

Part-I: Bootcamp on using PARAM Shakti supercomputer

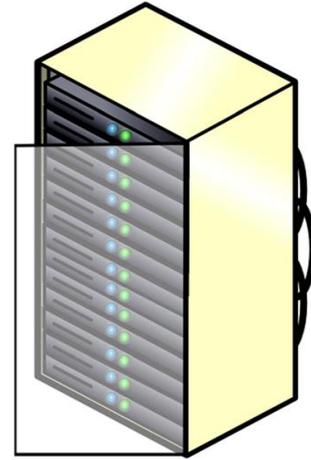
**Mr. Hemanta Kumar Sahu
Sr. System Engineer ,CCDS**

Date:18-Sep-2021

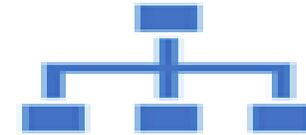


Param shakti supercomputing center IIT Kharagpur

Supercomputing cluster terminology



- What is the need of a cluster ?
 - **When one server is not enough** for the computational task or analysis to complete , clusters are used.
- Cluster is a group of machines interconnected in a way that they work together as a single system
- Terminology :
 - **Node** – individual machine in a cluster
 - **Head/Master node** – connected to both the private network of the cluster and a public network and are used to access a given cluster.
 - **Compute nodes** – connected to only the private networks of the cluster and are generally used for running jobs assigned to them by the login node(s).
- Compute nodes can be of different types:
 - CPU only nodes
 - GPU nodes
 - High memory nodes, etc



Supercomputing cluster types

Statefull (Diskfull) Cluster

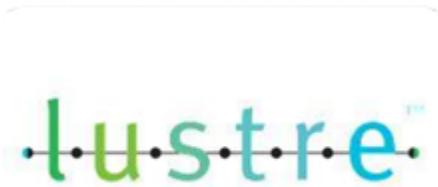
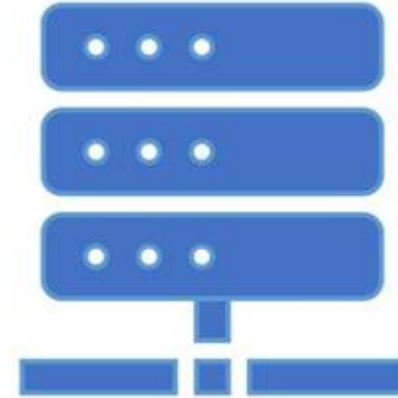
- Traditional cluster with OS on each nodes local disk.

Stateless (Diskless) Cluster

- Nodes are booted using RAMdisk os-image.

Supercomputing cluster hardware components

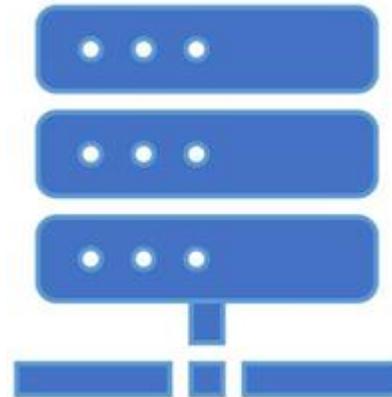
1. Nodes (servers)
2. Parallel file system (Storage)
3. Interconnect (Networking)
4. Accelerator cards for performance boosting (GPU)
5. Optimized compilers and libraries



Supercomputing cluster hardware components

•Nodes:

- Nodes are the actual server who will take part in computation which has its own CPU, memory, network card.
- They have multiple cores and supports hyper threading.



Supercomputing cluster hardware components

Lustre parallel file system

- Lustre is parallel file System where multiple clients can write to the different parts of same file, at the same time multiple clients can read the file.
- It supports High bandwidth Interconnects such as Mellanox, Omnipath, etc.
- It is a scalable file system.
- It supports HA and it is a POSIX compliant file system.
- ACL can be applied.



Supercomputing Cluster Hardware Components

Graphical processing Unit

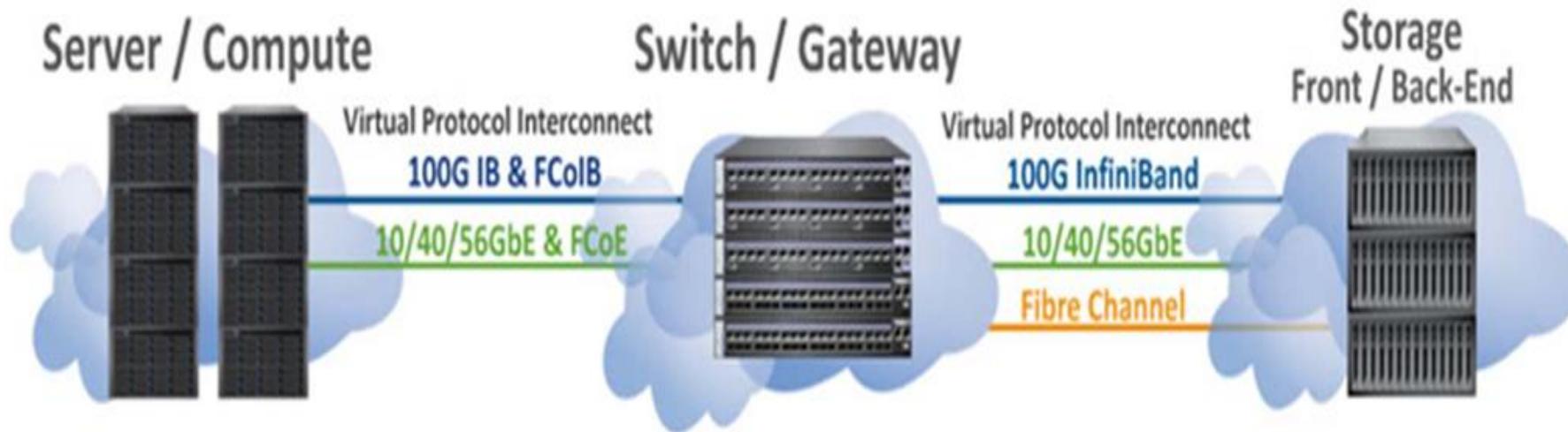
- A co-processor to accelerate general purpose scientific and engineering computing.
- It accelerates applications running on CPU by offloading compute intensive and time consuming portion of the code.
- GPU consists of thousands of smaller cores which together operate to crunch data in the application.



