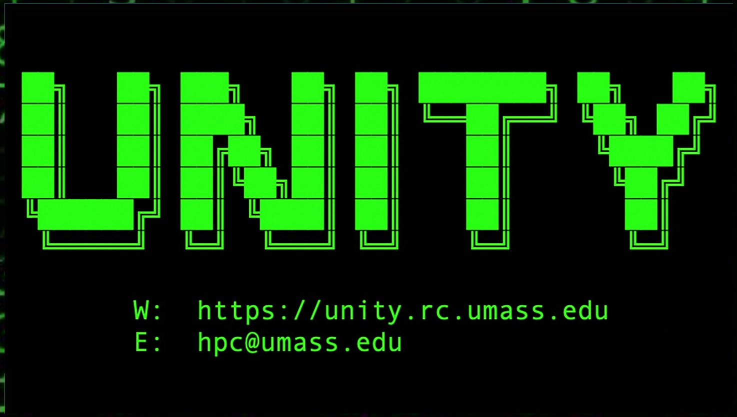
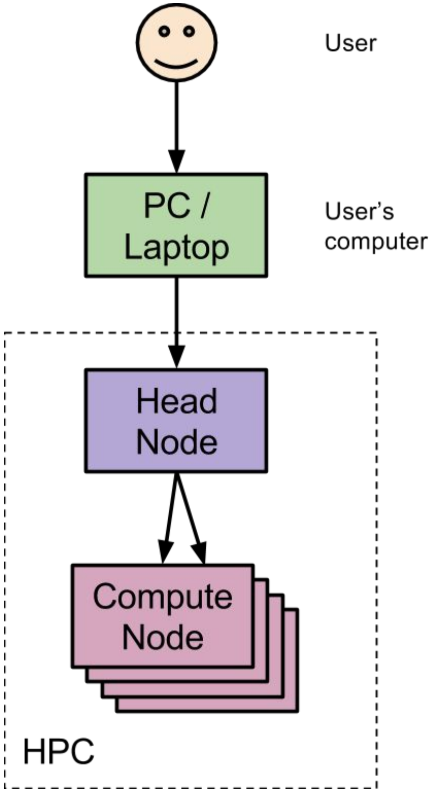
Introduction to Bioinformatics on Unity I

Cecile Cres & Anna Schrecengost
February 26 2024



```
W:  https://unity.rc.umass.edu
E:  hpc@umass.edu
```

High-Performance Computing and Biology



Outline

- Accessing Unity
- Data storage
- Transferring data to and from Unity
- Bioinformatics tools available on Unity
- Installing softwares via conda, apptainer and R
- Biological datasets
- Submitting jobs on Unity

Different ways to access Unity

Signing up for an account

Full tutorial/demo available on the Unity Onboarding video (Spring 2024)

1. Go to unity.rc.umass.edu
2. Log in/request account → Log in with institution SSO → My PIs → input PI NetID
    - Need to be a PI or part of PI group
    - If PI: Account settings → request PI account

Connecting on CLI via SSH

- SSH = secure shell
- Key-based authentication
- Can generate keys on Unity portal or on command line
- Unity username is based on email:

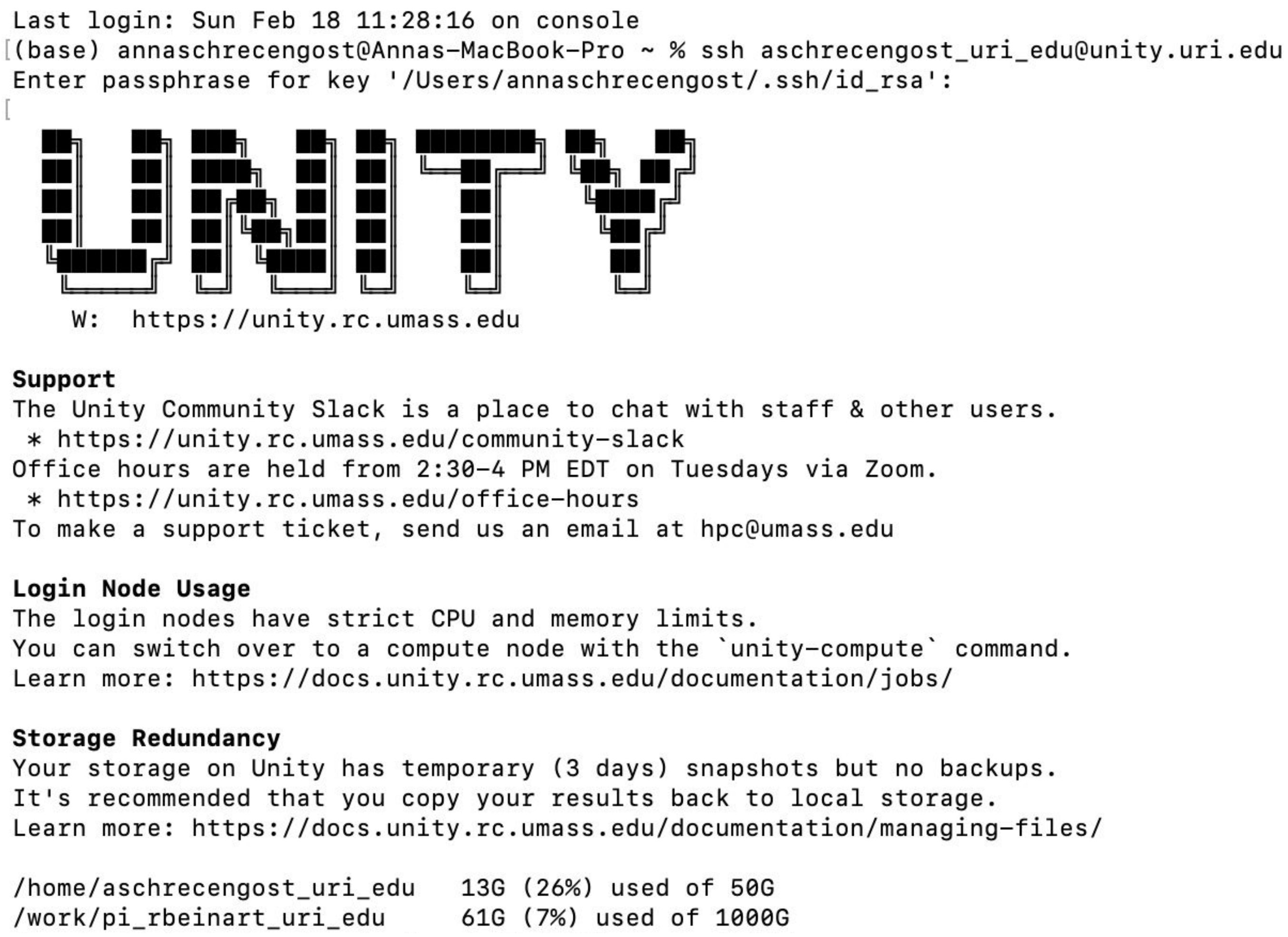
```
ssh <username>_<organization>_edu@unity.<organization>.edu
```

```
ssh <username>_uri_edu@unity.uri.edu
```

Connecting via SSH

- Create config file on your computer which contains information for connecting to unity
- Full details can be found here

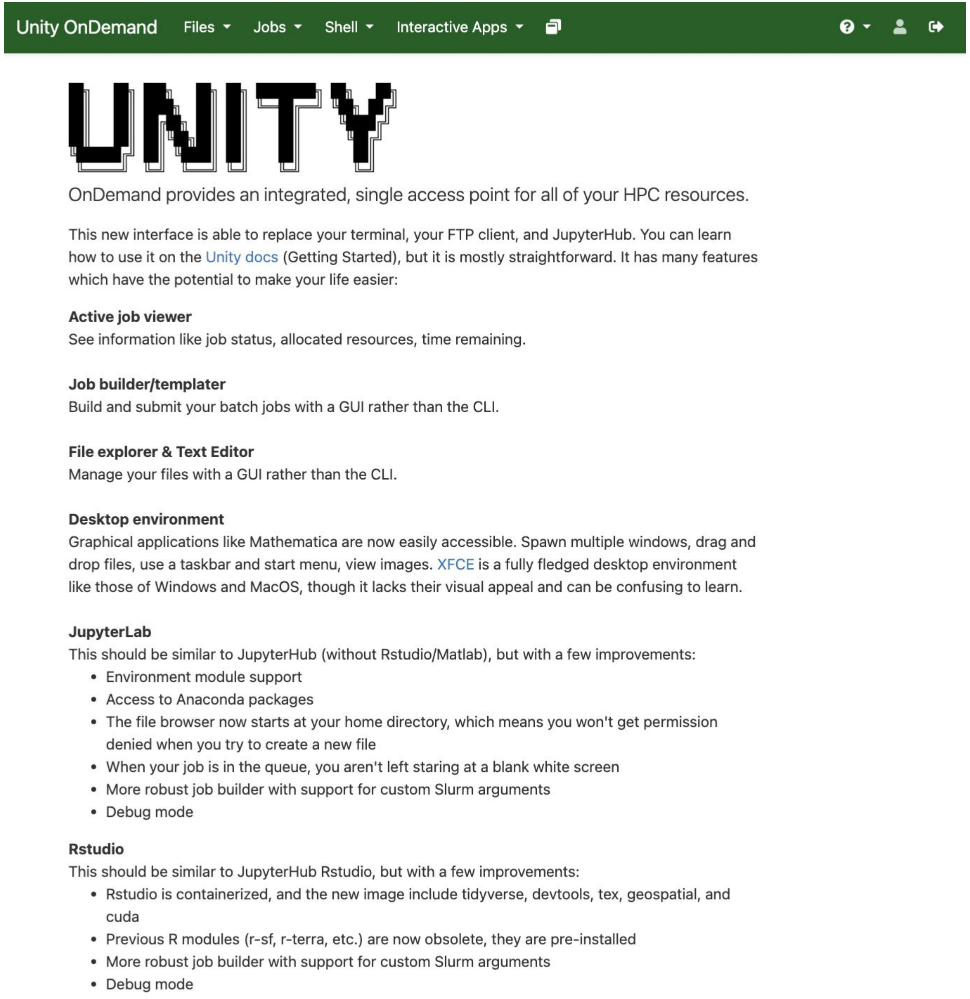
```
ssh unity
```



```
Last login: Sun Feb 18 11:28:16 on console
(base) annaschrecengost@Annas-MacBook-Pro ~ % ssh aschrecengost_uri_edu@unity.uri.edu
Enter passphrase for key '/Users/annaschrecengost/.ssh/id_rsa':

    W:  https://unity.rc.umass.edu

Support
The Unity Community Slack is a place to chat with staff & other users.
 * https://unity.rc.umass.edu/community-slack
Office hours are held from 2:30-4 PM EDT on Tuesdays via Zoom.
 * https://unity.rc.umass.edu/office-hours
To make a support ticket, send us an email at hpc@umass.edu

Login Node Usage
The login nodes have strict CPU and memory limits.
You can switch over to a compute node with the `unity-compute` command.
Learn more: https://docs.unity.rc.umass.edu/documentation/jobs/

Storage Redundancy
Your storage on Unity has temporary (3 days) snapshots but no backups.
It's recommended that you copy your results back to local storage.
Learn more: https://docs.unity.rc.umass.edu/documentation/managing-files/

/home/aschrecengost_uri_edu   13G (26%) used of 50G
/work/pi_rbeinart_uri_edu     61G (7%) used of 1000G
```

