# ilifu Online Training

Robin Hall

User Training Workshop – Basic Training

7 September 2021

# Topics

- Introduction to the ilifu research facility services
- Directory structure
- Software environment
- Accessing the ilifu services
- Using JupyterHub

# High Performance Computing

Combining power of distributed computers

- Collection of servers (computers)
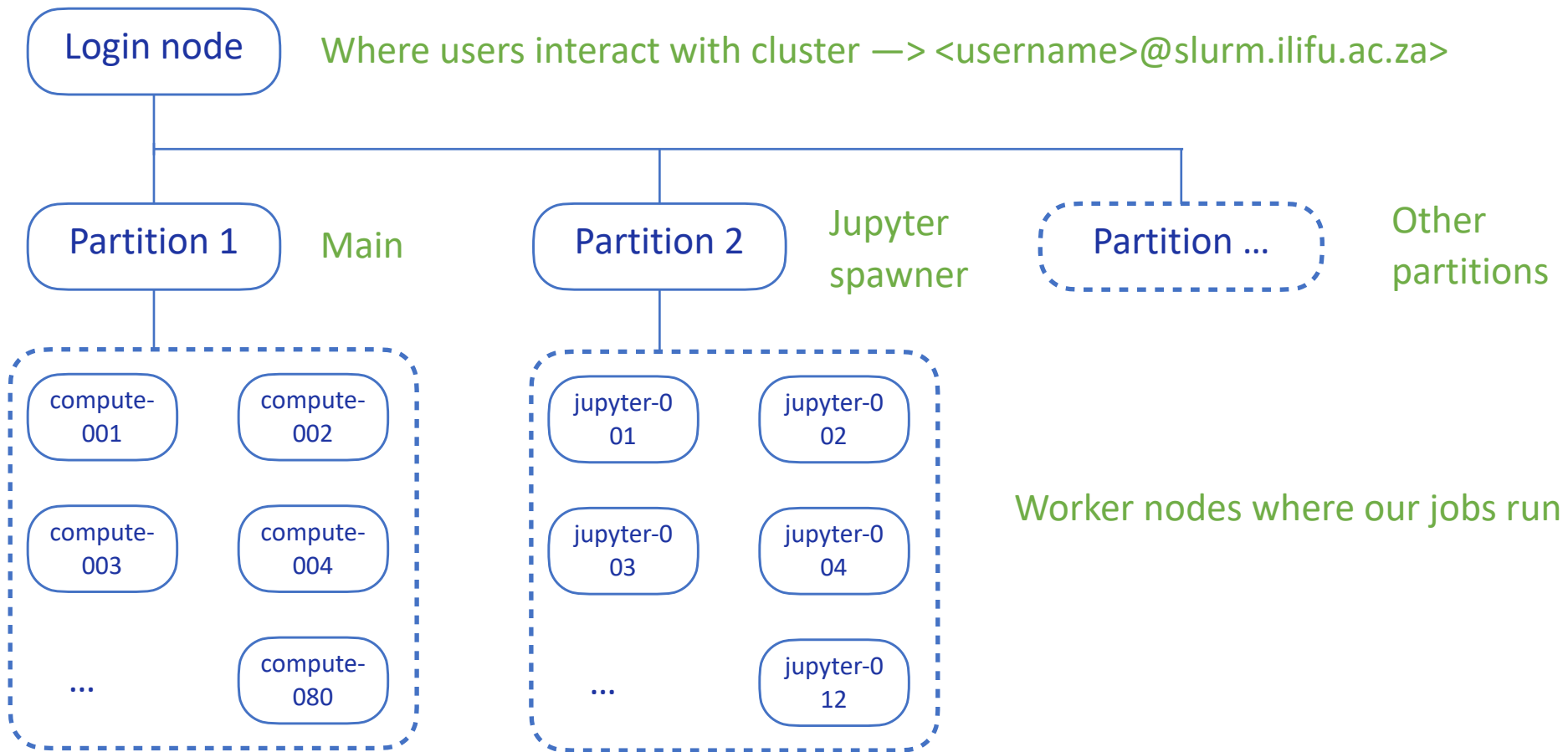- Connected by fast local network

Some terminology

- Servers also referred to as nodes
- Group of nodes is a cluster

# ilifu Research Facility

Cloud-based infrastructure for data-intensive research

- Support variety of different scientific projects and requirements
- Flexible compute environment
- Data management: storage, transfer

# ilifu Research Facility

# Software we use

- Job Scheduler to manage resources - Slurm

- JupyterHub service – development environ.

