16 March 2021

Scaling Up Research Computing: ARC Systems and Services

Matthew Brown
Advanced Research Computing
Division of Information Technology
Virginia Tech

VIRGINIA TECH

Some Common Scenarios

"Our team just completed the first run of our analysis and found that it took four hours to run on a laptop. The results are perfect, but we have 8,500 more of these to run and need finish in a few months. Even if we all had our laptops running 24x7, it would take more than a year!"

"I have a 80GB data set that I need to process using a colleague's program. I have done this with 3GB data sets in the past, but my computer crashes when I try to process the larger data set. I think I need more memory."

"I want to try out using this neural network to see if it provide insights into my problem. But training it on my data is taking weeks."

Overview

- What is ARC?
- What does ARC have?
- What services does ARC provide?
- How to get started

What is ARC?

VIRGINIA TECH

ARC's Mission

Advanced Research Computing (ARC) provides centralized support for research computing by building, operating and promoting the use of advanced cyberinfrastructure at Virginia Tech.

ARC delivers a comprehensive ecosystem consisting of advanced computational systems, large-scale data storage, visualization facilities, software, and consulting services.

ARC provides education and outreach services through conferences, seminars, and scientific computing courses.

ARC seeks to help maximize research productivity at Virginia Tech through interdisciplinary collaborations that connect researchers to new opportunities in computing and data driven research as they occur.

By fostering strategic partnerships with the public and private sector, ARC serves to cultivate an entrepreneurial spirit around advanced computing infrastructure as a platform for collaboration and helps secure the position of Virginia Tech as a leader in education and research.

VIRGINIA TECH

ARC's Mission

Advanced Research Computing (ARC) provides centralized support for research computing by building, operating and promoting the use of advanced cyberinfrastructure at Virginia Tech.

ARC delivers a comprehensive ecosystem consisting of advanced computational systems, large-scale data storage, visualization facilities, software, and consulting services.

ARC provides education and outreach services through conferences, seminars, and scientific computing courses.

ARC seeks to help maximize research productivity at Virginia Tech through interdisciplinary collaborations that connect researchers to new opportunities in computing and data driven research as they occur.

By fostering strategic partnerships with the public and private sector, ARC serves to cultivate an entrepreneurial spirit around advanced computing infrastructure as a platform for collaboration and helps secure the position of Virginia Tech as a leader in education and research.

VIRGINIA TECH

ARC's Mission

Advanced Research Computing (ARC) provides centralized support for research computing by building, operating and promoting the use of advanced cyberinfrastructure at Virginia Tech.

ARC delivers a comprehensive ecosystem consisting of advanced computational systems, large-scale data storage, visualization facilities, software, and consulting services.

ARC provides education and outreach services through conferences, seminars, and scientific computing courses.

ARC seeks to help maximize research productivity at Virginia Tech through interdisciplinary collaborations that connect researchers to new opportunities in computing and data driven research as they occur.

By fostering strategic partnerships with the public and private sector, ARC serves to cultivate an entrepreneurial spirit around advanced computing infrastructure as a platform for collaboration and helps secure the position of Virginia Tech as a leader in education and research.

VIRGINIA TECH

ARC's Mission

Advanced Research Computing (ARC) provides centralized support for research computing by building, operating and promoting the use of advanced cyberinfrastructure at Virginia Tech.

ARC delivers a comprehensive ecosystem consisting of advanced computational systems, large-scale data storage, visualization facilities, software, and consulting services.

ARC provides education and outreach services through conferences, seminars, and scientific computing courses.

ARC seeks to help maximize research productivity at Virginia Tech through interdisciplinary collaborations that connect researchers to new opportunities in computing and data driven research as they occur.

By fostering strategic partnerships with the public and private sector, ARC serves to cultivate an entrepreneurial spirit around advanced computing infrastructure as a platform for collaboration and helps secure the position of Virginia Tech as a leader in education and research.

VIRGINIA TECH

ARC's Mission

Advanced Research Computing (ARC) provides centralized support for research computing by building, operating and promoting the use of advanced cyberinfrastructure at Virginia Tech.

