

Introduction to the Amarel Cluster

W. Evan Johnson, Ph.D.

Professor, Division of Infectious Disease

Director, Center for Data Science

Co-Direcor, Center for Biomedical Informatics and Health AI

Rutgers University – New Jersey Medical School

2025-07-27

Introduction to the Amarel HPC Cluster

HPC at Rutgers

About Amarel

In July 2017, the [Office of Advanced Research Computing](#) (OARC) unveiled Amarel, a “condominium” style computing environment developed to serve the university’s wide-ranging research needs. The Amarel cluster provides a shared platform that optimizes resources for the benefit of all users. Named in honor of **Dr. Saul Amarel**, one of the founders of the Rutgers Computer Science Department and contributor to advanced computing and artificial intelligence research and methodologies, Amarel is designed to suit many different research applications.

Through Aberdeen (late 2023):

- 560 compute nodes
 - 27,872 Intel Xeon cores
 - 224 GPUs
 - Open OnDemand servers
 - InfiniBand FDR & EDR fabric

More information:

- Community-contributed software repository
- Spans three Rutgers data centers (Piscataway, Newark, and Camden) producing a unified compute, data, and storage system
- Rolling node/phase retirement with new node replacement, beginning in spring 2021

Key Features of Amarel

1. **Compute Nodes:** Amarel consists of multiple compute nodes with varying specifications, including CPUs and GPUs.
2. **Parallel File System:** A high-speed parallel file system for storing and accessing large datasets.
3. **Job Scheduler:** A job scheduling system for managing computing resources efficiently.
4. **Software Stack:** A comprehensive collection of software packages and libraries for various scientific computing tasks.
5. **Networking:** High-speed interconnects for fast communication between compute nodes.

Getting Started with Amarel

If you are off campus, you will need access to VPN:

[Click here for Rutgers VPN instructions for Windows Users](#)

[Click here for Rutgers VPN instructions for Mac Users](#)

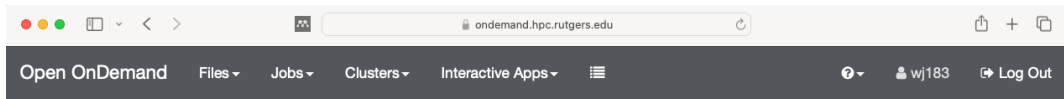
Getting Started with Amarel

1. **Request Access:** Contact OARC to request access to Amarel
2. **Training Workshops:** Attend training workshops offered by OARC to learn how to use Amarel effectively.
3. **Documentation and Support:** Explore documentation and seek support from OARC staff and the user community.
4. **Start Small:** Begin with small-scale experiments and gradually scale up as needed.

Introduction to OnDemand for HPC Clusters

What is OnDemand?

- ▶ OnDemand is a web-based platform that provides a user-friendly interface for accessing and managing (HPC) clusters.
- ▶ It simplifies the process of submitting and monitoring jobs, accessing software, and managing data on HPC systems.
- ▶ Go to: <http://ondemand.hpc.rutgers.edu> (VPN for off campus).



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

Key Features of OnDemand

1. **Web Interface:** Access HPC resources through a web browser from anywhere with an internet connection.
2. **Job Submission:** Submit and manage computational jobs without needing to use command-line interfaces.
3. **File Management:** Upload, download, and manage files and data on the HPC cluster directly from the browser.
4. **Interactive Sessions:** Launch interactive computing sessions for data analysis and exploration.
5. **Software Environment:** Access a variety of software packages and development tools installed on the HPC cluster.

Benefits of OnDemand

- ▶ **User-Friendly:** OnDemand provides a simplified interface, making HPC resources more accessible to a wider range of users.
- ▶ **Remote Access:** Users can access HPC resources remotely without needing to install any special software.
- ▶ **Increased Productivity:** Streamlined workflows and intuitive interfaces help users focus on their research instead of dealing with technical complexities.

Accessing Files

Dashboard

File Explorer - /home/wj183/

File Explorer v1.4.1

Go To...

