RIT

RIT Research Computing: Theory and Applications

COGS 621

2025 - 09 - 16

Prepared by: Viet Nguyen

Part 3 – RIT Research Computing Theory

RIT Research Computing Theory

The Grant – Available Resources

- 64 nodes for mainstream jobs (4-GPU block):
  - A total of 2,304 Intel Skylake CPU cores, 24 Tb of RAM
  - A total of two 2.7 GHz Intel Xeon Gold 6150 processors (36 cores), 384 Gb RAM
  - A total of 10 Pb of storage (1 Pb ~ 1,024 Tb)
    - Each account is allowed to have 2 Tb
    - Research groups can request for shared space, starting from 1 Tb
  - Two 100 Gb/s Ethernet connection
- 2 nodes for debugging requests: 144 cores and 2.24 TB of RAM
- 2 nodes for interactive purposes: two 2.2 GHz Intel Xeon Gold 6238R processors (56 cores) with 384 GB of RAM each
- The cluster leverages batch scheduling with SLURM to enable researchers to share access to our compute resources
- GPUs:
  - 15 nodes with four A100s
  - 20 nodes with two A100s
  - 1 node with four H100s
  - 2 nodes with one GH200 each
  - 2 debug nodes (1 with two A100s)
  - 2 interactive nodes with one A100 each

RIT Research Computing Theory

Filling Forms and Accessibility

- **[Demo]** Service Requests > Information Technology Services > Research Computing > Research Computing Project Request
  - https://help.rit.edu/sp?id=sc_category&sys_id=59a1dcbb1b0ac0d07cc34377cc4bcbe7&catalog_id

Research Computing Project Request

Questionnaire to request high throughput and data-intensive computing resources from the RIT Research Computing department.

RC offerings include engineering, managing, and supporting traditional high-performance computing (HPC) systems, software stack, and services, as well as resources for high throughput and data-intensive computing. Research Computing aims to know the researchers, their research, and their computational and research computing needs.

* Indicates required

*Name of the Project

*This Project relates to a:
-- None --

*Does this project have an SRS# in RAPID?
-- None --

*Does this project have a Data Management Plan?
Yes

*Project Abstract

*Does your work have a deadline?
-- None --

*Name of PI/Faculty Advisor(s)

RIT Investigators working on this research who will be utilizing RC resources

List the collaborators working on this research with you and who will be utilizing RC resources

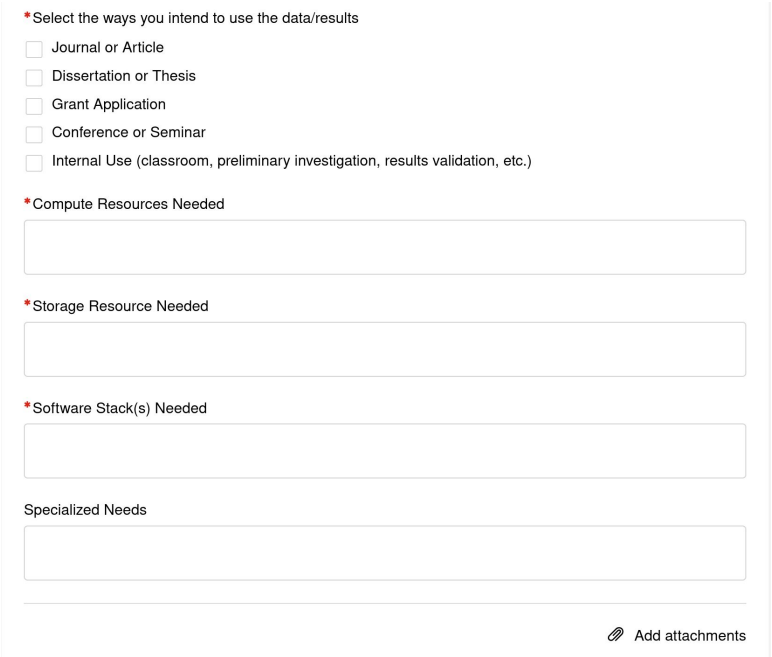
Other/not listed (or non-RIT email address)

