

HTCondor and the OSG Token Transition



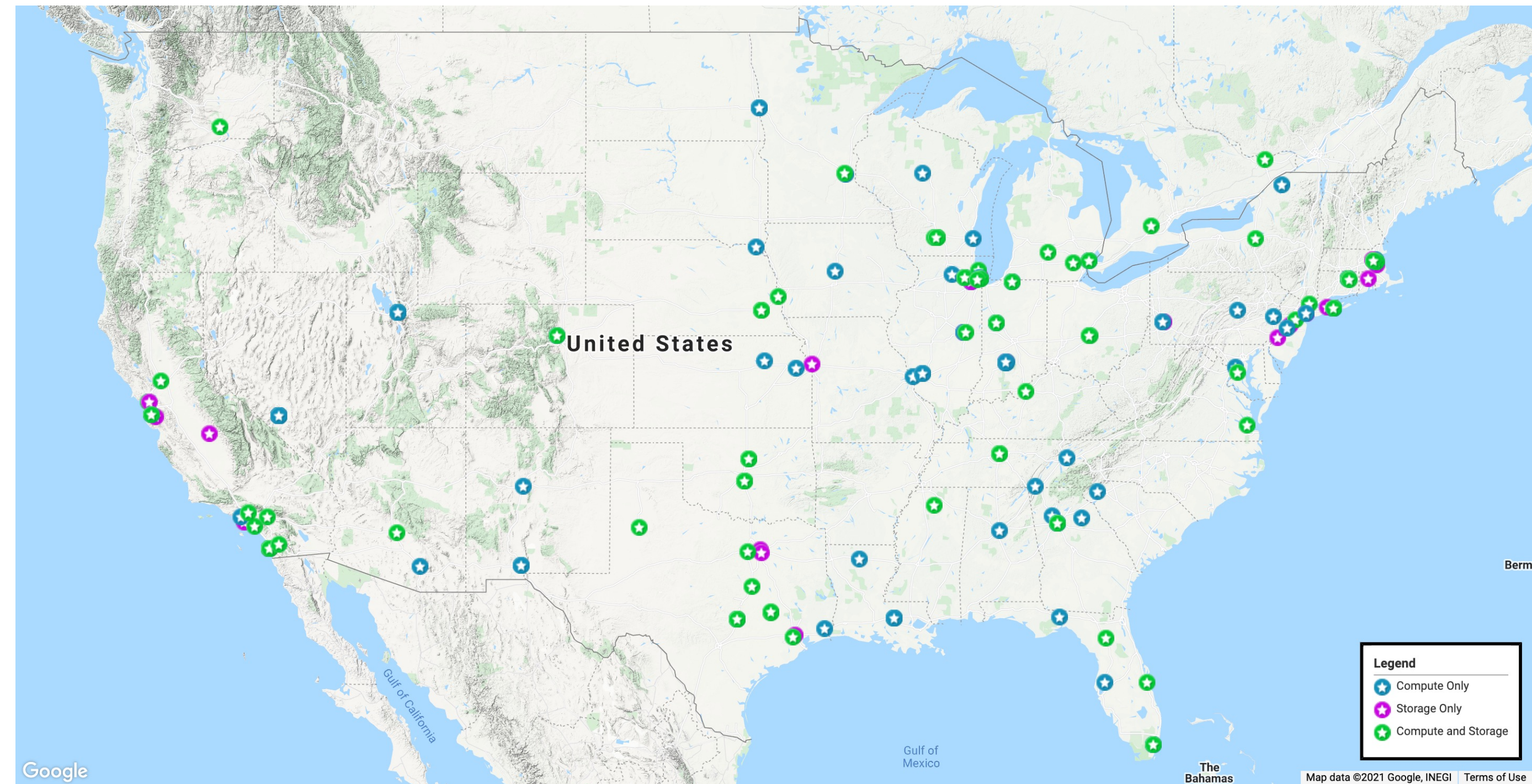
FEARLESS SCIENCE

DHTC: Distributed High Throughput Computing

High Throughput Computing:

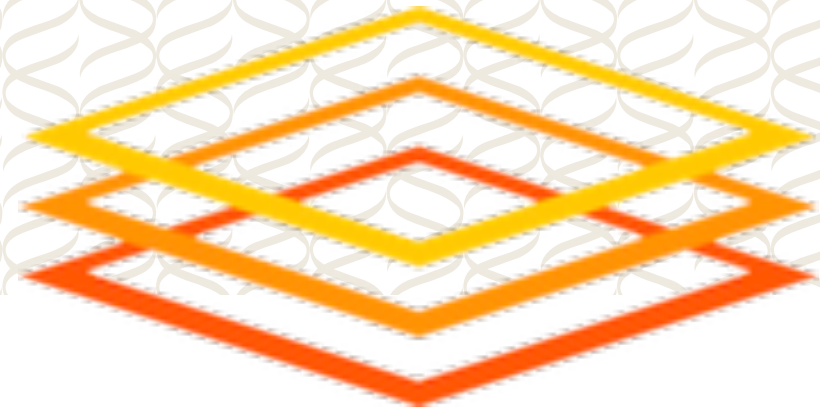
Maximizing the throughput of a computing resource toward a common problem.

