FACILITIES AND OTHER RESOURCES
University of Idaho



IIDS Research Computing and Data Services

Research Computing and Data Services (RCDS) is the computational backbone for research at the University of Idaho. It provides investigators with access to reliable, state-of-the-art high-performance computing (HPC), data storage, and AI infrastructure for use in analyzing and managing large volumes of multidisciplinary research data. RCDS provides the expertise and computational tools required for processing data across all stages of the scientific data lifecycle, including raw data acquisition, modeling and analysis, sharing, dissemination, and archival. RCDS technical staff include systems administrators, scientific programmers, web developers, and data managers that collaborate with researchers to transform scientific questions into impactful results. Users may run jobs that use hundreds of processors or require large amounts of memory or specialized hardware and can take weeks to complete. Typical high-end projects include mathematical/computational modeling, machine learning, phylogenetic analyses, interactive data dashboards and visualizations, genome assembly, protein structure modeling, and physics simulations.

Overview of Facilities

RCDS infrastructure is explicitly designed to manage the complex computational and storage requirements for UI researchers with very high performance and reliability. RCDS contains an advanced mix of high-performance computing clusters, powerful servers, and reliable data storage components as well as the knowledge and technical skills required to compress years of analysis into days. Funding for RCDS infrastructure was provided from a variety of sources, including the National Institutes of Health (NIH), the National Science Foundation (NSF), the U.S. Department of Agriculture (USDA), the U.S. Department of Energy (DOE), and the M.J. Murdock Charitable Trust. RCDS equipment is housed in a modern data center provided by the University of Idaho, a second campus data center in the basement of the UI Library, and in the Collaborative Computing Center (C3) data center at the Idaho National Laboratory (INL).

Data Centers

Last update: Oct 2026

McClure Data Center: The primary IIDS Research Computing and Data Services data center is a redesigned and renovated 1400 square-foot facility in Room 124 in McClure Hall on the University of Idaho campus. Optical fiber and copper interconnections provide up to 25 Gb/s data transfer rates within the data center, which is redundantly connected at 100Gb/s to the university backbone and from there to the Idaho Regional Optical Network (IRON). The McClure data center has a dedicated 80KVa UPS with three-phase power and backup diesel generator. There are four-forced air handlers attached to redundant university chilled water systems.

UI Library Data Center: The secondary RCDS data center resides within the basement of the UI Library. This facility is also used as the primary production data center for the campus Office of Information Technology (OIT) unit and provides the most direct uplink from the UI campus to IRON. RCDS currently maintains several racks within this data center which houses additional RCDS data storage, virtualization, and backup infrastructure. RCDS maintains in-rack UPS units here, while an attached diesel generator provides indefinite emergency backup power.

INL Collaborative Computing Center (C3): RCDS also manages server and storage infrastructure within the C3 facility in Idaho Falls. The C3 is a 64,000 sq. ft state-of-the-art facility with a 200-person occupancy and a large 7,000 sq. ft data center capable of supporting up to 200 water-cooled server racks with a maximum power load of 8.5MW. RCDS manages 5 equipment racks within the C3, housing servers, storage, and network hardware. The C3 currently hosts several large DOE supercomputers: Teton, Wind River, Sawtooth, Bitterroot, and Lemhi. The Lemhi supercomputer hosted at C3 are jointly operated by the three public Idaho Universities.

All three RCDS data centers have rigorous physical security and access controls. RCDS facility staff have office space in room 123 McClure Hall, 441D Life Sciences South, and room 416 in the UI Library.

Computing Systems

Statewide Supercomputing

RCDS, in collaboration with Boise State University and Idaho State University, now operate the Lemhi supercomputer. Lemhi has 20,160 cores and 94TB of overall system memory. 5 specialized GPU nodes are attached to this shared HPC infrastructure for accelerating various ML and AI workloads. These systems share a 1.1 PB Ceph file system used for storage and scratch space.

On-Premises HPC

RCDS manages one large computer cluster on the Moscow campus for research and data analysis and modeling. Our main cluster provides over 2,500 processor cores and over 8 terabytes (TB) of system memory. The servers that comprise the cluster are interconnected with 40Gb/sec QDR (Quad Data Rate) Infiniband for inter-node communication and 1Gb/sec ethernet

for management. This is a heterogeneous cluster with many specialized compute nodes with high RAM (up to 2TB). Components include Dell M1000e blade enclosures with various blade servers, Dell rack servers, and various Supermicro servers. We have 25 cluster nodes with various NVIDIA GPU accelerators ranging from T4 to A100s.

RCDS also maintains 12 servers (various Dell and Supermicro rack servers) that are not connected to the cluster systems for jobs that require very large, shared memory machines (such as distance-based phylogenetic analyses, genome assembly, and molecular simulations), for software development, and for investigators who are unfamiliar with or do not require a cluster environment. The most powerful servers in this group each contain 64 cores and 1 TB (1000GB) of system memory. These powerful servers are used heavily for applications such as hybrid sequence assembly of genomic data.

Virtualization and Hosting

RCDS manages a rich virtualization environment for hosting 100+ virtual machines (VMs) dedicated to specific applications or research projects. These VMs run web servers, applications, databases, GIS services, models-as-a-service (MaaS), custom web Application Programming Interfaces (APIs), and more. We use VMWare ESXi and OpenNebula hypervisors. The main VMWare environment is run on a Dell VRTX converged chassis with 4x M640 Dell PowerEdge blade servers, each with 1TB of RAM and 28/56 physical/hyperthreaded Xeon cores. The secondary OpenNebula VM environment uses 9 M630 blade servers each with 768 GB RAM and 32 cores.

