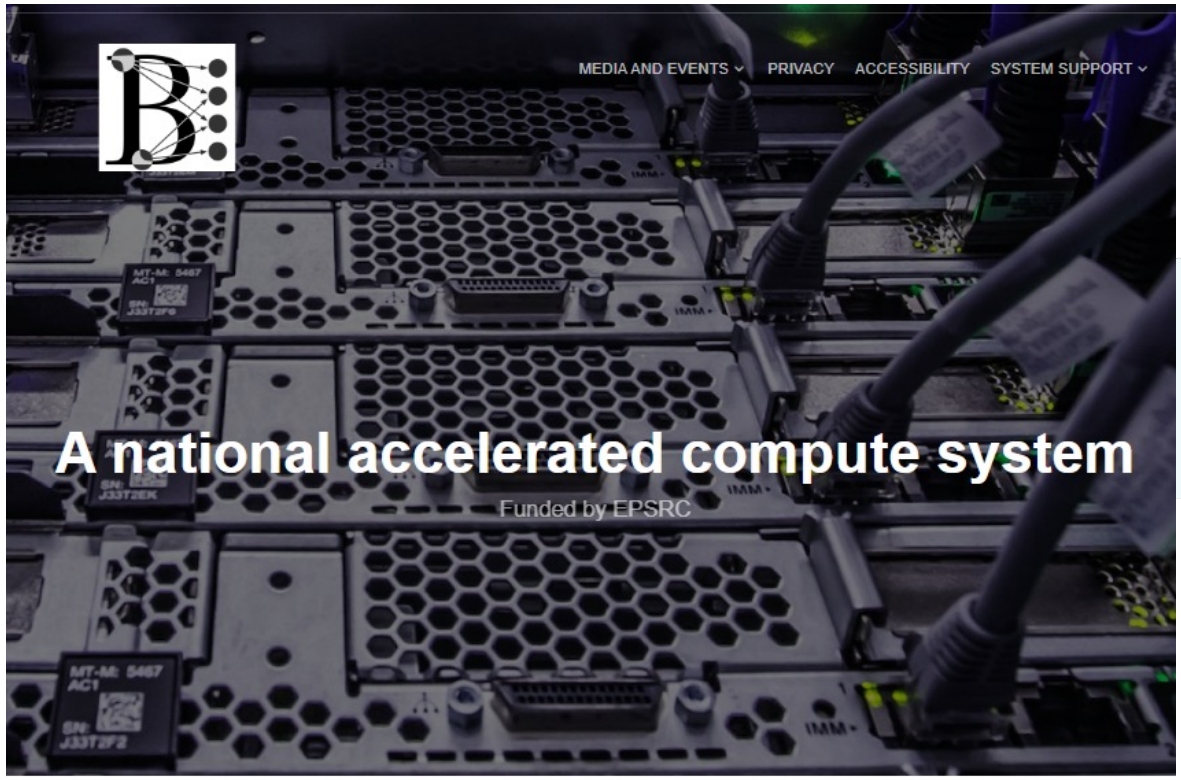
The Rosalind Franklin Institute

Logging In and Module Loading Session

Baskerville HPC



Baskerville launched to users in July 2021

Read details of our Baskerville launch event.

208
GPUs

52
Compute Nodes

5400
Storage (TB)

- 52 Compute Nodes, each:
  - 2x CPUs so total of 144 logical cores 512GB RAM
  - 4x NVIDIA A100s GPUs
    - 46 Nodes have the 40GB VRAM variety
    - 6 Nodes have the 80GB VRAM variety
  - Local Storage:
    - 1TB available as `/scratch-local` used from OS
  - Global storage:
    - 418x 16TB HDD available as `/bask` general storage for home directory and project bulk data storage
    - 48x 7.68TB SSD `/scratch-global` transient storage on SSD enclosures available on all compute systems

More details click here

A condition of access to Baskerville is that the service is acknowledged in the research output generated through the use of Baskerville. Here is a suggested form for this acknowledgment.

