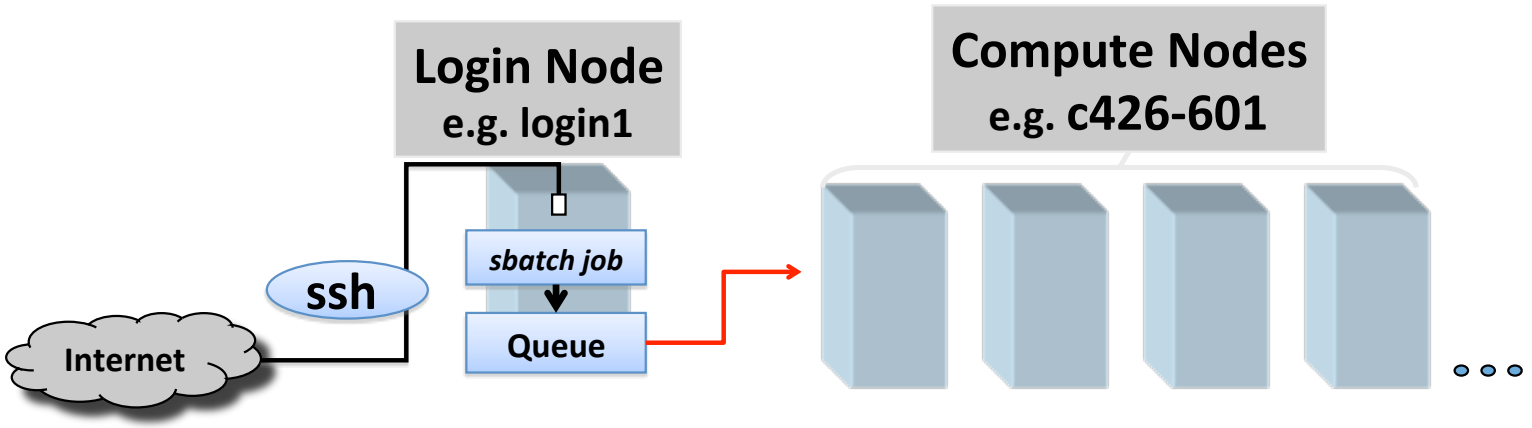
UNIX/Linux for remote access, TACC systems

(draft: still finishing the "Hands-on" exercises)

stampede2

- Note there is a user guide: https://portal.tacc.utexas.edu/user-guides/stampede
- and a ticket system: https://portal.tacc.utexas.edu/tacc-consulting

Login Nodes and Compute Nodes



"Front end" or "head node"

"Back end" Compute nodes

Do not run parallel programs on the login nodes!

Good Citizenship

- The Stampede cluster is a shared resource. Hundreds of users may be logged on at one time accessing the filesystem, hundreds of jobs may be running on all compute nodes, with a hundred more jobs queued up. All users must practice good citizenship and limit activities that may impact the system for other users. Stampede's four login nodes as well as the three Lustre file systems ($HOME, $WORK, and $SCRATCH) are shared among all users. Good citizenship can be boiled down to two items:

- Do not run programs on the login nodes

- Do not abuse the shared filesystem:

    - Avoid running jobs in the $HOME directory. Run jobs in $WORK or $SCRATCH.

    - Avoid too many simultaneous file transfers. Three concurrent scp or globus-url-copy sessions (see Transferring Files) is probably fine. One hundred concurrent file sessions is not.

    - Limit I/O intensive sessions (lots of reads and writes to disk)

stampede2 filesystem

Environmental Variable	User Size Limits	Characteristics
$HOME	10.0 GB	Not intended for parallel of high-intensity file operations Regular back ups
$WORK	1.0 TB	Not intended for parallel of high-intensity file operations Not purged, Not backuped
$SCRATCH	(30PB total)	Subject to purge after 10 days

Hands-on: ssh to stampede2

- `ssh username@stampede2.tacc.utexas.edu`
- Note the project balances, disk quotas and tips.
- Note your current directory `(pwd)`
- `cd $WORK`
  `cd $SCRATCH`

TACC's module system

- Developed from an earlier open source project “software modules”
- Load a module to make certain software available (including setting paths and environment variables)
  - Note: some software available both from the Linux installation and as module (e.g. cmake, python). The module is likely more up to date

Hands-on: module commands

- Give these a whirl

$ module list	lists currently loaded modules
$ module spider python	lists all modules with text "python"
$ module help	lists options
$ module avail	lists available modules
$ module load <module>	add a module
$ module help <module>	module-specific help
$ module unload <module>	remove a module
$ module swap <mod1> <mod2>	swap two modules
$ module spider	lists all modules
$ module reset	restore "factory settings"

Your ticket to Compute Nodes:

- Four ways to get to the back end (compute nodes):
  SLURM batch job: sbatch <batchfilename>
  SLURM interactive session: srun <flags>
  Also try idev (recommended methods)
  Run special application that connects to back end: e.g. ddt
  ssh to node on which you already have a job running

- If you don't use sbatch, srun, or equivalent, you're running on the front end (login nodes) – don't do this!

- Don't launch exe (e.g. ./a.out) on the command line
  One of the easiest ways to get your account suspended

Hands-on: idev

- (compute Pi using Monte Carlo method)

slurm

- (compute Pi using Monte Carlo method)

Hands-on: slurm

- To launch a batch job
- sbatch <batchfilename>
- To launch a one-node, sixteen core interactive session in the development queue
- $ srun --pty -n 16 -t 00:30:00 -p development -A 2016HPC /bin/bash –l
- # last char is lower case "el" (launches bash as login shell)
- # –A flag is optional unless you have multiple projects
- $ idev –n 16 –t 00:30:00 –p development –A 2016HPC
- To view all jobs in the queues: squeue | more or showq | more
- To view status of your own jobs: squeue –u <userid> or showq -u <userid>
- To delete a job: scancel <jobid>

vis portal and vnc access (will cover this later)