- Containerised software environment - Singularity

- Other services: data transfer, CARTA

# Computing environment - interface

## ssh – shell terminal



**ssh &lt;username&gt;@slurm.ilifu.ac.za**

## JupyterHub



https://jupyter.ilifu.ac.za

# Computing environment - ssh

Your SSH key

- Used in the SSH (Secure Shell) protocol
- Authentication method for gaining access to encrypted connection between systems
- Use connection to manage system remotely
- We need your SSH key so our system knows to let you in

# Computing environment - ssh

Generating SSH key
- If you don't already have one
- New computer

GitHub Docs on key generation:
[https://docs.github.com/en/github/authenticating-to-github/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent](https://docs.github.com/en/github/authenticating-to-github/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent)

# Directory Structure

- Common areas:
  - /users

    - only 40 TB shared among all users, for scripts and small files – **don't place data here, capping /users storage capacity can prevent access to the cluster for all users.**

  - /scratch3/users

    - storage space for processing data, temporary storage only, i.e. use this space during processing, and then clear all files immediately after processing. Remove unnecessary data and move data that you want to keep to project folder. /scratch3 has recently been increased from 100 TB to 400 TB

- Remaining storage is separated by group:
  - IDIA, CBIO, ILIFU

# Directory Structure

IDIA structure:

- ## /idia/users

  - user's private work directory, may store data products that are not ready to move to shared project space

- ## /idia/projects

  - project specific directories. These directories are for sharing data and resources within project groups. Raw data associated with a project will also be available from the project folder. Raw data should always be read-only.

- ## /idia/software

  - software containers and the IDIA Pipelines software is stored here

- ## /idia/software/containers

# Directory Structure

- Similar structure for /cbio and /ilifu groups,
  - /cbio/users
  - /cbio/projects
  - /cbio/soft

  - /ilifu/users
  - /ilifu/software
  - Exception for ilifu projects:
    - /ilifu/astro/projects
    - /ilifu/bio/projects

# Directory Structure - workflow

/users — Scripts and small files only

/n/projects/
/n/users/
/n/data/*

Data

read

Results

selective write

/scratch
/scratch3

compute/process

Results + intermediate files
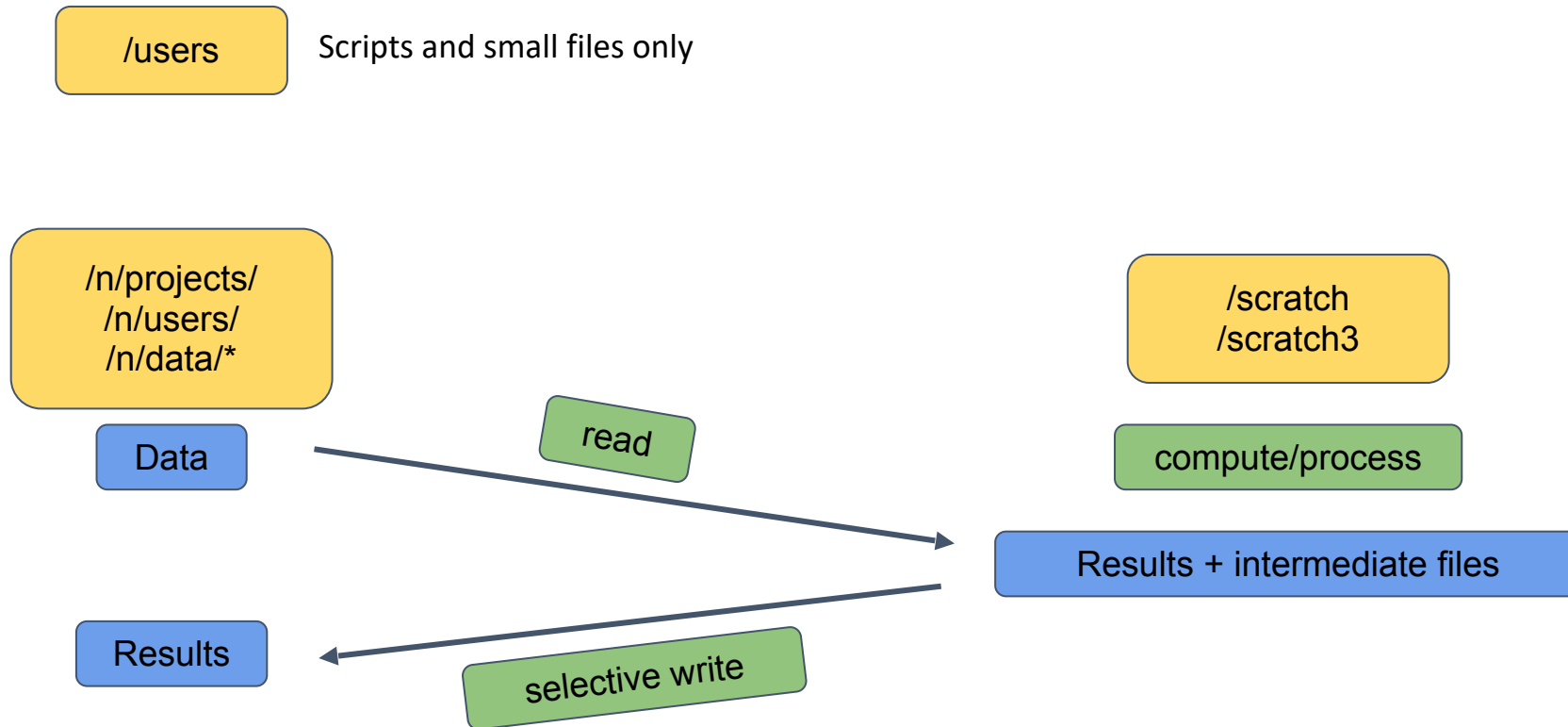
*/n/data generally read-only

