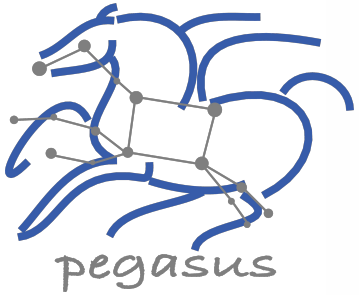


Running Pegasus Workflows on OSG

An Introductory Tutorial

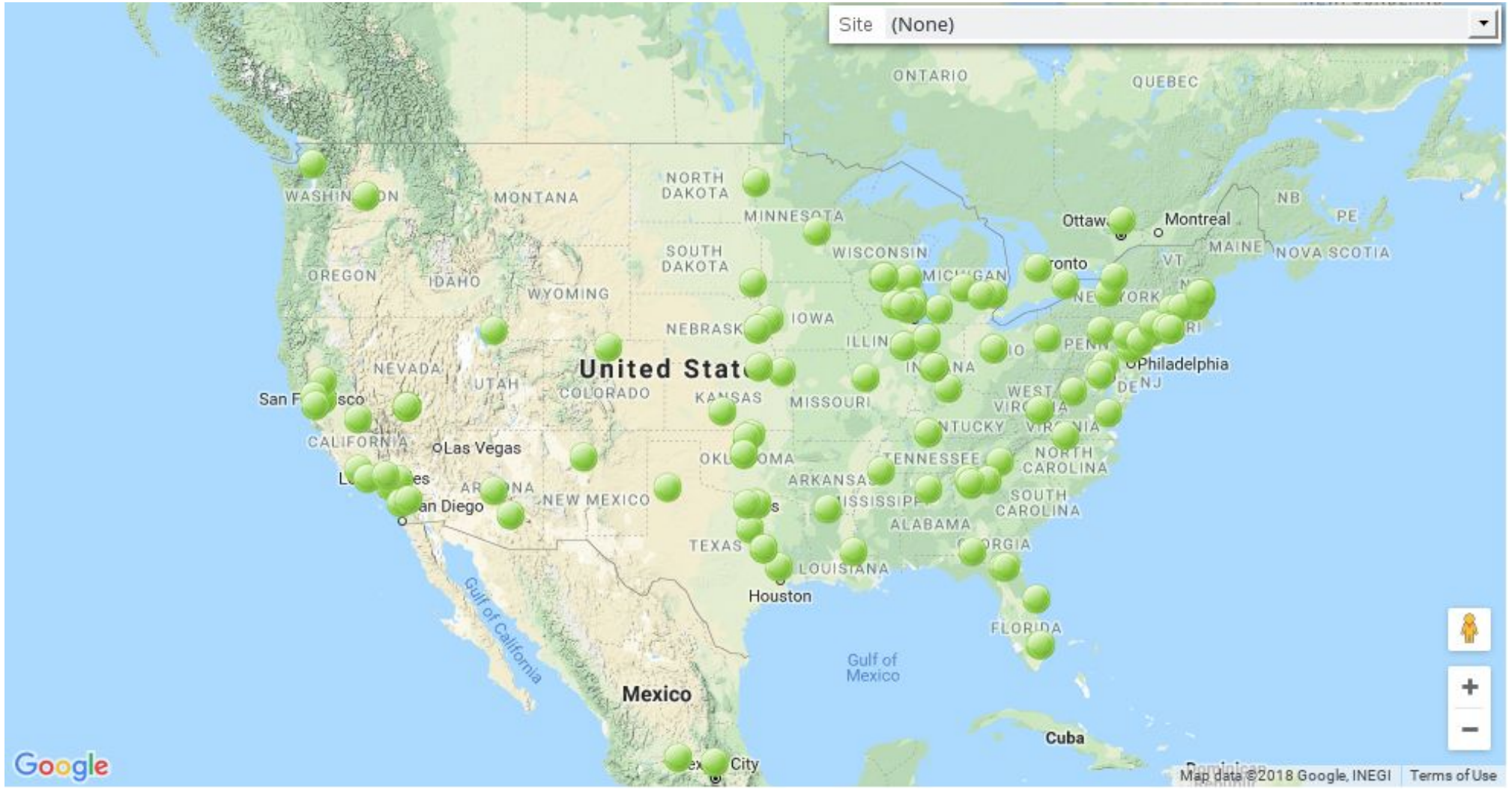
USC Viterbi School of Engineering Information Sciences Institute

NSF U.S. DEPARTMENT OF ENERGY NIH DARPA

The Open Science Grid (OSG)

- What is OSG?
  - Distributed computing network of mostly universities and national labs who contribute their own local compute resources
    - Some resources are dedicated to OSG
    - Other resources are contributed when they sit idle
- Who can use OSG?
  - Anyone affiliated with a United States research institution (university or research lab)
  - You can use OSG resources without contributing to its pool
- How do you access OSG?
  - ssh into a designated login node (e.g. login04.osgconnect.net)
  - submit HTCondor jobs or Pegasus workflows
- Where is it located?
  - All over the United States with computing and storage elements at over 100 individual sites
  - https://map.opensciencegrid.org/map/iframe#38.575019,-112.638144|3|terrain

The Open Science Grid (OSG) cont.