Generative AI Inference

In addition to the dozens of GPU nodes within our HPC environment, RCDS operates a dedicated generative AI inference cluster for University of Idaho research, teaching, and operations. The cluster comprises 15 GPU servers equipped with 54 NVIDIA GPUs and roughly 7 TB of combined GPU memory. This includes 8 NVIDIA B300 (Blackwell Ultra) GPUs, providing about 2.2 TB of GPU memory in a single server for frontier-scale open-weight models; 28 NVIDIA H200 (141 GB each) GPUs spread across seven 4-GPU servers, five of which are housed at the INL Collaborative Computing Center (C3) in Idaho Falls and connected to Moscow via a 100 Gbps research network; 5 NVIDIA H100 GPUs (3 H100 PCIe and 2 H100 NVL); and 13 workstation-class GPUs (6 Quadro RTX 8000, 4 RTX 4000 Ada, 2 RTX A5000, and 1 RTX A4000) reserved for speech services, smaller models, and sensitive-data scanning. The cluster currently serves more than 60 language, vision, embedding, and reranking models, ranging from compact models to large open-weight systems such as Kimi K3, Qwen 3.5/3.6/3.8, OpenAI gpt-oss-120b, Google Gemma 4, NVIDIA Nemotron 3 Super, and Mistral Large. Beyond standard text generation, the platform supports tool calling and reasoning for coding agents, vision and document understanding (including OCR), text embeddings and reranking for retrieval-augmented generation, text-to-image generation and editing (FLUX.2) with invisible watermarking, text-to-

video generation, image background removal, speech-to-text and text-to-speech, web search, and structured decision models. All models run on vLLM and Ollama, and because everything executes on university-owned hardware, prompts and data never leave university infrastructure.

MindRouter

The cluster is accessed through MindRouter (https://mindrouter.ai), an inference gateway and load balancer developed by RCDS and released as open source under the Apache 2.0 license. MindRouter provides the entire cluster behind a single endpoint and a single API key, routing each request to a healthy server running the requested model, balancing load across multiple replicas of popular models, and scheduling requests fairly so that no single user can monopolize capacity. It speaks the OpenAI, Anthropic, and Ollama APIs, including the OpenAI Responses API and the Model Context Protocol (MCP), so existing tools work without code changes, including coding agents such as Codex, OpenCode, and Goose; office add-ins; RAG frameworks; and custom applications. Users sign in with their university credentials and manage their own API keys, while administrators gain per-user quotas and usage reporting, full request audit logging, automated detection of sensitive data (PII and credentials), and per-server GPU power and energy monitoring. MindRouter also powers VandalChat, the university's homegrown web-based AI assistant, and handles tens of thousands of requests per day.

Because this scale of operation falls well outside typical University of Idaho information technology and computing services, we maintain our own support infrastructure. These include several servers for storage and authentication of user accounts (LDAP), domain name resolution (DNS), internet address assignment (DHCP) and secure connections to private networks (VPN). We also provide web and database services for online documentation and training.

Data Storage Systems

We have four distinct classes of data storage. The first group is our high-performance non-redundant storage (330TB available). This is a ZFS based storage system with large SSD caches for reading and writing and is intended for user software installation and hosting other easily-reproduced data. This system is composed of Supermicro equipment.

The second group is our commodity storage (6.0 PB gross, 2.0 PB realized). This storage group uses cheaper, but slower, disks to house most user data. We currently run a Ceph distributed file system, which offers horizontal scalability and reliability. Components currently include 5 monitor, 5 metadata, and 53 object storage servers (various Dell and Supermicro rack servers). All Ceph servers are redundantly connected to the data center networks. Additional storage servers are added to the Ceph storage pool regularly.

The third storage group comprises our backup storage systems (2.3PB realized on UI main campus, 830 TB at INL). We back up user data regularly to an offsite location on HDDs using ZFS snapshots. Components include Storinator and Supermicro rack servers.

Finally, our main virtualization and hosting infrastructure uses 1.2PB gross Ceph distributed file system and includes 72TB of SSD drives. This system is backed up to the INL C3 datacenter. Our legacy virtualization and hosting infrastructure uses a 30TB ultra-high-performance, in-chassis SSD RAID-5 array to host the virtual disks for production VMs. 500TB of slower production data storage is provided by a NetApp FAS 2554 server connected to the VRTX chassis at 10Gbps. A second 500TB NetApp FAS 2554 is located within the C3 data center in Idaho Falls, primarily for the purposes of off-site backup and Disaster Recovery (DR).

Data Management and Scientific Programming Services

RCDS provides research data management infrastructure and services that enable UI investigators to store, catalog, disseminate, and archive their research data outputs. RCDS helps researchers write comprehensive Data Management and Sharing Plans (DMSPs) to include in their funding proposals. RCDS collaborates with UI Library to manage and use the UI's official research data repository and is actively engaged with related national initiatives such as the Earth Science Information Partners (ESIP). RCDS can also provision Digital Object Identifiers (DOIs) for data through our institutional DataCite membership.

RCDS provides custom scientific programming services to help investigators design, develop, and host custom research applications and tools. These tools often include databases, web/mobile applications, and interactive geospatial mapping and data visualization. RCDS is increasingly involved in helping researchers with computational modeling, machine learning, agentic AI, and mobile application development.

Science DMZ

Working with the UI OIT Networking team, we have set up a Science DMZ network to allow for unfettered, high-throughput access to research data. This Science DMZ network hosts multiple Globus Data Transfer Nodes (DTNs) and perfSONAR servers. Globus allows for reliable, high-speed data transfers across IRON and the Internet, making it possible to connect safely and efficiently to computational cores at collaborators' institutions to share data. The perfSONAR servers provide a managed mechanism to test and improve latency and network throughput to other institutions.