Table of Contents – OCF User Manual – Maxwell HPC Cluster

1. OCF Document Control........................................................................................................................................ 4
   1.1 Latest ............................................................................................................................................................ 4
   1.2 History ........................................................................................................................................................... 4
   1.3 Notes ............................................................................................................................................................. 4

2. Cluster Summary .................................................................................................................................................. 5

3. Cluster Resources ................................................................................................................................................ 6
   3.1 Physical Resources ........................................................................................................................................ 6
   3.2 Storage Resources ........................................................................................................................................ 6
   3.3 Slurm Resources ........................................................................................................................................... 7
   3.4 Slurm Partition Priorities .............................................................................................................................. 8

4. Cluster Access .................................................................................................................................................. 9
   4.1 University of Aberdeen Users ..................................................................................................................... 10
      4.1.1 Command Line ................................................................................................................................... 10
      4.1.2 Graphical ........................................................................................................................................... 12
   4.2 National Decommissioning Centre Users ................................................................................................... 13
      4.2.1 Command Line ................................................................................................................................... 13
      4.2.2 Graphical ........................................................................................................................................... 13
   4.3 File Transfers ............................................................................................................................................... 14
      4.3.1 Standard Linux Utilities ...................................................................................................................... 14
      4.3.2 WinSCP ............................................................................................................................................. 15
      4.3.3 Remote Cluster Access ........................................................................................................................ 16
         4.3.3.1 Command Line Access ................................................................................................................... 16
         4.3.3.2 Graphical ....................................................................................................................................... 17
         4.3.3.3 File Transfers .................................................................................................................................... 18

5. Cluster Applications .......................................................................................................................................... 19
   5.1 Loadable Applications .................................................................................................................................. 19
   5.2 Installed Applications ................................................................................................................................... 19
   5.3 Galaxy ......................................................................................................................................................... 21
6. Workload Manager - Slurm ................................................................................................................ 23
   6.1 Common Slurm Commands ........................................................................................................... 23
   6.2 Common Slurm Parameters ........................................................................................................... 26
   6.3 Submitting Compute Jobs ............................................................................................................ 27
   6.4 Submitting GPU Jobs .................................................................................................................. 29

7. Getting Help .................................................................................................................................. 31

8. Index ........................................................................................................................................... 32
   8.1 Tables ......................................................................................................................................... 32
   8.2 Images ....................................................................................................................................... 32
   8.3 Code Samples ............................................................................................................................. 33
1. OCF Document Control

1.1 Latest

<table>
<thead>
<tr>
<th>Company:</th>
<th>University of Aberdeen</th>
<th>OCF Deal Number:</th>
<th>4755</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date Submitted:</td>
<td>16/10/2019</td>
<td>Last Updated by:</td>
<td>Matthew Brazier</td>
</tr>
<tr>
<td>Document Owner:</td>
<td>Chris Devine</td>
<td>Version No:</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 1 - version control - latest

1.2 History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Author</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>11/10/19</td>
<td>Matthew Brazier</td>
<td>Initial version</td>
</tr>
<tr>
<td>0.2</td>
<td>15/10/19</td>
<td>Matthew Brazier</td>
<td>Add GPU placement table, added support section and corrected a few minor spelling mistakes.</td>
</tr>
<tr>
<td>0.3</td>
<td>16/10/19</td>
<td>Matthew Brazier</td>
<td>Added remote cluster access sections</td>
</tr>
</tbody>
</table>

Table 2 - version control - history

1.3 Notes

This document will refer to University of Aberdeen as UoA and National Decommissioning Centre as NDC.
2. Cluster Summary

The cluster is built up of Lenovo hardware. The cluster has a 10G management backbone and a 100G InfiniBand high speed interconnect between each node. The cluster is accessible using command line or graphical interfaces.

The cluster comprises of 30 nodes with a mixture of different memory capacities. Six nodes are equipped with consumer grade GPUs suitable for single precision workloads and one node is equipped with an enterprise grade GPU optimised for double precision workloads.

Total cluster nodes: 30 nodes comprising of Lenovo ThinkSystem SD530, Lenovo ThinkSystem SR650 and Lenovo ThinkSystem SR850
Total cluster cores: 1240 using Intel Xeon Gold 6138 2.00Ghz processors
Total cluster memory: 10.5T using TruDDR4 2666MHz RDIMM
Total cluster storage: 991T block storage using 8T and 10T 7.2K near line SAS disk drives
Cluster interconnect: Mellanox InfiniBand HDR 100G
High performance shared file system: BeeGFS
3. Cluster Resources

3.1 Physical Resources

The cluster has the following computational resources available to users:

