the bold requests.] orderkey: [450], L partkey: [781], L suppkey: [82], L linenumber: [5], L quantity: [27], L extendedprice: [3363.56], L discount: [0.09], L tax: [0], L returnflag: [A], L linestatus: [F], L product [1005.93.21], L commitdate: [1005.93.21], L comment: [1005.93.21], L commit [1005.93.21],
 ipdate: [1995-03-11], l commitdate: [1995-05-21], l receiptdate: [1995-03-16], l shipinstruct: [COLLECT COD], l shipmode: [AIR], l comment: [y even pinto beans; qui]
 orderkey: [481], l partkey: [207], l_suppkey: [8], l̄ linenumber: [2], l quantity: [19], l extendedprice: [21036.8], l discount: [0.08], l tax: [0.01], l returnflag: [R], l linestatus: [F]
   shipdate: [1993-01-09], l commitdate: [1992-11-27], l receiptdate: [1993-01-14], l shipinstruct: [TAKE BACK RETURN], l shipmode: [AIR], l comment: [p blithely after t]
 orderkey: [481], L partkey: [1120], L suppkey: [93], L linenumber: [5], L quantity: [31], L extendedprice: [31654.7], L discount: [0.05], L tax: [0.01], L returnflag: [A], L linestatus: [F L shipdate: [1993-01-15], L commitdate: [1992-12-31], L receiptdate: [1993-01-21], L shipinstruct: [DELIVER IN PERSON], L shipmode: [AIR], L comment: [usly final packages believe. quick] orderkey: [548], L partkey: [568], L suppkey: [59], L linenumber: [4], L quantity: [21], L extendedprice: [30839.8], L discount: [0.08], L tax: [0.03], L returnflag: [A], L linestatus: [F] L shipdate: [1994-10-27], L commitdate: [1994-12-04], L receiptdate: [1994-11-21], L shipinstruct: [DELIVER IN PERSON], L shipmode: [AIR], L comment: [ engage quickly, regular theo]
 orderkey: [549], L partkey: [653], L suppkey: [54], L linenumber: [3], L quantity: [36], L extendedprice: [55931.4], L discount: [0.08], L tax: [0.04], L returnflag: [R], L linestatus: [F]
     shipdate: [1992-09-11], l commitdate: [1992-10-11], l receiptdate: [1992-09-12], l shipinstruct: [DELIVER IN PERSON], l shipmode: [AIR], l comment: [ts against the ironic, even theodolit
```

| Contenting | Con

# 5) gTest results:

- i. TEST(DBFileTest, Create\_Equal)
  - a. Check return value is 1 on successful creation of file.
- ii. TEST(DBFileTest, Create\_NotEqual)
  - **a.** Check program terminates with error message "BAD! Open did not work for ..." on passing empty file path function argument.
  - b. Check return value is 0 on passing wrong file type function argument.
- iii. TEST(DBFileTest, Open\_Equal)
  - a. Check return value is 1 on successfully opening file at the specified file path.
- iv. TEST(DBFileTest, Open\_NotEqual)
  - a. Check return value is 0 on passing incorrect file name.
- v. TEST(DBFileTest, Close Emty File)
  - a. Check return value is 0 on closing files that are not loaded or opened.

## vi. TEST(DBFileTest, Close\_ Nations\_DBFile) -

a. Check return type is not 0 on closing non empty opened files.

```
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-projl$ ./DBFileTest.out
** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
=======] Running 6 tests from 1 test case.
     -----] Global test environment set-up.
       --- 1 6 tests from DBFileTest
 RUN
         ] DBFileTest.Create Equal
       OK ] DBFileTest.Create Equal (1 ms)
          ] DBFileTest.Create NotEqual
       OK ] DBFileTest.Create_NotEqual (1 ms) | | File :: GetLength
 RUN
         | DBFileTest.Open Equal
       OK J DBFileTest.Open Equal (0 ms)
         ] DBFileTest.Open NotEqual
       OK | DBFileTest.Open NotEqual (0 ms)
          J DBFileTest.Close Empty File
       OK ] DBFileTest.Close Empty File (0 ms)
         J DBFileTest.Close Nations DBFile
       OK | DBFileTest Close Nations DBFile (13 ms)
        -- 1 6 tests from DBFileTest (16 ms total)
       -- I Global test environment tear-down
  ======= 1 6 tests from 1 test case ran. (16 ms total)
  PASSED ] 6 tests.
arpan@arpan-laptop:~/Desktop/Krutantak/Git/dbi-proj1$
```