



Intro to SLURM

& CS Clusters

Kevin Miles – Systems Engineer
UVa Computer Science

Talk Overview



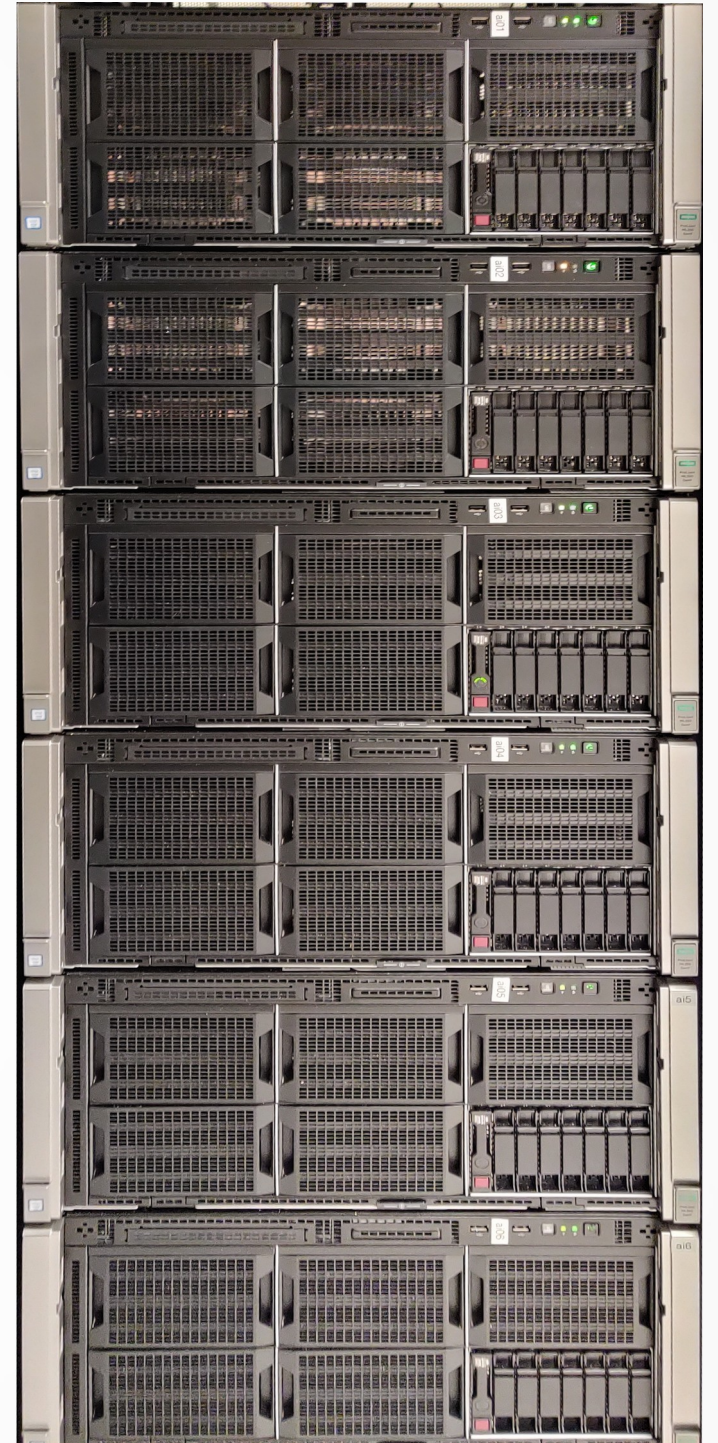
CS Computing

- What resources are available
- Where do I find software
- What OS am I using
- Using CUDA

SLURM

- What is SLURM
- Query cluster status
- Control jobs
- Scheduling
- Debugging tips

CS Computing



CS Computing



Resources

- Servers running CentOS Linux
 - Based on Red Hat Enterprise Linux (RHEL)
 - Shared “interactive” servers
 - portal.cs.virginia.edu
 - Submit jobs to cluster nodes from [portal.cs](http://portal.cs.virginia.edu) nodes
-
- Wide variety of available software & programming languages

