

# OSG Site Installation Overview

Brian Lin

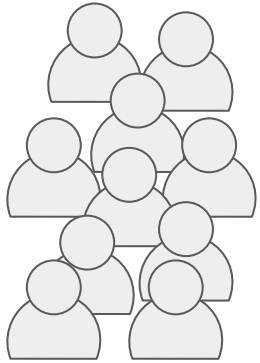
OSG Software

University of Wisconsin — Madison

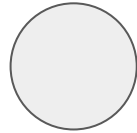


# Step 0: Is the OSG for you?

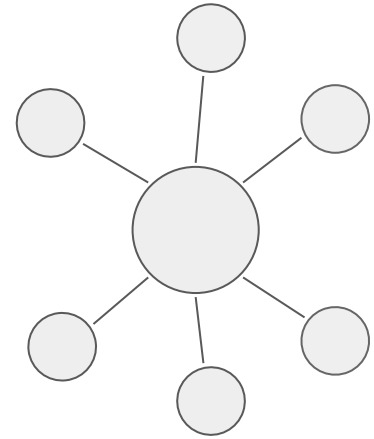
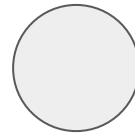
# The OSG Model



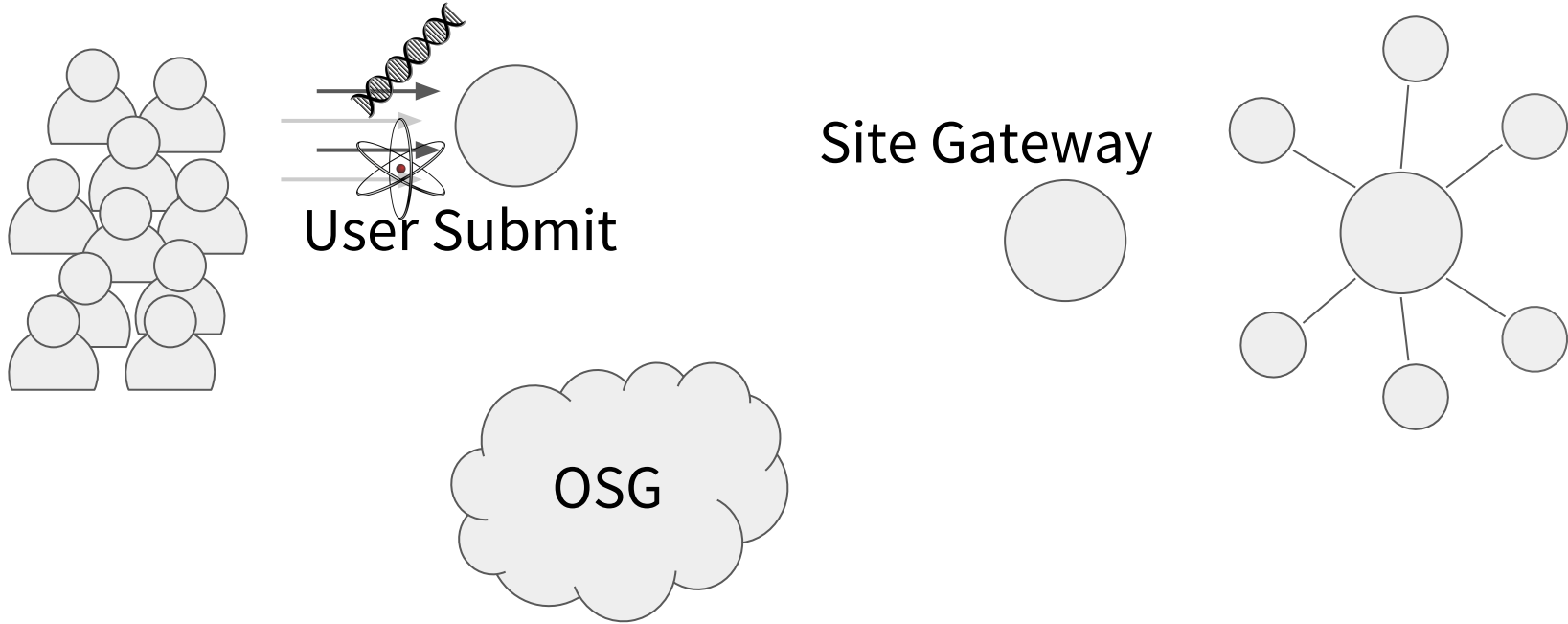
User Submit



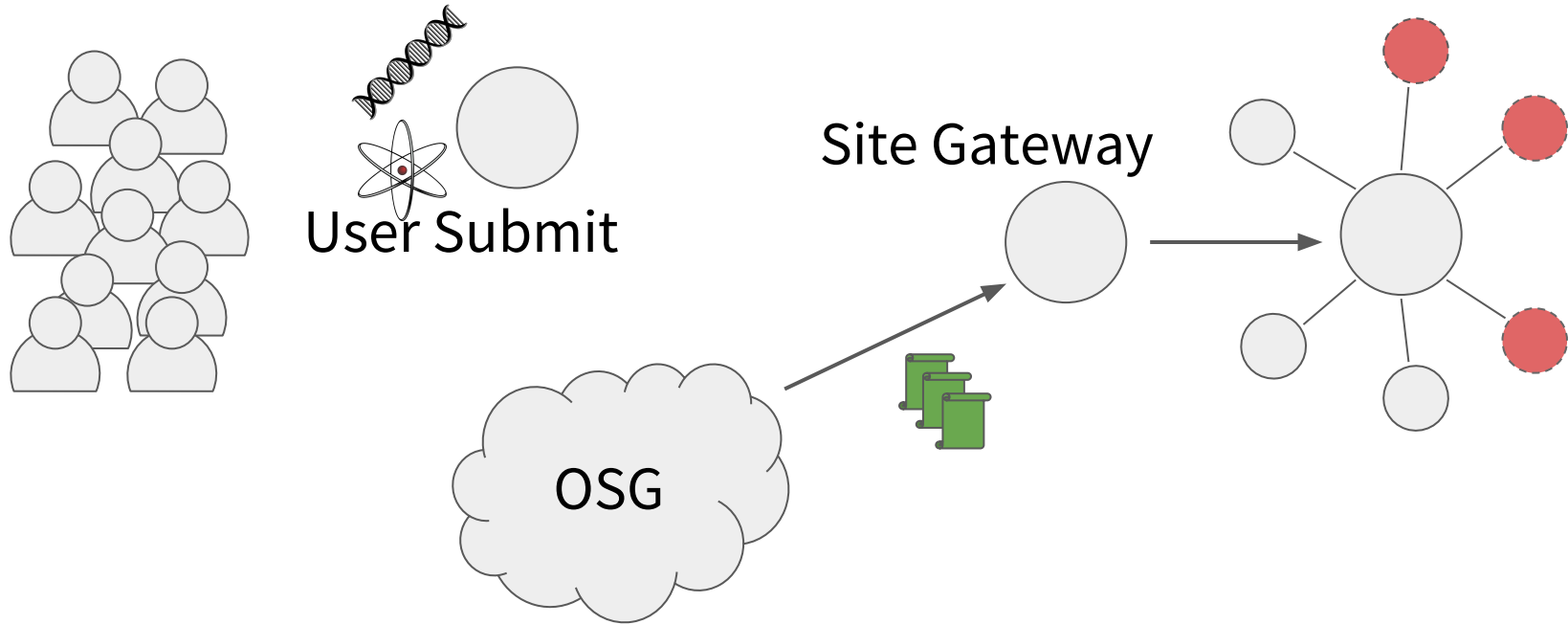
Site Gateway



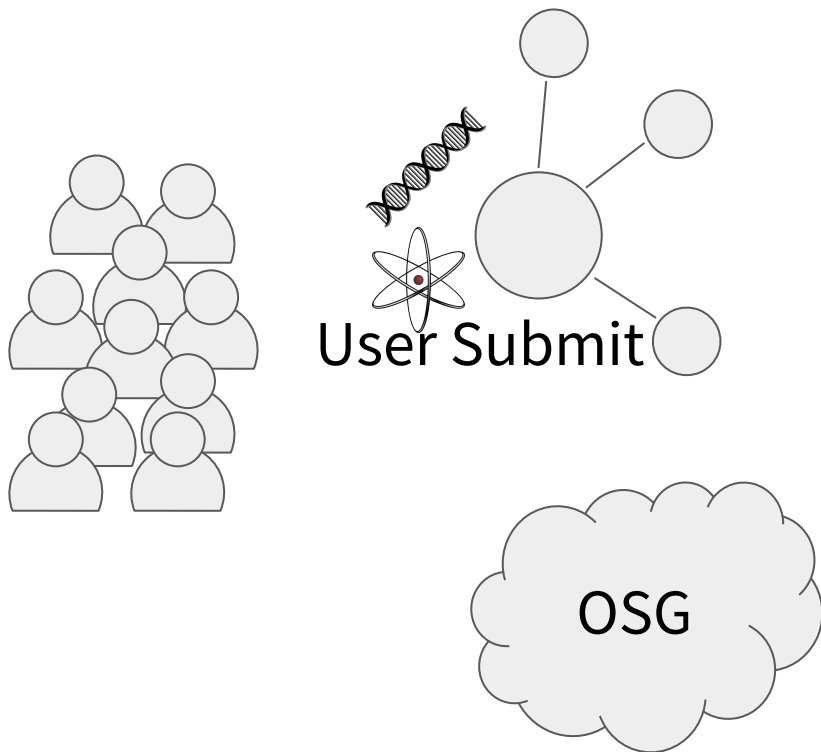