<table>
<thead>
<tr>
<th>Node Type</th>
<th>Node Name</th>
<th>Cores</th>
<th>Useable Memory</th>
<th>GPU Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Memory</td>
<td>hmem[01-11]</td>
<td>40 (2x Intel Xeon Gold 6138 2.00Ghz)</td>
<td>380G</td>
<td>2 x NVIDIA 2080Ti (12G GPU Memory)</td>
</tr>
<tr>
<td>Consumer GPU</td>
<td>cgpu[01-11]</td>
<td>40 (2x Intel Xeon Gold 6138 2.00Ghz)</td>
<td>185G</td>
<td></td>
</tr>
<tr>
<td>Compute</td>
<td>node[01-11]</td>
<td>40 (2x Intel Xeon Gold 6138 2.00Ghz)</td>
<td>185G</td>
<td></td>
</tr>
<tr>
<td>Very High Memory</td>
<td>vhmem01</td>
<td>80 (4x Intel Xeon Gold 6138 2.00Ghz)</td>
<td>3000G</td>
<td></td>
</tr>
<tr>
<td>Enterprise GPU</td>
<td>egpu01</td>
<td>40 (2x Intel Xeon Gold 6138 2.00Ghz)</td>
<td>185G</td>
<td>2 x NVIDIA V100 (32G GPU Memory)</td>
</tr>
</tbody>
</table>

Table 3 - cluster resources - physical resources

All nodes are running CentOS 7.6\(^1\) and have 1 x Mellanox InfiniBand HDR 100G.

3.2 Storage Resources

The cluster has the following storage resources available to users:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Storage Limit</td>
<td>File Limit</td>
</tr>
<tr>
<td>UoA</td>
<td>64G</td>
<td>200K</td>
</tr>
<tr>
<td>NDC</td>
<td>64G</td>
<td>200K</td>
</tr>
</tbody>
</table>

Table 4 - cluster resources - storage resources

Note: These limitations are the defaults on the cluster. Higher limits could be made to specific users with appropriate justification.

\(^1\) CentOS 7.6 installed as of 10/10/2019. Operating systems will be updated on a regular basis when updates are stable.
3.3 Slurm Resources

<table>
<thead>
<tr>
<th>Primary User Organisation</th>
<th>Slurm Partition</th>
<th>Nodes</th>
<th>Notes (default)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UoA</td>
<td>uoa-compute</td>
<td>hmem[01-11]</td>
<td></td>
</tr>
<tr>
<td>UoA</td>
<td>uoa-gpu</td>
<td>cgpu[01-06]</td>
<td></td>
</tr>
<tr>
<td>NDC</td>
<td>ndc-compute</td>
<td>node[01-11]</td>
<td></td>
</tr>
<tr>
<td>NDC</td>
<td>ndc-gpu</td>
<td>egpu01</td>
<td></td>
</tr>
<tr>
<td>NDC</td>
<td>ndc-hmem</td>
<td>vhmem01</td>
<td></td>
</tr>
<tr>
<td>UoA</td>
<td>spot-compute</td>
<td>hmem[01-11] cgpu[01-11]node[01-11] egpu01 vhmem01</td>
<td>Default partition for jobs unless a partition is specified at submission.</td>
</tr>
<tr>
<td>UoA</td>
<td>spot-gpu</td>
<td>egpu01</td>
<td>Active jobs will be requeued if a job submitted to ndc-gpu requires resource.</td>
</tr>
<tr>
<td>UoA</td>
<td>spot-vhmem</td>
<td>vhmem01</td>
<td>Active jobs will be requeued if a job submitted to ndc-hmem requires resource.</td>
</tr>
<tr>
<td>UoA</td>
<td>uoa-teaching</td>
<td>node[01-05]</td>
<td>Max nodes per job = 2 Max CPUs per node = 10 Max memory per node = 20G Users must be a member of “maxwell-users-teach”</td>
</tr>
</tbody>
</table>

Table 5 - cluster resources - Slurm resources

Note: All standard user accounts will have a limit of 200 allocated cores. If this limit is reached jobs will be queued until enough resources are available.
3.4 Slurm Partition Priorities

Slurm partitions are setup with priority tiers allowing a flexible and timely approach to usage of the cluster. Jobs submitted to higher priority tiers will cause lower priority jobs to be re-queued if resources are required to run the higher priority job.

<table>
<thead>
<tr>
<th>Priority Tier</th>
<th>Partitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>uoa-teaching</td>
</tr>
<tr>
<td>5</td>
<td>ndc-hmem</td>
</tr>
<tr>
<td>4</td>
<td><em>not in use</em></td>
</tr>
<tr>
<td>3</td>
<td>uoa-gpu</td>
</tr>
<tr>
<td>2</td>
<td>spot-gpu</td>
</tr>
<tr>
<td>1</td>
<td>spot-compute</td>
</tr>
</tbody>
</table>

Table 6 - cluster resources - Slurm partition priorities

```bash
[root@maxmgmt1 ~]# sinfo
PARTITION     AVAIL  TIMELIMIT  NODES  STATE NODELIST
uoa-gpu          up   infinite      6   idle cgpu[01-06]
spot-vhmem       up   infinite      1    mix vhmem01
ndc-gpu          up   infinite      1   idle egpu01
ndc-hmem         up   infinite      1   idle egpu01
ocf-benchmark    up   infinite     22   idle hmem[01-11],node[01-11]
spot-compute*    up   infinite      1   mix vhmem01
spot-compute*    up   infinite     29   idle cgpu[01-06],egpu01,hmem[01-01],node[01-11]
spot-gpu         up   infinite      1   idle egpu01
ndc-compute      up   infinite     11   idle node[01-11]
uoa-compute      up   infinite     11   idle hmem[01-11]
uoa-teaching     up   infinite     5   idle node[01-05]
```

Code 1 - cluster resources - Slurm sinfo
4. Cluster Access

Access to the cluster will vary based on the organisation you are part of. All access to the cluster for standard users is based on your Windows credentials used to access any University computer. To access the cluster, you must be a member of one of the HPC user groups. If you need to become a member of a group or are not sure whether you are already a member speak to your local HPC administration team. UoA users must be part of the “maxwell-users” or “maxwell-users-teach” to be able to access the cluster. NDC users must be part of the “ndc-hpc-users” group to be able to access the cluster.

