

# HiPerGator 2.0: Moving your research to the next generation of super computing

*Matt Gitzendanner*     *magitz@ufl.edu*



**UF** | Information Technology

OneIT for the #GatorGood


# HiPerGator

The University of Florida Supercomputer





**UF** | Information Technology


OneIT for the #GatorGood



**HiPerGator**  
The University of Florida Supercomputer

- 30,000 cores
- #115 in Nov 2015 Top 500 
  - #2 among US public universities
  - HPL RMAX 738 TFlops
- Dell PowerEdge
- Intel Xeon E5-2698v3 2.3GHz CPUs 

**UF** | Information Technology OneIT for the #GatorGood



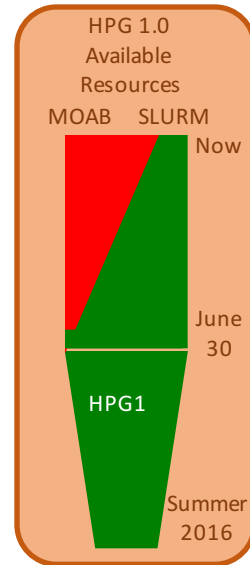
**CHANGE AHEAD**

- GatorLink Authentication
- SLURM Scheduler
- /ufrc unifies /scratch/lfs, /lts, /rlts

**UF** | Information Technology OneIT for the #GatorGood

## Transition Period

- Three places to send jobs:
  - HiPerGator 1.0:
    - MOAB/Torque
    - SLURM
  - HiPerGator 2.0:
    - SLURM

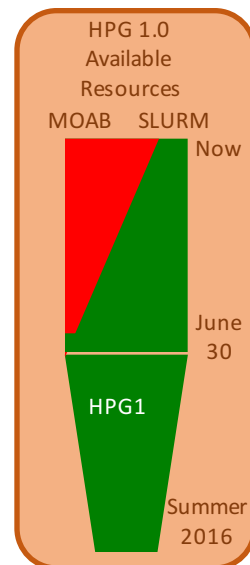


UF | Information Technology

OneIT for the #GatorGood

## Transition Period

- Three places to send jobs:
  - HiPerGator 1.0:
    - MOAB/Torque—No changes
      - Finishing project by 6/30
      - MPI applications not yet ready
    - SLURM
  - HiPerGator 2.0:
    - SLURM

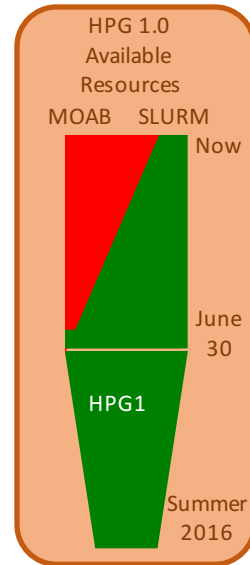


UF | Information Technology

OneIT for the #GatorGood

## Transition Period

- Three places to send jobs:
  - HiPerGator 1.0:
    - MOAB/Torque—No changes
    - SLURM
      - **module load slurm**
      - Rewrite job scripts to use SLURM
      - MPI applications need to be recompiled
  - HiPerGator 2.0:
    - SLURM

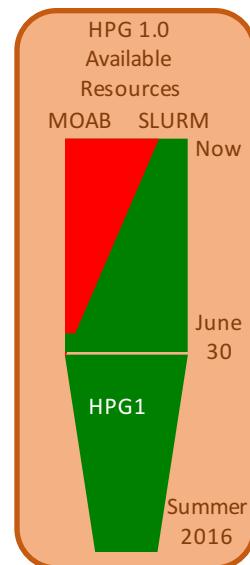


UF | Information Technology

OneIT for the #GatorGood

## Transition Period

- Three places to send jobs:
  - HiPerGator 1.0:
    - MOAB/Torque—No changes
    - SLURM
      - **module load slurm**
      - Rewrite job scripts to use SLURM
      - MPI applications need to be recompiled
  - HiPerGator 2.0:
    - SLURM



