

Q&A: Working Effectively with HPC Systems

2021-04-20

Write a question, you're welcome to provide a name/nick. We (NSC staff) will try to answer. After the seminar, the document (possibly edited), will be available on the event page; https://www.nsc.liu.se/support/Events/HPC_course_2020/

Example:

Q: Type your question, use as many lines as you need. You can give your name or alias within the parentheses (Name)

A: NSC staff will type a response (NSC staff1)

A: There might be several responses from different staff (NSC staff2)

Q:I hope it works to type in questions, or I messed up the document rights (Weine)

A: Yes, Weine, it works fine.

Q: For the ssh-config: I believe, if the system uses one-time passwords, there is no way to log in without typing in the one-time password explicitly. Or is there? (Thanks for the useful answer!)

A: No, you are correct. If using OTP is the only option, then that is what is always required. However, often a center can provide several options and then ssh keys are considered a good and quite secure option.

Q:For big datasets (many files and many Gb of data), what is more prone to transfer errors: an uncompressed folder, or a compressed folder?

A:I'd guess that the best option is to transfer compressed folders, for a faster transfer, assuming that statistically it give fewer errors, though I'm not expert here (Weine)

A: If you are transferring a folder with many files, you should use a tool like rsync that can resume from where it broke if something goes wrong with e.g. the connection.

Q: scp command stops entirely the transfer if the connection gets lost even for a second? or is there any waiting time to reestablish the connection? rsync is definitely a much better software for transfer. Is there any disadvantage of making rsync my default transfer software?

A: I don't think there are any disadvantages with rsync if you know how to use the command.

A: I'm not an expert, but I think ssh/scp can survive minor interruptions. Not sure about the details! However, scp cannot resume a broken transfer from the point where it stopped.

Q:What is the Operating system installed on the HPC cluster?

A:Linux. At NSC we are currently running CentOS7. <https://www.centos.org/>

Q:Is there a shortcut to logout?

A: ctrl-d

Q: Of the session-preserving tools (screen, tmux, ThinLinc), which of them have to be pre-setup on the HPC side, and which ones can be used on a generic HPC?

A: screen and tmux should work everywhere. ThinLinc needs to be provided by the HPC center.

Q: Which system do you use for taping data? I mean the software used for the actual taping jobs.

A: Do you mean what system we use for the storage we backup? I believe we use a system called spectrum scale from IBM.

Q: In which situations are simple bash scripts more advantageous compared to simple py scripts?

A: With shell scripts, such as bash it is very simple to interact with the file system (e.g. copy/rename files), run linux commands, interact with the queuing system, etc. However if you are very used to writing python, then that will work as well.

Q: Is there a command for backing up data (especially huge files) from HPC to Google Drive/iCloud/Dropbox/Onedrive?

A: I'm not aware of it, but if the question is for NSC systems, you can submit a question to support@nsc.liu.se since a system expert might have a good answer (Weine)

A: On our systems, you can probably use the standard web tools for interacting with your favorite cloud storage. Don't know how convenient it is, though.

Q: Is there any solution to backup data which is Confidential and you cant have backup on your laptop/hardisk? If you have permission to use an HPC system...for example patient data, but you cannot have it on your personal computer?

A: In general you should not have any confidential data on an HPC system unless the system is explicitly designed to handle such data! Patient data can only be used on special systems that are designed to handle such data and hopefully such systems also provide some type of snapshots or secure backup solution. I'm not aware of any generic solution, but perhaps there is some type of approved data encryption for such data, which would allow you to use some non-specialized system for backup.

Q: What are good tools to monitor how my job utilises OpenMP or MPI resources?

A: There are several tools, but it depends a little on what you are looking for. To monitor processes as the job is running, you can login to the compute nodes where the job is running and use top/htop to see the number of processes, threading, memory use. There are also tools such as 'cgstat', 'collectl', 'perf top' etc. but they might not be available everywhere. Post job, you might get useful information from the queuing system as well. For slurm, there is a command called 'seff'.

A: also see the Examples section today, the .pdf is now uploaded, here's the link: https://www.nsc.liu.se/support/Events/HPC_course_2020/Examples.pdf (Weine)

A: we will also see some examples in the "interacting with the queue system" session this afternoon (Hamish)

Q: I would have a couple of questions about backup.

Just to check if I understand correctly: in your slide named "Files and Folders" from the "HPC Systems Anatomy & Storage" presentation, you (Weine) give as a tip to create a link as follows:

In `-s /proj/ourproject/users/username ourproject`

Does it mean that anything which is saved in `/proj/ourproject/users/username/` will be backed up?

Inversely, if I have a few files in my home folder that do not need to be backed up (intermediate files which are used in several projects and are nice to have but which can be recreated in a reasonable amount of time), would doing `ln -s ~/myIntermediateFiles /proj/ourproject/users/username/` mean that my intermediate files are NOT backed up? Is there another place to keep personal files which do not need to be backed up?

A: No, the softlink will not change what file system is backed up! `/home` is backed up and `/proj` is not and a softlink between these systems doesn't change anything regarding this.