# The OSG Model



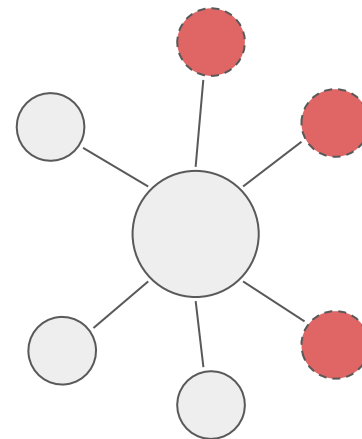
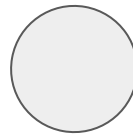
# The OSG Model



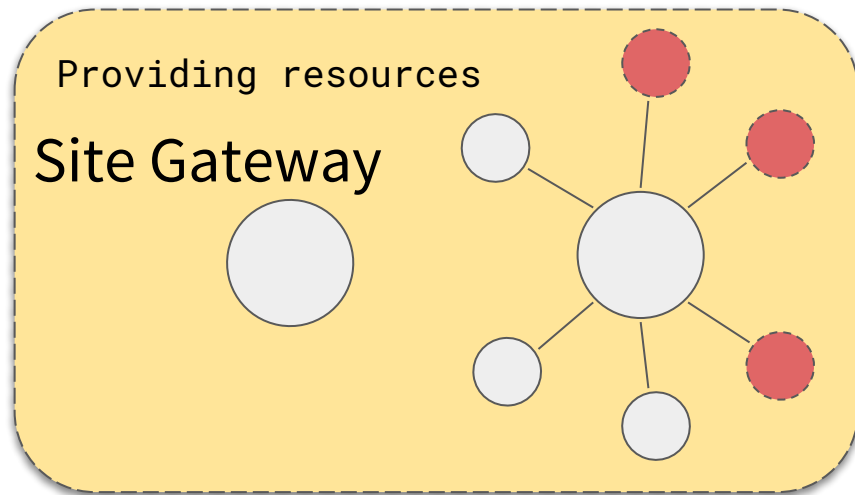
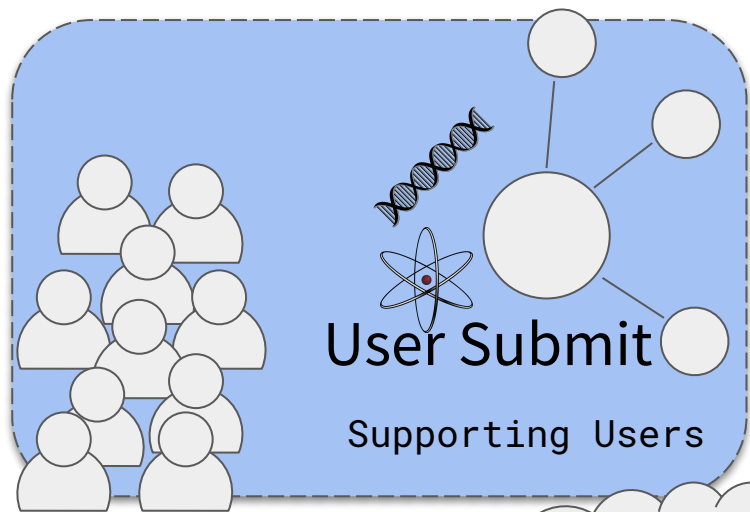
# The OSG Model