ARC delivers a comprehensive ecosystem consisting of advanced computational systems, large-scale data storage, visualization facilities, software, and consulting services.

ARC provides education and outreach services through conferences, seminars, and scientific computing courses.

ARC seeks to help maximize research productivity at Virginia Tech through interdisciplinary collaborations that connect researchers to new opportunities in computing and data driven research as they occur.

By fostering strategic partnerships with the public and private sector, ARC serves to cultivate an entrepreneurial spirit around advanced computing infrastructure as a platform for collaboration and helps secure the position of Virginia Tech as a leader in education and research.

VIRGINIA TECH

ARC's Mission

Advanced Research Computing (ARC) provides centralized support for research computing by building, operating and promoting the use of advanced cyberinfrastructure at Virginia Tech.

ARC delivers a comprehensive ecosystem consisting of advanced computational systems, large-scale data storage, visualization facilities, software, and consulting services.

ARC provides education and outreach services through conferences, seminars, and scientific computing courses.

ARC seeks to help maximize research productivity at Virginia Tech through interdisciplinary collaborations that connect researchers to new opportunities in computing and data driven research as they occur.

By fostering strategic partnerships with the public and private sector, ARC serves to cultivate an entrepreneurial spirit around advanced computing infrastructure as a platform for collaboration and helps secure the position of Virginia Tech as a leader in education and research.

VIRGINIA TECH

ARC's Mission

In a nutshell our goals are to:

- Engage with researchers across all colleges, centers, and disciplines at Virginia Tech

… and to …

- Advance the research programs of Virginia Tech faculty and students by providing the Research Computing systems (hardware, software, integration) and support (helpdesk, consultation, collaboration) that will help them excel.

VIRGINIA TECH

Who We Are

Associate VP for Research Computing:	Terry Herdman
Assist. Director, Development and Fiscal Admin:	Alana Romanella
Network Research Manager:	Mark K. Gardner
Director, Visualization:	Nicholas Polys
Visualization and Virtual Reality Systems Specialist:	Lance Arsenault
Computational Scientists/Software Engineers:	Justin Krometis, James McClure, Bob Settlage, Matt Brown, Nathan Liles
Director of ARC Operations:	Kevin Shinpaugh
Systems Engineers/HPC System Administrators:	Miles Gentry, Jeremy Johnson, Doug McMaster, William Marmagas, Jessie Bowman, Ben Sandbrook

Plus our student interns and Helpdesk GRAs!

VIRGINIA TECH

Division of IT Inter-Connections

- ARC systems are connected to and integrated with VT infrastructure and services
  - Networks and data centers
  - Authentication, directory services, 4Help, hosting
  - Logging, monitoring, alerting
- Deliver courses/workshops via TLOS's Prof. Development Network
- Direct or indirectly make use of wide range of DoIT systems/services

