Quick Intro to MSI

Ju Sun (Computer Science & Engineering)

Mar, 2025

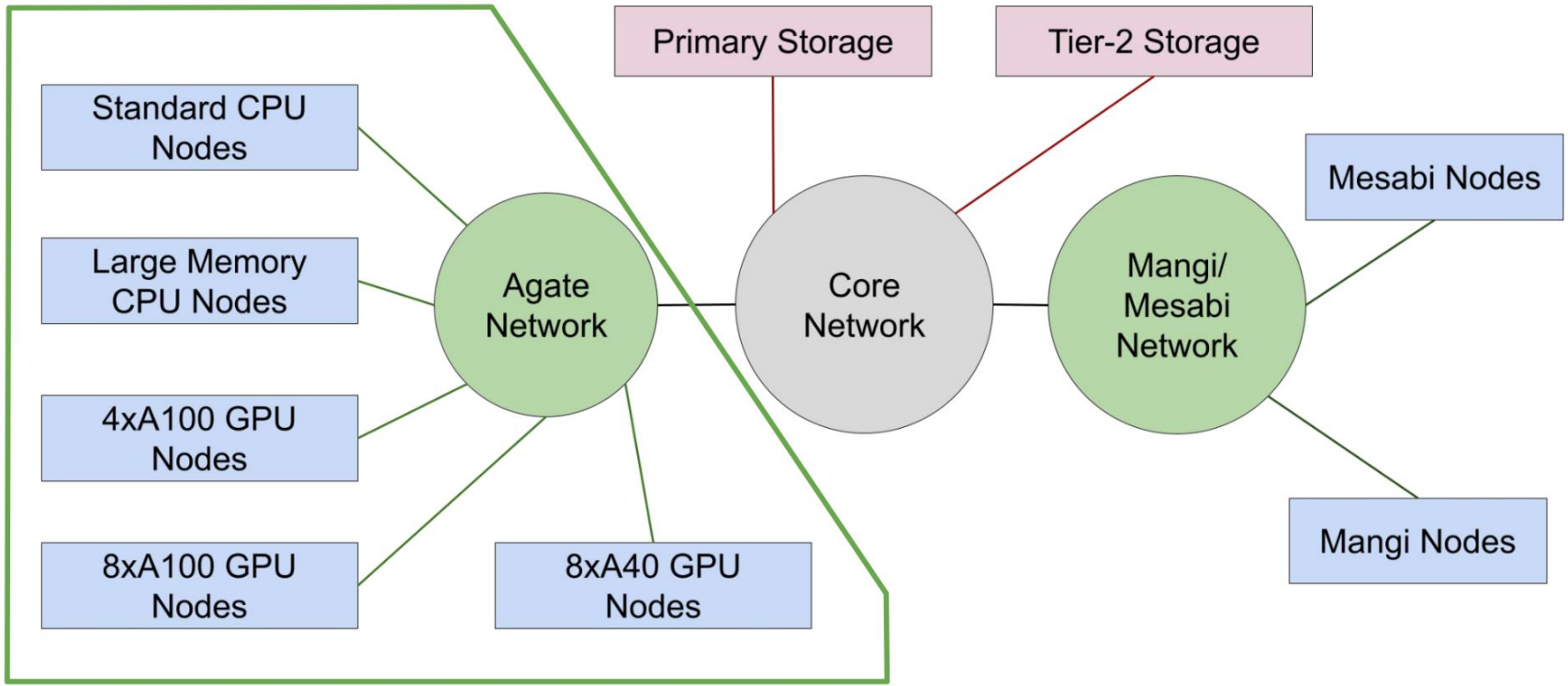
UNIVERSITY OF MINNESOTA
Driven to Discover℠

Disclaimer

Many of the slides are taken from MSI's tutorial slides below, as well introductory slides made by **Dr. Ben Lynch, MSI Director.**

- Introduction to Minnesota Supercomputing Institute (MSI) https://youtu.be/bNBfh2Q9l1E
- Job submission and scheduling at MSI https://youtu.be/r4m4WK4LfPY?si=BxtSrhz0fEcYCTH6
- Interactive computing at MSI https://youtu.be/D6D77_9Emhw?si=PlMJi9p-GYO_3bcF

What's MSI?


Primary Storage
Tier-2 Storage
Standard CPU Nodes
Large Memory CPU Nodes
4xA100 GPU Nodes
8xA100 GPU Nodes
Agate Network
8xA40 GPU Nodes
Core Network
Mangi/ Mesabi Network
Mesabi Nodes
Mangi Nodes

What's MSI?

