

Containers for DHTC

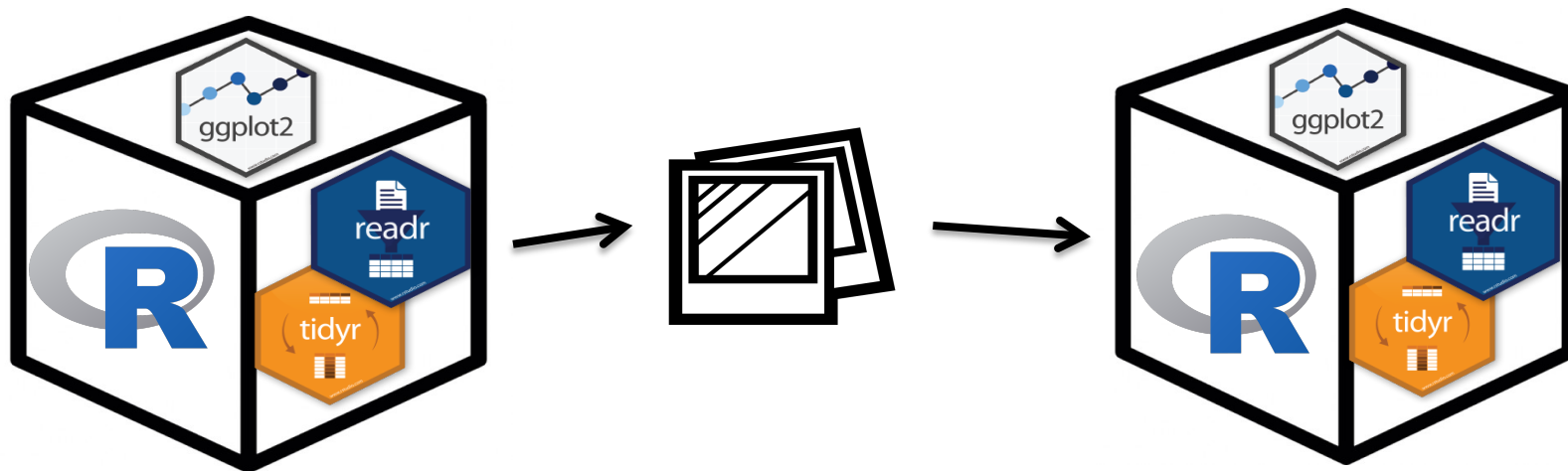
Christina Koch (ckoch5@wisc.edu)

Research Computing Facilitator

University of Wisconsin - Madison

Containers

- Containers are a tool for capturing an entire job “environment” (software, libraries, operating system) into an “image” that can be used again.



Container Types

- Two common container systems:

Docker

<https://www.docker.com/>



Singularity

<https://sylabs.io/>



The container itself will always be some version of Linux - but can be run on Linux / Mac / Windows if Docker or Singularity is installed

Why Containers?

Why use containers instead of the methods we discussed yesterday?

- **Complex installations:** software that has a lot of dependencies or components
- **Software that can't be moved:** do files or libraries have to be at a specific path?
- **Sharing with others:** one container can be used by a whole group that's doing the same thing
- **Reproducibility:** save a copy of your environment

DHTC Container Requirements

- Computer requirements:
 - Underlying container system (Docker or Singularity) needs to be installed on the computers where your job runs
 - Permissions on that system allow the use of containers
- Your requirements:
 - A container with your software

Using Containers in DHTC

- Who can support containers?
 - CHTC: Docker, some Singularity
 - OSG (from OSG Connect): Singularity
 - ask your local computing center
- Where do you find containers?
 - DockerHub: <https://hub.docker.com/>
 - SingularityHub: <https://www.singularity-hub.org/>
 - Can also create your own!
 - Talk to us if you're interested in this.

Container Workflow

1. Create a container or find one online
 - DockerHub: <https://hub.docker.com/>
 - SingularityHub: <https://www.singularity-hub.org/>

2. Place container into public or private registry
 - Docker Hub: <https://hub.docker.com/>
 - OSG Connect registry

Container Workflow

3. Create a customized submit file that runs your job inside the container.
 - Requirements to indicate that you need Docker or Singularity
 - Docker: Use HTCondor's docker universe
 - Singularity (from OSG Connect servers): add attribute to submit file

