HPC DOCUMENTATION

Account in HPC server can be given to every student on request and after faculty approval

1. Hardware Resource :-

Our HPC consists of Blade chassis with Seven blade servers and Five GPU rack servers.

- Total available cores for computing: - 284 cores.
- Cores reserved and dedicated to master node:- 04
- Cores reserved and dedicated to GPU node:- 20
- Cores reserved and dedicated to Matlab:- 72
- Cores reserved and dedicated to High priority:- 16
- Total cores (summation of all):- 172
- **GPU Server details: -**
- V100 Server : 20 core , 96 GB RAM and 2 V100 32 GB RAM
- V100 Server2 : 20 core, 96 GB RAM and 2 V100 32 GB RAM
- A100 Server : 64 core , 256 GB RAM and 2 A100 40 GB RAM
- A100 Server : 64 core , 256 GB RAM and 2 A100 40 GB RAM
- A100 Server : 128 core , 512 GB RAM and 4 HGX A100 40 GB RAM
- Total User Data area(SAN) :- 13 TB

2. Software Resources:-

a. Operating system:- CentOS Linux release 7.4.1708 (Core)
b. Compilers & Libraries:- Cuda compliers.
c. Application:- MATLAB 2019a, gromacs, python, caffe, vmd theano, keras
d. Scheduler :- Grid Engine
e. Monitoring tool:- Ganglia
f. mpirun (Open MPI) 2.1.1

3. Node Names and IP addresses:-

Node details with respect to their individual IP addresses are given below:-

Sl No.	Host Name	Node Type	IIITD LAN IP
1	hpc.iiitd.edu.in	(Mater Node)	hpc.iiitd.edu.in
2	compute-0-0	Compute node	ssh from master node
3	compute-0-1	Compute node	ssh from master node

4	compute-0-2	Compute node	ssh from master node
5	compute-0-3	GP GPU server	ssh from master node
6	Compute-0-4	Compute-0-4	ssh from master node
7	V100 server	Compute node	On request
8	V100 Server2	Compute node	On request
9	A100-1	Compute node	On request
10	A100-2	Compute node	On request
11	A 100 HGX	Compute node	On request

Note: - Access to compute node from master node is given only to privileged users who requires it.

4. Scratch area and User Data area details:-

Scratch area is mainly are of two types as given below:-

a. **/ssd-scratch:-** It is 1TB shared space allocated from SAN SSD drives. This space is commonly available to all nodes. All users are expected to use this common space for saving rapidly accessed temporary computational data.

b. **/scratch: -** It is of 900GB local disk space available to each compute nodes individually and 200GB from master node is allocated for this purpose. There are 5 compute nodes and 1 master as given above. Hence total space available will be around 3.8 TB across all nodes. Users are expected to save temporary data to this area. Data saved on this drive will be purged after every 15 days. Purging will be done on files which have been last accessed before 15 days. Hence users are advised to shift data from **/scratch** to respective data area like **/home/<username>.**

User Data area is mainly available for users to save their computational data permanently. A total space of 4TB is allocated from SAN for this purpose. This space is available globally across all nodes. **/home/<username>** is the location of the data area available to each user. There are some restrictions and Quota applied on users and group basis on the DATA area. Details are given in section 5 and 6. No users are allowed to delete any other users data area files.

5. Users and Groups:-

All HPC users are primarily divided into two groups namely **faculty** (which includes Ph.D students also) **and student.** Separate quota and restrictions have been applied to the individual groups.

6. Quota and Restrictions:-

Quota and some restrictions have been implemented on HPC. These restrictions can be broadly classified as in two categories: - **CPU cores restrictions and DATA restrictions.**

CPU cores restrictions

a. No user, in student group , will be able to submit more than 2 jobs at a time. (Unless special permission is granted)
b. No user, in faculty group, will be able to submit more than 4 jobs at a time. (Unless special permission is granted)
c. No user, irrespective of groups, will be allowed to use more than 30 CPU cores per user account except GPU queue users and unrestricted queue users. (Unless special permission is granted)
d. GPU queue users, irrespective of groups, can take at most 4 CPU cores and all CUDA cores. But they cannot fire more that 2 jobs for student group and 4 jobs in faculty group.   (Unless special permission is granted)

DATA restrictions

a. Faculty group users are allowed to use 40GB of user data space each.
b. Student group users are allowed to use 10GB of user data space each.
c. There is no quota or restriction imposed on **/ssd-scratch** and **/scratch**

