Running the workload

There are 6 steps to run workload against the database. The steps are used to run the single client server. Also this should be sufficient for small or medium cluster. For the larger cluster, we need run the multiple clients on the different server to generate enough workload for effective benchmarking of the system.

1. *Setup the database system to test*

Setup the configuration that needs to be benchmark. We must create or setup tables or keyspaces or storage bucket to store records. Before the YCSB client runs, the table must be created also client won't request to create the table during the benchmarking. Some database systems, the table creation happens manually. Created table depends on the workload. For the core workload, the Client will assume that there is a table called USERTABLE with flexible schema. If additional columns is required during the runtime.

2. *Choose the appropriate DB interface layer*

The DB interface layer is implemented using a java class which executes read, insert, update, scan call with help of YCSB client and calls against the DB's API. We have specify the class name of the layer in the command line and specify the properties in the command line or using parameter files. These property parameters will be DB interface instance and used to configure the layer for benchmarking.

3. *Choose the appropriate workloads*

We have two phases in the YCSB client. Loading phase – the data will be loaded into database (proper kind of record can be constructed and inserted). Transaction phase – the operation will be executed against the dataset (Correct the record ids and fields can be referred)

Also both the phases will use the parameter file. Workload Java class specify the properties will perform either Loading or Transaction phase. We can specify both java class and parameter files on the command line. The client will dynamically load the java class and pass the properties from the parameter file and execute the workload for both phases.

4. *Choose the appropriate runtime parameters*

Though we have java class and parameter file to specify the workload. YCSB client have additional settings that needed for specific benchmark run and this setting can be provided on the command line.

~threads – Number of client threads. By default, YCSB cilent uses single worker thread. Additional threads can be specified using this setting.

~target – Target number of operation per second. By default, the Client will try to perform as many operation as possible.

~*s* – Status. This can be used for the long running workload and it helps us to check whether system is running and crashed during the operations. By using this setting, the client will report the status every 10 seconds.

5. *Load the database*

Workloads have two phases loading and transaction phase. To load the data, we can run the YCSB client and execute the loading section.

To load the standard dataset

$ ./bin/ycsb load basic -P workloads/workloada

-P - This is used to load the parameter from the property files.

6. *Execute the workload*

Once the loading phase loads the dataset into the database, we can execute the workloads. We can specify the YCSB client to execute the transaction phase of the workload.

To execute the workload

$ ./bin/ycsb run basic -P workloads/workloada -P large.dat -s > transactions.dat

Reference:

https://github.com/brianfrankcooper/YCSB/wiki/Running-a-Workload