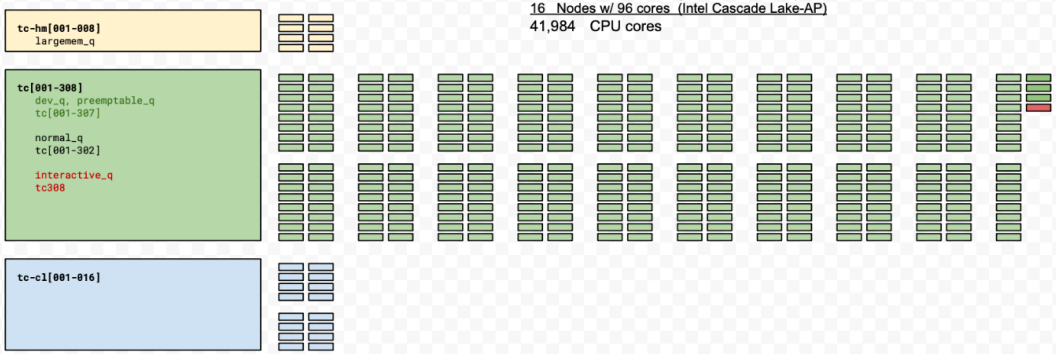
VIRGINIA TECH

ARC Systems

VIRGINIA TECH

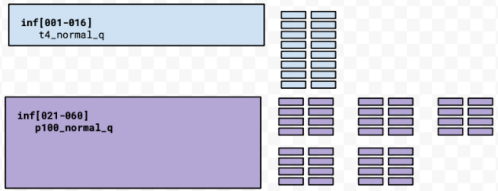
TinkerCliffs - Flagship CPU Cluster

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



Infer - Accelerating ML/DL and Inference

16 Nodes w/ 32 cores (Intel Skylake) + 1 NVIDIA T4 GPU (2560 CUDA + 320 tensor cores)
40 Nodes w/ 28 cores (Intel Broadwell) + 2 NVIDIA P100 GPUs (3580 CUDA cores)
1,632 CPU cores
96 GPU accelerators
184,160 CUDA cores
5,120 Tensor cores

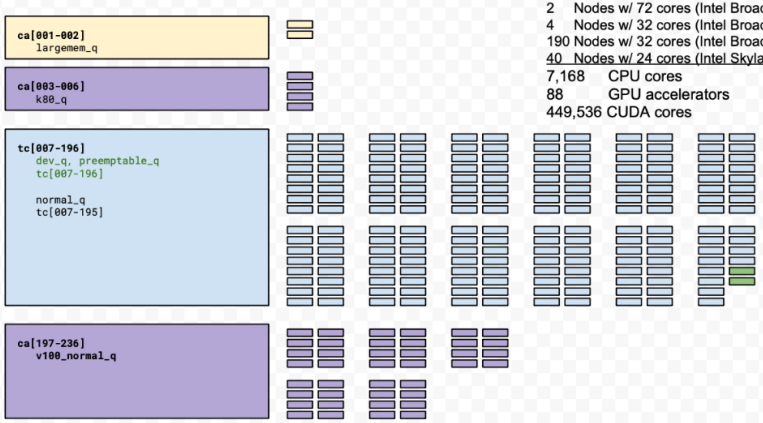


Coming 2021: Dense GPU for AI

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

VIRGINIA TECH

Cascades - Heterogeneous HPC + GPU



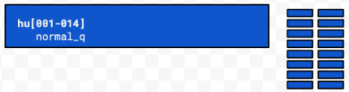
2 Nodes w/ 72 cores (Intel Broadwell)
4 Nodes w/ 32 cores (Intel Broadwell) + 2 NVIDIA K80 GPUs (4,992 CUDA cores)
190 Nodes w/ 32 cores (Intel Broadwell)
40 Nodes w/ 24 cores (Intel Skylake) + 2 NVIDIA V100 GPUs (5,120 CUDA cores, 640 tensor cores)
7,168 CPU cores
88 GPU accelerators
449,536 CUDA cores

Dragonstooth - HTC

48 Nodes w/ 24 cores (Intel Haswell)
1,152 CPU cores

Scheduler permits very long jobs (30 days)
Scheduler permits large volumes of small jobs

Huckleberry - Power for Deep Learning



14 Nodes w/ 120 cores (IBM Power8) and 4 NVIDIA P100 GPUs w/ NVLINK
1,680 CPU cores
56 GPU accelerators
200,704 CUDA cores

VIRGINIA TECH

Systems

Aggregated resources:

690	Compute nodes
54,128	CPU cores
272	GPU accelerators
1,055,584	NVIDIA Cuda cores

+ high speed Ethernet and low-latency Infiniband interconnecting networks
+ large scale and high-performance parallel storage (~10PB total)
+ various boutique/custom architecture systems

VIRGINIA TECH

Systems

Aggregated resources:

690	Compute nodes
54,128	CPU cores
272	GPU accelerators
1,055,584	NVIDIA Cuda cores

ARC has diverse architectures and capacity to meet the demands of a wide variety of projects and scale them up.

