

# Writing SLURM Submission Scripts

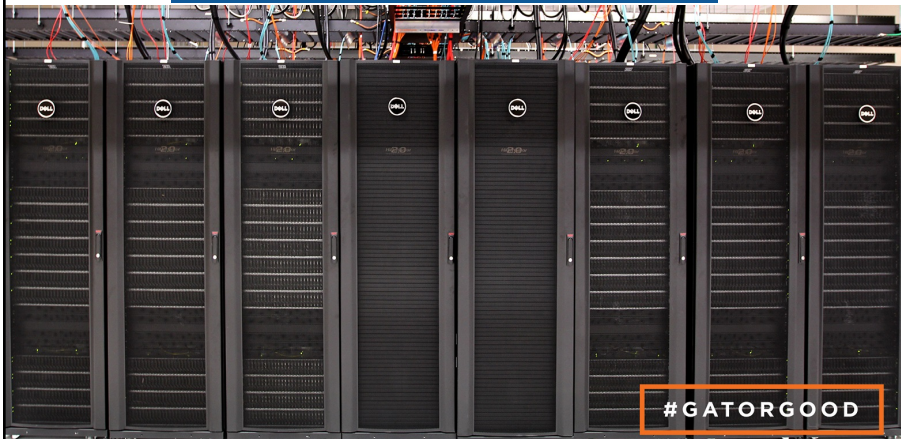
*Matt Gitzendanner*     *magitz@ufl.edu*



**UF** | Information Technology

OneIT for the #GatorGood


**HiPerGator**  
The University of Florida Supercomputer



#GATORGOOD

**UF** | Information Technology

OneIT for the #GatorGood

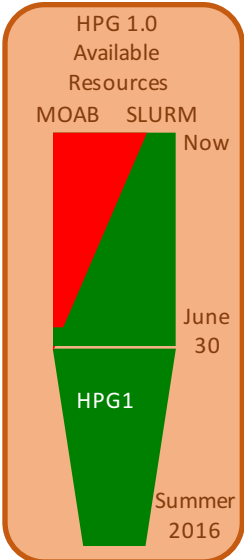


- GatorLink Authentication
- SLURM Scheduler
- /ufrc unifies /scratch/lfs, /lts, /rlts

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

## SLURM

- Why change?
  - Functionality
  - Flexibility
  - Cost



UF | Information Technology

OneIT for the #GatorGood

## Moab to SLURM

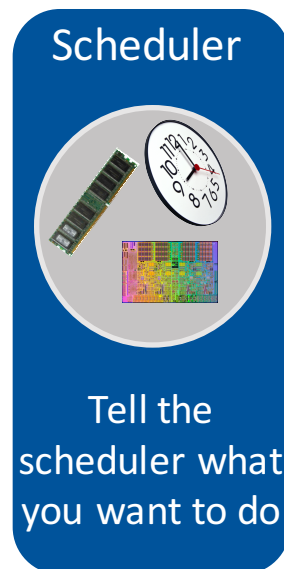
- Documentation
  - [PBS2Slurm Command Reference](#) wiki page
  - Other Wiki pages being developed

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:
  - **--cpus-per-task=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
  - **--ntasks=4** (4 MPI ranks)
  - **--cpus-per-task=8**
  - **(--nodes=4)** Not needed unless you really want 4 different nodes

## 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

## SLURM environment

- SLURM inherits your environment
  - This includes present working directory
    - Don't need `cd $SLURM_SUBMIT_DIR`
  - Modules that are loaded
  - **Be careful of conflicting modules**

## 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



UF | Information Technology

OneIT for the #GatorGood

## Emails

```

Job ID: 94392
Cluster: hipergator
User/Group: magitz/ufhpc
State: COMPLETED (exit code 0)
Nodes: 1
Cores per node: 4
CPU Utilization: 00:00:44
CPU Efficiency: 52.38% of 00:01:24 core-walltime
Memory Utilization 1.52 MB
Memory Efficiency: 0.04% of 4.00 GB
  
