

# Introduction to Linux

(Tsai) Wei Wu

09/10/2024

Command Line Interface (CLI)







## Help

- Training workshops

- User Guide

[arch.jhu.edu/guide](https://arch.jhu.edu/guide)

[arch.jhu.edu/support/faq/](https://arch.jhu.edu/support/faq/)

[arch.jhu.edu/support/](https://arch.jhu.edu/support/)

- Ticket

[help@rockfish.jhu.edu](mailto:help@rockfish.jhu.edu)

- User ID at Rockfish.
- Detailed explanation of the problem/issue.
- Add snapshots if possible.
- Path to the working directory, scripts and (slurm) output files.

- Google

<https://info-ee.surrey.ac.uk/Teaching/Unix/>



JOHNS HOPKINS  
UNIVERSITY

Advanced Research  
Computing at Hopkins

[github.com/tsaiweiwu/arch](https://github.com/tsaiweiwu/arch)





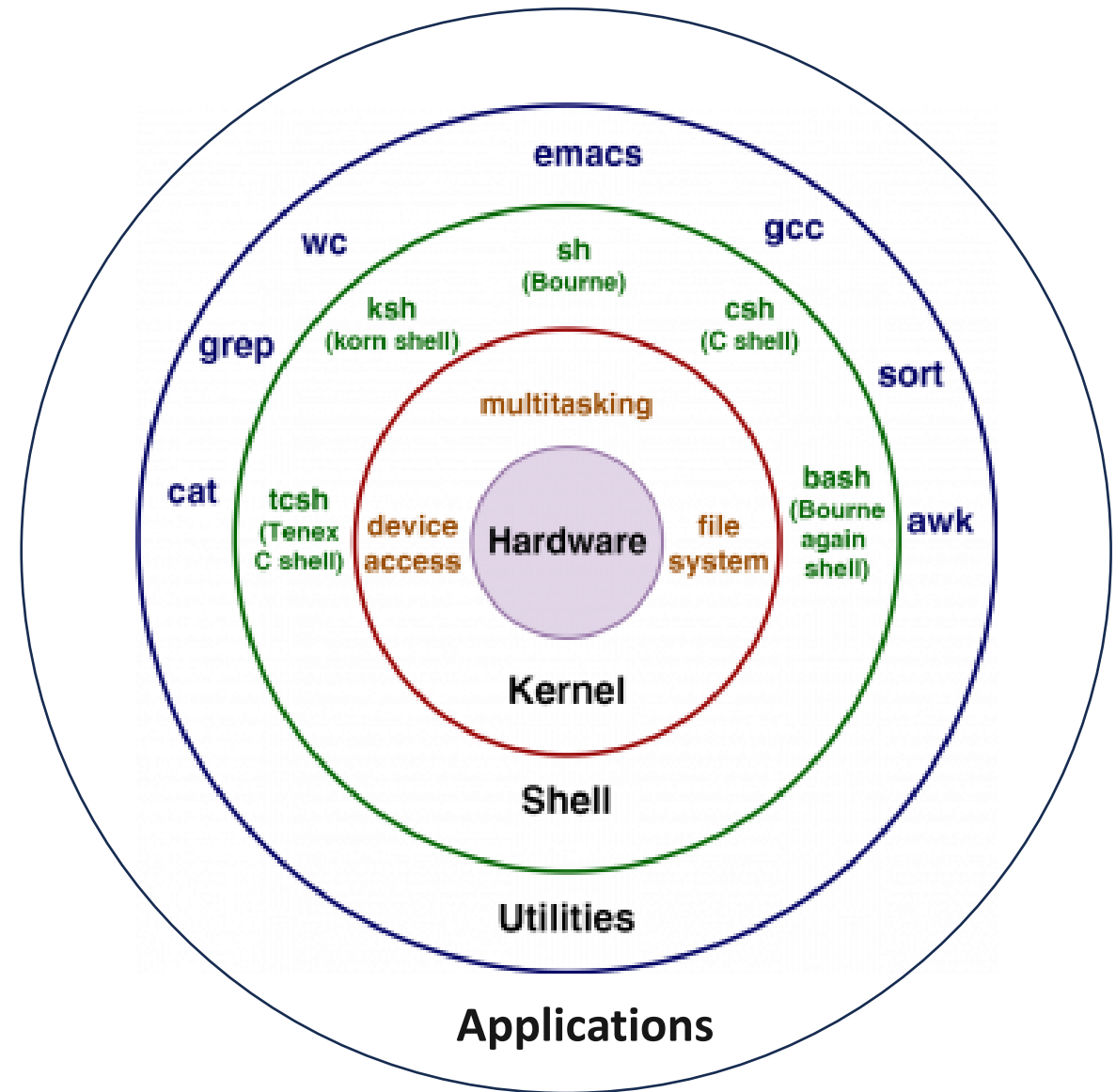
## Outline

- Linux Architecture
- Linux Distributions
- Linux File System
- Connection and File Transfer
- Basic Commands
- More Commands
- Demo
- Terminology



# Linux Architecture

- Hardware: CPU, RAM, IO, GPU
- Kernel: core of OS
- Shell: bash
- Utilities:
  - basic commands
  - text editor - vi, emacs, **nano**
  - compilers - gcc, intel, aocc, nvcc



<https://puneetpanwar.com/introduction-to-linux-command-line/>

# Linux Distributions

## Ubuntu Linux

One of the most popular Linux distributions, Ubuntu is based on Debian Linux.

## Linux Mint

A popular desktop Linux distribution, Linux Mint is based on Debian Linux.

## Puppy Linux

Puppy Linux is a lightweight Linux distro, intended to be run from removable media.

## Fedora

Fedora is a community-supported Linux distribution sponsored by Red Hat, an IBM subsidiary. Red Hat Enterprise Linux is based on Fedora.

## Debian Linux

One of the first Linux distributions, first published in 1993. Many other Linux distros, including Ubuntu and Kali, are based on Debian.

## SUSE Linux

Includes openSUSE, a community distribution; and SUSE Linux Enterprise, a commercial distribution designed for enterprise use.

## Red Hat Enterprise Linux (RHEL)

RHEL is a commercial enterprise Linux distribution, available through subscription. Customers receive patches, bug fixes, updates, upgrades and technical support.

## TAILS

TAILS, the amnesic incognito live system, is a lightweight distribution intended to preserve user privacy and anonymity. TAILS runs from removable media and is based on the Debian distribution.

## Kali Linux

Kali provides users with what amounts to a hacker's toolkit and is widely used for penetration testing by information security professionals.

<https://www.techtarget.com/searchdatacenter/definition/Red-Hat-Enterprise-Linux-RHEL>



