



Advanced Supercomputing Hub for OMICS Knowledge in Agriculture



**Advanced Supercomputing Hub for OMICS Knowledge
in Agriculture**

National Agricultural Bio-computing Portal

(<http://ashoka.cabgrid.res.in/>)

Centre for Agricultural Bioinformatics

ICAR - Indian Agricultural Statistics Research Institute

Library Avenue, Pusa, New Delhi - 110012 (INDIA)

website: www.cabgrid.res.in



Table of Contents:

- 1) [Client based access to HPC resources](#)
- 2) [Web based Access to HPC Resources](#)
- 3) [Access to CLC Workbench](#)

For further suggestions/queries:

Send an email to admin.hpc@icar.gov.in or hd.cabin@icar.gov.in

Phone: 91-11-25847121-24 (PBX) **Ext:** 4349, 4341

Fax: 91-11-25841564

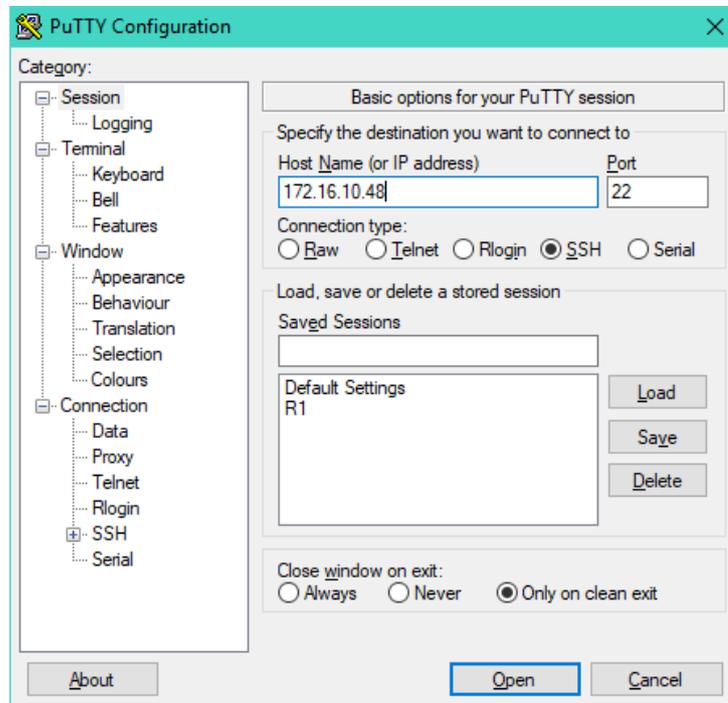
Note:

- If required software is not available in the applications category, kindly e-mail the details of the required software/tools at admin.hpc@icar.gov.in including its source and command line for its execution. We will be happy to prepare the GUI for the same and make it available to use.
- Kindly acknowledge the services received from ASHOKA as “Computing facility is availed from ASHOKA: Advanced Supercomputing Hub for Omics Knowledge in Agriculture, ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, Pusa Campus, New Delhi 110012 India” in your publication/reports. Further, it is also requested to kindly send a copy of your publication to us for our record.

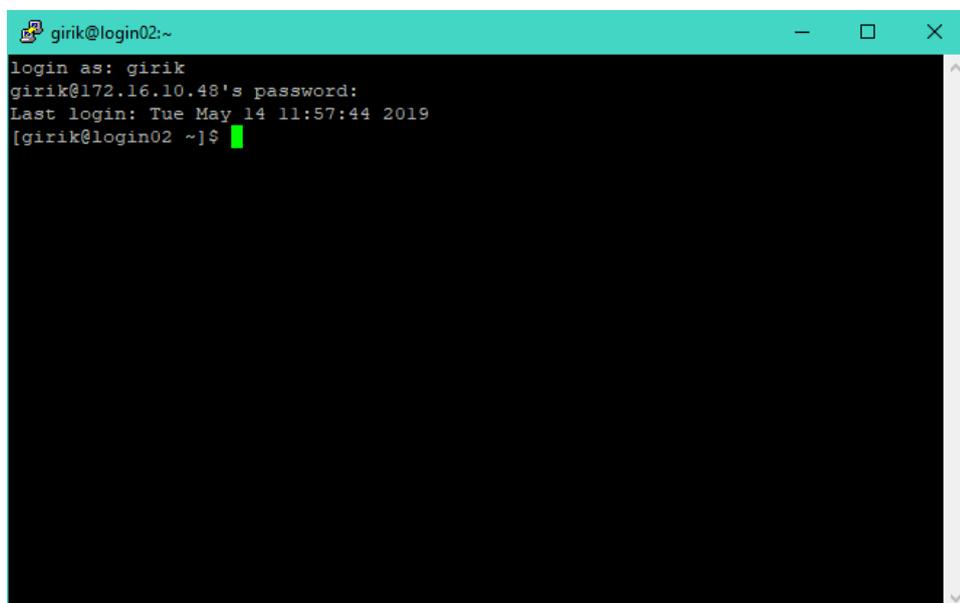
Client based Access of HPC Resources

The following steps are useful to access HPC resources for bioinformatics data analysis:

