Open Science Grid

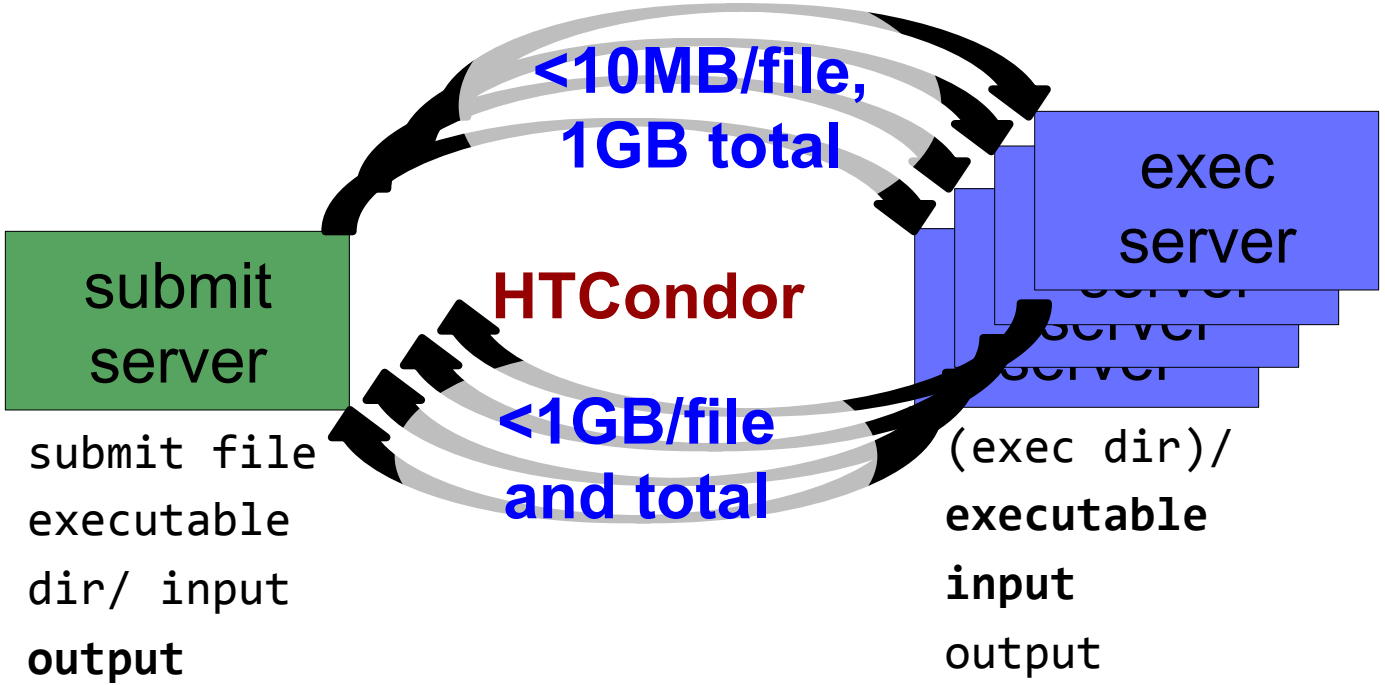
Large Input in DHTC

Thursday PM, Lecture 1

Lauren Michael

CHTC, UW-Madison

Hardware transfer limits

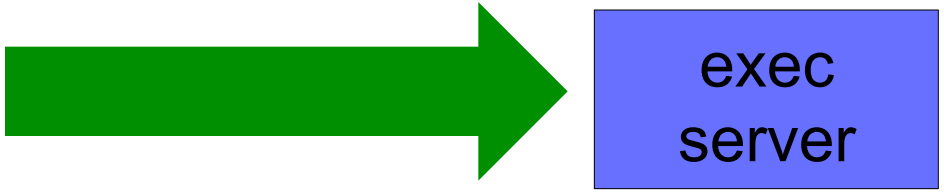


Reducing data needs

- ***An HTC best practice!***

- split large input for better throughput *and* less per-job data
- eliminate unnecessary data
- compress and combine files

