CVMFS for Data Federations

Derek Weitzel
University of Nebraska - Lincoln

Problem with Data Federations

- Users must know the exact filenames for each job.
- They have to use special tools they are unfamiliar with in order to use it (such as `xrdcp` or `stashcp`).
  - Applications may only talk POSIX.
- They are difficult to setup for opportunistic VO's; OSG has already created one StashCache.

Changes to CVMFS

- As discussed in Brian's talk on yesterday, changes in CVMFS developed by him and I have enabled CVMFS's use in data federations.
- CVMFS can now access data federations through HTTP gateways.
  - Metadata (catalogs) come from the normal OASIS Stratum-1 infrastructure.
  - Data files come from the data federation.

Changes to CVMFS

- File accesses can be redirected to another server
- Files that are retrieved from this other server are not in standard CVMFS hashed' format
- Rather, they are **uncompressed**.
- Instead they are pointers to a file on another server, i.e. a XRootD server.

Repositories

- **nova.osgstorage.org** - Repo from XrootD data source at FNAL

- **stash.osgstorage.org** - Repo built from user accessible storage at OSG-Connect

- **cms.osgstorage.org** - Repo of the CMS data federation

- **ligo.osgstorage.org** - Repo of LIGO data stored at Nebraska

Repositories

- **nova.osgstorage.org** - Repo from XrootD data source at FNAL

- **stash.osgstorage.org** - Repo built from user accessible storage at OSG-Connect

