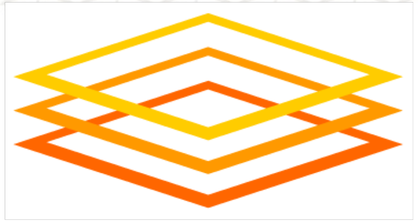
OSG Technology Updates for 2019

MORGRIDGE
INSTITUTE FOR RESEARCH



FEARLESS SCIENCE

iris hep

2018: Year In (Technology) Review

What did we do in the last year?

- Changed some technologies – such as retired GUMS & VOMS-Admin.
- Moved off of Globus Toolkit…
- … to the Grid Community Toolkit, a community-supported fork!
- Complete review of all our operational services: each one was migrated, re-envisioned, or shut down.

We were not bored in 2018

Why do we care about technology?

Why does OSG technology matter? Why evolve?

- Technology does not do the science.
- We have already shown we can compute and move data "at scale".

So why change?

- It *does* allow us to reach new communities and it *does* allow us to simplify operations.
- For the OSG-LHC community, OSG provides an on-ramp for technologies such as those coming from IRIS-HEP (SSL or DOMA areas) or the US LHC communities.

Standard Disclaimer:

The real magic is the facilitation staff, not the technology.

Thinking about Technology

OSG is an organization that is thinking deeply about what we offer.

Looking to move to a series of services that can be *allocated* for communities.

Open Science Grid

UCSD

Our Aspirational Goal:

A campus Research IT Organization should not have to learn anything ”non-standard” in order to have their researchers benefit from OSG, or have their resources be available via OSG.

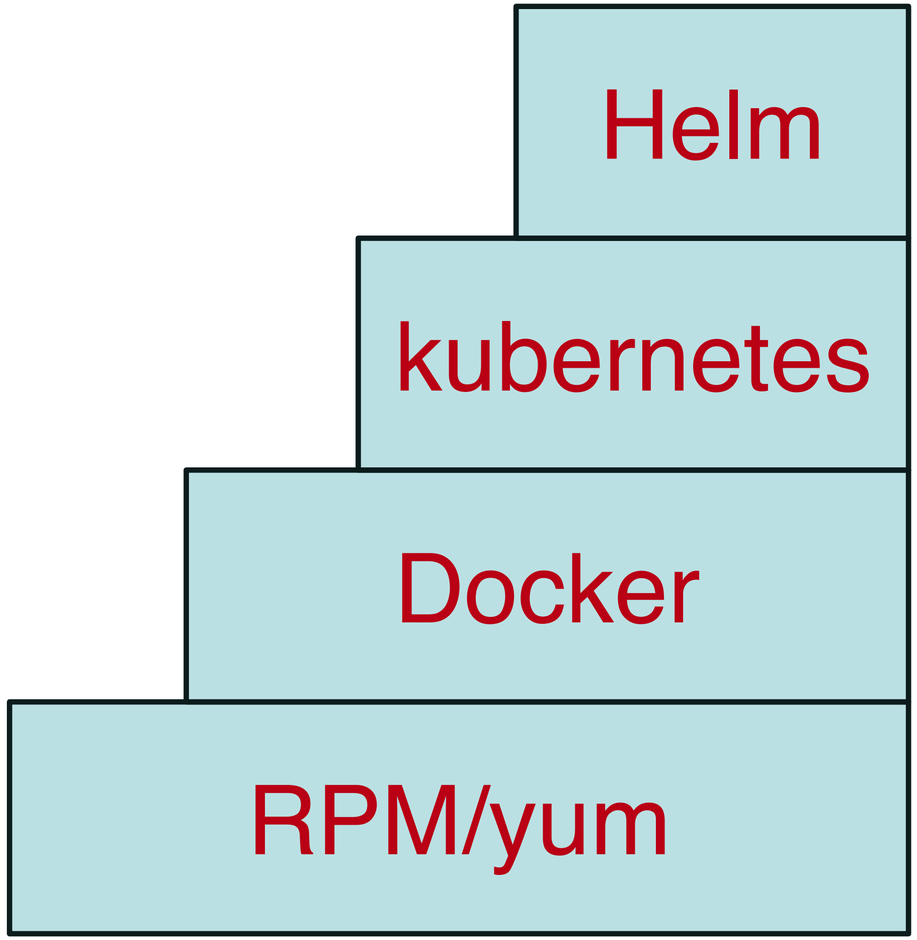
Well, …. we got some ways to go before we reach that goal …

Where Next? Packaging and Deployment

We have been happily building and handing off RPMs to OSG sysadmins for years. This has been a strength of our team.