Mesabi

- Over 700 nodes
  - Memmory configuration
    - 616 nodes have 64GB RAM
    - 24 nodes have 256GB RAM
    - 16 nodes have 1TB RAM
    - 40 k40 GPU nodes with 128GB RAM
- 17,784 cores provided by Intel Haswell Processors
- 480GB SSD available on 32 nodes

Agate

- 412 nodes
- AMD processors with 64-128 CPU cores per node
- 344 CPU compute node
  - 244 have 512G mem
  - 100 have 2TB mem
- 58 GPU compute nodes
  - 50 A100 512G mem
  - 8 A100 1TB mem
- 10 GPU interactive nodes
  - 8 A40 GPUs 512G mem each

Basic access

Prerequisite to connect

Campus VPN

https://it.umn.edu/services-technologies/virtual-private-network-vpn

- ➔ MSI services are only available to on-campus network addresses
- ➔ Especially during the pandemic, many users are working off-campus
- ➔ VPN clients and instructions are available from OIT for most operating systems and mobile devices

DUO Authentication

https://it.umn.edu/services-technologies/self-help-guides/duo-set-use-duo-security

- ➔ DUO 2-factor authentication is required for most MSI services
- ➔ DUO is also required to connect to the campus VPN

Two way to access

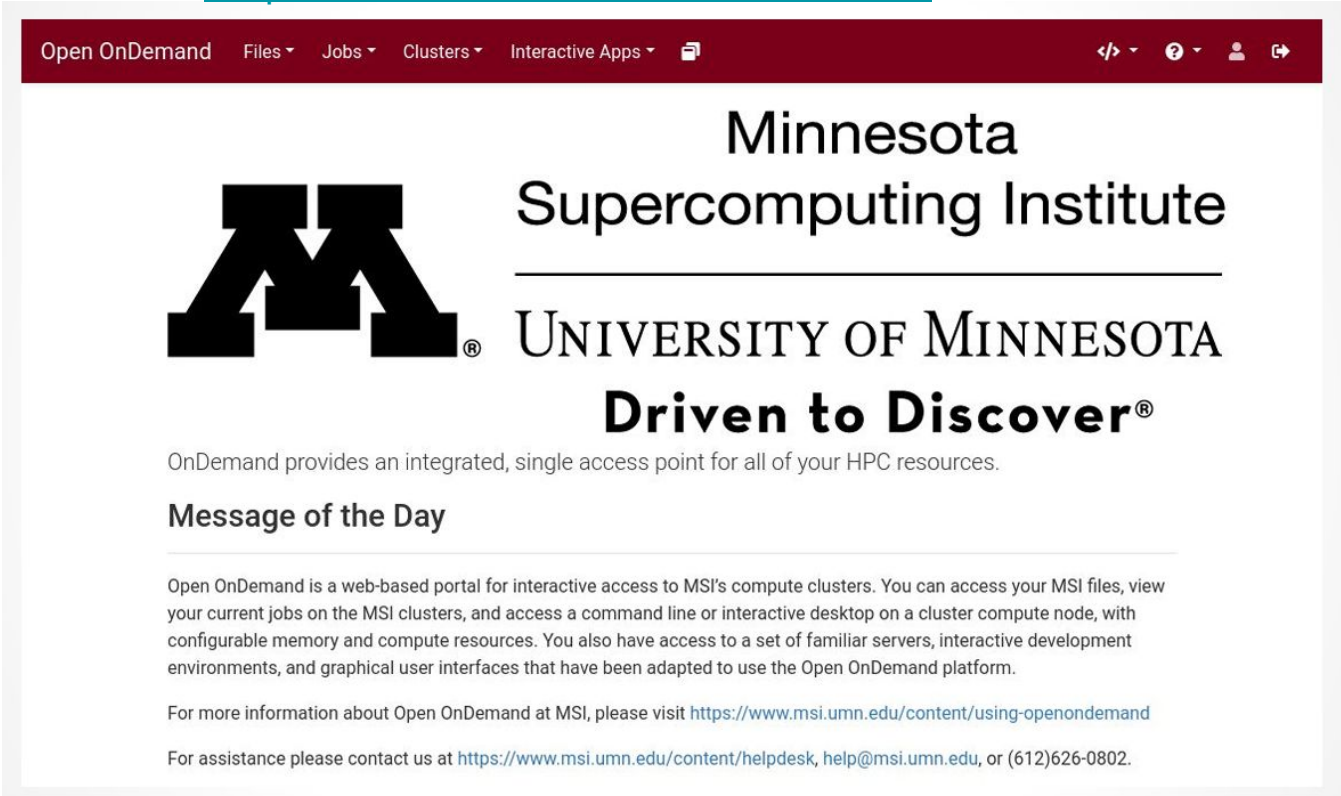
- GUI interactive access
- Command-line access (please check their tutorial details)

Job submission and scheduling at MSI
https://youtu.be/r4m4WK4LfPY?si=BxtSrhzofEcYCTH6

