

EECE 7398: Machine Learning with Small Data

Utilizing the Explorer Cluster for Advanced Research

Prof. Sarah Ostadabbas

Electrical and Computer Engineering @Northeastern University

September 8, 2025

Introduction to the Explorer Cluster

- What is the Explorer Cluster?
- Why is it critical for machine learning research?
- How has it evolved from the Discovery Cluster?

What is the Explorer Cluster?

- The Explorer Cluster is Northeastern's new **high-performance computing (HPC)** resource that supersedes Discovery.
- It features cutting-edge technologies:
 - **Rocky Linux 9.3** operating system for enhanced compatibility
 - State-of-the-art **NVIDIA H200 GPUs** with 140 GB memory per GPU
 - Advanced storage solutions and improved efficiency
- Designed for complex computational tasks including:
 - Machine learning and artificial intelligence
 - Large-scale data analysis and simulations
 - Bioinformatics and scientific modeling



Migration from Discovery to Explorer

- **All public GPU resources** have been moved from Discovery to Explorer
- Key improvements in Explorer:
 - More efficient operating system (Rocky Linux 9.3)
 - Enhanced GPU capabilities with NVIDIA H200s
 - Improved Open OnDemand interface
 - Better software compatibility and performance
- Your data from Discovery is preserved and available on Explorer
- Seamless transition for existing users

RC Migration: Discovery to Explorer

- **Research Computing is migrating from Discovery to Explorer Cluster**
- **GPU Migration Status:**
 - **All public GPU resources** (H200s) moved to Explorer
 - Explorer features state-of-the-art NVIDIA H200 GPUs
 - Discovery still hosts lower-tier GPUs: P100, V100, T4, A100, H100
- **Where to find GPUs:**
 - **Explorer:** H200 GPUs (gpu and multigpu partitions)
 - **Discovery:** P100, V100, T4, A100, H100 GPUs (gpu and multigpu partitions)
- **Future:** Eventually all machines will be moved to Explorer

GPU Types and Capabilities

GPU Model	Memory	FP32 TFLOPS	Tensor TFLOPS	Location
H200 SXM	141 GB	67	989 (FP8)	Explorer
H100 SXM	80 GB	67	989 (FP8)	Discovery
A100 SXM (80GB)	80 GB	19.5	312 (TF32)	Discovery
A100 SXM (40GB)	40 GB	19.5	312 (TF32)	Discovery
V100 SXM2	32 GB	15.7	125 (FP16)	Discovery
V100 SXM2	16 GB	15.7	125 (FP16)	Discovery
T4	16 GB	8.1	65 (FP16)	Discovery
P100 SXM2	16 GB	10.6	N/A	Discovery

- **Key Highlights:**

- H200 offers unprecedented 141 GB memory for large models
- Tensor performance varies by precision (FP8, TF32, FP16)
- Choose GPU based on model size and computational needs

Key Features of the Explorer Cluster

- High-Performance Computing Resources:
 - Advanced CPU nodes with multiple cores
 - **32 NVIDIA H200 GPU chips** for accelerated computing
 - 140 GB memory per GPU for large-scale ML models
- Flexible Environment:
 - Access via Open OnDemand (OOD) or SSH
 - Batch job scheduling using SLURM
 - Support for Conda, Python, and scientific libraries
- Enhanced Storage:
 - New **/projects** directory for performant storage
 - Improved file system performance

How to Access the Explorer Cluster (Part 1)

- There are several methods to access the Explorer Cluster:
- **Open OnDemand (OOD) Web Portal:**
 - Enhanced web-based interface for managing files, submitting jobs, and launching interactive apps
 - Access through: **`http://ood.explorer.northeastern.edu`**
 - Improved applications including JupyterLab, RStudio, MATLAB, and VSCode
- **SSH (Secure Shell) Access:**
 - Direct terminal access for command-line operations
 - Connect using: **`ssh your-username@login.explorer.northeastern.edu`**
 - Same authentication as your Northeastern credentials

How to Access the Explorer Cluster (Part 2)

- **VSCode Remote SSH:**

- Use Visual Studio Code with the ****Remote - SSH**** extension
- Connect to **login.explorer.northeastern.edu**
- Integrated development environment for coding and file management