However, additional mechanisms provides new “product channels” that offer new strengths to sysadmins.

Goal: treat it as a hierarchy. Keep a quality base and build up toward higher-level functionality.



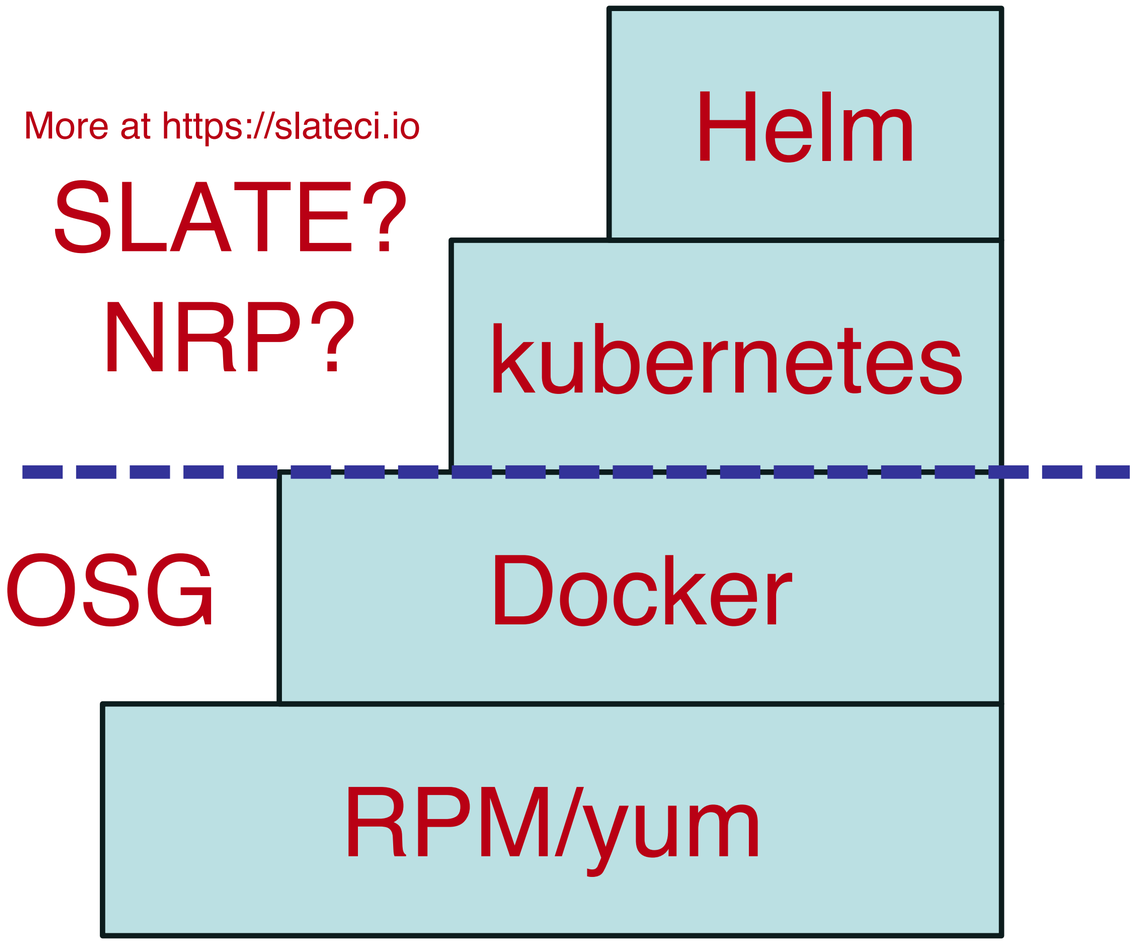
Where Next? Packaging and Deployment

We now have Docker images ready for use for XCache and the OSG worker node.

These are quality, stand-alone Docker images – but the first thing everyone does is run them in k8s.

Right now, the SLATE team is producing Helm charts; where to put the layer between “OSG” and “other” is unclear.