Supercomputing cluster hardware components

Infiniband

- High throughput and low latency technology that interconnects compute nodes and I/O nodes to form a system area Network.
- It uses RDMA (Remote Direct Memory Access) protocol to transmit data across the wire.



Supercomputing cluster software components

- **Operating System –**

- HPC clusters generally are build with Linux operating system as a base OS (Centos7.6)
- It includes all the device drivers for the H/W connected to each node.

- **Cluster Manager/Orchestrator**

- Tools in this category builds a centralized architecture where a controller node builds and manages the cluster.
- xCAT – Is a open source cluster Manager, developed by IBM, Maintained by community, is the widely used tool for HPC as well as cloud clusters.
- It provides flexibility to handle objects within the cluster with its easy manageable methods.
- It provide methods to deploy nodes with a very light weight stateless images.

Different parts of the data center



Server room

- 2 Master nodes
- 8 Login nodes
- 2 Service/Management nodes
- 442 (CPU+GPU) nodes
- Total peak computing capacity 1.66 PFLOPS
- Intel Xeon SKL 6148
- NVIDIA Tesla V100
- Mellanox (EDR) infiniband interconnect network which provides 100Gbps backbone.
- Lustre parallel file system.

Different parts of the data center

- Raw Power comes from two different transformer source through two separate electrical Panel (i.e Incomer-1 , Incomer2).
- 320 KVA online UPS (2 + 1) to provide uninterrupted power supply to IT load and 40KVA online UPS (1 + 1) to provide power to non-IT load until the DG comes into operation in case of power failure.
- Total peak Load = 370 KW .
- Total peak IT load =250 KW .



UPS and electrical panel room

Different parts of the data center



Chiller yard

- 80Tr Chiller (2 + 1) to provide cooling to the datacenter equipment including server, storage and UPS.
- Two Thermal storage tanks to provide chilled water for 15 mins in case emergency failure of all the chillers.

Different parts of the data center



DG Yard

- Two(2) 500 KVA DG sets running in parallel using the sync panel .
- Each DG set has approx. 700L diesel storage capacity to cater 8 -10 hours power during peak load.

Different parts of the data center

- Battery room consists of three(3) battery banks for each of the three 320 KVA UPS. There are two(2) small battery banks for each of the 40 KVA UPS.



