



The Matilda HPC Cluster

November 11, 2022

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 Red Hat Enterprise Linux (RHEL) environment accessed primarily via SSH (terminal, or terminal client)

Why Matilda HPC?

- Lots of processing cores (CPUs)
- 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

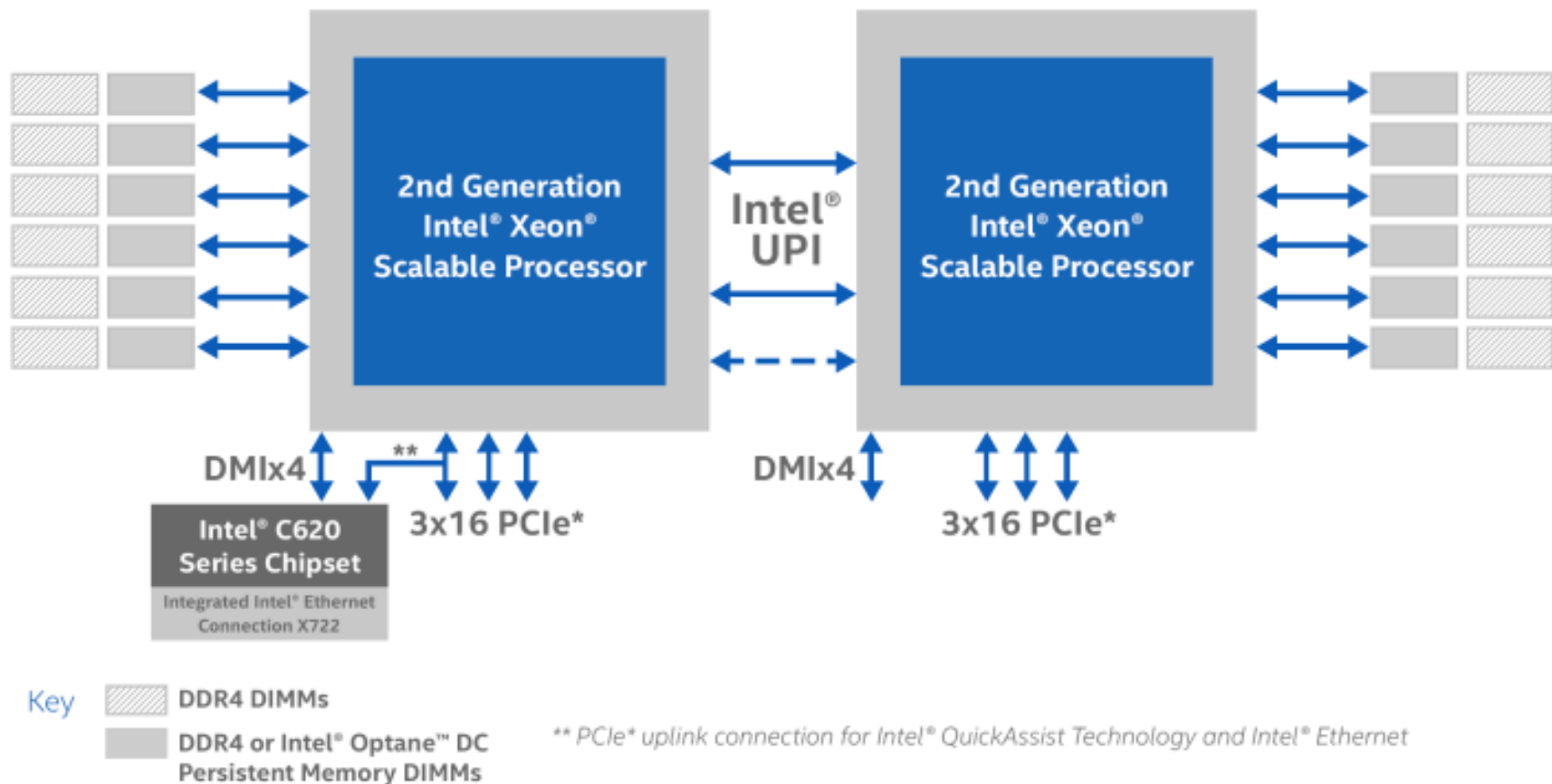
Why Matilda HPC (2)?

