Introduction to MSI

Jim Ferguson
msi.umn.edu → About → Staff

Overview

- Mission
- Organization
- Resources and Services
- Eligibility, Access, and Accounts

Minnesota Supercomputing Institute

- To provide advanced **research computing infrastructure** and **expertise** to the University of Minnesota research and scholarly community and the State of Minnesota in order to advance and accelerate research and foster innovation and discoveries through **advanced computing technologies**, scientific **computing and informatics**, **application development**, and **services**.
- Academic unit under the Office of the Vice President of Research
- 50 full time employees and ~6 students
  - Six working teams
- 900+ active Principle Investigator groups using MSI resources
- 4,800+ active users
- Significant resources and expertise provided for FREE to UMN PIs.
- UMN have been leaders in academic supercomputing since the 1980s.

Minnesota Supercomputing Institute:

Strong HPC History & Regular Upgrade Cycle

Major Upgrade 96 TeraFlops

Itasca
Sun 1/31

Major Upgrade 860 TeraFlops

Mesabi
Sun 3/1

Major upgrade 7000 TeraFlops

Agate
Mon 11/1

Agate+1
Wed 8/5

Planned: Major upgrade

2010 | 2010 | 2012 | 2014 | 2016 | 2018 | 2020 | 2022 | 2024 | 2026

Wed 9/1
Blackwell

HIPAA System 360 TeraFlops

Tue 8/13
Agate+1/2

Planned: Minor upgrade

Fri 9/13
Mangi

Minor Upgrade 1150 TeraFlops

UMN Supercomputing History

- Seymour Cray UMN graduate
- First university to own a Cray-1 supercomputer **-1981**
- Home to Army HPC Research Center - 1989-2006
- Top500 systems since the beginning of the Top500