Connecting via OnDemand

- Simplest way to connect – no SSH keys
- Sign in on browser using organization SSO
- Can access the Unity login shell this way
- Can also access the filesystem and interactive Apps
    - RStudio, JupyterLab, MATLAB, Mathematica, Unity Desktop



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

This new interface is able to replace your terminal, your FTP client, and JupyterHub. You can learn how to use it on the Unity docs (Getting Started), but it is mostly straightforward. It has many features which have the potential to make your life easier:

Active job viewer
See information like job status, allocated resources, time remaining.

Job builder/templater
Build and submit your batch jobs with a GUI rather than the CLI.

File explorer & Text Editor
Manage your files with a GUI rather than the CLI.

Desktop environment
Graphical applications like Mathematica are now easily accessible. Spawn multiple windows, drag and drop files, use a taskbar and start menu, view images. XFCE is a fully fledged desktop environment like those of Windows and MacOS, though it lacks their visual appeal and can be confusing to learn.

JupyterLab
This should be similar to JupyterHub (without Rstudio/Matlab), but with a few improvements:

- Environment module support
- Access to Anaconda packages
- The file browser now starts at your home directory, which means you won't get permission denied when you try to create a new file
- When your job is in the queue, you aren't left staring at a blank white screen
- More robust job builder with support for custom Slurm arguments
- Debug mode

Rstudio
This should be similar to JupyterHub Rstudio, but with a few improvements:

- Rstudio is containerized, and the new image include tidyverse, devtools, tex, geospatial, and cuda
- Previous R modules (r-sf, r-terra, etc.) are now obsolete, they are pre-installed
- More robust job builder with support for custom Slurm arguments
- Debug mode

Data storage

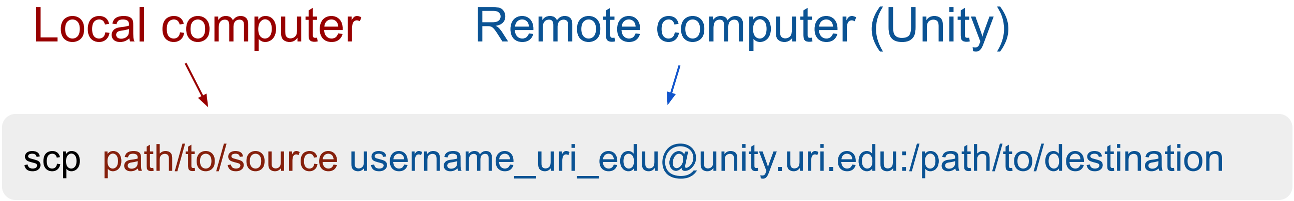
- /home directory → 50 GB quota
- /work/pi_ → PI's work directory → 1TB quota
- /project → available to PIs upon request → quota depends on storage needed
- /scratch → scratch space → no quota

More info on Unity docs

Transferring data to & from Unity

Transferring data using CLI - scp

- These CLI tools are best for small/quick data transfers
- Do not need to install scp
- Copy files or a whole directory to/from Unity
- Add “-r” flag after scp to copy contents of entire directory



Local computer

Remote computer (Unity)

```
scp  path/to/source username_uri_edu@unity.uri.edu:/path/to/destination
```

```
scp username_uri_edu@unity.uri.edu:/path/to/source path/to/destination
```

Transferring data using CLI - scp

- If have created config file can simplify:

Local computer Remote computer (Unity)

```
scp  path/to/source unity:/path/to/destination
```

```
scp unity:/path/to/source path/to/destination
```

Transferring data using CLI - rsync

- Mac and Linux users can install
- Add “-r” or “-a” flag after rsync to copy contents of entire directory
- Resume interrupted downloads using flag “--partial”
- Compress files during transfer with “-z”

Local computer

Remote computer (Unity)

```
rsync  path/to/source unity:/path/to/destination
```

```
rsync unity:/path/to/source path/to/destination
```

Transferring data using CLI - rclone

RCLONE

- Command-line utility for managing files in cloud storage
- Can use to transfer large files from Unity to a sequencing database such as ncbi SRA, for example
- Unity provides rclone - can use on CLI or also as an interactive app on Unity OnDemand
- Rclone documentation is [here]

Transferring data using GUI - Cyberduck

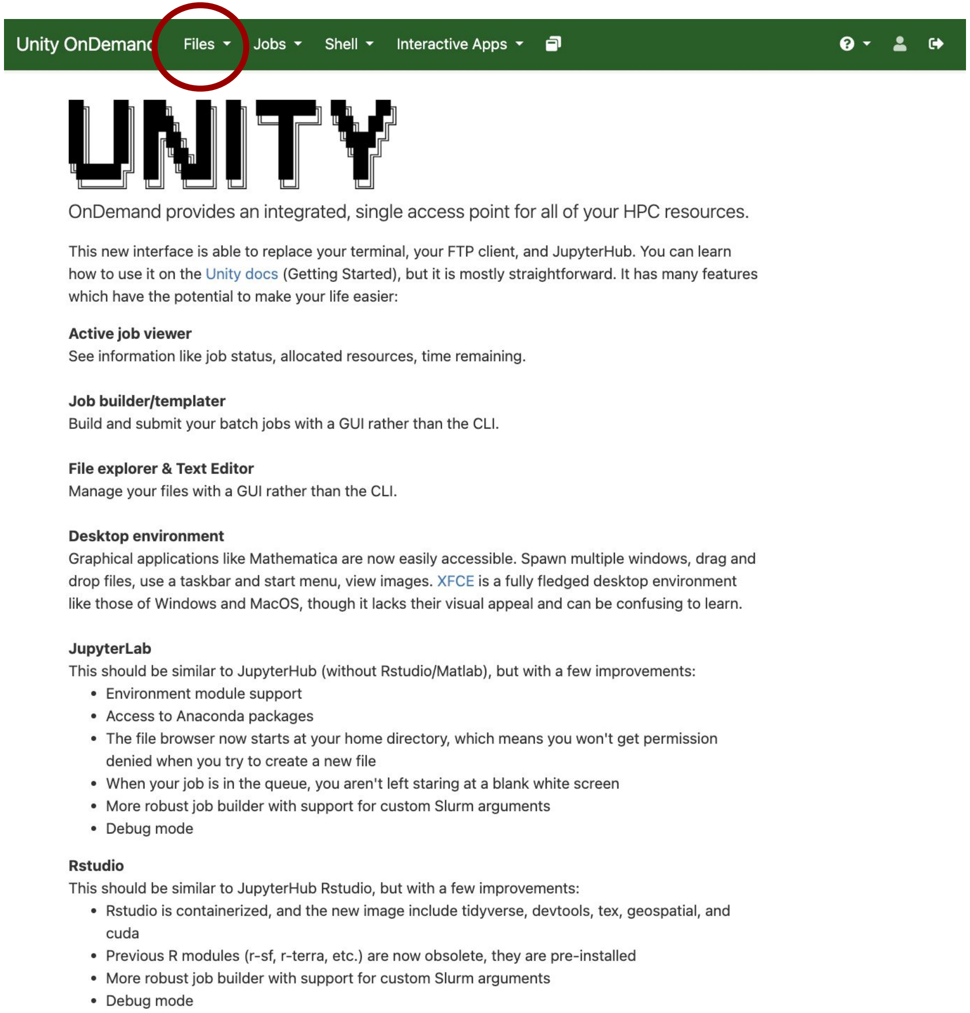
- Transfer small amounts of data to/from cluster via user-friendly interface
- Can also be used for most types of file transfer (e.g. local storage, cloud storage, etc.)
- Set-up tutorial is available here
- FileZilla is a similar client - tutorial is available here

Transferring data using Globus

- Best option for transferring large amounts of data between clusters
- Can also transfer between other Globus endpoints, or Unity and your local machine (e.g. backup to external hard drive)
  - To transfer to/from local machine, must install Globus Connect Personal
- To get started, go to app.globus.org
- Tutorial available here

Transferring data using Unity onDemand

- Simplest way to transfer files - does not require use of CLI or ssh configuration
- Cannot transfer large files - use for files less than 5GB



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

This new interface is able to replace your terminal, your FTP client, and JupyterHub. You can learn how to use it on the Unity docs (Getting Started), but it is mostly straightforward. It has many features which have the potential to make your life easier:

Active job viewer
See information like job status, allocated resources, time remaining.

Job builder/templater
Build and submit your batch jobs with a GUI rather than the CLI.

File explorer & Text Editor
Manage your files with a GUI rather than the CLI.

