

Introduction to High-Performance Computing (HPC)

Computer components

CPU : Central Processing Unit

cores : individual processing units within a CPU

Storage : Disk drives

HDD : Hard Disk Drive

SSD : Solid State Drive

Memory : small amount of volatile or temporary information storage



Computer components (my Macbook Pro)

Model Name: MacBook Pro

Number of Processors: 1

Total Number of Cores: 4

Memory: 16 GB

Data storage: 512 GB



*“High Performance Computing **most generally** refers to the practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business.”*

<http://insidehpc.com/hpc-basic-training/what-is-hpc/>

Computer resources required for NGS Data Analysis

100s of cores for processing!

100s of Gigabytes or even Petabytes of storage!

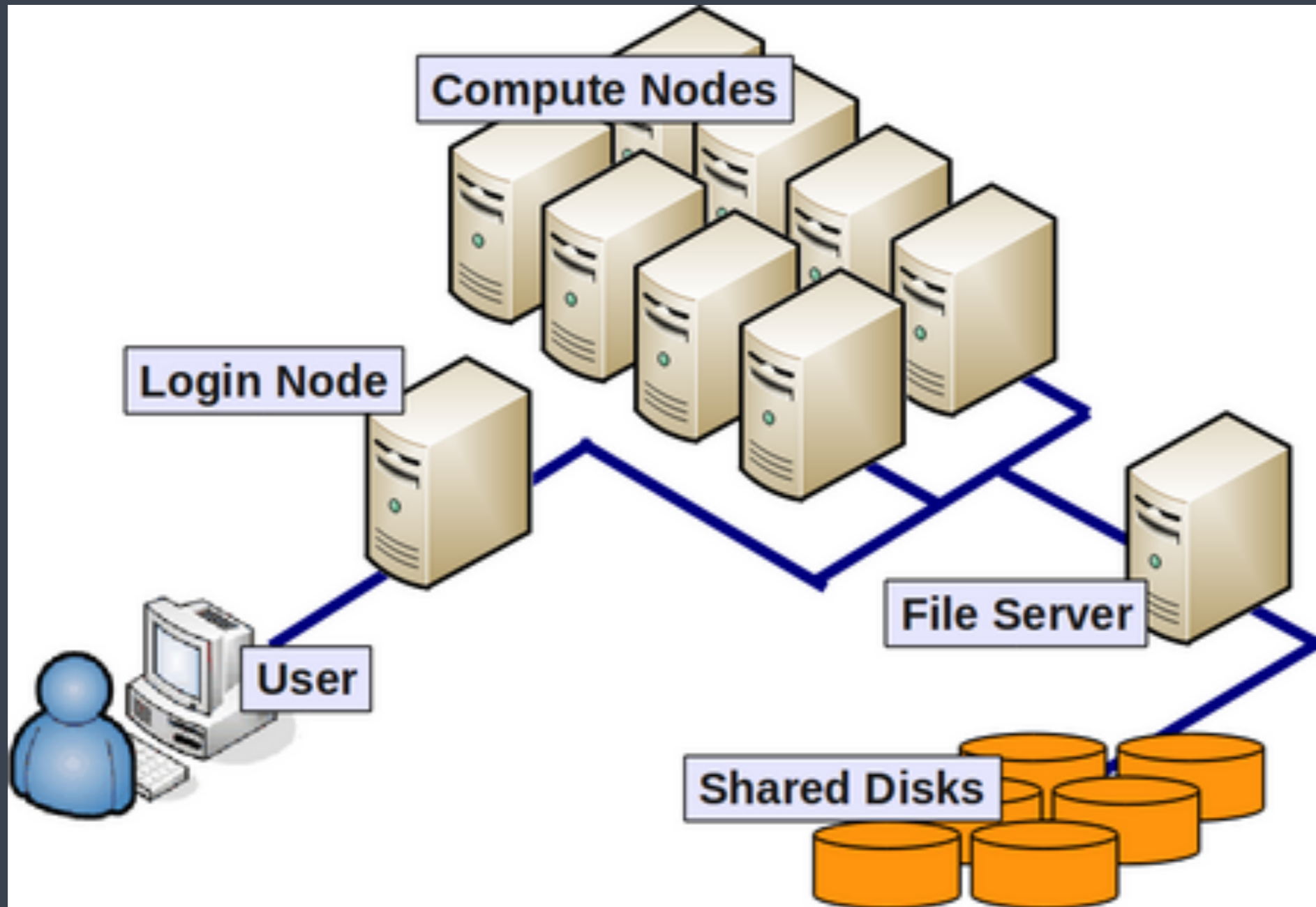
100s of Gigabytes of memory!