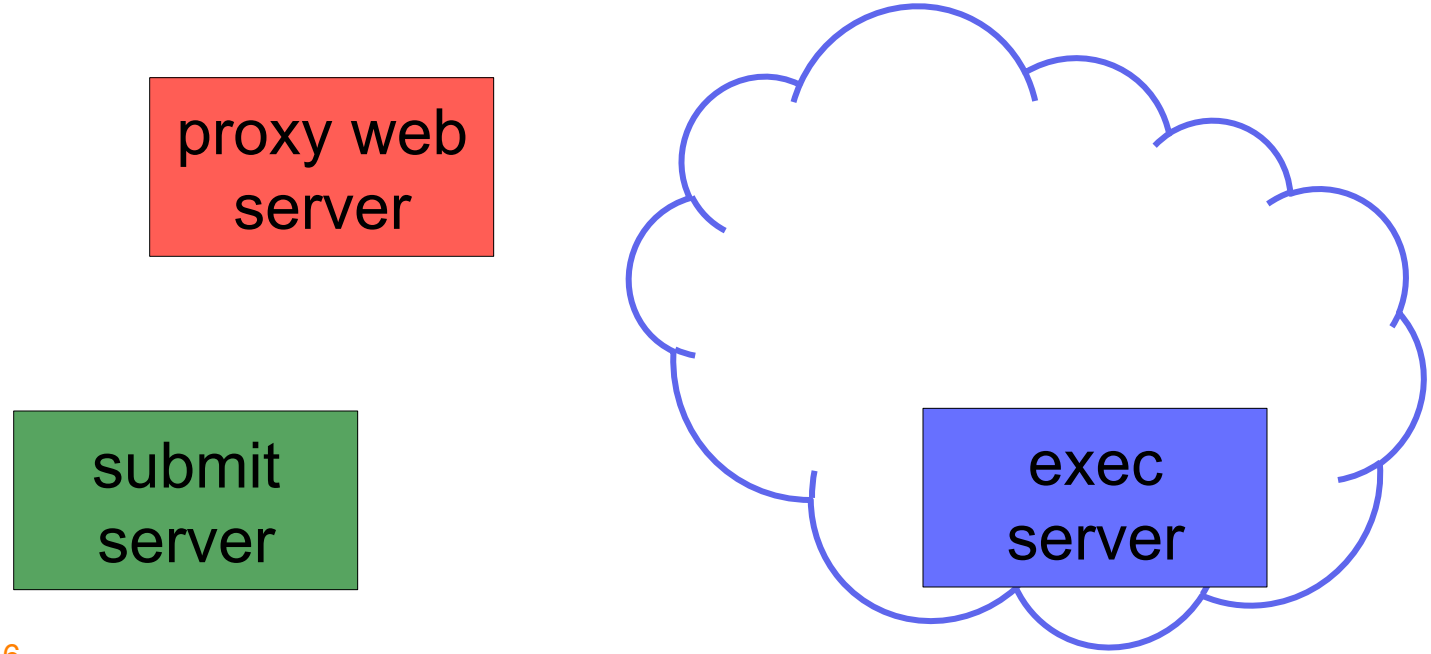
Large input in HTC and OSG



file size	method of delivery
words	within executable or arguments?
tiny – 10MB per file	HTCondor file transfer (up to 1GB total per-job)
10MB – 1GB, shared	download from web proxy (network-accessible server)
1GB - 10GB, unique or shared	StashCache (regional replication)
10 GB - TBs	shared file system (local copy, local execute servers)

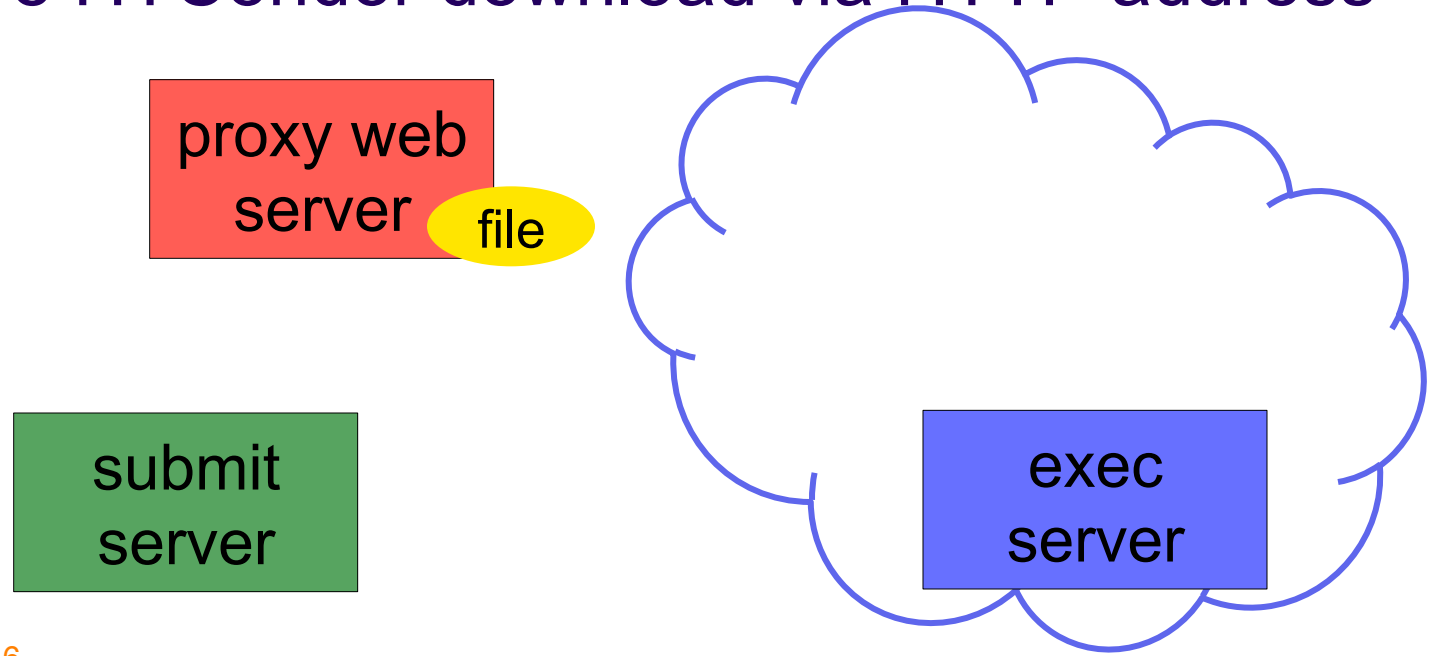
Using a Web Proxy

- Place the file onto a local, proxy-configured web server
- Have HTCondor download via HTTP address