A: I would simply put personal files that don't require backup in `/proj/ourproject/users/username/`

Q: Hi, when using a software provided as a module, can we upload new software data? for instance, I'm using `snpEff` and I would like to build a new database to use it with the species I'm studying, but the genome is not implemented in the version in the cluster. I downloaded the software on my own project and built the database on it, just wondering if possible to do it on the cluster one, for other users and later use. Very easy to do btw

A: If you need a special rebuild of an already installed software, then it is probably easiest to ask for help with getting a special build for you or your project. That said, I'm not familiar with `snpEff`, so I don't know how it works in this particular case.

Q: Are there good tools for managing groups of submissions/structured submissions (e.g. finishing job A launches jobs B and C etc etc) other than usual scripting with `bash/py`?

A: SLURM (`sbatch`) provides an option `--dependency` which might help (Hamish) see: <https://slurm.schedmd.com/sbatch.html> /search for "dependency")

Q: Is there the option to share user installed software with other users in the same group of the project application? If yes how?

A: Yes, if the software is installed under the project's `/proj` storage it can easily be used by everyone in the project.

Comment: You could also use containers (Singularity).

Q: My research involves improving the numerical models built in an existing simulation software. May I ask,

1) Is it possible to fork files from GitHub and push updates from HPC?

2) Is it possible to compile and run calculations with the new executable on HPC?

A: For NSC systems, the answer is yes on both 1 and 2. You can interact with GitHub from the system and you can absolutely compile software on the system. If you are asking about a continuous integration (CI) setup, then the answer is maybe a little different.

Q: About workflow management tools (such as Kepler, Pegasus, Galaxy): would any of these tools be suitable/user friendly to create a visual representation of a workflow (let's say a workflow that has multiple jobs of different softwares), but not really a workflow manager. Or

in another way: does anybody know of a tool that is great to create visuals that help keeping track of workflows (of data analysis, let's say) for reproducible data science?

A: I know that Snakemake can create a graph for your workflow but I am not sure if it suits your application/use case.

Q: Is it a good practice to develop code directly on HPC login nodes?

A: It is at least possible to do it. I would say that it is good practice to test the code on the system that you target for production jobs.

Q: How can we see the evolution of our simulation using a slurm command?

A: Evolution in terms of resource usage, or?

No, just to check the simulation, if we don't want to wait until the end

A: You can monitor the output. If the output is not written directly to a file system that you can access from the login environment, then you can (at least on NSC systems) login to the compute node running the job and check the output on the local file system. At this point you can also copy the created output to the global file system. You can even add a function to your batch script that copies output regularly during the job. Is this what you are asking? Yes, it's clear for me thank you.

Q: I might have missed it but where to find the slurm variable name associated with the different options (-c for instance)?

A: 'man sbatch' and other man pages will have sections on variables. Testing with a short job example that uses a version of the resource description you are considering and simply just prints the environment variables to a file with the env command is what I often do.

A: <https://slurm.schedmd.com/sbatch.html> gives all the details of the slurm sbatch options (Hamish)

Q: Are there ways to get ahead of the queue by asking for smaller jobs (in terms of nodes, CPUs, RAM etc)? If my job is flexible so if I ask for more CPUs it goes faster? Isn't the queue dependent on how much you ask for? Can I check what is running and in queue so I can dimension my job to 'fit' in a gap and get through faster? Thanks! Great link!

A: You can ask for fewer nodes or cores. For test jobs, there's a special queue at NSC for max. 1 node max. 1h, add #SBATCH --reservation=devel. Though, it should be used for testing. Hmmmm, I'm not sure one can ask for resources dynamically. Note that the time for your job to start also depends much on your priority, how many resources were already consumed in your compute project (check with "projinfo" command), but it also depends on the relative priority of the other users at the system. There are many users and jobs on the system, so it'll probably be tricky to figure out what is optimal. There's more info about the queue e.g. in the sections listed at the end of this page:

<https://www.nsc.liu.se/support/batch-jobs/tetralith/> (Weine)

Q: Regarding percentage of node usage: in CPU only nodes I believe that we are talking about the usage/workload of the total of CPUs in that node. What about GPU ones? Would we be able to see the usage of CPU and GPU, or what percentage will be reported with a e.g top command. So 'top' will anyway show CPU usage, correct?

A: I don't think that top shows you anything about the GPUs. For Nvidia GPUs, you use 'nvidia-smi' (e.g. 'nvidia-smi dmon') to check usage of the GPU. Yes, top still shows CPU usage also on compute nodes with GPUs.

Q: Is there any Jupyter notebook equivalent for the command line? brilliant! Thanks

A: Not sure I understand the question, but when you launch a Jupyter notebook you can also get a terminal in your browser window.

Q: Examples of tools that Peter used for examining node usage.

A: htop, collectl, collectl -sc, collectl -sm, collectl -sx, collectl -sd, collectl -scmxd, vmstat -l, mpstat, perf top --sort dso

Q: Is there any source about all these commands with their advantages over each other?

A: I will try to find a useful reference (Hamish)

Thank you.