MPhys Project - Machine Learned Potentials For MD Simulations

University of Exeter EMPS

Natan Szczepaniak
Supervisor: Prof. Saverio Russo

Week 3 (08/06/20-12/06/20)

Contents

Day 10 - 08/06/20

10.1 Email from Mastria Rosanna

After last week's email Prof. Russo sent an email to two PHD students to ask them whether we can use a University computer for our project. Mastria Rosanna responded saying there is a computer in G16 we could use by a VPN connection.

> Dear Saverio,
>
> if the computer is switched on, you can use remote desktop through VPN connection. You need to know the user name and password (if the computer has different users) and once connected you can do everything (install and run program etc.). Another possibility is to use team-viewer for remote connection and in this case you don't need to connect through VPN. Anyway, you need the account number (I can try to find it if you need) and at least for the first time a password that appears on the team viewer window. In order to have this psw you have two possibility: (i) use remote desktop (ii) ask somebody to go in G16. If you need further information, please let me know.
>
> Kind regards,
> Rosanna >
>
> > Dear Konstantinos and Rosanna
> >
> > My 3rd year MPhys students (Nathan, Natan and Max all Cc) are working on a neural network computing model. They would like to use a sufficiently powerful computer for running the codes. In G16 we have a very powerful computer which might be able to run the codes for neural networks. Is there a way for the students to access this computer?
> >
> > Kind regards,
> > Saverio

Nathan mentioned he has experience with VPN connections so we decided he should be in charge of the setting up of our environment. I proposed we use a Virtual Machine as to not damage anything on the computer.

10.2 Meeting with Prof. Russo

At 11:30 we held a meeting with Prof. Russo to discuss the progress of the project. I outlined the process with installing and running SIMPLE-NN and the process to follow using LAMMPS. We discussed the G16 computer and came to a conclusion we will not be able to use it as it is currently off. Prof. Russo mentioned he will bring this up to Prof. Hepplestone as he is to have a meeting later on in the day. Hopefully we can get access to a University Computer designed for these purposes. We arranged another meeting for Wednesday 11:30.

Our supervisor then later sent us an email about what he discussed with Prof. Hepplestone and the use of the Universities ISCA supercomputer was brought up. He also mentioned having a meeting with Steven Hepplestone later this week. Here is the email.

> Dear all
>
> Steven Hepplestone is suggesting to look at ISCA, which is a supercomputer. He says that we don't have enough computing capability on desktops. Perhaps it might be a good idea to arrange for a meeting with all of us, included Steve. I will try to arrange it.
>
> Kind regards,
> Saverio

10.3 Installing LAMMPS - MPI (Message Passing Interface) Version

As a follow up from last week, I am trying to install and run a LAMMPS simulation on my personal desktop computer via a virtual environment Running Ubuntu 18.04 in order to link this with the output of the SIMPLE-NN code.

I managed to install the lmp_serial (Single Core LAMMPS) version which was very easy however I decided to go for the Message Passing Interface (MPI) installation which directs one process per processor core. This is harder and I am literally making it harder for myself at this point but this is worth knowing for when we move on to more powerful computers. Also I gave the virutal machine 4 cores so im going to use all of them.

I had some difficulties installing the software at first due to the the two mpi packages installed were interfering with one another. I deleted mpich2 and installed libopenmpi-dev which seemed to fix the error.

The (lammps-directory) is the installation path of lammps

```
cd (lammps-directory)/src
sudo apt autoremove mpich2
sudo apt-get install libopenmpi-dev
make mpi
```

This ran the installation perfectly and built the program.

The executable used to run simulations lmp_mpi was not added to path initially so I had to run:

export PATH=$PATH:"/home/natan/lammps - 3Mar20"

10.4 Uniaxial Compression of Aluminium

After a lot of looking around, I found an example input script for LAMMPS that simulates a uniaxial compressive loading of single crystal aluminum. (Uniaxial Compression is compression along one axis. Imagine hydraulic press). Then I ran the simulation using an example input script.