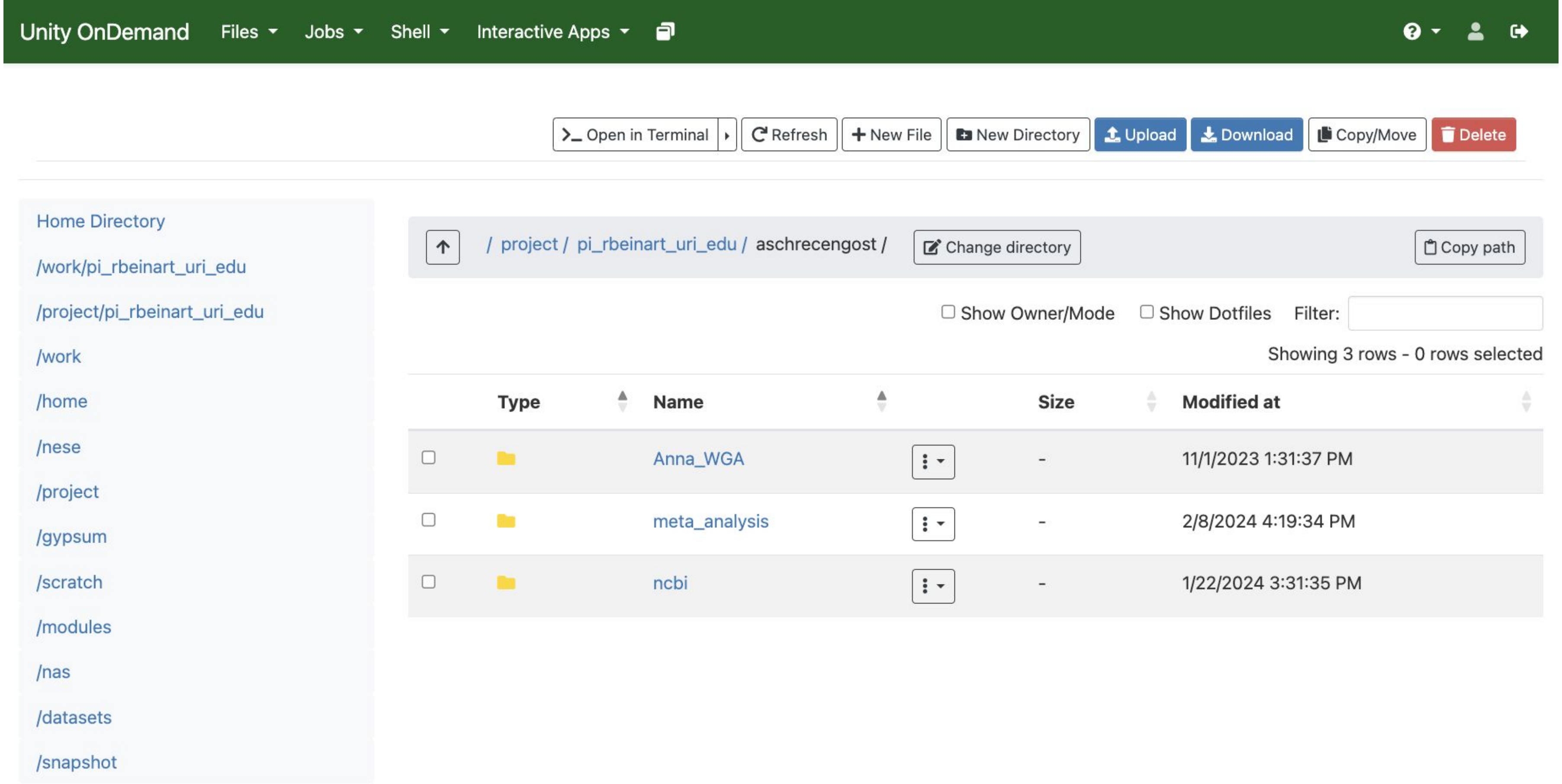
Desktop environment
Graphical applications like Mathematica are now easily accessible. Spawn multiple windows, drag and drop files, use a taskbar and start menu, view images. XFCE is a fully fledged desktop environment like those of Windows and MacOS, though it lacks their visual appeal and can be confusing to learn.

JupyterLab
This should be similar to JupyterHub (without Rstudio/Matlab), but with a few improvements:

- Environment module support
- Access to Anaconda packages
- The file browser now starts at your home directory, which means you won't get permission denied when you try to create a new file
- When your job is in the queue, you aren't left staring at a blank white screen
- More robust job builder with support for custom Slurm arguments
- Debug mode

Rstudio
This should be similar to JupyterHub Rstudio, but with a few improvements:

- Rstudio is containerized, and the new image include tidyverse, devtools, tex, geospatial, and cuda
- Previous R modules (r-sf, r-terra, etc.) are now obsolete, they are pre-installed
- More robust job builder with support for custom Slurm arguments
- Debug mode



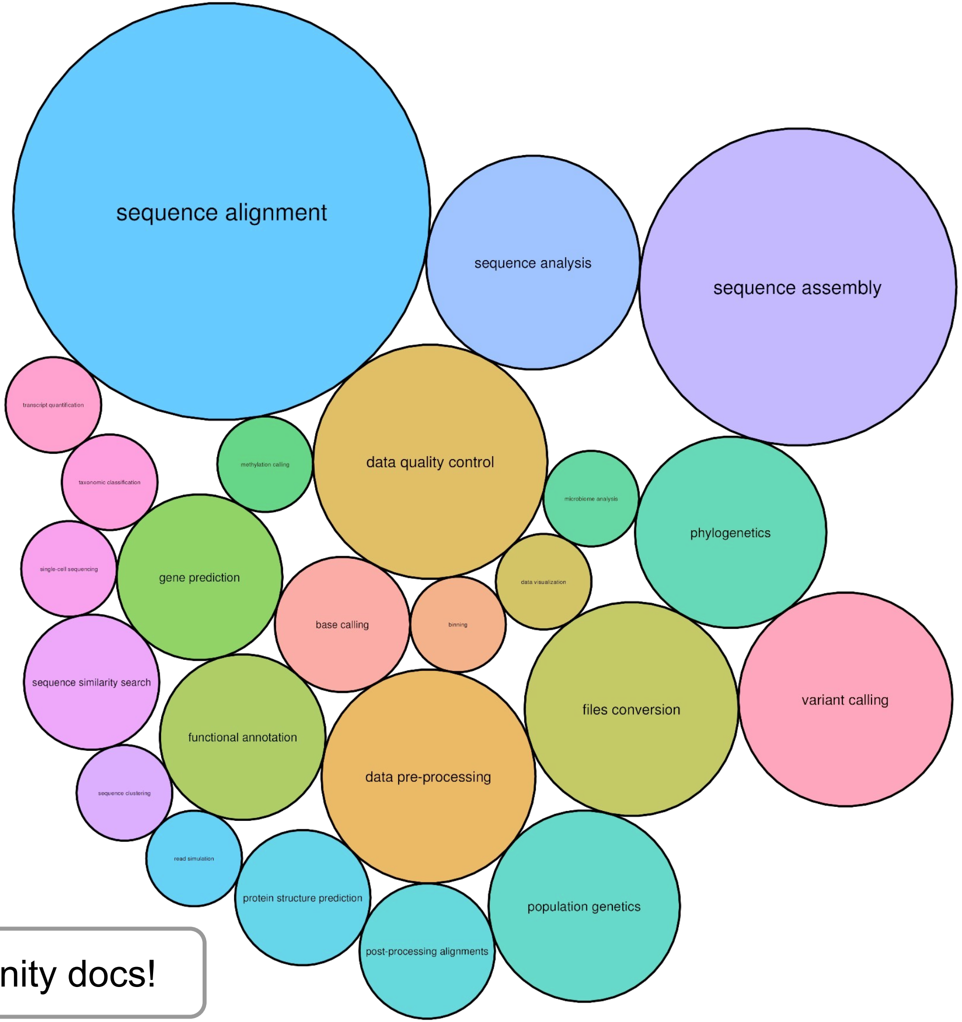
/ project / pi_rbeinart_uri_edu / aschrecengost /

Showing 3 rows - 0 rows selected

	Type	Name	Size	Modified at
□	📁	Anna_WGA	-	11/1/2023 1:31:37 PM
□	📁	meta_analysis	-	2/8/2024 4:19:34 PM
□	📁	ncbi	-	1/22/2024 3:31:35 PM

Bioinformatics tools available on Unity

Bioinformatics softwares on Unity



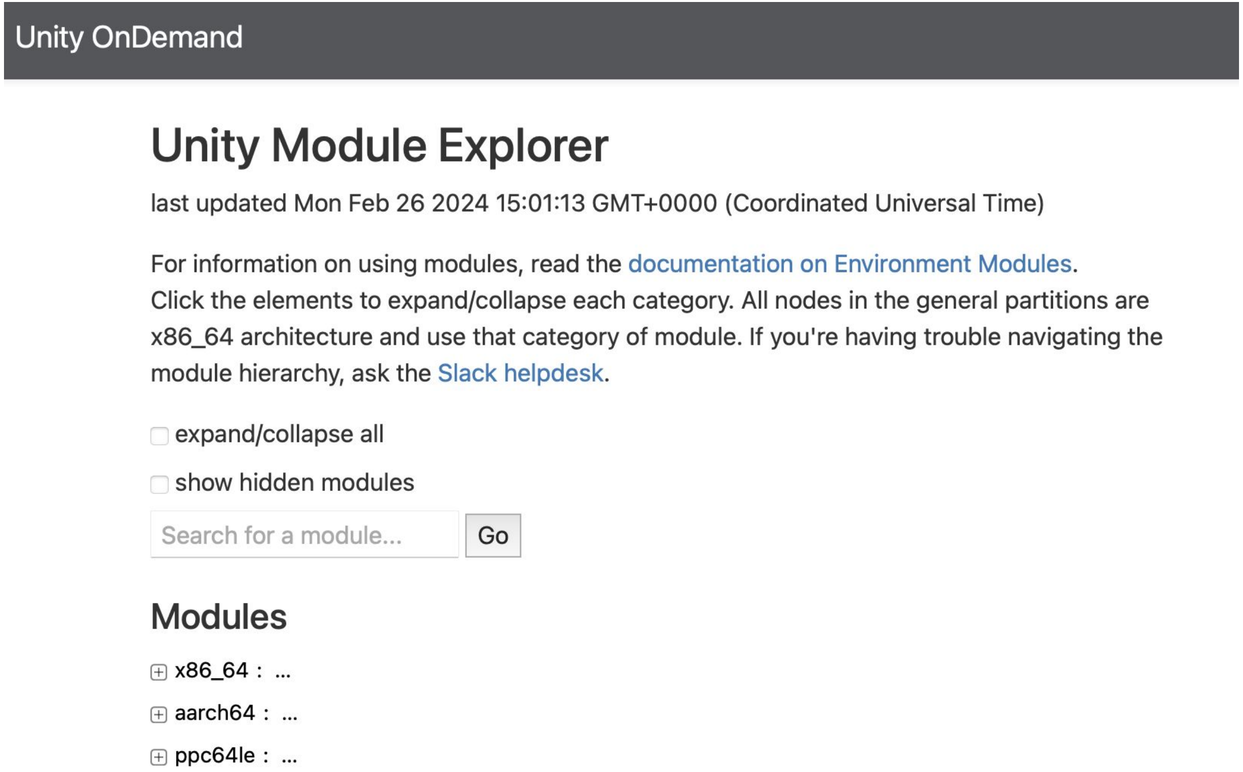
List of bioinformatics tools coming soon on the Unity docs!