# Software environment - Singularity containers

- Encapsulated software environments

- A software stack that contains everything required to run an application/workflow, including files, environmental variables, libraries and dependencies

- Containers accessible across platforms and services, allowing sharing of application environments

# Software environment - Singularity containers
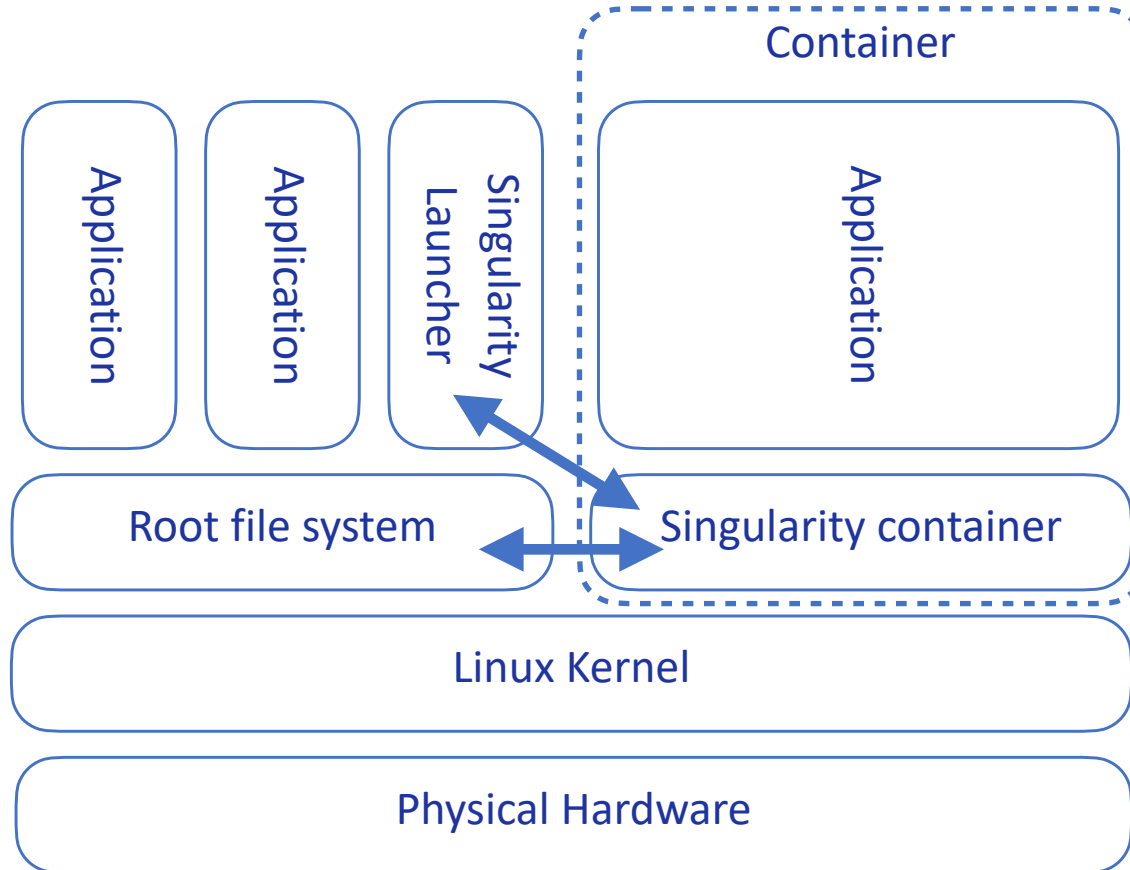
# Software environment - Singularity containers