In addition to being able to log on to the cluster you must also have a Slurm account to be able to submit jobs for processing. Slurm accounts are setup by your local HPC administration team and they will be your first point of contact should you find you do not have a Slurm account or your account is not working. See section 6.3 - Workload Manager - Slurm - Submitting compute jobs for more detail about the usage of Slurm.

There are two dedicated login nodes for each organisation for gaining access to the cluster. All login nodes can be accessed using SSH or x2go graphical client. To access the cluster you will need an SSH client on your computer. If you are a Windows user Putty is the recommended application.

X2GO client is required for accessing the graphical client and can be installed by IT Services if not already present on your computer. Supported session types for X2GO are XFCE, KDE and MATE. Other desktop window systems are not currently installed on the cluster due to incompatibility or insufficient justification.
### Cluster Access

#### 4.1 University of Aberdeen Users

Login node addresses:

- maxlogin1.abdn.ac.uk
- maxlogin2.abdn.ac.uk

#### 4.1.1 Command Line

SSH access to the cluster is the most straightforward from a Linux/Mac/Terminal emulator.

```
[ocf-engineer@OCF ~]$ ssh USERNAME@maxlogin1.abdn.ac.uk
USERNAME@maxlogin1.abdn.ac.uk's password:
[USERNAME@maxlogin1 ~]$  
```

**Code 2 - cluster access - UoA maxlogin1**

```
[ocf-engineer@OCF ~]$ ssh USERNAME@maxlogin1.abdn.ac.uk
USERNAME@maxlogin1.abdn.ac.uk's password:
[USERNAME@maxlogin1 ~]$  
```

**Code 3 - cluster access - UoA maxlogin2**
Cluster Access

If using Putty simply open Putty and enter the desired login node address into the Host Name box.

Image 1 - cluster access - UoA Putty

Image 2 - cluster access - UoA Putty logged in
4.1.2 Graphical

The following X2GO session server options are required to connect to a graphical instance. Remember the session type must be XFCE, KDE or MATE.

Host: maxlogin1.abdn.ac.uk or maxlogin2.abdn.ac.uk
Login: Your username
SSH Port: 22

You will be prompted for your password after the session has been setup.
Cluster Access

4.2 National Decommissioning Centre Users

Login node addresses:

- ndc1.abdn.ac.uk
- ndc2.abdn.ac.uk

4.2.1 Command Line

SSH access to the cluster is the most straightforward from a Linux/Mac/Terminal emulator.

```
[ocf-engineer@OCF ~]$ ssh USERNAME@ndc1.abdn.ac.uk
USERNAME@ndc1.abdn.ac.uk's password:
[USERNAME@ndc1 ~]$
```

Code 5 - cluster access - NDC ndc1

```
[ocf-engineer@OCF ~]$ ssh USERNAME@ndc2.abdn.ac.uk
USERNAME@ndc2.abdn.ac.uk's password:
[USERNAME@ndc2 ~]$
```

Code 6 - cluster access - NDC ndc2

*Please refer to section 4.1.1 Putty for Putty screenshots.*

4.2.2 Graphical

The following X2GO session server options are required to connect to a graphical instance. Remember the session type must be XFCE, KDE or MATE.

```
Host: ndc1.abdn.ac.uk or ndc2.abdn.ac.uk
Login: Your username
SSH Port: 22
```

Code 7 - cluster access - NDC x2go server settings

*Please refer to section 4.1.2 for x2go client screenshots.*
Cluster Access

4.3 File Transfers

Data can be transferred in and out of the cluster using standard Linux tools, such as "scp" or "rsync". Windows users can use WinSCP to transfer data using a graphical application.

Remember the hostname will be one of the following:

- maxlogin1.abdn.ac.uk
- maxlogin2.abdn.ac.uk
- ndc1.abdn.ac.uk
- ndc2.abdn.ac.uk

4.3.1 Standard Linux Utilities

SCP or rsync directly to your home directory and then work on the data when it is on the cluster.

```
scp jobdata.tar.gz USERNAME@maxlogin1:~/
rsync -avzh ./jobName/ USERNAME@maxlogin1:~/
```

Code 8 - cluster access - scp and rsync
4.3.2 WinSCP

Login to the cluster using the same connection details as you would use for logging into the command line or graphical interface. Once logged in you will be able to drag and drop files to the cluster.
Cluster Access

4.3.3 Remote Cluster Access

Access to the cluster can be obtained from outside the University allowing you to access the cluster from non-university computers or when not connected to university WiFi. Remote access to the cluster must be enabled for your account by the HPC team and does not come as part of a standard user account.