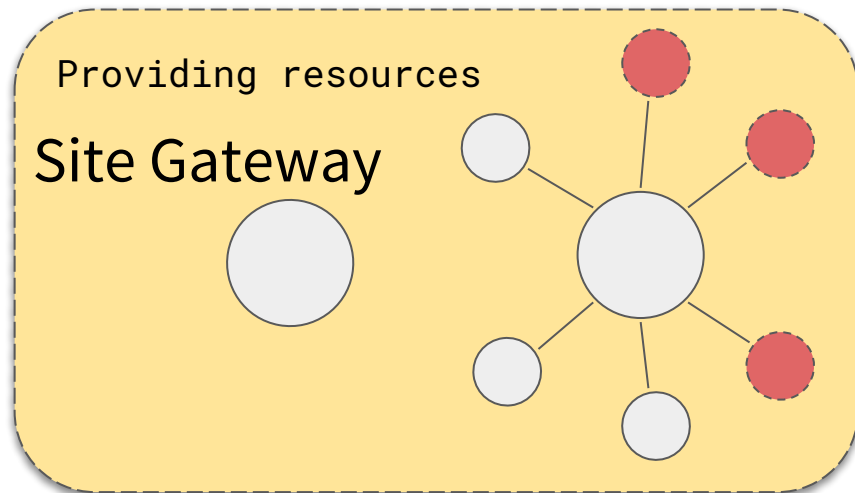
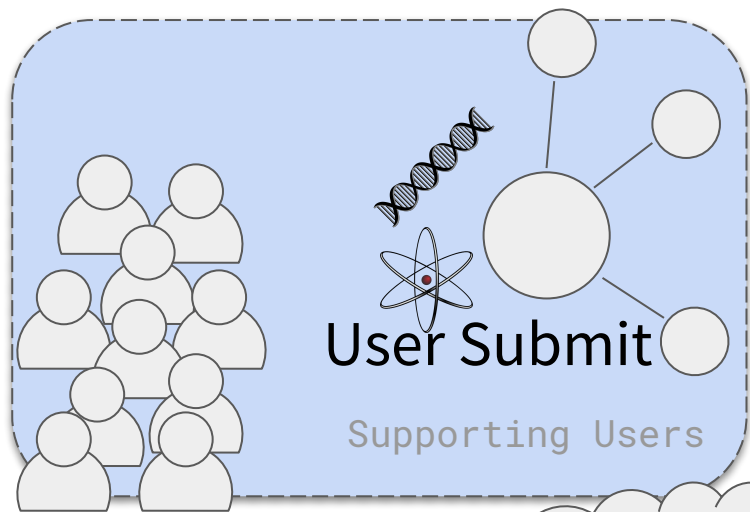
Site Gateway



# The OSG Model



# The OSG Model





# Minimum Requirements for Providing Resources

- Batch Systems: HTCondor, Slurm, Torque/PBS, LSF, SGE
- Operating Systems: Red Hat Enterprise Linux, CentOS, Scientific Linux
- Outgoing WAN access from worker nodes
- Firewall opened to OSG services on the site gateway or head node

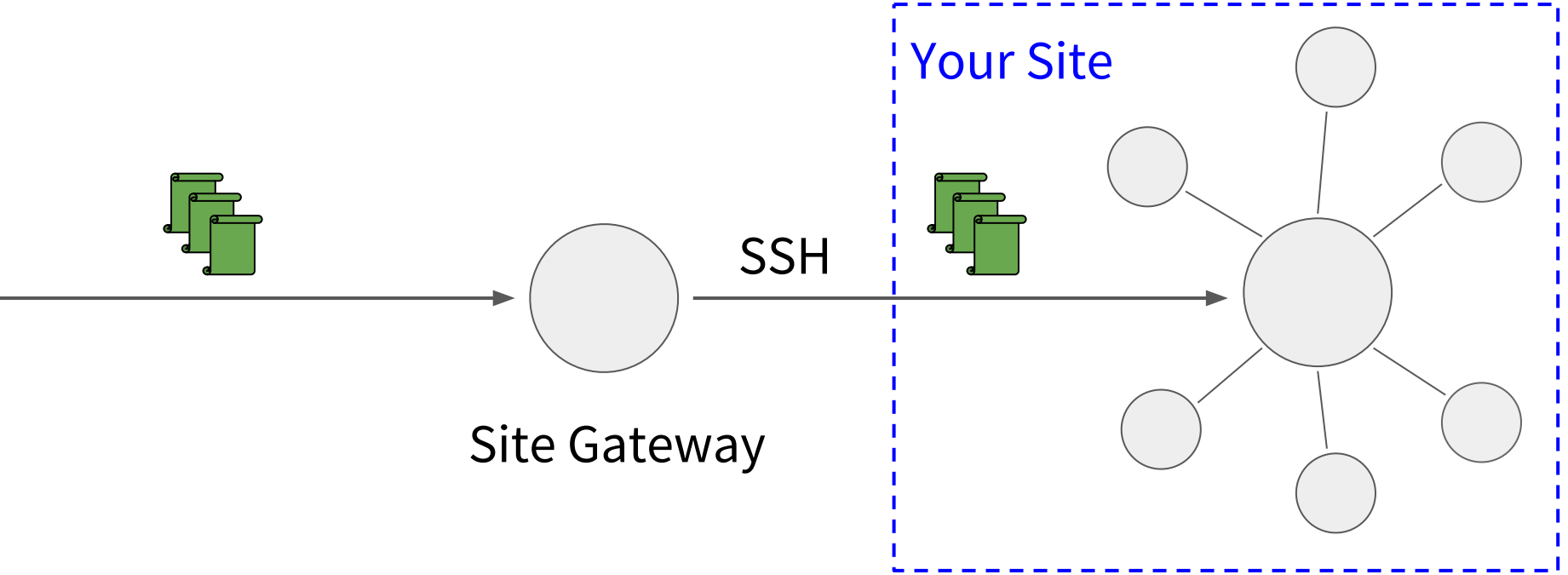
# Step 1: Hosted CE or HTCondor-CE?