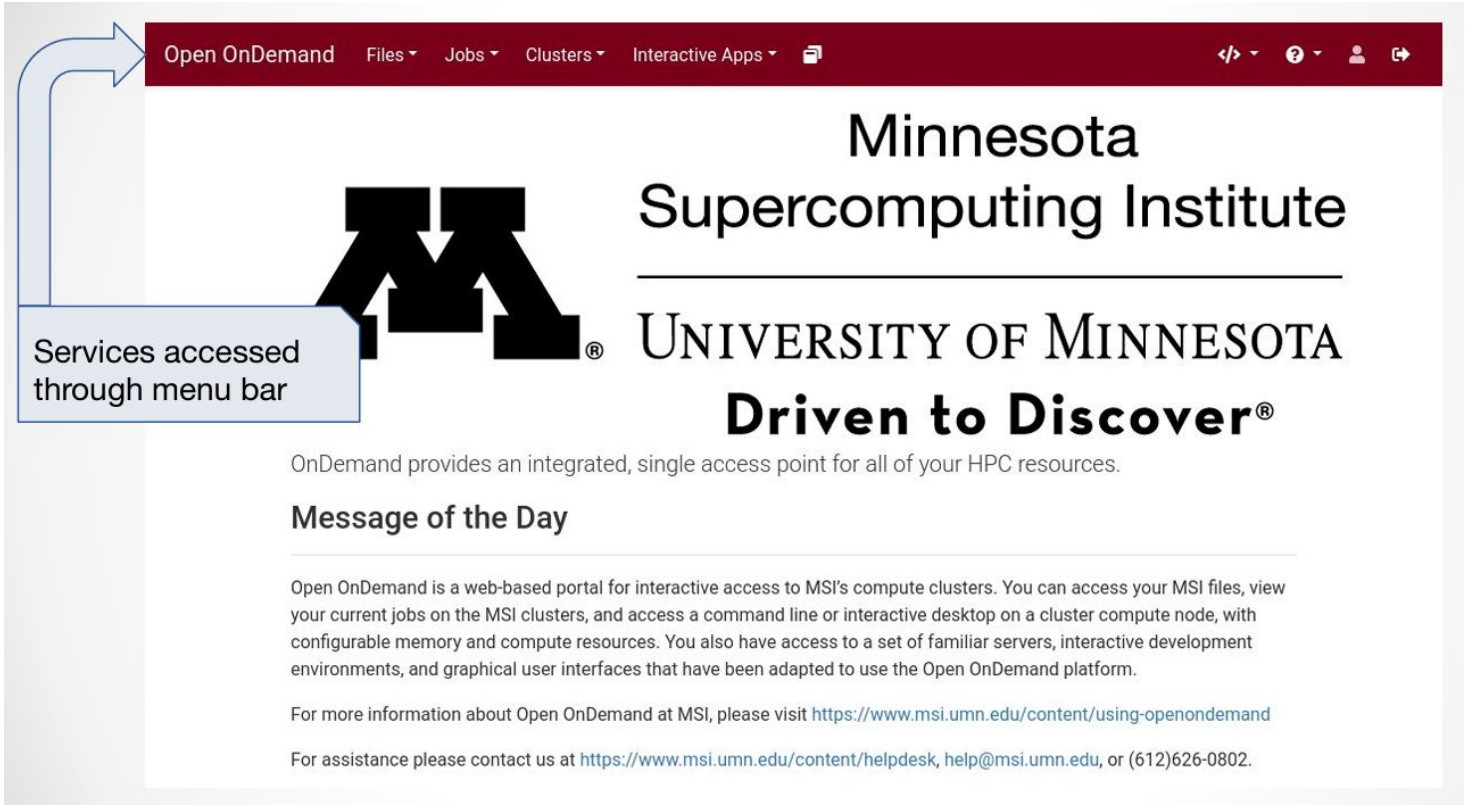
Interactive computing at MSI
https://youtu.be/D6D77_9Emhw?si=PlMJi9p-GYO_3bcF