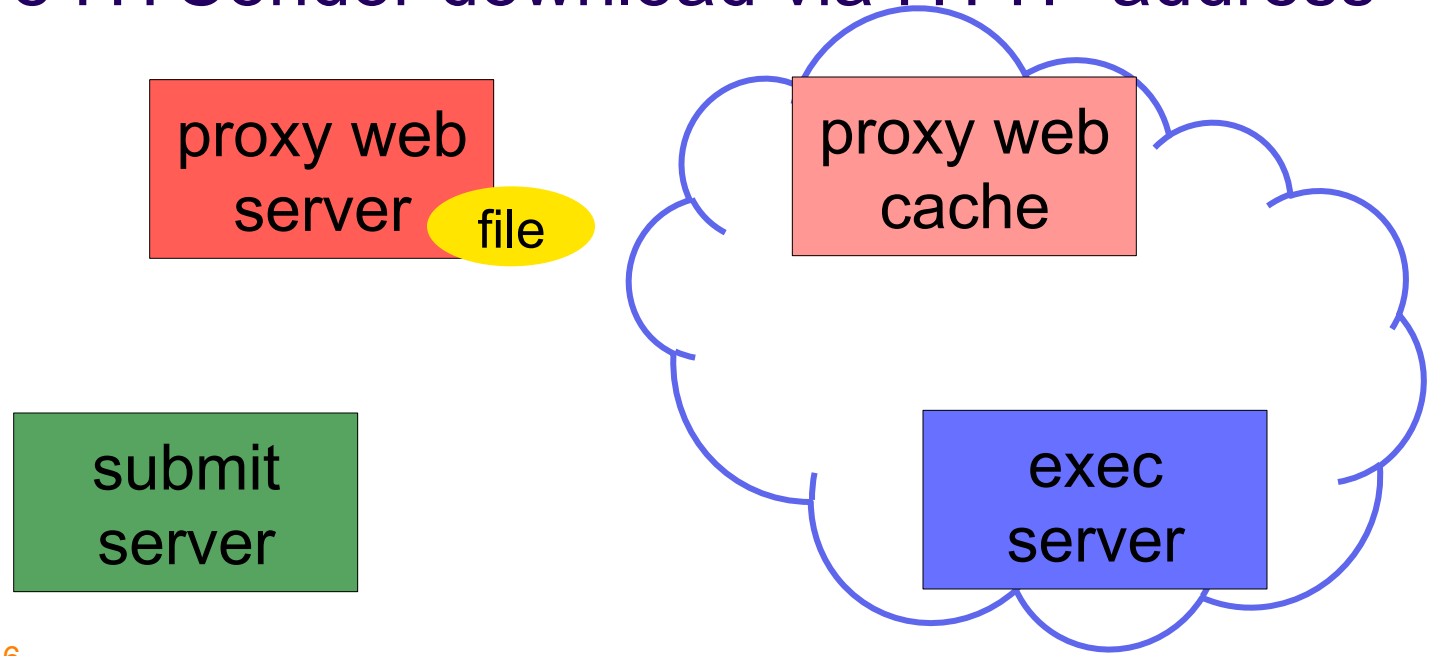
Using a Web Proxy

- Place the file onto a proxy-configured web server
- Have HTCondor download via HTTP address



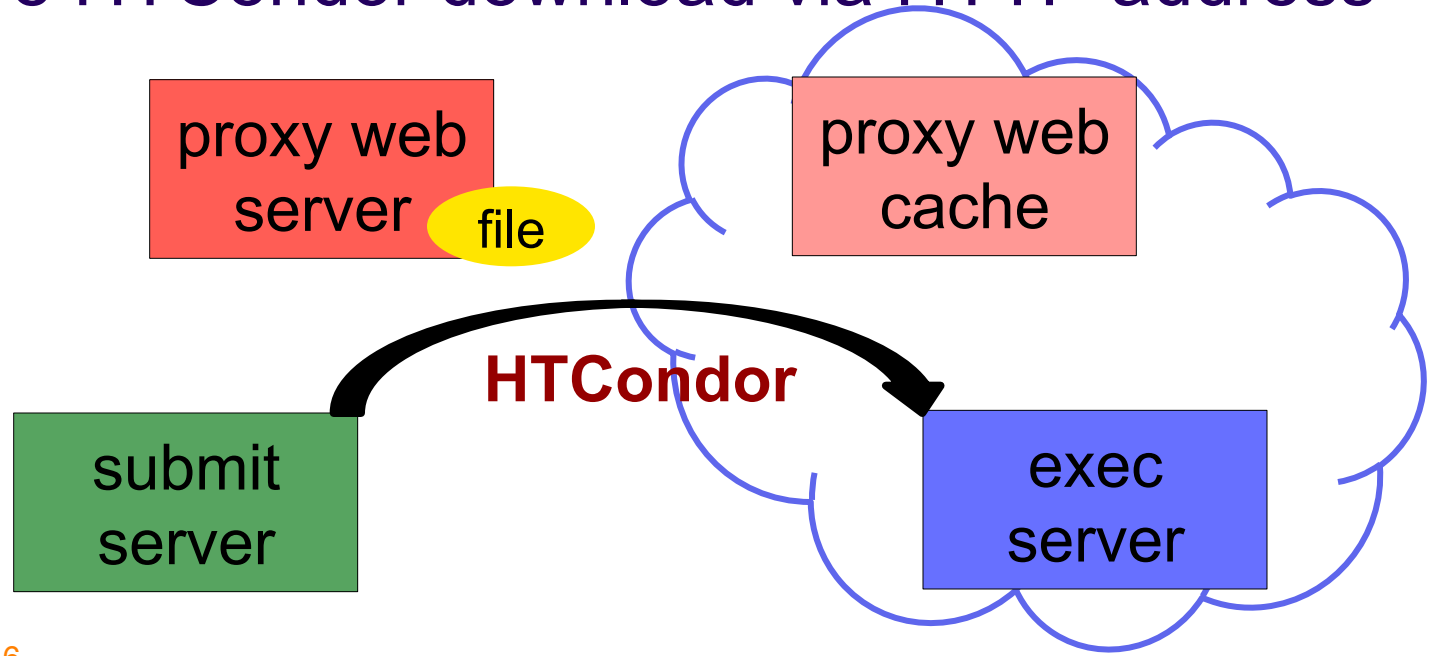
Using a Web Proxy

- Place the file onto a proxy-configured web server
- Have HTCondor download via HTTP address