# Hosted CE or HTCondor-CE?

- Do you want  $> O(10^4)$  OSG jobs?
- Are there special rules or policies for submitting jobs to your site?
- Do you want to change your configuration frequently?

If you answered no to the above questions, an OSG Hosted CE could work for you. Contact us at [help@opensciencegrid.org](mailto:help@opensciencegrid.org) for further information.

# OSG Hosted CE

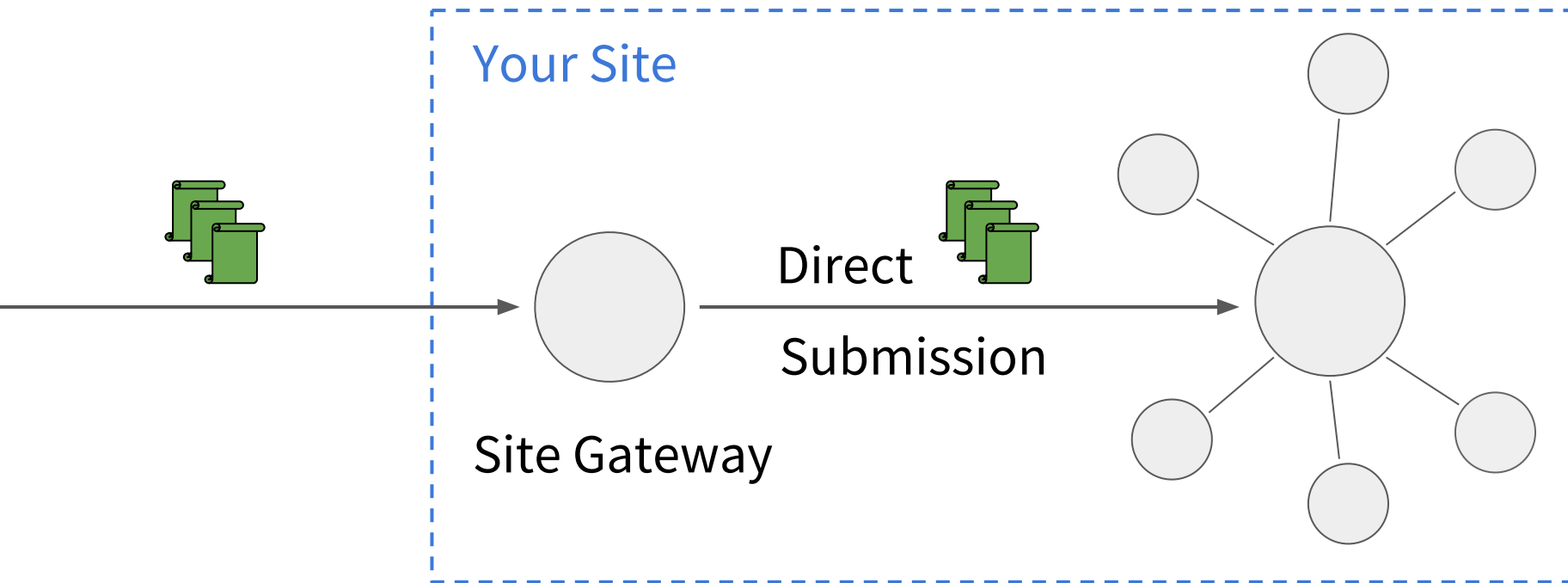


# Hosted CE

- In addition to the minimum requirements, you must:
  - Create a Unix account across your cluster for OSG jobs to run as
  - Allow SSH key access to your head node from the OSG to the above Unix account
- If you have an HTCondor batch system, you'll need a shared FS to share the user home directories or install the OSG worker node client across your cluster
- Optional: Support additional jobs with CVMFS and a Frontier Squid proxy
- See more details in our documentation