View from March 2019

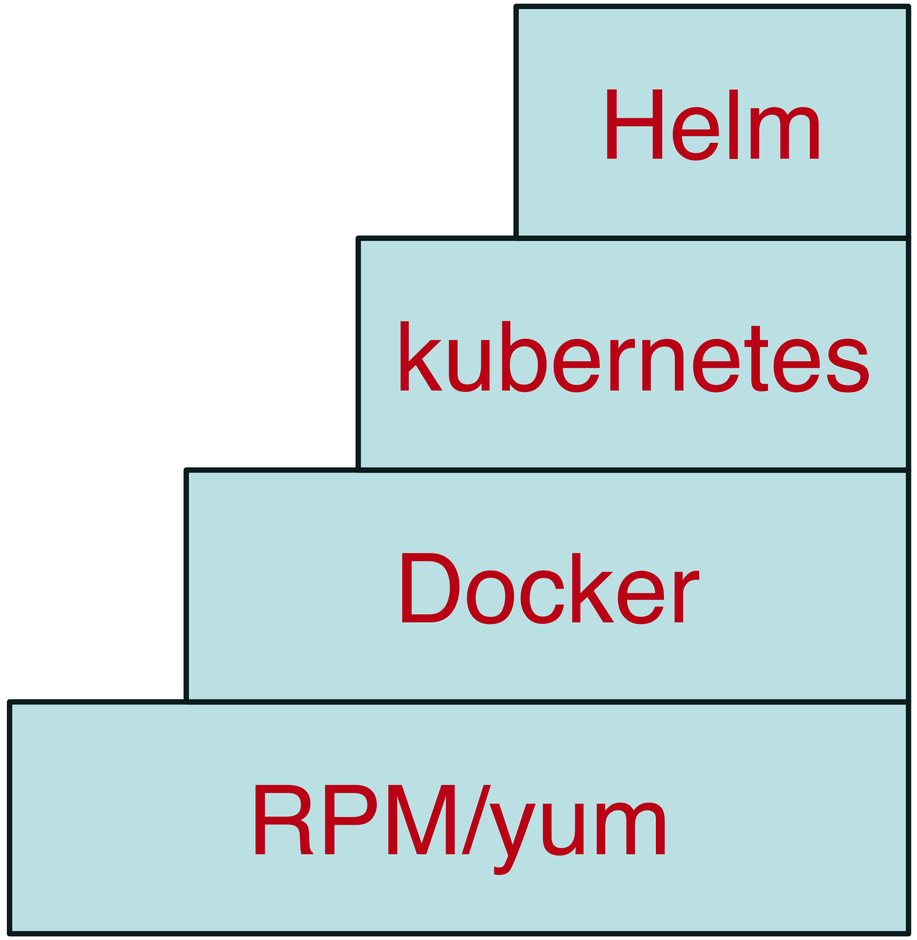


Moving from Packages to Applications

We are slowly migrating from having only a package repository to producing a series of applications.

This allows us to ship when the *application* is ready as opposed to all use cases.

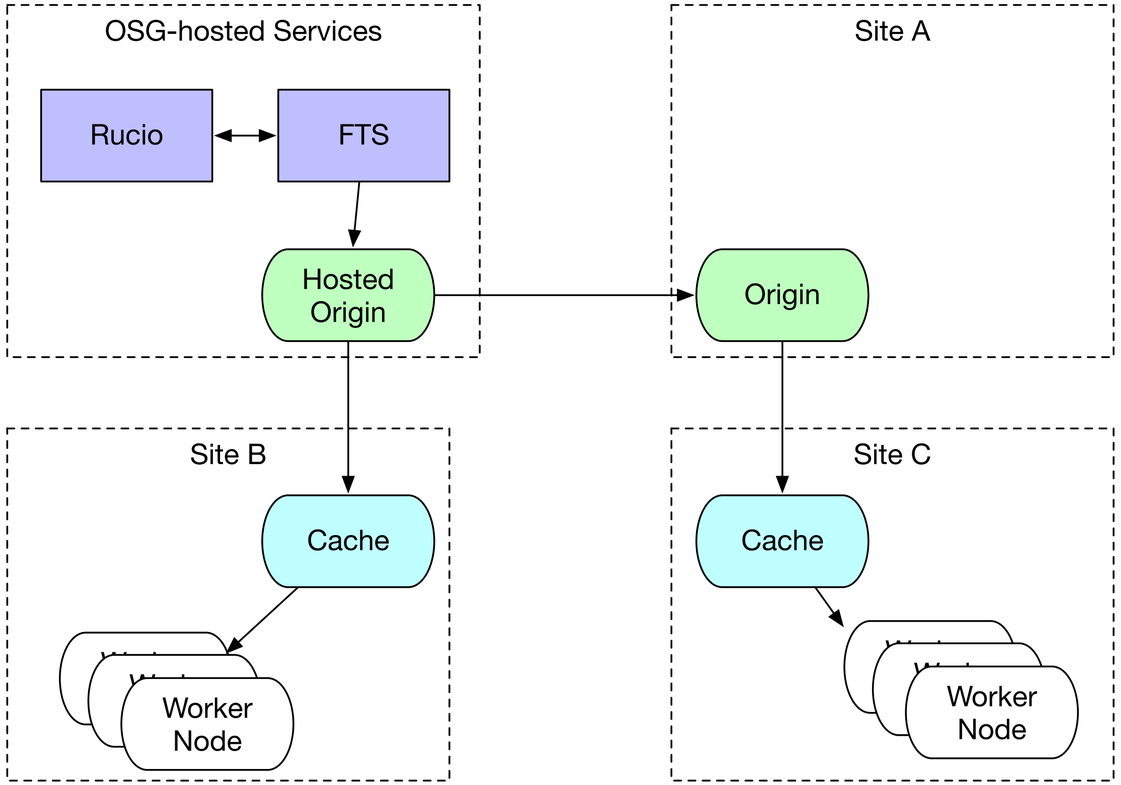
Example: we could ship an XCache container based on XRootD 4.9.0 despite the fact a GridFTP/XRootD bug blocks us from shipping in the yum repository.