The Open Science Grid (OSG) cont.

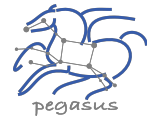
- Is my workload suitable?

	Ideal jobs!	**Still very advantageous**	**Less-so, but maybe**
CPU	1	< 8	> 8 (or MPI)
Walltime	< 12 hrs*	< 24 hrs*	> 24 hrs
RAM	< few GB	< 10s GB	> 10s GB
Input	< 500 MB**	< 10 GB**	> 10 GB
Output	< 1GB**	< 10 GB**	> 10 GB
Software	pre-compiled binaries, containers	Most other than ->	Licensed software, non-Linux

* Or checkpointable
** Per job; you can work with a large dataset on OSG if it can be split into pieces

- Caveats
  - You don't know exactly where your job will run
  - It is not guaranteed that you will have enough time for your job to finish
    - jobs could be preempted without warning (because we are using someone else's resource)

Outline



1. ~~The Open Science Grid~~
2. **Running Pegasus Workflows**
3. Running HTCondor Jobs
4. Additional Notes & Resources

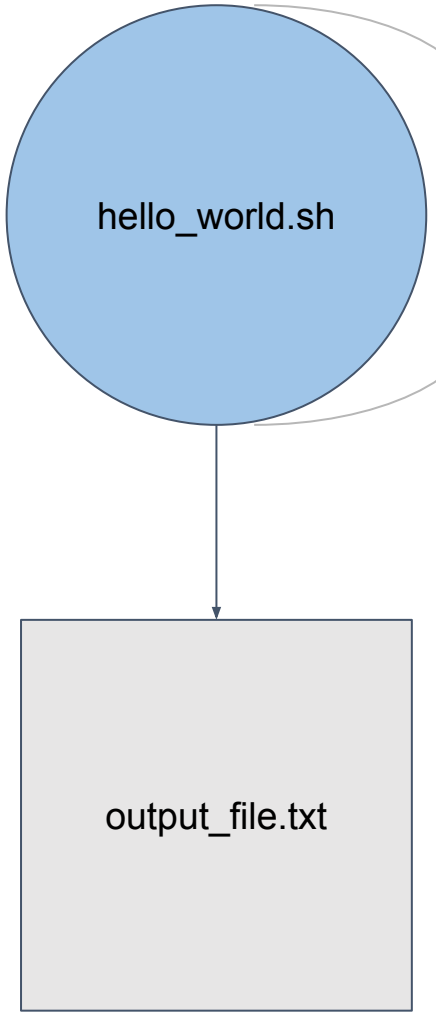
USC Viterbi School of Engineering Information Sciences Institute

Running Pegasus Workflows On OSG

- "Out of the box" Pegasus support through login nodes
  - pegasus4.9.3 installed by default (need to download rhel7 tarball to get the latest)
  - HTCondor is already installed
  - can use Pegasus' default generated SiteCatalog
  - can use Pegasus' default data staging configuration: condorio

Let's try running an example Pegasus workflow..

Example Workflow: What does it look like?




```
#!/bin/bash

set -e

printf "Hello OSG!\nI ran on this OS:\n" >> output_file.txt
# show OS info
cat /etc/os-release >> output_file.txt


if command -v nvidia-smi &> /dev/null
then
    printf "\nAnd here is some GPU info:\n" >> output_file.txt
    nvidia-smi >> output_file.txt
fi
```

Example Workflow: Steps to run

1. ssh into log-in node: ssh <username>@login04.osgconnect.net
2. setup Pegasus path: export PATH=/public/ryantanaka/pegasus-5.0.0dev/bin:$PATH
3. ensure path set correctly: which pegasus-version
4. setup Pegasus python path: export PYTHONPATH=$(pegasus-config --python)
5. get workflow: git clone https://github.com/ryantanaka/pegasus-osg-tutorial.git
6. go to pegasus-workflow: cd pegasus-osg-tutorial/pegasus-workflow
7. run: ./workflow.py

Once the workflow finishes, it will automatically print the contents of the produced output file if it exists.

