



The Matilda HPC Cluster

November 19, 2021

John B. Johnston
Linux Systems Administrator
Oakland University

Matilda HPC: What is it?

- Computational hardware and support for Oakland University faculty, students & researchers
- 2000+ processing cores connected via high speed networking fabric, containing 690TB of high performance storage, and 12 NVIDIA GPUs
- Advanced job scheduling and resource management
- A RedHat Enterprise Linux environment accessed primarily via SSH (terminal)

Why Matilda HPC?

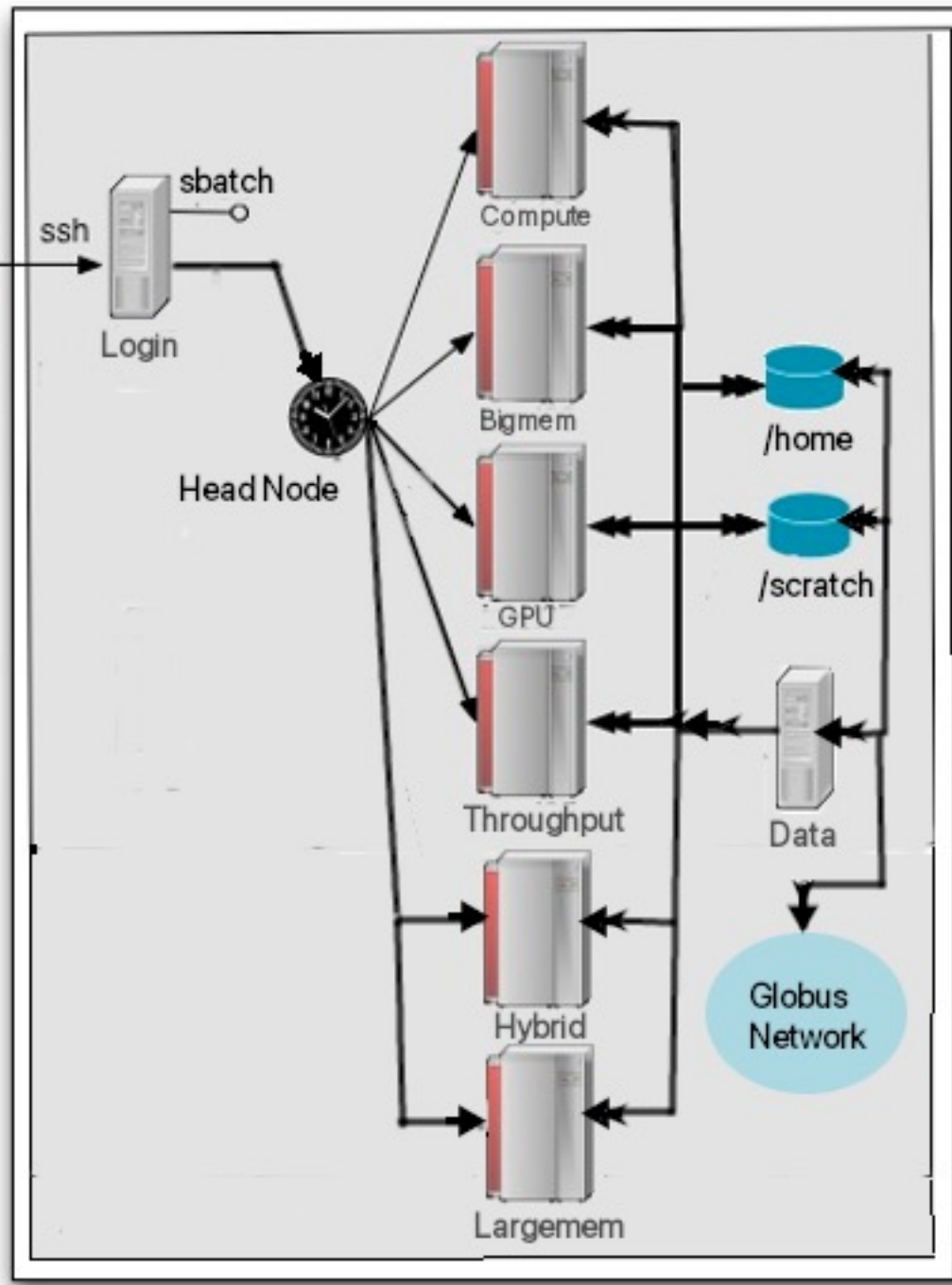
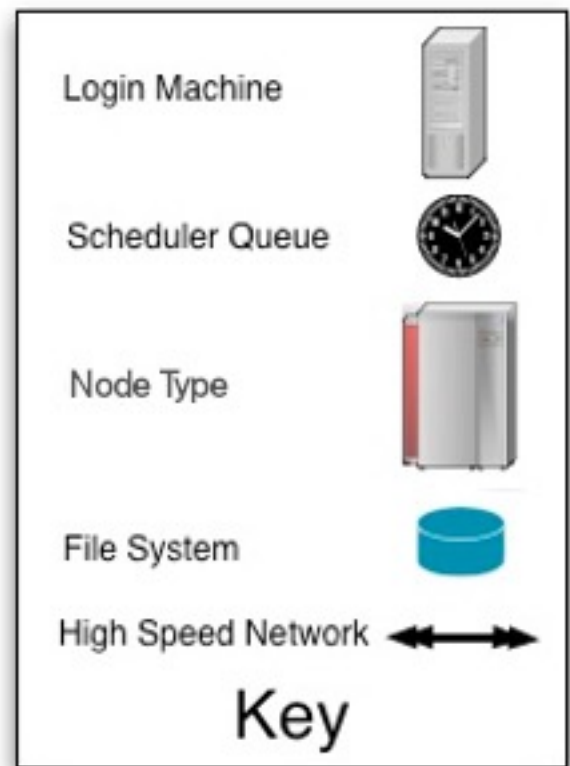
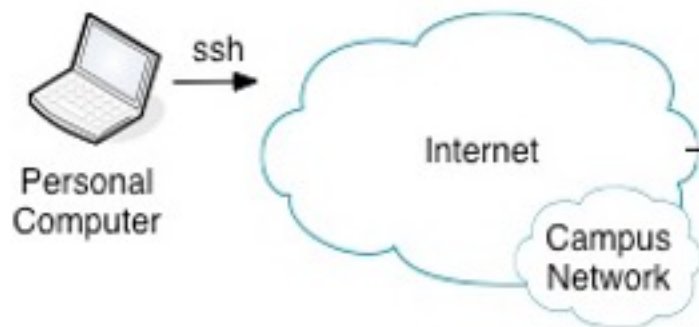
- Large data sets
- Lots of number crunching
- A need to run many simultaneous jobs with different data sets and/or configuration settings
- You need software you don't have, don't want to / can't setup
- Comprehensive ready-made development environment that is actively administered
- High-performance GPU processing capability

