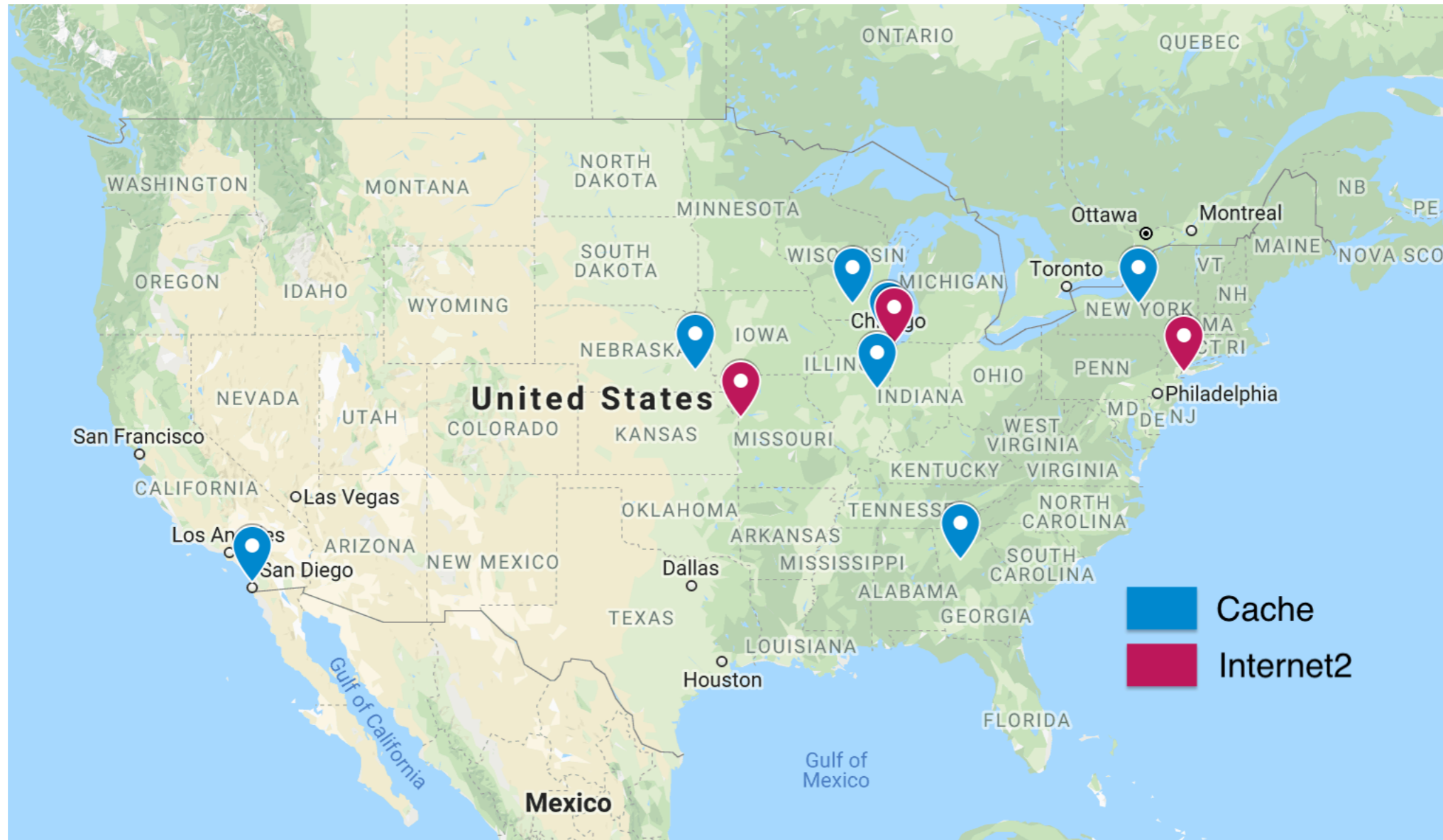


OSG Data Federation

StashCache

Derek Weitzel
University of Nebraska - Lincoln

Big Picture

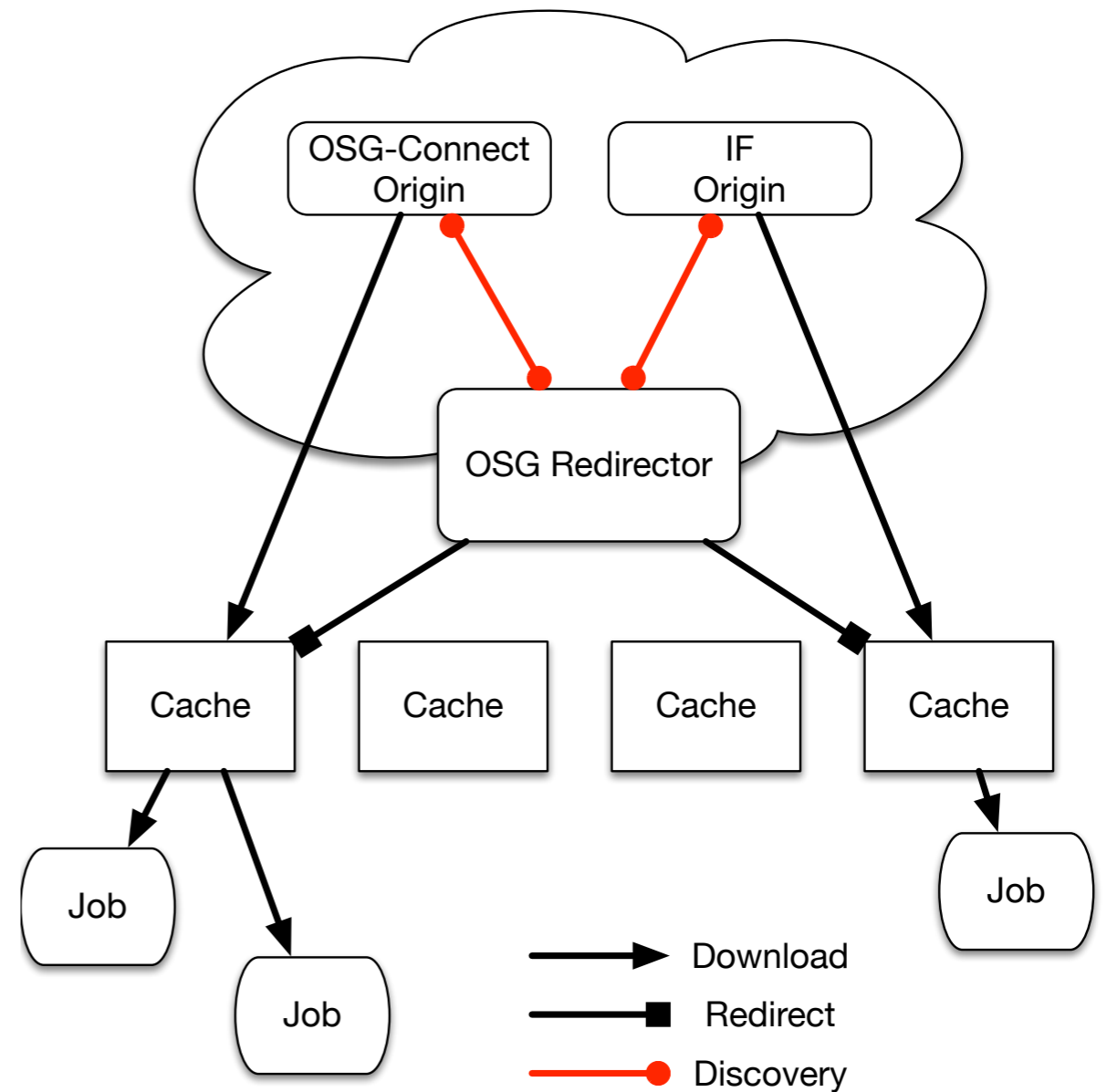


Big Picture

- General Science caching network
- Provides data delivery for multiple science domains
- Cache hosting sites are HEP (UCSD, Nebraska) and non-HEP (Internet2)

Overview

- Clients use GeoIP to determine “near” cache
- Origins for each science project



Clients

- CVMFS
 - Uses CVMFS's external data capabilities.
 - Metadata is in CVMFS, actual data is pulled through HTTP interface on XCache.
 - Data is available after a scan from CVMFS scanner
 - 30 min to 12 hours after data is on origin.
 - Secured access
- StashCP
 - Custom download script
 - Data is immediately available

Clients - CVMFS

- Meta-data is from CVMFS catalogs

```
$ ls -l /cvmfs/gwosc.osgstorage.org/README.gwosc  
-rw-r--r-- 1 cvmfs cvmfs 119 Mar 20 10:43 /cvmfs/gwosc.osgstorage.org/README.gwosc
```

- Checksums are kept in CVMFS
- Data is from XCache

```
$ sha1sum /cvmfs/gwosc.osgstorage.org/README.gwosc  
9bbca126599cbc032c031ed977aee76f4891b7c1 /cvmfs/gwosc.osgstorage.org/README.gwosc
```

