Week 8: Introduction to Xsede and Stampede2

Computational Tools and Techniques in STEM

Apr 2-4, 2019

Learning Goals

- **L1**: Formal introduction to unix shell.
- **L2**: Changing .bashrc.
- **L3**: Xsede account charging.
- **L4**: Batch jobs.
- **L5**: Batch submission script.
- **L6**: Monitoring queue.
- **L7**: Setting up Globus account.

Unix Shell

Motivation: Graphical user interfaces (GUI) are easy to learn and use for simple tasks. But when the tasks are complex, we need a command line interface (CLI). A shell provides a simple language and a CLI with which you can use run commands.

What is a unix shell?

A Shell is program that reads commands and run programs. Every bash command corresponds to a program stored on the computer and bash knows where to look for that program in your computer.

Are there different kinds of shell?

Yes! But the most popular one is Bash (Bourne Again SHell). It is the default shell on most unix systems.

Advantages of Using Shell

- Automation of routine tasks.
- Combining commands into single command (pipelining).
- Portable across any Unix-like operating system.
- Can handle large volume of data.
- High action-to-keystroke ratio.

Disadvantage: Can be hard for beginners and the commands can be cryptic.

How does a shell work

Let's see it with an example!

The three components are: commands, flags and arguments. They are separated by spaces. ls -a [path] E.g. ls -a lists all the files including the hidden files.

PATH

To show the path:

```
$echo $PATH
```

The $ sign before the variables is used to evaluate and output the value of the variable. By default, an initial set of path names are always assigned to PATH. But these can be removed and added from PATH either temporarily or permanently.

To add a path in your home dir permanently, open .bashrc and add the following lines:

```
# this adds anaconda to the path
export PATH="/home/prao/software/install/anaconda/bin:$PATH"
```

If you use this command in the terminal window, this path will be added to the PATH variable until the session expires.

As a precaution, to confirm with the user before deleting, add this to your .bashrc file:

alias rm='rm -i'

Xsede Account Charge

- The account is charged based on node hours.
- One service unit represents a single compute node used for one hour (a node-hour) rather than a core-hour.

SUs billed (node-hours) = (# nodes) x (job duration in wall clock hours) x (charge rate per node-hour)

Batch Jobs

Batch jobs do not start immediately.
They wait in the queue until the requested resources become available.
Stampede2 used SLURM for the workload management.
$idev

This will lead to a few messages. Wait until the command prompt ($)
shows up again.

Once you are on the interactive node, this is where you will do all your
compilation, test runs for the serial code, etc.

Production Queue and Job Submission

Queue Name	Node Type	Max Nodes per Job (assoc'd cores)*	Max Duration	Max Jobs in Queue*	Charge Rate (per node-hour)
development	KNL cache-quadrant	16 nodes (1,088 cores)*	2 hrs	1*	0.8 Service Unit (SU)
normal	KNL cache-quadrant	256 nodes (17,408 cores)*	48 hrs	50*	0.8 SU
large**	KNL cache-quadrant	2048 nodes (139,264 cores)*	48 hrs	5*	0.8 SU
long	KNL cache-quadrant	32 nodes (2,176 cores)*	120 hrs	2*	0.8 SU
flat-quadrant	KNL flat-quadrant	32 nodes (2,176 cores)*	48 hrs	5*	0.8 SU
skx-dev	SKX	4 nodes (192 cores)*	2 hrs	1*	1 SU
skx-normal	SKX	128 nodes (6,144 cores)*	48 hrs	25*	1 SU
skx-large**	SKX	868 nodes (41,664 cores)*	48 hrs	3*	1 SU

Figure: Picture source: `https: //portal.tacc.utexas.edu/user-guides/stampede2#table5`

Batch Submission Script

From the login nodes:
$sbatch myjobscript In this script, you specify:

1. computing resources requested
2. Job you want to run

When the job gets accepted, it will run in the environment which includes your currently loaded modules and the current working directory.

Slurm directives begin with "#SBATCH", followed by an option.

Monitoring Jobs and Queues

To show all jobs in all queues:

$squeue

To show all jobs owned by a specific user:

$squeue -u < *username* >

To get more info on squeue command:
$man squeue

The column labeled "ST" displays each job's status:

"PD" means pending, "R" means running, "CG" means completing.

To cancel a job:
$scancel < *jobid* >

Globus File Tranfer

Go to the page and follow instuctions to setup a globus account:
https://portal.xsede.org/data-management