

Facilities, equipment, and other resources

Expertise

Research Computing at CU Boulder consists of a group of computational scientists, high-performance computing specialists, and system administrators with the mission to provide leadership in developing, deploying, and operating an integrated cyberinfrastructure. This cyberinfrastructure consists of on-premise high-performance computing, commercial cloud computing, large-scale storage, and high-speed networking that supports research, collaboration and discovery. Research Computing contributes to the educational mission of the university by providing training workshops and consultation services for cyberinfrastructure related topics.

Compute

- *Alpine*: Research Computing hosts the Alpine supercomputer, which is available at no cost to students, staff and researchers at CU Boulder and partner institutions. Alpine is funded by contributions from the University of Colorado Boulder, Colorado State University, the University of Colorado Anschutz, and the National Science Foundation. Alpine has 390 standard CPU-based compute nodes, 24 high-memory nodes, 23 compute nodes with multiple GPUs on each (NVIDIA L40, NVIDIA A100, and AMD MI100), and 2 NVIDIA Grace-Hopper nodes. Nodes are connected through a high-performance network based on Mellanox Infiniband with a bandwidth of 200 Gb/s. A 2 PB high-performance IBM GPFS file system is provided. Additional hardware phases are added to Alpine each year.
- *Blanca*: Research Computing offers a condo computing service that enables researchers to purchase and own compute nodes that are operated as part of a cluster, named “Blanca.” The aggregate cluster is made available to all condo partners while maintaining priority for the owner of each node. Blanca has 242 nodes.
- *Cloud Computing*: Research Computing facilitates access to and provides support for offsite commercial cloud resources at negotiated rates.
- *Secure computing*: Research Computing manages *The Preserve*, an environment available to researchers who engage in projects that contain confidential or highly confidential data. It is highly controlled to specifically protect research that must eventually meet Cybersecurity Maturity Model Certification (CMMC) 2.0 Level 2, Controlled Unclassified Information (CUI) and certain export control regulations.

Networking

The current CU Boulder network is a 40 Gbps fiber core with Cat 5 or higher wiring throughout campus. Research Computing has created an 80 Gbps Science-DMZ to

connect the RMACC Summit supercomputer to storage and to bring individual dedicated 10 Gbps circuits to various locations as needed. CU Boulder participates in the Internet 2 higher education, government, and vendor research computing consortium, and is an active member of the Front-Range gigapop and other networks. Research Computing has begun to provide campus researchers with a leading-edge network that meets their needs and facilitates collaboration, high performance data exchange, access to co-location facilities, remote mounts to storage, and real-time communications.

File Transfer

For moving large volumes of data Research Computing has four nodes dedicated to file transfer via multiple protocols including scp, sftp, rsync, rclone, and Globus (GridFTP). Research Computing also supports desktop-based file transfer protocols including SMB and sshfs.

Storage

Each Research Computing user has a 2 GB home directory and a 250 GB projects directory, each of which are backed up regularly. Each user also has a 10 TB scratch directory that is hosted on a high-performance IBM GPFS filesystem.

PetaLibrary Storage Services

The PetaLibrary is a CU Research Computing service supporting the storage, archival, and sharing of research data. It is available at a subsidized cost to any researcher affiliated with the University of Colorado Boulder. The two main categories of service offered to customers of the PetaLibrary are Active storage for data requiring frequent access, and Archive storage for data that is accessed infrequently. Both tiers are hosted on spinning disk enclosures on the CU Boulder Campus. The cost for CU system members is \$45 TB/year for Active and \$25/TB/year for Archive. Redundant copies of data in non-colocated data centers may be purchased for an additional cost.

Through collaboration with the CU Libraries, the PetaLibrary can also host the publication and long-term archival of large datasets. The datasets are assigned unique digital object identifiers (DOIs) that are searchable and accessible via the “CU Scholar” institutional repository.

OnDemand

Open OnDemand is a browser-based, integrated, single access portal for CURC high performance computing (HPC) resources. It provides a graphical interface to view, edit, download, and upload files, manage and create job templates for CURC’s computing clusters, and access CURC interactive applications (visualization nodes, Matlab, VSCode, Rstudio, and Jupyter Notebooks), all via a web browser.

Center for Research Data and Digital Scholarship (CRDDS)

The Center for Research Data & Digital Scholarship (CRDDS) is a collaboration between Research Computing and University Libraries, offering a full range of data services for both university and community members. The aim of CRDDS is to provide support to community members in areas related to data intensive research. CRDDS fulfills this mission by providing education and support on such issues as data discovery, reuse, access, publication, storage, visualization, curation, cleaning, and preservation, as well as digital scholarship. CRDDS is located in Norlin Library on Main Campus at CU Boulder.

CRDDS offers many opportunities to students working with data. The expert staff work hand-in-hand with researchers via weekly office hours, one-on-one consultations, and group trainings in programming, data visualization and more. CRDDS serves as a resource for data management, manipulation and publication for trainees working through undergraduate and graduate coursework.

- Examples of workshops/trainings CRDDS has offered include:
- High performance computing
- Programming in R
- Programming in Python
- Software Containerization
- Data mining

Last updated: 2025-06-11