>_ Open in Terminal

New File

New Dir

Upload

Show Dotfiles

Show Owner/Mode

Home Directory

2023_03_ugandan_nasal_blood

Desktop

Documents

Downloads

Music

Pictures

Public

R

Templates

Videos

bin

for_howard

intel

my.R.libs

ondemand

rna_seq

/home/wj183/

View

Edit

A-z Rename

Download

Copy

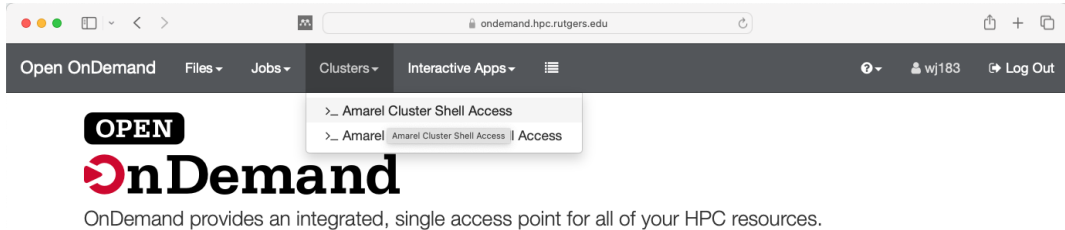
Paste

(Un)Select All

Delete

name	size	modified date
..	<dir>	
2023_03_ugandan_nasal_blood	<dir>	03/31/2023
Desktop	<dir>	03/27/2023
Documents	<dir>	03/21/2023
Downloads	<dir>	03/21/2023
Music	<dir>	03/21/2023
Pictures	<dir>	03/21/2023
Public	<dir>	03/21/2023
R	<dir>	03/31/2023
Templates	<dir>	03/21/2023
Videos	<dir>	03/21/2023
bin	<dir>	03/21/2023

Accessing the Terminal



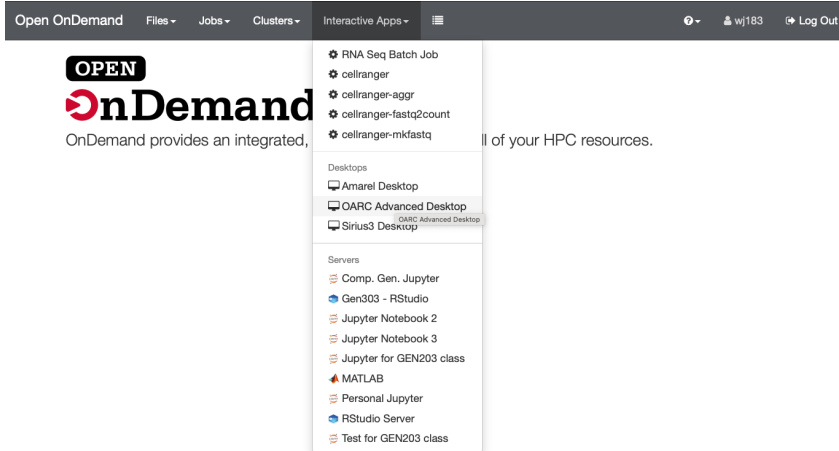
The screenshot shows a web browser window at `ondemand.hpc.rutgers.edu`. The navigation bar includes links for Open OnDemand, Files, Jobs, Clusters, and Interactive Apps. The Clusters dropdown menu is open, displaying two options: ">_ Amarel Cluster Shell Access" and ">_ Amarel Amarel Cluster Shell Access | Access". The main content area features the "OPEN OnDemand" logo and the text: "OnDemand provides an integrated, single access point for all of your HPC resources."

Accessing the Terminal

```

Dashboard wj183@amarel1:~
wj183@amarel.hpc.rutgers.edu's password:
#####
AMAREL AT RUTGERS
Welcome to the AMAREL research computing cluster managed by
the Office of Advanced Research Computing (OARC) at Rutgers University
http://oarc.rutgers.edu
Need help? Send an e-mail to help@oarc.rutgers.edu
User documentation: https://sites.google.com/view/cluster-user-guide
Do NOT run programs, tests, analyses, or pre/postprocessing on the shared
login nodes (amarel1, amarel2, etc.). That's what compute nodes are for.
#####
[wj183@amarel1 ~]$
```

Accessing the Amarel Desktop through OnDemand



The screenshot shows the Open OnDemand web interface. The top navigation bar includes links for Open OnDemand, Files, Jobs, Clusters, and Interactive Apps. The Interactive Apps menu is open, displaying a list of applications. Under the 'Desktops' section, the 'Amarel Desktop' is highlighted. Other desktops listed include OARC Advanced Desktop and Sirius3 Desktop. Under the 'Servers' section, various Jupyter and RStudio environments are listed.

OPEN
OnDemand
OnDemand provides an integrated, ... of your HPC resources.