Software

```
Last login: Tue Nov 12 09:29:51 2019 from kurma.cs.virginia.edu
ktm5j@portal01 ~ $ module avail

----- /sw/centos/Modules/systemModfiles -----
slurm-client

----- /sw/centos/Modules/modulefiles -----
afl                cudnn-7.0.5        java8u161          php
altera_pro         cudnn-7.1.1        java9              php7
anaconda3          cudnn-7.3.1        java9.0.4          php7.1.10
anaconda3-2019.10  cudnn-7.5.0        lammmps            python
antlr              dmd                lua                python2
atom               dmd-2.088.0        lua-5.3.4          python2.7.15
boost              doxygen            modelsim_ae        python3
boost-1.67.0       eclipse            modelsim_ase       python3.6.2
clang-llvm          emacs              nano               python3.8.0
clang-llvm-6.0.0   fio               nccl               qsys
clang-llvm-7.1.0   gcc               nccl-10.0          quartus
cmake               gcc-6.3.0          nccl-9.2           R
cmake-3.10          gcc-7.1.0          nios2eds           rstudio
cmake-3.15          gdb-8.1            nvtop              ruby
cs6620             ghex              openmpi            ruby2.5.1
cuda-toolkit        git               openmpi2.1.5       sbt
cuda-toolkit-10     gradle             openmpi4.0.0       scala
cuda-toolkit-8.0    hashcat           parallel           sqlite3
cuda-toolkit-9.0    java              perl               storm
cuda-toolkit-9.2    java11            perl-5.26          svn
cudnn               java8             perl-6             torch

----- /sw/linux-any/Modules/modulefiles -----
ant                cplex-cpooptimizer go                scons-3.0.1
antlr              cplex-opl          matlab           tmux
apktool            cplex-studio-all  ripgrep          tmux-2.7
cplex              dmd                sbt              vscode
cplex-concert      dmd-2.088.0        scons
ktm5j@portal01 ~ $
```

CS Computing

Software

- Software can be loaded/unloaded using `module` command
- Software stored in network volume
 - mounted to `/sw`
 - filesystem tuned for read access
 - noatime, zfs tweaks, async
 - don't waste space/bandwidth storing software in your home directory
- Send software requests to `csheelpdesk@virginia.edu`

CS Computing

Software -- Modules

- Modules control environment variables like
 - PATH
 - LD_LIBRARY_PATH
 - LD_INCLUDE_PATH
 - MANPATH
- Primarily designed for `bash`
- Support for alternate shells
 - `zsh`
 - `csh`
 - `tcsh`
- Support for scripting languages
 - `perl`
 - `python`
 - `ruby`

CS Computing

Software -- Modules

- If using a non-bash shell you must run init script manually
 - `/etc/profile.d/modules.sh` won't work
 - Init scripts found in `/sw/centos/Modules/current/init`
- More shells coming soon
 - Modules version 4.3 supports `fish`, `ksh`, `tcl` and `R`
- You can write your own modules!
 - append path to custom modfiles: `module use -a ~/local/dir`

Hardware



CS Computing

GPGPU

- GPU equipped nodes available inside & outside of Slurm
 - `gpusrv01` - `gpusrv06` available outside of Slurm
 - Cuda-toolkit available in multiple versions
 - Be sure that you have module loaded
 - `module load cuda-toolkit`
-
- cuDNN module also available

CS Computing

GPGPU

- Python3 and Anaconda3 both available
- Including GPU support
- Can create virtualenv's from either python base
 - inside you can use pip or conda to install your own packages
- Popular packages already available
 - torch
 - pytorch
 - tensorflow-gpu