Logging In and Module Loading Session

Request access to Baskerville

If you are new Baskerville user, and want access, go to https://docs.baskerville.ac.uk/request-access/#consortium-member-access to check how to request access based on the institution you belong.

In Particular,

- For Diamond Light Source, contact Tim Poon (timothy.poon@diamond.ac.uk) and upon review of your request, will grant you access.

- For Rosalind Franklin Institute, please email helpdesk@rfi.ac.uk with the following details
  - Your email address (if you have an RFI email please email us with this) and optionally your FedID
  - The Franklin themes you are associated with.
    Aka which Baskerville projects of our institution you should join.
    - 5 Baskerville projects for each RFI theme, (storage quota 50TB pre theme).
      1 General Access ‘core’ Baskerville project, shared amongst all (storage quota 20TB)

The Rosalind Franklin Institute

Logging In and Module Loading Session

How to login to Baskerville for the first time

- Follow the instructions as explained here https://docs.baskerville.ac.uk/logging-on/ **carefully**.
  In short:

1. **Important:** Install an authenticator app on your smartphone like Google Authenticator, Microsoft Authenticator, FreeOTP, etc.
   And have it ready to scan a QR code

2. Login to Baskerville authentication portal

3. Select the “New User / Forgot Password?” link

4. Enter ther email address you offered when requesting access to Baskerville

5. You will receive an email to (re)set your (strong) password.
   Once you click to (re)set your password do not close your browser or navigate elsewhere as there will be a step to scan a QR code for 2FA

6. (Re)set your password and **then Scan the QR code with the smartphone app you installed.**

7. In authentication portal set your first name and last name

8. There is a way to setup a ssh key (it is covered in the last slides)

The Rosalind Franklin Institute

Logging In and Module Loading Session

Introduction to Baskerville

In Baskerville there are 2 types of Nodes-Machines

- Login Nodes and
- Compute Nodes

Login Nodes

Login Nodes are accessible by connecting to Baskerville via the terminal

To login execute (replace <baskerville-username> with your Baskerville username):

```
ssh <baskerville-username>@login.baskerville.ac.uk
```

To login with X11 forwarding enabled (if you are sshing from Linux e.g. a VM and planning to run graphical apps) execute:

```
ssh -X <baskerville-username>@login.baskerville.ac.uk
```

- Do not have any software on them
- Baskerville modules cannot be loaded to them, only for display purposes
- Only to submit:
  - Non-interactive terminal jobs (are explained in later slides)
  - Interactive terminal jobs (are explained in later slides)

  These jobs are being executed on the Compute Nodes

The Rosalind Franklin Institute

Logging In and Module Loading Session
Introduction to Baskerville

It is recommended to connect (ssh) to Baskerville Login nodes using:

- A remote machine or VM (e.g. Guacamole VM)
- Or upon connecting use a terminal multiplexer like:
  - tmux (https://github.com/tmux/tmux/wiki/Getting-Started)
  - screen (https://linux.die.net/man/1/screen)

  This is way it safer to work in case of Internet disconnection

Furthermore, important to remember that Baskerville blocks non UK IPs.

Compute Nodes

Where terminal jobs run and where Interactive Apps from the Baskerville Portal run

The Rosalind Franklin Institute

Logging In and Module Loading Session

Useful terminal commands and Storage

To learn about the Baskerville projects and the Quality of Service (QoS) label that you have access execute:

```
my_baskerville
```

To learn about the current GPU availability execute:

```
baskstatus
```

Home Directory Storage

To learn what is the available quota in your Baskerville home directory (`/bask/homes/<first-letter-of-your-Baskerville-username>/<Baskerville-username>`)

```
my_quota
```

The total storage quota on your home directory is 20GB and it cannot be increased