Remote access is possible via any of the access methods described above and will require a proxy or jump host to be specified when connecting to the cluster. The proxy connection must use:

- Hostname: ssh-gateway.abdn.ac.uk
- Port: 22
- Proxy/connection type: ssh tunnel
- Credentials: same as you would use to login to the cluster from within the University network

4.3.3.1 Command Line Access

Log on to the SSH gateway prior to logging on to your login nodes as per sections 4.1.1 and 4.2.1.

```
[ocf-engineer@OCF ~]$ ssh USERNAME@ssh-gateway.abdn.ac.uk
-sh-4.2$ ssh maxlogin1.abdn.ac.uk
[USERNAME@ndc1 ~]$
```

Code 9 - cluster access - remote command line access
Cluster Access

4.3.3.2 Graphical

Image 9 - cluster access - remote x2go
4.3.3.3 File Transfers

Image 10 - cluster access - SSH tunnel WinSCP
5. Cluster Applications

5.1 Loadable Applications

All applications available on the cluster are installed as loadable modules. Loadable modules allow a vast array of software to be installed on the cluster without causing compatibility between applications that have dependencies on the same applications but of different versions.

Available modules can be displayed using “module avail”. At the time of writing the following applications are available on the cluster.

5.2 Installed Applications

See https://www.abdn.ac.uk/it/documents/Maxwell-Software.pdf for a list of applications and the “Loadable Module” identifier. Modules are loaded by issuing the command “module load <Loadable Module>”.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>module load &lt;Loadable Module&gt;</td>
<td>Load a module</td>
<td>module load intel-19.0.4.243</td>
</tr>
<tr>
<td>module list</td>
<td>List loaded modules</td>
<td>module list</td>
</tr>
<tr>
<td>module unload &lt;Loadable Module&gt;</td>
<td>Unload a loaded module</td>
<td>module unload intel-19.0.4.243</td>
</tr>
<tr>
<td>module spider &lt;search term&gt;</td>
<td>Search available modules</td>
<td>module spider intel</td>
</tr>
</tbody>
</table>
Cluster Applications

By default “slurm/19.05.1-2” or a newer version will be loaded and must be loaded to allow jobs to be submitted to Slurm.

```
[root@maxlogin1 ~]# module load intel-19.0.4.243
[root@maxlogin1 ~]# module list
Currently Loaded Modules:
  1) slurm/19.05.1-2   2) intel-19.0.4.243
[root@maxlogin1 ~]# module spider intel
----------
intel:
----------
Versions:
  intel/imkl-2019
  intel/impi-2019
  intel/2017.3
  intel/2019.3
Other possible modules matches:
  ffmpeg-4.1.1-intel-19.0.4.243 intel-19.0.4.243
  intel-mpi-09.01.04.03r-ce intelconda3 mothur-1.39.5-intel-19.0.4.243
  mothur-1.40.5-intel-19.0.4.243 netcdf-4.7.0-intel-19.0.4.243 ...
----------
To find other possible module matches execute:

   $ module -r spider '.*intel.*'
----------
For detailed information about a specific "intel" module (including how to load the modules) use the module's full name.
For example:

   $ module spider intel/impi-2019
----------
[root@maxlogin1 ~]# module unload intel-19.0.4.243
[root@maxlogin1 ~]# module list
Currently Loaded Modules:
  1) slurm/19.05.1-2
```
5.3 Galaxy

A Galaxy server is available on Maxwell. The Galaxy server uses the same login credentials as you would use to login to the command line or graphical desktop. Galaxy server is accessible by visiting:

https://maxgalaxy.abdn.ac.uk
Cluster Applications

Galaxy computation submits jobs to the Slurm cluster. All data used by Galaxy must be available on the cluster the same as any other Slurm job run on the cluster.
6. Workload Manager - Slurm

The cluster resources are managed by SchedMD Slurm. All computational jobs run on the cluster must be submitted via Slurm. There is no direct access to nodes jobs are not running on. Job data must be present in your home or scratch directory depending on dataset size. Jobs must be submitted via a login node. The environment on the login node is the same as what must be loaded on a compute node. For example, if you load a module using "module load" this must be in your submission script so it can be loaded by the node running the job.

6.1 Common Slurm Commands

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>sbatch [args] ./script</td>
<td>Slurm batch script submission</td>
<td>sbatch -p spot-compute ./script.sh</td>
</tr>
<tr>
<td>squeue</td>
<td>Show job queue</td>
<td>squeue</td>
</tr>
<tr>
<td>scontrol show partition &lt;name&gt;</td>
<td>Show partition details</td>
<td>scontrol show partition spot-compute</td>
</tr>
<tr>
<td>scontrol show job &lt;ID&gt;</td>
<td>Show job information</td>
<td>scontrol show job 205295</td>
</tr>
<tr>
<td>sinfo</td>
<td>Show summary information</td>
<td>sinfo</td>
</tr>
<tr>
<td>scancel &lt;ID&gt;</td>
<td>Cancel a running job</td>
<td>scancel 205295</td>
</tr>
</tbody>
</table>
Workload Manager - Slurm

