# Instructions

**Gem5 is a cycle-accurate full-system simulator and can incur very long simulation times. So, you may use a two-step method to finish the project.**

First, comment out Gem5-specific parts **(lines 11, 59, 68 and 71 in index.c and lines 11, 112, 121 and 124 in quantify.c in updated Docker Container)**, compile them using gcc, and run normally to check correctness of your implementation.

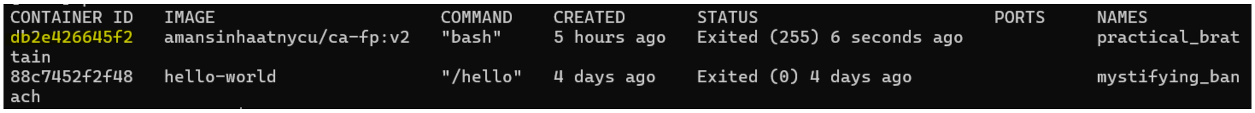
Second, once the correctness has been verified, uncomment the Gem5-specific code segments and use Makefile as per original instructions to compile and run on RISC-V architecture.

**Please follow the following steps to reduce the Gem5 simulation time (~105 minutes for *make run_index* and ~95 minutes for *make run_quantify*) for the default implementation (#transcripts = 4 and #queries = 100) in *evaluation* folder:**

- Pull the updated Docker Image. No function signature has been updated, so you can copy the code that you have implemented into the new container for evaluation folder.

  **sudo docker pull amansinhaatnycu/ca-fp:v4**

- Launch Docker Container from pulled Image. You can run as below to keep the container running in background and then connect to it through multiple shells
  1. Run container in background:
     **sudo docker run -dit amansinhaatnycu/ca-fp:v4 bash**
  2. Get the new container process ID (**CONTAINER ID** field):
     **sudo docker ps -a**



```
CONTAINER ID   IMAGE                    COMMAND     CREATED       STATUS                     PORTS     NAMES
db2e426645f2   amansinhaatnycu/ca-fp:v2 "bash"      5 hours ago   Exited (255) 6 seconds ago           practical_brat
tain
88c7452f2f48   hello-world              "/hello"    4 days ago    Exited (0) 4 days ago                mystifying_ban
ach
```

  3. Connect to the running container:
     **docker exec -it db2e426645f2 bash**

**IMPORTANT NOTES:**

1. Docker containers lose data when exited without explicit commits. So, remember to copy your whole evaluation folder to your local filesystem using below command:

```
sudo docker cp {CONTAINER ID}:/home/CA-FP1/evaluation/ .
```

2. After checking correctness of **make run_distribute**, you can modify the variables ***transcript_sequences_count*** and ***query_sequences_count*** in the files evaluation/transcripts_data.h and evaluation/query_data.h respectively to adjust the runtime.
3. The input datasets for evaluation will be released later, so design your solutions independently of the datasets.
4. The baselines folder has not been modified at all.
5. After you have implemented the functions ***distribute()*** in distribute.cpp, ***generate_index()*** in index.c, and ***quantify()*** in quantify.c, you need to submit the **whole evaluation folder**. You also need to submit a report, as mentioned in slide 22 of the original slides.
6. Submit additional performance-area tradeoff in the report for an extra 10 points, as mentioned in slide 22 of the original slides. Baseline performance will be released on 04/30.
7. You can email me (Dr. Aman Sinha) at amansinha.sw@gmail.com if you have any doubts. Please ensure that only one email is sent from a group. Please discuss among yourselves before emailing.