Office of the Vice President for Research

Research Computing

Minnesota Supercomputing Institute

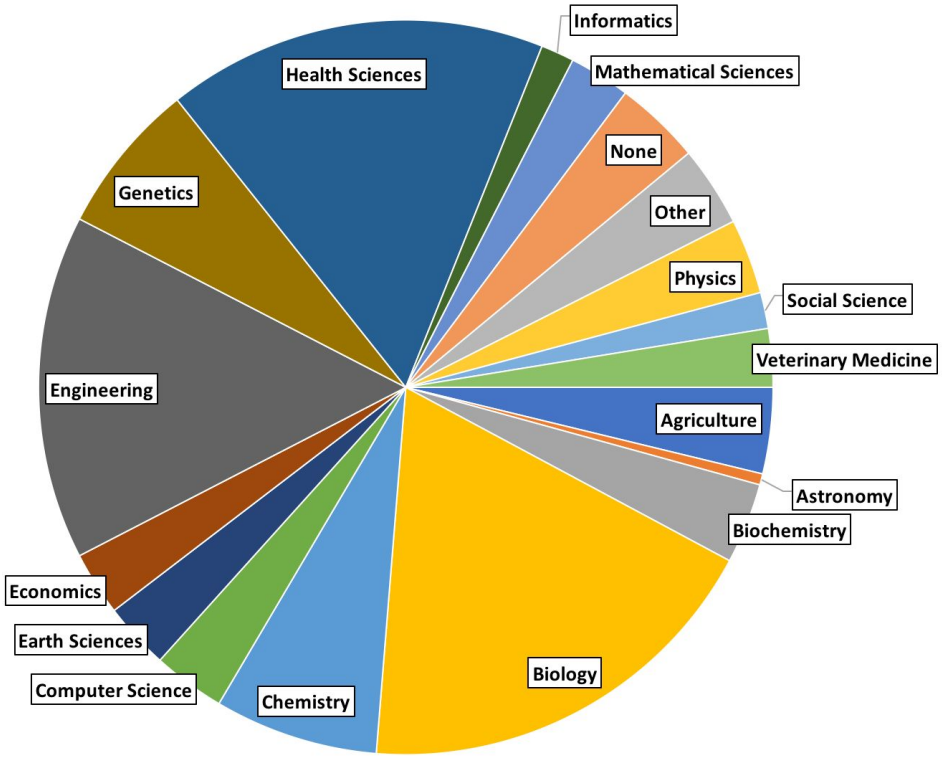
UM Informatics Institute

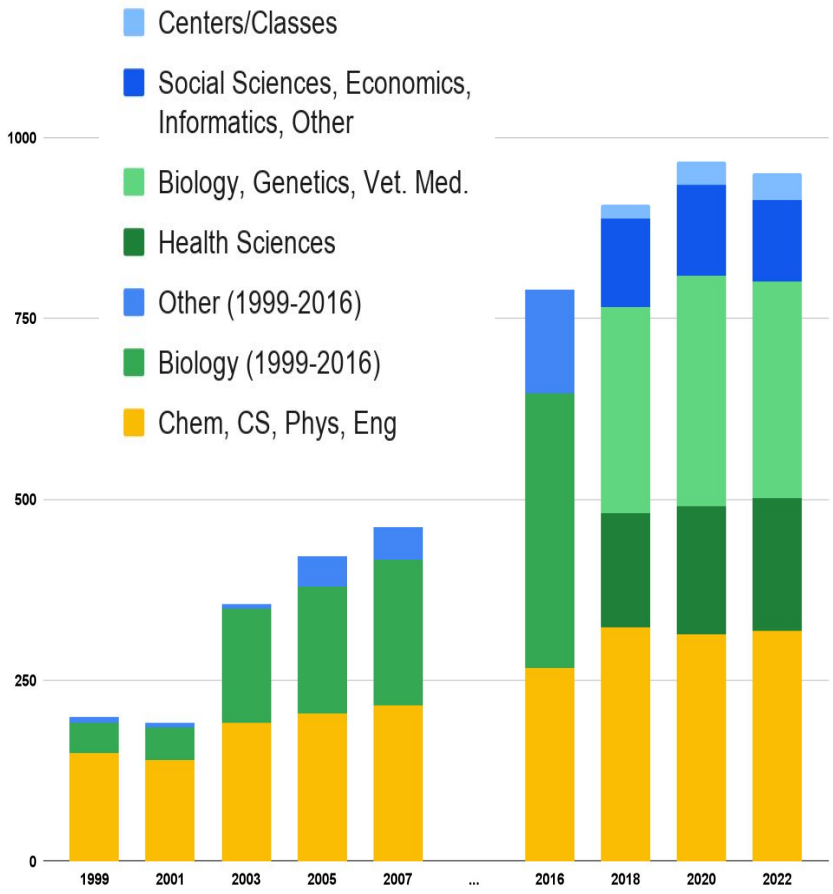
U-Spatial

User Gateway Group	Scientific Computing Solutions	Research Informatics Solutions	Application Development Solutions	Advanced Systems Operations
-- First Line User Support -- User Training -- On Boarding -Communications -- Outreach	-- Optimization -- Benchmarking -- HPC Research Workflow & pipeline Development	-Informatics education -Informatics research -Informatics services -Life Science Computing	- Custom App Dev - System Programming	- Common Services - HPC Systems - Storage Systems - Hosted Services

