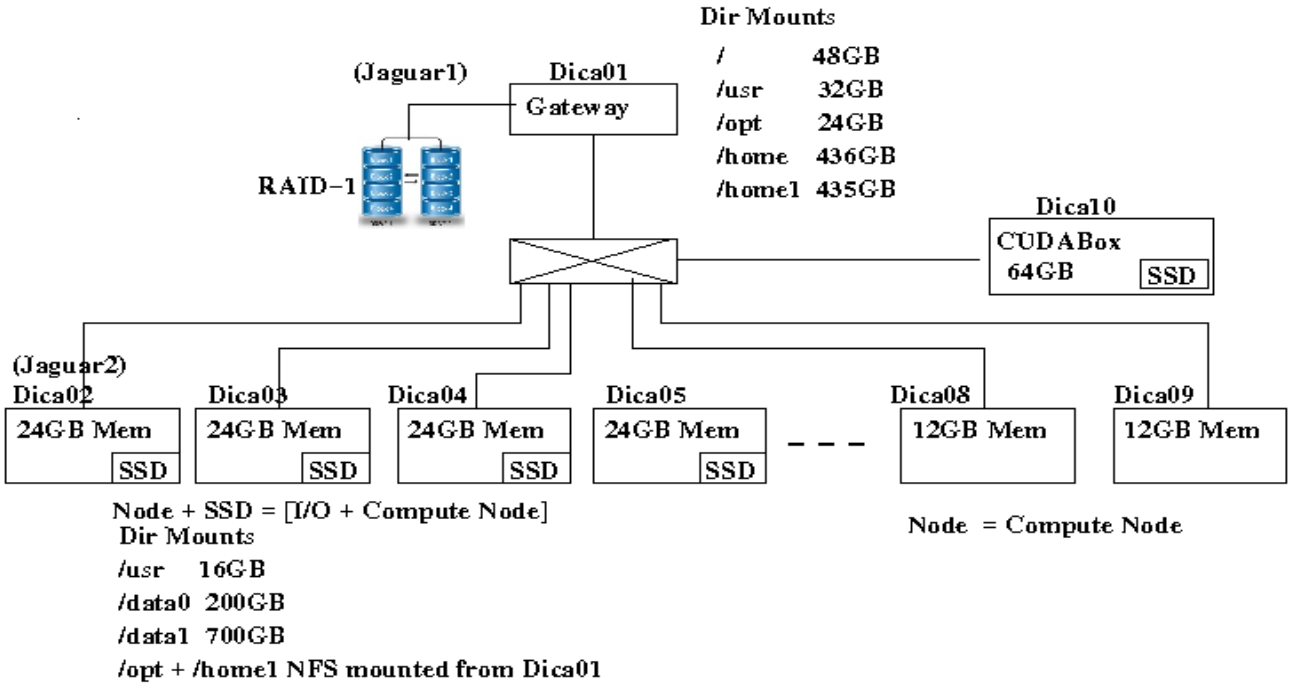
ELEN4020: Data Intensive Computing
Project's Cluster Usage Brief

April 18, 2019

System Architecture

A simple sketch of the cluster configuration, some details are shown in the diagram below.



The nodes have hostnames: dica01 (jaguar1) — the gateway node
dica02 (jaguar2) — kerberos authenticator
dica03
...
...
dica08
dica09
dica10 (cudabox)

You need to login ONLY onto **jaguar1** from any of the wits network domain and from outside networks. Internally within "eie.wits.ac.za" this is also seen as **dica01.eie.wits.ac.za (dica01)**. Your home directories will be on /home1 and the common pool of software will be on /opt on this machine. So there are two directories, **/opt** and **/home1** that are NFS mounted on all the other nodes.

