Hyak mox Overview

High Level Differences from ikt

1. Mox is an entirely separate cluster. They share nothing with one another.
2. You only get what you ask for, regardless of the resources available on the node. If you ask for 1 CPU, you'll only get one. If you ask for 1GB of RAM, you'll only get 1GB.
3. An allocation won't get the same set of nodes all the time, just access to the particular number of nodes to which they're entitled.
4. No occasional preemption in ckpt (formerly bf queue) for the moment.
5. Preempted jobs get 10s to do something smart before being killed and requeued.
6. Please report any problems to help@uw.edu with Hyak as the first word in the subject. Please also let us know you're using mox not ikt.

Connecting

SSH = mox.hyak.uw.edu

BBCP = mox1.hyak.uw.edu or mox2.hyak.uw.edu

File Transfers

Internal to Hyak systems

You can copy files at high speed without a password between the Hyak systems using commands like the ones below.

From ikt to mox

```
ikt1$ hyakbbcp myfile mox1.hyak.uw.edu:/gscratch/MYGROUP/
ikt1$ hyakbbcp -r mydirectory mox1.hyak.uw.edu:/gscratch/MYGROUP/
```

From mox to ikt

```
mox1$ hyakbbcp myfile ikt1.hyak.uw.edu:/gscratch/MYGROUP/
mox1$ hyakbbcp -r mydirectory ikt1.hyak.uw.edu:/gscratch/MYGROUP/
```

Filesystems

Scrubbed

There's a scrubbed temporary filesystem available at /gscratch/scrubbed. Files can be removed at any time, but they will be removed on a periodic basis based on creation date (files created 30 days ago or more). The maximum available space is 200TB. That space is shared among all users. Scrubbed storage is for temporary use only. Per month per TB persistent storage is available for purchase.

Quotas

Home Directory

```
mmlsquota gscratch:home --block-size G
```

Allocation Directory

```
mmlsquota -j <my short group> gscratch --block-size G
```

```
cat /gscratch/<my short group>/usage_report.txt
```

Slurm Primer

Show Queue

All Jobs

```
squeue
```

Jobs in Allocation

```
squeue -p <my short group>
```

All Jobs in ckpt (was bf)

```
squeue -p ckpt
```

Jobs in ckpt from Allocation

```
squeue -A <my short group>-ckpt
```

Submit

Own Allocation

Like ikt you're limited to the number of nodes your group purchased for your allocation.

```
sbatch -p <my short group> -A <my short group> test-job.sh
```

Checkpoint Allocation (formerly bf queue)

If you want access to all the nodes in the cluster, you need to use the ckpt partition.

```
sbatch -p ckpt -A <my short group>-ckpt test-job.sh
```

Interactive Session

Build Allocation - usage limited by core count and time

```
srun -p build --mem=100G --time=2:00:00 --pty bash -l
```

Own Allocation

```
srun -p <my short group name> -A <my short group name> --mem=100G --time=2:00:00 --pty bash -l
```

Show Allocation Information

Show all allocation information

```
hyakalloc
```

Show allocation and usage information for specified allocation

```
hyakalloc <my short group name>
```

Show Job Info

```
scontrol show job <jobid>
```

Show Node Info

```
scontrol show node <node>
```

Other Commands

Cancel Jobs

```
scancel <jobid>
```

```
scancel -u <username>
```

Show Job Resource Information

No Slurm Job Step (generally non-MPI jobs)

```
sstat <jobid>.batch
```

Slurm Job Step (generally MPI jobs)

```
sstat <jobid>
```

Sample Job Script

Change items that are bold and red. Check items that are bold and blue and change if necessary.

```
#!/bin/bash
## Job Name
#SBATCH --job-name=test-job
## Allocation Definition
#SBATCH --account=MYSHORTGROUP
#SBATCH --partition=MYSHORTGROUP
## Resources
## Nodes
#SBATCH --nodes=2
## Tasks per node (Slurm assumes you want to run 28 tasks, remove 2x # and adjust parameter if needed)
###SBATCH --ntasks-per-node=28
## Walltime (two hours)
#SBATCH --time=2:00:00
# E-mail Notification, see man sbatch for options


##turn on e-mail notification

#SBATCH --mail-type=ALL

#SBATCH --mail-user=your_email_address


## Memory per node
#SBATCH --mem=100G
## Specify the working directory for this job
#SBATCH --chdir=/gscratch/MYGROUP/MYUSER/MYRUN

module load icc_<version>-impi_<VERSION>
mpirun /gscratch/MYGROUP/MYMODEL/MYMODEL-BIN
```

Usage Reporting

The sreport command has many options for reporting. For more details refer to the man page. Below you will find two common examples. The usage is reported in TRES minutes which is the equivalent of core minutes. Using all the cores in one twenty-eight core node for one hour is 1680 TRES minutes.

Show formatted per-user utilization data for the month of June 2017 for specified accounts

```
sreport cluster AccountUtilizationByUser Accounts=<my short group name>,<my short group name>-ckpt Start=2017-06-
01 End=2017-07-01
```

Show easily parsable per-user utilization data for the month of June 2017 for all accounts

```
sreport --parsable --noheader cluster AccountUtilizationByUser Start=2017-06-01 End=2017-07-01
```

Other Documentation

How To

See articles on mox at below link (scroll down to links with prefix mox):

Hyak HOWTO

GNU parallel:

https://wiki.cac.washington.edu/display/hyakusers/Hyak+Serial+Job+Scripts

Hyak parallel-sql:

Hyak parallel-sql

Ikt Documentation

You can find additional documentation that applies to both mox and ikt on the main Hyak User Wiki.