The Long Tail of Research Computing

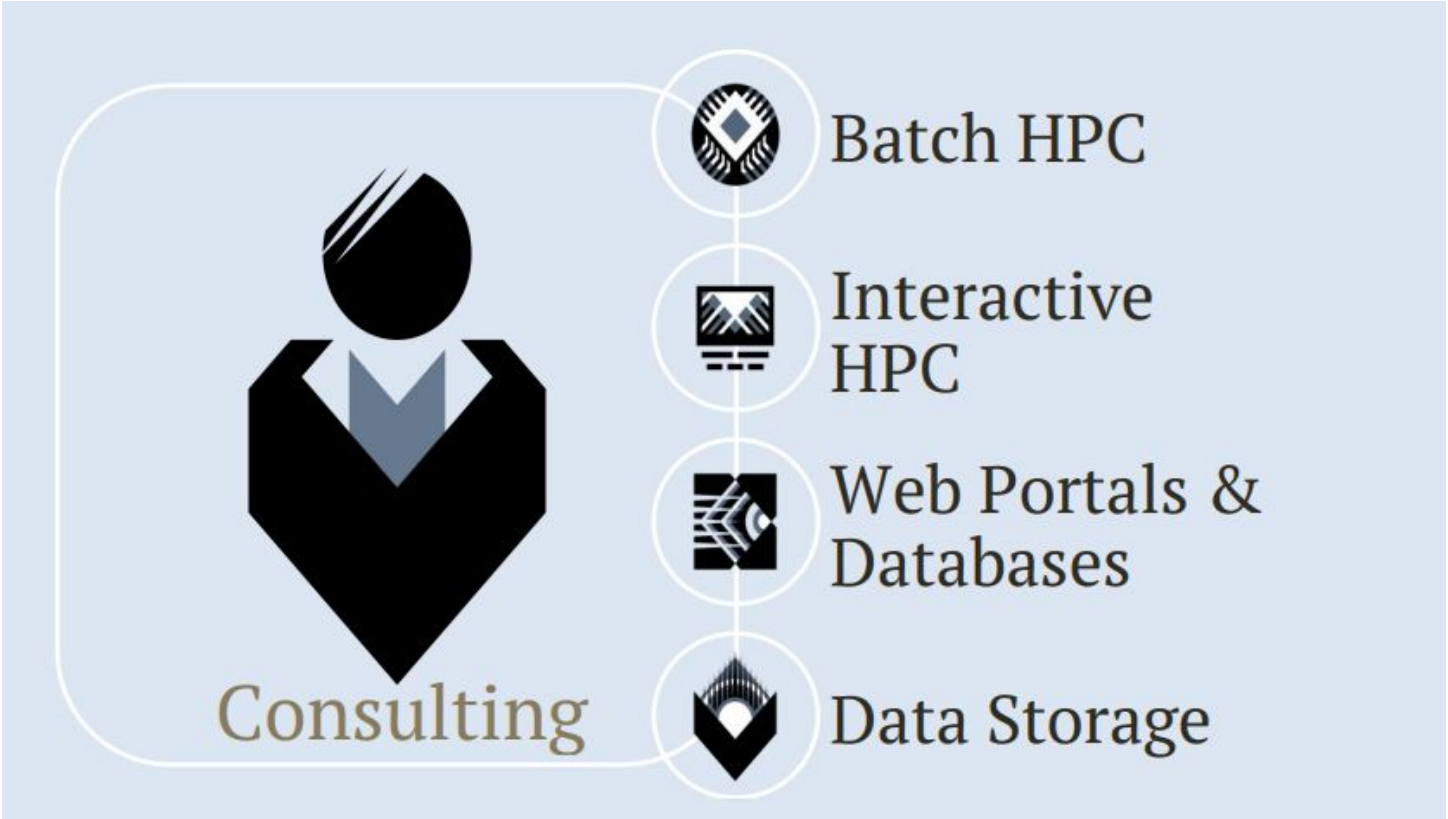
Groups in 2020: 900 User Groups
4,555 Active users