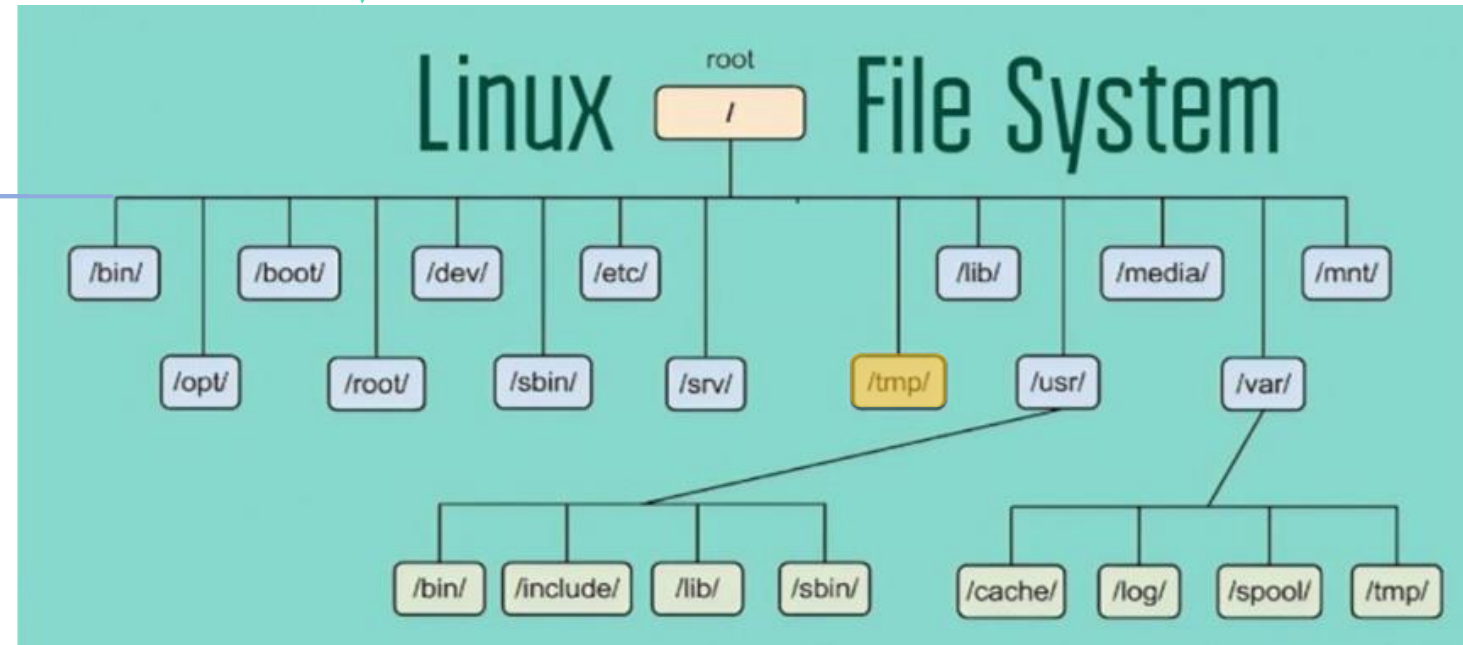
# Linux File System (Folder Structure)



- /scratch16/PI-userid/userid
- /scratch4/PI-userid/userid
- /data/PI-userid/userid
- /home/userid

node

login (01-03)  
compute  
dtn (data transfer node)  
management



<https://english.newstracklive.com/news/exploring-linux-file-system-hierarchical-structure--permissions-emc-sc71-nu384-ta384-1287279-1.html>

# Linux File System (Storage Quota)

## quota commands

```
$ du -shc /home/userid/
```

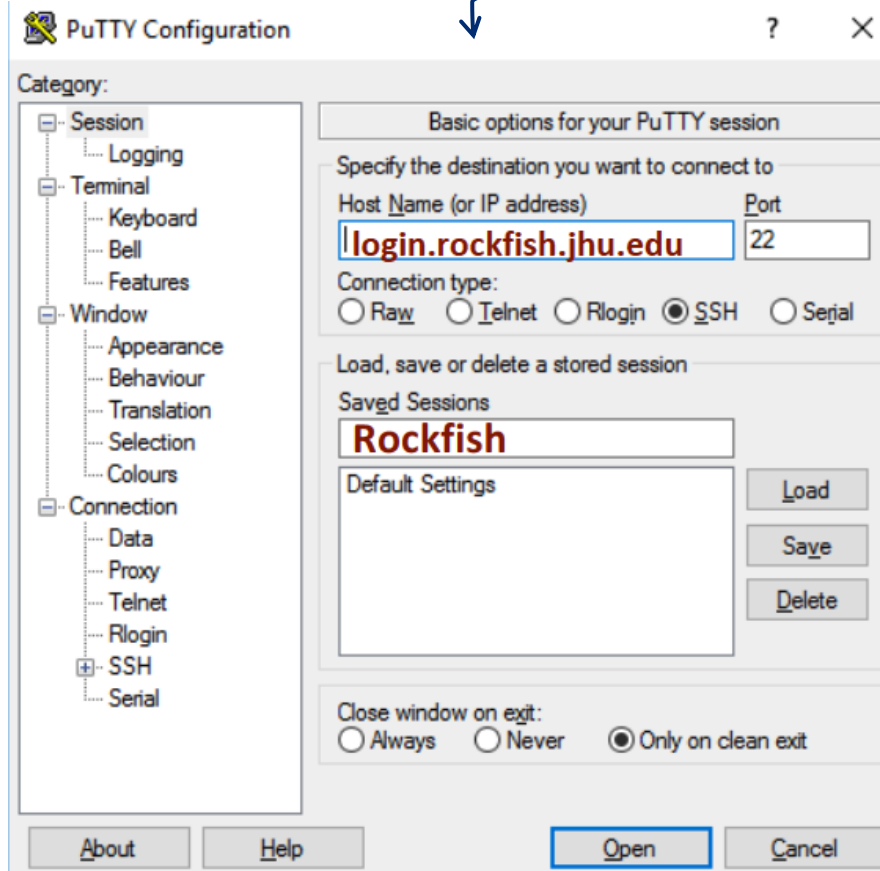
```
$ du -hc --max-depth=1
```

```
$ quotas.py
```

```
$ quotas.py [-group PI-userid]
```