Refocusing our Data Management

We have been slowly refocusing our data services from solely considering a storage element to one including *origins* and *caches*:

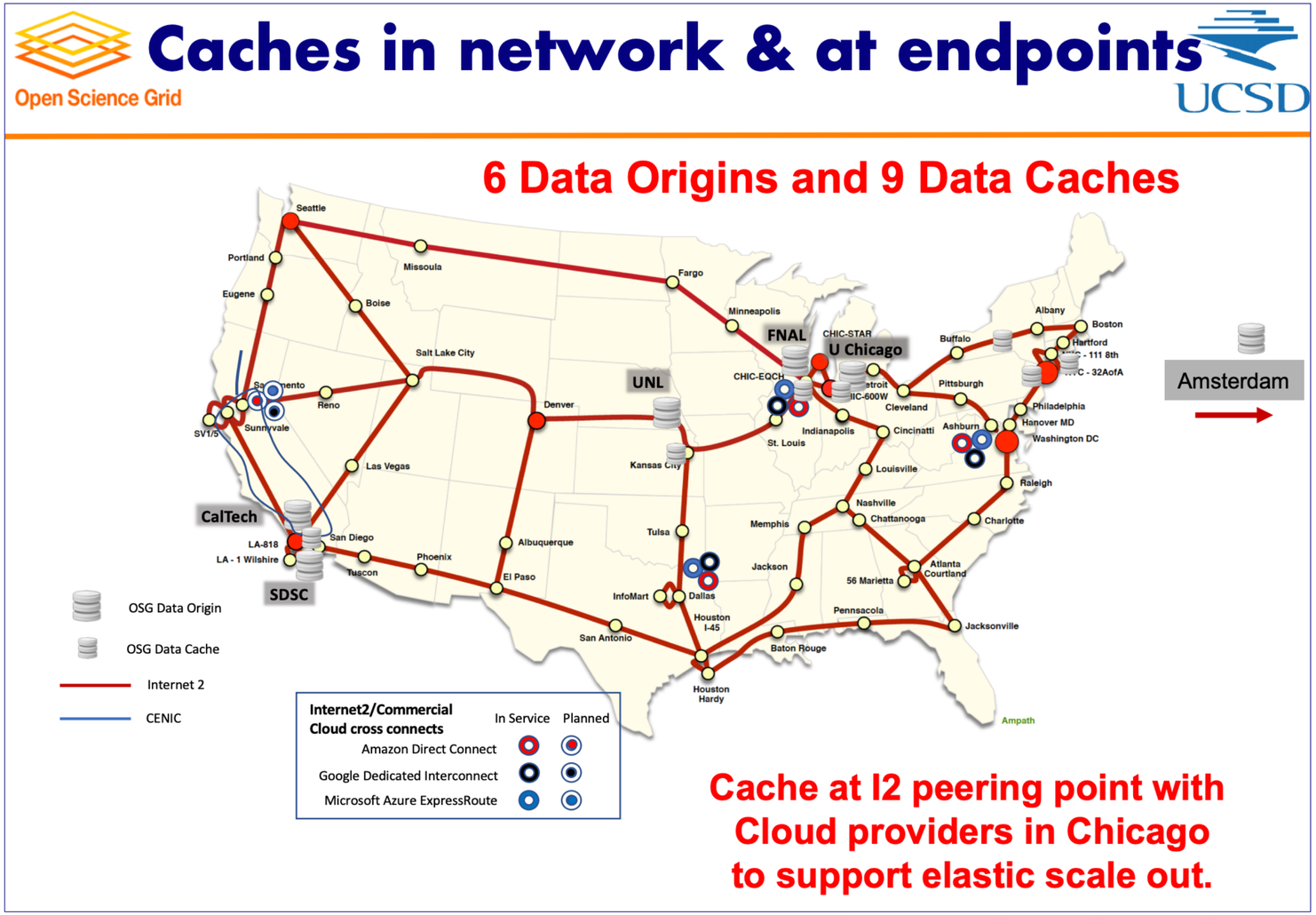
- *Origins* hold one's unique data and participate in a data federation to link their namespaces.
- *Caches* serve as a data access tier, delivering data to jobs.
- Origins are linked together by a data management tier, consisting of policy management (such as Rucio) and file transfer (such as FTS).