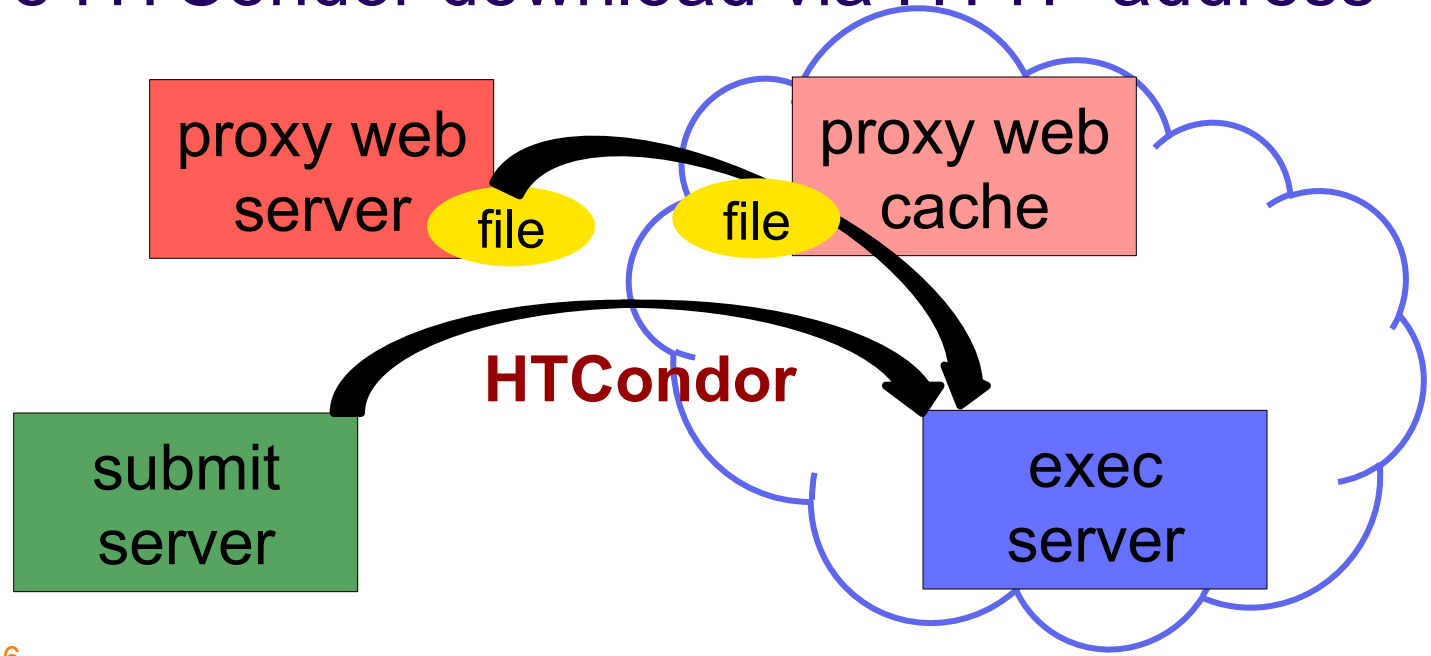
Using a Web Proxy

- Place the file onto a proxy-configured web server
- Have HTCondor download via HTTP address



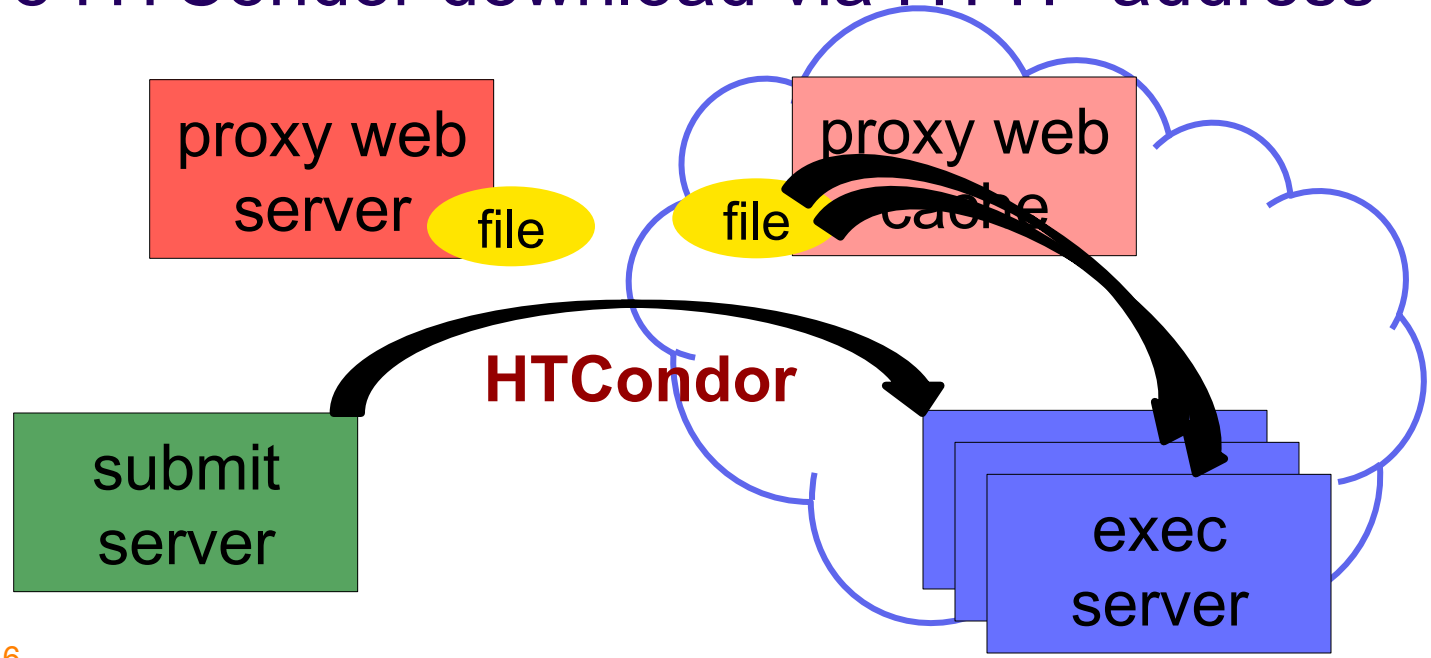
Using a Web Proxy

- Place the file onto a proxy-configured web server
- Have HTCondor download via HTTP address



