Optimization

Database Administration
Lab Guide 4

2022/2023

Consider now the optimization of the toy benchmark application. For each change, evaluate its the impact by (i) observing the query plan and (ii) running the benchmark for 1, 2, 4, 8, ... threads and plotting the corresponding scalability curve.

Steps

1. Change the *sell* id distribution by performing the following changes:

   - clientid = $random(2^N) \,|\, random(2^N)$;
   - productid = $random(2^N) \,|\, random(2^N)$.

   Using SQL, check how the new distribution behaves by obtaining the invoice count per client.

2. Analyze the plans generated by executing the *account* operation with the client with the most invoices, the 10th client with most invoices, and the client with least invoices.

3. Evaluate the impact of removing the prepared statements from the benchmark.[1]

4. Search and evaluate different plans by disabling operators.

5. Determine what resource is limiting the performance for each configuration. Consider changes to memory configuration and CPU vs. I/O weights in the server configuration file.

Questions

1. Does the engine generate different plans for the same query based on the arguments provided?

2. Does the engine generate different plans for the same **prepared** query based on the arguments provided?[2] What advantages do prepared statements provide?

[1] In each query, create a new Statement object with the query and hardcoded parameters. E.g.:
```
Statement s = this.c.createStatement();
s.executeUpdate(String.format("insert into invoice (productid,
clientid, data) values (%s, %s, '%s')", clientId, productId,
randomString(1000)));
```

[2] Prepared statements in `psql` can be created using the `PREPARE` command: `https://www.postgresql.org/docs/current/sql-prepare.html`

3. Considering all the optimizations that you have performed, to what extent have you improved the maximum throughput of your application? What was the improvement in response time?

4. What is the relative impact of redundancy/algorithmic changes vs. configuration changes?

Learning Outcomes Discuss trade-offs between different optimization decisions. Plan, conduct, and justify the steps to optimize the performance of a relational application.