This is the input script called Al.compression2.in. This is generally what an input script for LAMMPS looks like.

```
# Input file for uniaxial compressive loading of single crystal aluminum
# Mark Tschopp, November 2010

# ------------------------ INITIALIZATION ----------------------------
units       metal
dimension   3
boundary    p   p   p
atom_style  atomic
variable latparam equal 4.05

# ------------------------ ATOM DEFINITION ----------------------------
lattice     fcc ${latparam}
region      whole block 0 10 0 10 0 10
create_box  1 whole

region      upper block INF INF INF INF INF INF units box
lattice     fcc ${latparam} orient x 1 0 0 orient y 0 1 0 orient z 0 0 1
create_atoms    1 region upper

# ------------------------ FORCE FIELDS ------------------------------
pair_style  eam/alloy
pair_coeff  * * Al99.eam.alloy Al

# ------------------------ SETTINGS ----------------------------------
compute csym all centro/atom fcc
compute peratom all pe/atom

#######################################
# EQUILIBRATION
reset_timestep  0
timestep 0.001
velocity all create 300 12345 mom yes rot no
fix 1 all npt temp 300 300 1 iso 0 0 1 drag 1

# Set thermo output
thermo 1000
thermo_style custom step lx ly lz press pxx pyy pzz pe temp

# Run for at least 10 picosecond (assuming 1 fs timestep)
run 20000
unfix 1

# Store final cell length for strain calculations
variable tmp equal "lx"
variable L0 equal ${tmp}
print "Initial Length, L0: ${L0}"

#######################################
# DEFORMATION
reset_timestep  0
```

```
fix     1 all npt temp 300 300 1 y 0 0 1 z 0 0 1 drag 1
variable srate equal 1.0e10
variable srate1 equal "-v_srate / 1.0e12"
fix     2 all deform 1 x erate ${srate1} units box remap x

# Output strain and stress info to file
# for units metal, pressure is in [bars] = 100 [kPa] = 1/10000 [GPa]
# p2, p3, p4 are in GPa
variable strain equal "(lx - v_L0)/v_L0"
variable p1 equal "v_strain"
variable p2 equal "-pxx/10000"
variable p3 equal "-pyy/10000"
variable p4 equal "-pzz/10000"
fix def1 all print 100 "${p1} ${p2} ${p3} ${p4}" file Al_comp_100.def1.txt scree
n no

# Use cfg for AtomEye
dump        1 all cfg 250 dump.comp_*.cfg mass type xs ys zs c_csym c_peratom fx
fy fz
dump_modify     1 element Al

# Display thermo
thermo  1000
thermo_style    custom step v_strain temp v_p2 v_p3 v_p4 ke pe press

run     20000

########################################
# SIMULATION DONE
print "All done"
```

This would have worked however I got an error that the pair_style was not found/working. Also this script is made for Atomeye as seen in the "dump" section and it would have to be modified to work with Ovito.

After some more reading, i found that this script was from 2017 when LAMMPS worked a little bit differently and this script would no longer work without other adjustments. Basically the specific pair_style package used in this input cell was not installed and I was not able to install the package.

10.5 Installing Packages in LAMMPS

After some browsing around, i found the way to install the needed packages for the script. The way to do this would be to run a

```
make yes-[insert package name]
```

command and then rebuild the software using

```
make mpi
```

There is also a thing called CMake however I am yet to figure out how that works.

10.6 Planning for future of project

In today's meeting we also discussed future aims and milestones we aim to achieve by different stages of our project. The initial goal of these 3 weeks was to find a topic to research as our previous one was discarded and try to use someone elses implementation to demonstrate what we would like to recreate.

So far this has proven quite ambitious however I got very close to using the SIMPLE-NN code in order to simulate a SiO structure however I will not have enough time to do this. Instead, at end of these 13 working days days I would like to have run a NNP package sucessfully and attempted to run a simulation in LAMMPS (former has already been achieved). Integrating the two may or may not happen.