High-Performance Computing

Provides all the resources to run the desired Omics analysis in one place.

Provides software that is unavailable or unusable on your computer/local system

HPC cluster structure



HPC cluster components

Nodes: Individual computers in the cluster

Cores (threads): individual processing units available within each CPU of each Node

e.g. a "Node" with eight "quad"-core CPUs = 32 cores for that node.

Shared disk: storage that can be shared (and accessed) by all nodes

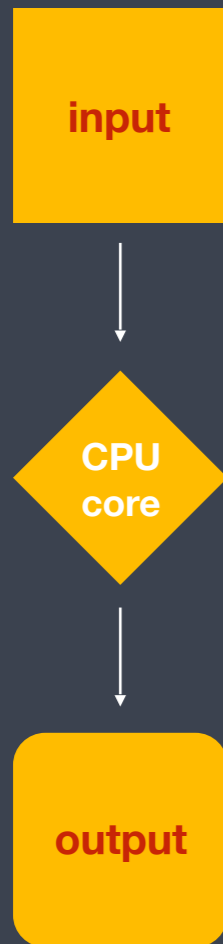
Parallel Computing

“Parallel computing is a form of computation in which many calculations are carried out simultaneously, operating on the principle that large problems can often be divided into smaller ones, which are then solved concurrently (“in parallel”).”