7. Scratch area deletion policy:-

Data saved under **/scratch** gets deleted after every 15 days. Purging will be done on files which have been last accessed before

15 days. Hence users are advised to shift data from **/scratch** to their respective data area like **/home/<username>.**

8. How to login to HPC:-

Generally every users has to login to master node for firing jobs in

any of the queues as explained above. Login to compute nodes directly is not allowed. If required users have to first login to master node and then they can login to any of the compute nodes. ***Steps to login to master node with respect to both MAC/Linux/Windows machines are explained below:-***

For Linux/MAC machine users

Please follow the steps as given below:-

      a.**ssh <username>@hpc.iiitd.edu.in**
      b. Please use the same **<username>** as used in IIITD Domain
      c. First time password will be first 3 characters of the username followed by @4321. For example, suppose username is **helpdesk** then password will be **hel@4321**
      d. Each user will be asked / forced to change their respective password when they login first time.

For windows machine users

Please follow the steps as given below:-

      a. Please use Putty and put hpc.iiitd.edu.in in place of host name and port selected should be 22.
      b. On command prompt please put your **<username>** when asked for. Please use the same **<username>** as used in IIITD Domain.
      c. Please key in your password when asked for. First time password will be first 3 characters of the username followed by @4321. For example, suppose username is **helpdesk** then password will be **hel@4321**
      d. Each user will be asked/ forced to change their respective password when they login first time.
.

3. Queue details: -

We have created following Queues to fire jobs through it. The details of number of cores allocated to respective queues are given below:-

Queue Name	TOTAL Cores
all.q	172
gpu.q	20
highpri.q	16
unrestricted.q	77

all.q :- This queue is available to both faculty and student groups. In this queue a total of 80 CPU cores are available but a restriction of 30 CPU cores per user account has been imposed on this queue. No user can submit more that 2 jobs at a time (unless specific permission is granted). All programming jobs are expected to be fired from this queue. All compute node CPU cores are dedicated for this queue. Waiting time in this queue is expected to be higher as this queue is being shared by all users.

gpu.q:- This queue is available to both faculty and student groups. In this queue there are 4 CPU core and 2496 cuda cores available for usage. There are no restrictions imposed on this queue except that no user will be allowed to fire more than 2 jobs at a time (unless specific permission is granted). All GPU /Cuda programming related jobs are expected to be fired from this queue. Compute node **Compute-0-3** is dedicated for this queue. Waiting time in this queue is expected to be much lesser as this queue is being shared only by GPU / Cuda programming users.

highpri.q:- This queue is available to only faculty group. In this queue a total of 16 CPU cores are available for usage. There are no restrictions imposed on this queue except that no user will be allowed to fire more than 2 jobs at a time (unless specific permission is granted). Users are expected to use this queue in case they want to run a very high priority job which requires less CPU cores but needs lesser waiting time.

unrestricted.q :- All 77 CPU cores are available in this queue. This queue is available only to faculty groups. Jobs that are being fired from this queue will have to wait till all required numbers of CPU

cores are freely available and are not being used by any other queue or jobs.

4.   Job Submission Steps: -

Steps for firing jobs are given below:-

How to Submit a Parallel Job (Open MPI)

Please create the following script and save it to any file name as shell file. For example let's save it as **Jobscript.sh**

```
#!/bin/bash

#$ -N <Parallel-Job>          // Parellel-Job is the name of the job //
#$ -cwd

#$ -S /bin/bash
#$ -e err.$JOB_ID.$JOB_NAME
#$ -o out.$JOB_ID.$JOB_NAME

#$ -q <queue name> // The queue where the users wants to
fire jobs//
#$ -pe mpi 8        // The maximum MPI users wants to use
its upper //
 unset SGE_ROOT            // limit is the number of cores available to
the //
 /opt/mpi/1.8.1/bin/mpirun -np $NSLOTS –hostfile          //respective
queue//
$TMPDIR/machines ./a.out > <output_file> // a.out is code file name and
//outputfile is redirection file name//
```

Please submit the said job using the following command:-

qsub <Jobscript name >

In our example we will use the following command:-
qsub Jobscript.sh

For checking status of various jobs fires please use the following commands:-

- **qstat (provides status listing of jobs)**
- **qstat –f**
- **qstat –g c**