Taking the project onto next year, we will look at different implementations similar to SIMPLE-NN and possibly different ML methods. Once we determine a method we are comfortable with we will use this method in order to boost the computing efficiency of a chosen molecule of interest hopefully from the departament. We will try to focus on specific properties of these molecules/materials in order to give the project however this is to be determined and needs more research into contemporary MD research.

Day 11 - 09/06/20

11.1 ISCA Supercomputer Access Discussion

I responded to Prof. Russo about the ISCA access and meeting with Steven Hepplestone, very promptly, we were CCd in an email sent to Prof. Hepplestone about a possible meeting later this week.

> Dear Steve
>
> as you know I have a group of 3 MPhys students (Max, Nathan and Natan -Cc) who had to change the direction of their project from experimental to computing aspects of neural networks. It would be really beneficial for the students to have a brief meeting with you in which they could seek some advice from an expert in the field. I have time this week on Thursday and Friday.
>
> Kind regards,
> Saverio

11.2 ARCHER UK National Supercomputing Service LAMMPS Tutorial

To get more familiarised with LAMMPS I decided to find more content online to learn how to run simulatons with LAMMPS. As learned yesterday, a lot has changed overtime with the innerworkings of LAMMPS which has resulted in some things being broken and not working properly such as the pair_coeffs parameter file. I found a good up-to date tutorial on LAMMPS from ARCHER UK which is UK's national supercomputing service.

The link to the tutorial and included files:
http://www.archer.ac.uk/training/courses/2019/10/lammps/index.php
(http://www.archer.ac.uk/training/courses/2019/10/lammps/index.php)

11.2.1 LAMMPS input script

The input script is what tells LAMMPS what to simulate, how to simulate it and for how long to simulate it.

General Structure:

1. **Set up simulation box**
2. **Defining interparticle interactions (Interatomic potentials) <-- Our area of interest**
3. **Define neighbour lists (cutoff etc.)**
4. **Define simulation parameters (Rules of the environment)**
5. **Final setup**

Each one of these sections has their own variables and functions that need to be specified. I will now compile of my understanding of these from the video in a text format for later reference. This is not an exhaustive list however

1. Set up simulation box

- units - *Units used in the simulation eg. "lj" (reduced)*

- atom_style - *types of particles being simulated eg. "atomic", "dipole" etc*
- dimension - *Simple, number of dimensions 1,2 or 3*
- boundary - *types of boundary conditons for the system eg. periodic, non-periodic or hybrid.(eg periodic is useful for bulk simulations)*
- lattice - *Types of lattice eg. sc (simple cubic) or fcc (face centered cubic) together with the lattice constant*
- region - *type of environment and the size*
- create_box - *Creates box and specifies where*
- create_atoms - *Creates atoms and where to place them, you can create multiple atoms of different species*

Note: setting up the simulation box can be replaced with an input file. These input files can be made using packages using "packmol".

2. Defining interparticle interactions

- pair_style - *Type of potential being used together with a manual cutoff*
- pair_modify - *Changes the potential eg. shifting it rotating it*
- pair_coeff - *This specifies the particle interactions first two numbers signify the types of atoms specified earlier and the next two are particle diameter σ and interparticle energy ϵ*
- mass - *Mass of each specific particle in the system. First number is particle type and the other one is mass*

Note: There are preinstalled pair_coeffs and pair_styles. Our project will aim to create these for a specific element.