- **cms.osgstorage.org** - Repo of the CMS data federation

- **ligo.osgstorage.org** - Repo of LIGO data stored at Nebraska

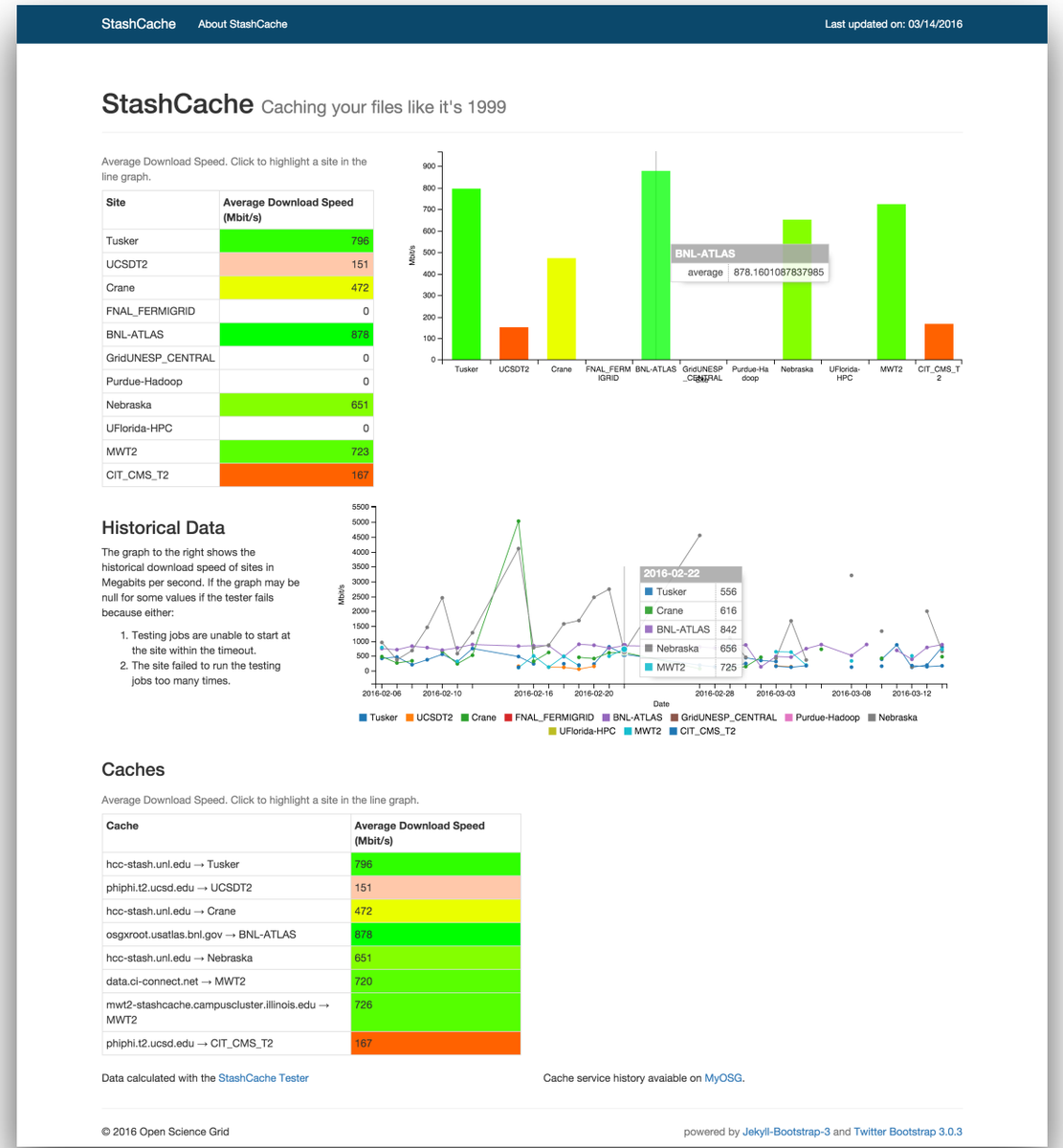
stash.osgstorage.org

1. At a CVMFS repo, HTCondor cron job scans the Stash filesystem at UChicago, recording differences since last scan.
    - This looks at the contents of `/stash/$USER/public` found on OSG-Connect.
2. Job puts records files' metadata (size, checksum) into the CVMFS repository server. Data stays on Stash.
3. CVMFS repository is published with new contents.

StashCache

- Managing data opportunistically at storage elements requires a CMS- or ATLAS-sized commitment.
- StashCache uses distributed caches across the country.
- Data origin is the Stash service on OSG-Connect.
- Users write data into Stash, and read the data from jobs through StashCache