- Within the OSG, we specialize distributed HTC - this involves many independent, collaborating administrative domains.
- There's a significant amount of resource sharing in the name of open science.
- This implies trust and authorization is of paramount importance!



OSG is a consortium dedicated to the advancement of all of open science via the practice of distributed High Throughput Computing (dHTC), and the advancement of its state of the art.

OSG Consortium



The OSG Consortium provides a fabric of services, including a software stack, that organizations can use to build dHTC environments.

- The OSG owns no clusters and pays no staff.
- It coordinates the efforts contributed by projects such as the NSF-funded IRIS-HEP and PATh.
- OSG also runs the “Open Science Pool” (OSPool), an environment for any scientist or group doing open science in the US.



In the last 24 Hours	
257,000	Jobs
5,801,000	CPU Hours
In the last 30 Days	
9,656,000	Jobs
183,849,000	CPU Hours
In the last 12 Months	
119,462,000	Jobs
2,069,366,000	CPU Hours
OSG delivered across 139 sites	

By connecting the resources of over 100 US sites and many others world, OSG is a national-scale resource for the NSF community.

OSG is a consortium funded by PATh, IRIS-HEP, other awards, and in-kind effort contributions.

Rethinking Identity and Capabilities for Authorization

How it started

OSG grew up in an **identity-centric** world. All our original services were developed around identity and our security activities were all around identities:

- How were they established?
- How are they revoked?
- How can they be delegated & impersonated?
- How can they be used for authorization?

The infrastructure was built around the idea of a users establishing a single, global identity and using that for accessing services they used.

- This was important as users would utilize a wide array of services – perhaps taking 1M jobs and submitting 1,000 jobs to each of 100 different worldwide sites.

How it worked out: users utilize almost no grid services!

The last 20 years have been marked by a continuous trend: users don't utilize grid services.

- Today, zero users submit jobs directly to sites on the OSG
- Rather, users utilize resources and the services provided by their local organization (such as the OSPool's "OSG Connect", CMS's CRAB, or ATLAS's PanDA).
 - In turn, those organizations use site services to organize and manage their allocated resources.

This should be entirely unsurprising:

- If you wanted to use AWS for LIGO, you wouldn't do it by handing 1,000 scientists each \$100 gift cards and tell them to have fun.
- In XSEDE community accounts, the resource provider interacts with the community and the community interacts with the user.

**We need an authorization model that matches
our usage mode.**

What's wrong with identity?