Search and load modules available on Unity

Using the command line interface:

- See list of modules available on Unity: `module avail`
- Searching for modules on Unity: `module spider <package name>`
- Load module: `module load <package name>`
- See list of modules loaded: `module list`

Search for modules using Unity OnDemand: Unity module explorer



Unity OnDemand

Unity Module Explorer

last updated Mon Feb 26 2024 15:01:13 GMT+0000 (Coordinated Universal Time)

For information on using modules, read the documentation on Environment Modules.
Click the elements to expand/collapse each category. All nodes in the general partitions are x86_64 architecture and use that category of module. If you're having trouble navigating the module hierarchy, ask the Slack helpdesk.

☐ expand/collapse all

☐ show hidden modules

Search for a module... Go

Modules

⊞ x86_64 : ...

⊞ aarch64 : ...

⊞ ppc64le : ...

Workflow management systems

- Snakemake

```
module load snakemake/7.22.0+py3.11.0
```

- Nextflow

```
module load nextflow/23.04.1
```

  - nf-core provides bioinformatics pipelines to use with Nextflow

Installing softwares on Unity

Installing softwares via
CONDA

- Package manager - allows you to install softwares without admin privileges
- Create conda environments and install all of the software that you need for a particular project/analysis
- Conda manages packages via conda channels - bioconda is a channel which contains 1000s of bioinformatics applications (see packages in bioconda)

Installing softwares via CONDA

- Anaconda and miniconda are available as software modules on Unity

```
module load anaconda/2022.10
```

```
module load miniconda/22.11.1-1
```

- Can install python into a conda environment when you create it or after
- Helpful cheat sheet for conda commands
- Check your software's documentation for how to install via conda, or Google to find its page on anaconda.org

Installing softwares via CONDA

1. Start an interactive session:

```
salloc -p cpu -c 1 --mem=5G --time=01:00:00
```

2. Load miniconda module:

```
module load miniconda/22.11.1-1
```

3. Create conda environment in your PI’s work directory:

```
mkdir -p /work/pi_name/$USER-conda/envs
```

```
conda create --prefix=/work/pi_name/$USER-conda/envs/chopper-env python=3.8
```

$USER → environment variable specifying your username

Installing softwares via CONDA

Why install conda environments in your PI’s work directory?

- Your home directory has a storage capacity of 50GB
- The conda environments are available to every member of the lab

Installing softwares via CONDA

4. Activate conda environment:

```
conda activate /work/pi_name/$USER-conda/envs/chopper-env
```

5. Install software:

```
conda install bioconda::chopper
```

or create conda environment and install packages via a yaml file:

```
wget https://github.com/bioconda/bioconda-recipes/blob/master/recipes/chopper/meta.yaml .
```

```
conda env create --file meta.yaml --prefix=/work/pi_name/$USER-conda/envs/chopper-env
```

Installing softwares via CONDA

6. Get usage message

```
(/work/pi_name/$USER-conda/envs/chopper-env)<username>@cpu001:~$ chopper -h
Filtering and trimming of fastq files. Reads on stdin and writes to stdout.

Usage: chopper [OPTIONS]

Options:
  -q, --quality <MINQUAL>      Sets a minimum Phred average quality score [default: 0]
      --maxqual <MAXQUAL>      Sets a maximum Phred average quality score [default: 1000]
  -l, --minlength <MINLENGTH>  Sets a minimum read length [default: 1]
      --maxlength <MAXLENGTH>  Sets a maximum read length [default: 2147483647]
      --headcrop <HEADCROP>    Trim N nucleotides from the start of a read [default: 0]
      --tailcrop <TAILCROP>    Trim N nucleotides from the end of a read [default: 0]
  -t, --threads <THREADS>      Use N parallel threads [default: 4]
  -c, --contam <CONTAM>        Filter contaminants against a fasta
      --inverse                Output the opposite of the normal results
  -h, --help                   Print help
  -V, --version                Print version
```

Installing softwares via

APPTAINER

- Apptainer (formerly Singularity) is an open source container platform created to run applications on HPC systems
- Container: lightweight executable built by packaging the software code together with the related configuration files, libraries, and dependencies
- Check your software’s documentation if it’s possible to install it via a container

Installing softwares via APPTAINER

- Do a search in the following container image registries:
  - Docker Hub
  - Quay Container Registry

Installing softwares via APPTAINER

Build a container from a given Uniform Resource Identifier (URI):

1. Start an interactive session:

```
salloc -p cpu -c 1 --mem=5G --time=01:00:00
```

2. Load Apptainer module:

```
module load apptainer/latest
```

3. Define location of /.apptainer directory used to store cache files:

```
export APPTAINER_CACHEDIR=<your PI’s work directory>/.apptainer
```

Installing softwares via APPTAINER

4. Pull the Docker image:

```
apptainer pull [pull options...] [output file] docker://<URI>:tag
```

Example:

```
apptainer pull chopper.sif docker://quay.io/biocontainers/chopper:0.7.0--hdcf5f25_0
```

Installing softwares via APPTAINER

5. Run a command within a container:

```
apptainer exec [exec options...] <container> <command>
```

Example:

```
apptainer exec chopper.sif chopper -h
```

Installing softwares via

APPTAINER

Start an interactive shell within a container:

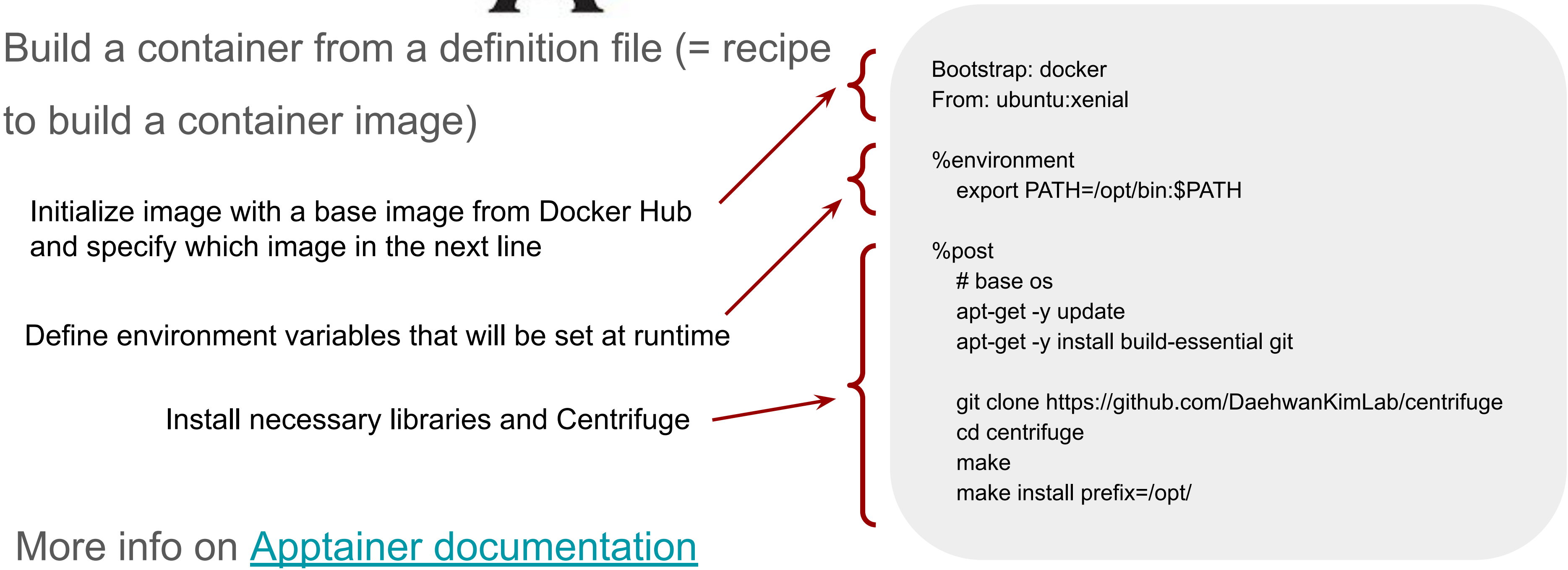
```
apptainer shell [shell options...] <container>
```

Example:

```
apptainer shell chopper.sif
```

Installing softwares via APPTAINER

Build a container from a definition file (= recipe to build a container image)



Initialize image with a base image from Docker Hub and specify which image in the next line

Define environment variables that will be set at runtime

Install necessary libraries and Centrifuge

```
Bootstrap: docker
From: ubuntu:xenial

%environment
   export PATH=/opt/bin:$PATH

%post
   # base os
   apt-get -y update
   apt-get -y install build-essential git

   git clone https://github.com/DaehwanKimLab/centrifuge
   cd centrifuge
   make
   make install prefix=/opt/
```

More info on [Apptainer documentation]

Installing softwares via
APPTAINER

Build a container from a definition file:

1. In an interactive session and after loading the Apptainer module, run:

```
unset APPTAINER_BINDPATH
```

2. Build container:

```
apptainer build centrifuge.sif centrifuge.def
```

3. Run command within container:

```
apptainer exec centrifuge.sif centrifuge -h
```

Installing softwares via APPTAINER

Sharing files between the host (Unity) and the container:

- By default, Apptainer binds your home directory ($HOME) and the current directory ($PWD) into the container
- Use --bind (-B) to specify additional bind mounts/paths. Syntax for using the bind option is “source:destination” (source = destination if not specified). Source has to be an absolute path.