Origins

- Origins export exclusive namespaces
- Namespaces are registered centrally

Cache	Namespace
Wisconsin	/chtc
OSG Connect Stash	/user
UConn GlueX	/GlueX
Dark Energy Survey	/desdm
Caltech LIGO Open Data	/gwdata
FNAL	/pnfs/ fnal.gov

Caches

- Caches located at contributing institutions and Internet 2 access points
- Caches are split between baremetal and Kubernetes

Authenticated Access

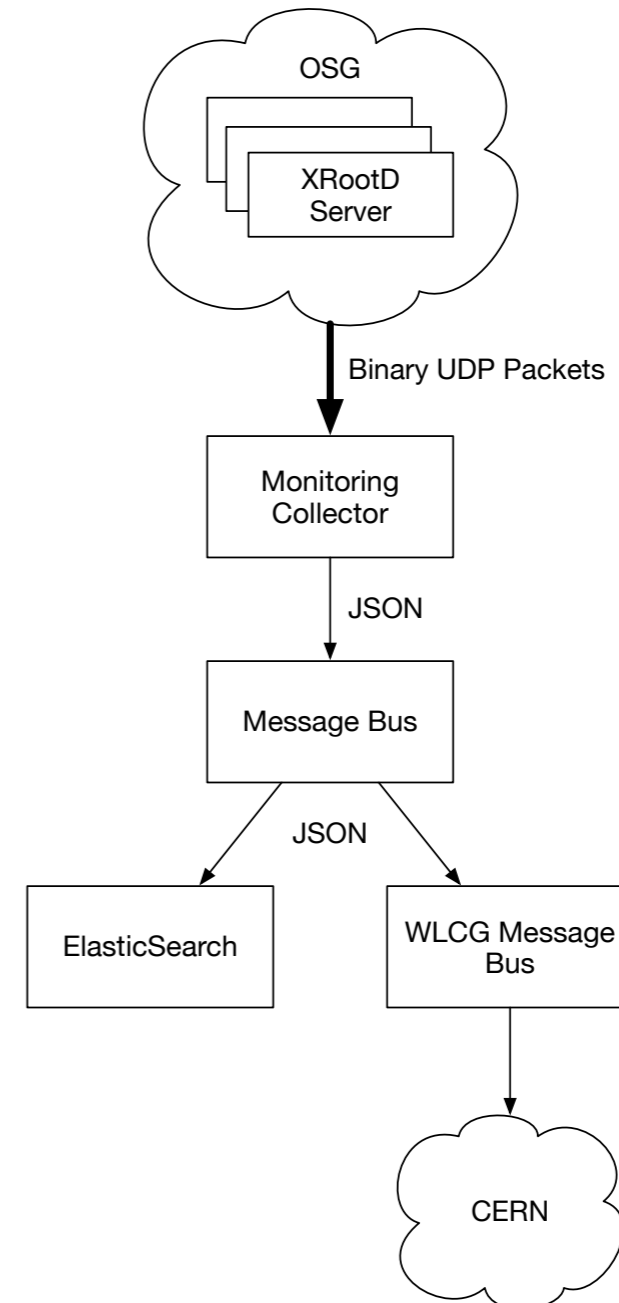
- Some experiments require restricted access
- CVMFS keeps a whitelist of accepted DN's
- Certificates are checked at the worker node, then passed to the cache for authentication
- XCache uses its own host cert to pull from origin.
- Added token based authentication to CVMFS, will begin using