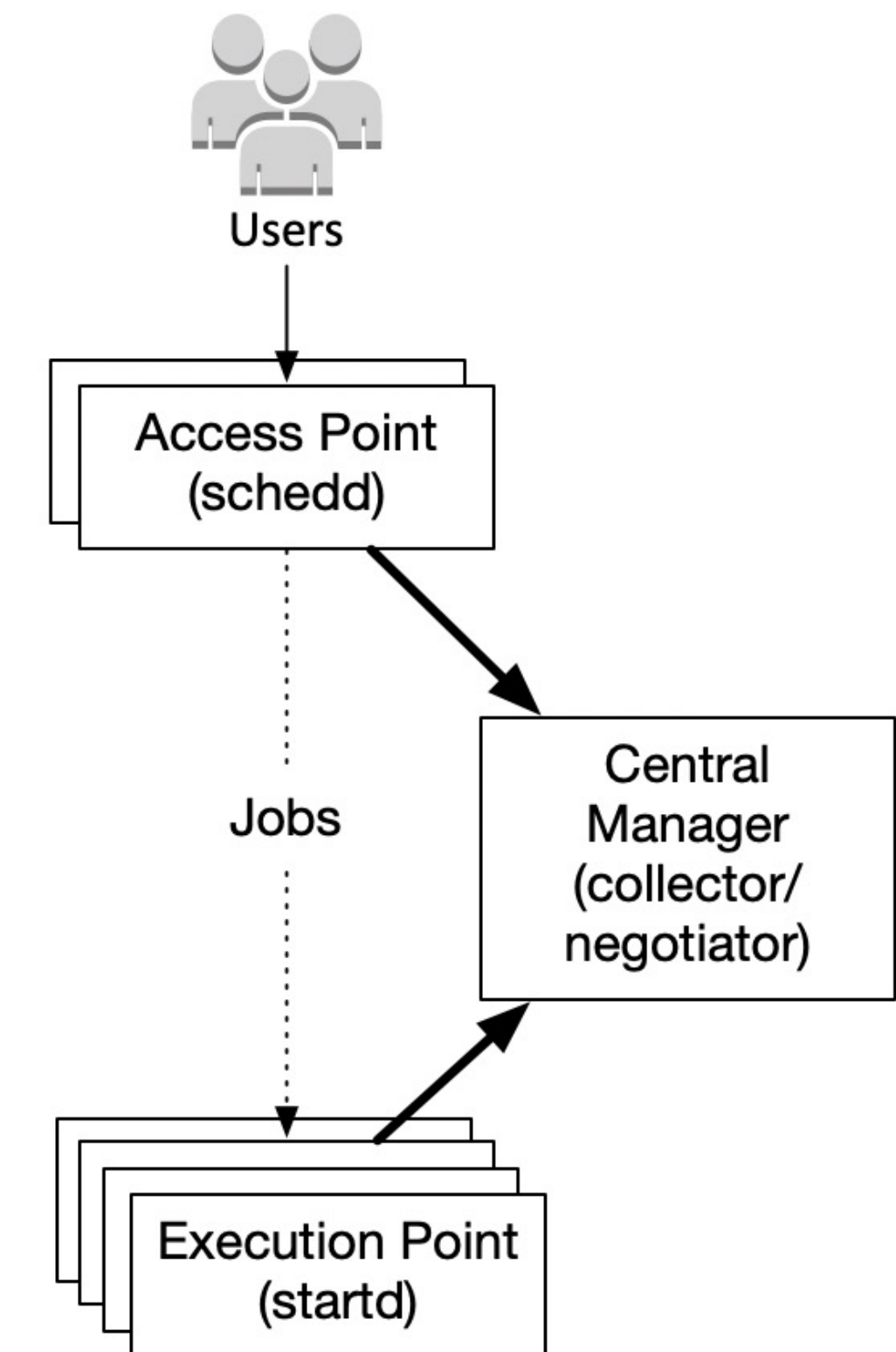
What's wrong with identity? Well, nothing as a concept. However,

- It's expensive to establish, especially in places where it's not needed.
 - If I'm a site, I may agree to give resources to the OSPool. I just care if the resource request is authorized on behalf of the OSG; I don't make the decision based on who it's from.
- Often what we need is simply whether two requests come from the same individual, not who those individuals are.
 - Traceability is important but orthogonal to establishing an identity.
- There **are** places where it's needed:
 - Requirement for individuals SSH'ing into a host.
 - Giving out home directory space to an individual.