File System	link/path	Quota	Comment
HOME	/home/userid	50GB	backed up
data	/data/PI-userid	1TB	long term files
scratch4	/scratch4/PI-userid	1TB	designed for small files
scratch16	/scratch16/PI-userid		designed for large files
tmp	/tmp	1TB	

# Connection



## Windows

- Putty
- MobaXterm
- SecureCRT

*gui*

- Open OnDemand (OOD)  
(requires VPN)

*web*

- portal.rockfish.jhu.edu

- PowerShell

*cli*

- Windows Sub-Linux System



# Connection and File Transfer with CLI

globus.org is for large data transfer



## Secure Shell

\$ ssh **(-l)** userid login.rockfish.jhu.edu

*small l*

\$ ssh userid@login.rockfish.jhu.edu

\$ ssh -X or ssh -Y

## Secure Copy for scripts or small data

\$ scp local\_file userid@login.rockfish.jhu.edu: **(~)**

\$ scp **(-r)** local\_dir userid@login.rockfish.jhu.edu: \$HOME

\$ scp userid@login.rockfish.jhu.edu:/path/to/file **(.)**

*environment variables*

- \$HOME
- ~
- /home/\$(id -un)
- /home/\$USER

.bashrc

.ssh

# File Transfer with Globus

use JHED to create an account  
on [coldfront.rockfish.jhu.edu](https://coldfront.rockfish.jhu.edu)



Log in to use Globus Web App

Use your existing organizational login

e.g., university, national lab, facility, project

Johns Hopkins

By selecting Continue, you agree to Globus [terms of service](#) and [privacy policy](#).

Continue



Globus uses CILogon to enable you to Log In from this organization. By clicking Continue, you agree to the [CILogon privacy policy](#) and you agree to share your username, email address, and affiliation with CILogon and Globus. You also agree for CILogon to issue a certificate that allows Globus to act on your behalf.

OR



Sign in with GitHub



Sign in with Google



Sign in with ORCID iD



Collection Search



FILE MANAGER



ACTIVITY



COLLECTIONS



GROUPS



CONSOLE



FLOWS



COMPUTE



SETTINGS



LOGOUT



HELP & SITMAP

Collection

Rockfish

## Rockfish AWS

Subscribed Mapped Collection (GCS) on RFDTN02

Owner: marcc@globusid.org

Domain: m-4c0e79.8b416.0ec8.data.globus.org

## Rockfish dtn01 user data

Subscribed Mapped Collection (GCS) on RFDTN01

Owner: a933ac1a-c486-473e-81f5-06baeedb08f2@clients.auth.globus.org

Domain: m-6971d6.5e9e0.bd7c.data.globus.org

Description: RFDTN01 endpoint for rockfish cluster data

## Rockfish dtn02 user data

Subscribed Mapped Collection (GCS) on RFDTN02

Owner: a7627ba9-e25e-44bd-9d48-d64aab883008@clients.auth.globus.org

Domain: m-0a4c68.8b416.0ec8.data.globus.org

Description: RFDTN02 endpoint for rockfish cluster data

## rbradle8 rockfish data

Guest Collection (GCS) on Rockfish dtn02 user data

Owner: rbradle8@johnshopkins.edu

Domain: g-0e4811.8b416.0ec8.data.globus.org

Description: Research data example for transfer to other sites.