Battery room

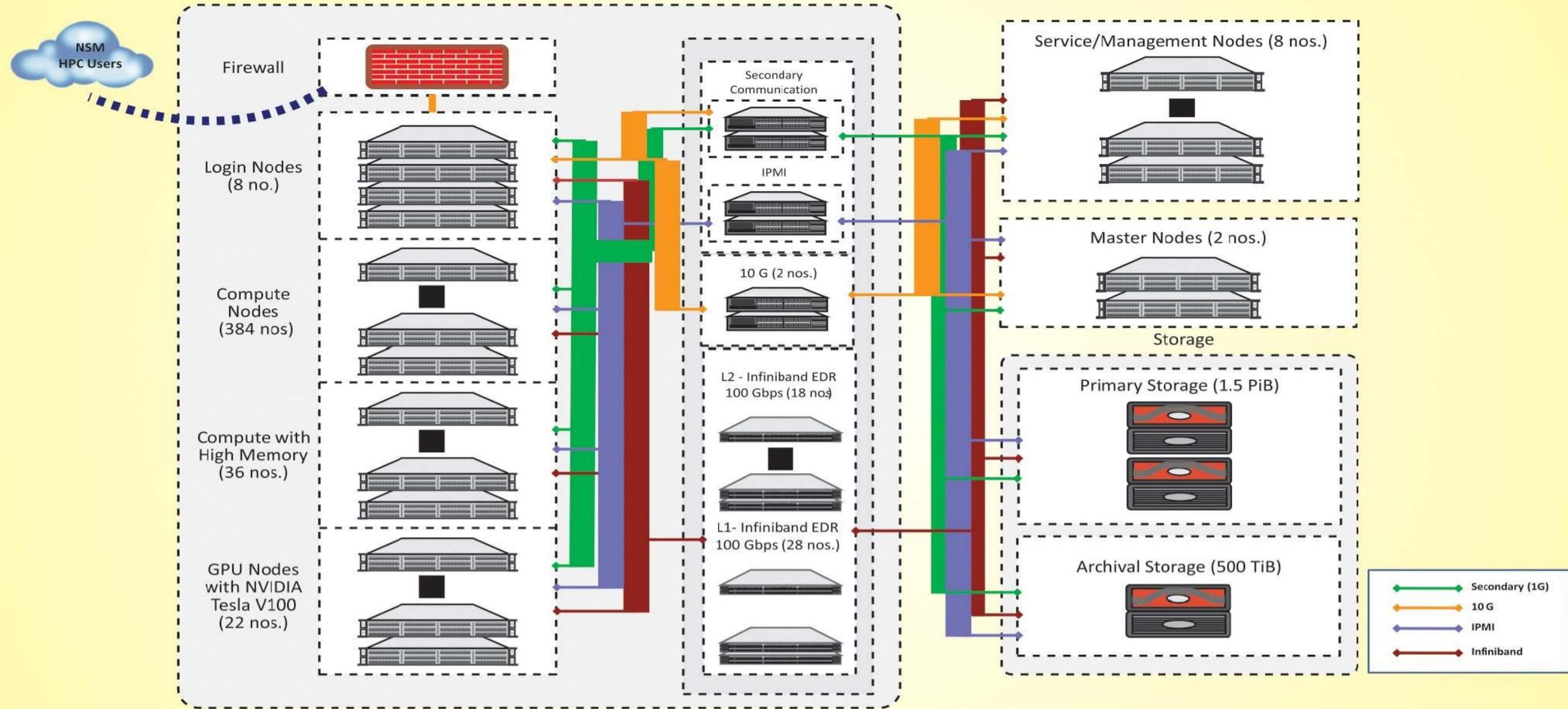
Different parts of the data center



BMS room

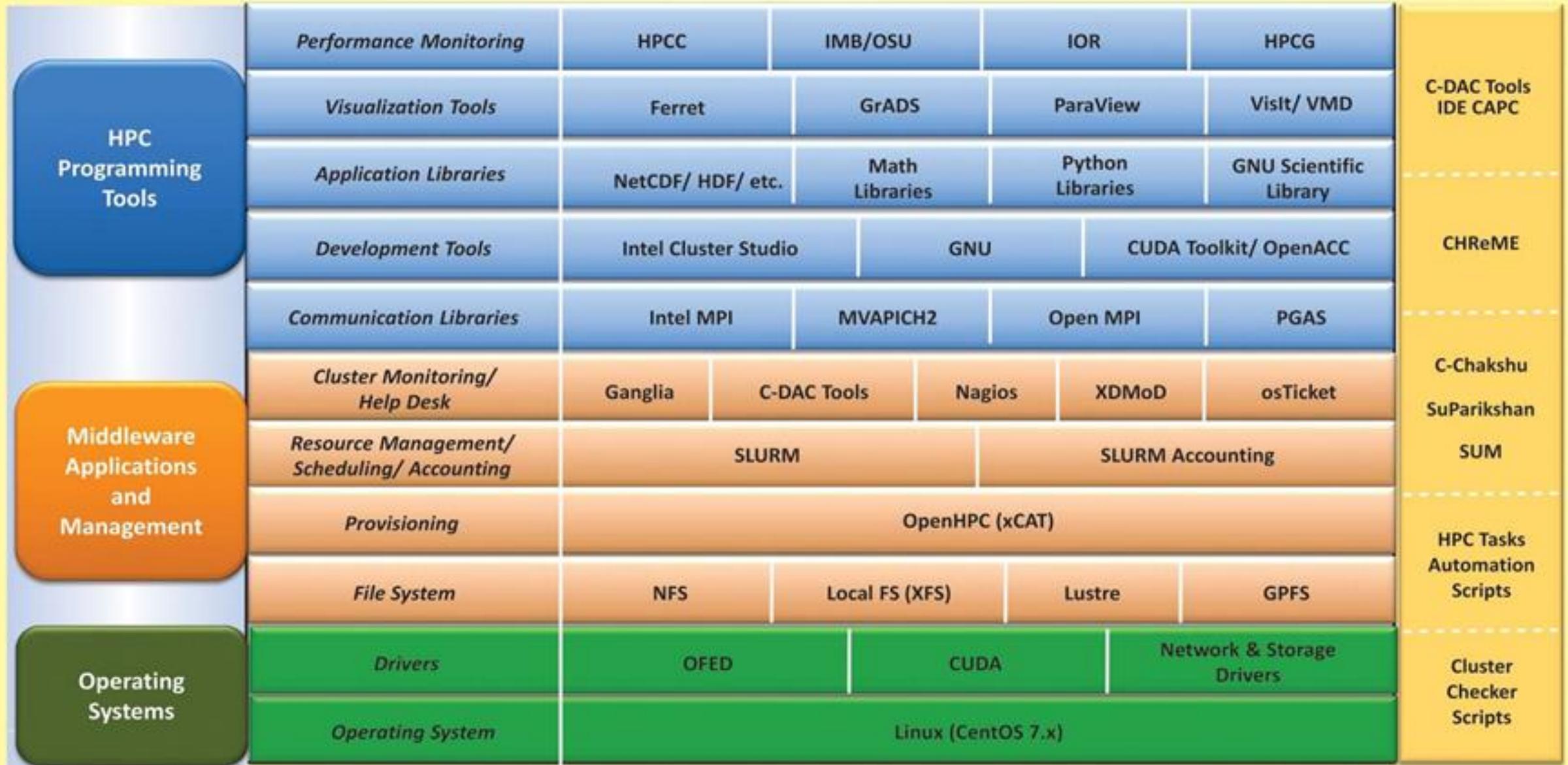
- The BMS room is central monitoring room where BMS engineer monitors health of all the datacenter equipment 24/7 and 365 days a year.
- Entire datacenter , Chiller and DG yard is under CCTV surveillance monitored by BMS engineer.
- We have VESDA, FAS and NOVAC Gas based fire suppression system for any fire emergency.
- We have also WLD system deployed to detect any water leakage in the chilled water pipeline.

Supercomputer architecture diagram(Hardware Stack)



**Final implementation : Primary storage 2.1PiB and Archival storage 500TiB . Service/Management Nodes (2 nos.)

Supercomputer architecture diagram (software stack)



Supercomputer compute nodes configuration

CPU only

- 384 nodes
- 15360 Cores
- Compute power - 1179 TFLOPS
- Each Node with
 - 2 x Intel Xeon SKL-6148, 20 cores, 2.4 GHz, processors
 - 192 GB memory
 - 480 GB SSD

High memory

- 36 nodes
- 1440 Cores
- Compute power - 110.59 TFLOPS
- Each Node with
 - 2 x Intel Xeon SKL G-6148, 20 cores, 2.4 GHz, processors
 - 768 GB memory
 - 480 GB SSD

Theoretical node performance in TFlops

(CPU speed in GHz) x (number of CPU cores) x (Floating point operations per cycle) x (number of CPUs per node) x Total Nodes