Example Workflow: Running jobs in containers

- How It Works
  - **OSG hosts singularity images on its CernVM File System** (CVMFS), a POSIX read-only file system where files are heavily cached in a manner transparent to the user
  - OSG provides pre-made containers with popular software packages
  - for a job to run in a container, two profile keys must be set:

```
Requirements = HAS_SINGULARITY == TRUE
+SingularityImage = "/cvmfs/singularity.opensciencegrid.org/opensciencegrid/osgvo-el7:latest"
```

  - this is telling OSG (not Pegasus) to use a container

- Using Your Own Custom Docker or Singularity Container (in brief)
  1. create your own custom container image using Docker and push it to Docker Hub
  2. add your Docker image to the OSG image repository
  3. use the container image in jobs (if your Docker image is updated, OSG will also update that image in its repository)

Example Workflow: Requesting GPU access

- How It Works
  - some of the compute nodes have nVidia GPUs
  - for your job to run on a compute node with an available GPU, the Condor profile: `request_gpus = 1` must be set
  - only a single GPU can be requested
  - you cannot specify a specific GPU without workarounds…
  - queuing times will be longer as there are far less GPU resources as there are CPU resources

```
condor_status -const 'gpus >= 1' -af CUDADeviceName | sort | uniq -c
                24 GeForce GTX 1080 Ti
                60 GeForce GTX 750 Ti
                 4 Quadro RTX 5000
                37 Quadro RTX 6000
                 4 Quadro RTX 8000
                15 Tesla K20m
               100 Tesla K40m
                 2 Tesla P100-PCIE-12GB
                 8 Tesla P100-PCIE-16GB
                 3 Tesla V100-PCIE-16GB
                15 Tesla V100-PCIE-32GB
```

What was the OS of the host where your job ran??

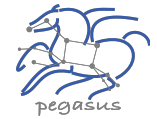
Example Workflow: Run w/Container + GPU

In workflow.py:

1. uncomment lines 28 and 29 so that jobs using the hello_world executable can be run in a container
2. uncomment line 32 so that jobs using the hello_world executable will run on a compute node that has a GPU

```
17 # --- Transformation Catalog ------------------------------------------------
18 hello_world = Transformation(
19                 name="hello_world",
20                 site="local",
21                 pfn=Path(".").resolve() / "hello_world.sh",
22                 arch=Arch.X86_64,
23                 os_type=OS.LINUX,
24                 is_stageable=True
25             )
26
27 # Require that this job is to be run in a specified container. (quotes are needed for image name)
28 hello_world.add_condor_profile(requirements='HAS_SINGULARITY == True')
29 hello_world.add_profiles(Namespace.CONDOR, key="+SingularityImage", value='"/cvmfs/singularity.opensciencegrid.org/opensciencegrid/tensorflow-gpu:2.3-cuda-10.1"')
30
31 # Require that this job needs a GPU.
32 hello_world.add_pegasus_profile(gpus=1)
33
34 tc = TransformationCatalog().add_transformations(hello_world)
```

Outline



1. ~~The Open Science Grid~~
2. ~~Running Pegasus Workflows~~
3. **Running HTCondor Jobs**
4. Additional Notes & Resources

USC Viterbi School of Engineering Information Sciences Institute

Submitting an HTCondor Job

htcondor-job
- hello_world.sh
- hello_world.submit

To submit, simply run:
condor_submit
hello_world.submit

```
# executable
executable = hello_world.sh

# file to save terminal output to
error = hello_world.err
output = hello_world.out

# output file to transfer back
transfer_output_files = output_file.txt

# request the resources that the job will need
request_cpus = 1
request_gpus = 1
request_memory = 8 MB
request_disk = 1 MB

# container settings
Requirements = HAS_SINGULARITY == TRUE
+SingularityImage = "/cvmfs/singularity.opensciencegrid.org/opensciencegrid/tensorflow-gpu:2.3-cuda-10.1"

# queue one job
queue 1
```

Outline

~~1. The Open Science Grid~~

~~2. Running Pegasus Workflows~~

~~3. Running HTCondor Jobs~~

4. Additional Notes & Resources

Other Notes

- Data Management

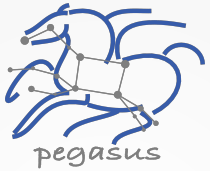
	Purpose	Network mounted	Backed Up	Initial Quota*
/home	Storage of submit files, input files <100MB each, and per-job output up to a 1GB. Jobs should ONLY be submitted from this folder.	No	No	50 GB
/public	Staging ONLY for large input files (100MB-50GB, each) for publicly-accessible download into jobs (using HTTP or stashcp, see below) and large output files (1-10GB).	Yes	No	500 GB

	Recommended File Sizes	Where Data is Placed	Command	Purpose	Details
HTCondor File Transfer	<100 MB per file; <500 MB per job	/home	transfer_input_files	General-use transfer of job input from within /home.	HTCondor File Transfer
HTTP	<1 GB per file	/public or non-OSG web server	http address in transfer_input_files	For input files 100MB-1GB, if shared across many jobs.	HTTP Access
OSG's StashCache	>1 GB per file; <10 GB per job	/public	stashcp	For the largest input files, staged within /public	StashCache
GridFTP	> 10 GB	/public	gfal-copy	input staged in /public	Typically used by experts with large work flows. Please contact us if you're interested.

Other Notes cont.

- Job Submission
  - jobs should only be submitted from /home
  - jobs should never be submitted from /public
- OSG Documentation
  - Containers
  - GPU Resources
- Running Jobs on the Open Science Grid video

What GPU did your job get??



Questions?