UF | Information Technology

OneIT for the #GatorGood

## SLURM

- Why change?
  - Functionality
  - Flexibility
  - Cost



UF | Information Technology

OneIT for the #GatorGood

## Moab to SLURM

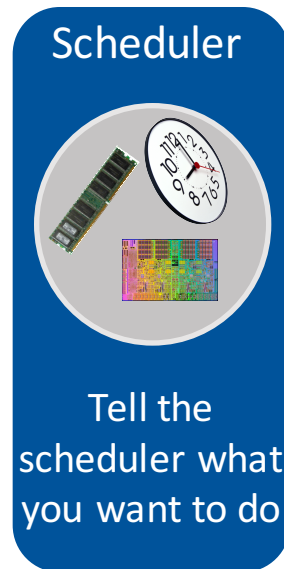
- Documentation
  - [PBS2Slurm Command Reference](#) wiki page
  - Other Wiki pages being developed

UF | Information Technology

OneIT for the #GatorGood

## Scheduling a job

- Need to tell scheduler what you want to do
  - **How many CPUs** you want and how you want them grouped
  - **How much RAM** your job will use
  - **How long** your job will run
  - The commands that will be run



## Basic SLURM job script

```
#!/bin/bash
#SBATCH --job-name=test      #A name for your job
#SBATCH -o job_%j.out       #Name output file
#SBATCH --mail-type=ALL     #What emails you want
#SBATCH --mail-user=<Email address> #Where
# } #SBATCH --ntasks=1       #Optional—single CPU
# } #SBATCH --mem-per-cpu=100mb #Per processor memory
# } #SBATCH -t=00:01:00      #Walltime in hh:mm:ss
#                               #or d-hh:mm:ss

hostname
module load python
python -v
```

## SLURM CPU Requests

- Nodes: **--nodes** or **-N**
  - Request a certain number of physical servers
- Tasks: **--ntasks** or **-n**
  - Total number of tasks job will use
- CPUs per task: **--cpus-per-task** or **-c**
  - Number of CPUs per task

HiPerGator 2.0 Compute Servers:

- 32 cores (2 X 16-core Intel Xeon CPUs)

## SLURM CPU Requests

- For single processor jobs
  - **--ntasks=1 (or omit)**
- For parallel jobs on a single node:
  - **--ntasks=8**

## SLURM CPU Requests

- For MPI jobs
  - **--ntasks=32**
  - Gets 32 cores for 32 MPI ranks
  - SLURM will determine node layout
- For Hybrid MPI/OpenMP jobs
  - **--nodes=4**
  - **--ntasks=4** (4 MPI ranks)
  - **--cpus-per-task=8**

## SLURM Memory Requests

- Memory: **--mem-per-cpu=1gb**
  - Can use mb or gb
  - Like Moab, no decimal values

### HiPerGator 2.0 Compute Servers:

- 128 GB total RAM (vs 256 GB on HPG1)
- Diskless servers: OS takes ~8GB RAM



## SLURM Time Request

- Time: `--time` or `-t`
  - 120 (minutes)
  - 2:00:00 (hh:mm:ss)
  - 7-0 (days-hours)
  - 7-00:00 (days-hh:mm)
  - 7-00:00:00 (days-hh:mm:ss)

## SLURM output/error files

- `#SBATCH -o output.file`
- `#SBATCH -e error.file`
- `#SBATCH -o output.file` #W/o -e  
combined
- Can also use `--output` and `--error`
  
- `#SBATCH -o JobFile.%j.out`
  - Use %j instead of \$SLURM\_JOBID

## SLURM

- Note that multi-letter directives are double-dash:
  - `--mail-type` `sbatch: error: distribution type 'ail-type=ALL' is not recognized`
  - `--ntasks`
  - `--mem-per-cpu`
- Do not use spaces with =
  - `--mail-user=magitz@ufl.edu` ✓
  - `--mail-user magitz@ufl.edu` ✓
  - not: `--mail-user= magitz@ufl.edu`

## Quality of Service (--qos)

- Each group has two QOS options
  - Investment QOS:
    - The NCUs the group has purchased
    - `--qos=group` (or leave off as this is default)
  - Burst QOS:
    - The burst capacity, available when idle resources are available on the cluster
    - `--qos=group-b`
- Unlike under MOAB/Torque the burst capacity is not automated