Interactive Apps ▾

- RNA Seq Batch Job
- cellranger
- cellranger-aggr
- cellranger-fastq2count
- cellranger-mkfastq

Desktops

- Amarel Desktop
- OARC Advanced Desktop
- Sirius3 Desktop

Servers

- Comp. Gen. Jupyter
- Gen303 - RStudio
- Jupyter Notebook 2
- Jupyter Notebook 3
- Jupyter for GEN203 class
- MATLAB
- Personal Jupyter
- RStudio Server
- Test for GEN203 class

Accessing the Amarel Desktop through OnDemand

Open OnDemand
Files
Jobs
Clusters
Interactive Apps

Home / My Interactive Sessions / Amarel Desktop

Interactive Apps

Desktops

Amarel Desktop

OARC Advanced Desktop

Sirius3 Desktop

Servers

Comp. Gen. Jupyter

Gen303 - RStudio

Jupyter Notebook 2

Jupyter Notebook 3

Jupyter for GEN203 class

MATLAB

Personal Jupyter

RStudio Server

Test for GEN203 class

Amarel Desktop

This app will launch an interactive desktop on one or more compute nodes. You will have full access to the resources these nodes provide. This is analogous to an interactive batch job.

Number of hours

Num Cores

Gigabytes of memory

Number of gigabytes of memory (larger values may mean longer wait)

Partition

Reservation

Launch

* All Amarel Desktop session data is generated and stored under the user's home directory in the corresponding [data root directory](#).

Accessing the Amarel Desktop through OnDemand

Open OnDemand Files ▾ Jobs ▾ Clusters ▾ Interactive Apps ▾   ▾  wj183  Log Out

Session was successfully created. 

[Home](#) / My Interactive Sessions

Interactive Apps

Desktops

 Amarel Desktop

 OARC Advanced Desktop

 Sirius3 Desktop

Servers

 Comp. Gen. Jupyter

 Gen303 - RStudio

 Jupyter Notebook 2

 Jupyter Notebook 3

 Jupyter for GEN203 class

 MATLAB

Amarel Desktop (36349078)

Queued

Created at: 2024-04-29 10:51:36 EDT

Time Requested: 12 hours

Session ID: [d6886f60-dd03-4dad-b856-8f02e081f209](#)

 Delete

Please be patient as your job currently sits in queue. The wait time depends on the number of cores as well as time requested.

Accessing the Amarel Desktop through OnDemand

Open OnDemand Files ▾ Jobs ▾ Clusters ▾ Interactive Apps ▾   wj183  Log Out

Session was successfully created. 

[Home](#) / My Interactive Sessions

Interactive Apps

Desktops

 Amarel Desktop

 OARC Advanced Desktop

 Sirius3 Desktop

Servers

 Comp. Gen. Jupyter

 Gen303 - RStudio

 Jupyter Notebook 2

 Jupyter Notebook 3

 Jupyter for GEN203 class

 MATLAB

Amarel Desktop (36349078)