Supported Containers:

- CASA
- Astronomy container (ASTRO-PY3)
- KERN suite
- Python 2.7, Python 3.6, R
- GPU container
- Project containers:
  - MeerLICHT,
  - LADUMA
  - HI intensity map

- /idia/software/containers
- /ilifu/software/containers

# Software environment - Singularity containers

**Open container as an interactive shell:**

```
singularity shell /path/to/container
```

```
Example:
$ singularity shell /idia/software/containers/ASTRO-PY3.simg
```

**Run a script/workflow using a container environment:**

```
singularity exec /path/to/container <software> <script/input_params>
```

```
$ singularity exec /idia/software/containers/casa-stable.img casa -c myscript.py
```

# Software environment - modules

**module avail**

```
~$ module avail

------------------------------- /software/modules/common -------------------------------------------
  LAPACK/3.9.0              anaconda3/login.old        drmaa/1.1.1          openBLAS/0.3.9        perlbrew/perlbrew        ruby/
2.6.6
  R/RStudio1.2.5042-R4.0.0  anaconda3/login            homebrew/2.4.13      openmpi/2.1.6         python/2.7.18
  R/RStudio1.2.5042-R4.0.4  anaconda3/2020.07     (D)  java/jre-1.8.0_261   openmpi/3.1.6         python/3.7.7
  R/3.6.3                   cuda/10.0.130_410.48       java/openjdk-14.0.1 (D)  openmpi/4.0.3      python/3.8.2
  R/4.0.0                   cuda/10.1.243_418.87.00    julia/1.5.3          openmpi/4.0.5         python/3.8.3
  R/4.0.2                   cuda/10.2.89_440.33.01     maven/3.6.3          openmpi/4.1.0    (D)  python/3.8.6
  R/4.0.3              (D)  cuda/11.0.2_450.51.05 (D)  mono/6.8.0.123       perl/5.33.0           python/3.9.0        (D)


------------------------------- /software/modules/astro --------------------------------------------
  casa/5.7.0    casa/5.7.2-4    casa/6.1.0-118-monolithic    casa/6.1.2.7-pipeline    casa/6.1.2.7-modular (D)    pybdsf/1.9.2


------------------------------- /software/modules/bio ----------------------------------------------
  bcbio/1.2.3       canvas/1.40.0.1613    genomestrip/2.00.1958    plink/2.00a2.3    prsice-2/2.3.1d    treePL/homebrew    vep/singularity
  bcftools/1.10.2   gemini/gemini         htslib/1.10.2            popgen/0.1        samtools/1.10      vcftools/0.1.16    vep/101.0        (D)


------------------------------- /usr/share/lmod/lmod/modulefiles ------------------------------------
  Core/lmod/6.6    Core/settarg/6.6

  Where:
  D:  Default Module
```

# Software environment - modules

- module avail

- module help <module>

```
~$ module help python

-------------------------- Module Specific Help for "python/3.9.0" --------------------------------------
This module configures Python 3.9.0 for use
```