= 2.4 x 20 x 32 x 2 x 420

~ 1290TFlops

Supercomputer compute nodes configuration

GPU

- 22 nodes
- 880 CPU cores
- Compute power - 67 TFLOPS + 308 TF
- Each node with
 - 2* Intel Xeon SKL G-6148, 20 cores,
 - 192 GB Memory
 - 480 GB SSD
 - 2 x Nvidia V100 SXM2 GPU cards

Theoretical peak performance of Pramshakti Supercomputer

~ **1290 TF + 67 TF + 308 TF**

~ **1665TF ~ 1.66PF**

Theoretical GPU Node performance in GFlops

= CPU core performance + CUDA core performance

~ $2.4 \times 20 \times 32 \times 2 \times 22 + 7 \times 2 \times 22$

~ **67 TF + 308 TF**

GPU FP32 Cores per node = $2 \times 5120 = 10240$

GPU Tensor Cores per node = $2 \times 640 = 1280$

Peak FP64 TFLOPS per GPU card = 7.8 (Ref:Nvidia datasheet *)

Supercomputer support staff

Datacenter Non-IT monitoring and support Staff

- The Non-IT equipment monitored 24/7 by following operation and maintenance staff on behalf of M/S IBM India Pvt. Ltd.
1. Nand Lal Sharma
 2. Tamaswam Chaudhury
 3. Sanjib Kumar Manna
 4. Kali Shankar Bag
 5. Amit Hela (Helper)