Example: bind your PI’s work directory from the host as /data inside the container:

```
apptainer exec --bind <your PI’s work directory>:/data centrifuge.sif centrifuge -h
```

Installing softwares via R

1. Start an interactive session:

```
salloc -p cpu -c 1 --mem=5G --time=01:00:00
```

2. Load R container module:

```
module load r-rocker-ml-verse/4.2.3+apptainer
```

3. Start an R interactive session

4. Get list of R packages already installed:

```
installed.packages()
```

5. Check if a package is already installed:

```
installed.packages(“<package name>”)
```

6. Install a package:

```
install.packages(“<package name>”)
```

Package name has to be surrounded by quotes!

Packages will automatically be installed in $HOME/R/x86_64-pc-linux-gnu-library/4.2

Installing softwares via R

Share R packages amongst members in the same lab

- Create folder to store packages in PI’s work directory:

```
mkdir -p <path to PI’s work directory>/R/4.2
```

- Install packages using:

```
install.packages(“<package name>”, lib='<path to PI’s work directory>/R/4.2')
```

- Add .libPaths('<path to PI’s work directory>/R/4.2') to ~/.Rprofile file

Installing softwares via R

- Bioconductor is an open software development for computational biology and bioinformatics
- It provides statistical and graphical methods for the analysis of genomic data
- Over 2000 packages

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Also available as a container, see on Docker Hub

Biological datasets available on Unity

- Biological databases are located in /datasets/bio
- List of databases:
  - NCBI databases (ncbi nr, ncbi nt, RefSeq Prokaryotic and Eukaryotic db)
  - Dfam → repetitive DNA families, sequence models, and genome annotations
  - EggNOG → biological information hosted by the EMBL
  - Protein databases: BFD, MGnify, pdb70 pdb_mmcif, pdb_seqres, uniclust30, UniProt, uniref90
  - TARA Oceans

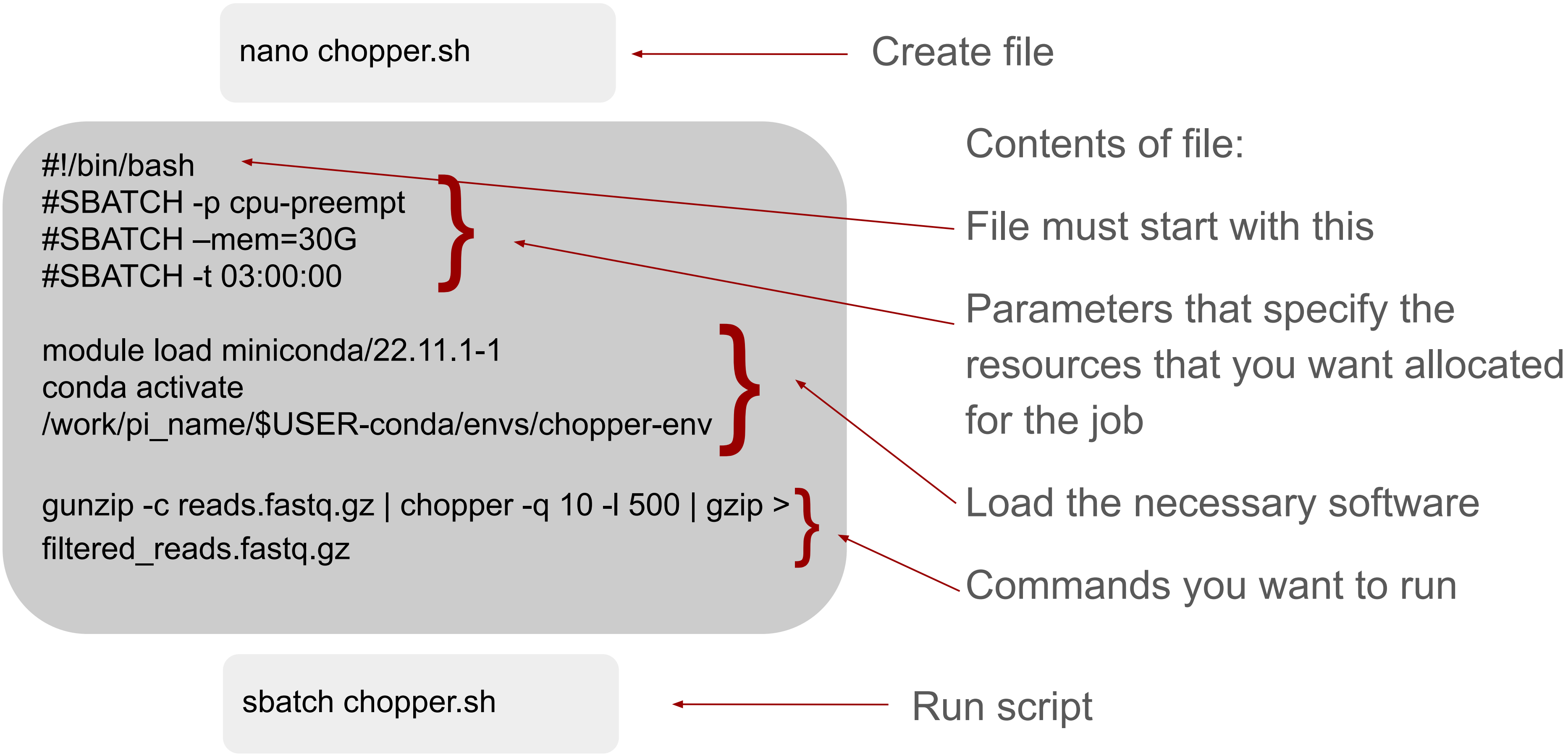
List of biological datasets coming soon on the Unity docs!

Submitting jobs on Unity

Submitting a batch script

- Submit a batch job to the job scheduler (slurm)
- The scheduler manages resources on Unity and executes the job when the resources become available
- The batch job is a task which can be executed without user intervention

Components of a batch script



```
nano chopper.sh
```

Create file

```
#!/bin/bash
#SBATCH -p cpu-preempt
#SBATCH –mem=30G
#SBATCH -t 03:00:00

module load miniconda/22.11.1-1
conda activate
/work/pi_name/$USER-conda/envs/chopper-env

gunzip -c reads.fastq.gz | chopper -q 10 -l 500 | gzip >
filtered_reads.fastq.gz
```

Contents of file:

File must start with this

Parameters that specify the resources that you want allocated for the job

Load the necessary software

Commands you want to run

```
sbatch chopper.sh
```

Run script

How to check on job status

`squeue --me`	Check on the jobs that you are currently running
`scancel <job_id>`	Cancel a specific job
`sinfo`	Check on the status of the cluster
`sacct`	Look at your entire slurm history
`scontrol show job <job_id>`	Show information for queued or running job
`scontrol update <job_id> partition=cpu,cpu-long`	If job is in queue for a long time, can update partition

Running an interactive job

- Allows you to access CLI on a compute node
- 8 hour time limit

```
salloc -p cpu-preempt --mem=30G --time=03:00:00
```

Parameters that specify the resources that you want allocated for the job

```
module load anaconda/2022.10
conda activate /work/pi_name/$USER-conda/envs/chopper-env
```

Load the necessary software

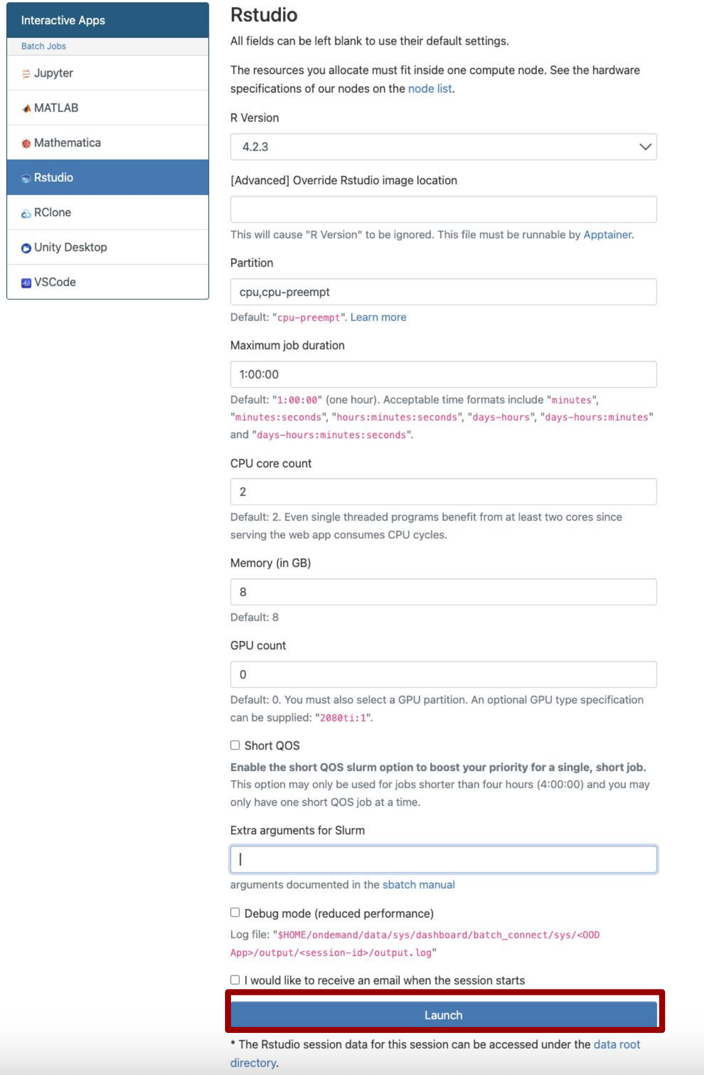
```
gunzip -c reads.fastq.gz | chopper -q 10 -l 500 | gzip >
filtered_reads.fastq.gz
```

Run your commands

Running interactive applications on Unity Ondemand

- This is a great option if you like to run your code in e.g. RStudio or in a Jupyter notebook and don’t need more than 8 hours
- Available interactive apps: Jupyter, MATLAB, Mathematica, RStudio, rclone, VScode, and Unity Desktop
  - Requests for new graphical apps are accepted, or we will tell you if it’s available via Unity Desktop
  - Tutorial for Jupyter OnDemand

Unity OnDemand - RStudio



Rstudio

All fields can be left blank to use their default settings.

The resources you allocate must fit inside one compute node. See the hardware specifications of our nodes on the node list.

R Version

4.2.3

[Advanced] Override Rstudio image location

This will cause "R Version" to be ignored. This file must be runnable by Apptainer.