3. Define neighbour lists

- neighbour - *Specifies the cut off distance at which particles will not consider another particle.*
- neighbour_modify - *Determines when the neigbour lists are updated. (should be each timestep)*

4. Define simulation parameters

- fix - *God option, you can control temperature, pressure, viscosity, gravity etc.*
- compute - *Compute the predetermined properties per particle.*

5. Final Setup

- velocity - *sets velocities for each particle*
- timestep - *Time between steps*
- thermo_style - *global properties that are being outputted (Energy, pressure, volume, density etc.)*
- run_style - *algorithm to run the simulation*
- run - *number of steps the simulation will run for*

11.2.2 MD Flowchart

The general flow of an MD simulation is the workflow:

```
1. From known velocities and forces, update positions
2. From known positions, update energies and forces
3. From known forces, update velocities
4. Go back to 1
```

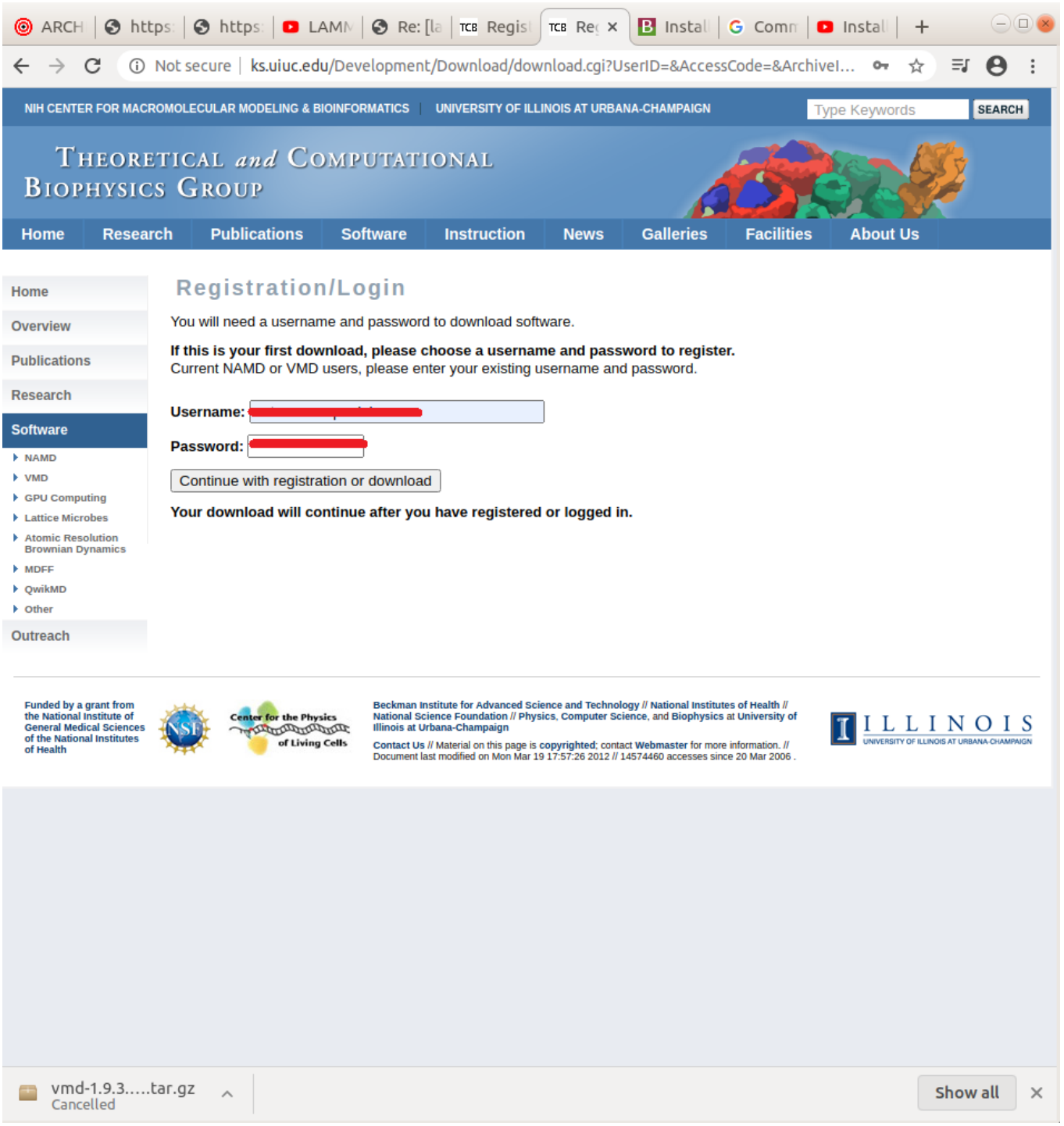
This can start at any number. The number of times this is done is the number of "steps"