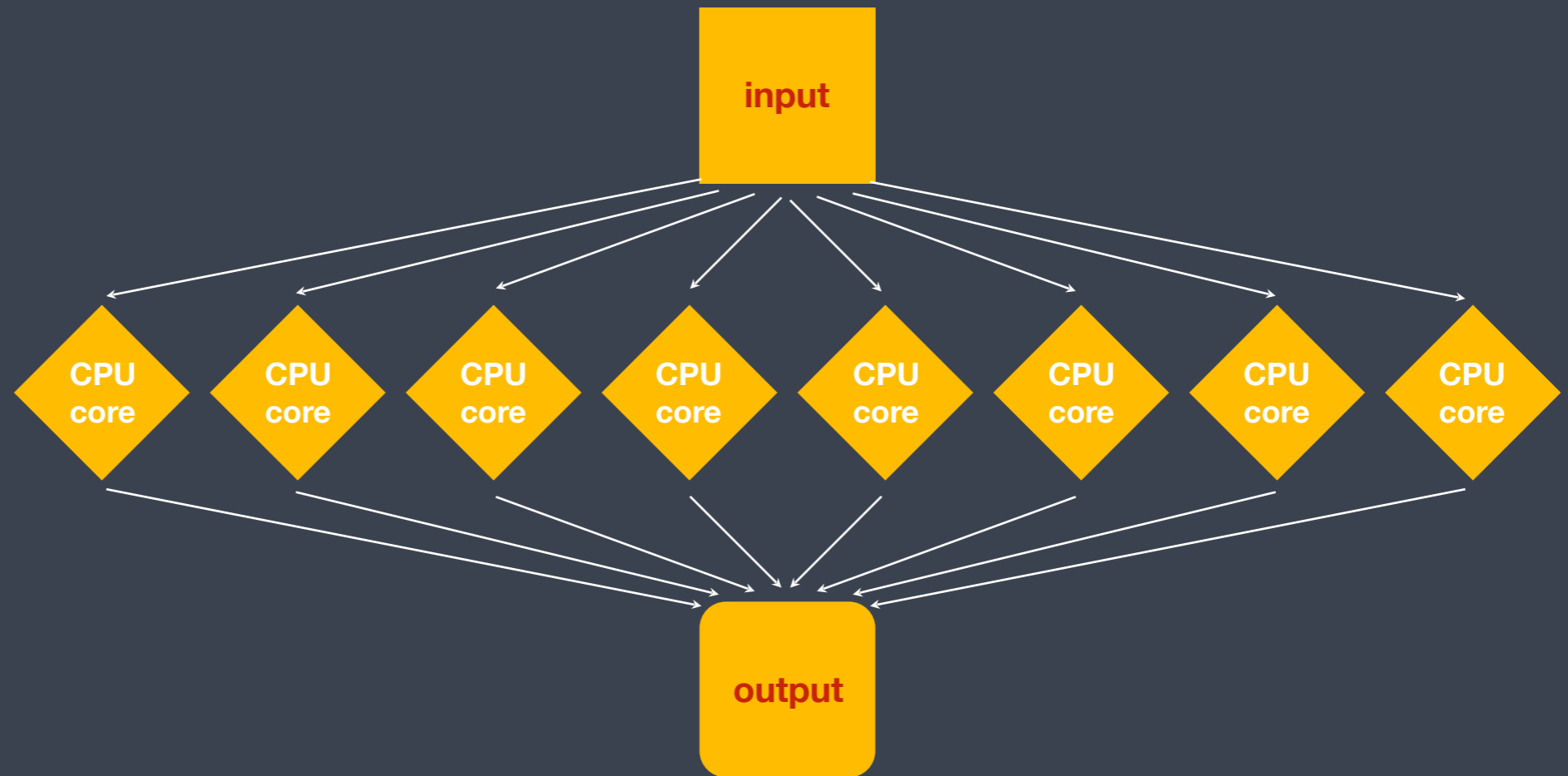
High-Performance Computing

For 1 sample

Serial



Multithreaded



Faster and more efficient...