Partition

cpu,cpu-preempt

Default: "cpu-preempt". Learn more

Maximum job duration

1:00:00

Default: "1:00:00" (one hour). Acceptable time formats include "minutes", "minutes:seconds", "hours:minutes:seconds", "days-hours", "days-hours:minutes" and "days-hours:minutes:seconds".

CPU core count

2

Default: 2. Even single threaded programs benefit from at least two cores since serving the web app consumes CPU cycles.

Memory (in GB)

8

Default: 8

GPU count

0

Default: 0. You must also select a GPU partition. An optional GPU type specification can be supplied: "2080ti:1".

□ Short QOS

Enable the short QOS slurm option to boost your priority for a single, short job. This option may only be used for jobs shorter than four hours (4:00:00) and you may only have one short QOS job at a time.

Extra arguments for Slurm

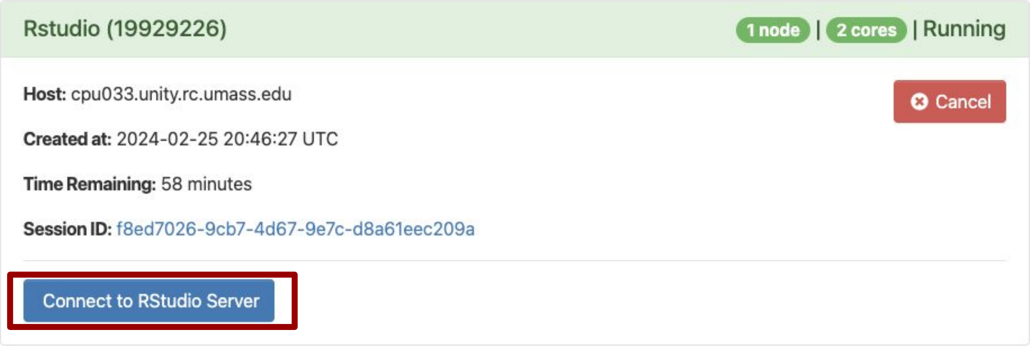
arguments documented in the sbatch manual

□ Debug mode (reduced performance)

Log file: "$HOME/ondemand/data/sys/dashboard/batch_connect/sys/<OOD App>/output/<session-id>/output.log"

□ I would like to receive an email when the session starts

* The Rstudio session data for this session can be accessed under the data root directory.



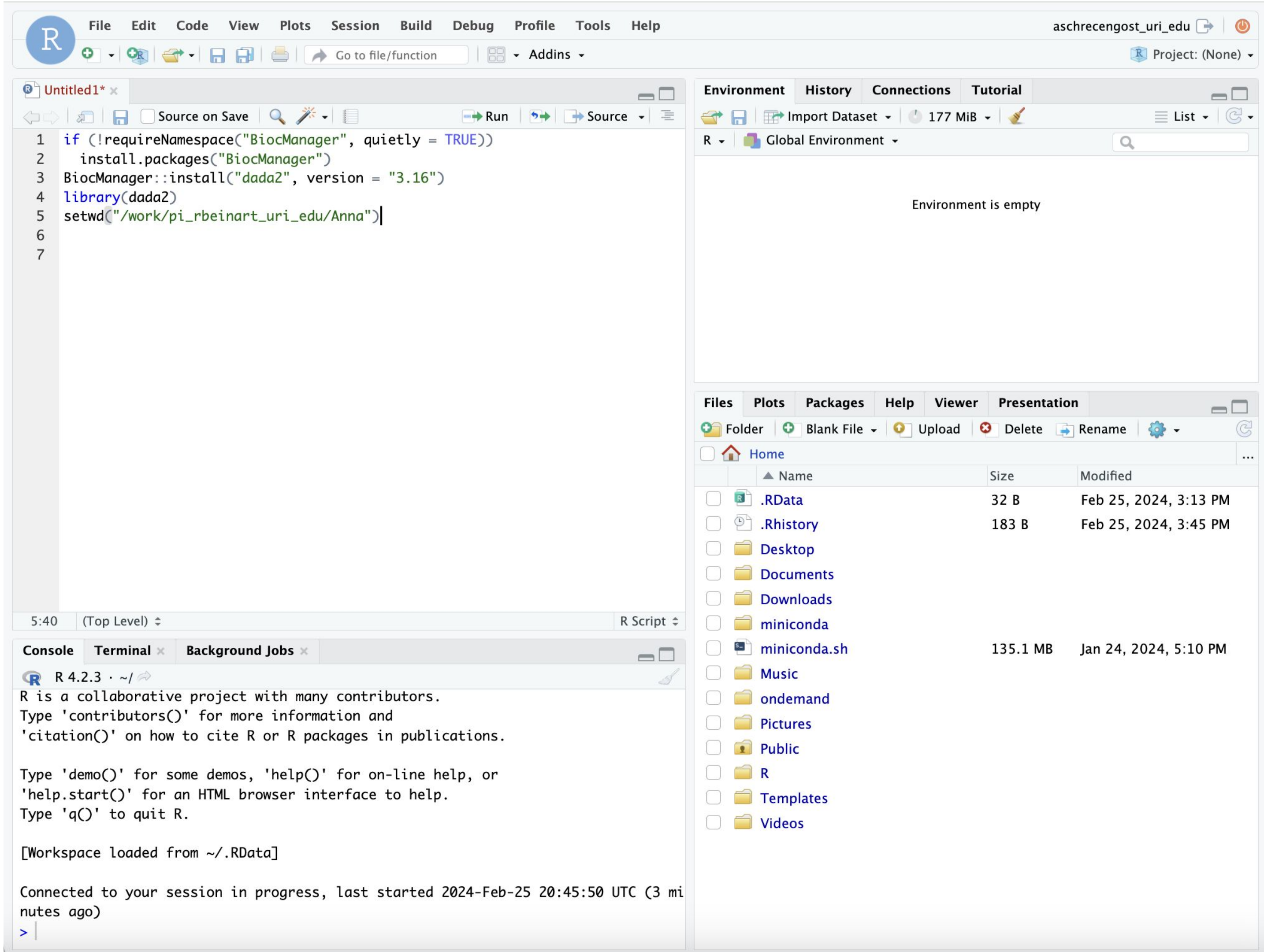
Rstudio (19929226) | 1 node | 2 cores | Running

Host: cpu033.unity.rc.umass.edu

Created at: 2024-02-25 20:46:27 UTC

Time Remaining: 58 minutes

Session ID: f8ed7026-9cb7-4d67-9e7c-d8a61eec209a



```r
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("dada2", version = "3.16")
library(dada2)
setwd("/work/pi_rbeinart_uri_edu/Anna")
```

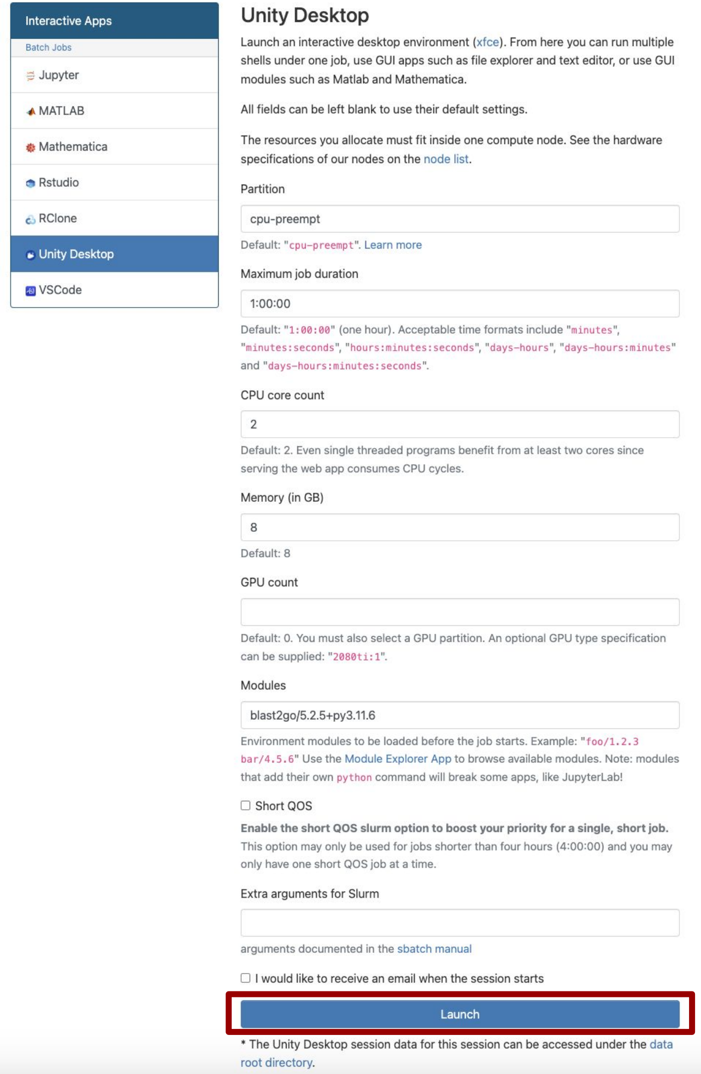
```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

Connected to your session in progress, last started 2024-Feb-25 20:45:50 UTC (3 mi
nutes ago)
>
```

Environment is empty

Name	Size	Modified
.RData	32 B	Feb 25, 2024, 3:13 PM
.Rhistory	183 B	Feb 25, 2024, 3:45 PM
Desktop		
Documents		
Downloads		
miniconda		
miniconda.sh	135.1 MB	Jan 24, 2024, 5:10 PM
Music		
ondemand		
Pictures		
Public		
R		
Templates		
Videos		

Unity OnDemand - Desktop

- If there is a GUI that you would like to run on Unity but which isn’t available as an interactive app, you may be able to run it via Unity Desktop



Unity Desktop

Launch an interactive desktop environment (xfce). From here you can run multiple shells under one job, use GUI apps such as file explorer and text editor, or use GUI modules such as Matlab and Mathematica.

All fields can be left blank to use their default settings.

The resources you allocate must fit inside one compute node. See the hardware specifications of our nodes on the node list.

Partition

cpu-preempt

Default: "cpu-preempt". Learn more

Maximum job duration

1:00:00

Default: "1:00:00" (one hour). Acceptable time formats include "minutes", "minutes:seconds", "hours:minutes:seconds", "days-hours", "days-hours:minutes" and "days-hours:minutes:seconds".