1. Download PuTTY from the official PuTTY link - <https://www.putty.org/>
2. Open PuTTY to SSH into the HPC cluster.
 - 2.1. Enter 172.16.10.47 or 172.16.10.48 in the "IP Address" field, select "Connection Type" as SSH and "Port" should be 22. No other changes required. Click on "Open".



- 2.2. At the login prompt, enter your username and password to get the shell.



3. Once you have the shell, load the required module (Software/Tool).
 - 3.1. To check all the available modules, enter the command "module avail".

```
[girik@login02 ~]$ module avail
----- /usr/share/Modules/modulefiles -----
bowtie2/2.3      gatk-4.1.2.0      intel2019_update3 mpiblast-1.6.0    samtools-1.9      tophat-2.1.1
cmake/3.14.3    gmap              module-git        ncbi-blast-2.9.0+ soap3             trinity-2.5.1
dot             gromacs/2019     module-info      null              soapdenovo2      use.own
fasta-36.3.8g  hmmer-3.2.1      modules          R-3.5.3          T-coffee          velvet
[girik@login02 ~]$
```

3.2. To check all the loaded modules, enter the command “module list”.

```
[girik@login02 ~]$ module list
Currently Loaded Modulefiles:
  1) intel2019_update3
[girik@login02 ~]$
```

Note – The “intel2019_update3” module is already loaded by default, as can be seen in the above screenshot.

3.3. To load a module, enter the command “module load <module_name>”. Do “module list” again to see the list of loaded modules.

```
[girik@login02 ~]$ module load mpiblast-1.6.0
[girik@login02 ~]$ module list
Currently Loaded Modulefiles:
  1) intel2019_update3  2) mpiblast-1.6.0
[girik@login02 ~]$
```

Note: If required software is not available in the module list, kindly email the details of the required software/tools at admin.hpc@icar.gov.in

4. Submit your job using the through PBS (Note - #PBS is NOT a comment, it is a directive to PBS).

4.1. Sample PBS script given below:

```
#!/bin/sh
### Set the job name (for your reference)
#PBS -N Your_job_name
### Specify email address to use for notification.
#PBS -M YourEmailAddress@icar.gov.in

### In case of serial application, specify no. of nodes as 1 and no. of
processors per node as 1
# PBS -l nodes=1:ppn=1
OR,
### In case of parallel application, select the number of nodes (1-5)
and the number of processors per node (1-30)
#PBS -l nodes=2:ppn=30

### Specify "wallclock time" required for this job, hhh:mm:ss
#PBS -l walltime=01:00:00
### Specify error file name and location
#PBS -e $PBS_O_WORKDIR/error.$PBS_JOBID
### Specify output file name and location
#PBS -o $PBS_O_WORKDIR/output.$PBS_JOBID

#PBS -l software=replace_with_your_software_name
```

```

# After job starts, must goto working directory.
# $PBS_O_WORKDIR is the directory from where the job is fired.
echo "======"
echo $PBS_JOBID
cat $PBS_NODEFILE
echo "======"
cd $PBS_O_WORKDIR

### You can load the relevant module in the PBS script itself, for eg:
module load ncbi-blast-2.9.0+

### Enter your serial command, for eg:
makeblastdb -in <input_file> -dbtype <database_type>

### For parallel program, use the appropriate module/command:
module load mpiblast-1.6.0
mpirun -np 60 mpiblast -p blastx -d <db_file> -i <input_file> -o
<output_file> -m 7

```

4.2. Submit the job using “qsub”.

```
$ qsub <script_name.sh>
```

This will return the job ID of the job you submitted in the format - <job_ID>.hn01

```

[girik@login02 ~]$ qsub pbsscript.sh
477.hn01
[girik@login02 ~]$ █

```

4.3. You can check the status of the job using the “qstat” command.

```
$ qstat <job_ID>
```

4.4. To check status historical data about a finished job:

```
$ qstat -x <job_ID>
```

```

[girik@login02 ~]$ qstat -x 477
Job id          Name                User                Time Use S Queue
-----
477.hn01       pbsscript.sh        girik                00:00:00 F workq
[girik@login02 ~]$ █

```

4.5. To check details of running jobs and its associated resources:

```
$ qstat -ans
```

4.6. You can delete any running job using the “qdel” command.

```
$ qdel <job_ID>
```

Web based Access of HPC Resources

The following steps are useful to access HPC resources for bioinformatics data analysis:

5. Open any java enabled web browser
6. Type the URL <http://ashoka.cabgrid.res.in/> and click on **HPC resources** or **direct link i.e., <https://ashoka.cabgrid.res.in:4443/pbsworks>** and local users (User working in Centre) can type URL <https://172.16.10.47:4443/pbsworks> and enter valid credentials i.e., UserId and Password. The request to get valid userId can be obtained by sending the signed copy of duly filled registration form (Available for download from the portal) through e-mail to admin.hpc@icar.gov.in with a copy to Head, CABIn (hd.cabin@icar.gov.in).