Biggest increasing in Life Sciences

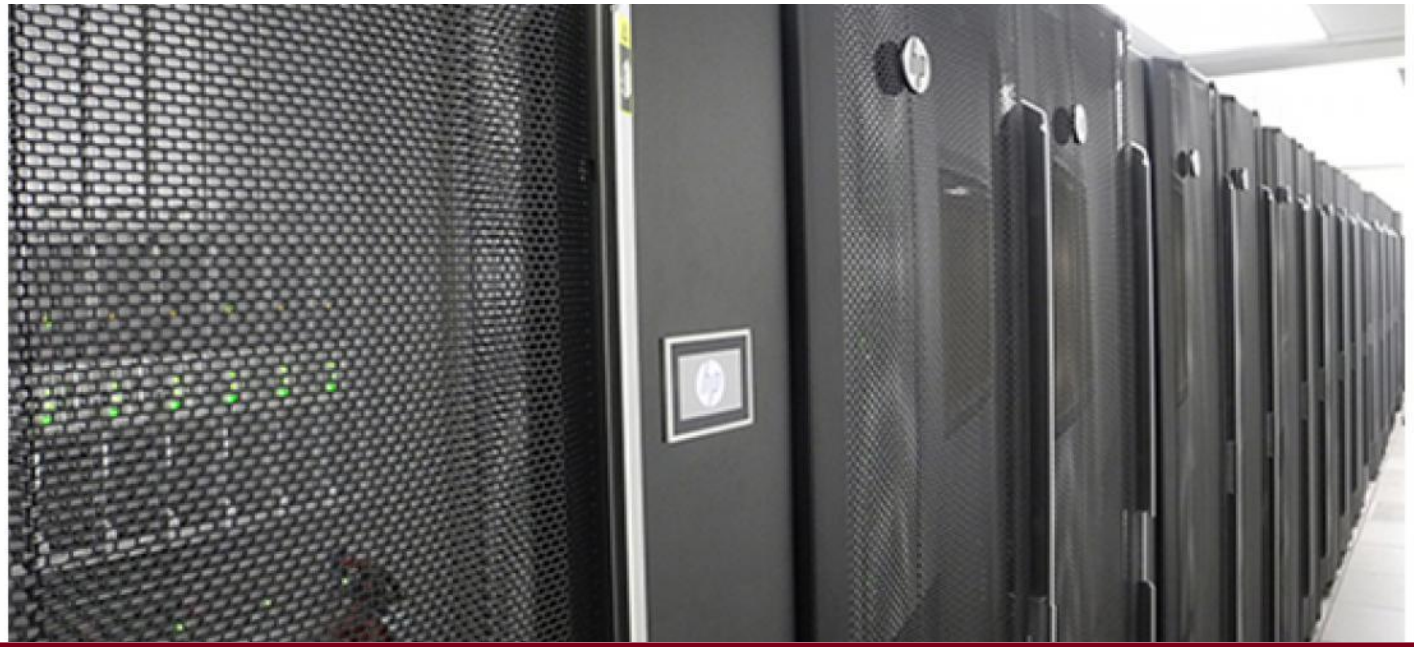
Resources and Services



High Performance Computing: Batch & Interactive

High Performance Computing Services

- Highly parallel computing environment with high-speed network connections, graphics processing units, high memory, and high-performance storage
- Offer both **batch** HPC and **interactive** HPC
- HPC access controlled by the “Fair share” SLURM scheduler



Mesabi/Mangi HPC System

Mesabi

- 741 nodes
- 17,784 cores
- 40 nodes with 2x NVidia Tesla GPUs
- 32 nodes with 480GB SSDs

Mangi

- 164 AMD ROME Nodes
- 20,992 cores
- Amd2tb Nodes
- v100 GPU nodes
- v100-4 GPU nodes
- v100-8 GPU nodes

Agate HPC System

- 264 NVIDIA A100 GPUs are available in two configurations.
  - 50 nodes have 4 A100 GPUs and 512 GB of memory
  - 8 nodes have 8 A100 GPUs and 1 TB of memory
  - Multi-GPU nodes are connected via NVLink

- 344 CPU-only compute nodes are available in two configurations
  - 244 with 512 GB of memory
  - 100 with 2TB of memory

- 10 GPU nodes equipped with 8 A40 GPUs, 128 cores, and 512 GB of memory are made available for interactive work though Jupyter or command line sessions.

Batch HPC

- Using batch scheduling is the primary way to make use of MSI's shared supercomputing resources.
- You will **'submit'** jobs to the cluster scheduler, and your analysis eventually runs on the computer without any further intervention.
  - Requires jobs to be written as **shell scripts** and submitted via the command line.
- Jobs get **queued** and run based on a combination of the size of the request, the current load of the systems, how many jobs your group has run recently.
  - This is called "FairShare" - details can be found at this link: https://www.msi.umn.edu/content/hpc
- This will probably be where you spend most of your time interacting with MSI, unless you have a specialized interactive application.

Batch HPC: SLURM

Job Submission and Scheduling (SLURM Scripts)

```
#!/bin/bash -l
#SBATCH --time=8:00:00
#SBATCH --ntasks=8
#SBATCH --mem=10g
#SBATCH --tmp=10g
#SBATCH --mail-type=ALL
#SBATCH --mail-user=sample_email@umn.edu
#
cd ~/program_directory
module load intel
module load ompi/intel
mpirun -np 8 program_name < inputfile > outputfile
```

Batch partitions

SLURM scripts generally have a line to designate a partition (or partitions) where your batch job can execute. It looks like this:

$SBATCH -p <NameOfPartition>

A table of the current partitions on all MSI HPC clusters can be found at

https://www.msi.umn.edu/partitions

Research groups who lease private nodes will have private partition names they can use to access their purchased nodes.

Interactive HPC

- MSI also offers interactive HPC access
  - Very useful for real-time data visualization and exploration
  - Allows use of specialized hardware and software without labs having to purchase it
- Types of interactive HPC offered:
  - srun: SLURM batch method to get an interactive shell on HPC queue
  - NICE: web-based desktop
  - CITRIX: remote Windows desktop on HPC hardware
  - NX NoMachine: remote Linux graphical desktop
  - Jupyter Notebooks (beta): interactive Python/R on MSI
  - *Coming Soon: OpenOnDemand*
- See https://www.msi.umn.edu/content/interactive-hpc

Interactive HPC

Interactive queue use with srun, using an applications that needs X windows:

```
ssh -X mesabi.msi.umn.edu
```

your_user_id@ln0004 [~] % **srun -N 1 --ntasks-per-node=4 --mem-per-cpu=1gb -t 1:00:00 -p interactive --x11 --pty bash**

your_user_id@cn0001 [~] % **module load a-module-you-need**
your_user_id@cn0001 [~] % **./my_program**

Software Resources

- Scientific and numerical software packages are built into “modules” and available to be loaded as you need them.
- Common software tools used in many fields are available on our systems
  - Sequence analysis
  - Genetics
  - Proteomics
  - Image processing
  - Computational chemistry
  - Etc…
- 500+ software applications
  - MSI Software Search
- We can work with your research groups to install software that you need
  - Contact the help desk!