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

Terminology

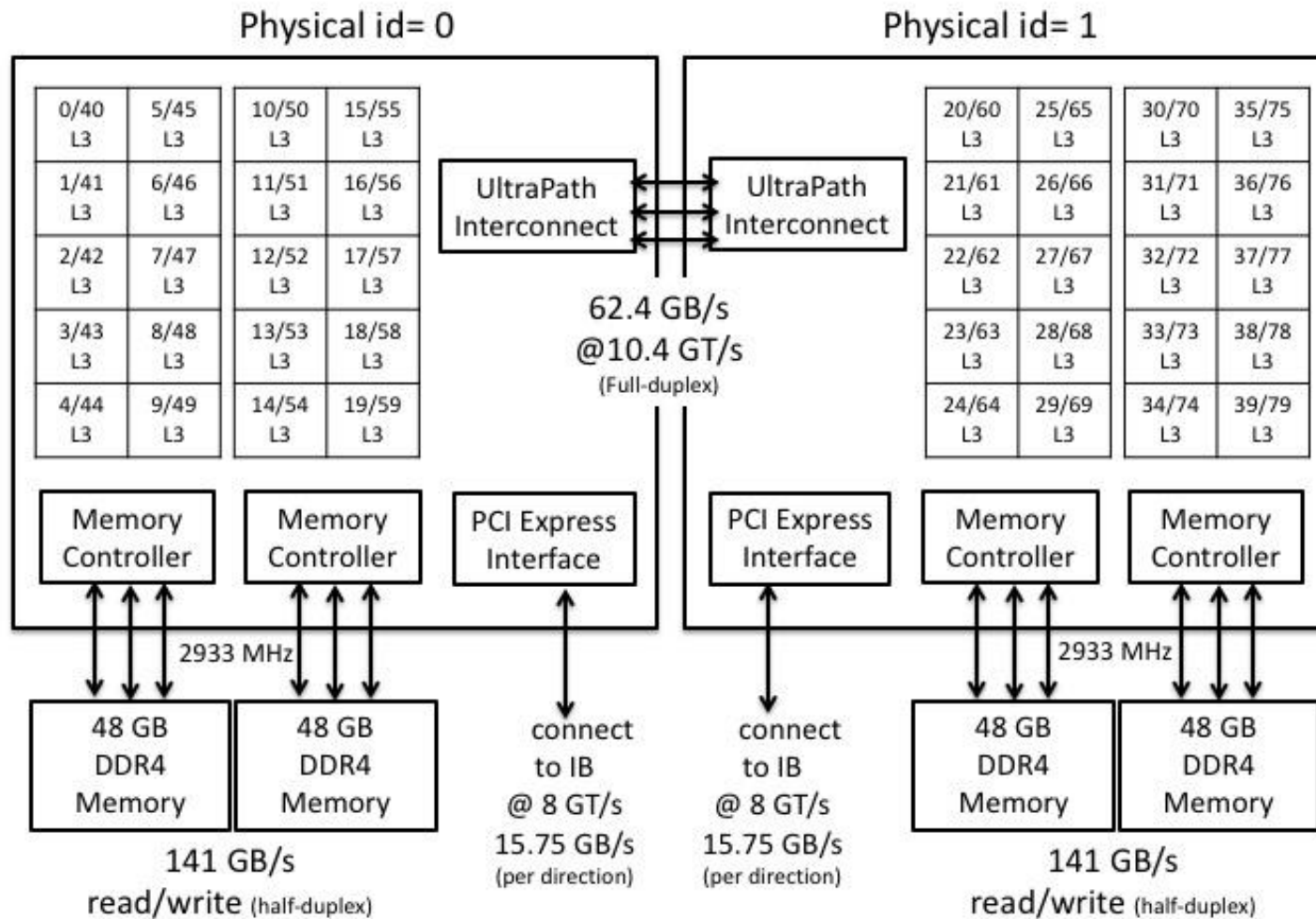
- Node – an enterprise-grade computer (server)
- CPU – the central processing unit (usually 2 per node)
- Core – a sub-processing unit contained within a CPU (usually 20 per CPU on Matilda)
- Job – a user task (program you run on data)
- Resource Manager (SLURM) - software that manages/controls your job(s)
- Memory – fast storage that holds data used by the cores (ephemeral).
- Storage – disk space where your data and programs live ("permanent" w/ qualifications).

Typical 2S Configuration



**Typical Matilda CPU
"Cascade" Processor
(source: Intel)**

Configuration of a Cascade Lake - SP Node



Cascade Lake CPU
Core Numbering
(source: NASA)

More Terminology

- Thread – a single sequential flow of program control
- Serial – a job that uses a single thread of control. Instructions are processed sequentially
- Multi-threaded - a job that is designed to run on more than one core on one node, with at least one thread per core. Instructions are processed in parallel
- MPI – a job that runs on multiple cores across more than one node. MPI jobs use distributed memory space. Messages must be passed between the nodes (MPI == Message Passing Interface) to coordinate the threads.

Important Take Aways...