Fig. 1: User Screen

7. The user will see the following page. The page contains three panes. Top pane contains Files, Jobs and sessions including New Jobs link. Left pane consists of Jobs details and description, Job status, Services (Name of server), Applications (Software module available in the GUI), Queue (Details of queues created), Time (description of jobs on periodic basis) and job types (specifies type of jobs) etc.

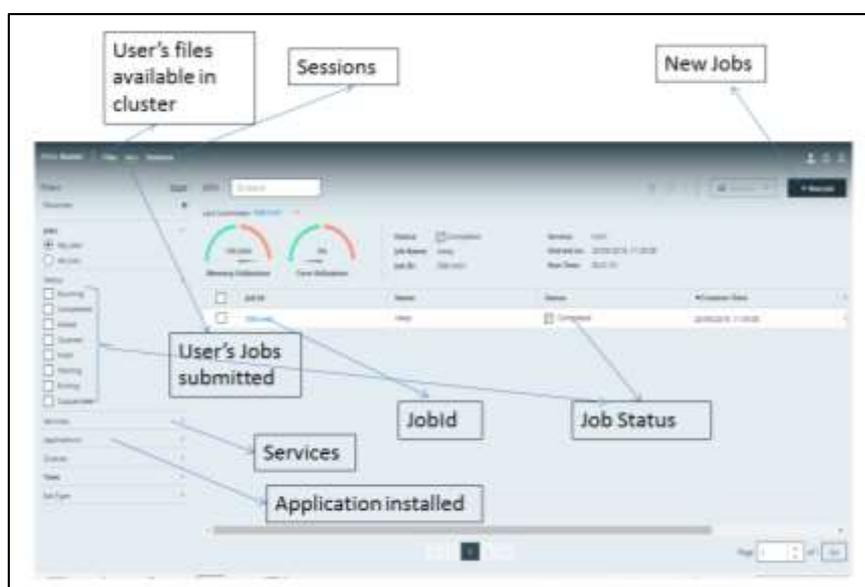


Fig. 2: Description of Job submission portal

- Now user can click on “New Job” link (Top right corner) and open the desired application for submission of data for required analysis.