GUI interactive access: OnDemand

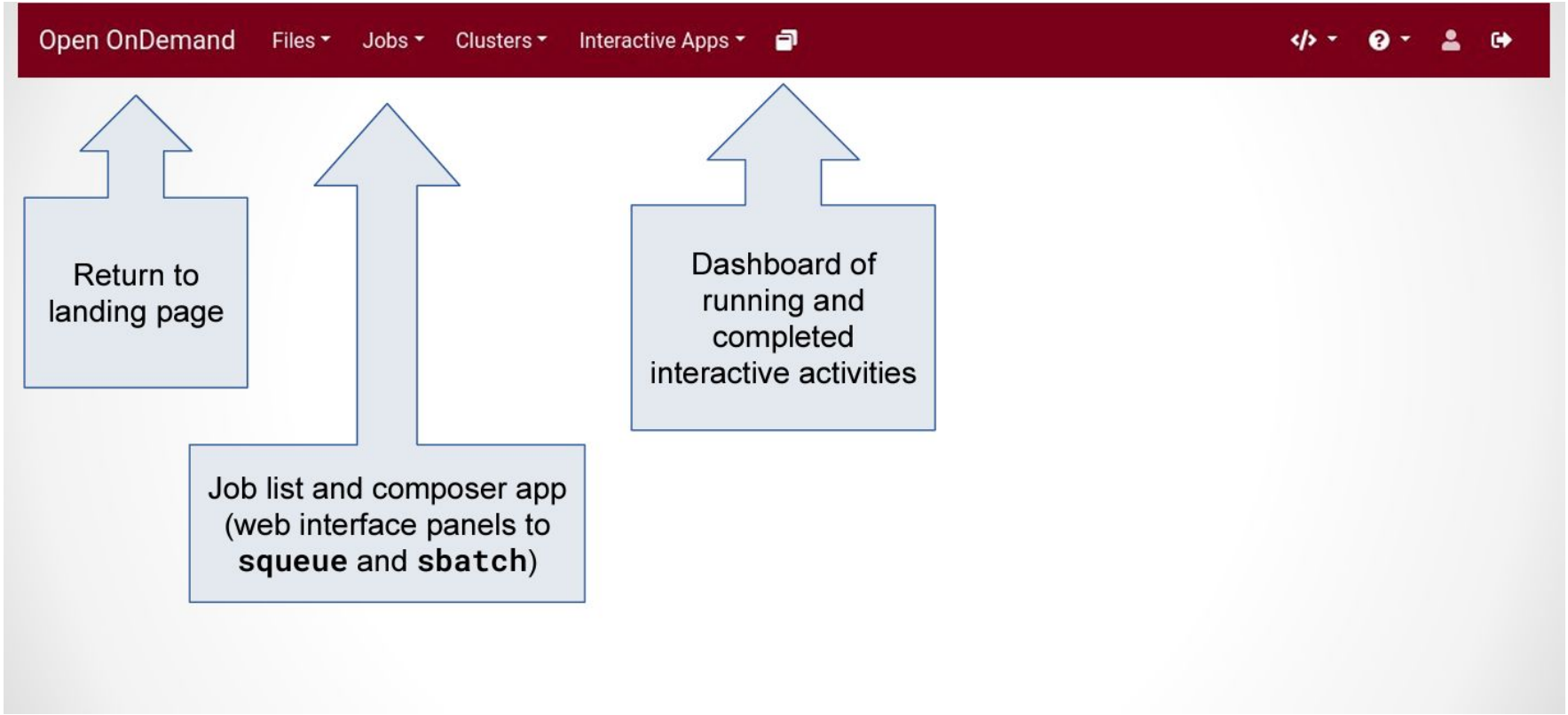
https://ondemand.msi.umn.edu/



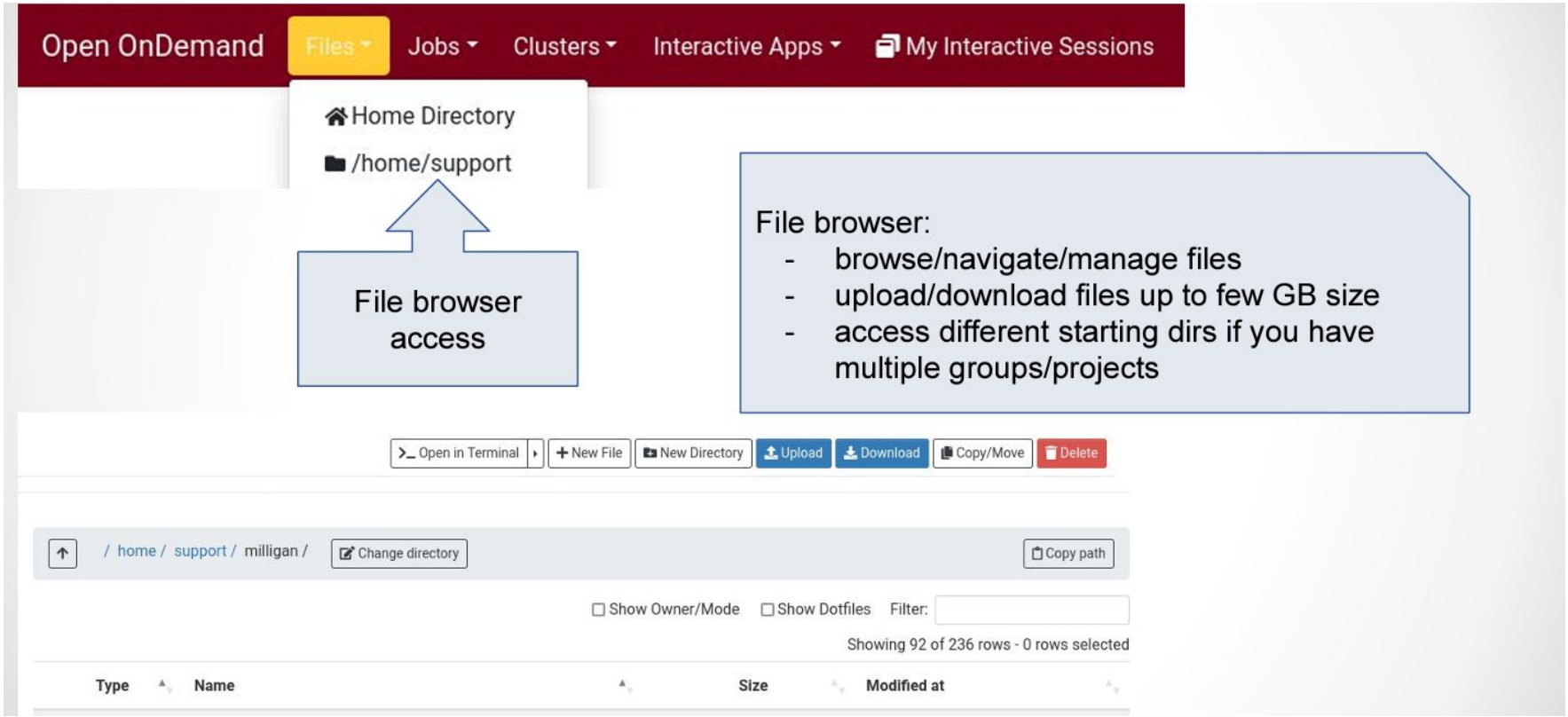
OnDemand walkthrough



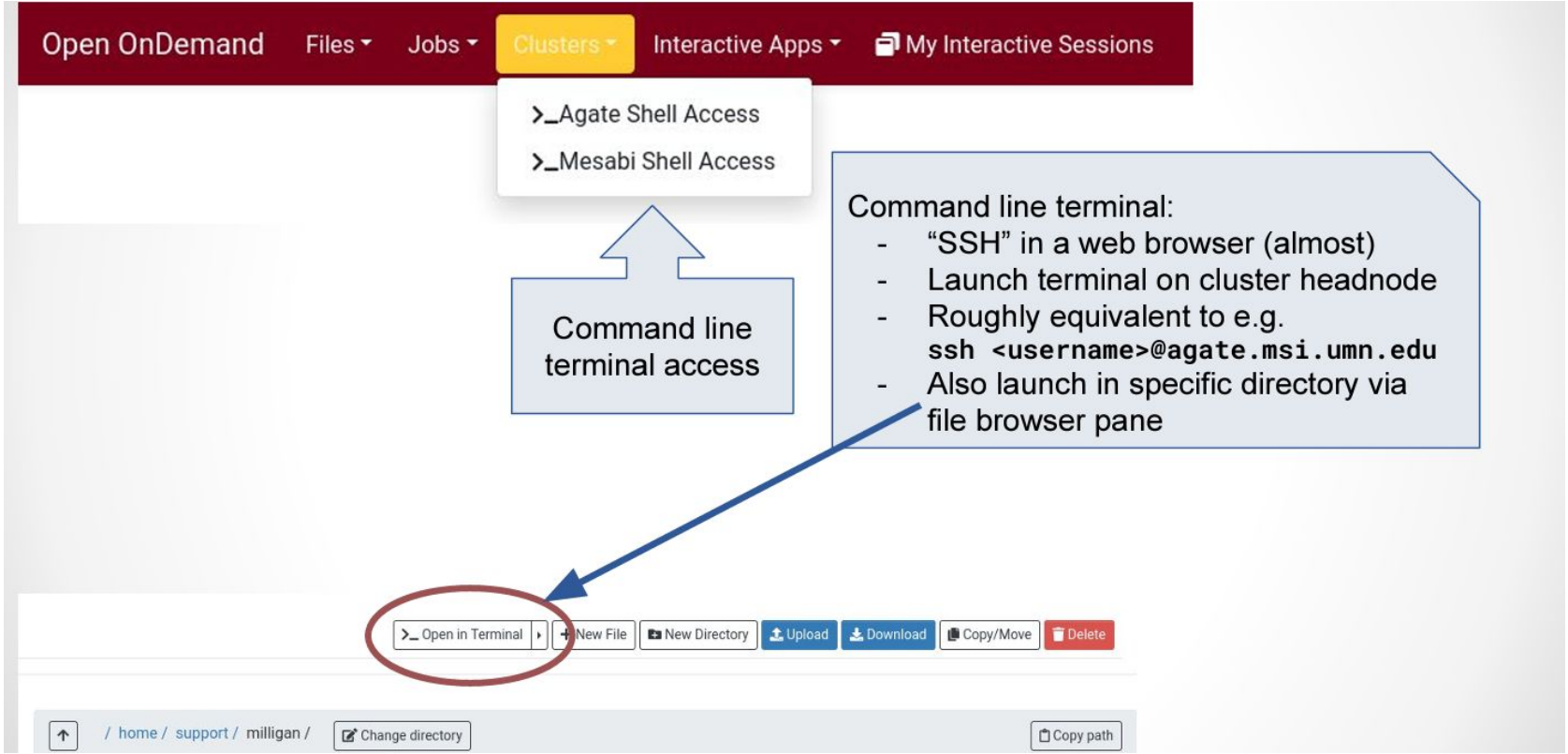