## rf-genemaprdata

Guest Collection (GCS) on Rockfish dtn01 user data

Owner: kesohku1@johnshopkins.edu

Domain: g-f8e93f.5e9e0.bd7c.data.globus.org

Description: GeneMAP research data on Rockfish cluster



# File Transfer from google drive

Rockfish OnDemand

AppsFilesJobsClustersInteractive AppsMy Interactive Sessions

Advanced Research Computing at Hopkins

JOHNS HOPKINS UNIVERSITY

OnDemand provides an integrated, single access point for all of your HPC resources.

Pinned Apps

A featured subset of all available apps

Job Composer

System Installed App

Home Directory

System Installed App

TotalView

System Installed App

Mathematica

System Installed App

Jupyter Server

System Installed App

Abaqus/CAE

System Installed App

VSCoDe Server IDE/Editor

System Installed App

Matlab

System Installed App

Rockfish HPC Cluster Shell Access

System Installed App

Active Jobs

System Installed App

Desktop

System Installed App

ParaView

System Installed App

IGV Desktop

System Installed App

Lumerical

System Installed App

RStudio Server

System Installed App

JBrowse Desktop

System Installed App

Desktop (11764647)1 node | 1 core | Running

Host: >\_c002

Delete

Created at: 2024-04-23 19:40:48 EDT

Time Remaining: 29 hours and 58 minutes

Session ID: f3b55076-ff2c-42ac-9a85-8f17425c575f

noVNC Connection

Native Instructions

Compression

0 (low) to 9 (high)

Image Quality

0 (low) to 9 (high)

Launch Desktop

View Only (Share-able Link)

A screenshot of a desktop environment with a dark background. A taskbar is visible at the bottom, containing several application icons including a terminal, a file explorer, and a web browser. A large red arrow points downwards towards the taskbar area.

Advanced Research Computing at Hopkins

JOHNS HOPKINS UNIVERSITY

## navigation commands

\$ cd

\$ PWD

## account commands

\$ whoami

\$ id

- \$USER

## environment variables

```
[twu73@login02 ~]$
```

whoami

## hostname

```
[twu73@login02 Desktop]$
```

```
tsaiwei@DESKTOP-UCTGA49:~$ ssh -Y twu73@login.rockfish.jhu.edu
twu73@login.rockfish.jhu.edu's password:
Warning: No xauth data; using fake authentication data for X11 forwarding.
Last login: Fri Apr 19 11:03:59 2024 from 73.103.78.70
```

```
[STATUS] loading software modules
[STATUS] modules are Lmod (https://lmod.readthedocs.io/en/latest/)
[STATUS] software is Spack (https://spack.readthedocs.io/en/latest/)
[STATUS] the default modules ("module restore") use GCC 9 and OpenMPI 3.1
[STATUS] you can search available modules with: module spider <name>
[STATUS] you can list available modules with: module avail
[STATUS] loading a compiler, MPI, Python, or R will reveal new packages
[STATUS] and you can check your loaded modules with: module list --terse
[STATUS] to hide this message in the future: touch ~/.no_rf_banner
[STATUS] restoring your default module collection now
[STATUS] quota information listed below, you can view by invoking "quotas.py"
```

# MOTD

Home Usage for user twu73 as of Mon Apr 22 18:00:08 2024			
Used	Quota	Percent	Files
13.88 GB	0.00 KB	N/A	273,244

GPFS Usage for Group rfdadmin: as of Mon Apr 22 18:30:13 2024						
FS	Used	Quota	Used %	Files	Files Quota	Files %
data	50.51 TB	150.00 TB	33.00%	29,756,870	61,440,000	48.00%
scratch4	7.40 TB	10.00 TB	74.00%	7,035,797	20,480,000	34.00%
scratch16	2.43 TB	10.00 TB	24.00%	3,035,106	10,240,000	29.00%