Using a Web Proxy

- Place the file onto a proxy-configured web server
- Have HTCondor download via HTTP address



Downloading Proxy Files

- HTCondor submit file: (recommended)

```
transfer_input_files = http://host.univ.edu/path/to/
shared.tar.gz
```

- Anywhere (in-executable, or test download)

```
wget http://host.univ.edu/path/to/shared.tar.gz
```

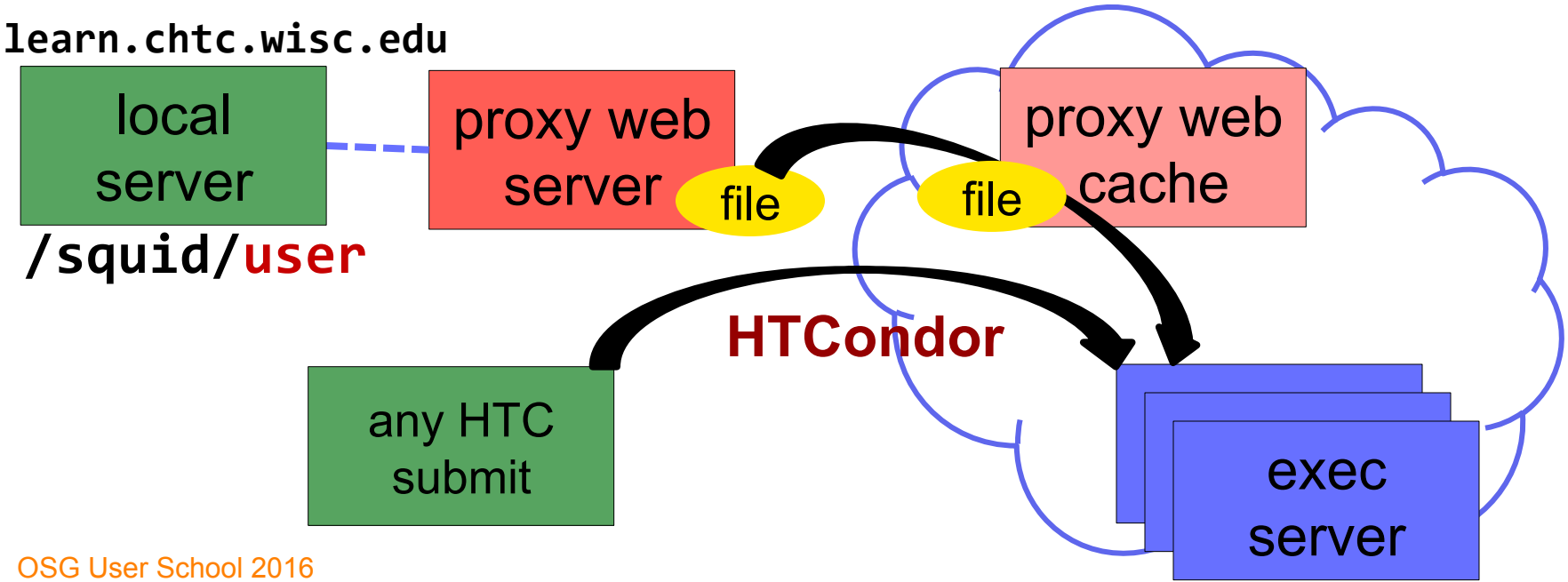
  - in-executable: make sure to delete after un-tar or at the end of the job!!! (HTCondor thinks it's 'new')

Web Proxy Considerations

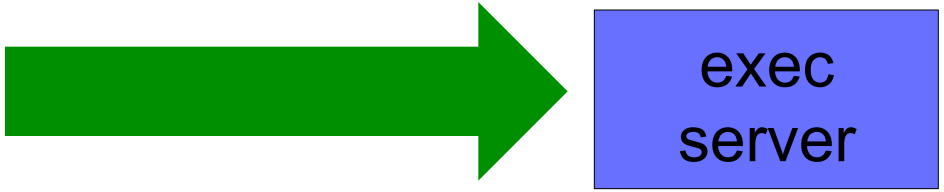
- Managed per-VO
- Memory limited, **max file size: 1 GB**
- Local caching at OSG sites
  - good for *shared* input files, only
  - perfect for software and common input
  - need to rename changed files!!!
- Files are downloadable by **ANYONE** who has the specific HTTP address
  - Will work on 100% of OSG sites, though not all sites will have a local cache

At UW-Madison (Ex. 3.1)

- place files in `/squid/user` on a local submit server
- address: `http://proxy.chtc.wisc.edu/SQUID/user/shared.tar.gz`



