

COP6726 – Database System Implementation

Project 3 Report

Team members:

1. Akshay Ganapathy (UFID: 3684-6922)
2. Riyaz Shaik (UFID: 4360-0170)

Steps to execute project code:

The code was executed using **Windows Subsystem for Linux (Ubuntu 18.04)** environment.

1. The project source code is in the folder “a3-test”. Its contents include the following:
 - a. Code needed to run test.cc.
 - b. All “.bin” files should be included in the “bin” folder.
 - c. Google test framework has been added in the directory “googletest”. This framework is needed to run gtest unit test cases.
 - d. All the gtest unit test cases have been added in “test” folder. RelOpTest.cc contains the unit test cases for methods in RelOp class, respectively. Make sure to include “partsupp.bin”, “part.bin”, “supplier.bin”, “partsupp.bin.meta”, “part.bin.meta”, “supplier.bin.meta” files in the directory, “bin”.
 - e. Makefile to compile test.cc and gtest unit test cases.
2. To run test.cc, from project root execute command “make test.out” and then execute with test case number i.e., “./test.out [1-5]”.

The following is the test run for a particular test case number:

Test Case Number	Description
1	Runs test case to compute query “select * from partsupp where ps_supplycost < 1.03”
2	Runs test case to compute query “select p_partkey(0), p_name(1), p_retailprice(7) from part where (p_retailprice > 931.01) AND (p_retailprice < 931.3);”
3	Runs test case to compute query “select sum (s_acctbal + (s_acctbal * 1.05)) from supplier;”
4	Runs test case to compute query “select sum (ps_supplycost) from supplier, partsupp where s_suppkey = ps_suppkey;”
5	Runs test case to compute query “select distinct ps_suppkey from partsupp where ps_supplycost < 100.11;”

3. To run gtest unit test cases, from project root execute command “make gtests” and then execute “./gtests”.

The following are the methods with their description:

Method	Description
void SelectFile::Run(DBFile &inFile, Pipe &outputPipe, CNF &selOp, Record &literal);	It initializes the class level variables and starts SelectFile class worker thread.
void *SelectFile::thread_work(void *args);	Inserts records matched according to CNF expression from a dbfile into output pipe.
void SelectPipe::Run(Pipe &inputPipe, Pipe &outputPipe, CNF &selOp, Record &literal);	It initializes the class level variables and starts SelectPipe class worker thread.
void *SelectPipe::thread_work(void *args);	It does a scan of the dbfile file provided, and for every tuple accepted by the CNF, it inserts the tuple into the pipe as output
void Sum::Run(Pipe &inputPipe, Pipe &outputPipe, Function &computeMe);	It initializes the class level variables and starts Sum class worker thread.
void *Sum::thread_work(void *args);	Sum computes the SUM SQL aggregate function over the input pipe, and puts a single tuple into the output pipe that has the sum.
void DuplicateRemoval::Run(Pipe &inputPipe, Pipe &outputPipe, Schema &mySchema)	It initializes the class level variables and starts DuplicateRemoval class worker thread.
void *DuplicateRemoval::thread_work(void *args)	It takes an input pipe, an output pipe, as well as the schema for the tuples coming through the input pipe, and does a duplicate removal.
void Project::Run(Pipe &inputPipe, Pipe &outputPipe, int *keepMe, int numAttsInput, int numAttsOutput)	It initializes the class level variables and starts Project class worker thread.
void *Project::thread_work(void *args)	It performs SQL Projection Operation and inserts only the attributes provided in the input from the tuples into the output pipe
void GroupBy::Run(Pipe &inputPipe, Pipe &outputPipe, OrderMaker &groupAtts, Function &computeMe)	It initializes the class level variables and starts GroupBy class worker thread.
void *GroupBy::thread_work(void *args)	It does grouping, and then puts one sum into the output pipe for each group. Every tuple put into the output pipe has a sum as the first attribute, followed by the values for each of the grouping attributes as the remainder of the attributes.

void WriteOut::Run(Pipe &inputPipe, FILE *outFile, Schema &mySchema)	It initializes the class level variables and starts WriteOut class worker thread.
void *WriteOut::thread_work(void *args)	It uses the schema to write text version of the output records to the file.
void Join::Run(Pipe &inputPipeL, Pipe &inputPipeR, Pipe &outputPipe, CNF &selOp, Record &literal)	It initializes the class level variables and starts Join class worker thread.
void *Join::thread_work(void *args)	It takes two input pipes, an output pipe, and a CNF, and joins all of the records from the two pipes according to that CNF
void WriteOut::WaitUntilDone()	It makes thread wait until other threads are completed.