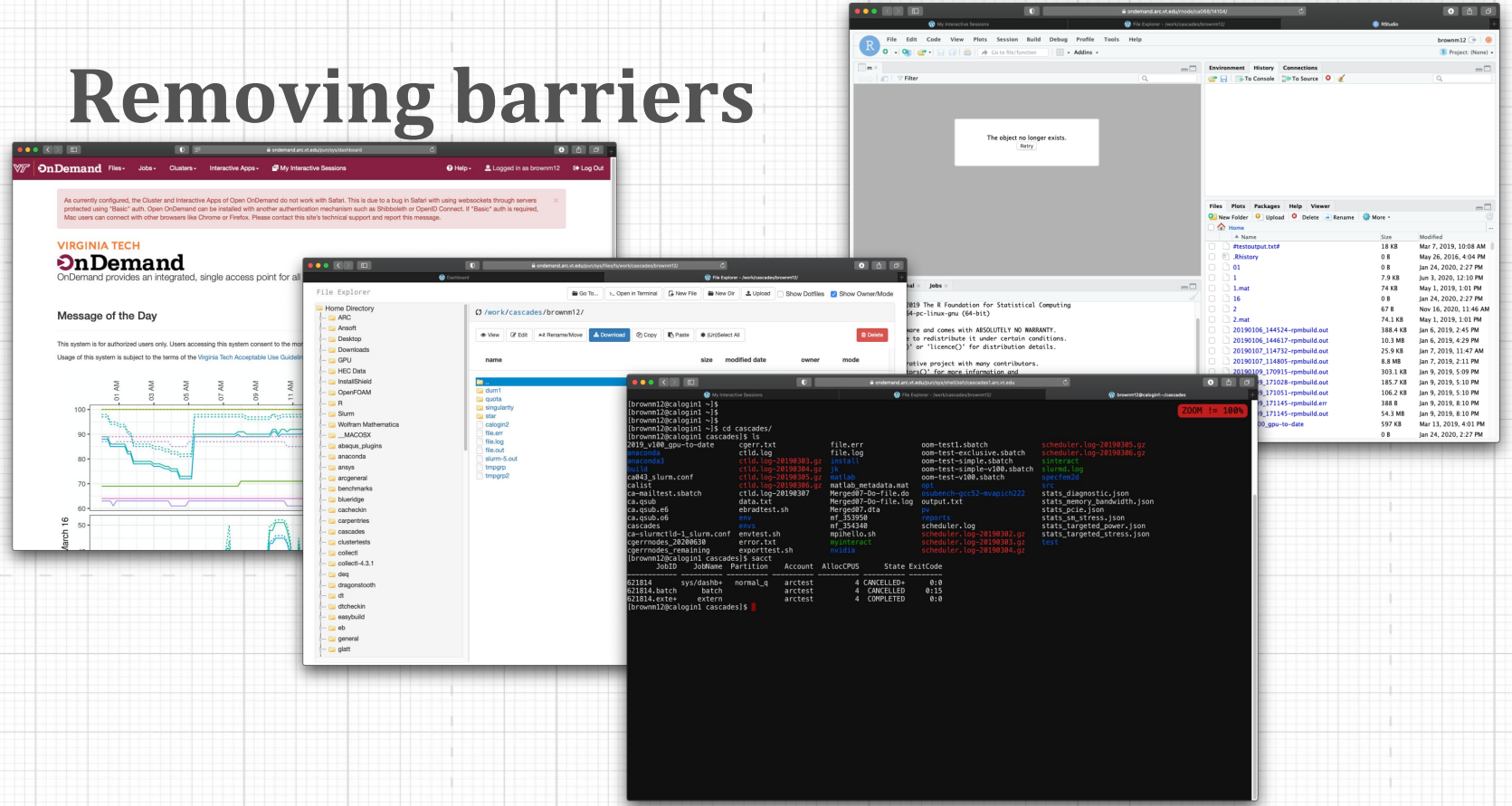
VIRGINIA TECH

Industry standard usage model

- Linux clusters, SLURM scheduler, Infiniband interconnect
- Connect to “login node” using SSH client, upload/download files
- Compose job script, submit to scheduler, job runs in batch mode on compute nodes

This is still a very productive model and the dominant mode of usage for many researchers, but has presented barriers to entry for others.

