

Advanced Cluster Computing

Presented by BioFrontiers IT

8/12/2020

If you haven't already checked out CC 101...

We walk through environment management with Lmod/modules

We submit a basic job

We enable ourselves to login to Fiji without a password

<https://bit.colorado.edu/biofrontiers-computing/workshops/cluster-computing-101/>

Presentation convention

This is informational text

```
$ this is a command
```

```
this is other command-line text such as a configuration file snippet  
(notice, there is no $)
```

You got an email that BIT killed your job...

Well, sorry, but we had to protect another user or the system

We try to tell you what was going awry with your job so you can fix it

We offer to help you in more depth if you submit a ticket!

Most commonly, though, you were either misusing the home directory system OR you were using too many CPUs (more than you requested)

Quick storage comparison

Home directories (/Users, /Shares) - slower, backed up!

Scratch directories (/scratch) - very fast, not backed up, ONLY should be used by running jobs

A well-written sbatch script will automatically move data between the two, and REMOVES any unnecessary source or intermediates

Automagically staging data

You can avoid the first scenario in the last slide entirely by staging all data to /scratch

```
rsync --stats $HOME/samples /scratch/$USER
```

```
cd /scratch/$USER/samples
```

```
zcat *.fastq.gz >> concat.fastq
```

```
rsync --stats /scratch/$USER/samples/concat.fastq $HOME/samples
```

```
rm -rf /scratch/$USER/samples
```

You got an email that your job failed...

There's a lot of reasons a job can fail, and we can't list them all here

What we can do is look at the common ones

```
$ sacct -j <jobid> [--format]
```

```
$ sacct -j <jobid> --format=ReqMem,MaxRSS
```

Watching a running job for potential issues

We have a few tools available to do this type of thing - some are available through the scheduler, some require you to login to the node with the job

While your job is running:

```
$ sstat -j <jobid> [--format]
```


Watching a running job for potential issues

Alternatively, you can login to a node to get better, more realtime information

Find the node the job is running on

```
$ squeue -u <your IdentiKey>
```

Login to the node

```
$ ssh <nodename> #(e.g. ssh fijinode-54)
```

Use htop or dstat

```
$ dstat -tcmsn -N bond0.100
```

A bit of etiquette as you scale-up your workload

Just because you got a single instance of some job type running doesn't mean you should jump into running it 5000 times

Think about how to scale the workflow most efficiently based on number of samples, shared inputs and outputs, and if you have a deadline to meet

Ask BIT if you're unsure about this step - sometimes we don't see when a single job isn't running properly, but we certainly see when 5000 aren't running properly

Sometimes it's okay for many to be running at once, but it's necessary to stagger when they actually start

Questions?

Email bit-help@colorado.edu

Check out our documentation - <https://bit.colorado.edu/biofrontiers-computing/fiji/>

Office hours TOMORROW, 2-3pm:

<https://bit.colorado.edu/calendar/biofrontiers-computing-office-hours/>