Service Units

- Service Units is a term MSI uses to reference the total compute resources consumed by a group or individual.
- Commercial and Industrial partners still purchase SUs, and their compute, storage, and memory usage all are accounted for.
- Formerly, MSI used SUs to allocate different resources for academic research groups.
- Allocation of resources is now done among groups via the SLURM “Fair share” algorithm.
- **Full deployment of SLURM commenced in January 2021, and the share of resources among groups is being monitored by an advisory group of MSI principal investigators.**

Data Storage at MSI

Primary Storage

- Primary storage is where you will do most of your work
  - GNU/Linux filesystem
- Watch your disk usage:
  - If you are up against file number or disk usage quota, **your entire group** cannot make new files
- Important directories in your primary storage space:

`/home/groupname/YOUR_X.500`
Your home directory. Store scripts and small files here

`/home/groupname/shared`
PI group shared directory. Store shared software, common datasets, reference databases, etc, here.

Primary Storage Scratch

- Quotas are large, so there should be leeway to make very large files here
  - Use it for intermediate/temporary files in analytical workflows
- Files are kept for 30 days, so be sure to copy important files to your group directory
  - Scratch space is also not backed up by snapshot
  - **We know the tricks that exist to keep data in scratch for longer than 30 days. Consider data on scratch as having no warranty.**

- Location:

  `/scratch.global/`

Second Tier Storage

- Sometimes referred to as “Ceph”, the name of the software defined storage utility in use in this storage layer.
- Second tier storage is separate from primary storage, it is an “object store” - you must use special tools to access it.
- From the command line, you can use the “s3cmd” software package to interact with your storage:
  - See: https://www.msi.umn.edu/support/faq/how-do-i-use-second-tier-storage-command-line
- For a graphical interface, you can use Globus
  - See: https://www.msi.umn.edu/support/faq/how-do-i-use-globus-transfer-data-msi-0
- **Not** backed up, so if you delete a file, it will be gone for good
  - Quotas are much more relaxed on second tier storage
  - **Resilient system to store data**

Data Storage Allocation

- Default Primary storage allocation: 150GB for each group
- **Requests of up to 1TB:**
  - No justification or review required
- **For requests of up to 5TB:**
  - Justification
  - Review by MSI staff
- **For requests of more than 5TB:**
  - Justification
  - Estimate of amount needed
  - Estimate of amount actively used
  - Estimate of duration of storage
  - Review by MSI staff
- ANNUAL renewals for all non-default storage allocations

PI Groups, Accounts and Support

Service Costs

- **Free** for UMN faculty and their associates
- **Free** for other MN academic institutions
- **Free** for expert consultation