After this there is also the "dump" command which outputs some of these into a file for visualisation/further analysis.

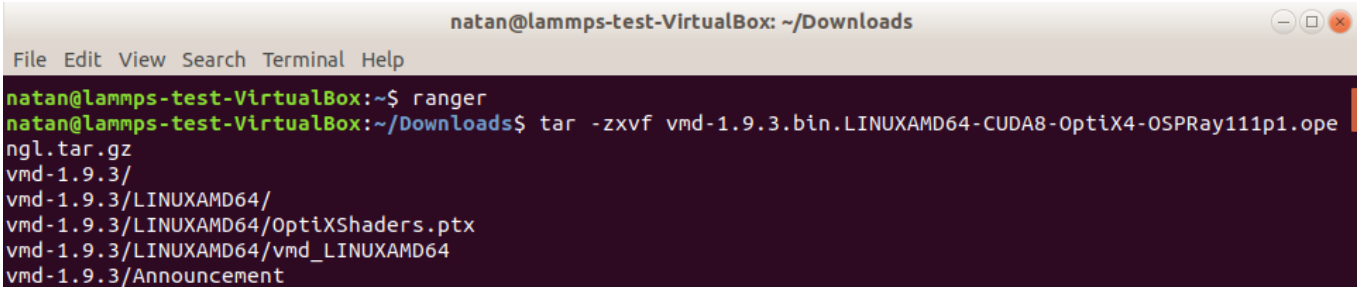
Day 12 - 10/06/20

12.1 Visualisation software

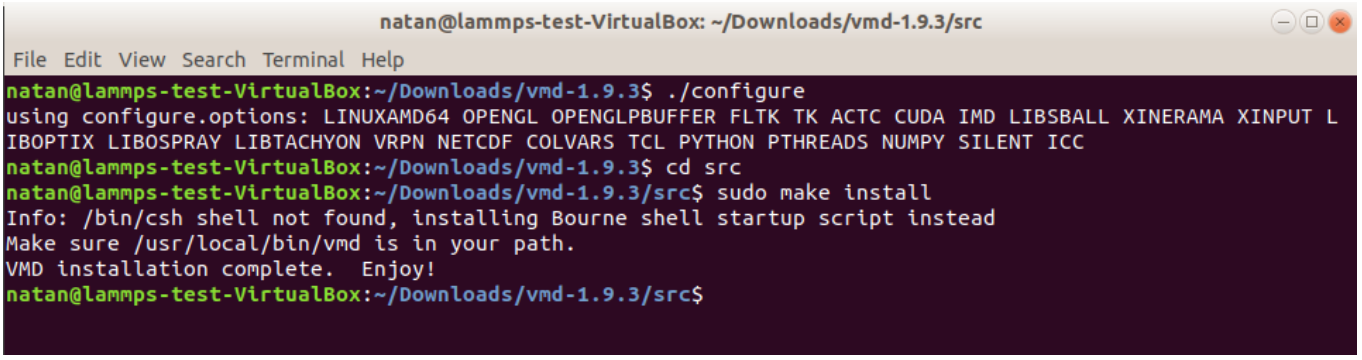
In order to visualise the simulation output generated using LAMMPS, i needed visualisation software. There are multiple including previously mentioned OVITO. On Ubuntu however, the most commonly used and reliable one seems to be VMD. Firstly, I had to download the package from the https://www.ks.uiuc.edu/ (https://www.ks.uiuc.edu/) website. I downloaded the Open-GL version after registering onto their website.