1 node | 1 core | Running

Host: hal0005.amarel.rutgers.edu

 Delete

Created at: 2024-04-29 10:51:36 EDT

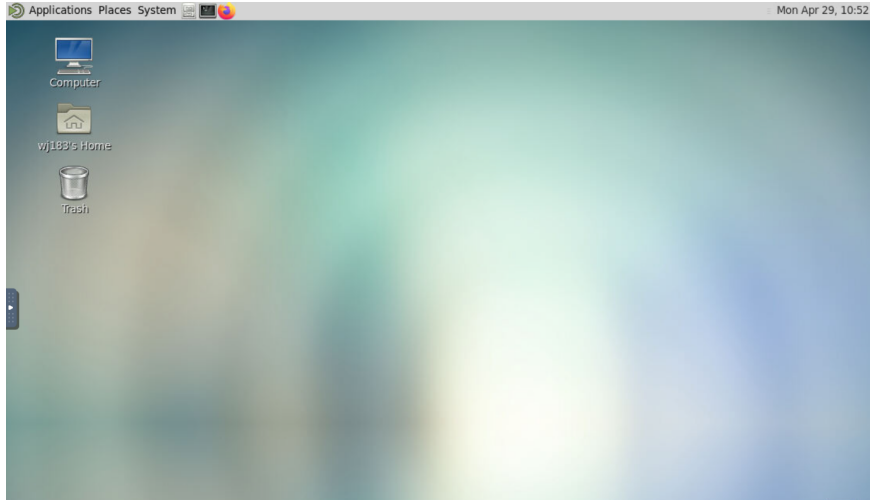
Time Remaining: about 12 hours

Session ID: d6886f60-dd03-4dad-b856-8f02e081f209

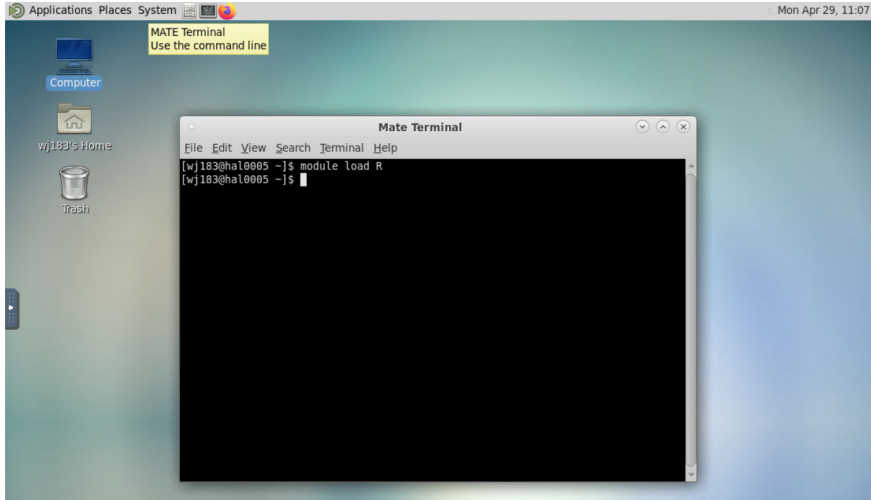
 Launch noVNC in New Tab

View Only (Share-able Link)

Accessing the Amarel Desktop through OnDemand



Accessing the Amarel Desktop through OnDemand

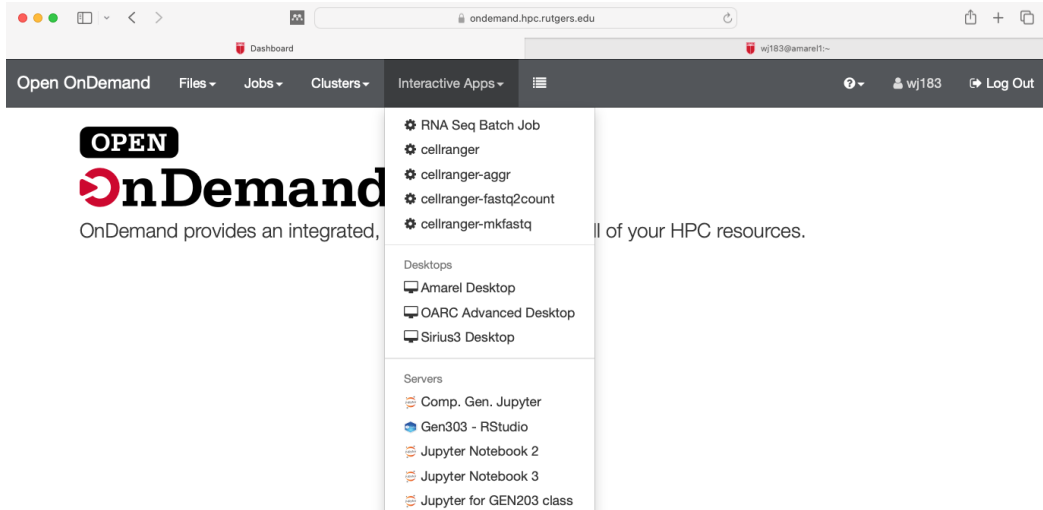


Accessing R through the terminal

You first need to attach R to your session (using Amarel Desktop):

```
[wj183@hal0005 ~]$ module load R  
[wj183@amarel1 ~]$ R
```

Accessing the RStudio through OnDemand



The screenshot shows a web browser window at `ondemand.hpc.rutgers.edu`. The dashboard includes a navigation bar with options: Open OnDemand, Files, Jobs, Clusters, and Interactive Apps. The 'Interactive Apps' menu is open, displaying a list of applications. Under the 'Servers' section, 'Gen303 - RStudio' is highlighted. The main content area features the 'OPEN OnDemand' logo and a description of the platform's capabilities.

OPEN OnDemand

OnDemand provides an integrated, ... of your HPC resources.

