1. History of Hyalite, Configuration, How to connect
2. How to get data to/from Hyalite
3. How to get software on Hyalite
4. How to run jobs on Hyalite
Overview

1. History of Hyalite, Configuration, How to connect
2. How to get data to/from Hyalite
3. How to get software on Hyalite
4. How to run jobs on Hyalite
Cluster History/Configuration

- Initial Purchase December 2014
  - 36 compute nodes, ~200TB
- Current Configuration
  - 65 compute nodes, ~600TB
- 10GbE Network
- ~18TFlops

- Funded by:
  - Faculty Startup Packages
  - Grants
  - UIT, VPR, Provost
Compute Node == Computer

- Intel Xeon CPUs
  - Gen 1: Sandy Bridge
  - Gen 2: Haswell
  - Gen 3: Broadwell
  - 16 (32HT) or 20 (40 HT)

- System Memory
  - 64GB
  - 256GB
  - 1.5TB

- Operating System
  - CentOS 6.8 Linux

- Workload Manager
  - SLURM

- Cluster Management
  - Bright Cluster Manager
Shape of the Cluster

User Workstation

Data Transfer Node (Globus)

Hyalite Head Node
hyalite.rci.montana.edu
$HOME

Lustre Filesystem
$SCRATCH
$WORK
$STORE

Compute Nodes
compute001..compute065
• Documentation Website: http://www.montana.edu/rci/hyalite/
• Search for “hyalite” at montana.edu
• Quickstart Guide (HTML)
• Cheatsheet PDF
• Link to XDMoD
Connecting to Hyalite - Windows

- Windows
  - Use MobaXterm: [http://mobaxterm.mobatek.net/](http://mobaxterm.mobatek.net/)
  - Free version is sufficient
  - SSH, SFTP, X11 all in one package
Connecting to Hyalite – Mac/Linux

- Mac OS X (or MacOS)
  - Use Terminal and system SSH for console access
  - Install XQuartz (https://www.xquartz.org/) for X11

- Linux
  - Use Terminal and system SSH for console access
  - You are on your own for the rest
Connecting to Hyalite

Windows / GUI Client
• Host: hyalite.rci.montana.edu
• User: Your NetID
• Password: Your NetID Password
  • Set at https://password.montana.edu

Terminal (Mac, Linux)
$ ssh netid@hyalite.rci.montana.edu

If off-campus, you need to Use VPN to Connect!
4 Options for Editing Files

1. Windows: MobaXterm Built-in Editor
2. Command-Line (CLI) Editor (vim, emacs, nano, pico, joe, etc.)
3. SFTP Client “Edit File” feature
4. SSHFS Folder Mount
Editing Files - Use SFTP Client

- Most SFTP clients have an “edit file” feature
- File is downloaded to be edited
  - May not be trivial for large files
Editing Files - SSHFS

- Use SSHFS to mount the directory, and use your own editor
  - https://www.digitalocean.com/community/tutorials/how-to-use-sshfs-to-mount-remote-file-systems-over-ssh
  - Google search for “sshfs digitalocean”
- Commercial Tools (easier and better, but costs money)
  - ExpanDrive, $49 (http://www.expandrive.com)
  - Mountain Duck, $39 (https://mountainduck.io)
Overview

1. History of Hyalite, Configuration, How to connect
2. How to get data to/from Hyalite
3. How to get software on Hyalite
4. How to run jobs on Hyalite
2 Ways to Transfer Data

1. Use Traditional Tool
   • SFTP GUI Client
   • CLI tool (scp, rsync, etc)

2. Use Globus
1. Traditional File Transfer Tool

• 5x Slower than Globus, but good for:
  • Data <5GB
  • Source or Destination is on Wifi (wifi is slow)

• MobaXterm has sftp built-in

• Rsync is best for resume-able transfers
  • GUI versions are not free
  • CLI version is OSS
2. Globus

- Globus GridFTP
  - Requires local client
  - Fast (line speed ~1000Mbps for wired)
  - Rsync-like features (mirror, sync only changed)

- Show Example of Globus
Storage Locations on Hyalite

• Run: `storage-help` on Hyalite
• 4 Storage Locations
• Lustre Filesystem Locations:
$HOME

• Good for:
  • Executables (compiled code, etc)
  • Configuration and Batch Scripts
• NFS mounted from head node
  • Slow for data I/O
  • Bottleneck

• How to get there:
  • Default place when you log in
  • cd command with no arguments
  • ~ is an alias for your home directory
Lustre Filesystem Locations

- **$SCRATCH**
  - No quota, files purged after 90 days

- **$WORK**
  - Large quota (5TB), permanent

- **$STORE**
  - Small quota (1TB), permanent
    - Will be backed-up eventually
    - Globus $HOME is actually $STORE (DTNs mount STORE as HOME)
Takeaway

- If you have <= 1TB, use: `$STORE`
  - Alias: `cdt`
  - Location: `/mnt/lustrefs/store/<first.last>`

- If you have more than that, use: `$WORK`
  - Alias: `cdw`

- If you need a lot of storage, use: `$SCRATCH`
  - But it’s temporary!
    - Alias: `cds`
Overview

1. History of Hyalite, Configuration, How to connect
2. How to get data to/from Hyalite
3. How to get software on Hyalite
4. How to run jobs on Hyalite
Two ways to use software on Hyalite:

1. Install software yourself in your $HOME (or on lustre if you have >10G executables)
2. Use the module system to select pre-installed software
Install software yourself

- Use `$HOME` unless you have >10GB executables
  - Good to use `$WORK` for source/build directories, and your `$HOME` for the executables
  - Lustre is not great for executables (but will probably work)
- Use the modules to select a compiler if you need something other than the system gcc.
- This is pretty common.
Use Module system to select software

- Show example of using `module`

- Custom `module-search` (aliased to `ms`) script for searching module names

- Common operations:
  - `module load` (or `add`)
  - `module unload` (or `rm`)
  - `module list`

- Modules may load other modules

- Be careful when adding to your initialization (`module initadd`)
Overview

1. History of Hyalite, Configuration, How to connect
2. How to get data to/from Hyalite
3. How to get software on Hyalite
4. How to run jobs on Hyalite
SLURM workload manager

• SLURM: Simple Lightweight Universal Resource Manager
• Manages the jobs on the cluster
• DO NOT RUN ON THE HEAD NODE
  • (without permission)
• DO NOT RUN WITHOUT SLURM
  • (without permission)
  • Running directly on the compute nodes without SLURM will throw a “rogue process” alert
Status of the Cluster: sinfo and partitions

- **sinfo**
  - By default only shows your partitions
  - Running `sinfo -a` shows all partitions
  - `man sinfo`
  - Example...

- **Partitions are pools of hardware that jobs can be run on**
  - Default partition: compute001 – compute010
  - Partition 1: compute011 – compute014
  - Partition 2: compute003 – compute012
  - Partition 3: compute015

- **Which Nodes, and:**
  - Priority
  - Job runtime
  - Other Restrictions

- A partition is not a queue, but is often described that way
### Partitions on Hyalite

<table>
<thead>
<tr>
<th>Name</th>
<th>Nodes</th>
<th>Max Time</th>
<th>Priority</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>defq</td>
<td>compute[003-064]</td>
<td>24 Hours</td>
<td>500</td>
<td>Default Partition, for everyone</td>
</tr>
<tr>
<td>admin</td>
<td>compute[001,002]</td>
<td>48 Hours</td>
<td>1000</td>
<td>Admin tasks</td>
</tr>
<tr>
<td>express</td>
<td>compute[001,002]</td>
<td>30 Minutes</td>
<td>500</td>
<td>Testing and Training</td>
</tr>
<tr>
<td>priority</td>
<td>compute[012-064]</td>
<td>72 Hours</td>
<td>10000</td>
<td>Cluster Contributors</td>
</tr>
<tr>
<td>bigjob</td>
<td>compute[012-064]</td>
<td>30 Days</td>
<td>500</td>
<td>Long Running Jobs (access by request)</td>
</tr>
<tr>
<td>unsafe</td>
<td>compute[012-065]</td>
<td>Unlimited</td>
<td>1</td>
<td>Jobs that can sustain interruption, but can run anywhere.</td>
</tr>
<tr>
<td>xlarge</td>
<td>compute065</td>
<td>7 Days</td>
<td>10000</td>
<td>Large RAM/CPU nodes (restricted access)</td>
</tr>
</tbody>
</table>
Status of running/waiting jobs - squeue

- `squeue`
  - Again, `-a` shows all partitions
  - Aliases:
    - `sq` (all jobs)
    - `squ` (your jobs)
  - `man squeue`
- Example...

- All jobs will go into a single queue
- If there are resources available, the job runs immediately
  - every 5 minutes, jobs reprioritized to determine the next job to run.
- If a job is submitted with resources that are not available, it will wait forever.
  - Requesting too much RAM, CPU, etc.
- Estimates on when a job will start are often wrong.
<table>
<thead>
<tr>
<th>Job Priority Factors</th>
<th>Description</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fair-Share</td>
<td>Cluster usage relative to other waiting jobs</td>
<td>10000</td>
</tr>
<tr>
<td>Partition</td>
<td>Base Partition Priority</td>
<td>10000</td>
</tr>
<tr>
<td>Age</td>
<td>Time spent pending relative to other jobs</td>
<td>1000</td>
</tr>
<tr>
<td>Job Size</td>
<td>Resource request relative to other waiting jobs</td>
<td>1000</td>
</tr>
</tbody>
</table>
Other tools available for data

- **sprio**
  - Information on the priority calculation for jobs

- **scontrol, sacct**
  - Information/control over running/pending/finished jobs

- **scancel**
  - Kill a pending/running job

- **sview**
  - Graphical squeue/sinfo if you have X-windows

- Use `man <command>` for more information on running these commands
Running Jobs

• `sbatch`
  • Define job options, environment, and command in an `sbatch` file.
  • Run with the `sbatch` command
  • Most common way to run jobs on the cluster

• Show example `sbatch` job
• Show example `sbatch` run
Editing and Running a Job

cdt

cp ~/slurm-examples/hello.slurm.sh .

<edit hello.slurm.sh>

sbatch hello.slurm.sh
Interactive Jobs

- **srun, salloc**
  - Use these if you know what you are doing
- **X Forwarding** (running a GUI job)
  - Module load srun.x11
    - Use standard option flags for runtime, etc
    - X session is forwarded to your local machine
Thanks

• Email hyalite-support@montana.edu for support.