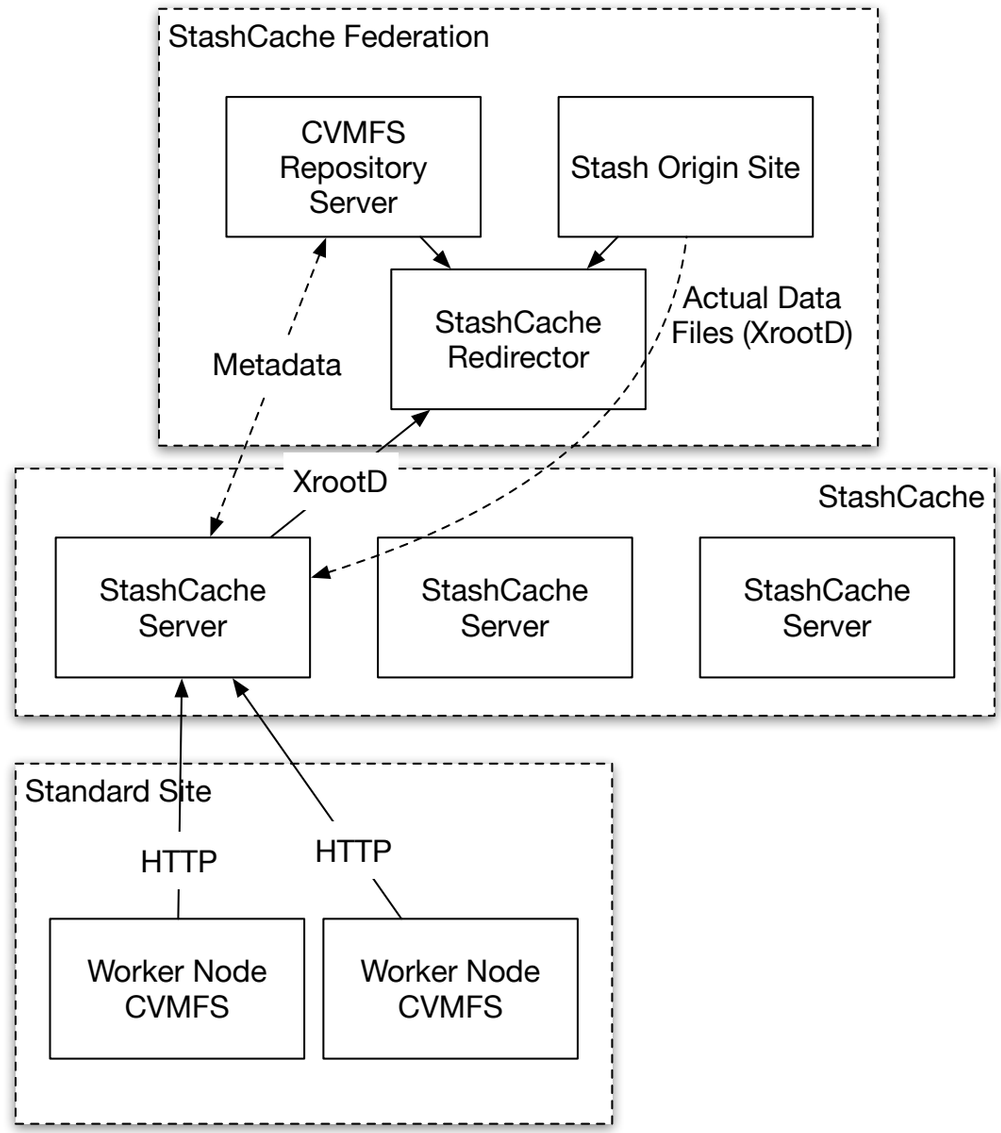
stashcache.github.io



For a full overview of StashCache, see Brian's talk from last years AHM.

Overview of CVMFS and StashCache

- Regular XrootD StashCache Federation
- CVMFS contacts the caching servers over HTTP
- Caching servers contact the federation for the data
- Worker nodes pull data from the caching servers



Uses

- Large datasets which cannot be cached with Squid
  - Full Blast Db's
  - Nova Flux Files…
- Targeting working set sizes* from 10GB to 10TB. Will work fine for smaller sizes, but OASIS may be more efficient for distribution.