Interactive Apps

- RNA Seq Batch Job
- cellranger
- cellranger-aggr
- cellranger-fastq2count
- cellranger-mkfastq

Desktops

- Amarel Desktop
- OARC Advanced Desktop
- Sirius3 Desktop

Servers

- Comp. Gen. Jupyter
- Gen303 - RStudio**
- Jupyter Notebook 2
- Jupyter Notebook 3
- Jupyter for GEN203 class

Accessing the RStudio through OnDemand

Open OnDemand
Files
Jobs
Clusters
Interactive Apps
Log Out

Home / My Interactive Sessions / RStudio Server

Interactive Apps

Desktops

Amarel Desktop

OARC Advanced Desktop

Sirius3 Desktop

Servers

Comp. Gen. Jupyter

Gen303 - RStudio

Jupyter Notebook 2

Jupyter Notebook 3

Jupyter for GEN203 class

MATLAB

Personal Jupyter

RStudio Server

Test for GEN203 class

RStudio Server

This app will launch [RStudio Server](#) an IDE for R on the [Amarel cluster](#).

Partition

nonpre

- nonpre** - nonpreemptible partition (a subset of nodes that are community owned)
- main** - all nodes, including owners
- graphical** - single node dedicated to small interactive jobs (2 cores, 4G RAM maximum)

R version

RStudio with R version 3.5.2

This defines the version of R you want to load.

Number of hours

12

Number of cores

1

Number of cores on node type (4 GB per core)

Launch

* All RStudio Server session data is generated and stored under the user's home directory in the corresponding [data root directory](#).

Accessing the RStudio through OnDemand

Open OnDemand Files Jobs Clusters Interactive Apps   wj183 Log Out

Session was successfully created.

[Home](#) / My Interactive Sessions

Interactive Apps

Desktops

 Amarel Desktop

 OARC Advanced Desktop

 Sirius3 Desktop

Servers

 Comp. Gen. Jupyter

 Gen303 - RStudio

 Jupyter Notebook 2

 Jupyter Notebook 3

 Jupyter for GEN203 class

 MATLAB

RStudio Server (36349030)

Queued

Created at: 2024-04-29 10:29:19 EDT

Time Requested: 12 hours

Session ID: [a26538a5-1a7d-41d8-b74c-093795747c2a](#)

 Delete

Please be patient as your job currently sits in queue. The wait time depends on the number of cores as well as time requested.

Accessing the RStudio through OnDemand

Open OnDemand Files ▾ Jobs ▾ Clusters ▾ Interactive Apps ▾ ⓘ wj183 Log Out

Session was successfully created. ✕

[Home](#) / My Interactive Sessions

Interactive Apps

Desktops

Amarel Desktop

OARC Advanced Desktop

Sirius3 Desktop

Servers

Comp. Gen. Jupyter

Gen303 - RStudio

Jupyter Notebook 2

Jupyter Notebook 3

Jupyter for GEN203 class

MATLAB

RStudio Server (36349030)