Refocusing our Data Management

This model allows us organizations wanting to process data to start simple – they only provide a single origin and we can help distribute data across the OSG-run network of caches.

The caches can potentially serve multiple data federations – our first foray into providing containers has been to produce an XCache container for ATLAS and OSG to utilize.

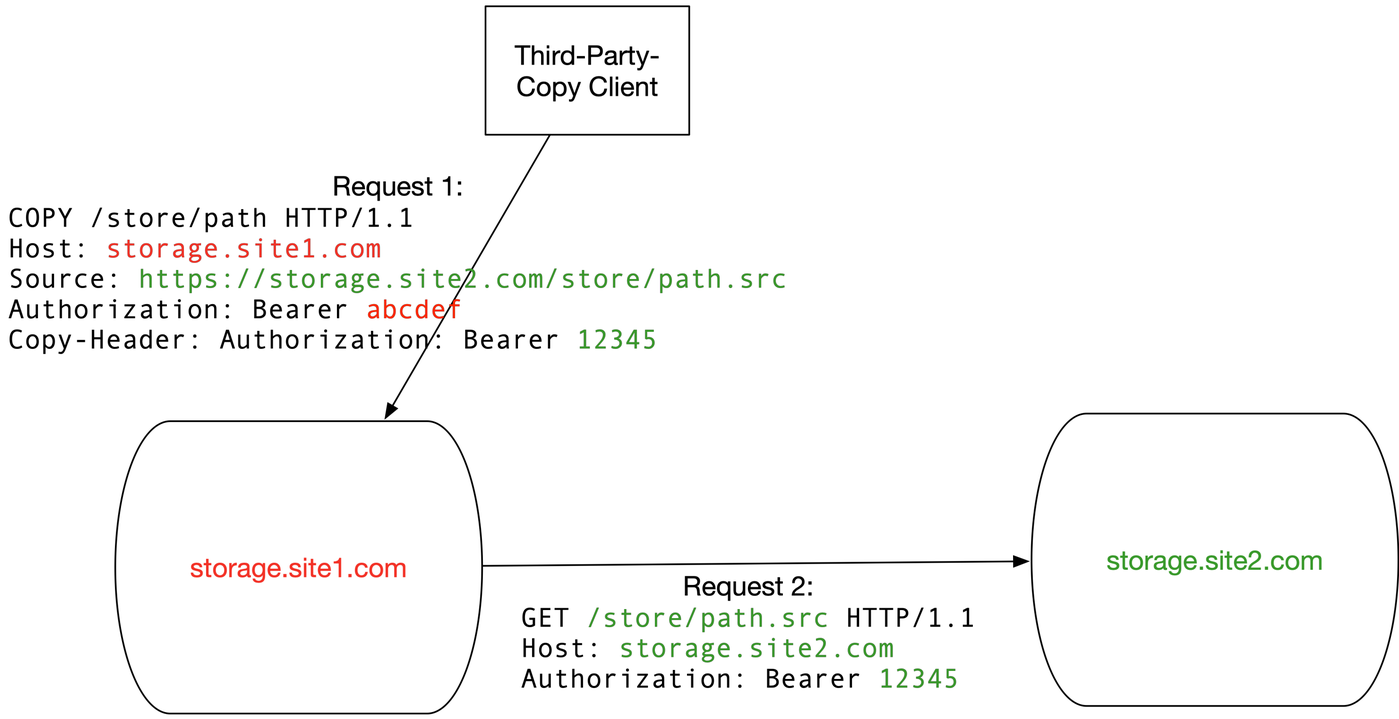


Refocusing our Data Management

The gradual shift away from the Globus Toolkit has reinvigorated work on alternate data movement protocols.