## SLURM Task Arrays

- **#SBATCH --array=1-200%10**
- Similar to Moab: range with % to limit number of jobs at a time
- **\$SLURM\_ARRAY\_TASK\_ID**
  
- Output file naming:
  - %A: job id
  - %a: task id
  - Output.%A\_%a.out

## HiPerGator 1.0

- Continue as now until June 30<sup>th</sup>
  - Dwindling MOAB/Torque resources available
- **module load slurm** and use SLURM
  - No need to move data or change login
  - Most applications will work unmodified






**HiPerGator**  
The University of Florida Supercomputer


- Making the switch to HiPerGator 2.0
  - Username needs to match GatorLink
    - If they match, you are set
      - [See information on web page](#)
    - If they are different
      - Open support request
      - Coordinate time to change RC name to GatorLink
        - No jobs running, logged out

UF | Information Technology OneIT for the #GatorGood



**HiPerGator**  
The University of Florida Supercomputer

- Making the switch
  - SSH Access:
    - **hpg2.rc.ufl.edu**
  - Data transfer to HPG2:
    - **sftp.rc.ufl.edu**



**Note:** Do not use login servers as data transfer hosts.

UF | Information Technology OneIT for the #GatorGood



**HiPerGator**  
The University of Florida Supercomputer

- Data transfer from HPG1
  - Do not transfer everything!
- Data transfer node
  - Login to **hpg2.rc.ufl.edu** then **ssh dtn1**
  - /scratch/lfs and /ufrc are mounted
  - **cp** or **rsync**
- Globus

UF | Information Technology OneIT for the #GatorGood

## Globus.org



- Fast transfer from /scratch/lfs/ to /ufrc
- Particularly good for large files
- ufrc#hpg2 and ufrc#go



The screenshot shows the Globus Connect Personal interface. At the top, there are navigation links: Manage Data, Publish, Groups, Support, and Account. Below that, there are tabs for Transfer Files, Activity, Endpoints, Bookmarks, and Console. The main area is titled "Transfer Files" and includes a prompt to "Get Globus Connect Personal" and "Turn your computer into an endpoint." There are two endpoint panels. The left panel has Endpoint "ufrc#hpg2" and Path "/ufrc/ufhpc/magitz/". The right panel has Endpoint "ufrc#go" and Path "/scratch/lfs/magitz/". Both panels show a list of files and folders with options to "select all", "select none", "up one folder", and "refresh list".

## Compilers

- HPG 2.0 and 1.0 under SLURM
  - intel/2016.0.109
    - openmpi/1.10.2
  - gcc/5.2.0
    - openmpi/1.10.2

Note that OpenMPI 1.10.2 on HPG1 is only compatible with SLURM. MOAB cannot run applications compiled with this module.

## Don't look back!

- 30,000 fast cores ready for you!



## End of free usage

- Previous policy
  - Up to 8-cores for free
- New policy
  - Research Computing has been told we can no longer offer any free access
  - Try-and-buy loans
    - 1-3 month loan of resources
    - Test the system
    - Verify needs are met
    - Become an investor

## Satisfaction Survey

- [training.it.ufl.edu](http://training.it.ufl.edu)

The screenshot displays the UF FIT Training website interface. At the top, there is a navigation bar with links for NEWS, CALENDAR, OFFICES & SERVICES, DIRECTORY, GIVING, UF HEALTH, and UF IFAS. A search bar on the right contains the text "Welcome to UF" and a search icon. Below the navigation bar, the main content area features a header with the UF FIT Training logo and navigation tabs for TRAINING, CANVAS BASICS, SERVICES (which is underlined), and CALENDAR. The main content area includes links for "UF Computing Help Desk" and "Contact Us". A prominent link for "Satisfaction Survey" is highlighted in orange. Below this, there is a section titled "NEW AND UPDATED" with the text "UFIT Training provides an extensive catalog of".

## Next Week:

- Writing SLURM Submission Scripts
  - 11:00am
  - NPB 2205

## Support

- Support requests



- [Web page](#) and [wiki](#)

### HiPerGator 2.0 Information

- [HiPerGator 2.0 Information](#)
- [SLURM Documentation](#)
- [Moab \(PBS\) to SLURM command reference](#)