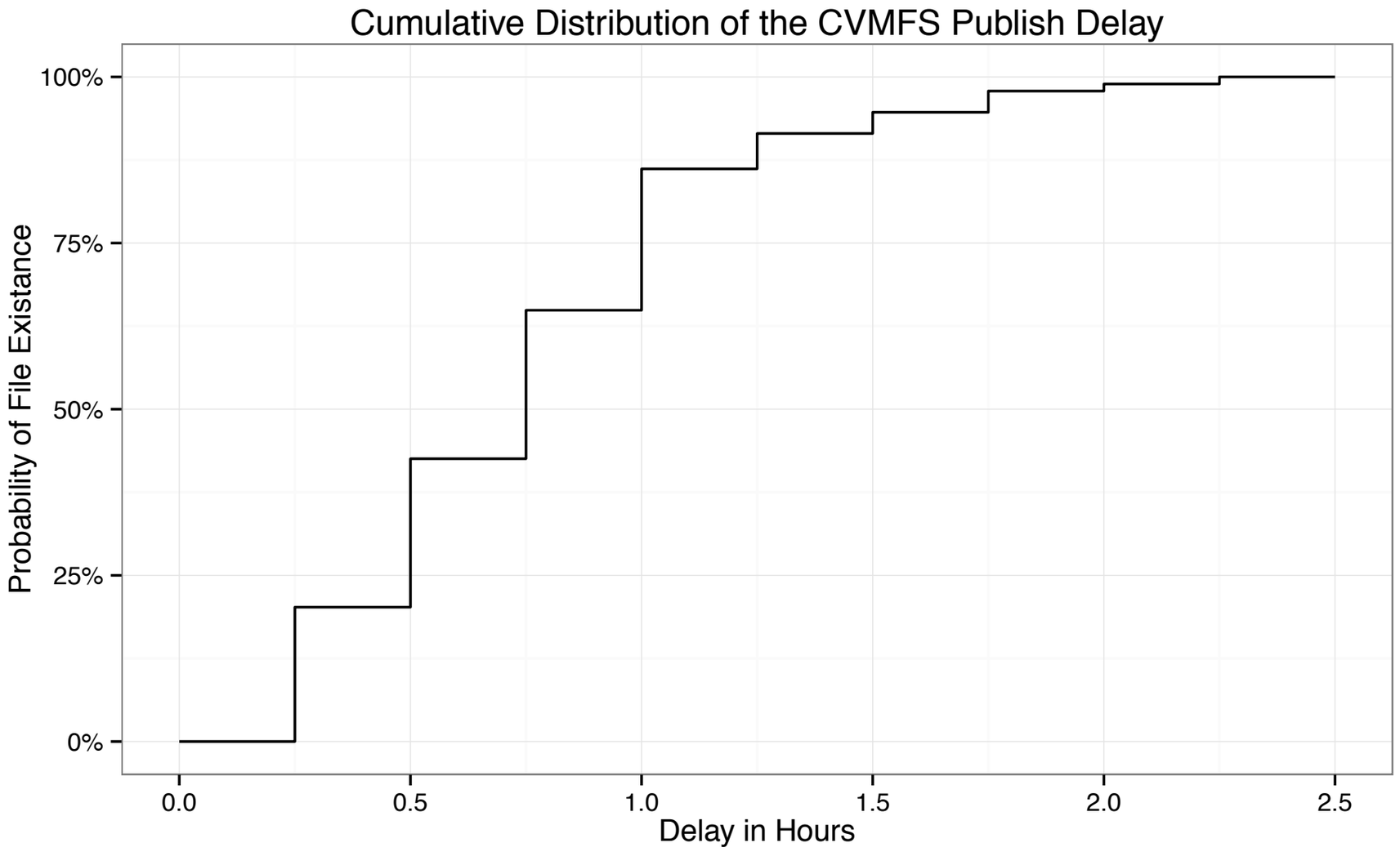
*Number of unique bytes touched in 24 hours

User Perspective

- Copies data onto OSG-Connect using `scp`, Globus Online - pick your favorite.
- Put data into `/stash/<user>/public`
- Wait for a while for the data to be published (~1 hr)
- Use data on the worker nodes!