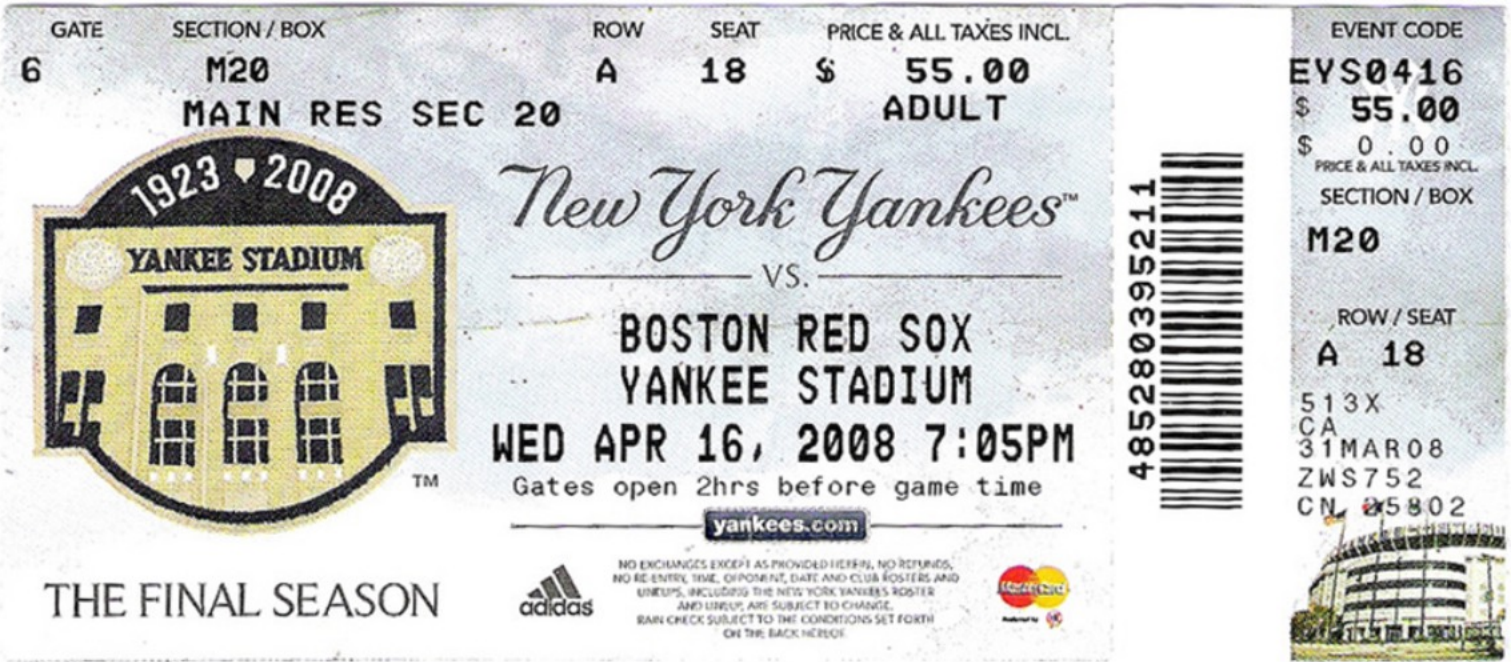
HTCondor and Identities

HTCondor does not rely on the concept of a global identity:

- Users establish an identity with the Access Point (AP) where they place their jobs.
 - For example, OSG staff do this with a video call for all new users.
- Daemons establish an identity with the Central Manager (CM) to advertise their existence and location.
 - The Execution Point (EP) provides the CM with a token and will trust any AP the CM gives that capability to.
 - The AP contacts the EP with the token; the AP does not establish its own identity with the EP.
 - Each job can be launched in its own container – EP does not know the identity of the user.



Moving from identity mapping to capabilities



Authorization on OSG was always based on identity mapping:

- A request was authenticated to a global identity.
- The global identity was mapped to a local identity.
- The request was authorized if the local identity was authorized to perform the action.

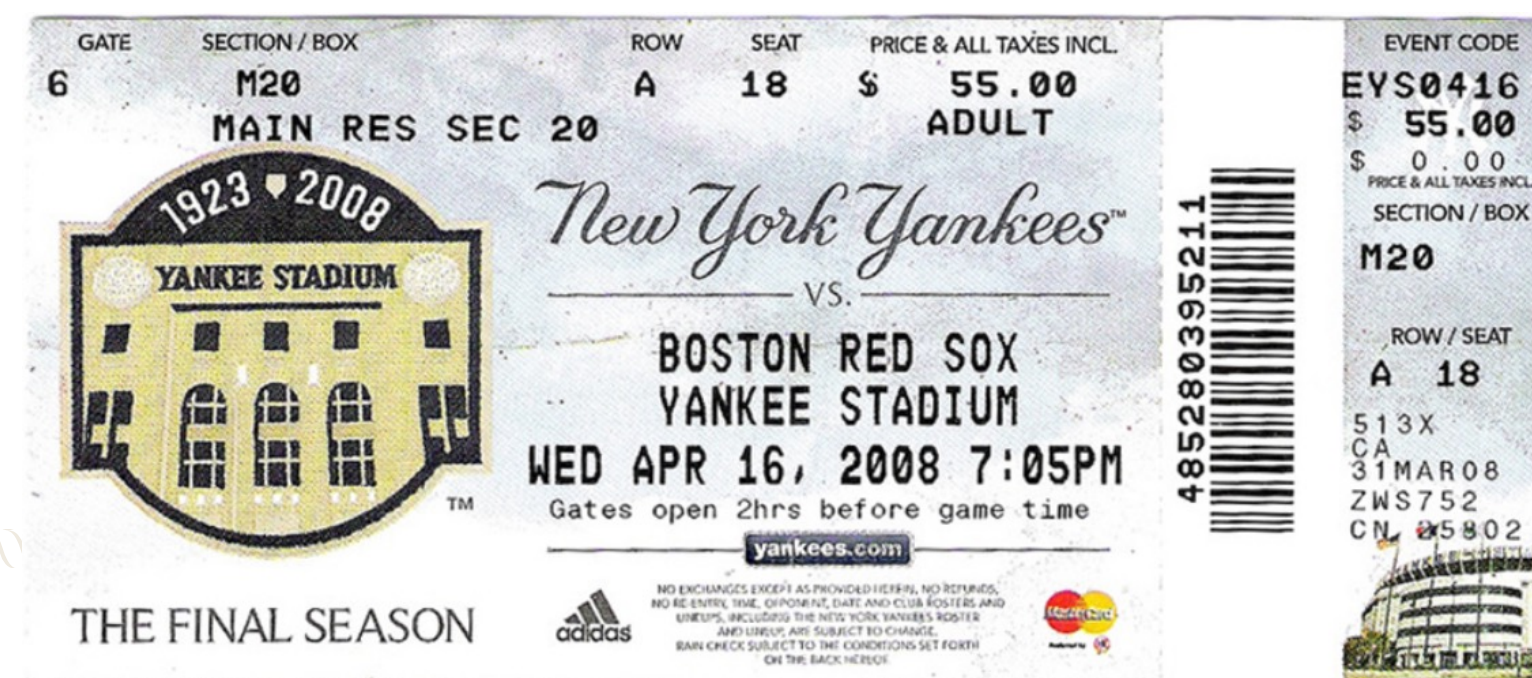
<u>Scheme</u>	<u>Credentials</u>	<u>Authentication</u>	<u>Authorization</u>
Gmail login	Password, 2FA	Username	Access to your inbox
Discovery Building access	ID card	Identity in HR database	Elevators
International Travel	Passport	Identity according to US Government	Enter Switzerland
Baseball Game	Ticket	NONE!	Sit in section 4, seat 34B
Webinar	WebEx URL	NONE!	Attend this wonderful seminar!