[root@maxlogin1 ~]# squeue

<table>
<thead>
<tr>
<th>JOBID</th>
<th>PARTITION</th>
<th>NAME</th>
<th>USER</th>
<th>ST</th>
<th>TIME</th>
<th>NODES</th>
<th>NODELIST(REASON)</th>
</tr>
</thead>
<tbody>
<tr>
<td>205295</td>
<td>uoa-gpu</td>
<td>fmnist_c</td>
<td>USER1</td>
<td>R</td>
<td>19:31</td>
<td>1</td>
<td>cgpu01</td>
</tr>
<tr>
<td>205294</td>
<td>spot-gpu</td>
<td>fmnist_c</td>
<td>USER1</td>
<td>R</td>
<td>19:50</td>
<td>1</td>
<td>egpu01</td>
</tr>
<tr>
<td>205298</td>
<td>spot-comp</td>
<td>MR125P5x</td>
<td>USER2</td>
<td>R</td>
<td>12:37</td>
<td>1</td>
<td>vhmem01</td>
</tr>
<tr>
<td>205297</td>
<td>spot-comp</td>
<td>MaR4P5nX</td>
<td>USER2</td>
<td>R</td>
<td>16:01</td>
<td>1</td>
<td>vhmem01</td>
</tr>
<tr>
<td>205296</td>
<td>spot-comp</td>
<td>MR125P4x</td>
<td>USER2</td>
<td>R</td>
<td>19:01</td>
<td>1</td>
<td>vhmem01</td>
</tr>
<tr>
<td>205291</td>
<td>spot-comp</td>
<td>MaR1P3nX</td>
<td>USER2</td>
<td>R</td>
<td>25:01</td>
<td>1</td>
<td>vhmem01</td>
</tr>
</tbody>
</table>

Code 11 - workload manager - squeue

[root@maxlogin1 ~]# scontrol show partition spot-compute

PartitionName=spot-compute
AllowGroups=maxwell-users,OCF-Admin-Users,galaxy AllowAccounts=ALL AllowQoS=ALL
AllocNodes=ALL Default=YES QoS=N/A
DefaultTime=41-15:00:00 DisableRootJobs=NO ExclusiveUser=NO GraceTime=0 Hidden=NO
MaxNodes=UNLIMITED MaxTime=UNLIMITED MinNodes=0 LLN=NO MaxCPUsPerNode=UNLIMITED
Nodes=node[01-11],cgpu[01-06],hmem[01-11],vhmem[01-01],egpu[01-01]
PriorityJobFactor=1 PriorityTier=1 RootOnly=NO PreemptMode=REQUEUE
State=UP TotalCPUs=1240 TotalNodes=30 SelectTypeParameters=NONE
JobDefaults=(null)
DefMemPerNode=UNLIMITED MaxMemPerNode=UNLIMITED

Code 12 - workload manager - scontrol show partition
Workload Manager - Slurm

```
[root@maxmgmt1 H2O-64]# scontrol show job 205290
JobId=205290 JobName=cp2k_h20_64_1_node
    UserId=root(0) GroupId=root(0) MCS_label=N/A
    Priority=6013 Nice=0 Account=root QOS=normal
    JobState=RUNNING Reason=None Dependency=(null)
    Requeue=1 Restarts=0 BatchFlag=1 Reboot=0 ExitCode=0:0
    RunTime=00:00:07 TimeLimit=41-15:00:00 TimeMin=N/A
    AccrueTime=2019-10-11T16:11:21
    StartTime=2019-10-11T16:11:36 EndTime=2019-11-22T06:11:36 Deadline=N/A
    PreemptEligibleTime=2019-10-11T16:11:36 PreemptTime=None
    Partition=spot-compute AllocNode:Sid=maxmgmt1:204311
    ReqNodeList=(null) ExcNodeList=(null)
    NodeList=cgpu01
    BatchHost=cgpu01
    NumNodes=1 NumCPUs=36 NumTasks=36 CPUs/Task=1 ReqB:S:C:T=0:0:*:*
    TRES=cpu=36,mem=150G,node=1,billing=36
    Socks/Node=* NtasksPerN:B:S:C=36:0:*:* CoreSpec=*
    MinCPUsNode=36 MinMemoryNode=150G MinTmpDiskNode=0
    Features=(null) DelayBoot=00:00:00
    OverSubscribe=OK Contiguous=0 Licenses=(null) Network=(null)
    Command=./h2o_64_1_node.sh
    WorkDir=/opt/software/benchmarking/cp2k/inputs/H2O-64
    StdErr=/opt/software/benchmarking/cp2k/inputs/H2O-64/slurm.205290.err
    StdIn=/dev/null
    StdOut=/opt/software/benchmarking/cp2k/inputs/H2O-64/slurm.205290.out
    Power=
```

Code 13 - workload manager - scontrol show job

```
[root@maxlogin1 ~]# sinfo
PARTITION   AVAIL TIMELIMIT NODES  STATE NODELIST
  uoa-gpu    up   infinite  1   mix  cgpu01
  uoa-gpu    up   infinite  5   idle cgpu[02-06]
  spot-vhmem up   infinite  1   mix  vhmem01
  ndc-gpu    up   infinite  1   mix  egpu01
  ndc-hmem   up   infinite  1   mix  vhmem01
  ocf-benchmark up  infinite  22 idle hmem[01-11],node[01-11]
  spot-compute* up  infinite  3   mix cgpu01,egpu01,vhmem01
  spot-compute* up  infinite  27 idle cgpu[02-06],hmem[01-11],node[01-11]
  spot-gpu    up   infinite  1   mix  egpu01
  ndc-compute up   infinite  11 idle node[01-11]
  uoa-compute up   infinite  11 idle hmem[01-11]
  uoa-teaching up   infinite  5   idle node[01-05]
```

