



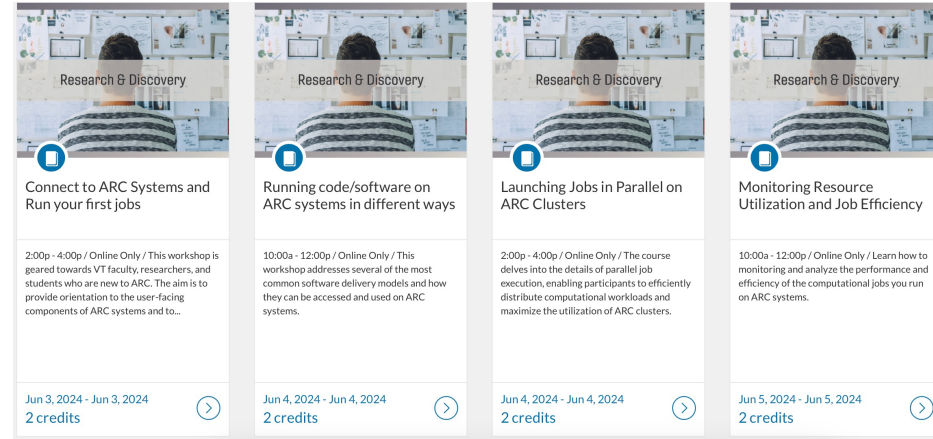
Connect to ARC systems and run your first jobs



Ayat Mohammed
Computational Scientist
Advanced Research Computing

ARC Course Offerings

- Who: VT faculty, researchers, and students who are new to ARC.
- What:
 - Connect via Open OnDemand
 - Connect via SSH
 - Cluster and Scheduler Orientation
 - Run Demo Jobs



Course Title	Time / Format	Duration	Credits
Connect to ARC Systems and Run your first jobs	2:00p - 4:00p / Online Only	Jun 3, 2024 - Jun 3, 2024	2 credits
Running code/software on ARC systems in different ways	10:00a - 12:00p / Online Only	Jun 4, 2024 - Jun 4, 2024	2 credits
Launching Jobs in Parallel on ARC Clusters	2:00p - 4:00p / Online Only	Jun 4, 2024 - Jun 4, 2024	2 credits
Monitoring Resource Utilization and Job Efficiency	10:00a - 12:00p / Online Only	Jun 5, 2024 - Jun 5, 2024	2 credits

Sign IN:

https://docs.google.com/document/d/1UC_1ifkOC4YWJ293nvs9C3CsXGII_NPBF8v9kMaJlkxU/edit?usp=share_link

<https://profdev.tlos.vt.edu/?query=ARC>

Expectations

This is an informal workshop

Mostly informational about Advanced research computing at VT

I want to hear your questions

Welcome to use chat to ask questions + some time at the end

Feedback needed to help improve future workshops

- One up / one down at the end

First Thing's First

VPN needed for connections from off-campus

- <https://www.nis.vt.edu/ServicePortfolio/Network/RemoteAccess-VPN.html>
- Nearly all ARC services require being on the campus network or VPN
- Use "VT Traffic over SSL VPN" connection
- ColdFront (accounting system) available with or without VPN

Get an ARC account:

- <https://coldfront.arc.vt.edu/account/create>
- Acceptable Use Policy

Getting Started

Needs Assessment

- Compute
- Storage
- Software
- Collaboration
- Visualization
- Lifecycle and data retention

Get an account

<https://arc.vt.edu/account>

- Get account for log-in

Register a Project and Get Allocations

<https://coldfront.arc.vt.edu>

- Create a “project”, add people, grants/pubs
- Request allocation for Compute to run jobs
- Request allocation for Project storage if desired

Where to get help

Website (<https://docs.arc.vt.edu>)

- FAQs
- Video demos
- Detailed instructions
- Examples

<https://github.com/AdvancedResearchComputing/examples>

Helpdesk (<https://arc.vt.edu/help>)

Office Hours (<https://arc.vt.edu/office-hours>)