Removing barriers



VIRGINIA TECH

https://ondemand.arc.vt.edu

ARC Services

VIRGINIA TECH

Use of Systems

Free Tier (Subsidized by VT/DoIT)

- Unlimited use (subject to availability) of older systems
- Generous monthly compute allowances on newer system
- 1TB storage per person, 25TB per principle investigator
- Support via ARC Helpdesk, consultation with Computational Scientists
- Archival data storage

ARC Cost Center

- Compute capacity and priority beyond the free tier
- More storage capacity available

https://arc.vt.edu/arc - investment - computing - and - cost - center/

VIRGINIA TECH

Consulting and Collaboration

ARC Computational Scientists

- Understand the Applications
- Provide Research Domain Expertise
- Offer Classes, Short Courses and Workshops
- Optimization of Codes

ARC Helpdesk (via 4Help)

- Getting started with ARC
- Basic questions and troubleshooting

VIRGINIA TECH

Visualization

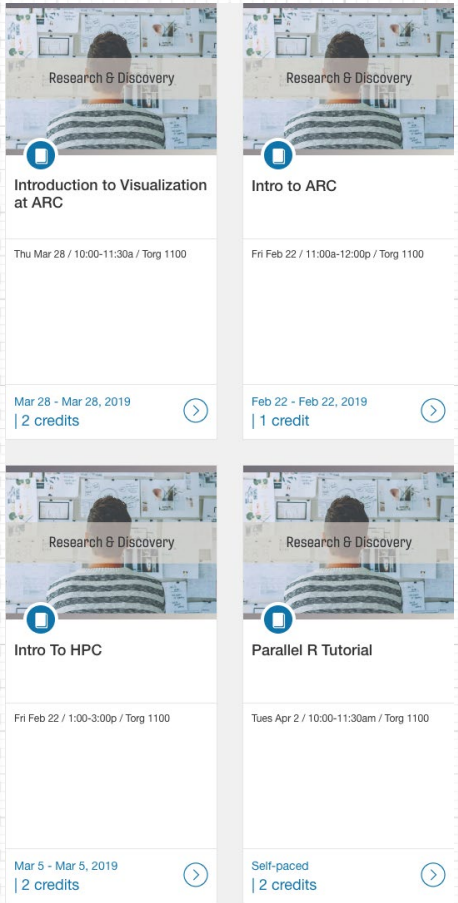
- Desktop Visualization
- HyperCube in the Visionarium Lab
- User support and consulting
- Research collaboration
- Trainings and classes
- Tours and field trips

VIRGINIA TECH

ARC Course Offerings

- ARC Personnel offer or guest lecture in regular courses (Math, CS, AOE, Statistics, CMDA)

- ARC Personnel offer short courses and workshops
  - Introduction to ARC systems
  - Introduction to High Performance Computing
  - Deep learning with NVIDIA Digits
  - Python for scientific computing
  - Parallel R
  - Numerical computing in Julia
  - Visual computing
  - Virtual Reality