Large input in HTC and OSG



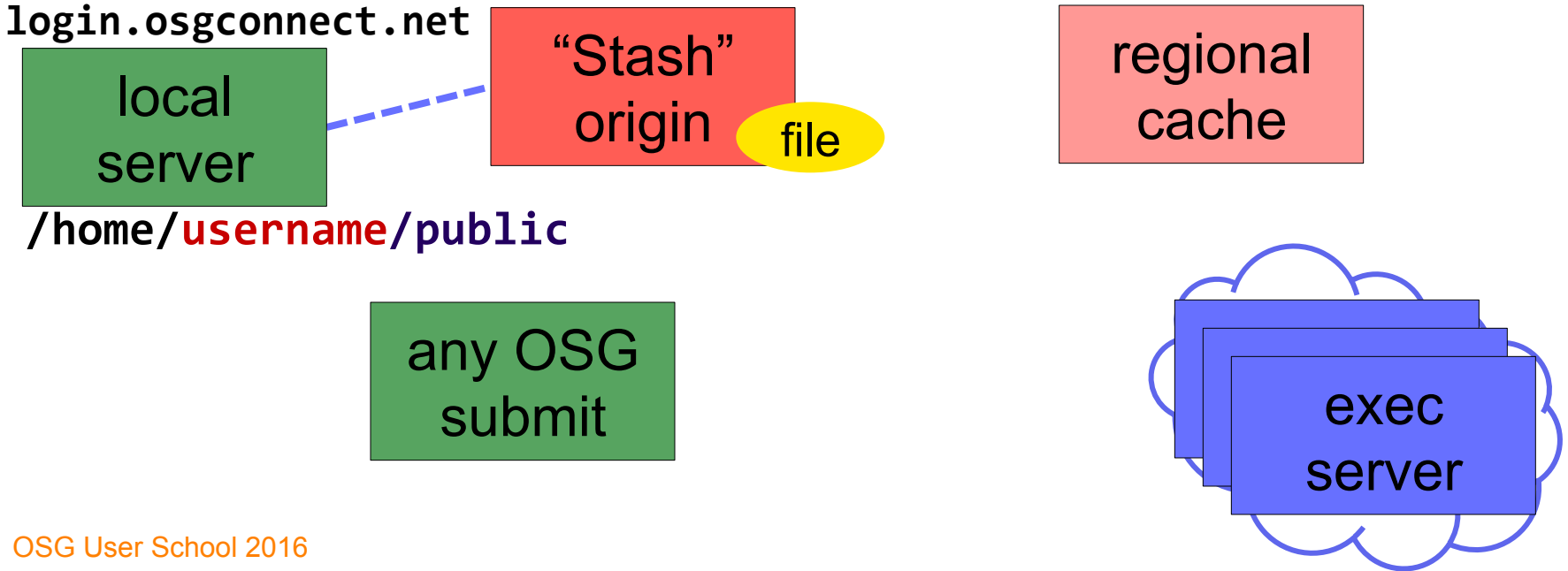
file size	method of delivery
words	within executable or arguments?
tiny – 10MB per file	HTCondor file transfer (up to 1GB total per-job)
10MB – 1GB, shared	download from web proxy (network-accessible server)
1GB - 10GB, unique or shared	StashCache (regional replication)
10 GB - TBs	shared file system (local copy, local execute servers)