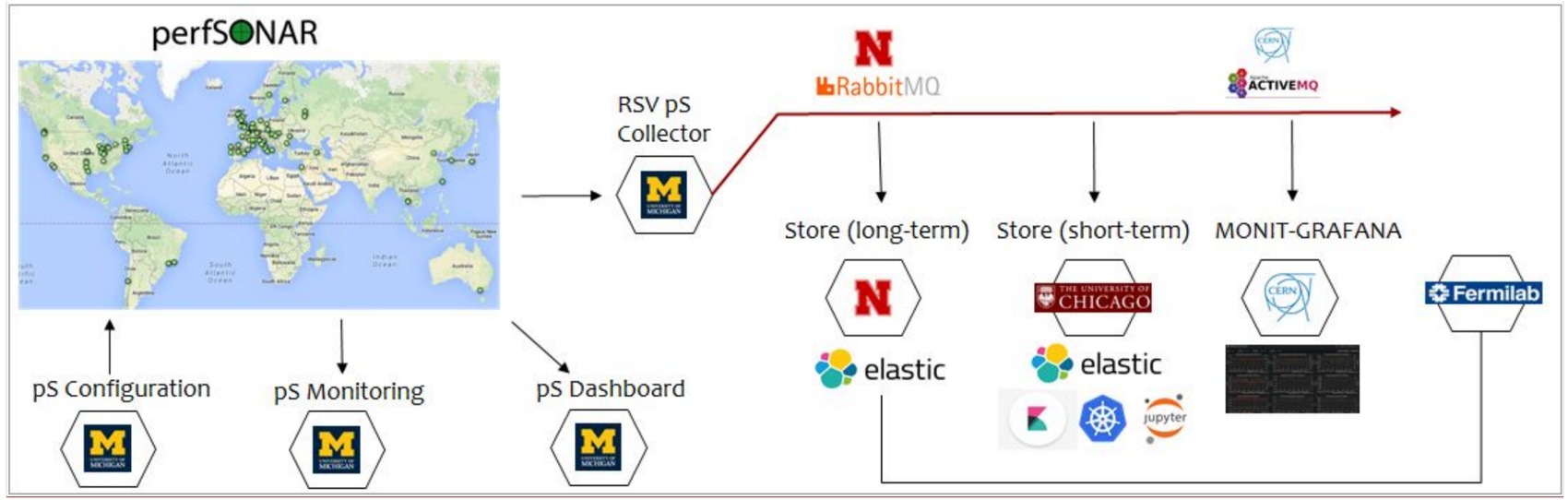
We have been participating in an active group (WLCG DOMA TPC) to grow the alternates (XRootD and HTTPS/WebDAV).

The technology pieces have been coming together early this year – expect rapid growth here!



Where next? Networking

- We continue to grow our capabilities in network monitoring and diagnostics with the SAND-CI project (NSF #1827116).
- SAND provides an improved infrastructure for aggregating our network metrics, pulling in over 1M datapoints an hour.
- The project is now working to turn toward analytics of our data store



Growing the HTCondor-CE community

The OSG introduced the HTCondor-CE about 6 years ago (2013 AHM); key observation is that HTCondor itself has all the functionality needed for a CE. For an organization that depends on HTCondor *anyway*, this allowed us to reduce our external dependencies.

We are incredibly pleased with how the approach has grown over the years.

There are no big technology announcements for today, but over the past year we have split the HTCondor-CE technology into a "base" and "OSG" variant.

This allows non-OSG communities to more easily adopt the technology and, potentially, allows the base HTCondor-CE to be owned by the HTCondor team.

HTCondor-CE

- Currently, Globus GRAM provides the abstraction, sandbox movement, and remote submission layers for the OSG-CE.
- In the April/May timeframe, we are targeting a new stack based on a HTCondor schedd.
  - Goals is to have HTCondor serve as a complete gatekeeper - only a special configuration, no additional OSG-maintained scripts.