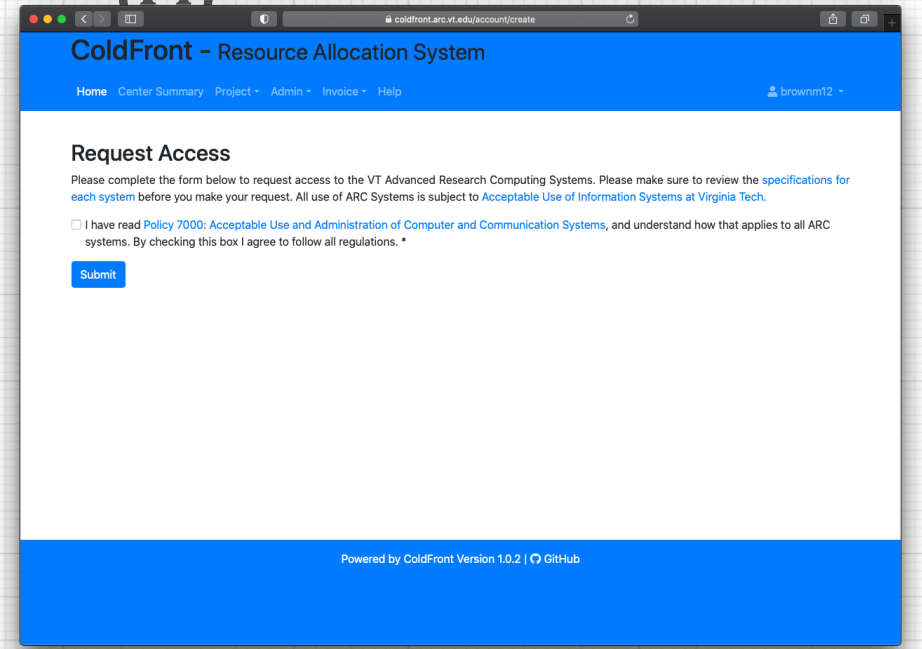
VIRGINIA TECH

https://profdev.tlos.vt.edu

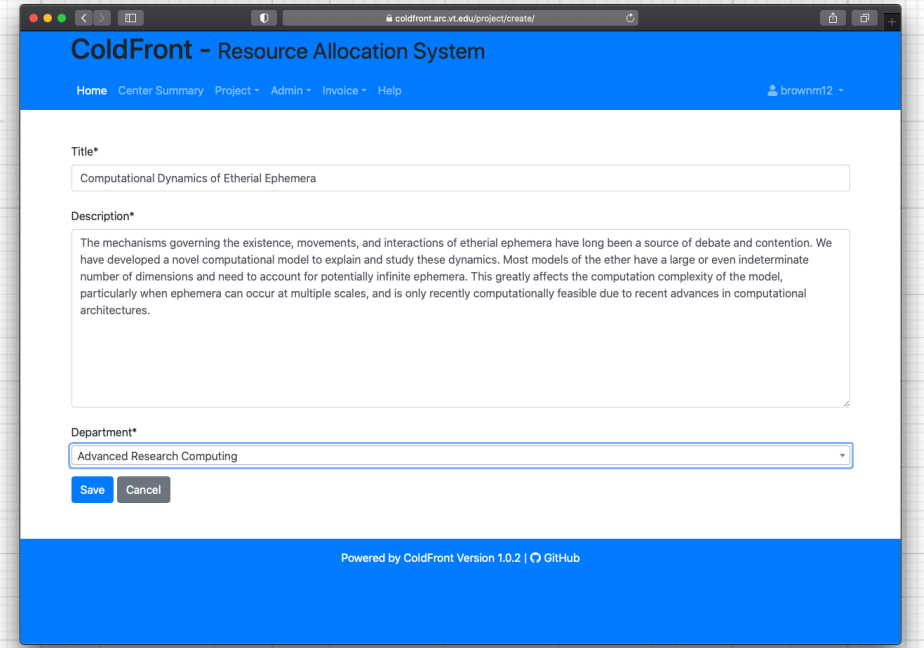
Getting Started

VIRGINIA TECH

Getting Started: Create an account (all) & project (PI)



https://coldfront.arc.vt.edu/account/create



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

VIRGINIA TECH

References

- ARC Website resources: https://arc.vt.edu/
- ARC has video tutorials
- OnDemand web interface for ARC systems https://ondemand.arc.vt.edu
- ColdFront - Account creation and management https://coldfront.arc.vt.edu
- ARC Training opportunities via TLOS PDN https://profdev.tlos.vt.edu
- General questions or help with ARC systems: https://arc.vt.edu/help
- Me (Matthew Brown): brownm12@vt.edu

VIRGINIA TECH