Why Matilda HPC?

- High speed data transfer between institutions with Globus
- CPU processing density
- Big memory machines up to 1.5 TB
- Resource management, scheduling (fire and forget)
- Storage capacity for big data
- Professionally built and maintained environment
- Secure

Matilda Cluster Nodes

Node Type	CPU Cores	RAM	Number Available	Other Features
compute	40	192GB	40	
throughput	8	192GB	10	
bigmem	40	768GB	4	
gpu	48	192GB	3	4 GPU Cards
hybrid	40	192GB	4	
largemem	40	1.5TB	1	
login	40	384GB	1	Gateway node
data	40	128GB	1	Globus data transfer node
head	40	384GB	1	Scheduling Control



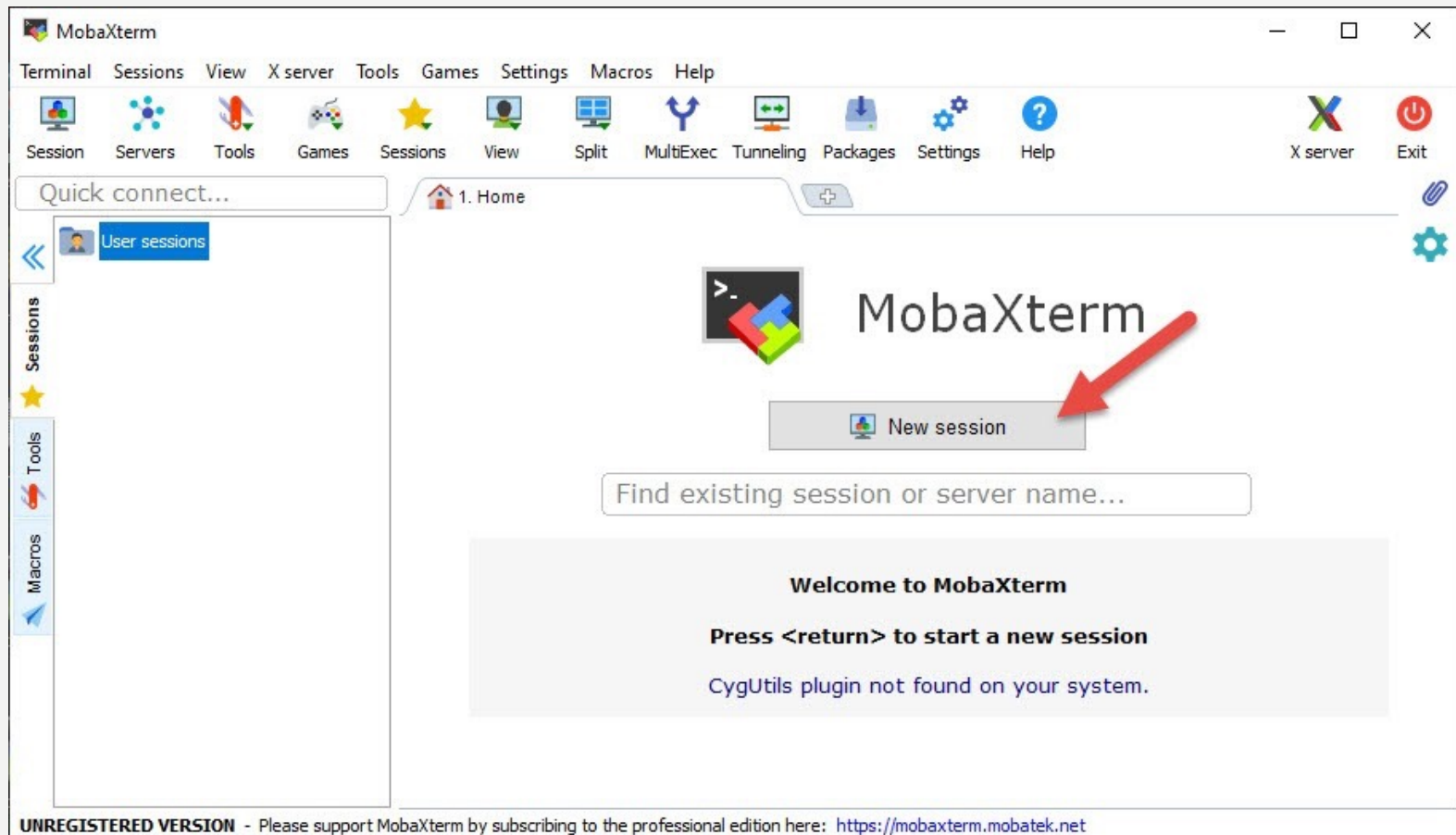
Interconnectivity

- The connections technology used between nodes, and filesystems
- Along with disk speed, this is one of the 2 primary bottlenecks in any computing system
- Gigabit – conventional ethernet technology, slow
 - 1 Gbs
- Infiniband – high throughput switched fabric
 - HDR100 – 100Gbs
 - All nodes are interconnected using infiniband
 - Scratch filesystem connected to all nodes using infiniband

