



Open Science Grid

Software Modules and Licenses

Christina Koch (ckoch5@wisc.edu)

Research Computing Facilitator

University of Wisconsin - Madison

Expanding Our Horizons

- Previously, we were using simple, open source code and building software ourselves.
- This presentation discusses two different cases:
 - Using software modules
 - Licensed software



MODULES

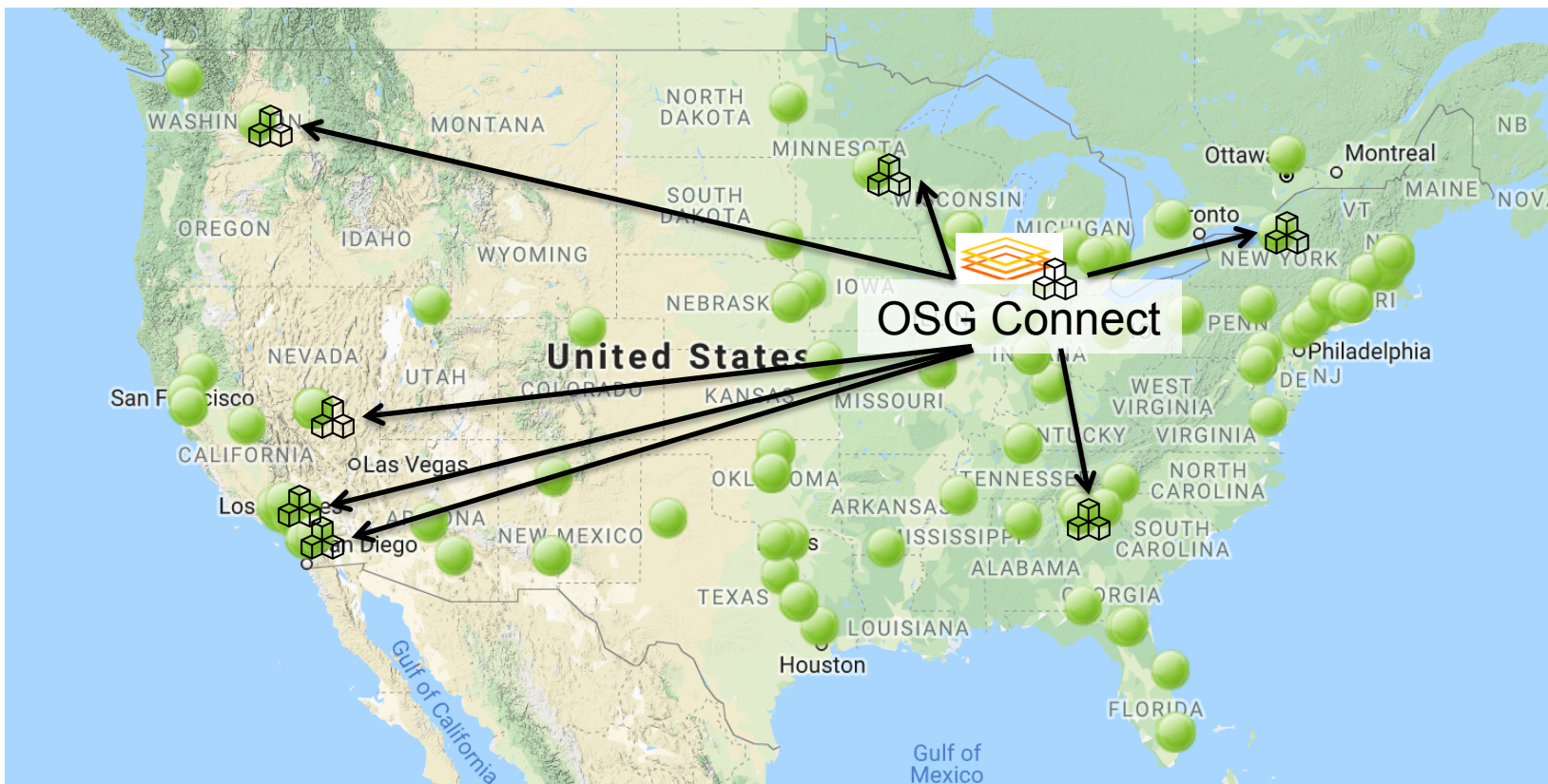
Pre-existing Software

- The ideal for DHTC is to package and bring along your own software, but...
- You can use pre-existing software installations **if** the computers you're running on have your software installed (or access to a repository with the software).