Wednesday, March 13, 13

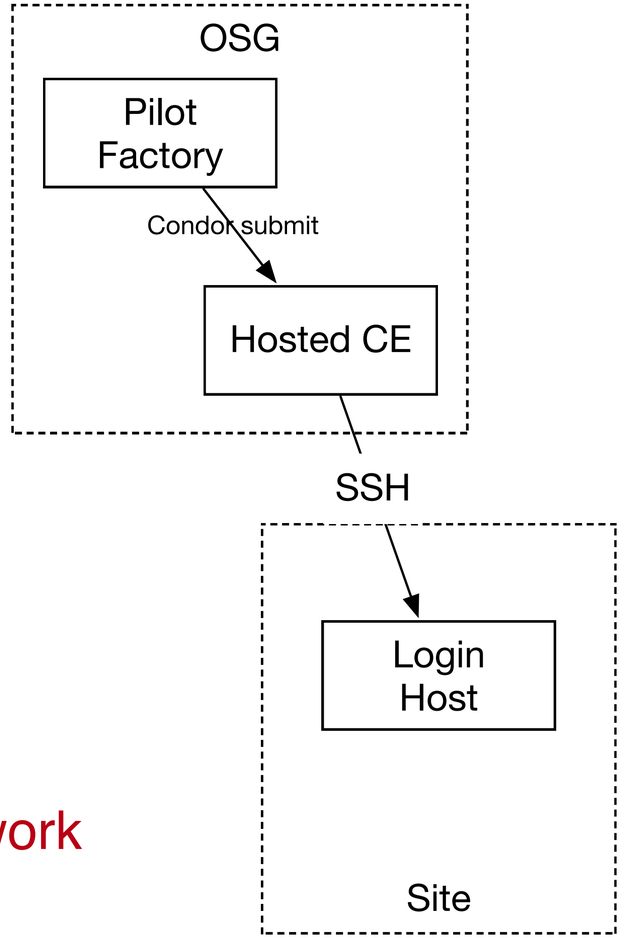
First mention - OSG AHM 2013

Hosted HTCondor-CE

The OSG is hosting nearly 20 CE instances on UChicago's OpenStack instance.

These CEs are part of our aspirational goal to not require sites to learn new, OSG-specific technologies.

However, they're still an operational burden for OSG. One of the major software projects in the next year is to distribute a "hosted CE" Docker image, making these services more homogeneous.



The first example of potentially-many allocable services from the OSG. Significant potential in tying our work with the Research Support and Collaborations area.

Lowering our Impact (on the worker nodes)

- Originally, the OSG stack was quite invasive on the worker node. You needed to install RPMs, keep them updated, etc.
- Over the years, we've slowly decreased our footprint. Many sites – especially WLCG – can get away with just installing Singularity and CVMFS.
- We can do better! With RHEL7.6, unprivileged containers are no longer considered a "preview". Many VOs can work with a Singularity **not installed by the sysadmin**.
- The Linux kernel moves on: newer kernels allow unprivileged mounts of filesystems. **Pilots can mount their own CVMFS *without*** sysadmin involvement.

Get back to the vision of "**execute one binary, no other dependency needed**"!

No RPMs

No CRLs

No cronjobs

No setuid

No CVMFS

Where next? Identity Management

Over the years, OSG has tried to steadfastly avoid managing user credentials. You've never had an OSG username and password!

This finally came true last year when the OSG CA shutdown. All credentials or identities you use on OSG are federated either through the IGTF or InCommon.

There's still a need, however, for identity and group management – e.g., who is allowed to submit downtimes for my site?

Over the next year, we plan to investigate keeping OSG identities and group information in COmanage.

Still: NO CREDENTIALS. This will support the InCommon federation; we are continue to move away from X509 client credentials.