Then I had to extract the .tar into the installation directory.



Then run the ./configure command, cd into the src folder and install.



After this, I had to add vmd to path. This was now ready for input files.

12.2 Running the an LJ Fluid Simulation using LAMMPS

The tutorial from yesterday provided a sample exercise input script for LAMMPS that simulates a LJ potential fluid for particles of same species. The input script looks like this:

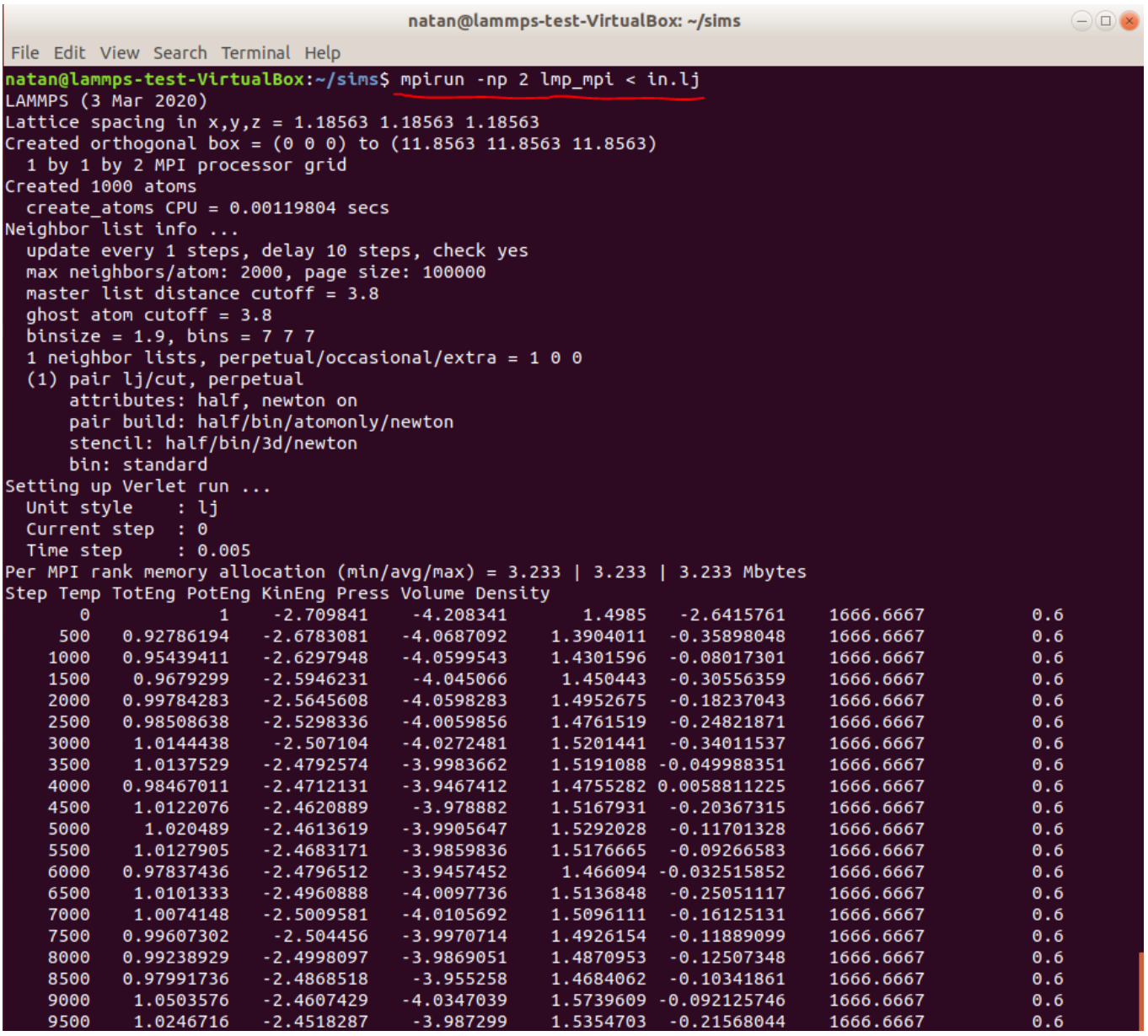
```
####################################
# Example LAMMPS input script      #
# for a simple Lennard Jones fluid #
####################################

####################################
# 1) Set up simulation box
#   - We set a 3D periodic box
#   - Our box has 10x10x10 atom
#     positions, evenly distributed
#   - The atom starting sites are
#     separated such that the box density
#     is 0.6
####################################

units       lj
atom_style  atomic
dimension   3
boundary    p p p

lattice     sc 0.60
region      box block 0 10 0 10 0 10
create_box  1 box
create_atoms    1 box

####################################
# 2) Define interparticle interactions
#   - Here, we use truncated & shifted LJ
#   - All atoms of type 1 (in this case, all atoms)
#     have a mass of 1.0
####################################

pair_style  lj/cut 3.5
pair_modify shift yes
pair_coeff  1 1 1.0 1.0
mass        1 1.0

####################################
# 3) Neighbour lists
#   - Each atom will only consider neighbours
#     within a distance of 2.8 of each other
#   - The neighbour lists are recalculated
#     every timestep
####################################