Using StashCache for Input

- regionally-cached repository managed by OSG Connect

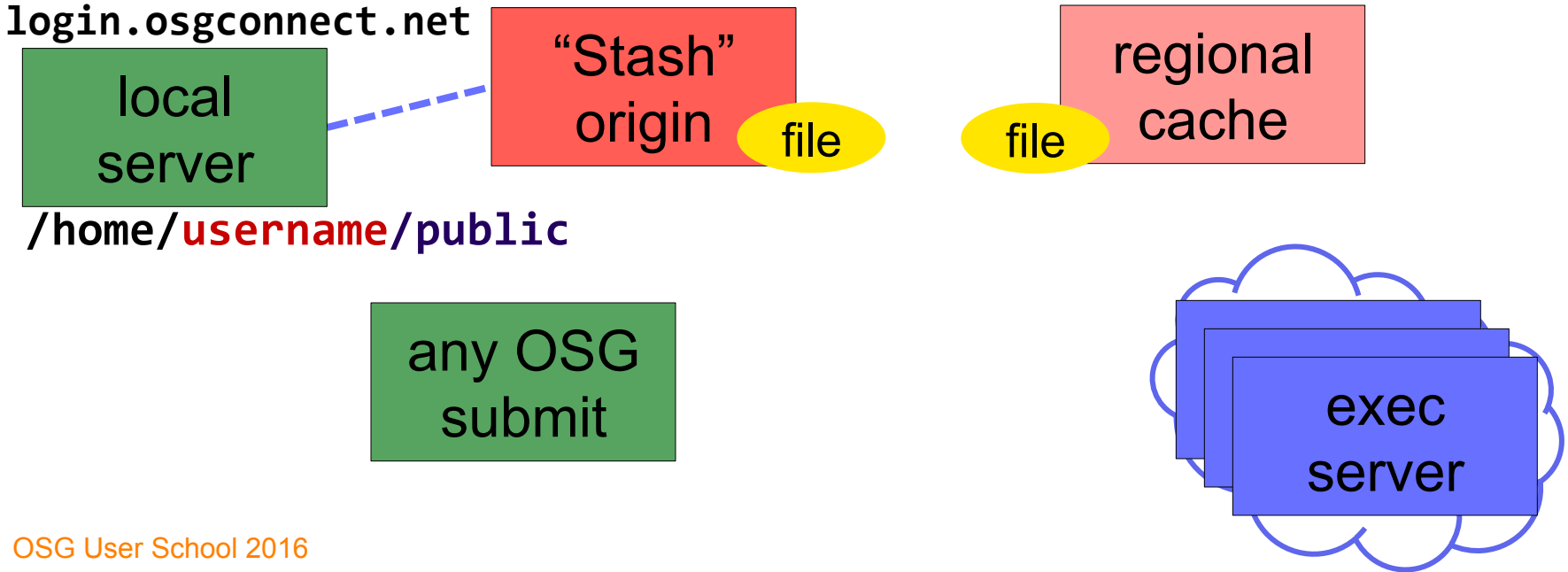
Placing Files in StashCache

- place files in `/home/user/public` on `login.osgconnect.net`



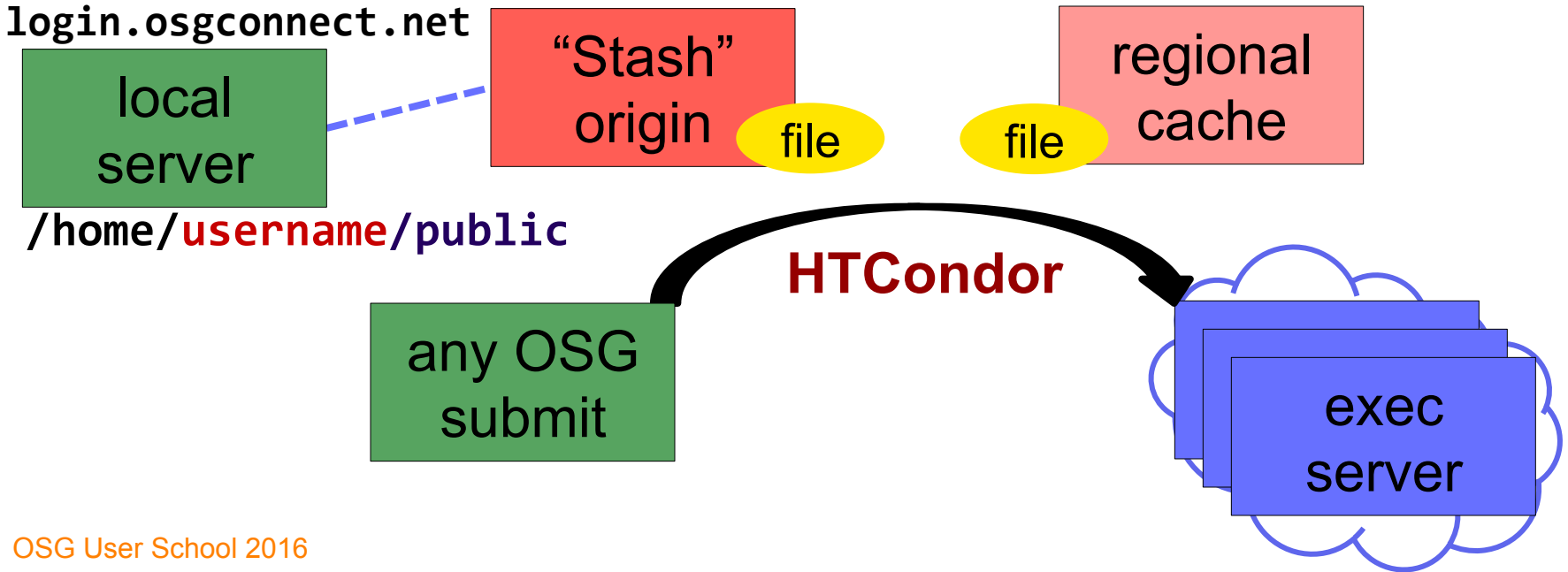
Placing Files in StashCache

- regional cache updates from origin every hour



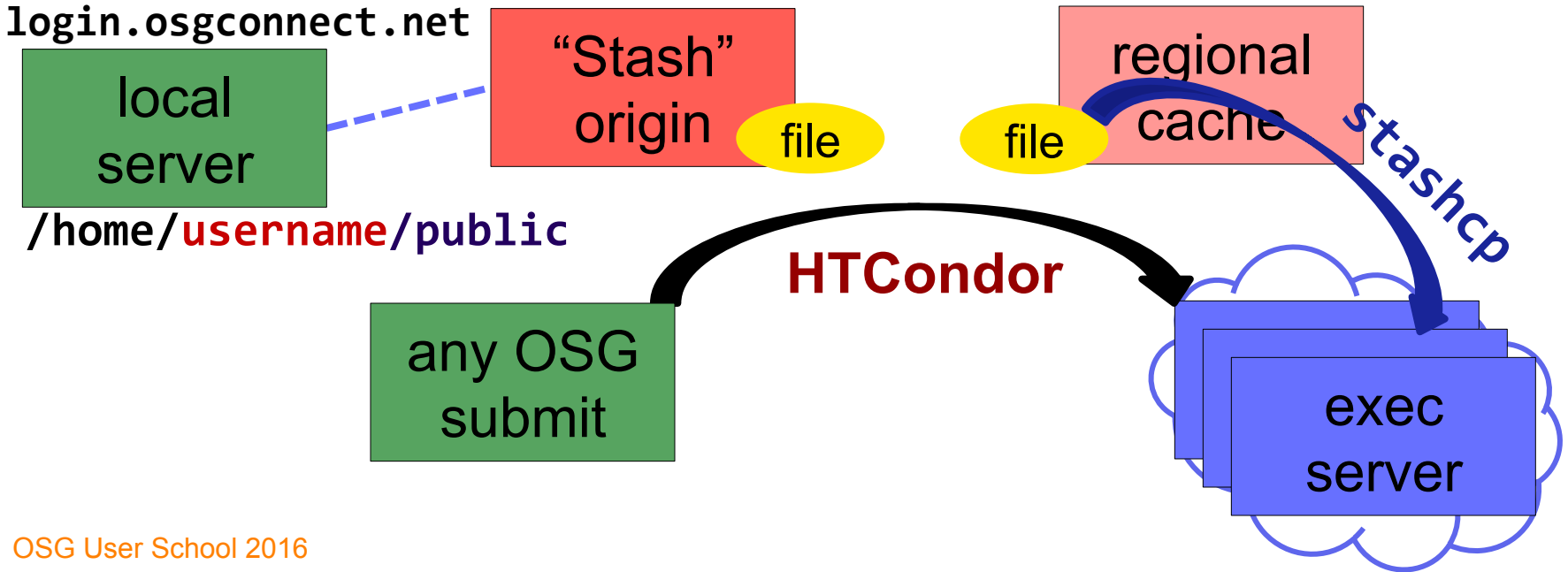
Obtaining Files in StashCache

- Use HTCondor transfer for other files



Obtaining Files in StashCache

- Download using `stashcp` command (available as an OASIS software module)