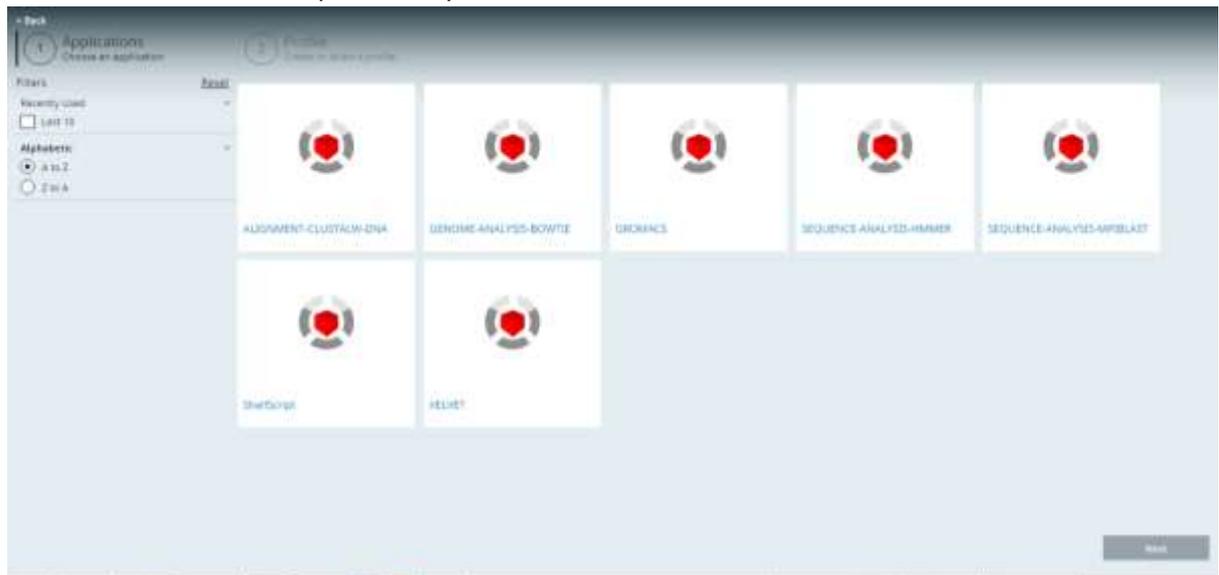


Fig. 3: List of Available Applications/Software tools

- After selecting a particular application (e.g. Shell Script). The user can select the developed shell script file from remote (server) or local (User’s desktop) system. The users can also select number of processors as per requirement. After filling up of required parameters, the user can click on Submit Jobs. After submission of jobs, the job Id will be displayed to the user. The user can track the status of the jobs by clicking on the jobID after clicking on Jobs tab (Top left pane).

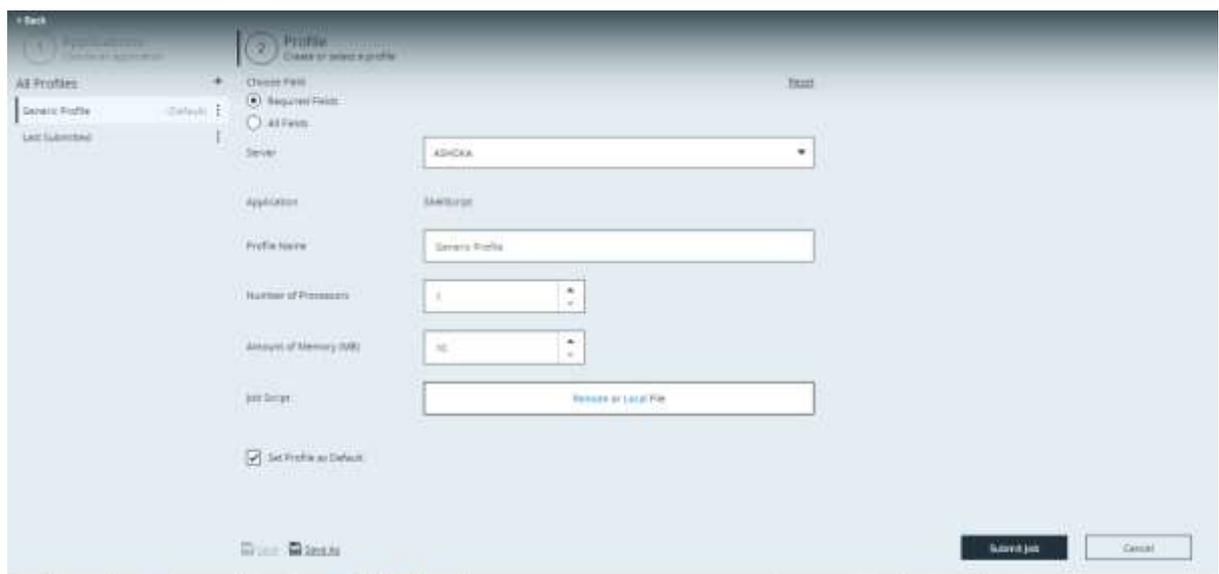


Fig. 4: Sample Job Submission Form

- The users are also having the option to choose the Required Field or All Fields (if required) to provide the details related to their Jobs.
- The other job submission portal specific help can also be obtained by clicking on user icon