CPU core count

2

Default: 2. Even single threaded programs benefit from at least two cores since serving the web app consumes CPU cycles.

Memory (in GB)

8

Default: 8

GPU count

Default: 0. You must also select a GPU partition. An optional GPU type specification can be supplied: "2080ti:1".

Modules

blast2go/5.2.5+py3.11.6

Environment modules to be loaded before the job starts. Example: "foo/1.2.3 bar/4.5.6" Use the Module Explorer App to browse available modules. Note: modules that add their own python command will break some apps, like JupyterLab!

☐ Short QOS

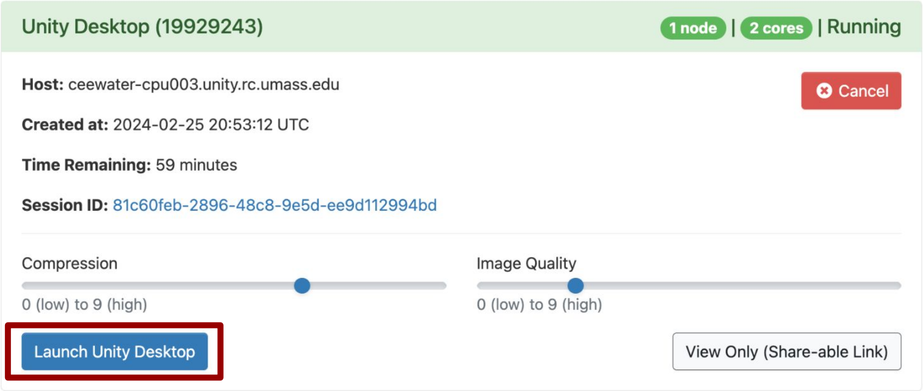
Enable the short QOS slurm option to boost your priority for a single, short job. This option may only be used for jobs shorter than four hours (4:00:00) and you may only have one short QOS job at a time.

Extra arguments for Slurm

arguments documented in the sbatch manual

☐ I would like to receive an email when the session starts

* The Unity Desktop session data for this session can be accessed under the data root directory.



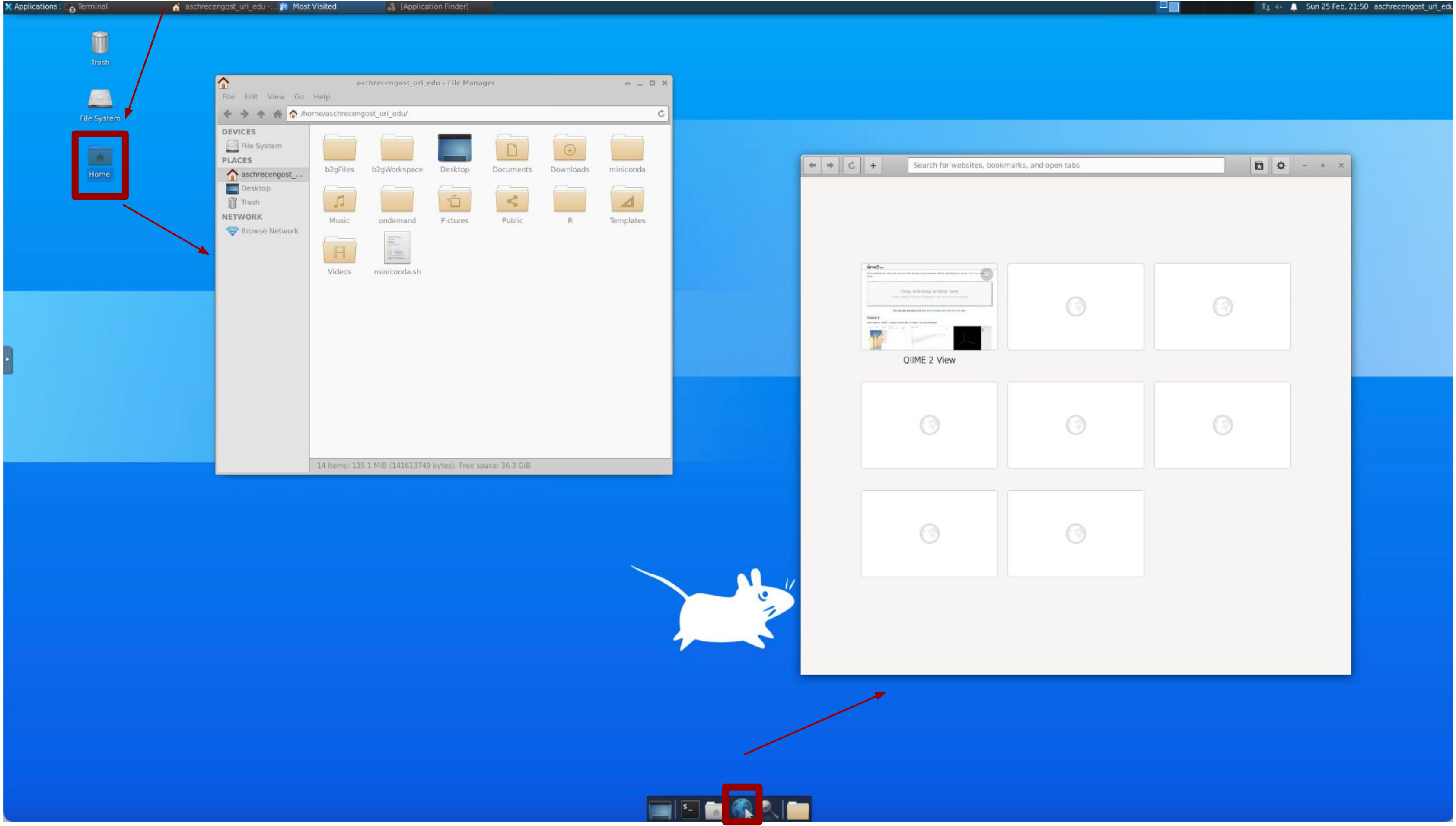
Host: ceewater-cpu003.unity.rc.umass.edu

Created at: 2024-02-25 20:53:12 UTC

Time Remaining: 59 minutes

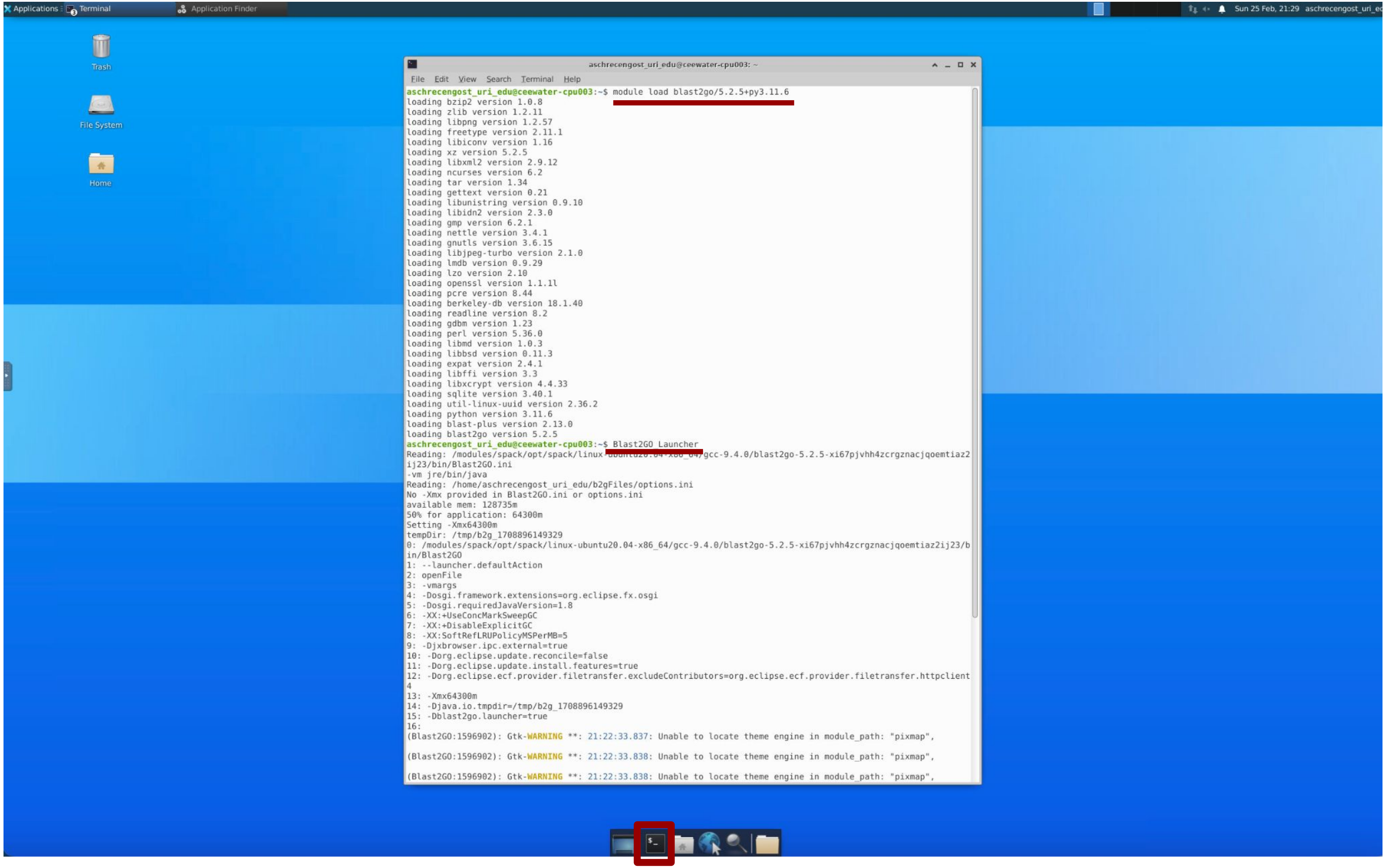
Session ID: 81c60feb-2896-48c8-9e5d-ee9d112994bd