NGS data analysis is very amenable to this strategy

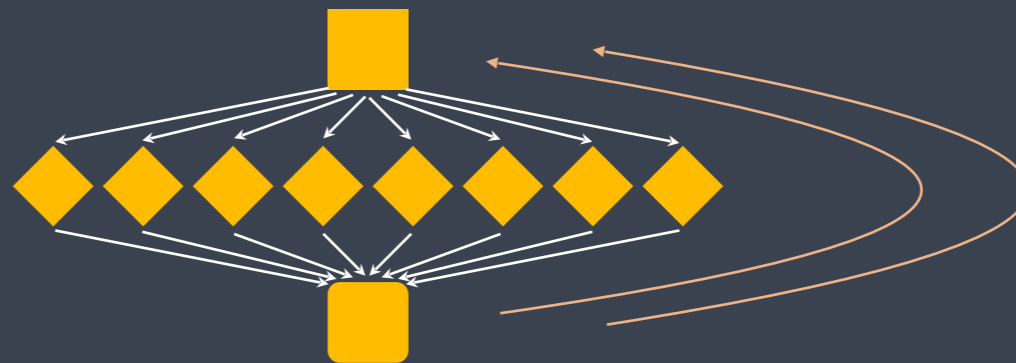
High-Performance Computing

For 3 samples

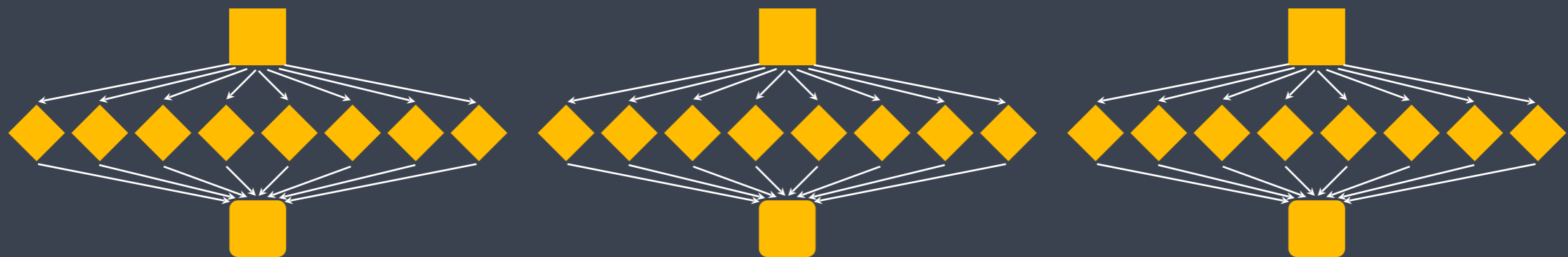
Serial



Multithreaded & Serial



Multithreaded and Parallel



HPC Cluster

- multi-user, shared resource
- lots of nodes = lots of processing capacity + lots of memory
- a system like this requires constant maintenance and upkeep, and there is an associated cost

Wiki page:

<https://wiki.rc.hms.harvard.edu/display/O2>



Tweets by @hms_rc

HMSResearchComputing @hms_rc
It's cake time for RC -- happy birthday to O2 !!



Sep 12, 2017

HMSResearchComputing @hms_rc
O2 is officially being launched today as the new RC production HPC cluster!! (thnx beta testers!) Get started at: hmsrc.me/O2docs

Sep 12, 2017

Introduction to High Performance Computing and O2 for New Users

HMS Research Computing

(Slides courtesy of Kris Holton & Kathleen Keating at HMS-RC)

Welcome to O2!

- HMS Research Computing's second High-Performance Compute cluster to enhance the compute capacity available to HMS Researchers
- Heterogeneous environment of newer, faster cores with high memory allocation to facilitate multi-core and parallelized workflows
- SLURM scheduler to efficiently dispatch jobs

O2 Tech Specs



- 11000 cores
- 32, 28, or 20 cores per node
- 256-160GB RAM (memory) per node (8-9GB/core)
- 9 756GB RAM and 1 1TB highmem nodes
- 32 GPUs (8 V100, 16 K80, 8 M40)
- CentOS 7 Linux
- SLURM job scheduler

2-Factor Authentication



- For logins using WiFi networks other than HMS Private/Harvard Secure
- Easiest: download Duo app to phone
- Similar to the setup for Harvard Key logins
- Setup details at:
<https://wiki.rc.hms.harvard.edu:8443/display/O2/Two+Factor+Authentication+on+O2>

Using O2!

1. Logging in to remote machines (securely)

- When logging in we used the “ssh” command,
ssh stands for Secure SHell
- **ssh** is a protocol for data transfer that is secure, i.e the data is encrypted as it travels between your computer and the cluster (remote computer)
- Commonly used commands that use the **ssh** protocol for data transfer are, **scp** and **sftp**

Logging Into Orchestra 2

- Open a terminal

```
ssh yourEcommons@o2.hms.harvard.edu
```

- If outside of “approved” internet sources (HMS Private/Harvard Secure):
Type 1/2/3 for DUO push/sms/phone

Welcome to O2!

Where are you in O2?

```
mfk8@login01: ~$
```

You are logged into a “**shell login server**”,
login01-05. These are not meant for heavy lifting!

```
mfk8@login01: ~$ pwd
```

You are in your home directory.