- A program must be written to use multiple cores on a single node (OpenMP)
- A program must be written to use multiple cores on multiple nodes (MPI library, eg. OpenMPI, MPICH)
- If a program wasn't written to use multiple cores on one or more nodes, it is "Serial".
- It does you no good to:
 - assign >1 node to a job that can only run on one node,
 - assign >1 cores to a job that can only use one core.
 - This wastes resources and delays the start of your job.
- Not all problems can be readily parallelized
- Know your program and its capabilities!

Matilda Cluster Nodes

Node Type	CPU Cores	RAM	Number Available	Other Features
compute	40	192GB	40	
throughput	8	192GB	10	Faster CPUs
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
Total Processing Cores:			2024	

File Space Review

- /home (user space)
 - 50GB
- /projects (group project space)
 - 1TB
- /scratch (high speed temp. space)
 - 10TB for users/groups. 45-day purge policy.
- Snapshot system (.snapshot) – available on /home and /projects. Hourly and daily (up to 1 month) available.
- Directories are mounted (available) everywhere
- Note: there are no backups!

Interconnectivity

- Matilda uses high-speed (fiber optic) connections between nodes and some filesystems - "Infiniband"
- Along with disk speed, this is one of the 2 primary bottlenecks in any computing system
- Comparison:
 - Gigabit Ethernet (e.g. ISP) - 1Gbs
 - Typical WiFi – 0.1Gbs (100Mbs)
 - Infiniband (HDR100) – 100Gbs
 - All nodes are interconnected using Infiniband
 - Scratch filesystem connected to all nodes using infiniband