1 node | 1 core | Running

Host: hal0100.amarel.rutgers.edu

Delete

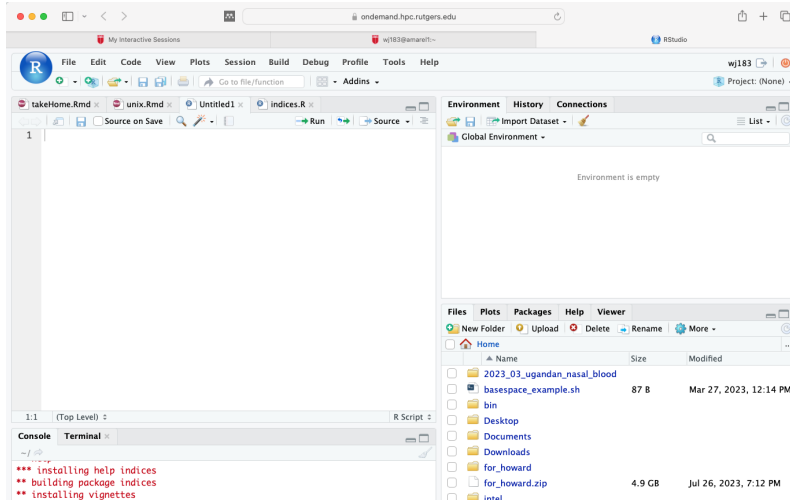
Created at: 2024-04-29 10:29:19 EDT

Time Remaining: about 12 hours

Session ID: a26538a5-1a7d-41d8-b74c-093795747c2a

® Connect to RStudio Server

Accessing the RStudio through OnDemand



The screenshot shows the RStudio OnDemand web interface in a browser. The address bar displays `ondemand.hpc.rutgers.edu`. The interface includes a menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help), a toolbar with icons for file operations and running code, and a main workspace area. The workspace shows several open files: `takeHome.Rmd`, `unix.Rmd`, `Untitled1`, and `indices.R`. The `indices.R` file is active, showing a single line of code: `1`. The Environment pane on the right indicates that the environment is empty. The Files pane at the bottom right shows a directory listing for the user's home directory, including folders like `2023_03_ugandan_nasal_blood`, `Desktop`, `Documents`, `Downloads`, `for_howard`, `for_howard.zip`, and `intel`, as well as a file named `basespace_example.sh`. The Console pane at the bottom left shows the output of the `install.packages()` function, indicating that help indices, package indices, and vignettes are being installed.

Environment is empty

Name	Size	Modified
2023_03_ugandan_nasal_blood		
basespace_example.sh	87 B	Mar 27, 2023, 12:14 PM
bin		
Desktop		
Documents		
Downloads		
for_howard		
for_howard.zip	4.9 GB	Jul 26, 2023, 7:12 PM
intel		

```

*** installing help indices
** building package indices
** installing vignettes

```

Introduction to Cluster Batch Job Scheduling with SLURM

What is SLURM?

- ▶ SLURM (Simple Linux Utility for Resource Management) is an open-source job scheduler and resource manager used on many high-performance computing (HPC) clusters.
- ▶ It allows users to submit, manage, and monitor batch jobs on computing clusters efficiently.

Key Concepts

1. **Partition:** A partition is a logical grouping of compute nodes with similar hardware characteristics or usage policies.
2. **Job Script:** A job script is a text file containing instructions for the scheduler, including job parameters, resource requests, and executable commands.
3. **Resource Allocation:** Users specify the required resources for their jobs, such as CPUs, memory, and runtime, in the job script.
4. **Job Status:** Monitor the status of their submitted jobs using SLURM commands or monitoring tools.

Submitting a Job with SLURM

1. **Create a Job Script:** Write a job script using a text editor, specifying job parameters and commands.
2. **Submit the Job:** Use the `sbatch` command to submit the job script to SLURM.
3. **Monitor Job Status:** Check the status of the submitted job using commands like `squeue` or monitoring tools provided by SLURM.
4. **Retrieve Results:** Once the job completes, retrieve the results and data generated by the job from the designated output directory.

Example Job Script

```
#!/bin/bash  
#SBATCH --job-name=my_job  
#SBATCH --nodes=1  
#SBATCH --ntasks-per-node=4  
#SBATCH --mem=8G  
#SBATCH --time=1:00:00  
  
# Execute your command here  
./my_executable input_file.txt > output_file.txt
```

Best Practices

- ▶ Resource Requests: Specify accurate resource requirements to ensure optimal resource allocation and job performance.
- ▶ Job Prioritization: Understand the partition structure and prioritize jobs accordingly to optimize resource utilization.
- ▶ Job Dependencies: Use job dependencies to sequence jobs and ensure proper execution order when necessary.
- ▶ Error Handling: Include error handling and logging mechanisms in job scripts to facilitate troubleshooting and debugging.

Conclusion

SLURM is a powerful tool for managing batch job scheduling and resource allocation on cluster computing systems. By understanding key concepts and following best practices, users can effectively utilize SLURM to submit, manage, and monitor computational tasks on HPC clusters. Explore documentation and seek support from the cluster administrators or user community to learn more about SLURM and optimize your workflow.

Session info

```
sessionInfo()
```

```
## R version 4.5.1 (2025-06-13)
## Platform: aarch64-apple-darwin20
## Running under: macOS Sequoia 15.5
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRlapack.dylib;  LAPACK version 3.12.1
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## time zone: America/New_York
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## loaded via a namespace (and not attached):
## [1] compiler_4.5.1    fastmap_1.2.0     cli_3.6.5         tools_4.5.1
## [5] htmltools_0.5.8.1 rstudioapi_0.17.1 yaml_2.3.10        rmarkdown_2.29
## [9] knitr_1.50         xfun_0.52         digest_0.6.37     rlang_1.1.6
## [13] evaluate_1.0.4
```