Moving from identity mapping toward capabilities

We are moving toward capabilities: credential describes what you can do, not who you are!

- Request carries a token issued by an organization.
- The token can be verified to have come from the organization.
- The token indicates whether the request is authorized by the organization.
- The request is authorized if the organization has the required authorization and the request is authorized by the organization.

Note the organization has identity but the requester does not.



Rethinking authorization in OSG

The fundamental concepts I describe have been true for quite some time.

Why are we suddenly making progress on this now?

- We are tying the move to capabilities to externally-driven transitions in software and a change in credential technology (JWT).

In 2020, the NSF funded the Partnership to Advance Throughput Computing (PATh), a partnership between OSG and the Center for High Throughput Computing (CHTC) at UW-Madison which produces the HTCondor Software Suite.

- The two entities – as well as other partner projects like IRIS-HEP, SciTokens, and SciAuth – are working hard with the community to move this forward!

OSG Token Transition



OSG Token Transition – Put into High Gear in 2017

High-level milestones from 2019 plan.

- **May 2017:** Globus announces the [Jan 2018 end-of-support](#) for the Globus Toolkit.
- **July 2017:** NSF funds the SciTokens project, which perform do the R&D necessary for the token transition.
- **Nov 2017:** OSG helps form the [Grid Community Forum](#) which forks the Globus Toolkit.
 - This stabilizes the software platform while we could put together a technology plan
- **Dec 2019:** OSG announces the [transition plan](#) for the OSG Software Stack.

Date	Milestone or Deliverable	Completed
Aug 2019	Beginning of OSG 3.5 release series (last release series depending on GCT)	✓
Aug 2019	Including HTCondor 8.9.2 in the 'upcoming' repository (first HTCondor version with SciTokens support).	✓
Oct 2019	OSG no longer carries OSG-specific patches for the GCT. All patches are upstreamed or retired.	✓
Mar 2020	"GSI free" site demo. Show, at proof-of-concept / prototype level, all components without use of GCT.	✓
Sep 2020	All GCT-free components are in OSG-Upcoming.	✓
Feb 2021	OSG series 3.6, without GCT dependencies, is released.	✓
Feb 2022	End of support for OSG 3.5.	

We are in the end stage for the Globus transition and ~middle for tokens.

OSG Token Transition Workshop

We are about 5 months away from end-of-life of GCT for OSG – are we ready?




Last week we had the OSG Token Transition Workshop to take the community's pulse on the overall transition.

- Day 1 contained an overview from the OSG and plenary talks from other communities on where they are (IceCube, LIGO, ATLAS, CMS, FNAL, BNL).
- Day 2 covered data services (Rucio+FTS), open discussion about community concerns, token basics for sysadmins, and a short hacking session to configure sites' HTCondor-CEs to accept tokens.

Basic message: progress across the board, February 2022 looks realistic.

HTCondor Software Suite (HTCSS) Token Transition

HTCSS has been carefully coordinating with OSG throughout the transition and has recently announced [their own schedule](#):

-  **April 2021:** HTCondor 9.0.0 released; drops support for GRAM and CREAM.
-  **May 2021:** HTCondor 9.1.x adds support for ARC-CE's REST interface.
-  **July 2021:** HTCondor 9.1.x implements proxy delegation with only OpenSSL; emits warnings when GSI authentication is used.
- **October 2021:** HTCondor 9.3.0 released without GSI authentication.
- **March 2022:** new long-term stable (LTS) version of HTCondor, 10.0.0, without GSI.
- **September 2022:** Support for old LTS release (9.0.x) ends.