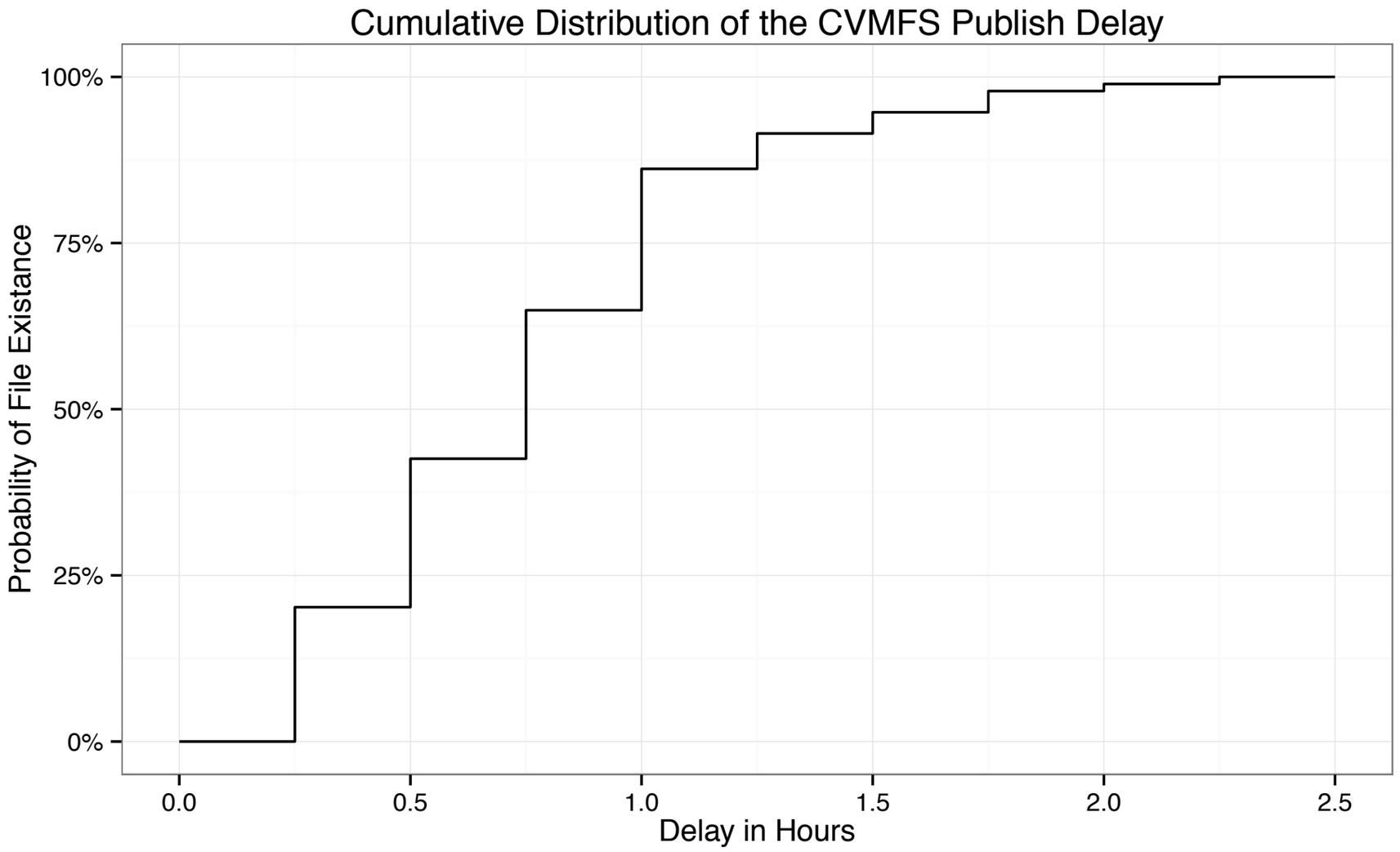
Stash -> CVMFS Delay

- There is a delay between when the file has been created, and when the it appears in CVMFS.



Stash -> CVMFS Delay

- In 1 hour, the files are largely available



CVMFS + StashCache

- This creates a global read-only filesystem
- Originally, a select few could put data into cvmfs using services such as Oasis
- Now, everyone with can add their own files and software into CVMFS
- Access at /cvmfs/stash.osgstorage.org/

ligo.osgstorage.org

- stash.osgstorage.org is unauthenticated access to public files

- LIGO has very specific rules about data access and even namespace visibility

- Therefore, had to develop new features in CVMFS to enable VOMS authentication.

Secure CVMFS

- Pull certificate from the user's environment
- Namespace is protected by authenticated access to CVMFS HTTP(S) server.
- Data is authenticated with XRootD HTTP(S) client authentication.

Secure CVMFS

- Special setup of HTTP server for authenticate setup (mod_gridsite)

- XrootD serves data directly from data servers, cannot currently proxy authenticated access.

What you can do!

- Update CVMFS on your worker nodes to 2.2 preview:

```
yum install --enablerepo=osg-upcoming cvmfs cvmfs-config-osg
```

- Feel free to install this locally and test the interface on OSG-Connect.

- Requires sites to upgrade their CVMFS client: widespread availability will probably occur in July.