Connecting to Matilda – The Very Basics for PC Users

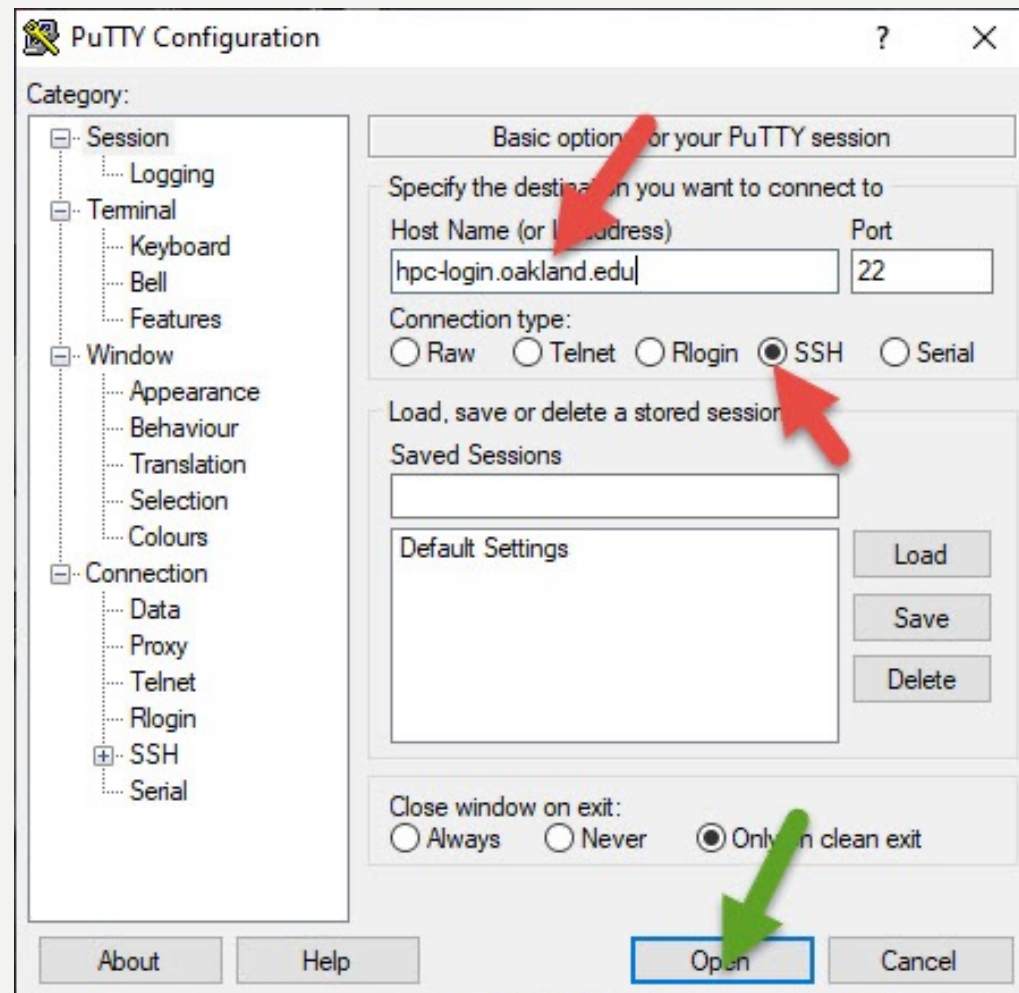
- MobaXterm:

<https://kb.oakland.edu/uts/HPCMobaXterm>



Connecting to Matilda – The Very Basics for PC Users

- PuTTY: <https://kb.oakland.edu/uts/HPCPuTTY>



Connecting to Matilda – The Very Basics for Mac Users

- Mac OS X is built atop a Unix derivative
- SSH, SCP, FTP, rsync ready
- Just open a terminal window!
- GUI clients available for file transfer – Filezilla, Cyberduck, etc.
- For X-windows (GUI) connections, may need to install XQuartz
- XQuartz:

<http://tinyurl.com/hpcc-xquartz>

The User Environment

- Primarily controlled by “Environmental Variables”
- Environmental variables are variables that are named in a way that are recognizable by a variety of applications, and which persist for the shell in which they are set.
- The “modules” system modifies these for you
- Useful for pre-installed applications
- “module av” – to see what’s available
- “module load <modulename> - to load a program into the environment.
- “module list” – to see what modules you have loaded

Software

- Matilda currently has over 240 software applications available for your use through the “modules” system.

<https://kb.oakland.edu/uts/HPCTModules>

- Installing Software:
 - If you don't see what you want (or you'd like a newer version), you can fill out a request via the UTS Help Desk (place “Matilda” somewhere in the subject):
<https://kb.oakland.edu/uts/HPCTModules>
 - You are free to install your own software in your /home or /project space
 - We have conda available for conda self-installations:
module load miniconda3

Software

- Some software cannot be run on Matilda HPC – for example, anything distributed as a Docker container
- In some cases, we can “unpack” the Docker distribution and install the software, although this can be challenging.
- Most software distributed exclusively as “conda” packages is best installed by the user in their working space. We have adapted some conda-based software to standard installs, but this isn’t also possible/practical
- If in doubt, just ask and we can evaluate your software request
- We do offer access to Jupyter Notebooks on Matilda

<https://kb.oakland.edu/uts/HPCJupyter>

Matilda Batch System

The Matilda HPC job system consists of three primary components:

1. Resource Manager – SLURM
2. Login Node – for submitting jobs
3. Compute Nodes – hpc-compute, hpc-gpu, hpc-throughput, etc.

The SLURM Resource Manager

A Resource Manager:

- Tracks and monitors cluster resources (compute nodes)
- Manages queues
- Monitors and controls jobs for adherence to global and job-specific limits within a single cluster
- Contains a facility for job launches, holds and terminations
- Provides a user command interface for interacting with jobs
- Has an interface to the cluster high-speed interconnect

Matilda Batch System Characteristics

- Single batch queue (defq)
- Single resource manager for domain
- A job script is generally required:

A job script is simply a text file containing instructions for the resource manager, and commands you would normally run on the command line

- You submit your job using the command:

```
sbatch <jobscriptname>
```

- You can check the status of your job using:

```
squeue
```


File Space Review

- /home

20GB of storage space

- /projects

Group project space with a quota of 1TB.

- /scratch

10TB for users and project groups. Lustre file system, somewhat faster. 45 day purge policy based on time file last accessed.

- Note: there are no backups!

Transferring Files

- “rsync” command - initiated from the local workstation, or a remote machine (like MSU HPCC). Only transfers files that aren’t already present in destination
- “scp” command – similar to rsync, but this “copies”, so it will overwrite files even if they exist at destination
- “Filezilla” – an SFTP GUI program for Windows and Mac. Drag and drop files.
- Globus – high-speed file transfer for large data sets. You can establish a connection directly between Matilda and MSU HPCC

<https://kb.oakland.edu/uts/HPCGlobus>

Matilda Online Resources

- <http://kb.oakland.edu/uts/ResearchComputingHPC> – UTS research computing home
- <http://help.oakland.edu> – file a ticket and request help, application installation, etc.
- https://join.slack.com/t/ouhpc/shared_invite/zt-yz508joh-mQXJbT5ohRmE0HG4~tGkBQ - Slack channel for real-time assistance

QUESTIONS?

John B. Johnston
jbjohnston@oakland.edu