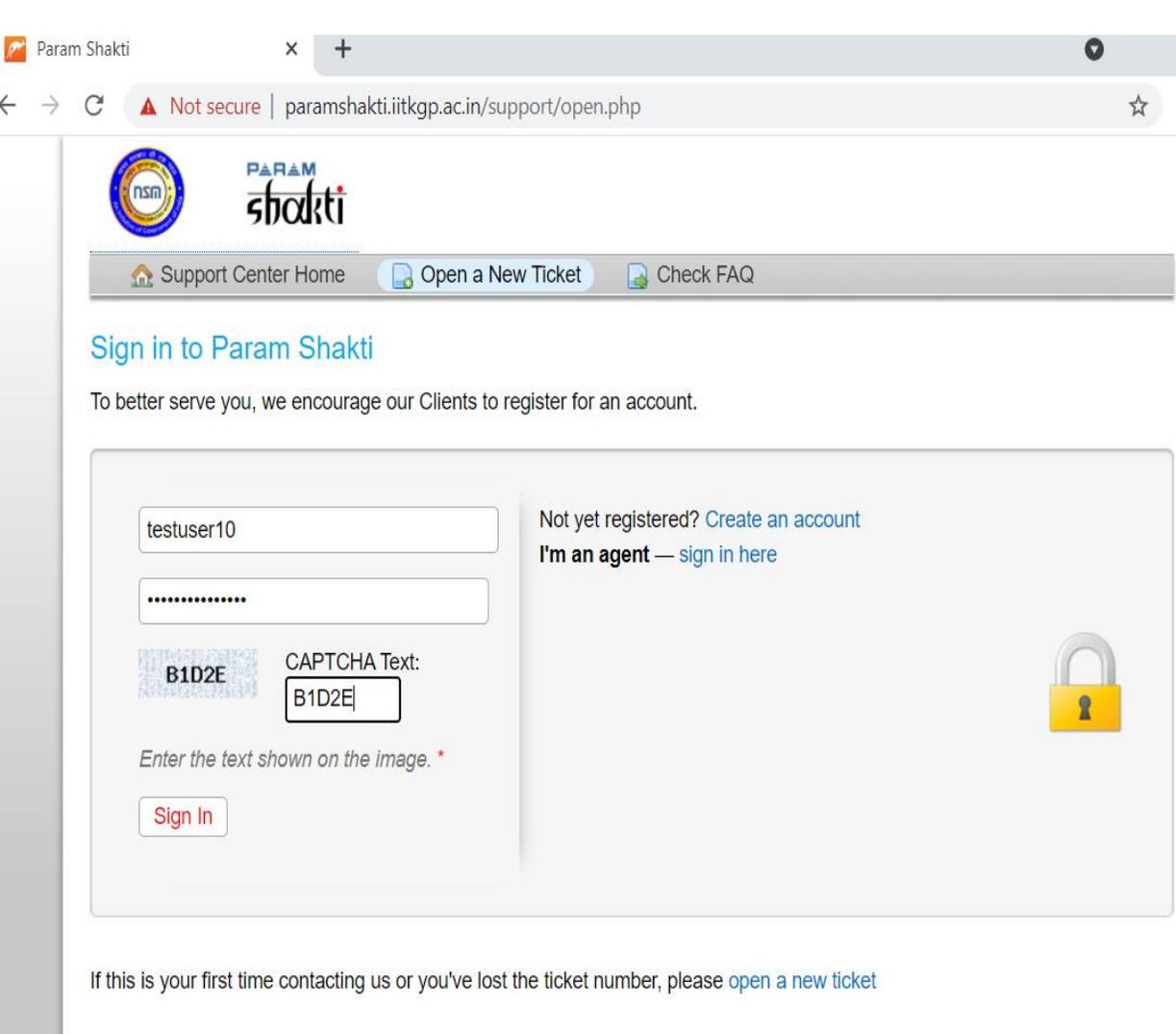
Supercomputer support staff

Datacenter IT monitoring and support Staff

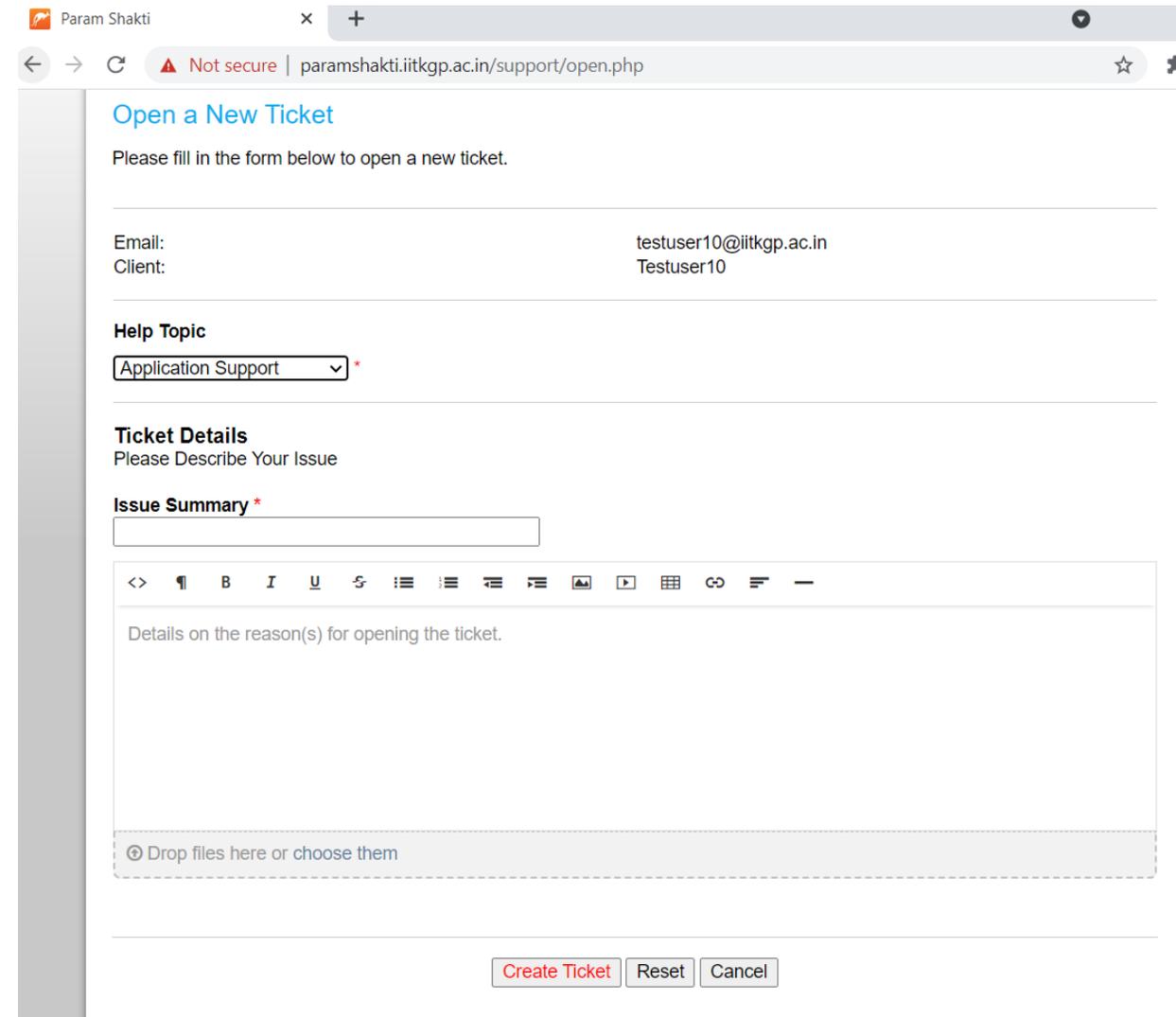
- Datacenter IT equipment monitored by following server support engineers on behalf of M/S ATOS India Pvt. Ltd.
 - 1) Mr. Anand Singh
 - 2) Mr. Rakesh Kumar
- Provide L1 support to external and internal users via web based ticketing tool.
<https://paramshakti.iitkgp.ac.in/support/open.php>
- Users are requested open a ticket for any issue or query .
- Do Not send email to shaktisupport@iitkgp.ac.in
- Note: If a support representative requests for any information *please don't reply to the shaktisupport@iitkgp.ac.in email , instead update the necessary information in the ticket itself so that the support representative will have all the information at one place .

Supercomputer support request and ticketing tool

Support portal for opening a ticket : <https://paramshakti.iitkgp.ac.in/support/open.php>



The screenshot shows the login page of the Param Shakti support portal. At the top, there is a navigation bar with the Param Shakti logo and three menu items: "Support Center Home", "Open a New Ticket", and "Check FAQ". Below the navigation bar, there is a "Sign in to Param Shakti" section with a message: "To better serve you, we encourage our Clients to register for an account." The main login area contains a form with a username field (containing "testuser10"), a password field (masked with dots), and a CAPTCHA section. The CAPTCHA image shows the text "B1D2E" and the CAPTCHA text input field also contains "B1D2E". There is a "Sign In" button and a "Not yet registered? Create an account" link. A yellow padlock icon is visible on the right side of the login area.