Commands cheatsheet here https://docs.baskerville.ac.uk/cheatsheet/

The Rosalind Franklin Institute

Logging In and Module Loading Session

Useful terminal commands and Storage

Project Storage

- `/bask/projects/<first-letter-of-project-name>/<project-name>`

To check the available project storage run:

```
df -h /bask/projects/<first-letter-of-project-name>/<project-name>
```

or go to https://admin.baskerville.ac.uk/project/<project-name>

Example:

```
df -h /bask/projects/t/test-project
```

To learn how much storage is being used by one of your folders, run:

```
du -sh /bask/projects/<first-letter-of-project-name>/<project-name>
/<your-folder>/<folder-in-question>
```

Example:

```
du -sh /bask/projects/t/test-project/myBaskID/Job-1
```

The Rosalind Franklin Institute

Logging In and Module Loading Session

Storage

Temporary (Scratch) Storage

- `/tmp`: local scratch storage
  When inside a Slurm job the `/tmp` directory is constructed in a per-job privately mounted namespace and therefore not visible to the users of other jobs running on the node.

- `/scratch-global`: shared (network) scratch storage, available across all the compute nodes.

It is highly recommended for jobs with significant and frequent I/O (Input/Output) to drive activity, local scratch space (`/tmp`) to be utilised. To do so, your job submission scripts should:

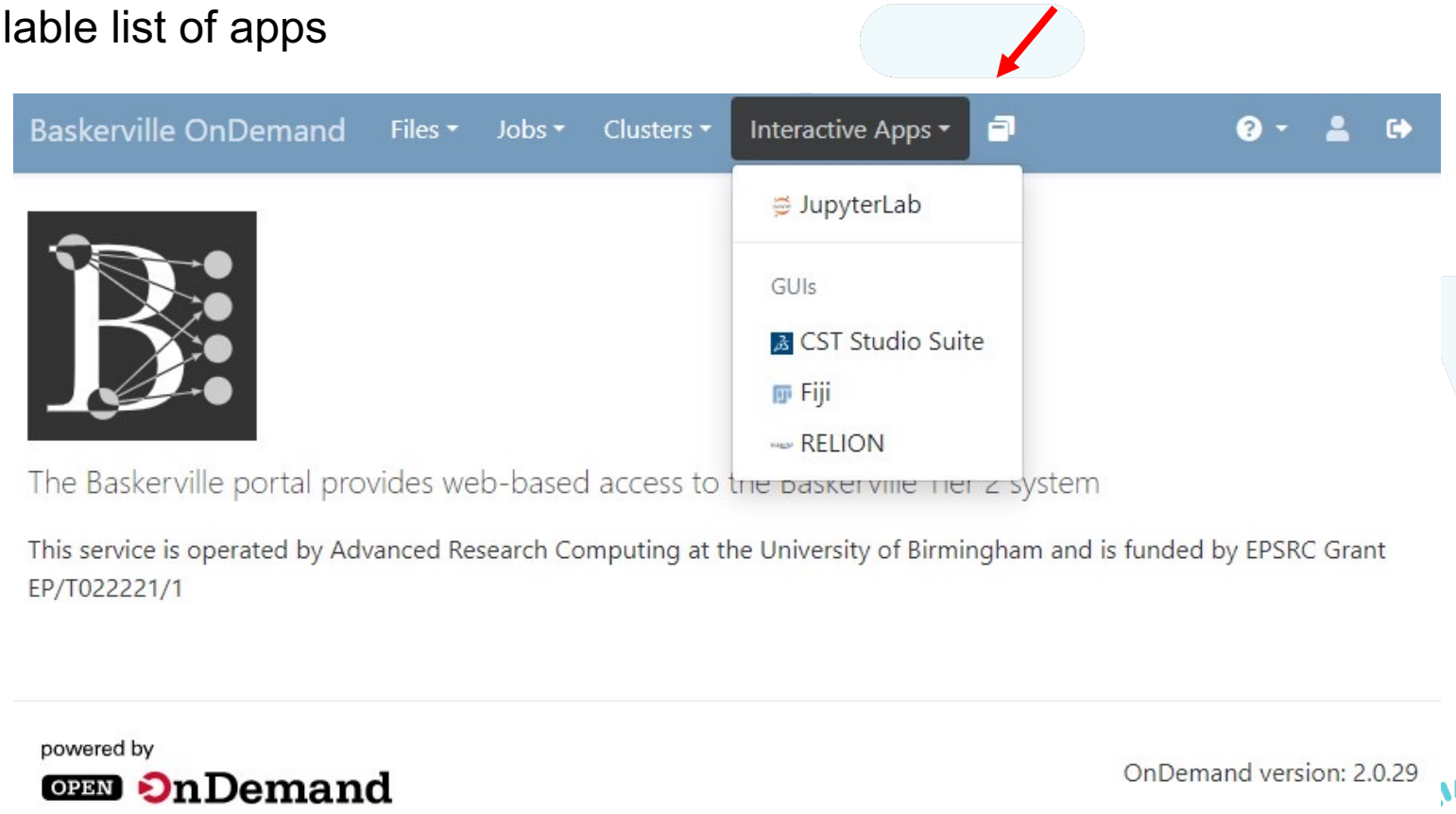
- Move data if necessary to `/tmp` (from Project Storage)
- Run main submission script commands that make use of `/tmp` for I/O to drive
- Move data to Project Storage