Interactive Sessions

- The login servers are not designed to handle intensive processes, and CPU usage is throttled.
- Start by entering your first job! This will (usually) log you into a “**compute** node!”

```
mfk8@login01:~$ srun --pty -p interactive -t 0-12:00  
--mem 8G bash
```

“srun --pty” is how interactives are started

“-p interactive” is the partition

“-t 0-12:00” is the time limit (12 hours)

“--mem 8G” is the memory requested

```
mfk8@compute-a:~$
```

2. Using & installing software

LMOD: Software Modules

- Most “software” on O2 is installed as an environment module.
- LMOD system adds directory paths of software into \$PATH variable, to make sure the program runs without any issues.
- Allows for clean, easy loading, including most dependencies, and switching versions.

LMOD: Software Modules

Most software is compiled against something called “gcc-6.2.0” — so, we need to load that before loading other programs that depend on it.

```
$ module load gcc/6.2.0
```

```
$ module avail #to see software now available to load
```

```
$ module spider #verbose list of all software available
```


Loading/Unloading Modules

Check module status (e.g. the alignment tool bowtie2)

```
$ module list
```

```
$ echo $PATH
```

```
$ bowtie2
```

Load the module

```
$ module load bowtie2/2.2.9
```

```
$ bowtie2
```

Which module version is loaded (if at all)?

```
$ which bowtie2
```

```
$ module list
```

```
$ echo $PATH
```

Loading/Unloading Modules

Need help with the module?

```
$ module help bowtie2/2.2.9
```

Unloading modules

```
$ module unload bowtie2/2.2.9
```

Dump all modules

```
$ module purge
```

3. The Job Scheduler, SLURM

Simple Linux Utility for Resource Management (SLURM)

- Fairly **allocates** access to resources (computer nodes) to users for some duration of time so they can perform work
- Provides a **framework** for starting, executing, and monitoring batch jobs
- **Manages** a queue of pending jobs; ensures that no single user or core monopolizes the cluster

Choosing the proper resources for your job with
the appropriate `SBATCH` options

Submitting Jobs

In an “interactive session”, programs can be run directly, however your computer will have to remain connected to the cluster for the duration of this run.

```
mfk8@compute-a:~$ bowtie2 -c 4 hg19 file1_1.fq
```

What if you wanted to run the program, close your computer and come back later to check on it?

A script with the required commands can be submitted to O2 (SLURM) using the sbatch command.

```
mfk8@compute-a:~$ sbatch mybowtiejob.sh
```

Creating a job submission script

```
#!/bin/sh

#SBATCH -p short
#SBATCH -t 0-03:00
#SBATCH -c 4
#SBATCH --mem=8G
#SBATCH -o %j.out
#SBATCH -e %j.err
#SBATCH -J bowtie2_run1
#SBATCH --mail-type=ALL
#SBATCH --mail-user=mfk8@med.harvard.edu

module load gcc/6.2.0
module load bowtie2/2.2.9

bowtie -c 4 hg19 file1_1.fq
```