Pre-existing software via OSG Connect

- On the Open Science Grid, jobs submitted from OSG Connect have access to a software repository maintained by OSG Connect staff.
- The software repository is available across the OSG.
- Software is accessed using “modules”.

Software across the OSG



Module Commands

- See what modules are available

```
[~]$ module avail
```

```
[~]$ module spider lammmps
```

- Load a module

```
[~]$ module load lammmps/20180822
```

- See loaded modules

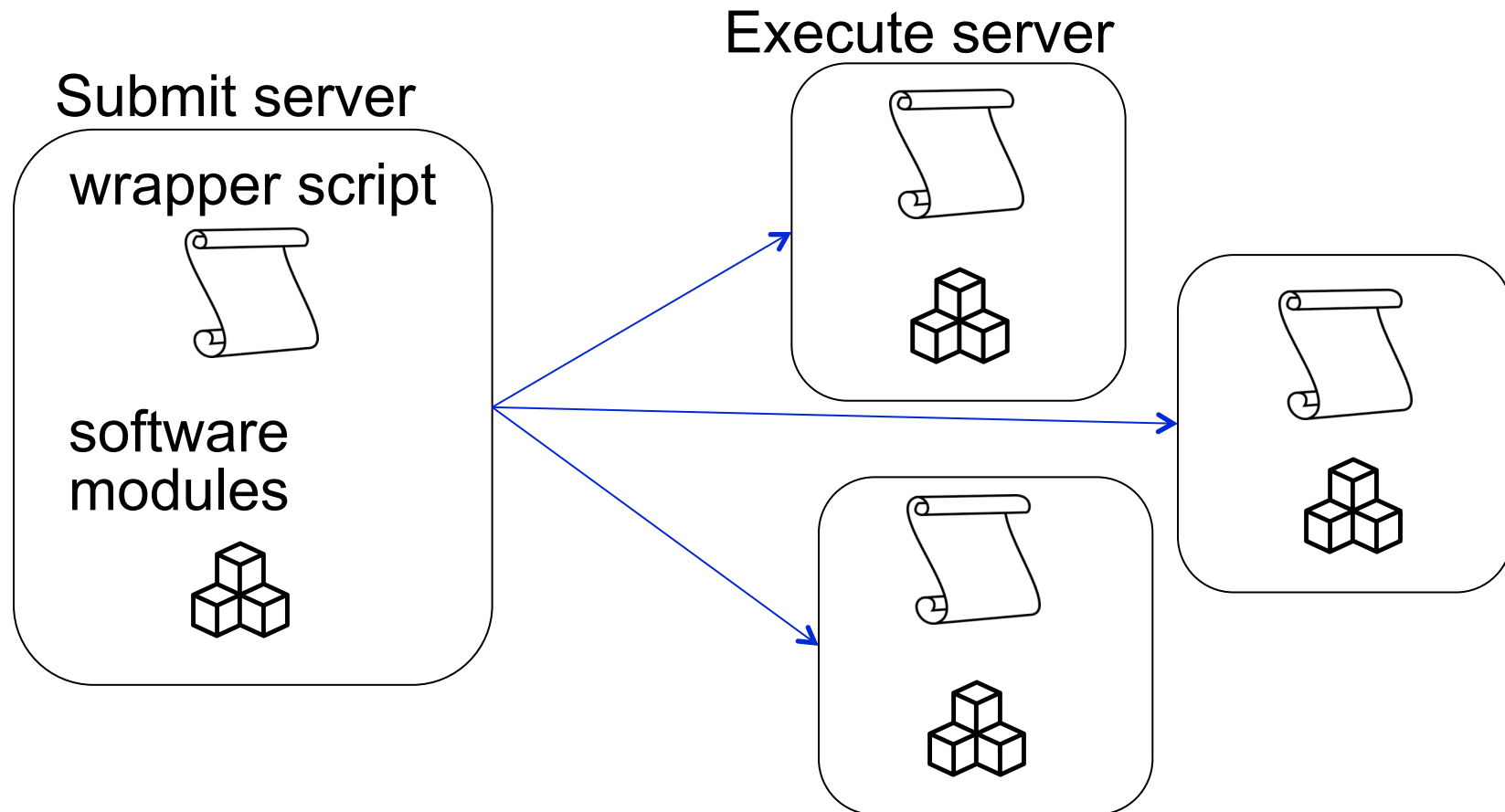
```
[~]$ module list
```

Module Workflow

1. Find a module for your software
2. Write a wrapper script that loads the module and runs your code
3. Include requirements to ensure that your job has access to modules

```
requirements = (HAS_MODULES ==?= true) &&  
(OSGVO_OS_STRING == "RHEL7") && (OpSys == "LINUX")
```