The following are the screenshots of the output after executing the 5 testcases.

The following results are obtained by using 1 GB TPCB data generated from tpch-dbgen.

```

output1.txt - Notepad
File Edit Format View Help

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

ps_partkey: [158093], ps_suppkey: [5639], ps_availqty: [3751], ps_supplycost: [1], ps_comment: [iously unusi
ps_partkey: [102497], ps_suppkey: [2498], ps_availqty: [6491], ps_supplycost: [1], ps_comment: [fully final
ps_partkey: [76994], ps_suppkey: [9502], ps_availqty: [9712], ps_supplycost: [1], ps_comment: [. carefully :
ps_partkey: [155112], ps_suppkey: [5113], ps_availqty: [7635], ps_supplycost: [1], ps_comment: [refully bol
ps_partkey: [122543], ps_suppkey: [5056], ps_availqty: [5753], ps_supplycost: [1], ps_comment: [e the quick
ps_partkey: [20468], ps_suppkey: [469], ps_availqty: [6884], ps_supplycost: [1], ps_comment: [furiously amo
ps_partkey: [193659], ps_suppkey: [6179], ps_availqty: [6606], ps_supplycost: [1.01], ps_comment: [can hagg
ps_partkey: [193981], ps_suppkey: [3982], ps_availqty: [619], ps_supplycost: [1.01], ps_comment: [ic account
ps_partkey: [60953], ps_suppkey: [954], ps_availqty: [8611], ps_supplycost: [1.01], ps_comment: [ully even i
ps_partkey: [6333], ps_suppkey: [6334], ps_availqty: [3711], ps_supplycost: [1.01], ps_comment: [s use slyl
ps_partkey: [139711], ps_suppkey: [9712], ps_availqty: [4286], ps_supplycost: [1.01], ps_comment: [ully unu
ps_partkey: [71984], ps_suppkey: [6999], ps_availqty: [6016], ps_supplycost: [1.01], ps_comment: [eodolites
ps_partkey: [43172], ps_suppkey: [685], ps_availqty: [6600], ps_supplycost: [1.01], ps_comment: [ites integ
ps_partkey: [9097], ps_suppkey: [4098], ps_availqty: [3012], ps_supplycost: [1.01], ps_comment: [s the bold
ps_partkey: [43764], ps_suppkey: [1277], ps_availqty: [2344], ps_supplycost: [1.02], ps_comment: [; furious
ps_partkey: [61707], ps_suppkey: [1708], ps_availqty: [3178], ps_supplycost: [1.02], ps_comment: [ide of th
ps_partkey: [93653], ps_suppkey: [3654], ps_availqty: [4473], ps_supplycost: [1.02], ps_comment: [ of the c
ps_partkey: [51671], ps_suppkey: [4177], ps_availqty: [3399], ps_supplycost: [1.02], ps_comment: [iously. b
ps_partkey: [27115], ps_suppkey: [9618], ps_availqty: [7966], ps_supplycost: [1.02], ps_comment: [e regular
ps_partkey: [74375], ps_suppkey: [6883], ps_availqty: [864], ps_supplycost: [1.02], ps_comment: [lithely ex
ps_partkey: [34494], ps_suppkey: [9501], ps_availqty: [7438], ps_supplycost: [1.02], ps_comment: [egular ex

query1 returned 21 records

*****
Ln 1, Col 1    100%    Unix (LF)    UTF-8

```

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

catalog location: catalog
tpch files dir: tpch-dbggen/
heap files dir: bin/

int: [10021], string: [blush steel green sienna snow], double: [931.02]
int: [11020], string: [plum khaki powder beige peru], double: [931.02]
int: [3028], string: [puff slate tomato moccasin azure], double: [931.02]
int: [8023], string: [mint salmon moccasin blanched beige], double: [931.02]
int: [2029], string: [midnight brown dim violet almond], double: [931.02]
int: [4027], string: [white ivory moccasin coral puff], double: [931.02]
int: [9022], string: [peru misty sandy dark drab], double: [931.02]
int: [5026], string: [blanched blush pink light wheat], double: [931.02]
int: [6025], string: [purple medium light aquamarine dark], double: [931.02]
int: [7024], string: [forest rosy peach antique midnight], double: [931.02]
int: [1030], string: [orange floral olive ivory lace], double: [931.03]
int: [31], string: [slate seashell steel medium moccasin], double: [931.03]

query2 returned 12 records

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

catalog location: catalog
tpch files dir: tpch-dbggen/
heap files dir: bin/

double: [9.24623e+07]