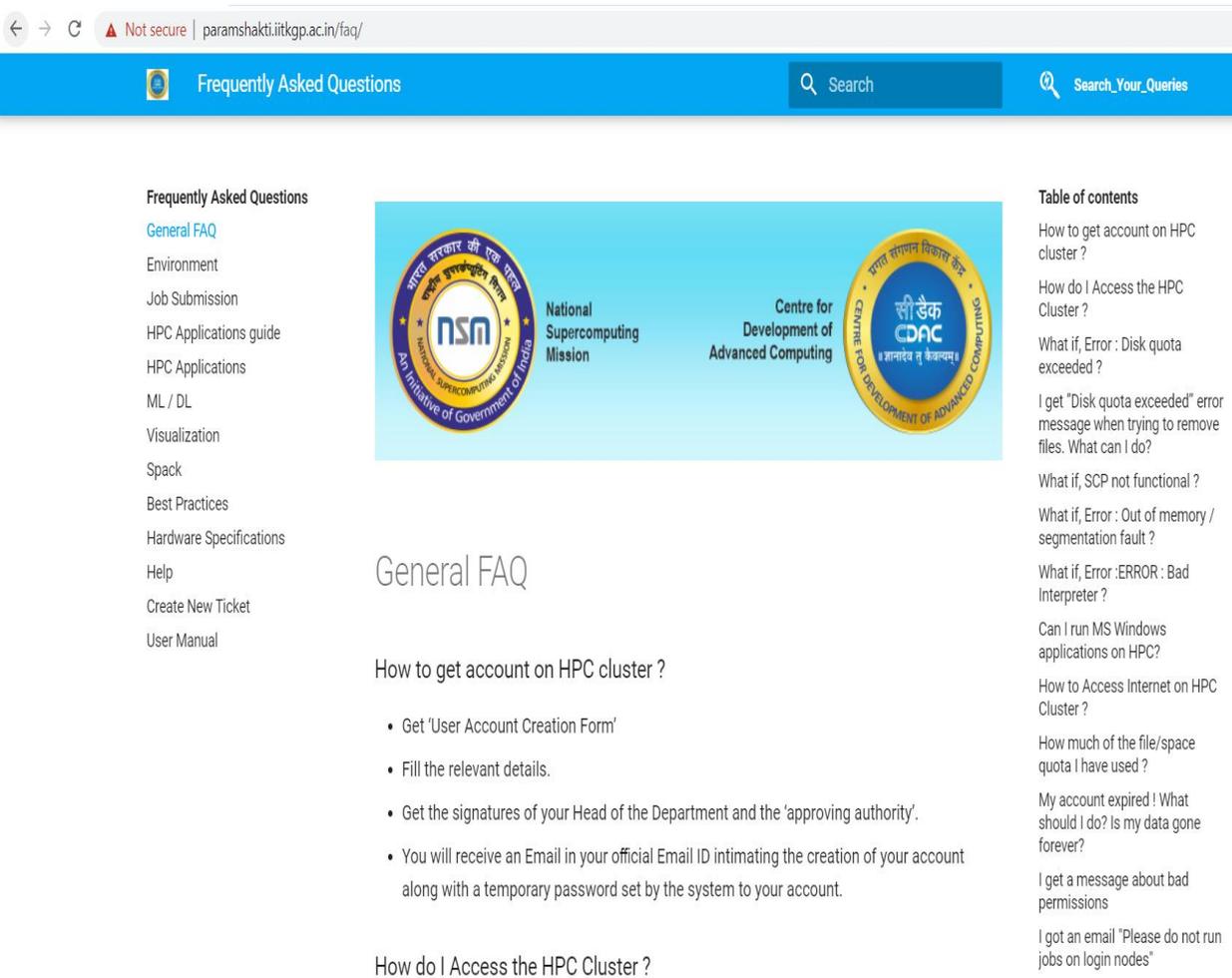
The screenshot shows the "Open a New Ticket" form in the Param Shakti support portal. The form is titled "Open a New Ticket" and includes the instruction "Please fill in the form below to open a new ticket." The form fields are: "Email:" (testuser10@iitkgp.ac.in), "Client:" (Testuser10), "Help Topic:" (Application Support), "Ticket Details:" (Please Describe Your Issue), and "Issue Summary:" (a text area). Below the issue summary is a rich text editor with a toolbar and a "Drop files here or choose them" area. At the bottom of the form, there are three buttons: "Create Ticket", "Reset", and "Cancel".

Supercomputer support request and ticketing tool

- Users are requested to go through the FAQ in HPC website as well as support portal before opening a ticket.

<https://paramshakti.iitkgp.ac.in/faq/>

<http://www.hpc.iitkgp.ac.in/HPCF/faqs>



The screenshot shows the 'Frequently Asked Questions' page on the Paramshakti website. The page has a blue header with a search bar and a navigation menu on the left. The main content area is divided into three columns: a left sidebar with a list of topics, a central 'General FAQ' section with a table of contents, and a right sidebar with a search bar. The table of contents lists various HPC-related questions and their corresponding answers.