Ask for consultation

- Workflow design
- Optimization
- Sponsored Projects

HPC Resources @ ARC

Cluster	Description	Since
CUI	Dense GPU + some CPU for projects with controlled data/software	c. 2021
Tinkercliffs	HPC/HTC Flagship CPU HPE Dense GPU nodes (A100) DGX Dense GPU nodes (A100)	c. 2020 c. 2021 c. 2022
Infer (nearing end of life)	Accelerating inference and ML workloads (T4 GPU) Added P100 GPUs from Newriver Added V100 GPUs from Cascades	c. 2021 c. 2016 (EOL) c. 2018 (EOL)
OWL (coming soon)	Water-cooled latest generation AMD CPU high mem-per-core DDR5	c. 2024
Falcon (later in 2024)	GPU node expansion L40S GPUs (20 nodes x4 GPUs) A30 GPUs (32 nodes x4 GPUs)	c. 2024

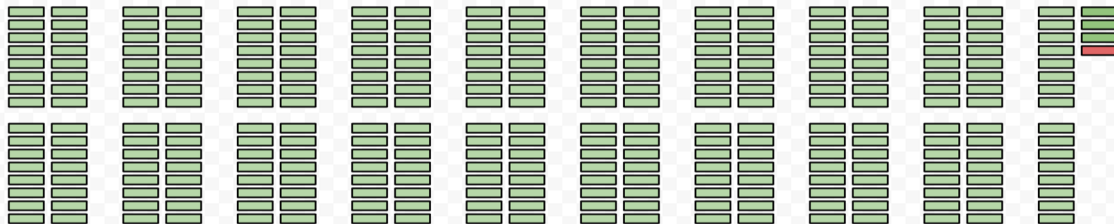
TinkerCliffs - Flagship CPU Cluster

316 Nodes w/ 128 cores(AMD EPYC Rome)
 16 Nodes w/ 96 cores (Intel Cascade Lake-AP)
 41,984 CPU cores

tc-hm[001-008]
 largemem_q



tc[001-308]
 dev_q, preemptable_q
 tc[001-307]
 normal_q
 tc[001-302]
 interactive_q
 tc308



tc-intel[001-016]



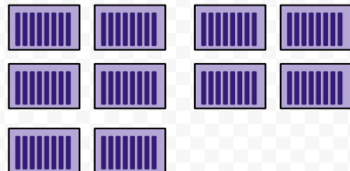
w/ dense GPU

ai[001-04]
 a100_normal_q



4 Nodes w/ 128 cores (AMD Epyc Rome 7742)
 + 8 NVIDIA A100-80GB GPUs (6912 CUDA)
 512 CPU cores
 32 GPU accelerators
 221,184 CUDA cores

ai[001-04]
 a100_normal_q



10 Nodes w/ 128 cores (AMD Epyc Rome 7742)
 + 8 NVIDIA A100-80GB GPUs (6912 CUDA)
 1280 CPU cores
 80 GPU accelerators
 552,960 CUDA cores

Web Access to Clusters

Open OnDemand

- Connect to VPN first
- Manage data: graphical interface to browse, transfer, view, and edit files
- Access the shell command-line interface (CLI) on ARC
- Interactive jobs, especially applications with graphical user interfaces (Matlab, Rstudio, Jupyter Notebooks, ...)

<https://ood.arc.vt.edu>

Overview of SSH/CLI Connection

-- *Must first be on campus network or VPN*

Linux or MacOS

- Use built-in "terminal" application to get CLI with standard tools

Access Windows SSH Client and SSH Server

- Recent versions include a built-in SSH server and client that are based on OpenSSH
- Git/BASH (Git for Windows with a shell emulator)
- Windows Subsystem for Linux (WSL), Microsoft VS-Code

Windows: pick and install an SSH client

Windows built-in SSH client

- <https://learn.microsoft.com/en-us/windows/terminal/tutorials/ssh>
- “The latest builds of Windows 10 and Windows 11 include a built-in SSH server and client that are based on OpenSSH”
- “You can ... check that it is installed in Windows Settings > Apps > Optional features, then search for "OpenSSH" in your installed features.”

