Tutorial 4 (2019)

From ACE Lab

Follow the instructions below to complete this tutorial. The tutors are available to answer valid, well thought out questions. The tutors will not touch your keyboard or complete the tutorials for you. Answer questions provided in a text document and email to dmacleod@csir.co.za. Pay special attention to the instructions in red, you will need to demonstrate this to a tutor in order to complete that task.

GROMACS Application Benchmark

GROMACS is a versatile package to perform molecular dynamics, i.e. simulate the Newtonian equations of motion for systems with hundreds to millions of particles. It is primarily designed for biochemical molecules like proteins, lipids and nucleic acids that have a lot of complicated bonded interactions, but since GROMACS is extremely fast at calculating the nonbonded interactions (that usually dominate simulations) many groups are also using it for research on non-biological systems, e.g. polymers.

Ask a tutor to copy the the files required for this tutorial to your headnode.

Part 1: Installation

Detailed installation instructions can be found at: http://manual.gromacs.org/documentation/2018.3/install-guide/index.html

Installation overview:
1. Ensure you have an up-to-date cmake available on your system.
2. You will also require a compiler (eg. GCC, Intel or other) and MPI (OpenMPI, MPICH, Intel MPI or other) be installed on your system and in your PATH & LD_LIBRARY_PATH.
3. Compile GROMACS with MPI support from source using cmake.

Part 2: Benchmark

You have been provided two GROMACS benchmarks. The first benchmark (adh_cubic) should complete within a few minutes and has a small memory footprint, it is intended to demonstrate that your installation is working properly. The second benchmark (1.5M_water) uses more memory and takes considerably longer to complete. The metric which will be used to assess your performance is the ns/day (number of nanoseconds the model is simulated for per day of computation), quoted at the end of the simulation output. Higher is better.

Benchmark 1 (adh_cubic):

Ensure that your GROMACS /bin directory is exported to your PATH. You should be able to type “gmx_mpi --version” and have application information displayed correctly. The first task is to preprocess the input data into a usable format, using the ‘grompp’ tool:

```
gmx_mpi grompp -f pme_verlet.mdp -c conf.gro -p topol.top -o md_0_1.tpr
```

The script “run_gromacs.sh” has been provided which simplifies the process of launching the simulation. Modify the variables to appropriate values and edit the file “hosts.txt”. Execute the script which will in turn launch the simulation using MPI and write output to the log file “gromacs_log...”. You may modify the mpirun

command to optimise performance (significantly) but in order to produce a valid result, the simulation must run for 5,000 steps. Quoted in the output as:

```
“5000 steps,     10.0 ps.“
```

You need to show a tutor the output of “gmx_mpi --version” as well as “gromacs_log...”.

QUESTION 1:
From your experience, how does the adh_cubic benchmark scale on your cluster?

_____

_____

_____

Benchmark 2 (1.5M_water):

Pre-process the input data using the ‘grompp’ command

```
gmx_mpi grompp -f pme_verlet.mdp -c out.gro -p topol.top -o md_0_1.tpr
```

Modify “run_gromacs.sh” and “hosts.txt” and run the benchmark. You may modify the mpirun command to optimise performance (significantly) but in order to produce a valid result, the simulation must run for 5,000 steps. Quoted in the output as:

```
“5000 steps,     10.0 ps.“
```

You need to show a tutor the output of “gmx_mpi --version” as well as “gromacs_log...”.

QUESTION 2:
From your experience, how does the 1.5M_water benchmark scale on your cluster?

_____

_____

_____

Part 3: Protein Visualisation

You are able to score bonus points for this tutorial by submitting a visualisation of your “adh_cubic” benchmark run. Follow the instructions below to accomplish this and show a tutor the visualisation.

You will need to work on the lab computer (or laptop) to complete this section.

Download and install the VMD visualization tool suitable for the lab PC from the link below, selecting the correct version for the operating system: eg. for a Windows machine with an Nvidia GPU, select the “Windows OpenGL, CUDA” option.

```
https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD
```

Use the WinSCP application for Windows, or the ‘scp’ command for Linux to copy the output file “md_0_1.gro” of the “adh_cubic” benchmark from your cluster to your PC. Attempting to visualise the larger “1.5M_water” simulation is not necessary and not recommended due to memory limitations of most PCs.

1. Open VMD, select “File” then “New Module...”, click “Browse...” and select your .gro file.
2. Ensure the filetype was detected as “Gromacs GRO” then click “Load”. In the main VMD window you will see that 134177 particles have been loaded. You should also see the Display window has been populated with your simulation particle data. You can manipulate the data with your mouse cursor: zoom with the mouse wheel or rotate it by dragging with the left mouse button held down. This visualisation presents a naturally occurring protein (blue/green) found in the human body, suspended in a solution of water molecules (red/white).
3. From the main VMD window, select “Graphics” then “Representations...”
4. Under “Selected Atoms”, replace “all” with “not resname SOL” and click apply. You will notice the water solution around your protein has been removed, allowing you to better examine the protein.
5. In the same window, select the dropdown “Drawing Method” and try out a few different options. Select “New Cartoon” before moving on.
6. From the main VMD window, once again select “Graphics” then “Colors”. Under “Categories”, select “Display”, then “Background”, followed by “8 white”.
7. Finally, you are ready to render a snapshot of your visualisation. From the main window, select “File” then “Render...”, ensure “Snapshot...” is selected and enter an appropriate filename. Click “Start Rendering”.

QUESTION 3:
Which compiler did you use to benchmark GROMACS? Did you compare different compilers/versions, what did you observe?

_____
_____
_____

Simulations like this are used to to develop and prototype experimental pharmaceutical drug designs. By visualising the output, researchers are able to better interpret simulation results. You can now show the output .bmp file to a tutor for bonus points.