SLURM



SLURM



Using Slurm -- Overview

- Viewing cluster status
- Viewing job status
- Submitting jobs
- Debugging

SLURM



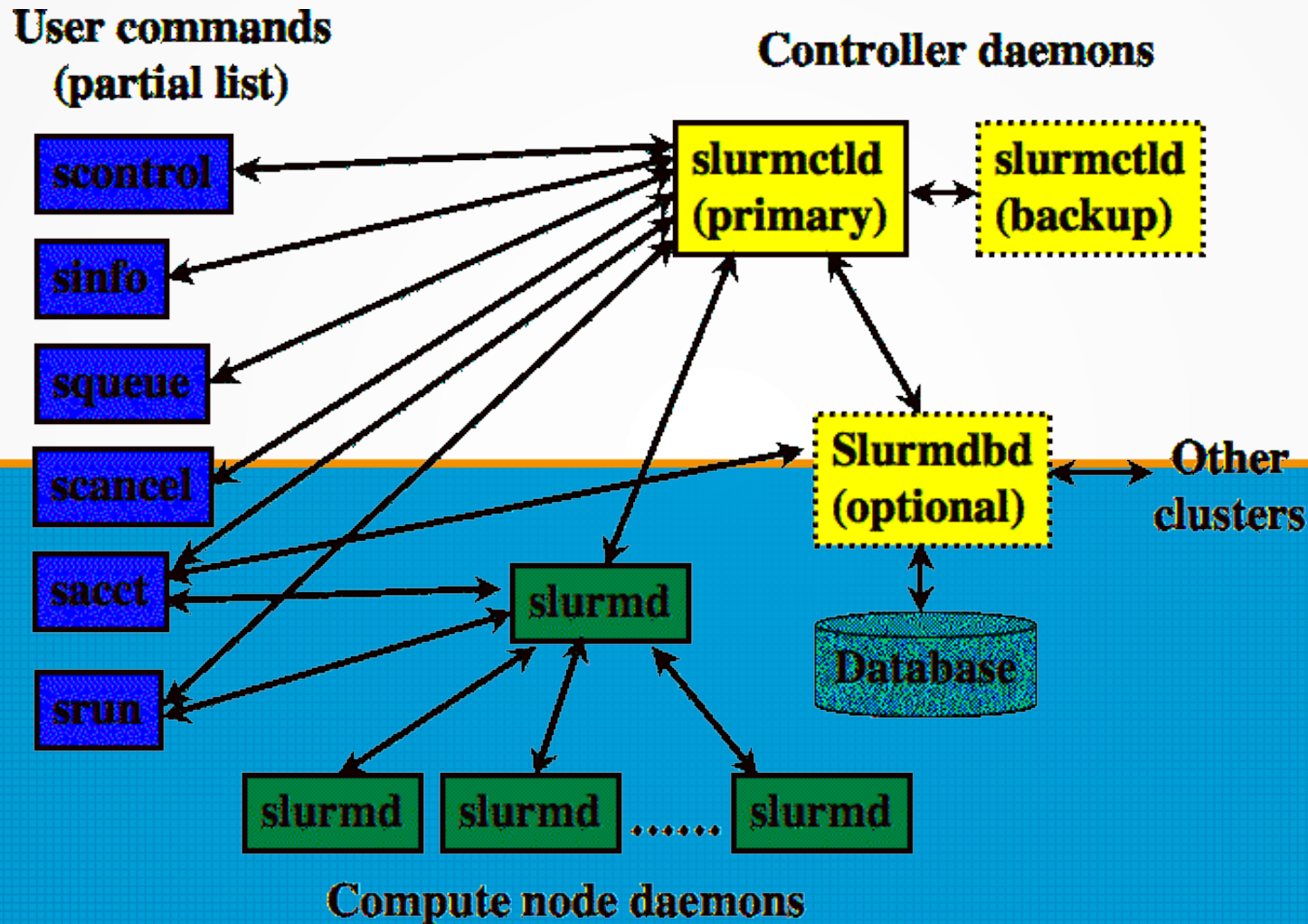
Using Slurm -- Access Model

- Compute servers are non-accessible via `ssh`
 - login access to server subverts resource allocation
- Interactive session still possible through `srun /bin/bash`
- Need GPUs for testing? Use `gpusrv` servers