neighbor        0.3 bin
neigh_modify    delay 10 every 1

####################################
# 4) Define simulation parameters
#   - We fix the temperature and
```

```
#     linear and angular momenta
#     of the system
#   - We run with fixed number (n),
#     volume (v), temperature (t)
####################################

fix     LinMom all momentum 50 linear 1 1 1 angular
fix     1 all nvt temp 1.00 1.00 5.0
#fix        1 all npt temp 1.0 1.0 25.0 iso 1.5150 1.5150  10.0

####################################
# 5) Final setup
#   - Define starting particle velocity
#   - Define timestep
#   - Define output system properties (temp, energy, etc.)
#   - Define simulation length
####################################

velocity    all create 1.0 199085 mom no

timestep    0.005

thermo_style    custom step temp etotal pe ke press vol density
thermo      500

run_style   verlet

run     50000

####################################
# 6) Advanced inputs
####################################

#dump       2 all custom 1000 positions.lammpstrj id x y z vx vy vz
#dump_modify    2 sort id

#compute        RDF all rdf 150 cutoff 3.5
#fix        RDF_OUTPUT all ave/time 25 100 5000 c_RDF[*] file rdf_lj.out mode ve
ctor

#compute        MSD all msd
#fix        MSD_OUTPUT all ave/correlate 1 5000 5000 c_MSD[*] file msd_lj.out av
e running

#compute        VACF all vacf
#fix        VACF_OUTPUT all ave/correlate 1 2500 5000 c_VACF[*] file vacf_lj.out
ave running

#run        5000

#write_restart  restart2.lj.equil
```