Tokens in HTCSS

There's three majors uses of tokens in HTCSS:

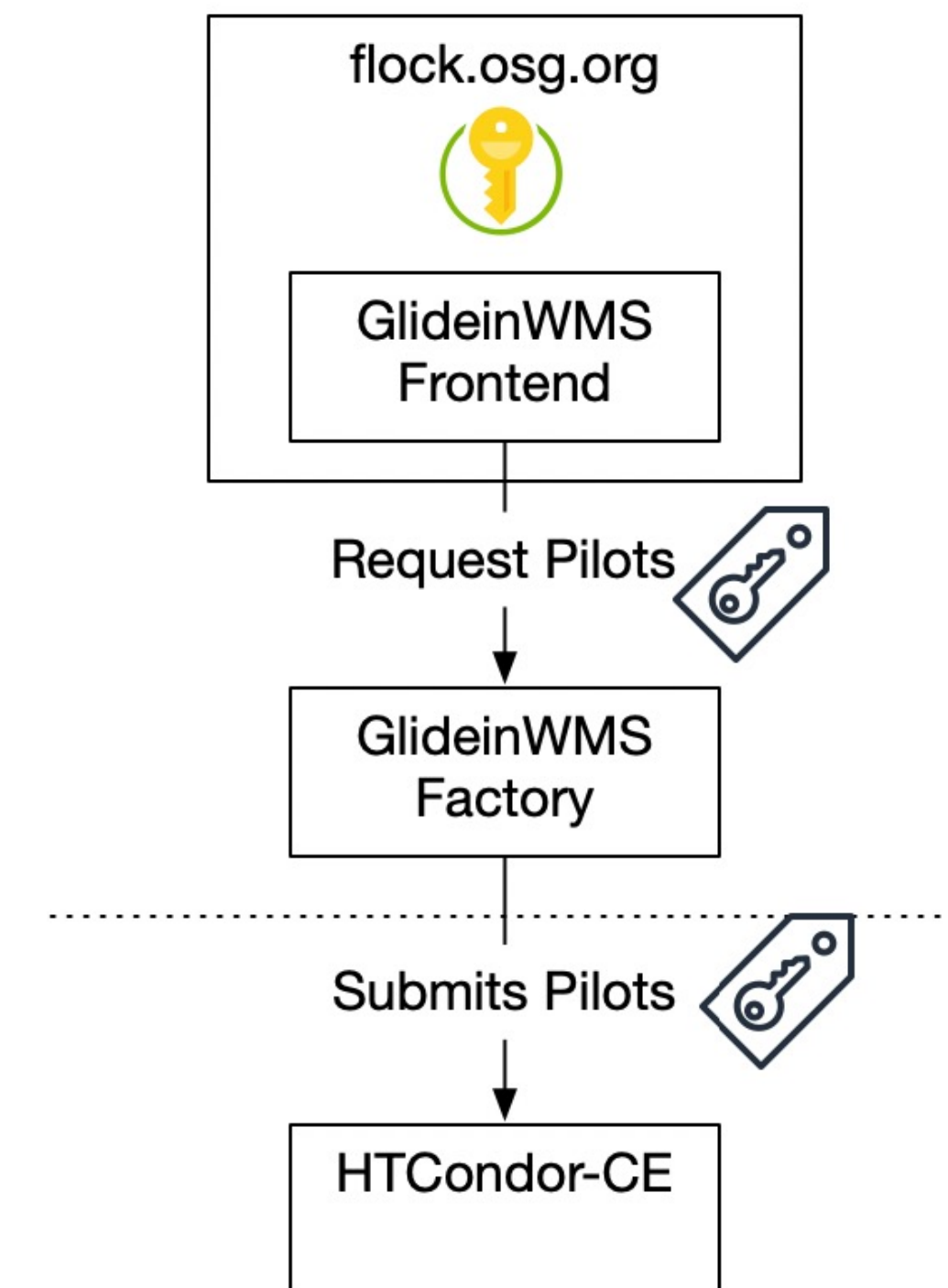
1. **SCITOKENS authorization:** Using a SciToken (or WLCG token) to establish client capabilities. Done over a TLS connection (server needs a host cert.) and JWT's signed with RSA or EC key; public key discovered using OAuth2 metadata discovery.
2. **IDTOKENS authentication:** Uses a JWT as a shared secret to establish client capabilities & any limits. Uses symmetric encryption (SHA256 HMAC-based signatures); server needs to have the signing key (no certificate).
3. **Job credential management:** The credd & credmon daemons will acquire and manage tokens on behalf of a job.
 - Ensures running jobs get renewed tokens as needed.
 - Plugin-based architecture provides multiple ways to acquire tokens including:
 - Creating them directly based on a locally-available private key.
 - Performing an OAuth2 code flow (with web server) to get AT + refresh token.

The OSPool uses all three of these!