The Rosalind Franklin Institute

Logging In and Module Loading Session

Running Interactive Apps

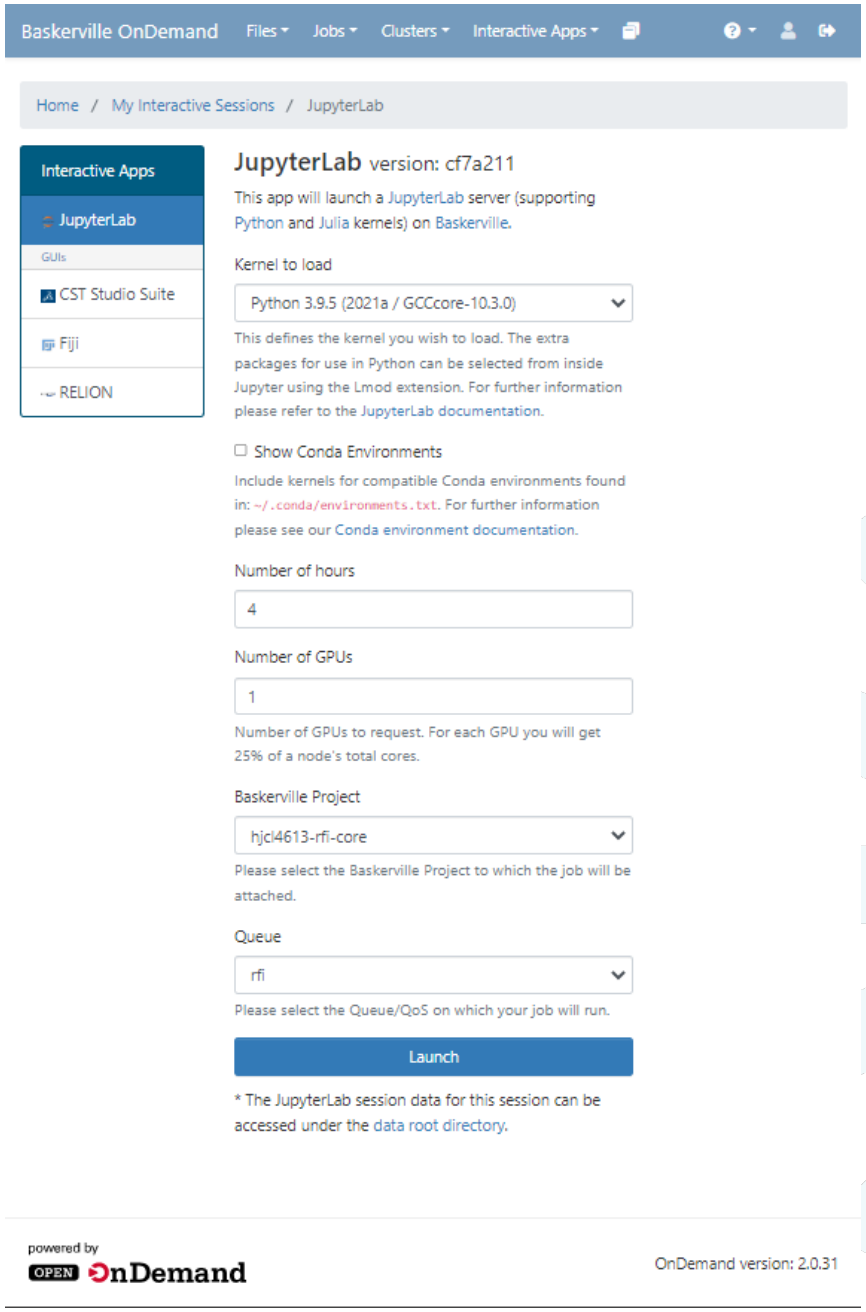
After connecting to [Baskerville Portal] go to the Interactive Apps selection to start a new Interactive App from the available list of apps