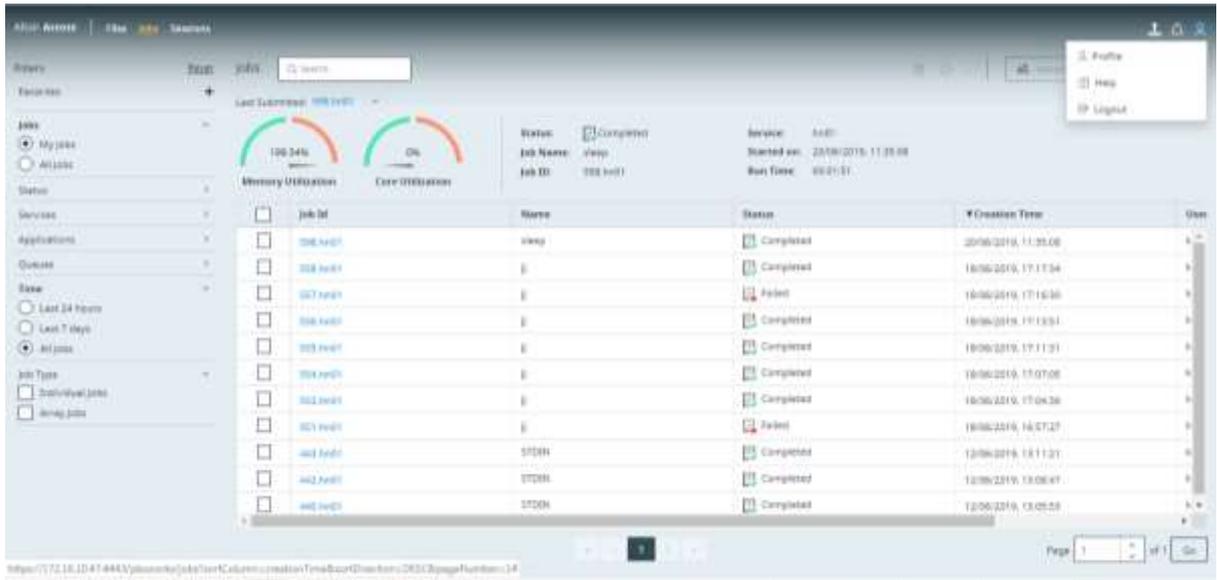


Fig. 5: Job list and its status

12. The progress of the Job can be viewed by clicking on the JobId. The details of Input (Fig. 6a), Output (Fig. 6b), Running Folder (Fig. 6c) and Details (Fig. 6d) are displayed on user's screen. The user can see the contents of any files by double clicking on the file or link (Fig. 6e).