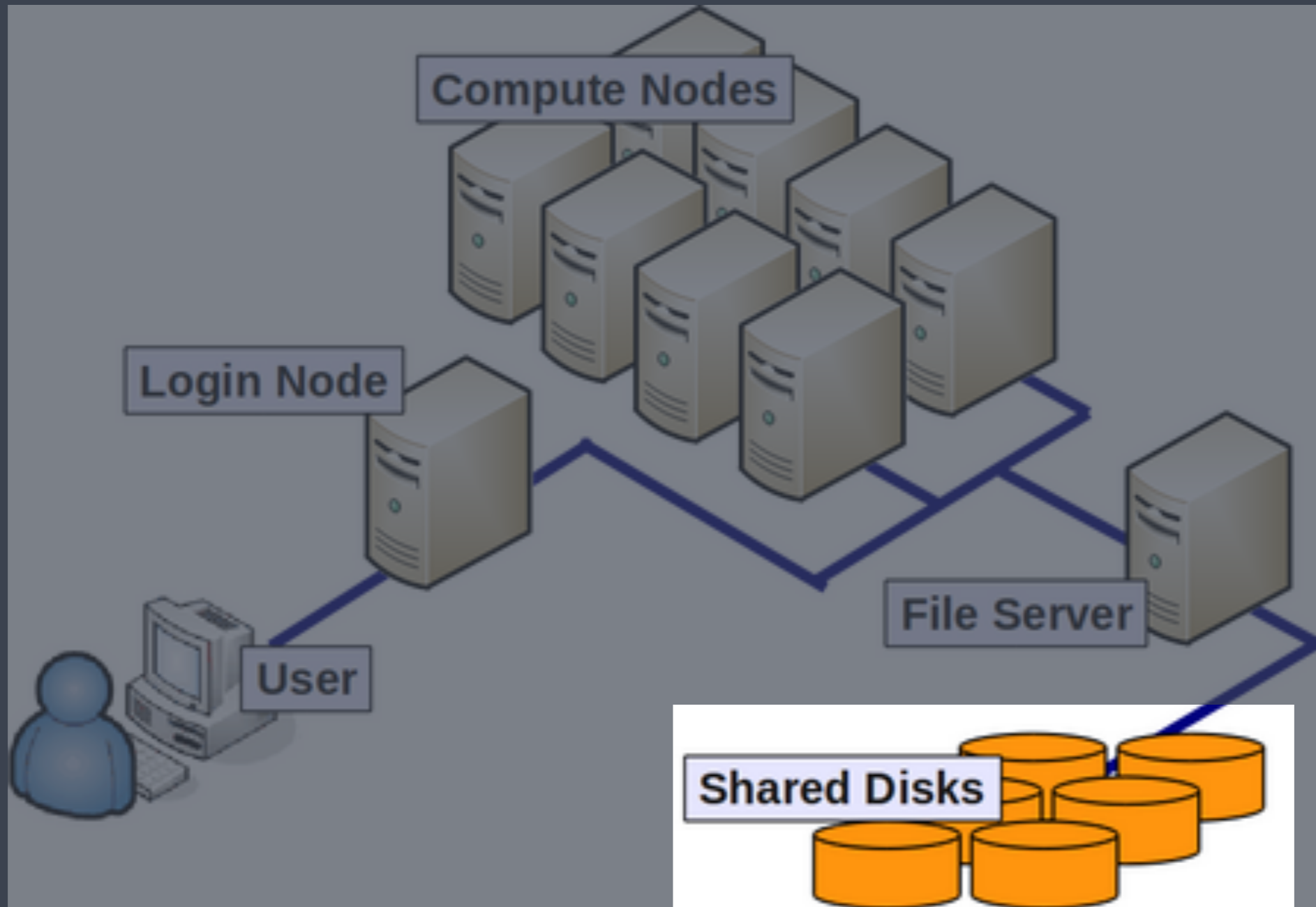
Save script as myJobScript.run and run it as follows:

```
$ sbatch myJobScript.run
```

