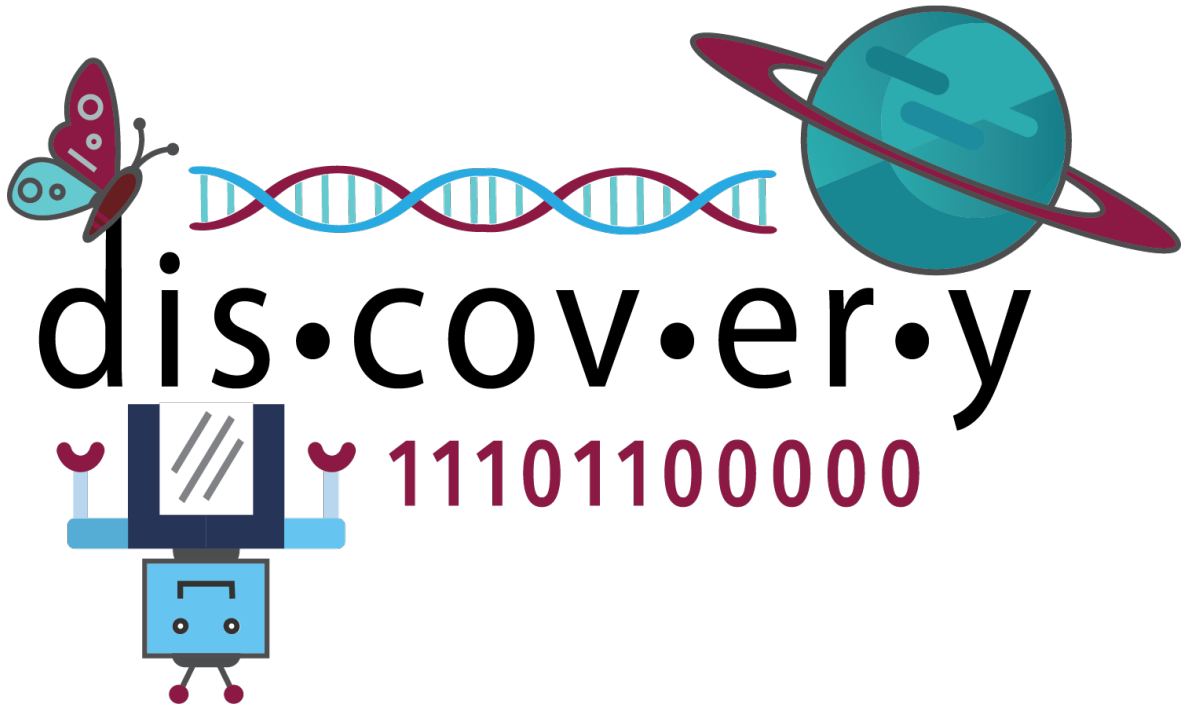
HPC Job Submission

BIOL 435/535: Bioinformatics

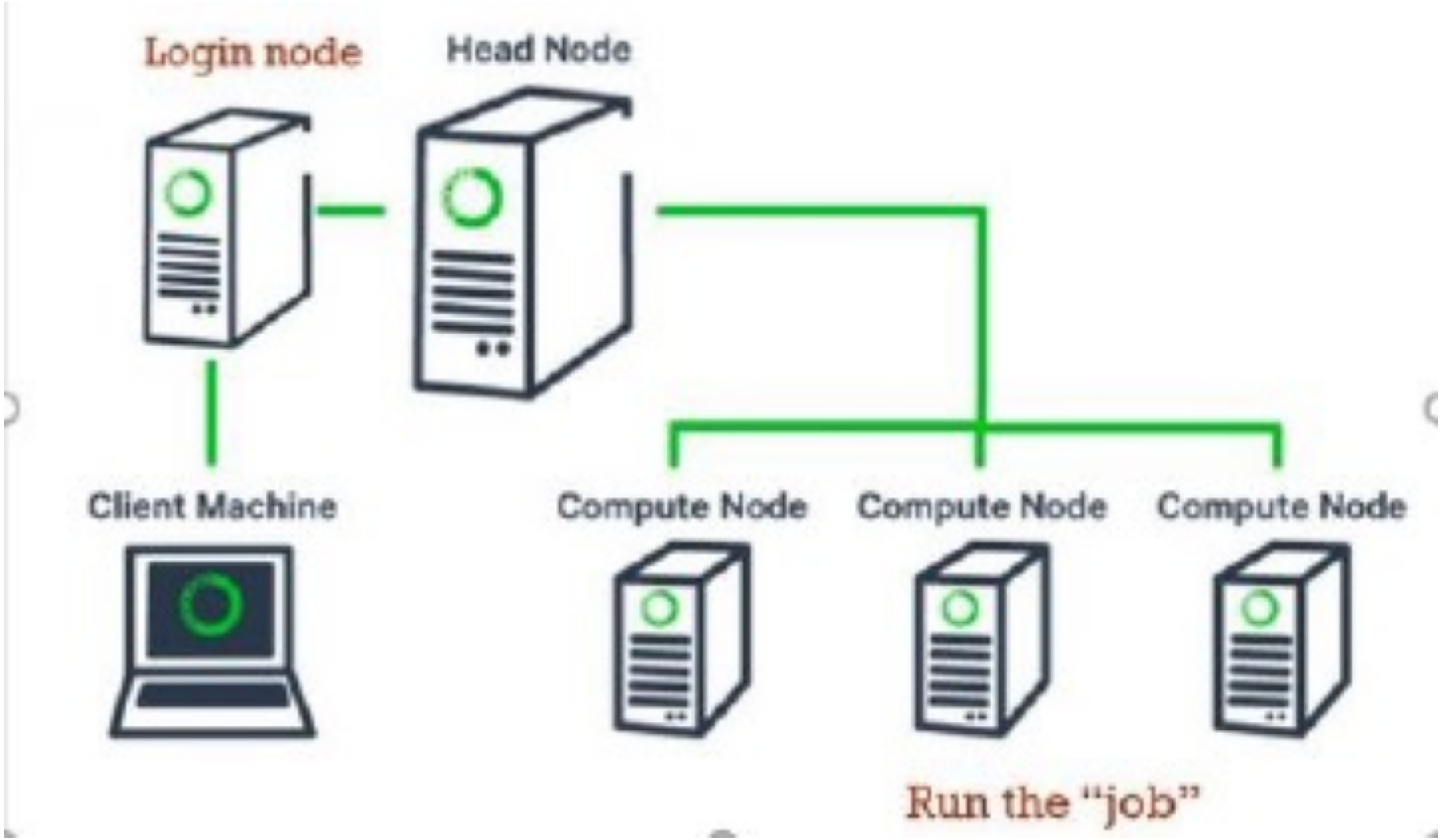
March 22, 2022

Discovery HPC at NMSU

https://hpc.nmsu.edu/discovery/



Job Submission



Job Submission

What language is this script in?

```
#!/bin/bash
#SBATCH --array=0-999
#SBATCH --requeue=
#SBATCH --time=24:00:00
#SBATCH --mem=1G
#SBATCH --partition=class
#SBATCH --job-name=blastn
#SBATCH --ntasks=24
#SBATCH --nodes=1
#SBATCH --cpus-per-task=1
#SBATCH --mail-user=joel.sharbrough@nmt.edu
#SBATCH --mail-type=ALL
#SBATCH --error=blastn.err
#SBATCH --output=blastn.out

module load blast-plus/2.9.0-gcc-9.2.0-5bhhtpv

blastn -db mtDNA.fasta -query query.fasta -outfmt 6 -max_target_seqs 1 -out mtDNA.out -num_threads 24
```

Resource request

Load required modules

Command you want to run on a compute node

Job Submission

```
[user@discovery-l2 cwd]$

[user@discovery-l2 cwd]$

[user@discovery-l2 cwd]$ sbatch job.sh
Submitted batch job 547353

[user@discovery-l2 cwd]$
```

Unix basic commands

command -options standard_input > standard_output

standard_error will print to the screen

Some additional considerations

~/.bash_profile (or ~/.bash_rc) file controls your environment

Object-oriented programming, and variable assignment

- $HOME
- $USER
- $PATH

Don’t use spaces or weird characters in file names

- Escape character allows you to deal with them