< >

Ln 1, Col 1

100%

Unix (LF)

UTF-8

output1.txt - Notepad

File Edit Format View Help

query2 returned 12 records

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

catalog location: catalog

tpch files dir: tpch-dbggen/

heap files dir: bin/

double: [9.24623e+07]

query3 returned 1 records

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

catalog location: catalog

tpch files dir: tpch-dbggen/

heap files dir: bin/

query4

double: [4.00421e+08]

query4 returned 1 recs

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

<

Ln 1, Col 1

100%

Unix (LF)

UTF-8

>

```
output1.txt - Notepad
File Edit Format View Help
heap files dir:      bin/

double: [9.24623e+07]

query3 returned 1 records

*****

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

query4
double: [4.00421e+08]
query4 returned 1 recs

*****

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

query5 finished..output written to file ps.w.tmp

Ln 1, Col 1    100%    Unix (LF)    UTF-8
```

The following are the results of gtest unit test cases:

Result: All the unit test cases passed

RelOpTest.cc contains the unit test cases for methods in RelOp class, respectively.

To run gtest unit test cases, from project root execute command “make gtests” and then execute “./gtests”.

Make sure to include “partsupp.bin”, “part.bin”, “supplier.bin”, “partsupp.bin.meta”, “part.bin.meta”, “supplier.bin.meta” files in the directory, “bin”.

```

akshayg1996@LAPTOP-ACV240Q1:/mnt/d/Gitlab/dbi/Project 3$ ./gtests
Running main() from gmock_main.cc
[=====] Running 4 tests from 1 test suite.
[-----] Global test environment set-up.
[-----] 4 tests from RelOpTest
[ RUN      ] RelOpTest.TEST1

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

ps_partkey: [20468], ps_suppkey: [469], ps_availqty: [6884], ps_supplycost: [1], ps_comment: [furiousl
y among the slyly ironic instructions. final, unusual packages wake slyly. final accounts cajole. depo
sits above the i]
ps_partkey: [43172], ps_suppkey: [685], ps_availqty: [6600], ps_supplycost: [1.01], ps_comment: [ites
integrate blithely above the slyly regular instructions. asymptotes besides the regular, even accounts
haggle carefully slyly bold requests. even pinto beans ]
ps_partkey: [60953], ps_suppkey: [954], ps_availqty: [8611], ps_supplycost: [1.01], ps_comment: [ully
even dolphins wake carefully about the slyly final pinto beans]
ps_partkey: [43764], ps_suppkey: [1277], ps_availqty: [2344], ps_supplycost: [1.02], ps_comment: [; fu
rious, ironic requests nag furiously against the silent packages-- furiously pending pinto beans use b
lithely careful]
ps_partkey: [61707], ps_suppkey: [1708], ps_availqty: [3178], ps_supplycost: [1.02], ps_comment: [ide
of the unusual, regular excuses. unusual, special packages are carefully across the even theodolites:
furi]
ps_partkey: [102497], ps_suppkey: [2498], ps_availqty: [6491], ps_supplycost: [1], ps_comment: [fully
final accounts. even accounts after the carefully final accounts haggle according to the blithely spec
ial requests. carefully unusual]
ps_partkey: [93653], ps_suppkey: [3654], ps_availqty: [4473], ps_supplycost: [1.02], ps_comment: [ of
the carefully final requests. bold deposits are slyly. instructions nod furiously instructions. carefu
l]
ps_partkey: [193981], ps_suppkey: [3982], ps_availqty: [619], ps_supplycost: [1.01], ps_comment: [ic a
ccounts after the unusual, regular instructions grow carefully around the blithely unusual dependencie
s. pending accounts along the bl]
ps_partkey: [9097], ps_suppkey: [4098], ps_availqty: [3012], ps_supplycost: [1.01], ps_comment: [s the

```

```

ps_partkey: [34494], ps_suppkey: [9501], ps_availqty: [7438], ps_supplycost: [1.02], ps_comment: [egular excuses. final, regular deposits wake. pinto beans according to th]
ps_partkey: [76994], ps_suppkey: [9502], ps_availqty: [9712], ps_supplycost: [1], ps_comment: [. carefuly ironic platelets cajole furiously among the furiously regular asymptotes. furiously express asymptotes wake caref]
ps_partkey: [27115], ps_suppkey: [9618], ps_availqty: [7966], ps_supplycost: [1.02], ps_comment: [e regular, ironic dugouts. slyly special requests cajole quickly across the blithely express requests. deposits unwind carefully pending theodolites. pinto beans about the even, regular theodolite]
ps_partkey: [139711], ps_suppkey: [9712], ps_availqty: [4286], ps_supplycost: [1.01], ps_comment: [ully unusual escapades sleep along the special instructions. final, bold ideas across the slyly ironic ideas sleep dependenc]