SLURM



Using Slurm -- Access Model

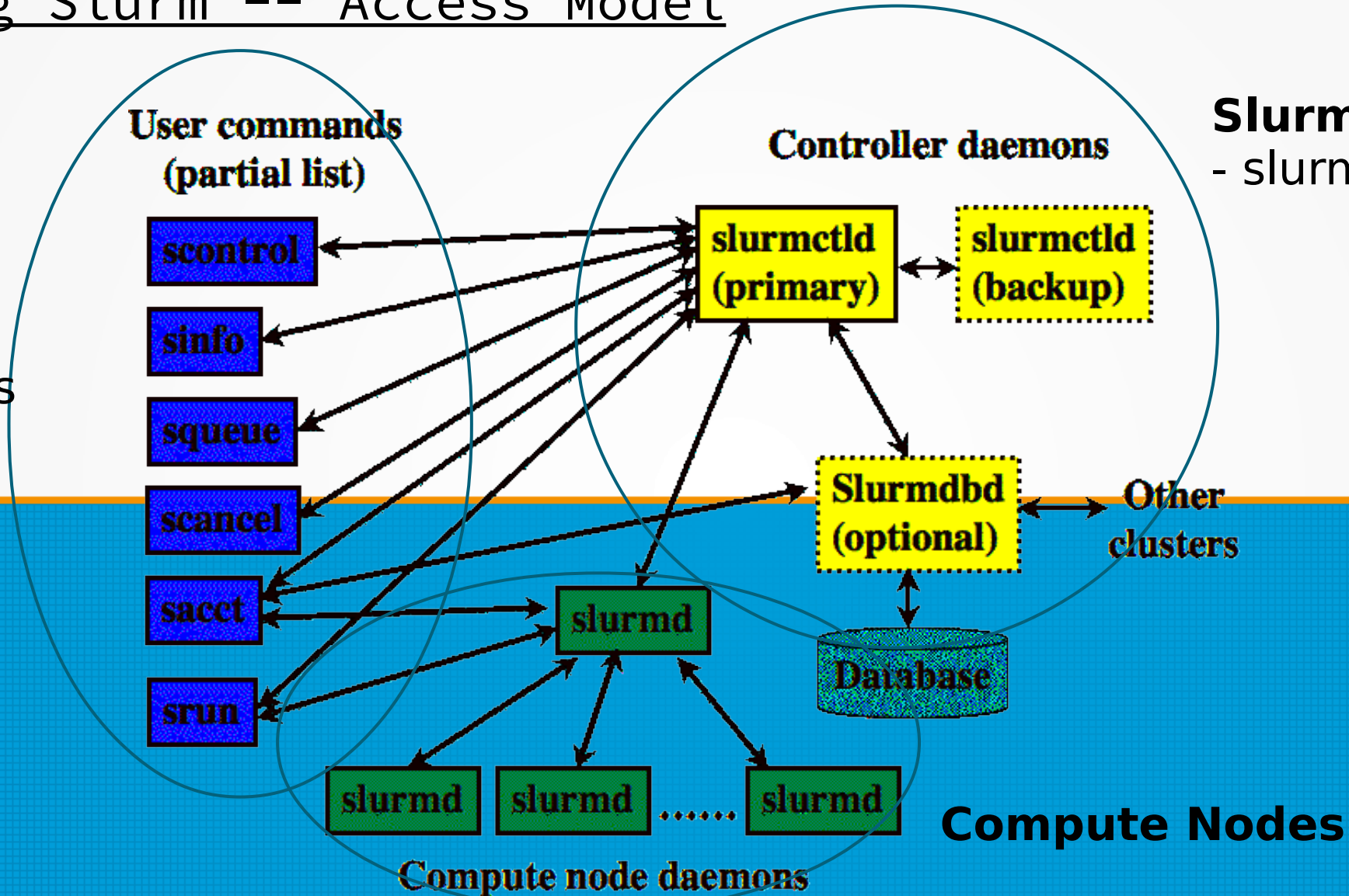


SLURM



Using Slurm -- Access Model

Slurm Clients
- portal.cs



Slurm Controller
- slurm-master.cs

SLURM



Using Slurm -- Commands Overview

- `sinfo` - list information about cluster nodes
- `squeue` - list currently queued jobs
- `srun` - quickly submit single command job
- `sbatch` - submit batch script job
- `scontrol` - admin tool, some commands for non-root users