(e.g., dont\ name\ your\ files\ like\ this\ .txt)

Unix basic commands (and Google is your friend)

cd – change directory **(cd ~ = ET Go Home)**
ls – lists files inside a directory (ls -l)
pwd – print working directory
cp – copies a file from one place to another
mv – moves a file from one place to another (including renaming)
rm – removes a file/directory (-r option required for removing a directory **BE VERY CAREFUL!!**)
mkdir – make a new directory (aka folder)
echo – print to the terminal
head – shows the top part of a file
tail – shows the bottom of a file
wc -l – List the number of lines in a file
cat – print the ENTIRE file's contents
wget – retrieve a file from a url
watch – watch a command in action, updates every 2 seconds
ssh – open a secure shell (encrypted connection between two computers)
screen – open a virtual ssh session – **won't get disconnected!**
scp – secure copy protocol (encrypted file transfer)
ftp – file transfer protocol (unencrypted file transfer)
more – read a file, little-by-little

chmod – modify file/folder permissions

unzip – unzip a compressed zip file

gzip – compress a file

gunzip – unzip a .gz file

tar – make/unpack a tarball

du -h – List your disk usage

diff – Compare two files, line by line

env – Print current environmental variables to the screen

vi/vim – text editor BE VERY CAREFUL!!

nano – text editor BE VERY CAREFUL!!

emacs – Text editor BE VERY CAREFUL!!