GPFS Usage for Group jaimecomb: as of Mon Apr 22 18:30:13 2024						
FS	Used	Quota	Used %	Files	Files Quota	Files %
data	34.47 GB	1.00 TB	3.00%	28,613	409,600	6.00%
scratch4	695.76 GB	10.00 TB	6.00%	41,988	1,048,576	4.00%
scratch16	67.92 GB	5.00 TB	1.00%	71,203	5,120,000	1.00%

```
[twu73@login02 ~]$
```



# More commands

ls	shows files
mv	move or rename files/dir
touch	create a file
mkdir	create a folder
cp [-r]	copy files or folder
rm [-r]	remove files or folder
wget	download from a link
curl	download file
scp [-r]	secure copy files from/to another host

hostname	shows the name of the node
who	list users on the node
top -u \$USER	displays a user's processes
htop	
history	shows the history of your commands
cat	view files
more / less	
head / tail	
cal [month] [year]	
date	

# More commands

```
$ ls -l
```

```
drwxr-xr-x. 4 root root   68 Jun 13 20:25 tuned
-rw-r--r--. 1 root root 4017 Feb 24  2022 vimrc
```

## list files

`$ ls -l`      long format

`$ ls -a`      hidden files

`$ ls -h`      human-readable

`$ ls -d`      folders

`$ ls -r`      reverse order

`$ ls -t`      modification time

`$ ls -alrt`   combination

## find files

`$ find . -name file_name`

`$ find . -type d -name dir_name`

`$ find . -size +500M file_name`

`$ find . -iname file_name`

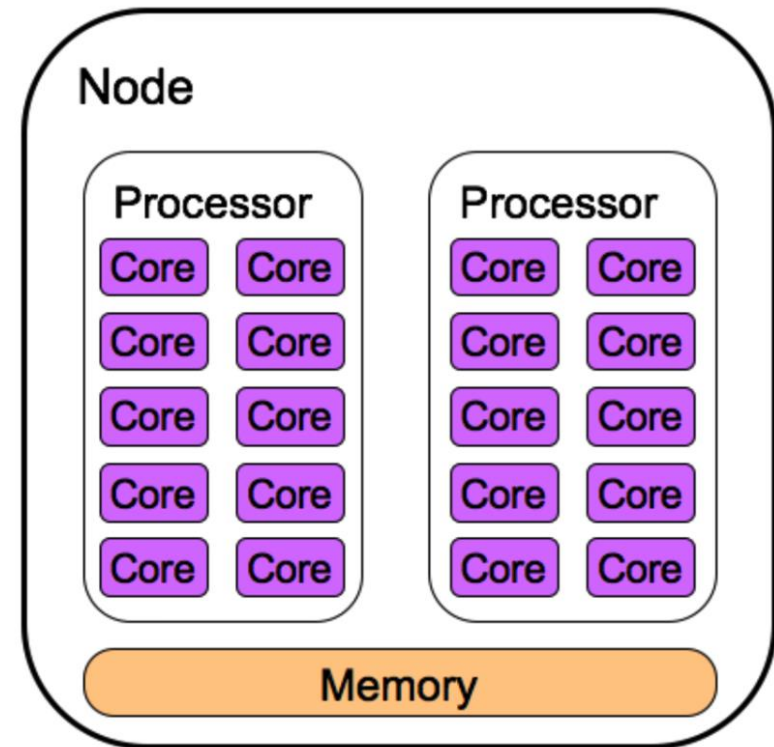


# Demo

- list folders
- storage quota with du
- file permissions
- find files
- text editor nano
- calculate folder space
- history
- tmux

# Terminology

- **NODE** on a cluster is a single computing unit.
- A node is comprised of many logical processors (=cpus=cores).
- A SLURM **JOB** can contains several **TASKS**. And each task can use one or more of cores.





The background of the slide is a photograph of a large, multi-story brick building with a series of prominent arches. The building is constructed of red brick with white mortar. To the left of the building, there are lush green trees and a clear blue sky with some light clouds. A person is visible standing on a balcony or walkway of the building, looking out. The overall scene is bright and sunny.

Questions?

THANK YOU.

[help@rockfish.jhu.edu](mailto:help@rockfish.jhu.edu)