Code 14 - workload manager - sinfo
# Workload Manager - Slurm

## 6.2 Common Slurm Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>-N</td>
<td>--nodes</td>
<td>Number of nodes to run a job on</td>
</tr>
<tr>
<td>-n</td>
<td>--ntasks</td>
<td>Number of tasks to allocate</td>
</tr>
<tr>
<td>--mem</td>
<td></td>
<td>Amount of memory required specified as M, G or T</td>
</tr>
<tr>
<td>--mail-user</td>
<td></td>
<td>Send an email upon a state change</td>
</tr>
<tr>
<td>-o</td>
<td>--output</td>
<td>Output file name</td>
</tr>
<tr>
<td>-e</td>
<td>--error</td>
<td>Error file</td>
</tr>
<tr>
<td>%%j</td>
<td></td>
<td>Job ID</td>
</tr>
<tr>
<td>%%s</td>
<td></td>
<td>Step ID</td>
</tr>
<tr>
<td>%%J</td>
<td></td>
<td>Job ID.state ID (eg 2001.0)</td>
</tr>
<tr>
<td>%%x</td>
<td></td>
<td>Job name</td>
</tr>
<tr>
<td>--gres=TYPE:X</td>
<td></td>
<td>Generic resources - GPUs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Type=gtx2080ti or v100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X = 1 or 2</td>
</tr>
</tbody>
</table>

**Table 7 - Workload manager - common Slurm parameters**

All commands are available on [https://slurm.schedmd.com/sbatch.html](https://slurm.schedmd.com/sbatch.html)
6.3 Submitting Compute Jobs

The basic steps to submitting a compute job are:

1. Determine what application you are using
2. Write a job script similar to below. The job script will need to contain:
   a. Number of nodes
   b. Number of tasks or cores
   c. Tasks per node
   d. Amount of memory
3. Copy your data set to your home directory or scratch directory
4. Submit the job using “sbatch”
5. After submission you will get a job ID. The job ID can be used to monitor the job using “squeue” or “scontrol”. Alternatively, if you specified “--mail-user” you will receive an email when the job begins and ends

Here is an example compute Slurm script.

```
#!/bin/bash
#SBATCH -N 1 # number of nodes
#SBATCH -n 36 # number of cores
#SBATCH --mem 150G # memory pool for all cores
#SBATCH -o slurm.%j.out # STDOUT
#SBATCH -e slurm.%j.err # STDERR
#SBATCH --ntasks-per-node=36
#SBATCH --job-name=cp2k_h20_64_1_node

cd ${SLURM_SUBMIT_DIR}
module load openmpi/3.1.4-intel-19.0.4.243-ocf
module load cp2k/cp2k-6.1-gcc-4.8.5-vsc5ptz
mpirun --allow-run-as-root cp2k.popt -i ./H2O-64.inp >
   ./ocf_uat_1.6_hmem_cp2k_h20_64.out-$(SLURM_JOB_ID)
```

Code 15 - workload manager - Slurm script

```
[root@maxmgmt1 H2O-64]# sbatch --partition=spot-compute ./h2o_64_1_node.sh
[root@maxmgmt1 H2O-64]# squeue
JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
205290 spot-comp cp2k_h20 root  R       0:03      1 cgpu01
```