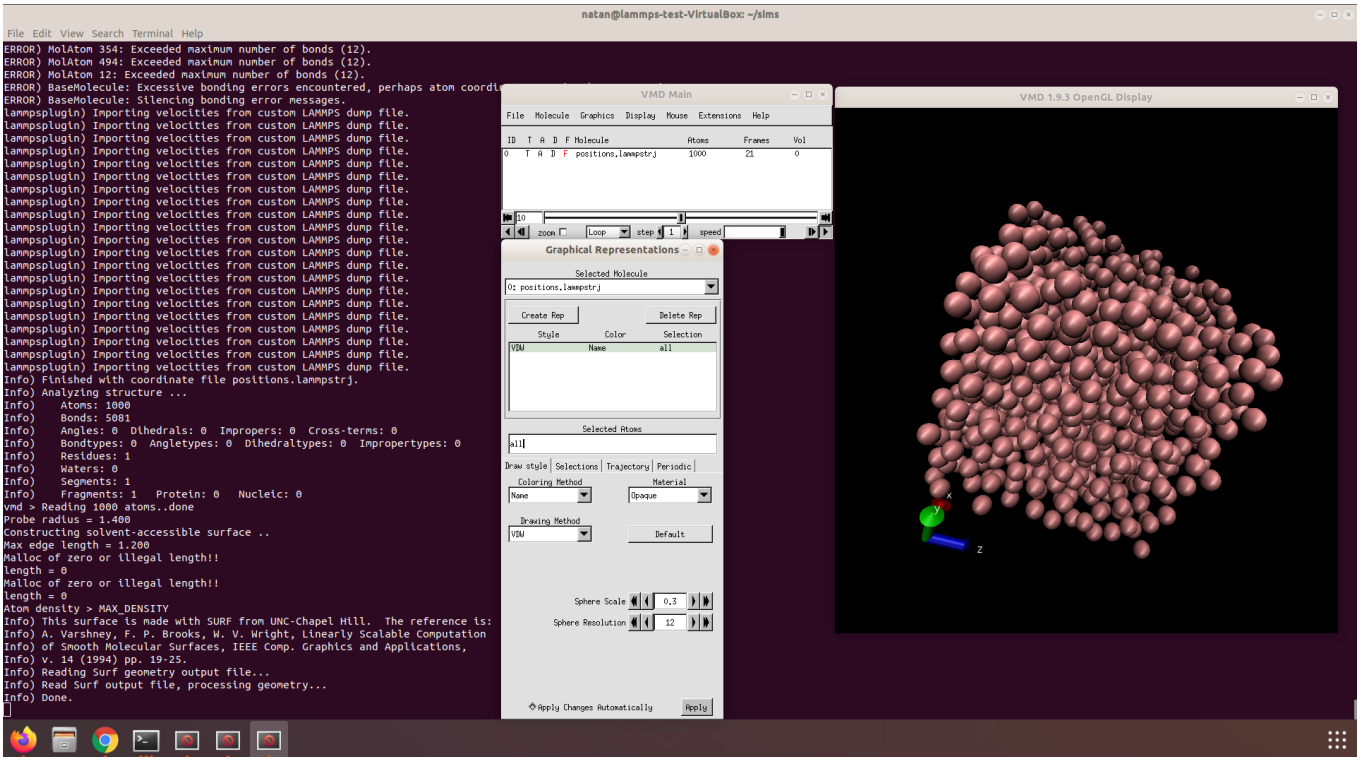
After uncommenting the "dump" command and changing the run time, I fed this input script into LAMMPS using this command from the terminal.

```
mpirun -np 2 lmp_mpi < in.lj
```

where mpirun is the mpi manager, -np specifies the number of processors used, lmp_mpi is the LAMMPS executable and in.lj is the input script. This ran the code sucessfully.



Next, the file that was outputted by LAMMPS "positions.lammpstrj" I took and run it into vmd to visualise the simulation. Of course there are other properties that can be extracted but this is a nice visual reminder that progress has been made.



Day 13 - 11/06/20

13.1 Finalising 3 Weeks

Today was used to fully go through and edit the Laboratory Book for assessment as well as finding a good way to export it to a PDF file without losing out too much on the formatting, making sure the images work and making sure the equations work.

I also caught up with my PHD colleague in the field to ask him some more questions I had and update him with my progress. He was quite impressed with the amount of work I have done and said he can help me integrate the SIMPLE-NN potential into LAMMPS as he would like to learn about the software also.

Then I prepared a summary in order to make the flow of the viva easier.

The method I used for exporting the lab book out of Jupyter Notebook is downloading it as a HTML, moving it to the file where my img folder is located and converting it to a PDF. Not ideal since some things are still not perfect but this will do. I also tried to upload this project onto github into a private repository however it did not convert .pynb into .md properly.