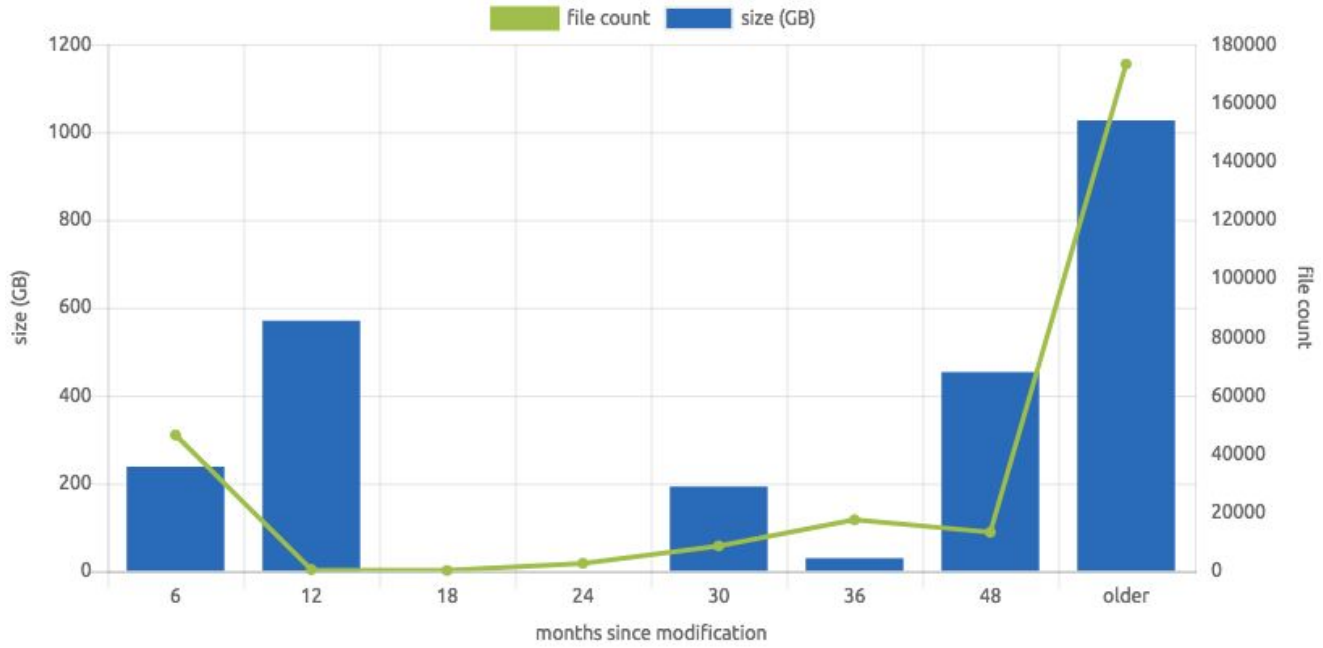
- Fee for resources over allocations limits
- Fee for expert collaboration
- Fee for service for external organizations

PI Group Structure

- **PI**: Controls everything
- **Group Administrators:** Users within the group designated to do everything a PI can do except add other Group Administrators
- **Users**: Added by PI, usually lab members (postdocs, techs, undergrads, grad students)
  - Users may be part of multiple accounts if they perform analysis for multiple PIs
  - The PI is a user, too, and their SU and disk usage counts the same as their users'

Checking Group Usage

- Check disk usage: https://www.msi.umn.edu/group/GROUPNAME/storage



Accounts and Renewals

- UMN faculty members, department staff, and researchers at other MN state institutions are eligible for MSI accounts

- **PI accounts are renewed annually**
  - Renewal period starts in October
  - Extra storage allocations all must be checked and justified

Research Support Services

- MSI has ~25 consulting staff, most with an advanced degree.
- We provide 20-30 workshops and tutorials annually
- 1-2 hour meetings on computational or informatics methodologies
- Informational sessions for proposal development
- Expertise including:
  - Research Computing; Informatics
  - Big Data Analysis
  - Parallel Algorithm Optimization and Development
  - Custom Analysis Pipeline Development
  - Genomics
  - Custom Application Development

Training and Tutorials

- We host tutorials on a range of topics https://www.msi.umn.edu/tutorials
- **Introductory tutorials**
  - Intro to MSI
  - Intro to Linux
  - MSI Batch Job Submission
- **Bioinformatics tutorials**
  - Proteomic and mass spec data analysis
  - RNAseq data analysis
- **Advanced tutorials**
  - Parallelization
  - Python for scientific computing
- **And many more!**
- **We also collaborate with The Carpentries and LATIS at UMN, have a look at our Events page:** https://www.msi.umn.edu/events

Links to Resources

- MSI Homepage
  - https://www.msi.umn.edu
- PI eligibility and access policies
  - https://www.msi.umn.edu/content/eligibility-getting-access
- SLURM batch partition limits page
  - https://www.msi.umn.edu/partitions
- Tutorials page
  - https://www.msi.umn.edu/tutorials
- Check Primary storage
  - https://www.msi.umn.edu/group/GROUPNAME/storage
- Software tools
  - https://www.msi.umn.edu/software

Questions?

- MSI Help Desk
- Every Day help available using Email or Video
- Walk-Ins
  - Monday, Wednesday, Friday (587 Walter Library, UM campus)
    9:00 AM – 4:00 PM
  - Tuesday, Thursday (Zoom virtual walk-in)
    9:00 - 11:00 AM, and 1:00 - 3:00 PM
    Zoom Meeting Room Link: https://umn.zoom.us/my/msi.helpdesk

- Call x6-0802 (1-612-626-0802)
- Email **help@msi.umn.edu**
- Schedule a Zoom Consultation
  https://www.msi.umn.edu/content/helpdesk-video-consultations