Code 16 - workload manager - sbatch and squeue
<table>
<thead>
<tr>
<th><strong>Field</strong></th>
<th><strong>Value</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>JobId</td>
<td>205290</td>
</tr>
<tr>
<td>JobName</td>
<td>cp2k_h20_64_1_node</td>
</tr>
<tr>
<td>UserId</td>
<td>root(0)</td>
</tr>
<tr>
<td>GroupId</td>
<td>root(0)</td>
</tr>
<tr>
<td>MCS_label</td>
<td>N/A</td>
</tr>
<tr>
<td>Priority</td>
<td>6013</td>
</tr>
<tr>
<td>Nice</td>
<td>0</td>
</tr>
<tr>
<td>Account</td>
<td>root</td>
</tr>
<tr>
<td>QOS</td>
<td>normal</td>
</tr>
<tr>
<td>JobState</td>
<td>RUNNING</td>
</tr>
<tr>
<td>Reason</td>
<td>None</td>
</tr>
<tr>
<td>Dependency</td>
<td>(null)</td>
</tr>
<tr>
<td>Requeue</td>
<td>1</td>
</tr>
<tr>
<td>Restarts</td>
<td>0</td>
</tr>
<tr>
<td>BatchFlag</td>
<td>1</td>
</tr>
<tr>
<td>Reboot</td>
<td>0</td>
</tr>
<tr>
<td>ExitCode</td>
<td>0:0</td>
</tr>
<tr>
<td>RunTime</td>
<td>00:00:07</td>
</tr>
<tr>
<td>TimeLimit</td>
<td>41-15:00:00</td>
</tr>
<tr>
<td>TimeMin</td>
<td>N/A</td>
</tr>
<tr>
<td>SubmitTime</td>
<td>2019-10-11T16:11:21</td>
</tr>
<tr>
<td>EligibleTime</td>
<td>2019-10-11T16:11:21</td>
</tr>
<tr>
<td>AccrueTime</td>
<td>2019-10-11T16:11:21</td>
</tr>
<tr>
<td>StartTime</td>
<td>2019-10-11T16:11:36</td>
</tr>
<tr>
<td>EndTime</td>
<td>2019-11-22T06:11:36</td>
</tr>
<tr>
<td>Deadline</td>
<td>N/A</td>
</tr>
<tr>
<td>PreemptEligibleTime</td>
<td>2019-10-11T16:11:36</td>
</tr>
<tr>
<td>PreemptTime</td>
<td>None</td>
</tr>
<tr>
<td>SuspendTime</td>
<td>None</td>
</tr>
<tr>
<td>SecsPreSuspend</td>
<td>0</td>
</tr>
<tr>
<td>LastSchedEval</td>
<td>2019-10-11T16:11:36</td>
</tr>
<tr>
<td>Partition</td>
<td>spot-compute</td>
</tr>
<tr>
<td>AllocNode:Sid</td>
<td>maxmgmt1:204311</td>
</tr>
<tr>
<td>ReqNodeList</td>
<td>(null)</td>
</tr>
<tr>
<td>ExcNodeList</td>
<td>(null)</td>
</tr>
<tr>
<td>NodeList</td>
<td>cgpu01</td>
</tr>
<tr>
<td>BatchHost</td>
<td>cgpu01</td>
</tr>
<tr>
<td>NumNodes</td>
<td>1</td>
</tr>
<tr>
<td>NumCPUs</td>
<td>36</td>
</tr>
<tr>
<td>NumTasks</td>
<td>36</td>
</tr>
<tr>
<td>CPUs/Task</td>
<td>1</td>
</tr>
<tr>
<td>TRES</td>
<td>cpu=36,mem=150G,node=1,billing=36</td>
</tr>
<tr>
<td>Socks/Node</td>
<td>*</td>
</tr>
<tr>
<td>NtasksPerN:B:S:C</td>
<td>36:0:0:<em>:</em></td>
</tr>
<tr>
<td>CoreSpec</td>
<td>*</td>
</tr>
<tr>
<td>MinCPUsNode</td>
<td>36</td>
</tr>
<tr>
<td>MinMemoryNode</td>
<td>150G</td>
</tr>
<tr>
<td>MinTmpDiskNode</td>
<td>0</td>
</tr>
<tr>
<td>Features</td>
<td>(null)</td>
</tr>
<tr>
<td>DelayBoot</td>
<td>00:00:00</td>
</tr>
<tr>
<td>OverSubscribe</td>
<td>OK</td>
</tr>
<tr>
<td>Contigious</td>
<td>0</td>
</tr>
<tr>
<td>Licenses</td>
<td>(null)</td>
</tr>
<tr>
<td>Network</td>
<td>(null)</td>
</tr>
<tr>
<td>Command</td>
<td>./h2o_64_1_node.sh</td>
</tr>
<tr>
<td>WorkDir</td>
<td>/opt/software/benchmarking/cp2k/inputs/H2O-64</td>
</tr>
<tr>
<td>StdErr</td>
<td>/opt/software/benchmarking/cp2k/inputs/H2O-64/slurm.205290.err</td>
</tr>
<tr>
<td>Stdin</td>
<td>/dev/null</td>
</tr>
<tr>
<td>StdOut</td>
<td>/opt/software/benchmarking/cp2k/inputs/H2O-64/slurm.205290.out</td>
</tr>
<tr>
<td>Power</td>
<td></td>
</tr>
</tbody>
</table>

**Code 17 - workload manager - scontrol job**
6.4 Submitting GPU Jobs

GPU jobs are very similar to compute jobs with the exception of an additional required Slurm parameter. When submitting GPU jobs "--gres" must be used to allow a GPU resource to be allocated. GPU jobs must be submitted to one of the GPU partitions. If a GPU job is submitted to a non-gpu partition it will not run and likely fail when being submitted and give an error or sit in the queue waiting for resources.

To determine which GPU type you are best using see the following table:

<table>
<thead>
<tr>
<th>Org.</th>
<th>Job requirements</th>
<th>Slurm Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Job Interruptible?</td>
<td>GRES Attributes</td>
</tr>
<tr>
<td>UoA</td>
<td>Single &lt;12G No</td>
<td>gtx2080ti:1</td>
</tr>
<tr>
<td>UoA</td>
<td>Single &lt;12G Yes</td>
<td>gtx2080ti:1</td>
</tr>
<tr>
<td>UoA</td>
<td>Single &lt;24G No</td>
<td>gtx2080ti:2</td>
</tr>
<tr>
<td>UoA</td>
<td>Single &lt;24G Yes</td>
<td>gtx2080ti:2</td>
</tr>
<tr>
<td>UoA</td>
<td>Single &gt;24G No</td>
<td>gtx2080ti:X</td>
</tr>
<tr>
<td>UoA</td>
<td>Double &lt;32G Yes</td>
<td>v100:1</td>
</tr>
<tr>
<td>UoA</td>
<td>Double &lt;64G Yes</td>
<td>v100:2</td>
</tr>
<tr>
<td>NDC</td>
<td>Single/Double &lt;32G N/A</td>
<td>v100:1</td>
</tr>
<tr>
<td>NDC</td>
<td>Single/Double &lt;64G N/A</td>
<td>v100:2</td>
</tr>
</tbody>
</table>

Table 8 - workload manager - GPU placement
**Workload Manager - Slurm**

It is important to remember the NVIDIA GeForce GTX 2080ti are not optimised for double precision workloads but are still capable of running them. Use `sinfo` to determine if there are any jobs running on spot-gpu partition to see if you can run your job there quicker than using a GTX 2080ti GPU. NDC users have priority access to V100 GPUs and can only submit jobs to this resource type.