OnDemand walkthrough



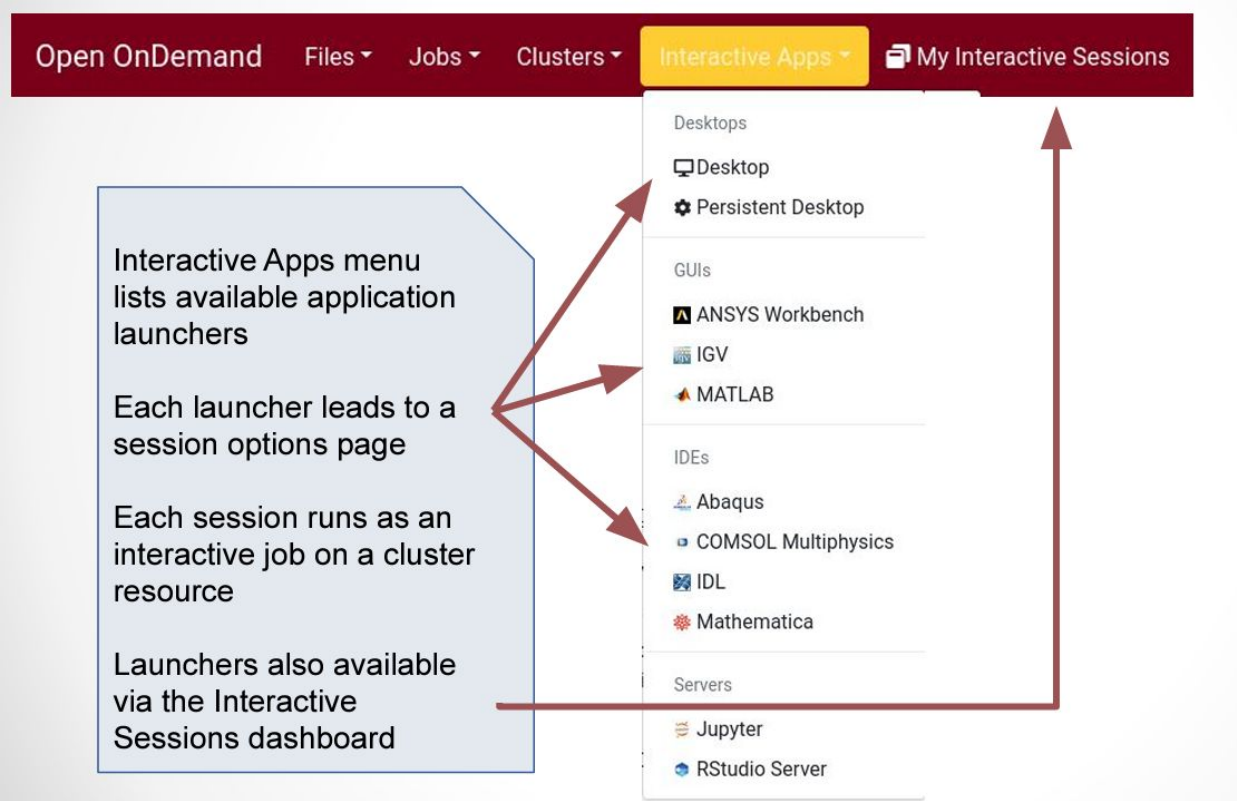
OnDemand walkthrough



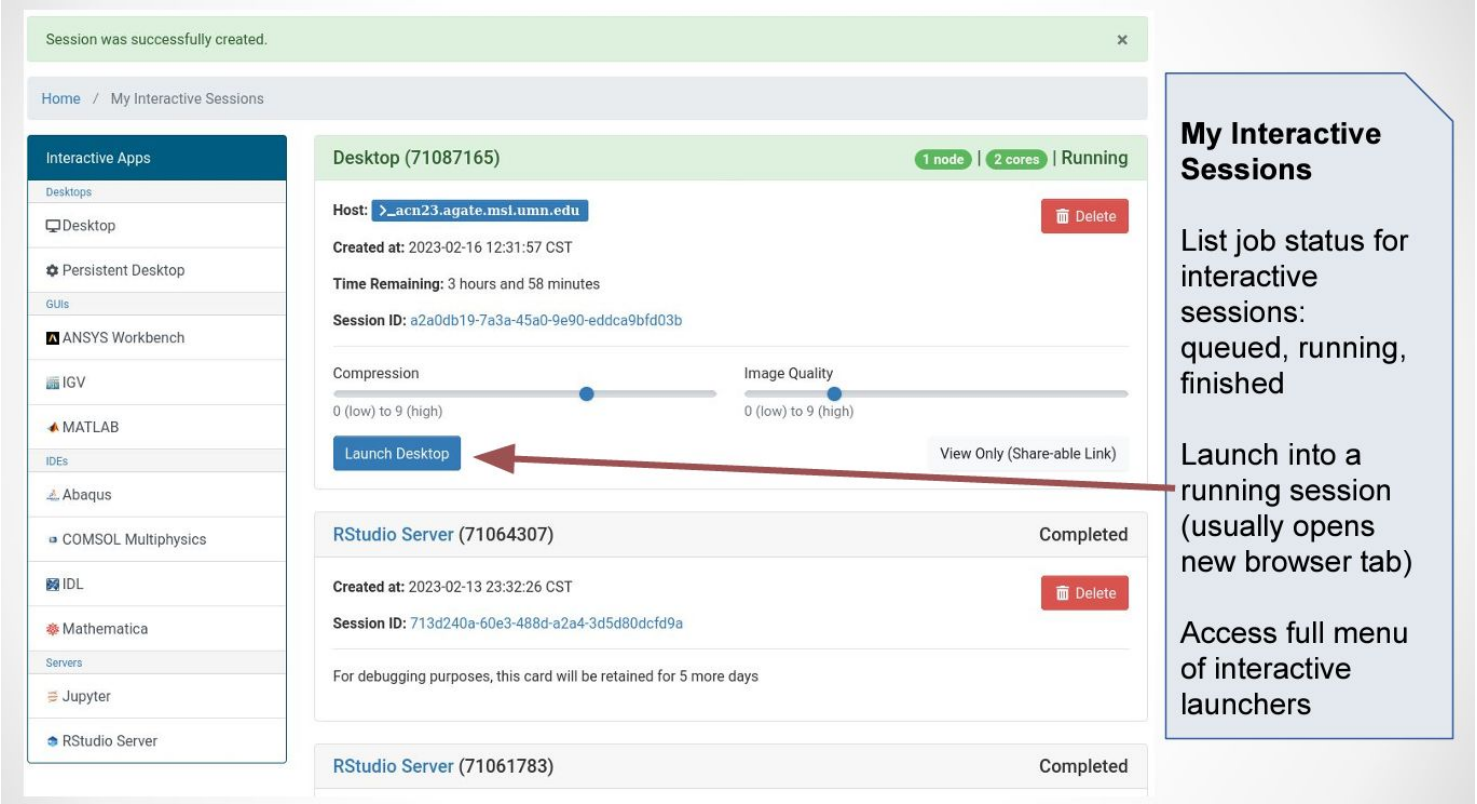
OnDemand walkthrough



OnDemand walkthrough



OnDemand walkthrough



Setting up your deep learning environment

Put everything inside a virtual environment

Conda

Package, dependency and environment management for any language—Python, R, Ruby, Lua, Scala, Java, JavaScript, C/ C++, Fortran, and more.