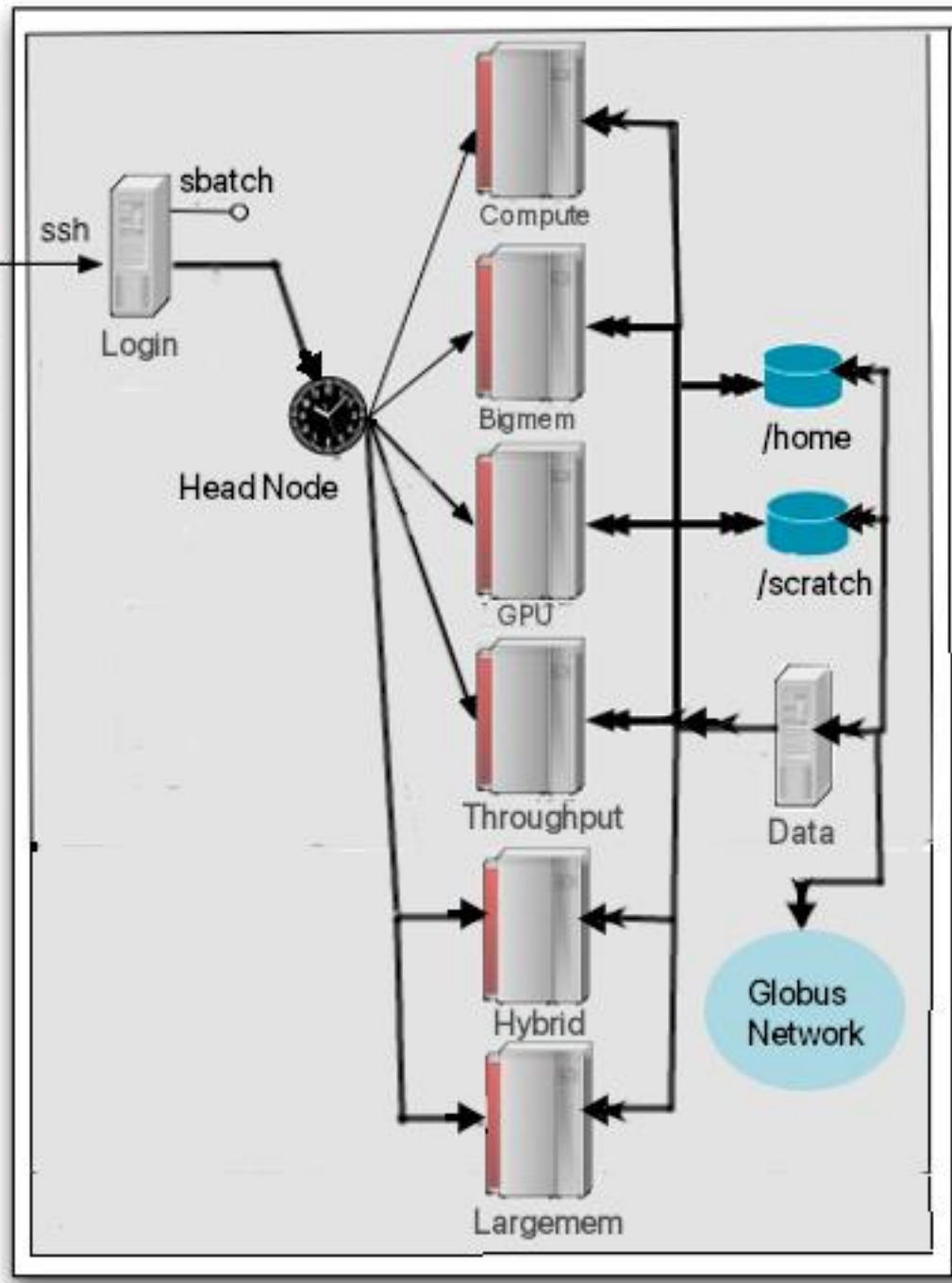
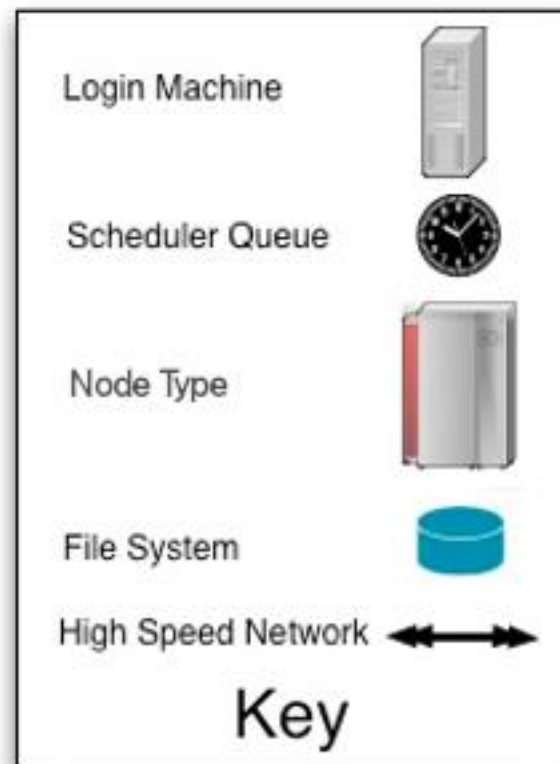
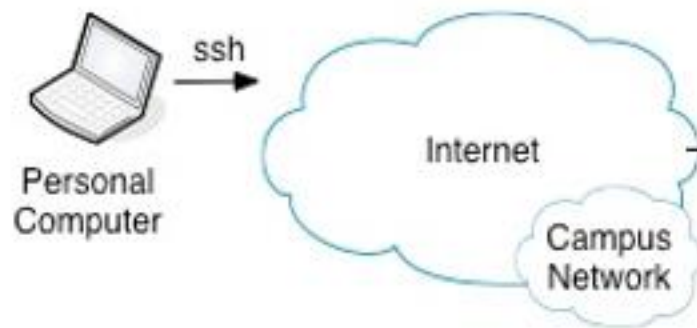
The SLURM Resource Manager

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

- SLURM Resource Manager
- Queue – the "line" for jobs.
 - Queues for short & long jobs, and "buy-in" users
 - Selected automatically for you based on job specifications.
- Single resource manager for domain
- Job Script - 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 -u <username>`



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

- We have >240 software titles available through the "modules" system (including MATLAB)
- Conda for customized user environments (DIY)
 - <https://kb.oakland.edu/uts/HPCConda>
- Jupyter Notebooks
 - <https://kb.oakland.edu/uts/HPCJupyter>
- Build your own in /home or /projects
- Some software cannot be run on Matilda – for example, Docker containers. We offer "Singularity" as a potential alternative.
- If in doubt, just ask: uts@oakland.edu

Connecting to Matilda

- Windows Clients (e.g. MobaXterm):
<https://kb.oakland.edu/uts/HPCMobaXterm>
- Mac (terminal)
- Linux (terminal, tux, etc.)
- SFTP Clients for file transfers (FileZilla, CyberDuck, etc.)

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. Many others.
- Globus – high-speed file transfer for large data sets. You can upload data, or share data with other institutions.

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

Matilda Online Resources

- <http://kb.oakland.edu/uts/ResearchComputingHPC> – UTS research computing home
- uts@oakland.edu – file a ticket and request help, application installation, etc.
- https://join.slack.com/t/ouhpc/shared_invite/zt-1ig4x3cvm-z33Sh_jC8HVmm0emG9ia0g - Slack channel for real-time assistance

QUESTIONS?

John B. Johnston
jbjohnston@oakland.edu