Using 10x Genomics Cell Ranger

OAsis cluster has Cell Ranger pre-installed. Users may load it from Lmod. The following is an example of converting a tiny sample from BCL format to FASTQ using Cell Ranger with a single node and multinode (cluster mode).

First of all, we will download the sample file we need:

```
# set up our working directory
mkdir -p ~/mrotest
cd ~/mrotest

# download and extract the sample files from 10xgenomics
wget https://cf.10xgenomics.com/supp/cell-exp/cellranger-tiny-bcl-1.2.0.tar.gz
wget https://cf.10xgenomics.com/supp/cell-exp/cellranger-tiny-bcl-simple-1.2.0.csv
tar -zxvf cellranger-tiny-bcl-1.2.0.tar.gz
rm cellranger-tiny-bcl-1.2.0.tar.gz
tree -L 2 cellranger-tiny-bcl-1.2.0/
```

Now, we are ready to convert the sample to the FASTQ format.

We support running Cell Ranger in three different ways. We suggest reading through the following options first before trying it out.

Option1. Local mode (single node, run interactively for testing purposes)

```
# request a node, and run Cell Ranger interactively (suitable for troubleshooting issues)
srun -p batch -c16 --mem 128G --pty bash

# inside the shell, load the respective modules
module load GCC/11.3.0 bcl2fastq2 CellRanger

# run the case
rm -rf test
```

```
cellranger mkfastq --id=test \
  --run=./cellranger-tiny-bcl-1.2.0 \
  --csv=./cellranger-tiny-bcl-simple-1.2.0.csv

# result will be located in the test folder in the current directory
```

Option2. Local mode (single node, scheduled)

You can run Cell Ranger in local mode for a practical but relatively simple case. Following is an example job script, we may name it **run.sh**.

```
#!/usr/bin/env bash

#SBATCH -J mkfastq
#SBATCH -o mkfastq.out
#SBATCH -e mkfastq.out
#SBATCH -p batch
#SBATCH -n 1 -c 16
#SBATCH --mem-per-cpu=8G

module load GCC/11.3.0 bcl2fastq2 CellRanger

rm -rf test

cellranger mkfastq --id=test \
  --run=./cellranger-tiny-bcl-1.2.0 \
  --csv=./cellranger-tiny-bcl-simple-1.2.0.csv
```

Then you can submit the job by running: **sbatch run.sh**

Option3. Cluster mode (multinode)

Multiple nodes may be needed for a larger case. First, we have to set up a job script template. Cell Ranger will then generate and submit jobs based on it. Set up a new file called **slurm.template** as follows.

```
#!/bin/bash
#SBATCH -p batch
#SBATCH -J __MRO_JOB_NAME__
#SBATCH -o __MRO_STDOUT__
#SBATCH -e __MRO_STDERR__
```

```
#SBATCH -N 1
#SBATCH -n 1
#SBATCH -c __MRO_THREADS__
#SBATCH --mem=__MRO_MEM_GB__G
#SBATCH --export=ALL
#SBATCH --signal=2
#SBATCH --time=8:00:00

__MRO_CMD__
```

Then, we may launch Cell Ranger in cluster mode.

```
module load GCC/11.3.0 bcl2fastq2 CellRanger

rm -rf test

cellranger mkfastq --id=test \
  --run=./cellranger-tiny-bcl-1.2.0 \
  --csv=./cellranger-tiny-bcl-simple-1.2.0.csv \
  --jobmode=slurm.template \
  --maxjobs=3 --jobinterval=1000 --mempercore=4

# Here, we restrict Cell Ranger to launch at most 3 concurrent jobs.
# A 1-second interval will be waited between each job.
# 4 GB of memory is requested per core.
# You may play around with these parameters based on your need.
```

Reference links:

https://www.10xgenomics.com/support/software/cell-ranger/latest/tutorials/cr-tutorial-fq

https://www.10xgenomics.com/support/software/cell-ranger/latest/advanced/cr-cluster-mode

---

Revision #7
Created 25 January 2024 06:59:36 by Loki Ng
Updated 25 January 2024 08:24:55 by Loki Ng