Module Workflow





LICENSING

Licensing

- Many scientific software programs are licensed.
- Licenses are restrictive, particularly for high-throughput computing

License Variations

- Per machine or 'single-install'
- Per *running* instance of the software (per “job”)
- Per username / user
- Via a license server
 - can support 1 - 1000s of concurrently running processes (“seats”)

Licensing implications for DHTC

- **Per machine or 'single-install'**: can't be used for DHTC
- **Per job**: restrictive, limits the number of jobs you can have running, how do you access licenses from execute servers?
- **Username**: restrictive, could only run jobs on one system where your jobs run as *your username*

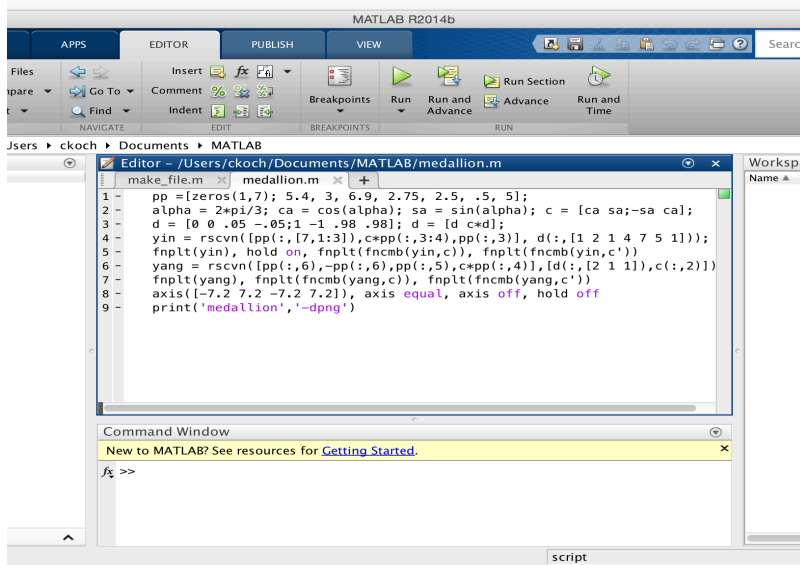
Matlab

- Wait a minute...isn't Matlab licensed?
- Yes, when interpreted on your computer using a normal Matlab installation.
- However, Matlab code can also be compiled.
- Once compiled, the code can be run without a license using a (free) set of files called the Matlab runtime (which acts like the interpreter).

Matlab contrast

Running Matlab on your computer
Uses license per instance

Running Matlab on DHTC
Uses license once, runs
many instances for free



Matlab script(s)

compiled w/ Matlab
compiler (uses license)

Compiled binary

interpreted by

Matlab Runtime (free)



Matlab on DHTC

1. Compile Matlab code using the Matlab compiler (mcc) on the same operating system (Linux)
 - requires a license
2. Prepare a copy of the Matlab runtime
 - download for free from Mathworks
3. Write a script that “installs” the runtime
 - The Matlab compiler actually writes most of this script for you
4. Use the runtime install to run the compiled Matlab code

Approaches

- Seek out open source alternatives
 - Python or R packages that emulate specific software behavior
 - If you can't replace entire workflow, substitute free software where you can
- License-free workarounds (Matlab) or special research agreements
- Choose the least restrictive license possible

Exercises

- Exercise 1.1: Try an OSG Connect software module
- Exercise 1.2: Compile and run Matlab code
- Feel free to catch up with materials from Monday/Tuesday or start implementing your own software

Questions?

- Now: Hands-on Exercises
 - 9:30-10:30am
- Next:
 - 10:30-10:45am: Break
 - 10:45am-12:15pm: Software Containers
 - 12:15-1:15pm: Lunch