There is a **/home1/Miscellaneous directory** on /home1 that contains the following four files: Dot_Bash_Profile, Dot_Bashrc and Dot_Bash_logout and Dot_Bash_Aliases. These should

be copied to replace the following respective files, .bash_profile, .bashrc and .bash_logout and .bash_aliases in your home directory the **first time only** after your login. After that, source the ".bashrc" file as follows after you login.

```
$ source  .bashrc  <-- This should also
                        source ".bash_profile" and
                               ".bash_aliases"
$ source  .bash_logout
```

SSH to Jaguar1

You should not login to the cluster using "ssh." The login accounts are the same as the accounts provided for **hornet01**

Passwordless SSH

Step 1: Check that you can do a read/write to your home directory and that you can read the content of /opt. To test if you can write to you home directory do the follow:

```
$ cd
$ touch tmpFile
```

A file called tmpFile should be created in your home directory.

Step 2: If you did not source your ".bashrc" file to set up your environment change directory (cd) to your home directory and do:

```
$ source  .bashrc  <-- This should also
                       source ".bash_profile" and ".bash_aliases"

$ source  .bash_logout
```

Step 3: Check that you are invoking the correct compilers and libraries and things are OK with commands like

```
    $ which gcc
    $ which mpicc
    $ which mpiexec
    $ ssh dica02 date    #You will be calling date on dica0???
OR
    $ rsh dica02 date
and
    $ ssh dica03 hostname
```

Finally check in the Miscellaneous directory for some sample codes, machinefiles and instructions on how to run simple MPI programs. These are also illustrated at the end of this document. The account sedic014 is setup identically as your group accounts and is used to test

that your environments work as expected. MachineFile can contain any where from 1 up to 9 dica?? hosts. The illustration shows all 6 working nodes; effectively you have 48 cores in total but can run a reasonable number of processes (say *le* 256). To execute the code a.out compiled from hellow.c, the following is done

```
$ mpicc hellow.c  ---> generates a.out
$ mpiexec -n 20 -f machineFile ./a.out
```

```
You could also check running
     $ mpiexec -n 9  -f machineFile hostname
This lists the names of the machines in machineFile
```

Some limitations in the setup

1. All hosts; dica01 to dica09 should be fully functional.
2. The MPI compiler being used is mpich3.2.1 (stable release) and hydra-3.2.1 (stable release) Hydra (mpiexec).
3. Setup a /scratch directory in your home for now. A common /scratch will be setup later .
4. The parallel file systems "BeeGFS" is not ready yet.
5. Similarly Parallel HDF5 (PHDF5) and Parallel NetCDF (PNetCDF) will be ready after the setup of BeeGFS.
6. The complete options list used in building the mpich compiler will placed in the /Miscellaneous directory.
7. The Python on these nodes are versions 2.7, and 3.6.5
8. **The Above Configuration is Fequently Changing. Please Check with the Lecturer for the Latest Changes**

```
sedic014@dica01:~/TestProgs$ ll
total 16
-rwxrwxr-x 1 sedic014 sedic014 7856 May  6 21:01 a.out
-rw-rw-r-- 1 sedic014 sedic014  788 May  6 21:00 hellow.c
-rw-rw-r-- 1 sedic014 sedic014    7 May  7 21:01 machineFile
sedic014@dica01:~/TestProgs$

sedic014@dica01:~/TestProgs$ cat machineFile
dica02:2
dica03:3
dica04:2
dica06:2
dica07:4
dica08:2
sedic014@dica01:~/TestProgs$

sedic014@dica01:~/TestProgs$ mpiexec -n 20 -f machineFile ./a.out
Hello from task 5 on dica04!
```

```
Hello from task 8 on dica06!
Hello from task 0 on dica02!
MASTER: Number of MPI tasks is: 20
Hello from task 15 on dica02!
Hello from task 13 on dica08!
Hello from task 14 on dica08!
Hello from task 16 on dica02!
Hello from task 1 on dica02!
Hello from task 9 on dica07!
Hello from task 7 on dica06!
Hello from task 2 on dica03!
Hello from task 10 on dica07!
Hello from task 11 on dica07!
Hello from task 17 on dica03!
Hello from task 12 on dica07!
Hello from task 3 on dica03!
Hello from task 4 on dica03!
Hello from task 18 on dica03!
Hello from task 19 on dica03!
Hello from task 6 on dica04!
sedic014@dica01:~/TestProgs$
```