***O2 will notify you when the job is done, or if there is an error*

4. Filesystems and storage

Filesystems and storage



Filesystems and storage

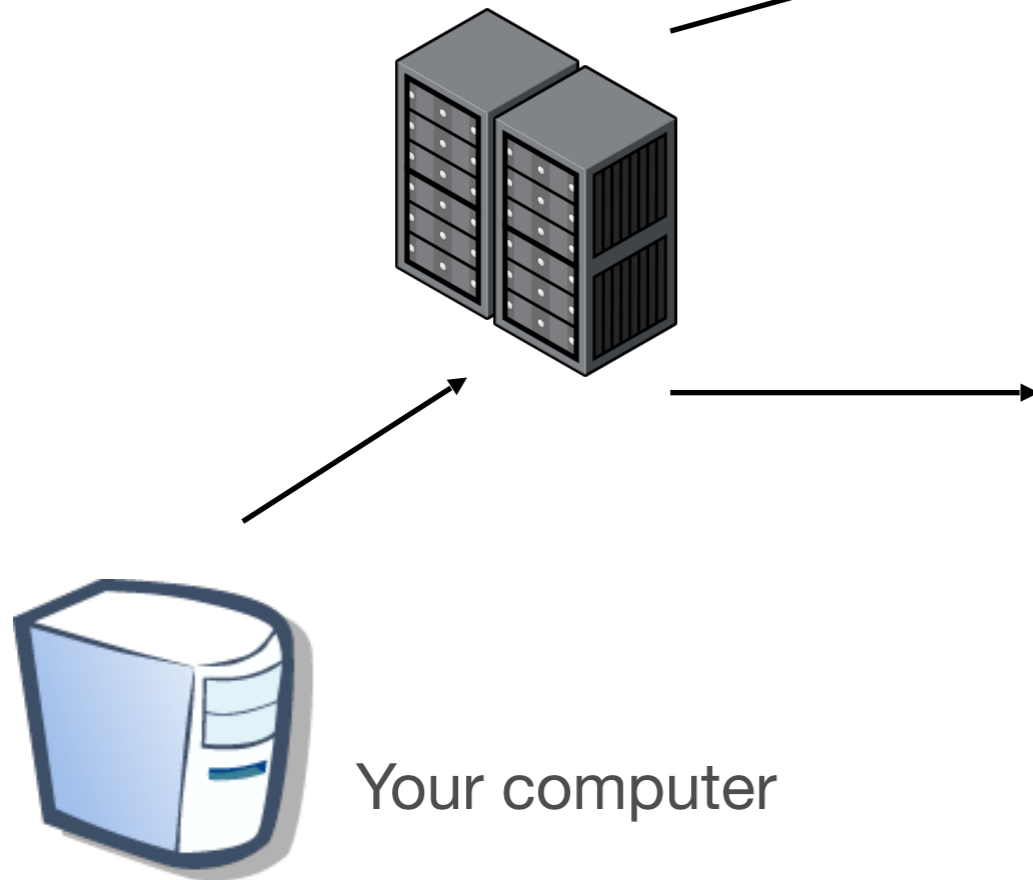
- Storage on HPC systems is organized differently than on your personal machine
- Physical disks are bundled together into a virtual volume; this volume may represent a single filesystem, or may be divided up, or partitioned, into multiple filesystems
- Filesystems are accessed over the internal network

O2 Primary Storage



O2 Cluster

- 11000+ cores
- SLURM batch system



/home

- [/home/user_id](#)
- quota: 100GB per user
- Backup: extra copy & snapshots:
 - daily to 14 days, weekly up to 60 days

/n/data1, /n/data2, /n/groups

- [/n/data1/institution/dept/lab/your_dir](#)
- quota: expandable
- Backup: extra copy & snapshots:
 - daily to 14 days, weekly up to 60 days

Temporary “Scratch” storage



- **/n/scratch2**
 - **For data only needed temporarily during analyses.**
 - **Each account can use up to 10 TB and 1 million files/directories**
- **Lustre** --> a high-performance parallel file system running on DDN Storage.
 - More than 1 PB of total shared disk space.
 - No backups! Files are automatically deleted after unaccessed for 30 days, to save space.
 - More info at: <https://wiki.rc.hms.harvard.edu/display/O2>

Tier 2 Storage

- Private cloud storage offering
- Data not frequently accessed moved to a cheaper platform
- Still visible in Isilon (/n/data1-/n/data2-/n/groups)
- Access to Tier 2 data is subject to initial delays ranging from seconds to hours based on the size of data.
- <http://it.hms.harvard.edu/services/storage-data-management/tiered-storage>

For more direction

- <https://wiki.rc.hms.harvard.edu/display/O2>
- <http://rc.hms.harvard.edu>
- RC Office Hours: Wed 1-3p Gordon Hall 500
- rchelp@hms.harvard.edu