**Submitting a job to a consumer GPU node:**

```bash
#!/bin/bash
#SBATCH -N 1 # number of nodes
#SBATCH -n 2 # number of cores
#SBATCH --mem 50G # memory pool for all cores
#SBATCH -o slurm.%j.out # STDOUT
#SBATCH -e slurm.%j.err # STDERR
#SBATCH --ntasks-per-node=2
#SBATCH --gres=gtx2080ti:1
#SBATCH --partition=uoa-gpu

module load cudatoolkit-10.1.168

nvidia-smi # show GPU status
./cuda-program
```

*Code 18 - workload manager - consumer GPU job*

**Submitting a job to an enterprise GPU node:**

```bash
#!/bin/bash
#SBATCH -N 1 # number of nodes
#SBATCH -n 2 # number of cores
#SBATCH --mem 50G # memory pool for all cores
#SBATCH -o slurm.%j.out # STDOUT
#SBATCH -e slurm.%j.err # STDERR
#SBATCH --ntasks-per-node=2
#SBATCH --gres=v100:1
#SBATCH --partition=spot-gpu

module load cudatoolkit-10.1.168

nvidia-smi # show GPU status
./cuda-program
```

*Code 19 - workload manager - enterprise GPU job*

If the job is from an NDC user the job should be submitted to “ndc-gpu” to use the V100 GPU and start the job with higher priority than any UoA jobs running on egpu01 node.
For support and assistance or to report a fault with the Maxwell cluster, contact the Digital Research Infrastructure team via email - digitalresearch@abdn.ac.uk

For urgent matters please email the details of your issue to the address below and follow up with a phone call to ensure the Digital Research Infrastructure team can give the matter the appropriate priority.

- hpc@abdn.ac.uk
- Naveed Khan: 01224 27-4126
- Michael Chung: 01224 27-2750

Where possible, please include the following information to ensure the issue/request can be dealt with as soon as possible:

1. Identify Maxwell HPC in the call
2. Include a description of the problem or request
3. Identify username(s) or groups the call relates to
4. Provide example job output files clearly identifying the issue
5. A method for replicating the problem
8. Index

8.1 Tables
Table 1 - version control - latest ................................................................................................................................... 4
Table 2 - version control - history ................................................................................................................................... 4
Table 3 - cluster resources - physical resources ........................................................................................................... 6
Table 4 - cluster resources - storage resources ............................................................................................................. 6
Table 5 - cluster resources - Slurm resources ............................................................................................................... 7
Table 6 - cluster resources - Slurm partition priorities ................................................................................................. 8
Table 7 - workload manager - common Slurm parameters ........................................................................................ 26
Table 8 - workload manager - GPU placement ........................................................................................................... 29

8.2 Images
Image 1 - cluster access - UoA Putty ............................................................................................................................ 11
Image 2 - cluster access - UoA Putty logged in ............................................................................................................... 11
Image 3 - cluster access - UoA x2go XFCE desktop ..................................................................................................... 12
Image 4 - cluster access - UoA x2go session preferences .......................................................................................... 12
Image 5 - cluster access - WinSCP login ...................................................................................................................... 15
Image 6 - cluster access - WinSCP logged in ............................................................................................................... 15
Image 7 - cluster access - remote access Putty remote ............................................................................................. 16
Image 8 - cluster access - remote access Putty ........................................................................................................... 16
Image 9 - cluster access - remote x2go ......................................................................................................................... 17
Image 10 - cluster access - SSH tunnel WinSCP ........................................................................................................ 18
Image 11 - workload manager - Galaxy login ............................................................................................................ 21
Image 12 - workload manager - Galaxy logged in ...................................................................................................... 22
8.3 Code Samples

Code 1 - cluster resources - Slurm sinfo ................................................................. 8
Code 2 - cluster access - UoA maxlogin1 ............................................................... 10
Code 3 - cluster access - UoA maxlogin2 ............................................................... 10
Code 4 - cluster access - UoA x2go server settings ............................................. 12
Code 5 - cluster access - NDC ndc1 ................................................................. 13
Code 6 - cluster access - NDC ndc2 ................................................................. 13
Code 7 - cluster access - NDC x2go server settings ........................................ 13
Code 8 - cluster access - scp and rsync ............................................................. 14
Code 9 - cluster access - remote command line access ..................................... 16
Code 10 - cluster applications - lmod ............................................................... 20
Code 11 - workload manager - squeue ............................................................... 24
Code 12 - workload manager - scontrol show partition ................................ 24
Code 13 - workload manager - scontrol show job ............................................. 25
Code 14 - workload manager - sinfo ............................................................... 25
Code 15 - workload manager - Slurm script ....................................................... 27
Code 16 - workload manager - sbatch and squeue ........................................... 27
Code 17 - workload manager - scontrol job ...................................................... 28
Code 18 - workload manager - consumer GPU job ........................................ 30
Code 19 - workload manager - enterprise GPU job ........................................ 30