The Rosalind Franklin Institute

Logging In and Module Loading Session

Running JupyterLab (Python or Julia)



If you select to start a JupyterLab you will see what is on the left.

You can select the time you want to use it, number of GPUs and under what project / QoS

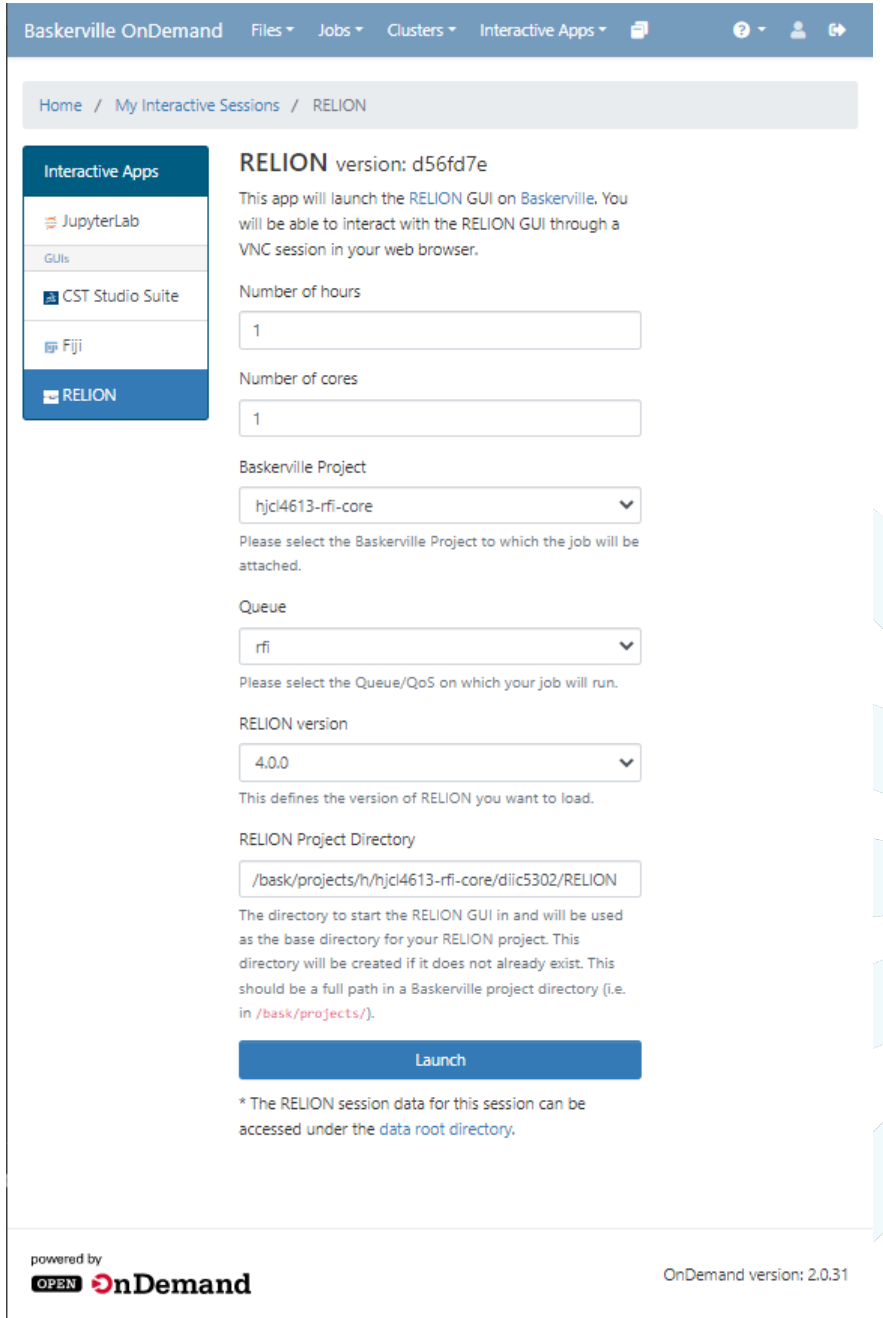
You can read more about how to use JupyterLab after its Interactive App starts here.