Miniconda **Highly recommended!**

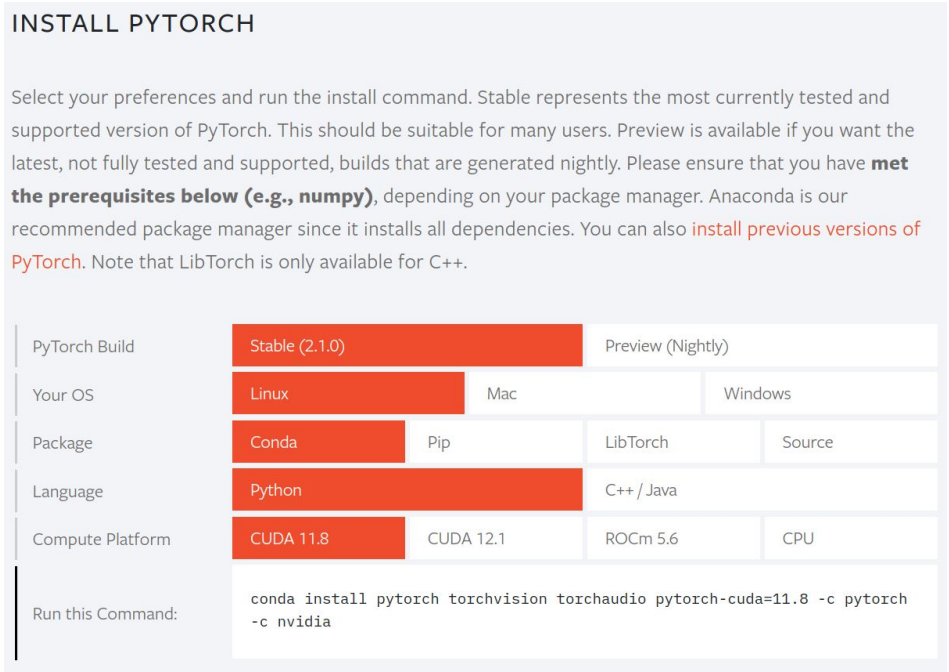
Miniconda is a free minimal installer for conda. It is a small bootstrap version of Anaconda that includes only conda, Python, the packages they both depend on, and a small number of other useful packages (like pip, zlib, and a few others). If you need more packages, use the `conda install` command to install from thousands of packages available by default in Anaconda's public repo, or from other channels, like conda-forge or bioconda.

Is Miniconda the right conda install for you? The Anaconda or Miniconda page lists some reasons why you might want one installation over the other.

https://docs.conda.io/projects/miniconda/en/latest/

Check out how to use conda to manage virtual environments, e.g., https://conda.io/projects/conda/en/latest/user-guide/getting-started.html

Install PyTorch inside Conda



https://pytorch.org/

Be careful of the CUDA version!

** **A100**: this gpu **ONLY WORKS** with cuda $\geq$ 11.3. (

** **V100**: this gpu **DOES NOT** work with cuda $\geq$ 11.3.

Prepare a job script

```
#!/bin/bash -l
#SBATCH --time=8:00:00   ## Specify your request computing time
                         ## 24:00:00 max for public queues
#SBATCH --nodes=1        ## Specify number of computing nodes to request
                         ## For most jobs in our research, we just need 1
#SBATCH --ntasks-per-node=4   ## Number of cpu cores to request.
                              ## Request more if:
                              ## 1) you need many cpu cores
                              ## for multi-thred computing
                              ## (e.g., parallelize for loops
                              ## for data pre-processing)
                              ## 2) multi-GPU (there may be min number of
                              ## cpu core requires for gpu connection
                              ## visit MSI website for detailed info.)

#SBATCH --mem=10g
#SBATCH --tmp=10g             ## Requesting storage (like cache and RAM)
                              ## to support your job computation

#SBATCH --mail-type=ALL       ## Receive all job notifications by email
#SBATCH --mail-user=$your_x500$@umn.edu   ## Your email address
                                          ## to view the notifications
```

```
#SBATCH --partition=interactive-gpu     ## Job queue to submit your job
#SBATCH --gres=gpu:a40:1    ## Number and type of gpu to request
                            ## Without this line,
                            ## you are not requesting any gpu
```

```
cd /home/csci5527/your_x500/your_working_folder
                            ## Direct your working directory

export PATH=/home/csci5527/$your_x500$/$your_conda$/
envs/$your_venv_name$/bin:$PATH    ## Load your conda env

python autoencoder_train.py    ## Execute your job
```

https://www.msi.umn.edu/content/job-submission-and-scheduling-slurm for more details