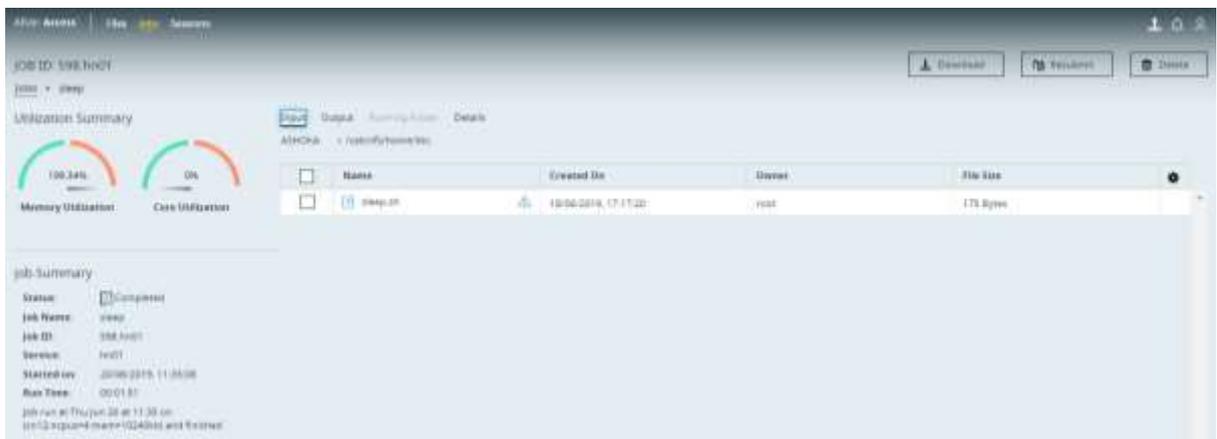


Fig. 6a: Input folder of particular JobId

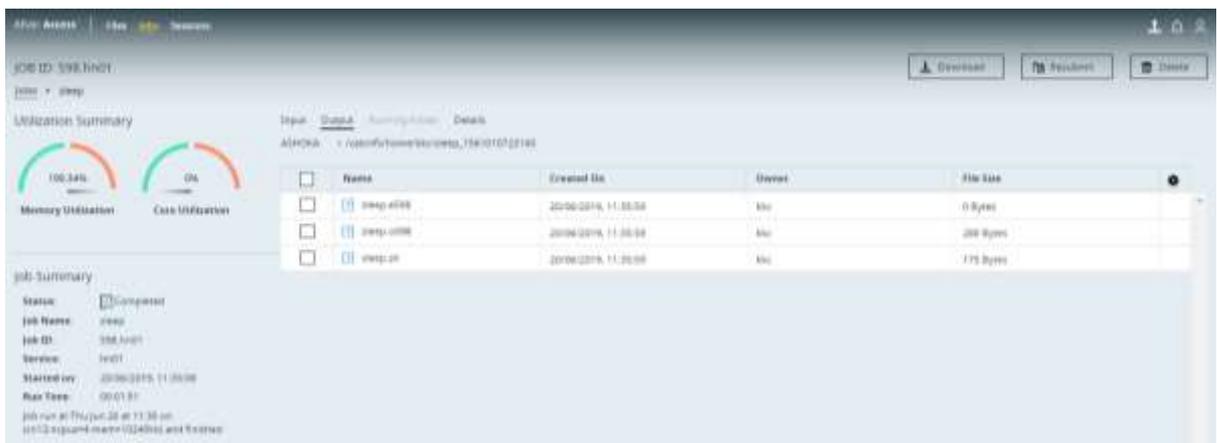


Fig. 6b: Output folder of particular JobId