*Does your research require a secured computer research environment for government compliance?
-- None --

RIT Research Computing Theory

Filling Forms and Accessibility

- [Demo] Service Requests > Information Technology Services > Research Computing > Research Computing Project Request



RIT Research Computing Theory

Scheduled Processing On Research Computing (SPORC)

- Submit node (sporcsubmit.rc.rit.edu)
  - A shared resource
  - 36-core 360 GB (Gigabytes) RAM machine with 1 GPU
  - Researchers are limited to at most two cores and 8 Gb of memory
  - DO NOT RUN JOBS ON THE SUBMIT NODE!
    - To debug, submit a job to `sinteractive` or `debug` queue.
- General production jobs should be submitted to `tier3` queue.
- `sinteractive` queue: use for interacting directly with the system, e.g., running programs. Maximum time is 12 hours
- `debug` queue: used for running job for debugging. Maximum time is 24 hours

RIT Research Computing Theory

Simple Linux Utility for Resource Management (SLURM) Tutorial

- **[Demo]** To tell Slurm what resources you need, you will have to create an sbatch script (also called a Slurm script)
- **[Demo]** Configs
  - Node
  - Task
  - GPU
  - Memory
  - Etc.
- **[Demo]** Debugging
- **[Demo]** Interactive jobs
- **[Demo]** SLURM commands

Practical Applications and Demonstrations

Demo Multi-Trial Job Submission Script

- [Demo]

Practical Applications and Demonstrations

Demo Sharded Data Parallel Job Submission Script

- [Demo]

Part 4 – Tips and Tricks

Tips and Tricks

RIT RC's ondemand and related tools

- **[Demo]** On demand web app (ondemand.rc.rit.edu)
  - Jobs
  - Cluster
  - Interactive apps

Tips and Tricks

How queuing system works

- `sprio`: current information about the queue
  - -t PD
  - -u <username>
  - -o "%i %u %Y %f": %i = Job ID, %u = User, %Y = Priority, %f = Fairshare factor
  - | sort -k4 -nr: sort descending
- Priority = (Fairshare * weight) + (Age * weight) + (TRES * weight) + (Partition * weight) +
  - Age: how long job has been waiting (higher is better)
  - Fairshare: your historical usage vs. others (balances fairness) (higher is better)
  - TRES (trackable resources): number of nodes/CPUs/GPUs requested (lower is better)
  - Partition: partition priority (higher is better)
- **[Demo]** RIT's current queue information

Tips and Tricks

Conda – the powerful python package management system

Conda environment

- Conda is a package and environment manager widely used in data science, ML, and scientific computing. It helps you:
  - Create isolated environments → different projects can have different Python versions and packages without conflict.
  - Install/manage packages → not just Python, also C/C++/Fortran-based libraries.
- Install:
  - Miniconda (lightweight, just Conda + Python) ← Just use this
  - Anaconda (full packages with other programs) ← This one is very heavy, most of the time you don't need it, especially on SPORC
- **[Demo]** Basic commands
  - Conda create -n <name> python=3.11
  - Conda env remove -n <name>
  - Conda activate <name>
- Best practices: use one conda environment per project

Tips and Tricks

File transfer using FileZilla and SFTP

- **[Demo]** Using FileZilla for transferring files using the SFTP protocol

Questions??

Thank you for your attention!!

Appendix: Resources

Resources

- Linux bash tutorial:
  - https://www.w3schools.com/bash/
- Regular expression theory and tester:
  - https://regex101.com/
  - https://regexr.com/
- RIT RC documentation: https://research-computing.git-pages.rit.edu/docs/index.html
- Github repository for this session: https://github.com/rxng8/hpc-workshop
- RIT RC Service Request: https://help.rit.edu/sp?id=sc_category&sys_id=59a1dcbb1b0ac0d07cc34377cc4bcbe7&catalog_id=e0d08b13c3330100c8b837659bba8fb4
- VSCode: https://code.visualstudio.com/download
- FileZilla: https://filezilla-project.org/