```

```

query1 returned 21 records
[ OK ] RelOpTest.TEST1 (1250 ms)
[ RUN ] RelOpTest.TEST1_FAIL

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

```

```

ps_partkey: [20468], ps_suppkey: [469], ps_availqty: [6884], ps_supplycost: [1], ps_comment: [furiously among the slyly ironic instructions. final, unusual packages wake slyly. final accounts cajole. deposits above the i]
ps_partkey: [43172], ps_suppkey: [685], ps_availqty: [6600], ps_supplycost: [1.01], ps_comment: [ites integrate blithely above the slyly regular instructions. asymptotes besides the regular, even accounts haggle carefully slyly bold requests. even pinto beans ]
ps_partkey: [60953], ps_suppkey: [954], ps_availqty: [8611], ps_supplycost: [1.01], ps_comment: [ully even dolphins wake carefully about the slyly final pinto beans]
ps_partkey: [43764], ps_suppkey: [1277], ps_availqty: [2344], ps_supplycost: [1.02], ps_comment: [; furious, ironic requests nag furiously against the silent packages-- furiously pending pinto beans use blithely careful]
ps_partkey: [61707], ps_suppkey: [1708], ps_availqty: [3178], ps_supplycost: [1.02], ps_comment: [ide of the unusual, regular excuses. unusual, special packages are carefully across the even theodolites: furi]

```

```

ps_partkey: [27115], ps_suppkey: [9618], ps_availqty: [7966], ps_supplycost: [1.02], ps_comment: [e regular, ironic dugouts. slyly special requests cajole quickly across the blithely express requests. deposits unwind carefully pending theodolites. pinto beans about the even, regular theodolite]
ps_partkey: [139711], ps_suppkey: [9712], ps_availqty: [4286], ps_supplycost: [1.01], ps_comment: [ully unusual escapades sleep along the special instructions. final, bold ideas across the slyly ironic ideas sleep dependenc]

```

```

query1 returned 21 records
[ OK ] RelOpTest.TEST1_FAIL (958 ms)
[ RUN ] RelOpTest.TEST2

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

```

```

int: [31], string: [slate seashell steel medium moccasin], double: [931.03]
int: [1030], string: [orange floral olive ivory lace], double: [931.03]
int: [2029], string: [midnight brown dim violet almond], double: [931.02]
int: [3028], string: [puff slate tomato moccasin azure], double: [931.02]
int: [4027], string: [white ivory moccasin coral puff], double: [931.02]
int: [5026], string: [blanched blush pink light wheat], double: [931.02]
int: [6025], string: [purple medium light aquamarine dark], double: [931.02]
int: [7024], string: [forest rosy peach antique midnight], double: [931.02]
int: [8023], string: [mint salmon moccasin blanched beige], double: [931.02]
int: [9022], string: [peru misty sandy dark drab], double: [931.02]
int: [10021], string: [blush steel green sienna snow], double: [931.02]
int: [11020], string: [plum khaki powder beige peru], double: [931.02]

```

```

query2 returned 12 records
[ OK ] RelOpTest.TEST2 (300 ms)
[ RUN ] RelOpTest.TEST3

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

```



```
int: [8023], string: [mint salmon moccasin blanched beige], double: [931.02]
int: [9022], string: [peru misty sandy dark drab], double: [931.02]
int: [10021], string: [blush steel green sienna snow], double: [931.02]
int: [11020], string: [plum khaki powder beige peru], double: [931.02]
```

```
query2 returned 12 records
[      OK ] RelOpTest.TEST2 (300 ms)
[  RUN   ] RelOpTest.TEST3
```

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

```
double: [9.24623e+07]
```

```
query3 returned 1 records
[      OK ] RelOpTest.TEST3 (87 ms)
[-----] 4 tests from RelOpTest (2595 ms total)

[-----] Global test environment tear-down
[=====] 4 tests from 1 test suite ran. (2595 ms total)
[  PASSED ] 4 tests.
akshayg1996@LAPTOP-ACV240Q1:/mnt/d/Gitlab/dbi/Project 3$
```