File manager



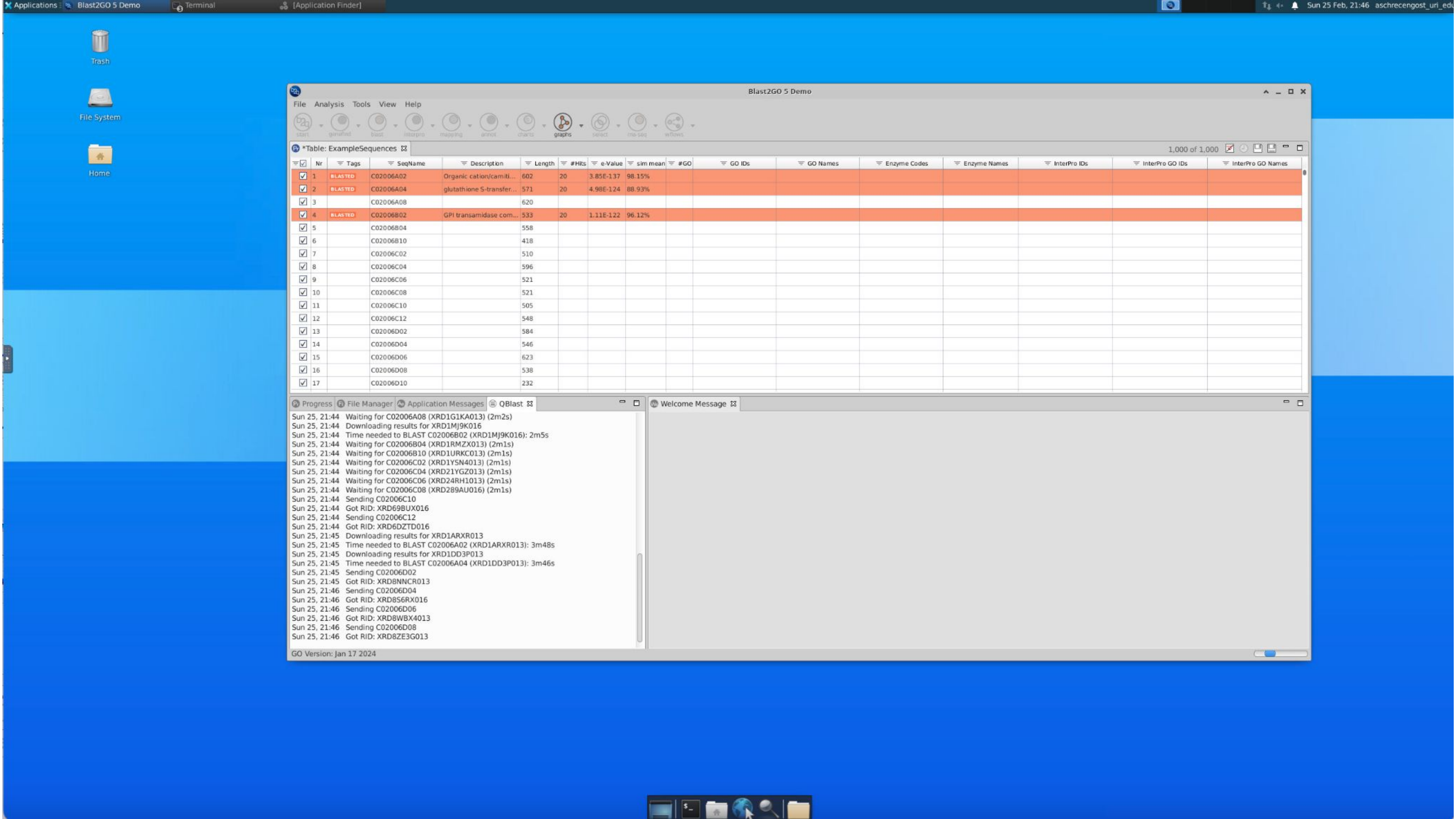
14 items: 135.1 MiB (141613749 bytes), Free space: 36.3 GiB

Web Browser



```
aschrecengost_uri_edu@ceewater-cpu003:~$ module load blast2go/5.2.5+py3.11.6
loading bzip2 version 1.0.8
loading zlib version 1.2.11
loading libpng version 1.2.57
loading freetype version 2.11.1
loading libiconv version 1.16
loading xz version 5.2.5
loading libxml2 version 2.9.12
loading ncurses version 6.2
loading tar version 1.34
loading gettext version 0.21
loading libunistring version 0.9.10
loading libidn2 version 2.3.0
loading gmp version 6.2.1
loading nettle version 3.4.1
loading gnutls version 3.6.15
loading libjpeg-turbo version 2.1.0
loading lmdb version 0.9.29
loading lzo version 2.10
loading openssl version 1.1.1l
loading pcre version 8.44
loading berkeley-db version 18.1.40
loading readline version 8.2
loading gdbm version 1.23
loading perl version 5.36.0
loading libmd version 1.0.3
loading libbsd version 0.11.3
loading expat version 2.4.1
loading libffi version 3.3
loading libxcrypt version 4.4.33
loading sqlite version 3.40.1
loading util-linux-uuid version 2.36.2
loading python version 3.11.6
loading blast-plus version 2.13.0
loading blast2go version 5.2.5
aschrecengost_uri_edu@ceewater-cpu003:~$ Blast2GO_Launcher
Reading: /modules/spack/opt/spack/linux-ubuntu20.04-x86_64/gcc-9.4.0/blast2go-5.2.5-xi67pjvhh4zcrgznacjqoemtiaz2
ij23/bin/Blast2GO.ini
-vm jre/bin/java
Reading: /home/aschrecengost_uri_edu/b2gFiles/options.ini
No -Xmx provided in Blast2GO.ini or options.ini
available mem: 128735m
50% for application: 64300m
Setting -Xmx64300m
tempDir: /tmp/b2g_1708896149329
0: /modules/spack/opt/spack/linux-ubuntu20.04-x86_64/gcc-9.4.0/blast2go-5.2.5-xi67pjvhh4zcrgznacjqoemtiaz2ij23/b
in/Blast2GO
1: --launcher.defaultAction
2: openFile
3: -vmargs
4: -Dosgi.framework.extensions=org.eclipse.fx.osgi
5: -Dosgi.requiredJavaVersion=1.8
6: -XX:+UseConcMarkSweepGC
7: -XX:+DisableExplicitGC
8: -XX:SoftRefLRUPolicyMSPerMB=5
9: -Djxbrowser.ipc.external=true
10: -Dorg.eclipse.update.reconcile=false
11: -Dorg.eclipse.update.install.features=true
12: -Dorg.eclipse.ecf.provider.filetransfer.excludeContributors=org.eclipse.ecf.provider.filetransfer.httpclient
4
13: -Xmx64300m
14: -Djava.io.tmpdir=/tmp/b2g_1708896149329
15: -Dblast2go.launcher=true
16:
(Blast2GO:1596902): Gtk-WARNING **: 21:22:33.837: Unable to locate theme engine in module_path: "pixmap",

(Blast2GO:1596902): Gtk-WARNING **: 21:22:33.838: Unable to locate theme engine in module_path: "pixmap",

(Blast2GO:1596902): Gtk-WARNING **: 21:22:33.838: Unable to locate theme engine in module_path: "pixmap",
```

Terminal emulator



☑	Nr	Tags	SeqName	Description	Length	#Hits	e-Value	sim mean	#GO	GO IDs	GO Names	Enzyme Codes	Enzyme Names	InterPro IDs	InterPro GO IDs	InterPro GO Names
☑	1	BLASTED	C02006A02	Organic cation/carniti...	602	20	3.85E-137	98.15%								
☑	2	BLASTED	C02006A04	glutathione S-transfer...	571	20	4.98E-124	88.93%								
☑	3		C02006A08		620											
☑	4	BLASTED	C02006B02	GPI transamidase com...	533	20	1.11E-122	96.12%								
☑	5		C02006B04		558											
☑	6		C02006B10		418											
☑	7		C02006C02		510											
☑	8		C02006C04		596											
☑	9		C02006C06		521											
☑	10		C02006C08		521											
☑	11		C02006C10		505											
☑	12		C02006C12		548											
☑	13		C02006D02		584											
☑	14		C02006D04		546											
☑	15		C02006D06		623											
☑	16		C02006D08		538											
☑	17		C02006D10		232											

```
Sun 25, 21:44   Waiting for C02006A08 (XRD1G1KA013) (2m2s)
Sun 25, 21:44   Downloading results for XRD1MJ9K016
Sun 25, 21:44   Time needed to BLAST C02006B02 (XRD1MJ9K016): 2m5s
Sun 25, 21:44   Waiting for C02006B04 (XRD1RMZX013) (2m1s)
Sun 25, 21:44   Waiting for C02006B10 (XRD1URKC013) (2m1s)
Sun 25, 21:44   Waiting for C02006C02 (XRD1YSN4013) (2m1s)
Sun 25, 21:44   Waiting for C02006C04 (XRD21YGZ013) (2m1s)
Sun 25, 21:44   Waiting for C02006C06 (XRD24RH1013) (2m1s)
Sun 25, 21:44   Waiting for C02006C08 (XRD289AU016) (2m1s)
Sun 25, 21:44   Sending C02006C10
Sun 25, 21:44   Got RID: XRD69BUX016
Sun 25, 21:44   Sending C02006C12
Sun 25, 21:44   Got RID: XRD6DZTD016
Sun 25, 21:45   Downloading results for XRD1ARXR013
Sun 25, 21:45   Time needed to BLAST C02006A02 (XRD1ARXR013): 3m48s
Sun 25, 21:45   Downloading results for XRD1DD3P013
Sun 25, 21:45   Time needed to BLAST C02006A04 (XRD1DD3P013): 3m46s
Sun 25, 21:45   Sending C02006D02
Sun 25, 21:45   Got RID: XRD8NNCR013
Sun 25, 21:46   Sending C02006D04
Sun 25, 21:46   Got RID: XRD8S6RX016
Sun 25, 21:46   Sending C02006D06
Sun 25, 21:46   Got RID: XRD8WBX4013
Sun 25, 21:46   Sending C02006D08
Sun 25, 21:46   Got RID: XRD8ZE3G013
```

GO Version: Jan 17 2024

Additional Resources

- Unity Onboarding video (Spring 2024)
- Snakemake workshop
- Conda docs on Unity
- AI lab workshops available this semester
- Unity community Slack
- More contact information

Next workshop

- More details about job submission, resource allocation, and how to ensure your code is running efficiently on Unity
- Monday, March 25 @ 11 am - Zoom link here