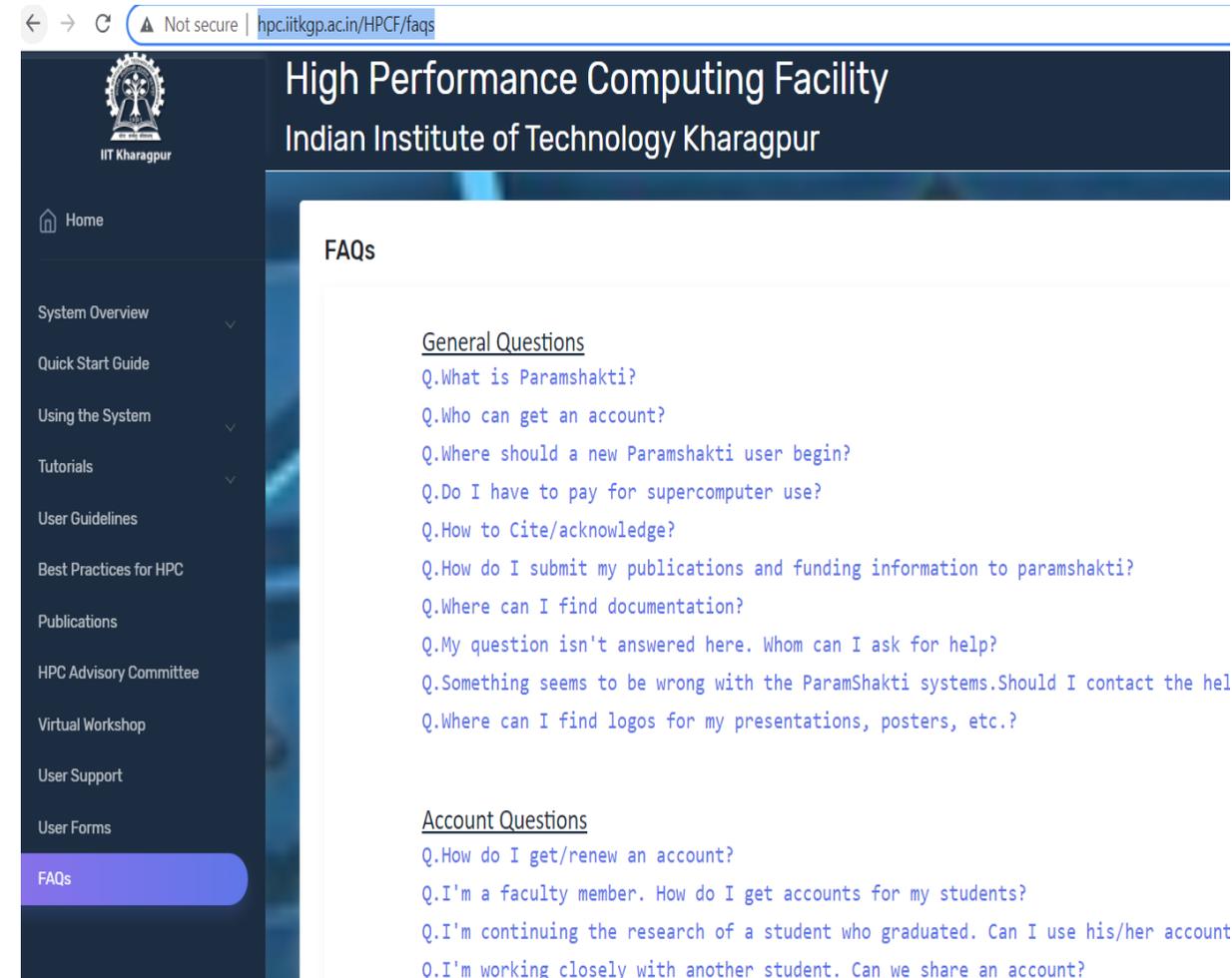
Frequently Asked Questions

- General FAQ
- Environment
- Job Submission
- HPC Applications guide
- HPC Applications
- ML / DL
- Visualization
- Spack
- Best Practices
- Hardware Specifications
- Help
- Create New Ticket
- User Manual

General FAQ

Table of contents

- How to get account on HPC cluster ?
- How do I Access the HPC Cluster ?
- What if, Error : Disk quota exceeded ?
- I get "Disk quota exceeded" error message when trying to remove files. What can I do?
- What if, SCP not functional ?
- What if, Error : Out of memory / segmentation fault ?
- What if, Error :ERROR : Bad Interpreter ?
- Can I run MS Windows applications on HPC?
- How to Access Internet on HPC Cluster ?
- How much of the file/space quota I have used ?
- My account expired ! What should I do? Is my data gone forever?
- I get a message about bad permissions
- I got an email "Please do not run jobs on login nodes"



The screenshot shows the 'High Performance Computing Facility' (HPCF) FAQ page on the IIT Kharagpur website. The page has a dark blue header with the IIT Kharagpur logo and name. The main content area is divided into two columns: a left sidebar with a navigation menu and a right sidebar with a search bar. The table of contents lists various HPC-related questions and their corresponding answers.

High Performance Computing Facility
Indian Institute of Technology Kharagpur

FAQs

General Questions

- Q.What is Paramshakti?
- Q.Who can get an account?
- Q.Where should a new Paramshakti user begin?
- Q.Do I have to pay for supercomputer use?
- Q.How to Cite/acknowledge?
- Q.How do I submit my publications and funding information to paramshakti?
- Q.Where can I find documentation?
- Q.My question isn't answered here. Whom can I ask for help?
- Q.Something seems to be wrong with the ParamShakti systems.Should I contact the hel
- Q.Where can I find logos for my presentations, posters, etc.?

Account Questions

- Q.How do I get/renew an account?
- Q.I'm a faculty member. How do I get accounts for my students?
- Q.I'm continuing the research of a student who graduated. Can I use his/her account
- Q.I'm working closely with another student. Can we share an account?

Navigating through the supercomputing website

- All the important information discussed so far has been captured in the website.

<http://www.hpc.iitkgp.ac.in/>

Let me browse through few important tabs of the website as given below.