Check the wiki for command line cheat sheet

Gathering Info



SLURM



Gathering Info -- sinfo

```
ktm5j@portal01 ~ $ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
main*      up      infinite    3     mix granger[5,8],trillian1
main*      up      infinite   22     idle artemis2,granger[2-4,6-7],lynx[08-12],nibbler[1-4],slurm[1-5],trillian[2-3]
falcon     up      infinite    8     idle falcon[1-6,8-9]
intel      up      infinite    2     mix granger[5,8]
intel      up      infinite   19     idle granger[2-4,6-7],lynx[08-12],nibbler[1-4],slurm[1-5]
amd        up      infinite    1     mix trillian1
amd        up      infinite    4     idle artemis[2,4],trillian[2-3]
centos     up      infinite    3     mix granger[5,8],trillian1
centos     up      infinite   22     idle artemis2,granger[2-4,6-7],lynx[08-12],nibbler[1-4],slurm[1-5],trillian[2-3]
share      up      infinite    5     idle lynx[08-12]
gpu        up      infinite    1  drain* ai06
gpu        up      infinite    2  drain artemis[6-7]
gpu        up      infinite    9     mix ai[01,03-05],lynx[01-02,05-07]
gpu        up      infinite    1  alloc lynx03
gpu        up      infinite    6     idle artemis[4-5],lynx04,ristretto[01,03-04]
```

SLURM



Gathering Info -- squeue

```
ktm5j@portal01 ~ $ squeue
```

JOBID	PARTITION	NAME	USER	ST	TIME	NODES	NODELIST(REASON)
1245851	main	bbv_gen	ab4cd	R	10-02:18:38	1	trillian1
1249230	gpu	bash	ab4cd	R	1-17:11:09	1	ai01
1249232	gpu	bash	ab4cd	R	1-17:01:12	1	ai03
1249233	gpu	bash	ab4cd	R	1-16:59:24	1	ai04
1249234	gpu	bash	ab4cd	R	1-16:54:41	1	lynx05
1249235	gpu	bash	ab4cd	R	1-16:52:43	1	lynx06
1249236	gpu	bash	ab4cd	R	1-16:51:34	1	lynx07
1249237	gpu	bash	ab4cd	R	1-16:49:07	1	ai05
1249808_28	main	PagingGr	wxy7z	R	8:02:39	1	granger8
1249808_30	main	PagingGr	wxy7z	R	8:02:39	1	granger8
1249808_31	main	PagingGr	wxy7z	R	8:02:39	1	granger8
1249808_96	main	PagingGr	wxy7z	R	8:02:39	1	granger5

SLURM



Gathering Info -- partitions

- Slurm partitions separate compute nodes into groups
- GPU enabled nodes are found in the **gpu** partition

```
ktm5j@portal01 ~ $ scontrol show partition | grep PartitionName  
PartitionName=main  
PartitionName=falcon  
PartitionName=intel  
PartitionName=amd  
PartitionName=centos  
PartitionName=share  
PartitionName=gpu
```

- **srun** commands: **-p gpu**
- **sbatch** commands: **#SBATCH --partition=gpu**

Using SLURM



SLURM



Using Slurm -- Modules

- Slurm will execute jobs outside of **login shell**
- Modules won't be available unless init called
 - Add to scripts: `source /etc/profile.d/modules.sh`

SLURM



Using Slurm -- srun

```
ktm5j@portal01 ~ $ srun --help
Usage: srun [OPTIONS...] executable [args...]
```

Common Options

-c, --cpus-per-task=<ncpus>

Request that ncpus be allocated per process.

--gpus-per-task=[<type>:]<number>

Specify the number of GPUs required for the job on each task to be spawned

--gres=<list>

Specifies a comma delimited list of generic consumable resources.

-N, --nodes=<minnodes[-maxnodes]>

Request that a minimum of minnodes nodes be allocated to this job.

SLURM



Using Slurm -- srun interactive

```
ktm5j@portal01 ~ $ srun -w granger2 --pty bash -i -l -  
ktm5j@granger2 ~ $ hostname  
granger2.cs.virginia.edu
```

- Interactive sessions are useful for debugging
- Please don't leave open sessions

SLURM



Using Slurm -- sbatch

```
ktm5j@portal01 ~ $ batch --help
Usage: sbatch [OPTIONS...] executable [args...]
```

Common Options

-c, --cpus-per-task=<ncpus>

Request that ncpus be allocated per process.

--gpus-per-task=[<type>:]<number>

Specify the number of GPUs required for the job on each task to be spawned

--gres=<list>

Specifies a comma delimited list of generic consumable resources.

