Ocelote Quick Start

Overview

This is an abbreviated version of information mostly found on other pages intended for existing users who wish to use the new cluster.

Ocelote is our newest cluster. Lenovo's Nextscale M5 technology was the winner of the RFP mainly on price, performance and meeting our specific requirements. This cluster is actually the next generation of the IBM cluster we call ElGato. Lenovo purchased IBM's Intel server line in 2015.

Features:

- Intel Haswell V3 28 core processors
- 192GB RAM per node
- FDR infiniband for fast MPI interconnect
- New DDN 12KX storage array (all HPC storage is being integrated into one array)
- One large memory node with 2TB RAM, Intel Ivy Bridge V2 48 cores
- vSMP large memory image

One very different feature of this cluster is that there are 330 compute nodes that are the identical as opposed to the previous clusters that had different configurations for different functions. That makes some of the usage less complex.

Access

There are two ways to access Ocelote:

1. We have a web interface to Ocelote which is available at https://OnDemand.hpc.arizona.edu. There is more detail about this feature
2. We have a command line interface which is more familiar to existing users. We use a bastion host now for access to all clusters. This is required to access all clusters. There is a useful screenshot below.

When you log in for the first time, your home directory is created (see below).
Bastion Host Access:

~ > ssh <insert NetID>@hpc.arizona.edu
Password:
Duo two-factor login for netid

Enter a passcode or select one of the following options:

1. Duo Push to XXX-XXX-3614
2. Phone call to XXX-XXX-3614
3. SMS passcodes to XXX-XXX-3614 (next code starts with: 1)

Passcode or option (1-3):
Success. Logging you in...
Creating home directory for cmitts.
This is a bastion host used to access the rest of the environment.

Shortcut commands to access each resource
-----------------------------------------
Ocelote:     El Gato:
$ ocelote     $ elgato

We have also provided a menu. It can be enabled by typing:
$ menuon
[netid@login2 ~]$  

Allocations

1. **Compute Time.** You will use the same allocation of CPU hours that is available to you already. Use the command "va" to display the hours available to your group.
2. **Storage Space.**
   - Your home directory is increased from 6GB to 15GB and is backed up.
   - /extra/<NetID> is a new feature of the expanded HPC Storage system. Your allocation is 200GB and it is not backed up.
   - "uquota" will display your available space
3. **Xdisk.** xdisk has been simplified and greatly increased from a ~200GB maximum to 1TB. Using xdisk has all the details.
   - The expiration function of xdisk is "on" for Ocelote. Use /extra for long term storage.
4. **Buy-in storage.** Your existing filesets from the old clusters have been mounted on Ocelote. You can access them as before.
5. **File limits** have been implemented on some filesystems. /extra and /sgrps are limited to 600 files / GB

PBS

More details are at Running Jobs

We are still using PBS so much of what you are used to is the same. Here the significant differences
1. The common resources are 28 cores and 192GB RAM. We use pcmem=6gb. So for example, a job that needs a whole node uses:

```bash
#PBS -l select=1:ncpus=28:mem=168gb:pcmem=6gb
```

168GB is a lot less than 192GB of physical memory. You can go up to 188GB if you really need to, but 168GB lets the OS use more memory for file handling and other uses that support job performance.

2. We use modules but the cluster manager comes prebuilt with some modules and the names are different, so use the command "module avail" and note the different sections, to choose which module you need. There are variations on some of the common compiler, so name them uniquely when you load them. For example, "module load intel/mkl/64" is unique - you don't need the rest.

3. gcc 5.2 automatically loads for you. If you "module unload gcc", you will have gcc 4.4.7 available.

4. If you use a line like this in your scripts:
   ```bash
   source /usr/share/Modules/init/csh
   ```
   Leave it out. It does not apply

5. Do not use "#PBS -l jobtype". All the nodes are the same. You may get a "job rejected by all possible destinations" error if you include it.

6. For jobs which require more than 168GB of memory there is one Large Memory (2TB) node on Ocelote (see below).

**GPU Nodes**

Ocelote now has 46 nodes with Nvidia P100's. They are available for use on the "standard" and "windfall" queues. The standard queue will draw from the group's allocation of hours in the same way as the non-accelerated nodes. More details are at Running Jobs.

**Large Memory Node aka "Fat Node"**

All of the nodes have the same resources resources (28 cpus and 188GB of memory) except one which has 48 cpus and 2TB of memory. This is designed for those jobs that need a lot of memory but are not MPI-enabled.

The Large Memory node is restricted to the "standard" queue to better increase its availability. It was possible previously for it to be consumed by windfall jobs and standard jobs would wait a long time.

More details are here at: Large Memory Node