Which partition?

```
#SBATCH --partition=interactive-gpu
```

Partition Name	Max cores per node	Maximum walltime	Max available node memory[3] (GB)	Memory per core (MB)	Local scratch per node (GB)	Max nodes per job	Max jobs per user
msismall	128	96:00:00	248-755	3900	850	1	4000 [4]
msilarge	128	24:00:00	248-755	3900	850	32	4000 [4]
msibigmem	128	24:00:00	1995	15950	850	1	4000 [4]
msigpu [1]	24-128	24:00:00	374-1002	3900-8000	850	4	4000 [4]
msilong	32	37-0	248-755	3900	850	*** [5]	*** [5]
interactive	128	24:00:00	499	3900	850	2	1
interactive-gpu [1,6]	64-128	24:00:00	499-755	5120	850	1	1
interactive-long	64-128	37-0	499	3900	850	1	1
preempt [2]	128	24:00:00	248-755	3900	850	1	4000 [4]
preempt-gpu [1,2]	64-128	24:00:00	499-755	3900-8000	850	1	4000 [4]

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

Frequently used commands

Step 5: Submit your job

```
sbatch XXX.slurm  # XXX.slurm is your job script name
```

Step 6: Useful commands after submitting your job

(1) Checking the status of the MSI queue:

```
squeue -p interactive-gpu  # Lists all job statuses on the
                           # interactive-gpu partition (not only you)
```

Or simply checking your own queue status by:

```
squeue --me  # chek your own jobs
```

(2) Canceling a submitted job:

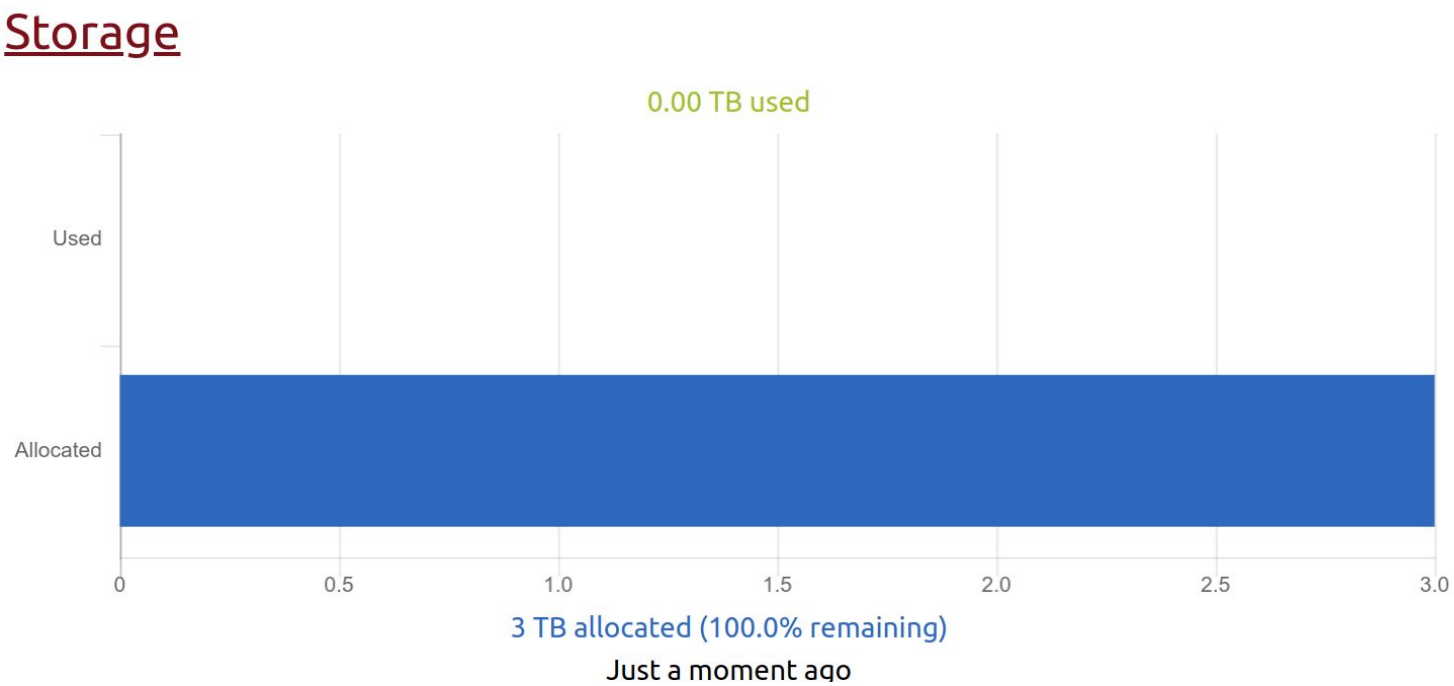
```
scancel jobIDnumber
```

(3) Canceling all your submitted jobs:

```
scancel --user=your_x500
```

Save the class storage

Storage/file counts shared for the whole class



https://www.msi.umn.edu/content/scratch-storage

Three levels of storage

Primary Storage

- Primary storage is where you will do most of your work
  - GNU/Linux filesystem
- Watch your disk usage:
  - [illegible]
- Impo[illegible]

/hom
Your

/hom
PI gr
refer

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

[illegible]pace is also not backed up by snapshot
[illegible] the tricks that exist to keep data in scratch for [illegible]an 30 days. Consider data on scratch as having no

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**

[illegible]obal/

Required: use scratch storage to store your large datasets and immediate results