- module load <module>
- module purge
- module --help

# JupyterHub

- https://jupyter.ilifu.ac.za

# JupyterHub

- Session size

## Server Options

**Nodes Free**

as at Wed Aug 25 14:18:01 SAST 2021

21 Minimum
9 Small
3 Medium
1 Large
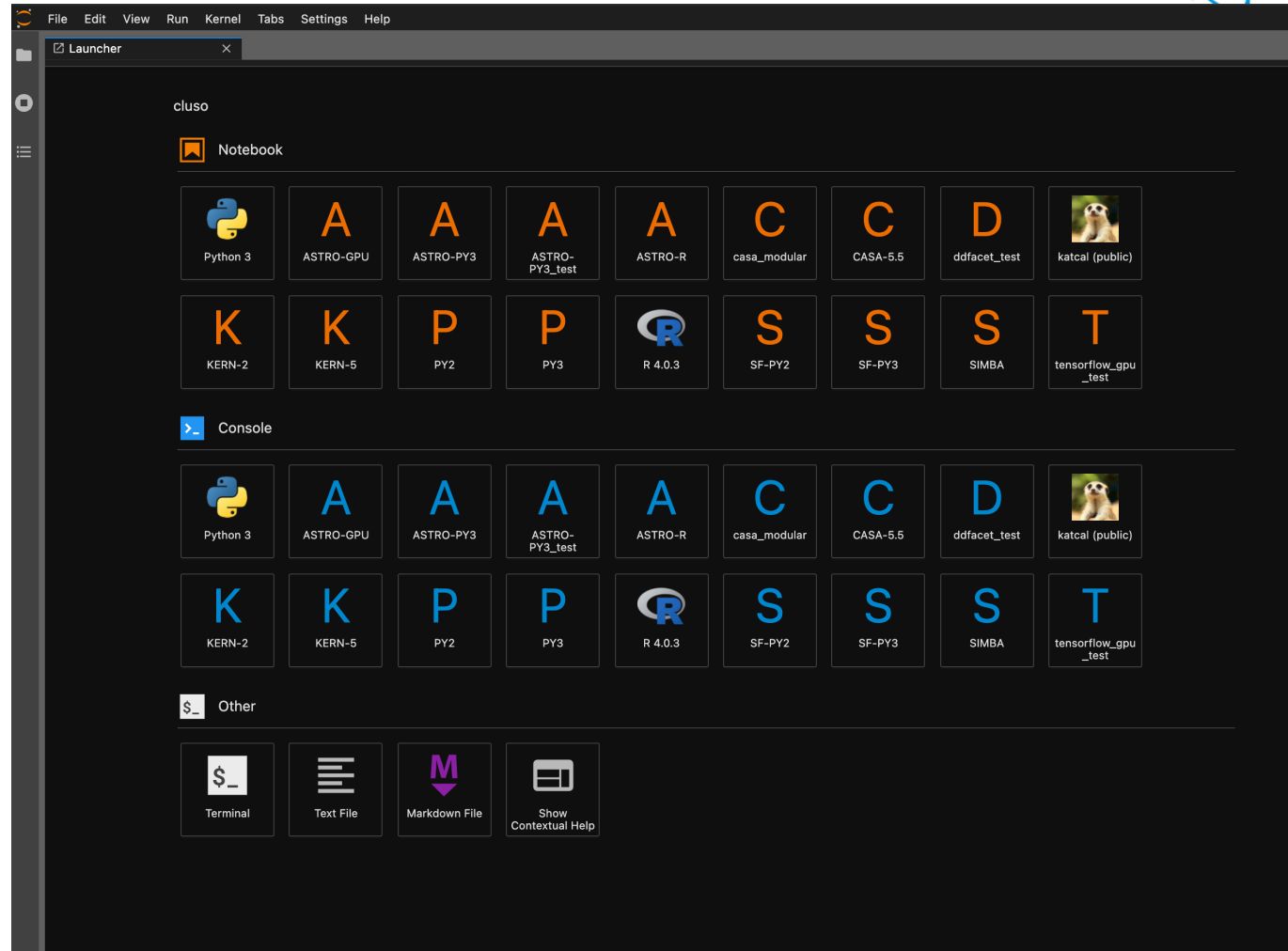0 Half-Max
0 Max
4 GPU

**Select a job profile:**

Minimum Node - 1 core, 7 GB, 18 hours idle timeout, max 5 days lifespan ⌄

**Start**

# JupyterHub

- Choose kernel in launcher

# Getting help

- Support contact
  support@ilifu.ac.za

- User documentation
  http://docs.ilifu.ac.za/#/

- Ilifu System Status
  https://status.ilifu.ac.za/

- Training videos
  http://www.ilifu.ac.za/il/accessing-facilities/training

# Demo time