Quickstart Guide: <http://www.hpc.iitkgp.ac.in/HPCF/quickGuide>

Best Practices: <http://www.hpc.iitkgp.ac.in/HPCF/pracHPC>

User Policy : <http://www.hpc.iitkgp.ac.in/HPCF/userpol>

User Forms : <http://www.hpc.iitkgp.ac.in/HPCF/userforms>

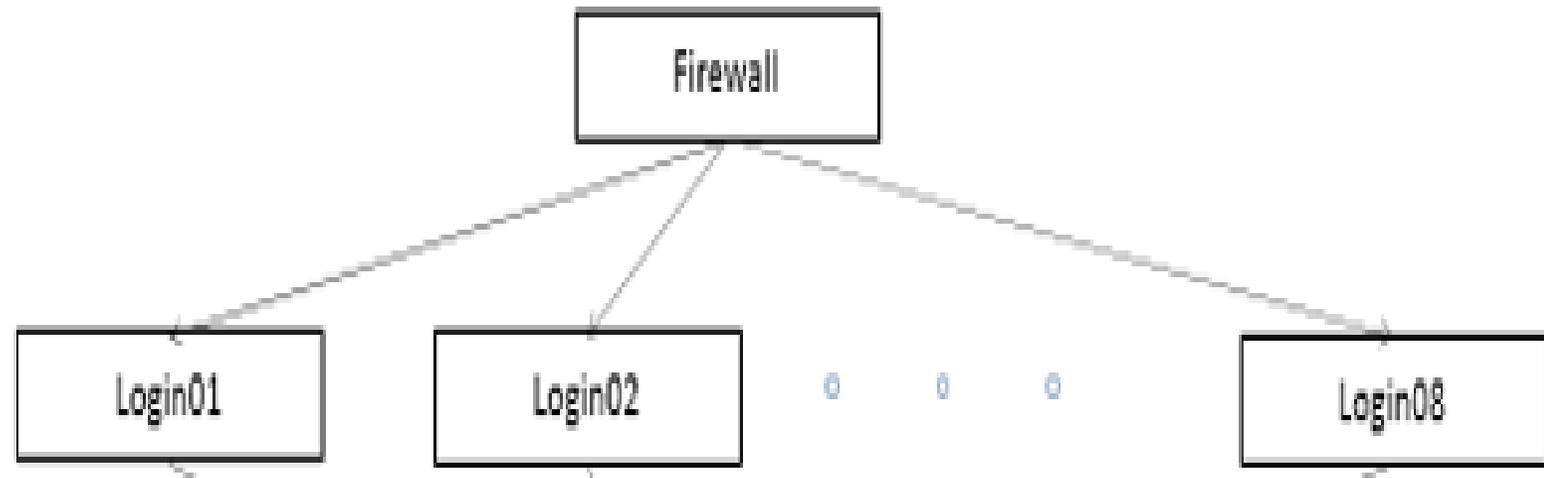
Publication : <http://www.hpc.iitkgp.ac.in/HPCF/publications>

File transfer: <http://www.hpc.iitkgp.ac.in/HPCF/fileTransfer>

Accessing Supercomputing Machines

Access Policy

- Access to login nodes are in round-robin mode.
- Login Nodes are not meant for running Jobs. After successful login an user will land in his home directory `"/home/<loginname>"`
- All users are given 50G quota in their "home" directory . Users should submit their computational job only from respective "scratch" directory.
- Login nodes are used for,
 - submit jobs to Compute Nodes and monitoring
 - compilation/installation of an application
 - editing and saving the files



Accessing Supercomputing Machines

How to access PARAM Shakti ?

- Linux users can use OS native SSH client.
- If you are using windows you can access via(SSH Clients):
 - MobaXterm
 - Putty, etc
 - Native SSH Client(Windows10)
- **Within IITKGP campus:**
 - ssh username@paramshakti.iitkgp.ac.in
- **Outside IITKGP campus:**
 - ssh username@paramshakti.iitkgp.ac.in -p 4422

Accessing Supercomputing Machines

Linux Users:

Open the Terminal, and type the ssh command as below:

```
$ ssh <your_user_name>@paramshakti.iitkgp.ac.in (accessing inside campus)
```

```
$ssh -p 4422 <your_user_name>@paramshakti.iitkgp.ac.in (accessing outside campus)
```

It will prompt for captcha and then password, after successful login, You are on to your home directory. You can use “ssh -X” to access graphics interface of the application.

```
[root@centos7vm1 ~]# ssh -X -p 4422 testuser10@paramshakti.iitkgp.ac.in
If you truly desire access to this host, then you must indulge me in a simple challenge
.
-----
Observe the picture below and answer the question listed afterwards:

( a | t | t | D | R | p | B | s )

Type the string above: attDRpBs
Password:
Last login: Wed May  5 17:44:06 2021 from 110.224.166.223
  Filesystem            used      quota   limit  grace  files   quota   limit  grace
   /home                136k      40G     50G    -      29      0      -      -
  /scratch              92k       2T      0k     -      13      0      -      -
[testuser10@login06 ~]$
[testuser10@login06 ~]$ xclock &
[1] 4160
[testuser10@login06 ~]$ Warning: Missing charsets in String to Fonts
```



Supercomputer software environment and using software modules

- Every time an user logs into the Paramshakti Supercomputer by default there is no application or software environment is set.
- So how to set software Environment ?
- **module** is the utility (also command name) to enable use of applications / libraries / compilers available on the HPC cluster.
- Applications / libraries / compilers available on the HPC cluster are organized as per the hierarchy given below.
 - apps/<application name>/version : Applications available on the cluster
 - compiler/<compiler name>/version : Compilers available on the cluster
 - lib/<library name>/version : Available libraries
 -

Supercomputer software environment and using software modules

- Some Important commands :
 - **module avail** To see the available software installed on HPC system
 - list of precompiled applications
 - different compilers and libraries (compilers include GNU, Intel, PGI)
 - **module list** Shows the currently loaded modules in your shell
 - **module load <Name of the module>**
 - module load compiler/intel/2018.2.199 (to set Intel compilers version 2018 in your environment)
 - module load apps/namd/2.12/impi2018/cpu (to set NAMD app version 2.12 in your environment)
 - **module unload <Name of the module>** : This Loads the module and sets the software environment.
 - **module purge** To clear all the loaded modules.

Supercomputer software environment and using software modules

- I need a conda package(s) that isn't installed, what do I do?
- You need to create your own conda virtual environment. Please follow the steps below to create and activate your own conda virtual environment .

```
$ module load python/conda-python/3.7_new
```

```
$ conda create -n myenv
```

```
$ source activate myenv
```

```
$ conda install <package-name>
```

Supercomputer software environment and using software modules

- Example given below how to create your local conda environment with name “mycudnn” for cudnn==7.6.0=cuda10.0_0 with python version=3.6

```
$ module load python/conda-python/3.7_new
```

```
$ conda create -n mycudnn cudnn==7.6.0=cuda10.0_0 python=3.6
```

→ Activate the new environment stored at “/home/<loginname>/.conda/

```
$ source activate mycudnn
```

OR

```
$ conda activate mycudnn
```

Tip :

- Install all the package(s) that you want in the environment at the same time. Installing 1 package at a time can lead to dependency conflicts. In slurm batch script must include below two statements

```
$ module load python/conda-python/3.7_new
```

```
$ conda activate mycudnn
```