- **File Transfer:**

- Transfer node remains: **xfer.discovery.neu.edu**
- Use scp, rsync, or Globus for data transfers
- OOD File Explorer for small file uploads/downloads

Storage Configuration on Explorer

- **Important Storage Changes:**

- Home directory: **/home/\$USER** (75GB limit)
- Performant storage: **/projects** (replaces /work from Discovery)
- Temporary storage: **/scratch** (for running jobs only)

- **Migration Notes:**

- Update scripts to change paths from **/work** to **/projects**
- All Discovery data is preserved and accessible
- Request new /projects directories as needed

- **Best Practices:**

- Keep /home under quota by cleaning regularly
- Use /projects for research data and active work
- Use /scratch only for temporary job files

Setting Up Your Research Environment

- **Module System Updates:**

- All modules on Explorer are new with updated versions
- Use `module avail` to see available software
- Update your module load commands from Discovery scripts

- **Conda Environments:**

- Existing Discovery conda environments may work on Explorer
- Recommend testing and rebuilding if needed
- Clean unused environments: `conda clean --all`

- **Software Compilation:**

- Recompile software that was built on Discovery
- Take advantage of Rocky Linux 9.3 improvements

GPU Utilization: NVIDIA H200 GPUs

- **New H200 GPU Features:**

- 32 powerful NVIDIA H200 GPU chips
- 140 GB memory per GPU (significant increase from previous generations)
- Available on **gpu** and **multigpu** partitions

- **PyTorch and Deep Learning:**

- Install PyTorch with CUDA support for optimal performance
- Leverage increased memory for larger models and datasets
- Monitor resources with `nvidia-smi`

- **High Demand:**

- H200 GPUs are highly popular and heavily utilized
- Plan job submissions accordingly

Running Experiments Efficiently on Explorer

- **SLURM Job Scheduling:**

- Same SLURM system as Discovery for familiar workflow
- Request appropriate resources (CPUs, GPUs, memory)
- Update partition names in your sbatch scripts

- **Interactive Development:**

- Enhanced Jupyter Notebooks via Open OnDemand
- Interactive sessions with `srun`
- Desktop applications available through OOD

- **Course-Specific Resources:**

- Dedicated **courses** and **courses-gpu** partitions
- Spring 2025 course materials only accessible via Explorer

Monitoring and Experiment Tracking

- **Experiment Tracking:**

- Weights and Biases (or similar tools) for ML experiments
- Track metrics, model versions, and hyperparameters
- Analyze training progress with enhanced GPU memory

- **Resource Monitoring:**

- Use NVIDIA tools to monitor H200 GPU usage
- Track memory utilization (up to 140GB per GPU)
- Monitor temperature and performance metrics

- **Job Management:**

- Use `squeue` to monitor job status
- Check job efficiency with `seff`

Best Practices and Transition Tips

- **Login Node Guidelines:**

- Avoid compute-intensive jobs on login nodes
- Use compute nodes for development and testing

- **Storage Management:**

- Keep /home under 75GB quota
- Clean conda caches: `conda clean --all`
- Update script paths from /work to /projects

- **Transition Support:**

- Test existing conda environments on Explorer
- Recompile Discovery-built software for optimal performance
- Contact RC team for migration assistance

Getting Help and Support

- **Research Computing Support:**

- Email: rchelp@northeastern.edu
- Documentation: Comprehensive guides at rc-docs.northeastern.edu
- ServiceNow tickets for technical issues

- **Office Hours (Same Times):**

- Wednesday: 3 - 4 p.m. ET
- Thursday: 11 a.m. - 12 p.m. ET
- Short consultations (10-15 minutes)

- **Consultations:**

- Schedule one-on-one assistance
- Help with optimization and troubleshooting
- Training sessions available

- Questions about migrating from Discovery to Explorer?
- Issues with accessing the new cluster or updating scripts?
- Specific questions about H200 GPUs and enhanced capabilities?
- Research project requirements and optimization strategies?

Conclusion and Next Steps

- **Key Takeaways:**

- Explorer offers enhanced performance with H200 GPUs and Rocky Linux 9.3
- Update connection methods and storage paths
- Take advantage of improved Open OnDemand interface

- **Next Steps:**

- Access Explorer and test your existing workflows
- Update scripts to use /projects instead of /work
- Explore new GPU capabilities for your ML projects

- **Remember:** RC team is available for transition support!