Authenticated Access

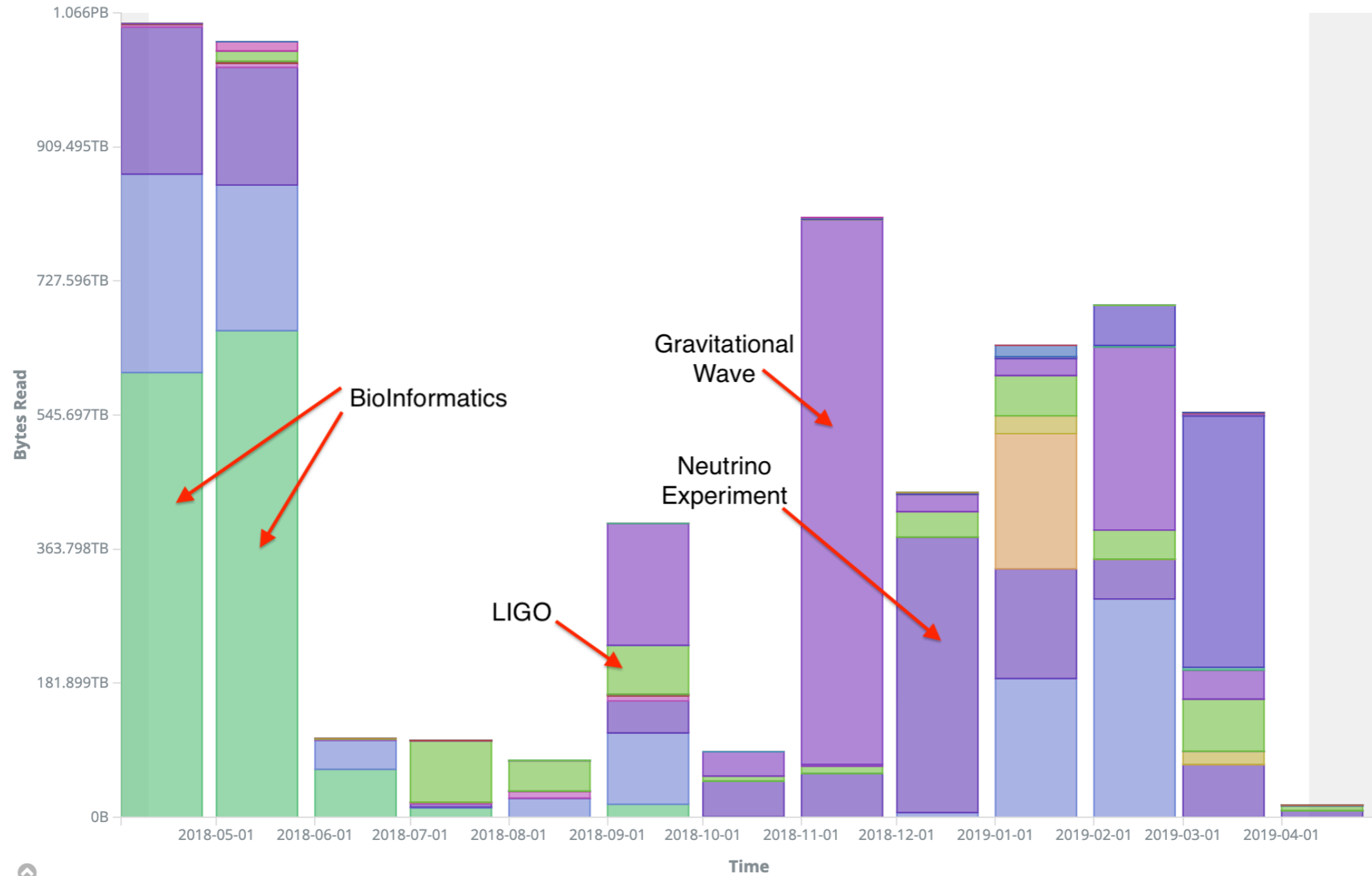
- Every StashCache Cache and Origin is registered in the OSG Topology service
- Authfile is generated on-demand from registration information.
 - What VOs and DNs are allowed at the cache.

Monitoring

- Use XRootD “f” stream
- Monitoring collector keeps state and sends file close events to message bus
- Stores all accesses in Elasticsearch



Operational Experience



Working Set Sizes

Directory ↕	Working Set ↕	Total Read ▼	Caching Ratio
/pnfs/fnal.gov/usr/minerva	655.068GB	615.889TB	940x
/user/ligo	25.146TB	454.02TB	18x
/user/dteam	3.287TB	390.025TB	118x
/user/bbockelm	4.096GB	186.331TB	46000x (testing)
/pnfs/fnal.gov/usr/nova	162.644GB	60.065TB	370x
/user/jeanjack	24.177GB	30.102TB	1250x
/user/cwalter	54.492GB	15.253TB	
/user/dweitzel	59.454GB	2.883TB	
/pnfs/fnal.gov/usr/sbnd	2.234GB	2.008TB	
/user/irisqing	545.226MB	1.442TB	

Future & Requests

- Token authorized caching (SciToken or equivalent)
- Protocol specific redirections
 - HTTP request will redirect to HTTP, xroot to xroot

Summary

- We see infrequent but very significant use of the caches
- The most heavily used datasets in the caches gives a high reuse ratio
- Significant effort to package XCache for StashCache, ATLAS, and CMS
- We run weekly XCache meetings to coordinate effort between ATLAS/CMS/OSG cache infrastructure and provide feedback to XRootD developers <http://bit.ly/xcache-meetings>
- StashCache is part of the OSG software stack releases (see BrianL talk)