CILogon

COmanage™

More at https://cilogon.org

Where next? Authorization

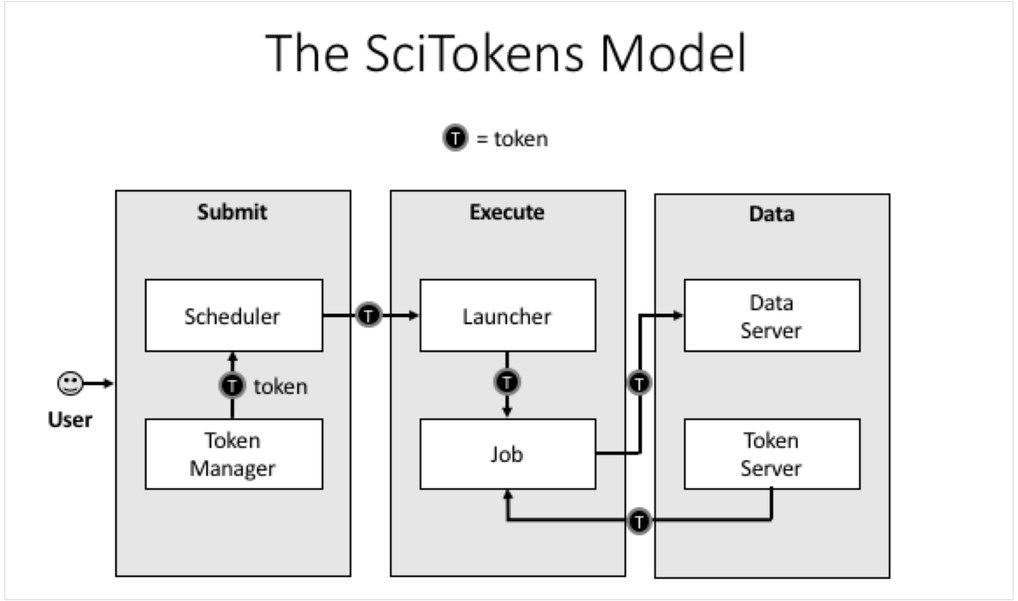
SCITOKENS

For too long, the OSG has tied authorization to identity mapping. You establish your global identity (X509 credential), it is confirmed by the remote service (CE), and you get mapped to a local identity (Unix account).

We want to change the infrastructure to focus on *capabilities*!

- The tokens passed to the remote service describe what authorizations the bearer has.
- For traceability purposes, there may be an identifier that allows tracing of the token bearer back to an identity.
- Identifier != identity. It may be privacy-preserving, requiring the issuer (VO) to provide help in mapping.

Example: "The bearer of this piece of paper is entitled to write into `/castor/cern.ch/cms`".



SciTokens is supported by NSF #1738962;
More at https://scitokens.org

Where next? Authorization

The decoded token contains multiple scopes – basically filesystem authorizations.

- The **aud**ience narrows who the token is intended for.
- The **iss**uer identifies who created the token; value used to locate the public keys needed to validate signature.
- The **sub**ject is an opaque identifier for the resource owner. In this case, it also happens to be the identity.
- The **exp**iration is a Unix timestamp when the token expires. A typical lifetime is 10 minutes.

eyJ0eXAiOiJKV1QiLCJhbGciOiJFUzI1NiIsImtpZCI6ImtleS1lczI1NiJ9.eyJpc3MiOiJodHRwczovL2RlbW8uc2NpdG9rZW5zLm9yZyIsImV4cCI6MTU1MzA5NDk2MywiaWF0IjoxNTUzMDk0MzYzLCJuYmYiOjE1NTMwOTQzNjMsImp0aSI6ImM5YzA2YWFmLWJlNTEtNDkwNi05M2UyLTE3ZjhlNjNmYmU0NSJ9.ciEE81Jqw8V_O6yO_0O-g0rQ22U1w3cTGomoVsN8KVuGHEeGwEMwOHYDNrjiCDsuoUHmm5VeTZk_GBHB0yr-Hw

HEADER: ALGORITHM & TOKEN TYPE

```
{
  "typ": "JWT",
  "alg": "RS256"
}
```

PAYLOAD: DATA

```
{
  "scope": "read:/protected",
  "aud": "https://demo.scitokens.org",
  "iss": "https://demo.scitokens.org",
  "exp": 1507686830,
  "iat": 1507686230,
  "nbf": 1507686230,
  "sub": "bbockelm@cern.ch",
  "jti": "abcdef12345"
}
```

Put it all Together

Today, we have all the technology pieces needed to replace the use of X509 certificates:

- For user access to data.
- For pilots to authenticate with the VO.
- For factories to authenticate with the CE.

The challenge this year will be to put all these pieces together!

Parting Thoughts

- Like every year, we have the usual churn of technologies; this happened last year and will happen next year!
- Particularly exciting is the progress to evolve Globus Toolkit technologies such as GridFTP and X509.
- *Don't be fooled by the implementation details*: a bigger sea change is the movement toward applications and services.
- Hope this allows our technologies to better service existing – and reach new – communities for years to come.

morgridge.org