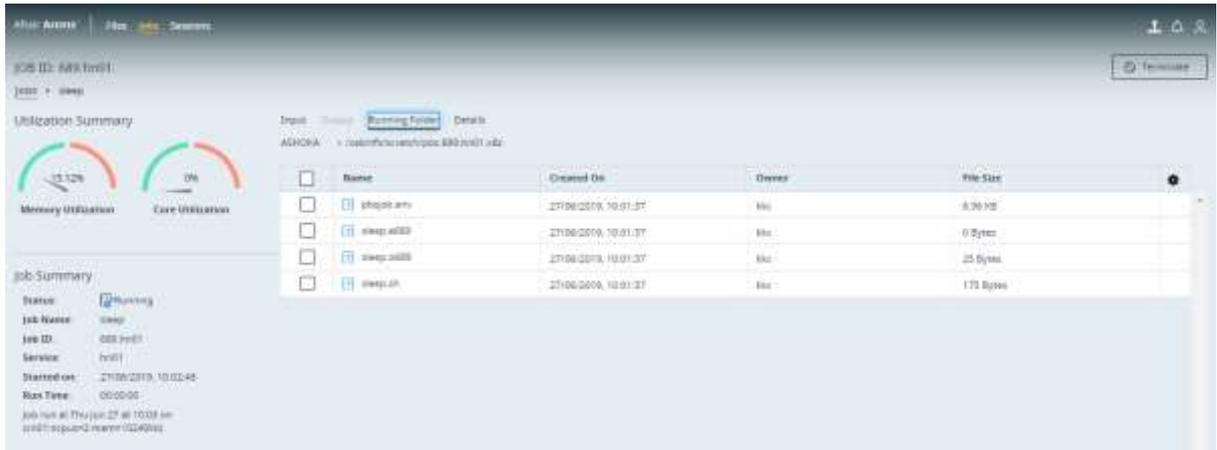


Fig. 6c: Progress of running jobs under Running Folder Tab

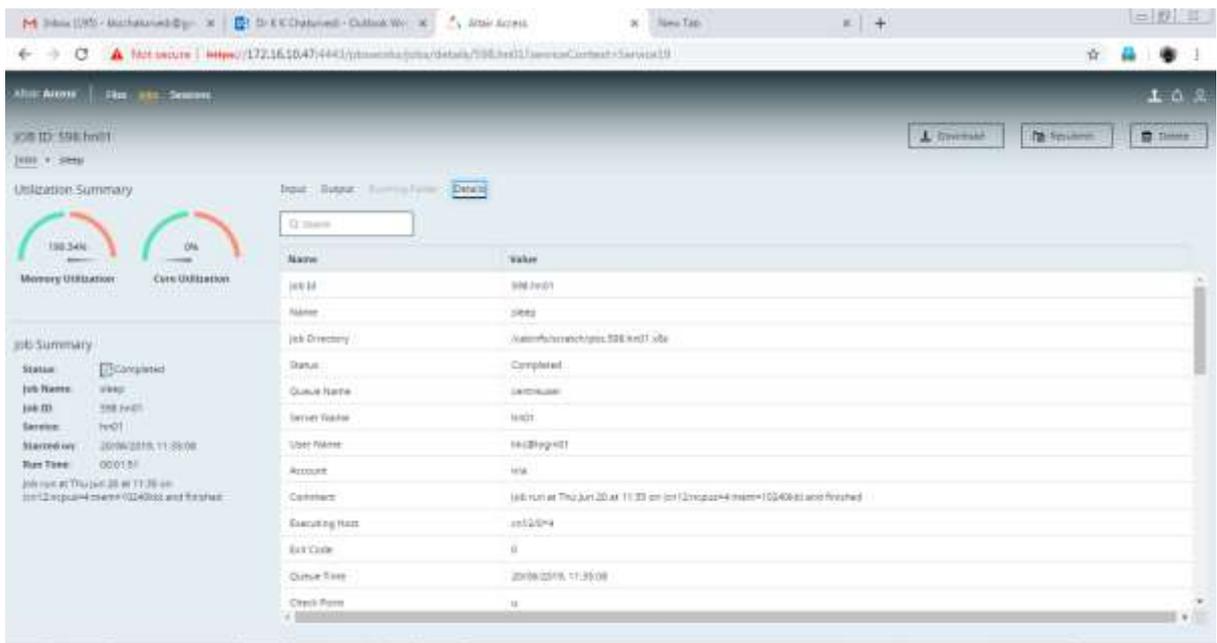
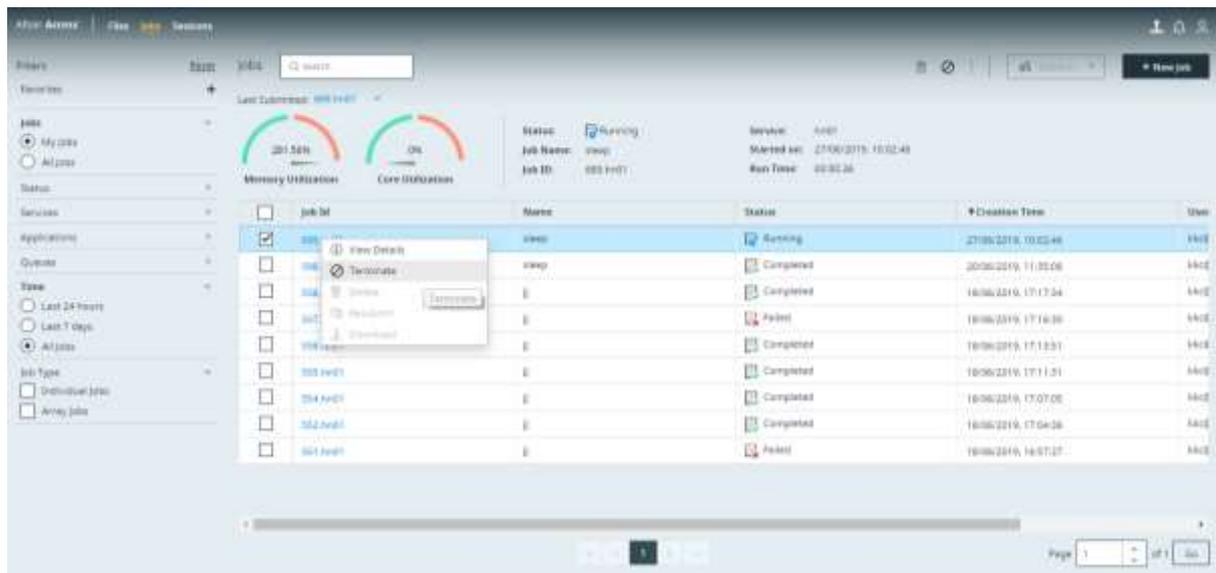


