

Chris Bora(cdb239)  
Richie Henwood(rbh228)  
Han Wen Chen(hc844)

### Project 4 Experiments

#### 1. Queries

```
1.Select S.A, S.B, S.C
   From Sailors S
   Where S.A > 900 AND S.B < 300;
2.Select S.A, S.B, S.C
   From Sailors S
   Where S.A < 600 And S.A > 550;
3.Select S.A, S.B, S.C
   Sailors S
   Where S.A = 400;
```

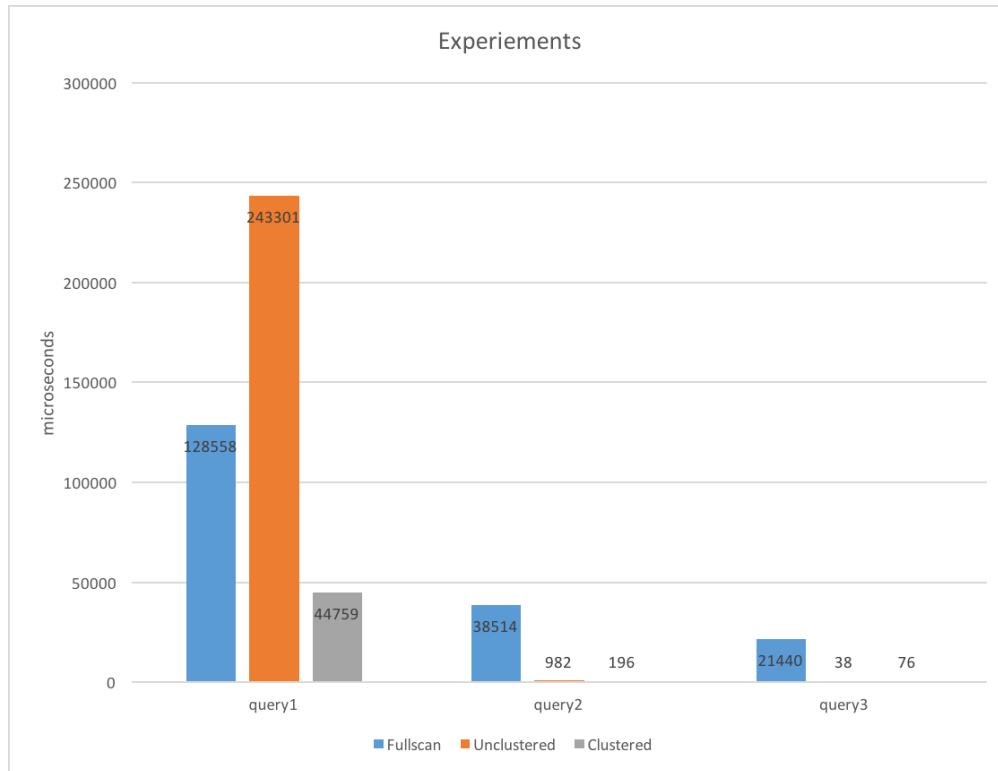
#### 2. Description of Data used

Table name	Number of columns	Column names	Number of tuples
Sailors	3	A B C	12,000

Generation of tuples: Column A was at first generated in order from 0 to 11,999. The values of the remaining columns were generated using a random number generator with a range [0,1000] using java's Random::randint(). We then shuffled everything using Collections.shuffle() so that the data would no longer be sorted on the first column.

#### 3. Indexed on column A with order (D) = 15

4.



Query 1: This query put a very low selectivity on the indexed column (our selection condition was  $S.A > 900$  where  $A$  ranges from 0 to 11,999). For this query, the unclustered index did much worse than a full table-scan. This happened because (since the index was unclustered) we had to jump around the data file thousands of times. The unclustered index still outperformed the full-scan though. This is because the clustered index allowed us to quickly find where to start our scan but did not require us to jump around the data file.

Query 2: Here we used a relatively high selectivity on the index column (our selection condition was  $550 < S.A < 600$ ). The index scans now both drastically outperformed the full-table scan. The clustered index still outperformed the unclustered index because the clustered index did not require us to jump around the data file.

Query 3: Here we used the highest possible selectivity, i.e. an equality condition. Both index scans performed hundreds of times better than the full-scan. The clustered and unclustered index performed about the same though. This is because we performed an equality selection, so the unclustered index did not require us to jump around the data file much.