<https://opensciencegrid.org/docs/compute-element/hosted-ce/>

# HTCondor-CE



# HTCondor-CE

- In addition to the minimum requirements, you must open port 9619 on your HTCondor-CE
- Installation details: <https://opensciencegrid.github.io/docs/compute-element/install-htcondor-ce/>
- Choose which virtual organizations (VOs) you'd like to support and create accounts <https://opensciencegrid.org/docs/security/lcmaps-voms-authentication/#supporting-mapped-vos-and-users>
- Register your HTCondor-CE <https://opensciencegrid.org/docs/compute-element/install-htcondor-ce/#registering-the-ce>
- Additional documentation for configuring your HTCondor-CE:
  - <https://opensciencegrid.org/docs/compute-element/install-htcondor-ce/#configuring-htcondor-ce>
  - <https://opensciencegrid.org/docs/compute-element/job-router-recipes/>
- Documentation for troubleshooting your HTCondor-CE: <https://opensciencegrid.org/docs/compute-element/troubleshoot-htcondor-ce/>

# Step 2: Preparing your worker nodes



# OSG Worker Node Client

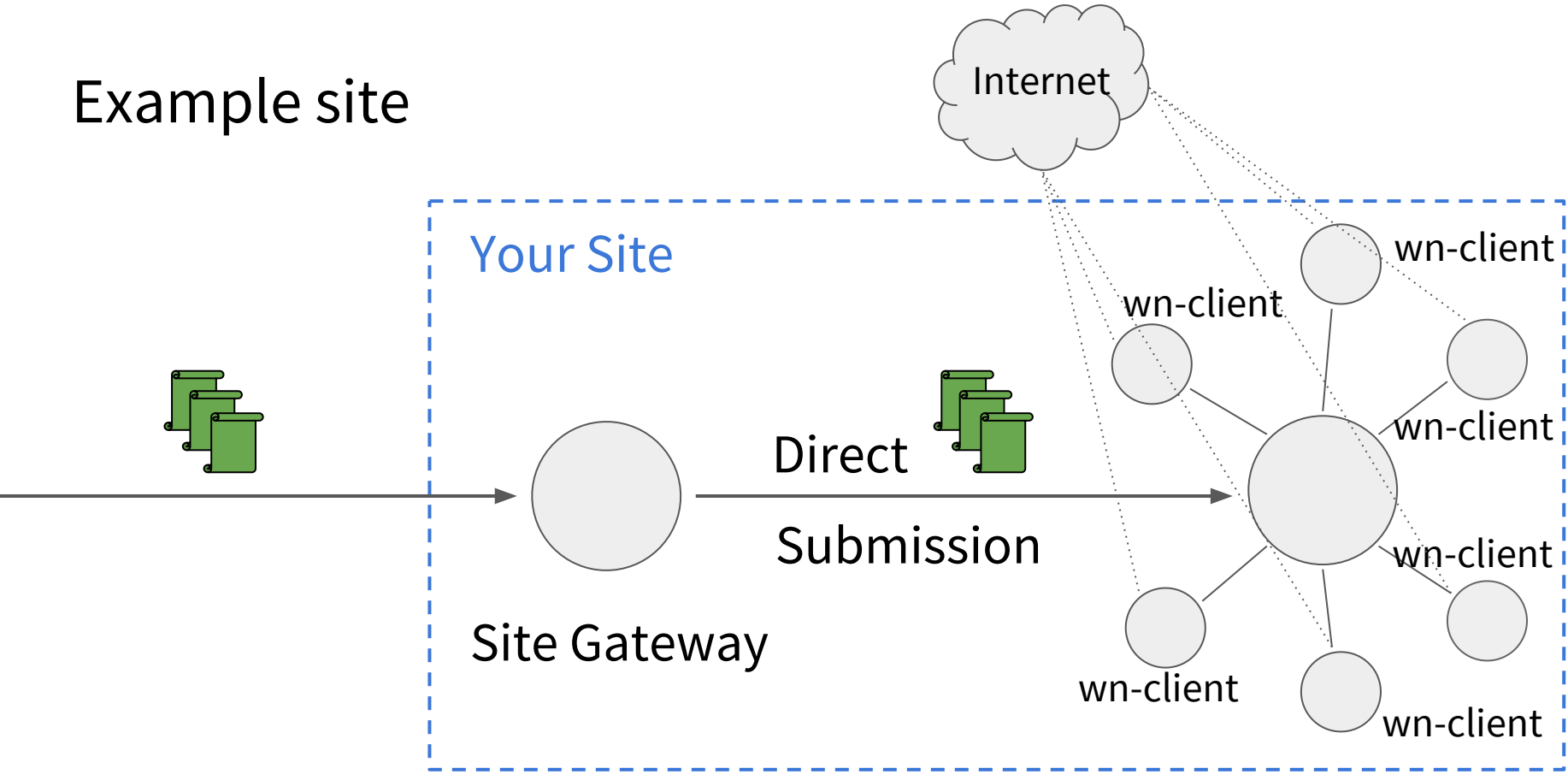
- Thin collection of software necessary for pilot job execution
- Installation options via RPM package, tarball, and OASIS
  - RPM: <http://opensciencegrid.github.io/docs/worker-node/install-wn/>
  - Tarball: <http://opensciencegrid.github.io/docs/worker-node/install-wn-tarball/>
  - OASIS: <http://opensciencegrid.github.io/docs/worker-node/install-wn-oasis/>