In the Submit File

- Require StashCashe sites in the submit file

```
+WantsStashCache
```

- Require sites with OASIS modules (for `stashcp`)

```
Requirements = <other &&> (HAS_MODULES =?= true)
```

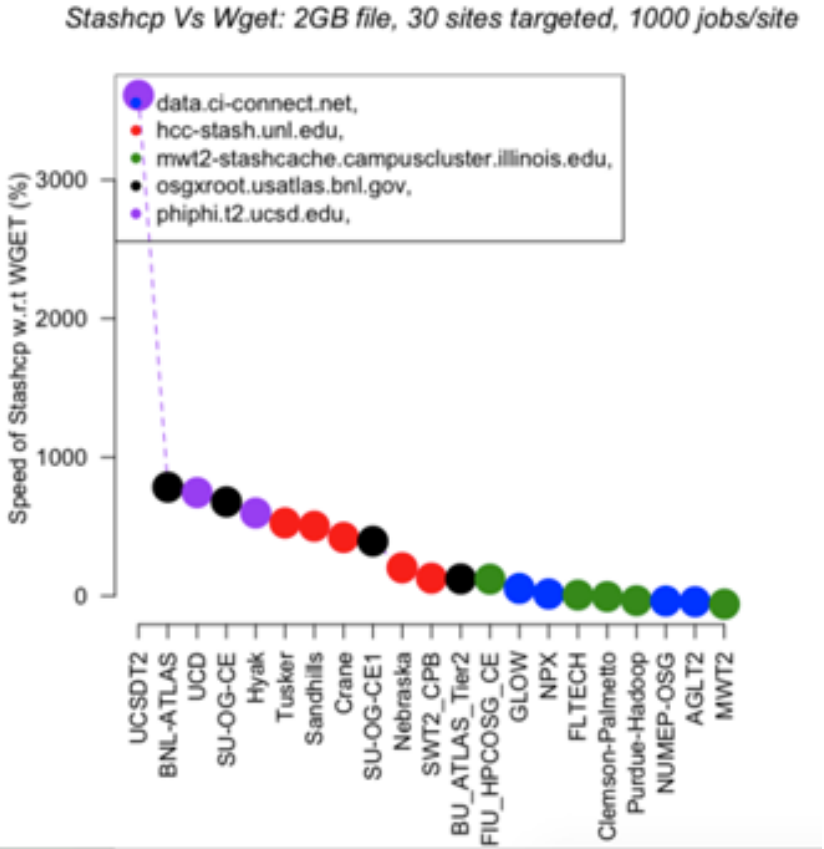
In the Job Executable

```
#!/bin/bash
# setup:
. /cvmfs/oasis.opensciencegrid.org/osg/modules/lmod/current/init/bash
module load stashcp
stashcp /user/username/public/file.tar.gz ./

<untar, then remove the file>
<job commands>

<remove all files from StashCache>
# END
```

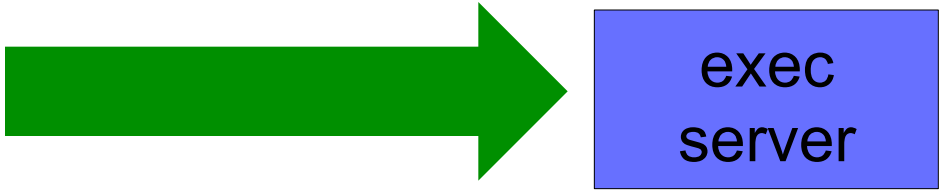
StashCache Considerations

- Available at ~90% of OSG sites
- Regional caches on *very fast* networks
  - **Max file size: 10 GB**
  - *shared* OR *unique* data
- Caches are updated ~hourly
  - rename files to update them (safest)
- Currently in transition to a new method, but `stashcp` will stay around!!

StashCache Speed



Large input in HTC and OSG



file size	method of delivery
words	within executable or arguments?
tiny – 10MB per file	HTCondor file transfer (up to 1GB total per-job)
10MB – 1GB, shared	download from web proxy (network-accessible server)
1GB - 10GB, unique or shared	StashCache (regional replication)
10 GB - TBs	shared file system (local copy, local execute servers)

Other Options?

- Some distributed projects with LARGE, shared datasets may have project-specific repositories that exist only on certain sites
  - (e.g. CMS, Atlas, LIGO?, FIFE?, others?)
  - Jobs will require specific sites with local copies and use project-specific access methods
- OASIS?
  - Best for lots of small files per job (e.g. software)
  - StashCache and Proxies better for fewer larger files per job

Cleaning Up Old Data

- For StashCache *AND* web proxies:

make sure to delete data when you no longer need it in the origin!!!

- StashCache and VO-managed web proxy servers do NOT have unlimited space!
  - Some may regularly clean old data for you. Check with local support.

Other Considerations

- Only use these options if you MUST!!
  - Each comes with limitations on site accessibility and/or job performance, and extra data management concerns

file size	method of delivery
words	within executable or arguments?
tiny – 10MB per file	HTCondor file transfer (up to 1GB total per-job)
10MB – 1GB, shared	download from web proxy (network-accessible server)
1GB - 10GB, unique or shared	StashCache (regional replication)
10 GB - TBs	shared file system (local copy, local execute servers)

Open Science Grid

Exercises

- 3.1 Using a web proxy for shared input
  - place the blast database on the web proxy
- 3.2 StashCache for shared input
  - place the blast database in StashCache
- 3.3 StashCache for unique input
  - convert movie files

Questions?

- Feel free to contact me:
  - lmichael@wisc.edu

- Next: Exercises 3.1-3.3
- Later: Large *output* and shared filesystems