-N, --nodes=<minnodes[-maxnodes]>

Request that a minimum of minnodes nodes be allocated to this job.

- Command arguments can be placed in command line, or in batch file header
- sbatch and srun use the same arguments

SLURM



Using Slurm -- sbatch files

```
#!/bin/sh  
  
#SBATCH --time=1  
  
/bin/hostname
```

- Sbatch files are shell scripts
- Can include Slurm arguments in file
 - argument lines begin with #SBATCH
- Can also include arguments in command line

SLURM



Using Slurm -- sbatch example

```
ktm5j@portal03 ~/slurm $ module load java
ktm5j@portal03 ~/slurm $ ls
batch.sh  hello.java
```

SLURM



Using Slurm -- sbatch example

```
ktm5j@portal03 ~/slurm $ module load java
ktm5j@portal03 ~/slurm $ ls
batch.sh  hello.java
ktm5j@portal03 ~/slurm $ cat hello.java
class Hello {
    public static void main(String args[]) {
        System.out.println("Hello from SLURM!");
    }
}
ktm5j@portal03 ~/slurm $ javac hello.java
```

SLURM



Using Slurm -- sbatch example

```
ktm5j@portal03 ~/slurm $ module load java
ktm5j@portal03 ~/slurm $ ls
batch.sh  hello.java
ktm5j@portal03 ~/slurm $ cat hello.java
class Hello {
    public static void main(String args[]) {
        System.out.println("Hello from SLURM!");
    }
}
ktm5j@portal03 ~/slurm $ javac hello.java
ktm5j@portal03 ~/slurm $ cat batch.sh
#!/bin/bash

#SBATCH --odelist=ai02
#SBATCH --partition=gpu

source /etc/profile.d/modules.sh

module load java

java Hello
```


SLURM



Using Slurm -- sbatch example

```
ktm5j@portal03 ~/slurm $ module load java
ktm5j@portal03 ~/slurm $ ls
batch.sh  hello.java
ktm5j@portal03 ~/slurm $ cat hello.java
class Hello {
    public static void main(String args[]) {
        System.out.println("Hello from SLURM!");
    }
}
ktm5j@portal03 ~/slurm $ javac hello.java
ktm5j@portal03 ~/slurm $ cat batch.sh
#!/bin/bash

#SBATCH --odelist=ai02
#SBATCH --partition=gpu

source /etc/profile.d/modules.sh

module load java

java Hello

ktm5j@portal03 ~/slurm $ sbatch batch.sh
Submitted batch job 1249913
```

SLURM



Using Slurm -- sbatch example

```
ktm5j@portal03 ~/slurm $ module load java
ktm5j@portal03 ~/slurm $ ls
batch.sh  hello.java
ktm5j@portal03 ~/slurm $ cat hello.java
class Hello {
    public static void main(String args[]) {
        System.out.println("Hello from SLURM!");
    }
}
ktm5j@portal03 ~/slurm $ javac hello.java
ktm5j@portal03 ~/slurm $ cat batch.sh
#!/bin/bash

#SBATCH --odelist=ai02
#SBATCH --partition=gpu

source /etc/profile.d/modules.sh

module load java

java Hello

ktm5j@portal03 ~/slurm $ sbatch batch.sh
Submitted batch job 1249913
ktm5j@portal03 ~/slurm $ ls
batch.sh  Hello.class  hello.java  slurm-1249913.out
ktm5j@portal03 ~/slurm $ cat slurm-1249913.out
Hello from SLURM!
ktm5j@portal03 ~/slurm $
```

SLURM



Using Slurm -- filesystem access

- Job files (sbatch, scripts, executables, input data) must be on shared storage
 - eg. /u, /p or /bigtemp
 - Your job can use local storage for tmp files

Final Notes

Final Notes



Online Resources

- CS Wiki
 - <https://www.cs.virginia.edu/wiki/>
 - https://www.cs.virginia.edu/wiki/doku.php?id=compute_slurm
- Slurm Website
 - <https://slurm.schedmd.com>
 - <https://slurm.schedmd.com/tutorials.html>
- Web Search is your friend!

SLURM Problems?

- Submit a Helpdesk Ticket via cshelpdesk@virginia.edu