Submit File Requirements

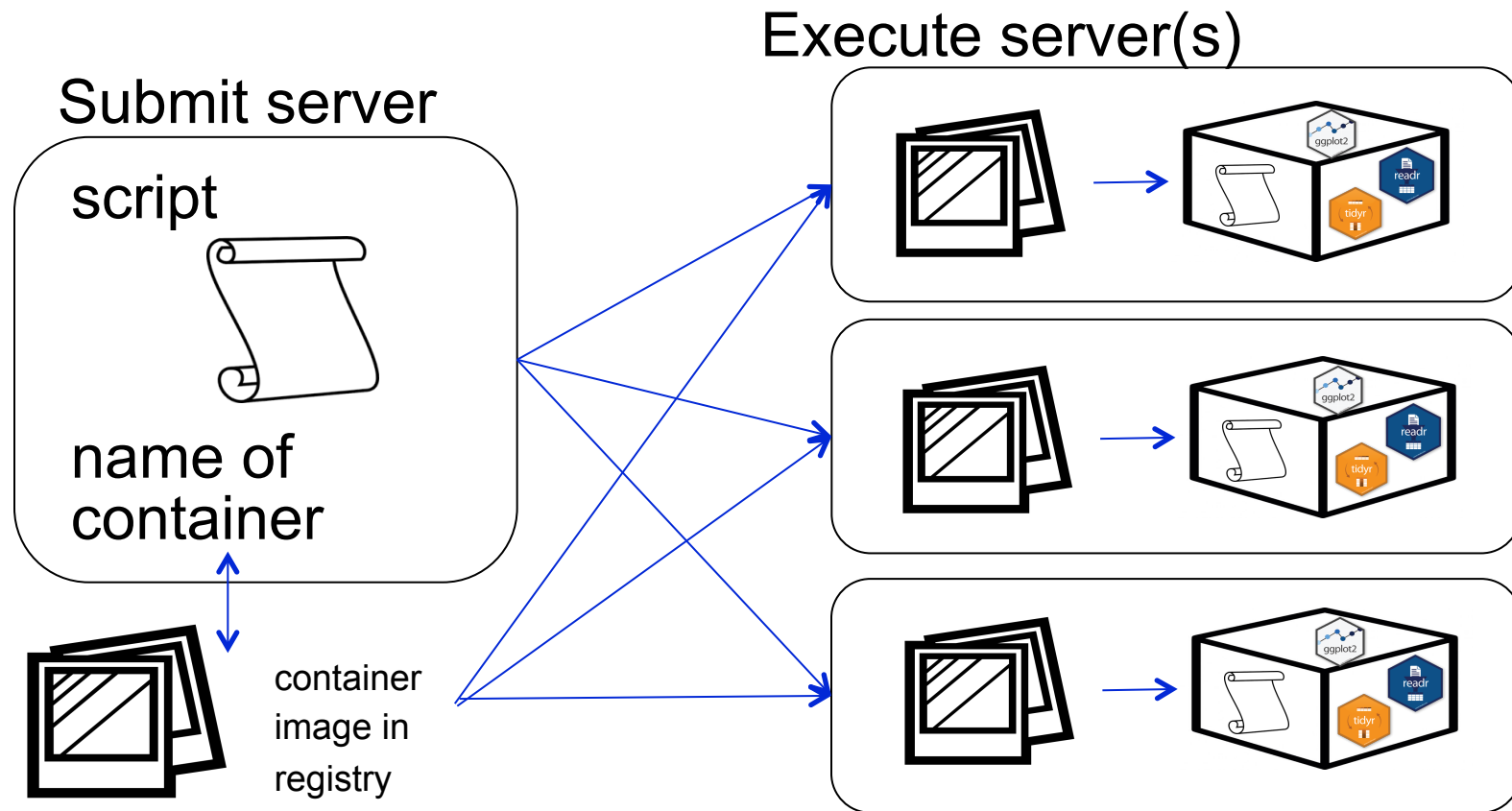
- Docker

```
universe = docker
docker_image = python:3.7.0
requirements = (HasDocker == true)
```

- Singularity (from OSG training server)

```
+SingularityImage = "/cvmfs/
singularity.opensciencegrid.org/centos/
python-34-centos7:latest"
requirements = (HAS_SINGULARITY == true)
```

Container Workflow



Conclusion

To use any software in a DHTC system:

1. Create/find environment/software package
 - download pre-compiled code, compile your own, build your own, use distributed software modules, create/find a container
2. Write a script to set up the environment when the job runs
3. Account for all dependencies, files, and requirements in the submit file

Exercises

- Exercise 2.1: Use Singularity from OSG Connect
- Exercise 2.2: Use Singularity to Run Tensorflow (Optional)
- Exercise 2.3: Using Docker

Questions?

- Now: Hands-on Exercises
 - 11:15am-12:15pm
- Next:
 - 12:15-1:15pm: Lunch
 - 1:15 onward: free time