Use Case 1: Submitting pilots to HTCondor-CE

The OSPool GlideinWMS frontend generates a **SciToken per CE**.

- The frontend measures the job pressure in the OSPool and requests resources from clusters across the OSG.
- If pilots are needed, then the frontend sends a pilot request and encrypted token to the factory.
- The factory submits individual pilots to the CEs using the SciToken via HTCondor-G.
 - The factory->HTCondor-CE connection uses the CEDAR protocol with the SCITOKENS authentication method.



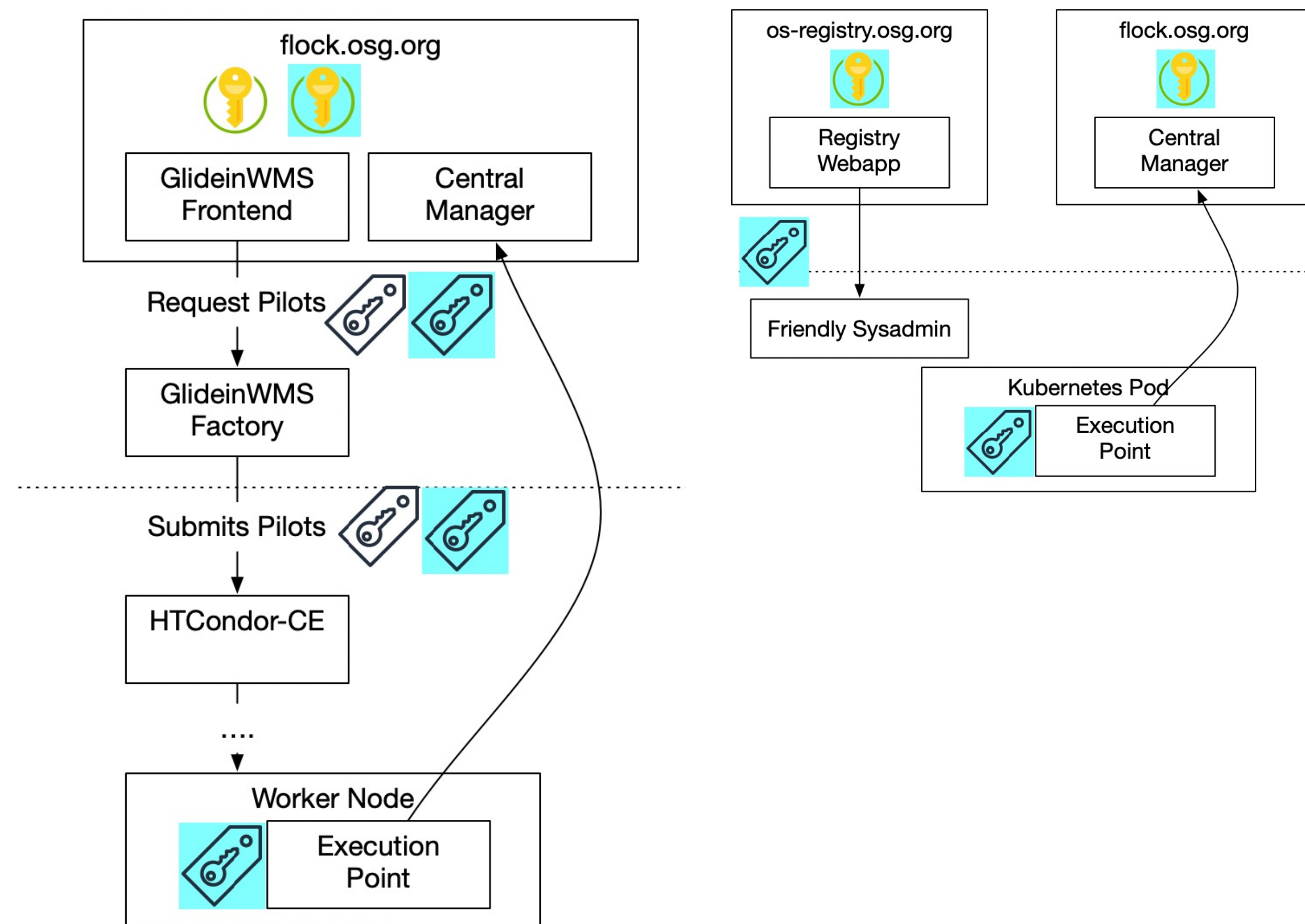
Use Case 2: Connecting the “execution point” to the central manager.

IDTOKENS are used internal to the OSPool. GlideinWMS can generate a token and send it along with the pilot.

- The IDTOKEN, not the SciToken, travels all the way to the worker node.
- Note the IDTOKEN must remain valid for **however long the pilot is in queue**.

Alternately, trusted administrators can request a token for their site and approve it through the webapp using their CILogon identities.