# OSG Worker Node Requirements

- **Outgoing WAN access!**
- OSG worker node client
- Pilot job temp space (OSG\_WN\_TMP)
  - Set by `worker_node_temp` configuration in `/etc/osg/config.d/10-storage.ini` on the CE
  - 2GB disk/core minimum; 10GB disk/core recommended
  - Site responsible for cleanup, e.g. `tmpwatch`
- Cleanup `/tmp` (recommendation)
- Further requirements and recommendations

<https://opensciencegrid.org/docs/worker-node/using-wn/>

# Example site



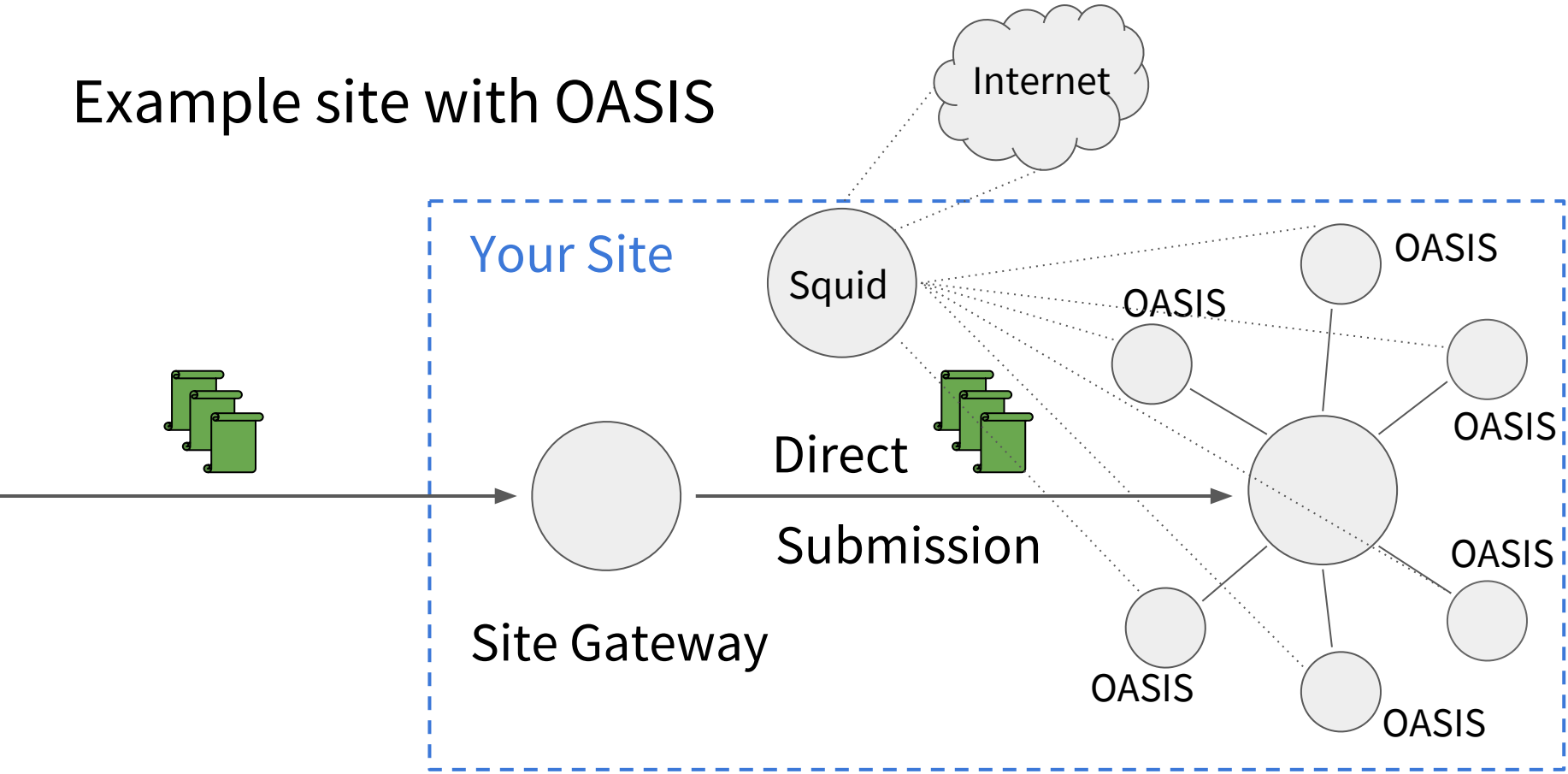
Validation: Request test pilot jobs

[osg-gfactory-support@physics.ucsd.edu](mailto:osg-gfactory-support@physics.ucsd.edu)

