RM-Replay for Cluster Scheduling Project Report 2/29-3/14

Zhen Huang, Blake Ehrenbeck

This two week we are trying to solve two main tasks: (1)After the meeting with Boyang from last Thursday, we found the metrics.log was incorrect, so we need to identify and fix the bug. (2)After repairing RM-Replay, migrate the dataset from Argonne Theta Machine to the RM-replay.

Progress This Week

After our debugging, we found that the trace_metrics.c in the container that we build every single time, some setting for the metric.log are hardcoded, so we needed to adjust the parameters including Nnodes_mc, Nnodes_gpu,Nnodes to be consistent with Jarvis configuration to make sure the metric.log shows up correctly. Also, the metrics builder doesn't consider jobs less than 2 hours, so we modified it to read our jobs.

Modification

After building the Docker images, go to /slurm/slurm-replay/submitter, and open the trace_metrics.c

1. Change the Nodes_mc to be 8, Nodes_gpu to be 3, and Nodes to be Nnodes_mc+Nnodes_gpu. (This is the configuration of the Jarvis machine we ran our tests on).

```
244     double dispersion;
245     double slowdown;
246     int Nnodes_mc  = 8;
247     int Nnodes_gpu  = 3;
248     int Nnodes  = Nnodes_mc+Nnodes_gpu; // all daint + data movers
249     size_t nlistjobs;
250     unsigned long long idx_jobid;
251     unsigned long* listjobs;
252     int list_file = 0;
253     long start_range, end_range;
```

2. Change the time_wait_arr[j] to be bigger than 2 minutes, this will change change the ignoring job duration from 3 hours to 2 minutes which allowed our testing job to not get ignored.

These modifications produced a sensible metrics output.

We are now in the process of rebuilding the Slurm database, adding in the data from Argonne so the trace builder from RM-Replay can produce a trace.
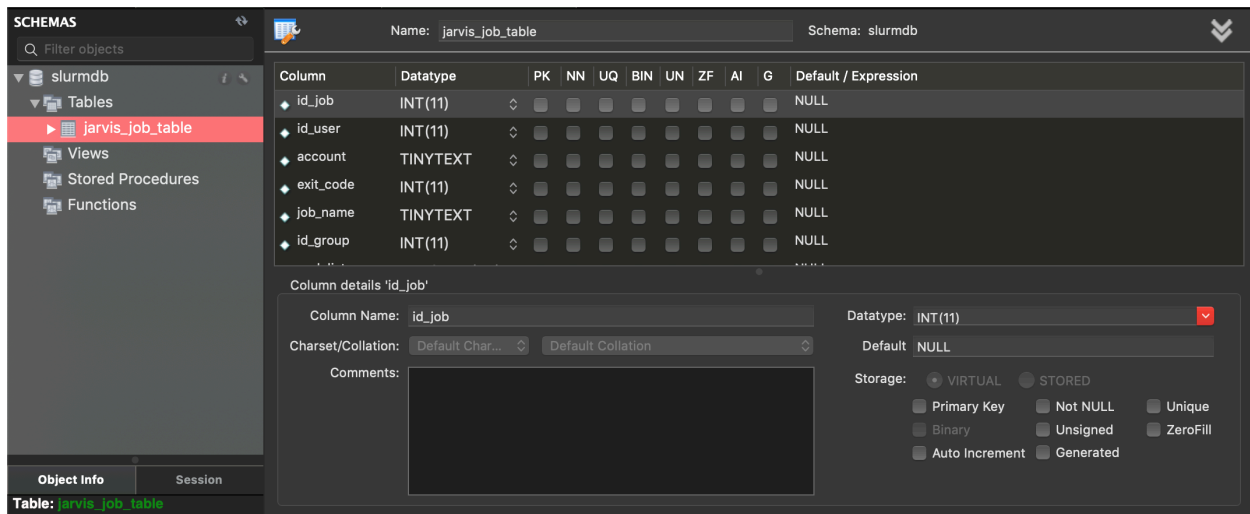
Building the jarvis_job_table

We will migrate the CSV table to this database.