```

UF | Information Technology

OneIT for the #GatorGood

## Emails

```

Job ID: 5019
Cluster: hpg1
User/Group: magitz/ufhpc
State: CANCELLED (exit code 0)
Cores: 1
CPU Utilization: 00:00:00
CPU Efficiency: 0.00% of 00:00:00 core-walltime
Memory Utilization 1.26 MB
Memory Efficiency: 126.17% of 1.00 MB

```

Job error file:

```

slurmstepd: Job 5019 exceeded memory limit (1292 > 1024), being
killed
slurmstepd: Exceeded job memory limit
slurmstepd: *** JOB 5019 ON dev1 CANCELLED AT 2016-05-16T15:33:27
***


```

**UF** | Information Technology

OneIT for the #GatorGood

## Getting Logged in

- `ssh user@hpg2.rc.ufl.edu`



```

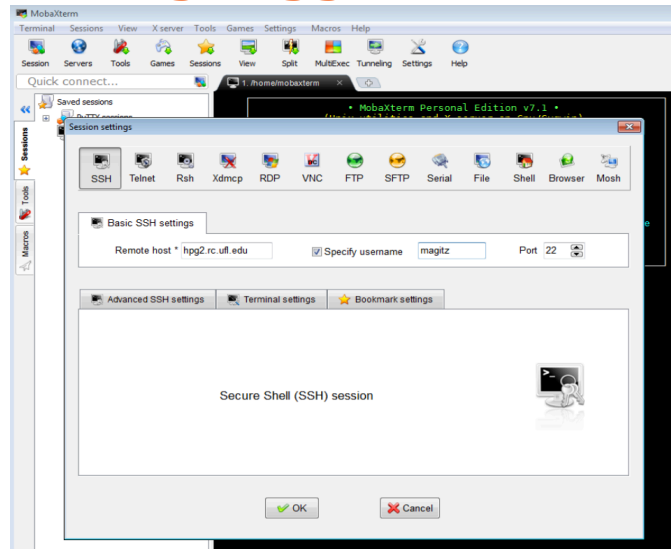
FLMNH-SOL-MAC1:~ gitz$ ssh magitz@hpg2.rc.ufl.edu
magitz@hpg2.rc.ufl.edu's password:

```

**UF** | Information Technology

OneIT for the #GatorGood

## Getting Logged In



## Development sessions

- Either:
  - `module load ufrc`
  - Followed by
    - `srundev`
    - `srundev -t 60:00`
- Or
  - `srun -p hpg2-dev --pty -u bash -i`
  - `srun -p hpg2-dev -t 60:00 --pty -u bash -i`

## Checking on jobs

- `squeue`
- `sacct`
  
- See <http://slurm.schedmd.com/>

## Example files

```
cd /ufrc/group/user/  
mkdir SLURM_examples  
cd SLURM_examples  
cp /ufrc/data/training/SLURM/*.sbatch .
```



# Satisfaction Survey

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

The screenshot shows the UF FIT Training website. The top navigation bar includes links for NEWS, CALENDAR, OFFICES & SERVICES, DIRECTORY, GIVING, UF HEALTH, and UF IFAS. Below this is a secondary navigation menu with TRAINING, CANVAS BASICS, SERVICES (underlined), and CALENDAR. The main content area features a 'Satisfaction Survey' link in orange text, along with 'UF Computing Help Desk' and 'Contact Us'. A footer section contains the UF Information Technology logo and the slogan 'OneIT for the #GatorGood'.

## Next Week:

- Open Q&A session
  - 11:00am
  - NPB 2205

The slide is titled 'Next Week:' and lists an 'Open Q&A session' with the following details: 11:00am and NPB 2205. The slide footer includes the UF Information Technology logo and the slogan 'OneIT for the #GatorGood'.

## Support

- Support requests



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

### HIPerGator 2.0 Information

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