Fig. 6d: Details of JobId

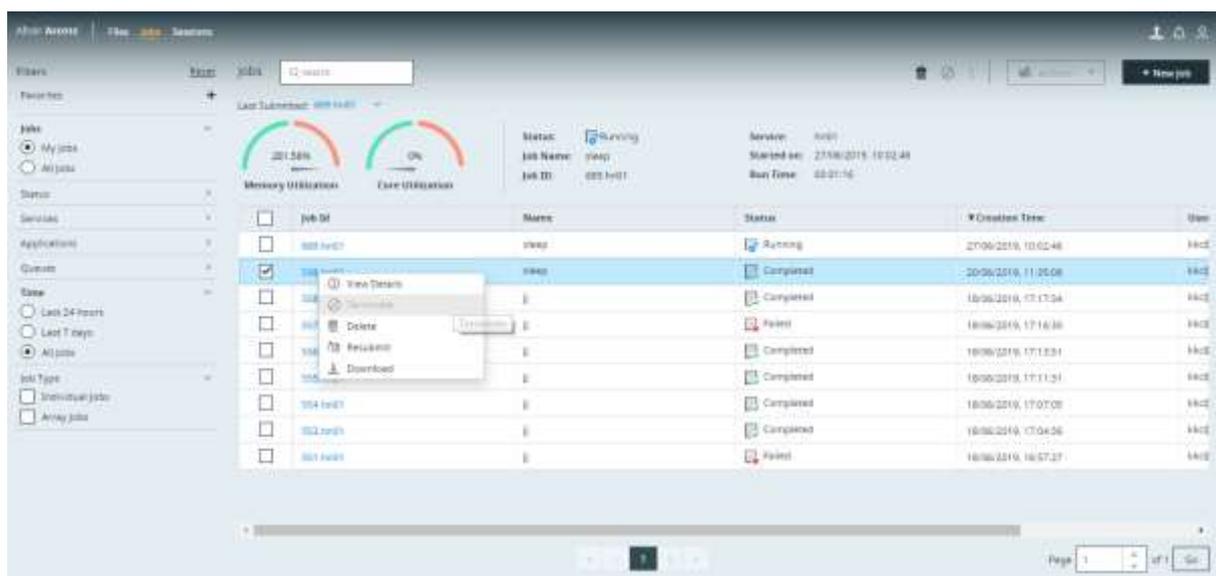


Fig 6e: Display and view the contents of the file

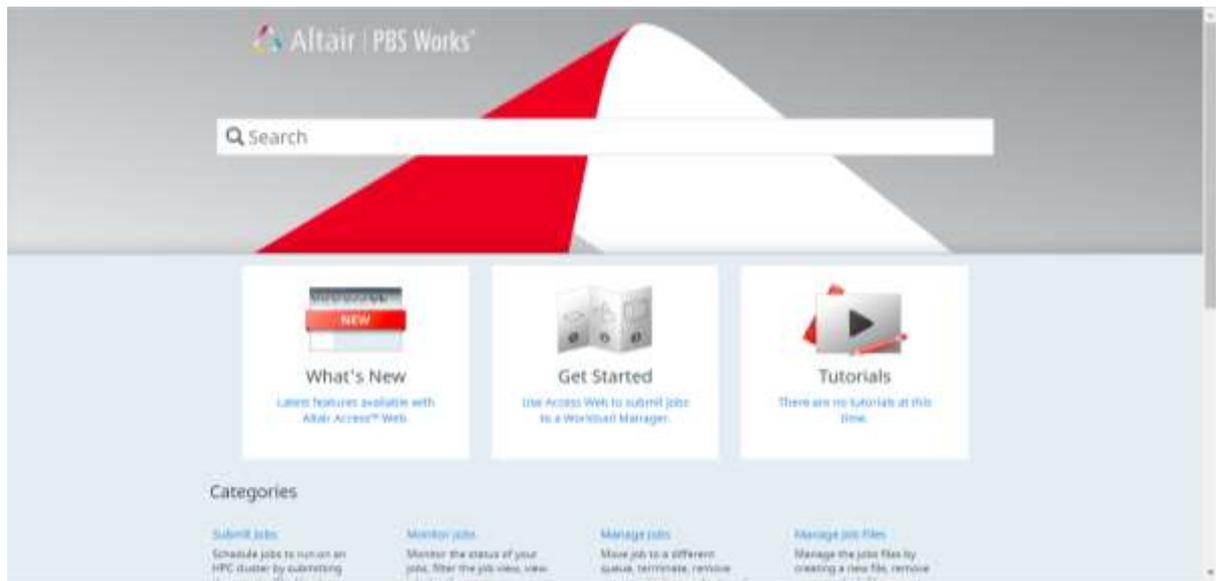
- If the user clicks on Jobs, he/she will see the summary of their Jobs and then one can choose a particular JobId listed under Jobs, right click on JobId and then click on either “Details” for detail description of Job or “Terminate” to terminate the running jobs.



- If any user wanted to submit the already completed jobs with the same parameters (Input and other required parameters), the user can check the particular JobId and then RightClick on JobId to see the various options related to “Resubmit” for resubmit the jobs, “Delete” for deleting the jobs and “Download” to download all the files and folders associated with this job in user’s desktop.



- Suitable interface and help will be available to the users. The user can get the detailed help instructions by clicking on “Ge Started” icon or on any desirable item.



16. The query/concern regarding working of particular software/tools can be posted through “**Help**” link available in the National Agricultural Biocomputing Portal.

Note:

- If required software is not available in the applications category, kindly e-mail the details of the required software/tools at admin.hpc@icar.gov.in including its source and command line for its execution. We will be happy to prepare the GUI for the same and make it available to use.
- Kindly acknowledge the services received from ASHOKA as “Computing facility is availed from ASHOKA: Advanced Supercomputing Hub for Omics Knowledge in Agriculture, ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, Pusa Campus, New Delhi 110012 India ” in your publication/reports. Further, it is also requested to kindly send a copy of your publication to us for our record.

Access to CLC-Bio Workbench

1. Open National Agricultural Biocomputing portal (<http://ashoka.cabgrid.res.in/>) and Click on CLC-Bio or local user (Working in Centre) type the URL <http://172.16.10.47:7777/CLCServer/>



2. The user can send their request to admin.hpc@icar.gov.in and copy to hd.cabin@icar.gov.in to get the valid credentials to access the CLC Bio Workbench software.
3. After receiving the valid credentials from Admin, the users can submit their jobs to our cluster through web based interface.
4. The user can also run their CLC jobs at their own desktop/workstation by obtaining the license details from admin.hpc@icar.gov.in. Licenses usage details will be shared through email. After obtaining the valid license and credentials, the user is suggested to the Download CLC-BIO Client Software and install on their PC. The links for the material is downloaded from <http://ashoka.cabgrid.res.in/biocomp/CLCBio/clc-bio.html>.
5. Now the user is ready to utilize the services being available in CLC Bio Workbench.