You can read how to use conda environments in the JupyterLab here

The Rosalind Franklin Institute

Logging In and Module Loading Session

Running RELION



If you select to start RELION you will see the following.

You can select the time you want to use it, number of GPUs (cores) and under what project / QoS

You can also set a project directory to avoid saving on the home directory (in case you did not perform the tasked mentioned in an earlier slide)

More on the 14:45 - 15:45 Breakout Session

The Rosalind Franklin Institute

Logging In and Module Loading Session
Baskerville modules

In a login node, to display the available modules (live applications), execute:

```
module avail
```

Baskerville-modules.txt

To load a module (in a Compute node) either

- include the following command to a submission script
- run it after you successfully login to a Compute node (with `srun`)
  More about but in the 10:45 - 12:15 Session

```
module load <module-name>/<module-version>
```

A number of Baskerville test modules is also available and to be able to load them, execute:

```
module purge
module load bask-apps/test/test
```

Baskerville-test-application-modules.txt

These modules are Baskerville test applications and these may change or be removed without warning.

To unload any loaded modules and be able to load from start the Baskerville (live application) modules execute:

```
module purge
module load baskerville
```

For dependencies not covered by the existing modules, new ones can be requested or they can be fullfiled via self-installed software (more on the 13:15 - 14:30 session)

The Rosalind Franklin Institute

Logging In and Module Loading Session

How to setup an SSH key

1. In your local machine generate a private-public key pair
   In Linux or MacOs machines this can be done in the terminal by executing:
   `ssh-keygen -t rsa -b 4096`

2. You can set any name for the key.
   By default in Linux machines it is: `id_rsa` and stored in `/home/<username>/.ssh` directory
   If you already have a key with this name, give it a new name. In the instructions below it will be referred as <key> Passphrase is not needed

3. In your local machine, in the `/home/<username>/.ssh` create a file named `config` and write the following (append at the end of the file if this file already existed):

```
Host login.baskerville.ac.uk
  User <baskerville-username>
  Hostname login.baskerville.ac.uk
  PreferredAuthentications publickey,keyboard-interactive
  IdentityFile ~/.ssh/<key>
```

The Rosalind Franklin Institute

Logging In and Module Loading Session

How to setup an SSH key

4. Execute: `chmod 600 /home/<username>/.ssh/config`
   This means that they owner of this file has read-write permissions and the group or other do not have any permissions

5. Login to the Baskerville authentication portal

6. Open/Read the <key>.pub (`/home/<username>/.ssh/<key>.pub`) file,
   copy its contents and paste them to the ssh field in the Baskerville authentication portal, and apply/save.

7. Try again to ssh to Baskerville from your local machine and it should ask only for your OTP

The Rosalind Franklin Institute