Windows: pick and install an SSH client

-- Free, open-source software. Includes Git for source control --

Git/BASH from Git for Windows

- Lightweight unix shell and tools
- <https://gitforwindows.org>
- Lots of questions to answer during install, but most relate only to the "Git" configuration.
- Default working directory is your Windows user directory: `/c/Users/Myaccount`

Git is the current standard tooling for version control and collaborative development, so this package has added value.

Also used by Software Carpentries



Windows: pick and install an SSH client

Windows Subsystem for Linux

- Essentially a Linux virtual machine running inside Windows
- Bigger download (~250MB)
- Highest level of features, options, etc. A more native Linux experience.

Installation:

- Settings > Apps > Programs and Features > Turn Windows features on or off > Windows Subsystem for Linux OR (latest versions) "wsl --install"
- Reboot
- Install Linux distribution using Microsoft Store (Ubuntu)
- Your Windows files are available in the VM at /mnt/c

<https://docs.microsoft.com/en-us/windows/wsl/install>

Connect via SSH

-- must be on campus network or connect to VPN

In a terminal window on your local machine (e.g., laptop) type:

```
"ssh myusername@tinkercliffs2.arc.vt.edu"
```

- On initial connection, type "yes" to accept the authenticity of the host and continue connecting
- Password, then DUO for 2-factor authentication

SLURM Intro

- SLURM = scheduler and cluster resource manager
- Request resources
 - **salloc** – get an allocation and wait for further commands interactively
 - **sbatch** – submit job script for non-interactive execution
 - **srun** – run command in allocation if the context already exists OR request an allocation and run the command
- Queued until resources are available (enhanced FIFO)
- Partial nodes can be allocated, ie. nodes can be shared by multiple jobs. Each job is contained to its allocation with linux cgroups

Commands: "--help", "man sbatch", "srun hostname", squeue

Anatomy of a Batch Script

```
#!/bin/bash
#SBATCH -J hello-world
#SBATCH --account=arcadm
#SBATCH --partition=normal_q #-P
#SBATCH --nodes=1 # -N
#SBATCH --ntasks-per-node=1
#SBATCH --cpus-per-task=1 # this requests 1 node, 1 core.
#SBATCH --time=0-00:10:00 # 10 minutes
##SBATCH --gres=gpu:1
echo "hello world from..."; hostname
echo "Working directory is: "; pwd
echo "Changing directory to $SLURM_SUBMIT_DIR"
cd $SLURM_SUBMIT_DIR
module reset
```

Commands:
"sbatch <scriptname>",
"man sbatch",
scancel <jobid>

Accounting

- Tinkercliffs
 - monthly limit per PI on usage (600,000 units/month)
 - PI may own multiple accounts, but limit will be imposed on aggregate usage
 - enables cost center meeting federal requirements
- Billing weights reflect node/component costs
- Also have policy limits per [user,job,account]

Commands: quota, scontrol show part, squeue, showusage, tcgetusage <accountname>, showqos

<https://coldfront.arc.vt.edu>, <https://www.arc.vt.edu>

Support, Consultation and Collaboration

ARC Helpdesk: <https://arc.vt.edu/help>

<https://4help.vt.edu> (mention ARC, provide details)

ARC Helpdesk GRAs work as a team to handle most incoming questions/problems.

"How do I setup SSH keys for authentication?" *"What can I do to get my job to launch faster?"* *"Why did my job stop?"*

"Is MATLAB available on Huckleberry?" *"How can I share my files with my collaborator?"*

Escalate to ARC Computational Scientists as needed.

Office Hours (<https://arc.vt.edu/office-hours>)

Feedback: How can we best advertise?

How can we better get the word out about our presentations and our capabilities?

- https://docs.google.com/document/d/1f39CKojG64M9CDguBM3UdcwiXBOVGJdNVniVe0CcPd4/edit?usp=share_link

Thanks for attending and listening!

- ARC Website: www.arc.vt.edu
- My contact info: Ayat Mohammed
maaayat@vt.edu