# CernVM File System (CVMFS)

- Software distribution service with a POSIX interface accessed over HTTP.  
<https://opensciencegrid.org/docs/worker-node/install-cvmfs/>
- OSG packaging includes the OSG Application Software Installation Service (OASIS) repository. Many jobs require OASIS or other CVMFS repositories to run!
- Mostly used by VOs to provide software for their users but it can also be used as an alternative source for OSG Software (e.g., `osg-wn-client`, `singularity`)
- Requirements:
  - At least one Frontier Squid (your CE can do double duty):  
<http://opensciencegrid.org/docs/data/frontier-squid/>
  - FUSE installed (brought in by the RPM)
  - ~25GB of cache space on its own partition
- Optional but recommended

# Example site with OASIS



# Singularity

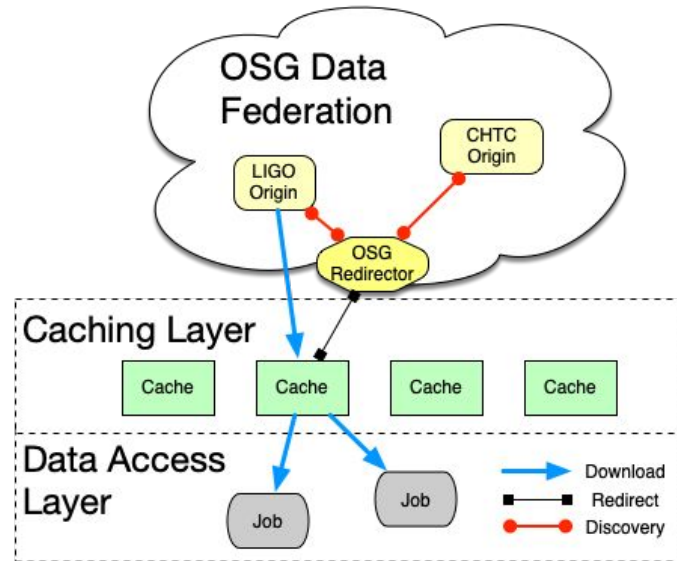
- Singularity is a container technology used by the OSG for payload job isolation:  
<http://opensciencegrid.org/docs/worker-node/install-singularity/>
- Get the most out of Singularity by installing it alongside OASIS, since some VOS distribute their images this way
- Optional but recommended

# Step 3: Storage services



# StashCache

- The OSG data federation for scalable distribution of “large” data <https://opensciencegrid.org/docs/data/stashcache/overview/>
- Origin: Provides a place for your users to put “large” data for their jobs in the OSG
- Cache: if you’d like to reduce OSG WAN usage, or a VO you support has their data in an origin. CVMFS should be installed on your WNs.



# VO Storage

If your site is funded by a particular VO, they may have data storage requirements:

- “Doors” to pre-existing storage:
  - Load-balanced GridFTP  
<http://opensciencegrid.org/docs/data/load-balanced-gridftp/>
  - XRootD  
<https://opensciencegrid.org/docs/data/xrootd/install-standalone/>
- XRootD as a storage solution:  
<https://opensciencegrid.org/docs/data/xrootd/install-storage-element/>

# Summary

# Decision points

- OSG-Hosted CE vs HTCondor-CE; if hosted CE, you can be done!
- What VOs do you want to support?
- osg-wn-client installation method
- Want more jobs? Install OASIS and Singularity on your worker nodes
- VO-specific storage requirements? StashCache/Load Balanced GridFTP/XRootD

# Networking

- Open outbound WAN access from worker nodes
- Allow SSH access to your head node from an OSG Hosted CE
- Open port 9619 (TCP) on HTCondor-CE
- Need to reduce WAN usage? Frontier Squid. Need more reductions? StashCache

Interested in an OSG Hosted-CE?

[help@opensciencegrid.org](mailto:help@opensciencegrid.org)

Ready for pilot jobs?

[osg-gfactory-support@physics.ucsd.edu](mailto:osg-gfactory-support@physics.ucsd.edu)

Issues?

[help@opensciencegrid.org](mailto:help@opensciencegrid.org)

# Questions?