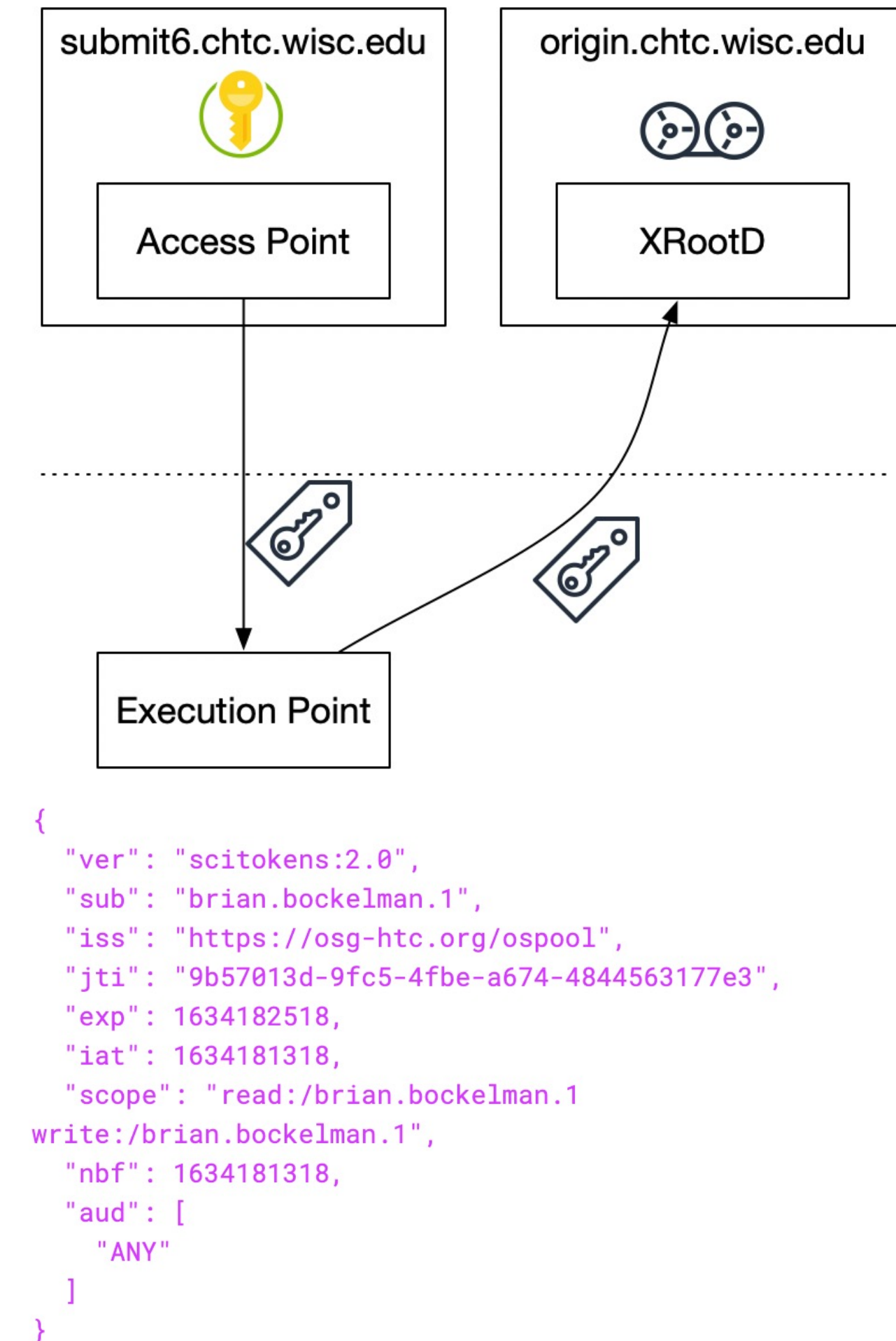
- These tokens can be inserted as a secret into a Kubernetes pod.
- The site admin then makes the decision on how many “backfill” containers to launch.
- [See the OSPool Containers documentation.](#)



Use Case 3: Job access to storage

The OSG Connect APs utilize the credd & credmon to enable access to OSG Connect storage.

- By default, user jobs are marked as needing storage tokens (otherwise, would need lines in the condor submit file).
- The AP runs the condor_credmon which automatically generates a token before a job is started.
 - A renewed token is periodically requested by the EP; ensures a valid token is always available.
- The token is usable at the (XRootD-based) origin server associated with the AP for writing...
 - Or any of the XRootD-based cache servers for reading.



Adoption of tokens on the OSPool

Here's the overall status of token use on the OSPool:

1. **SciToken-based pilot submission:** 52 CEs accept our SciTokens for job submission; 95% of the 38 operated by OSG itself.
2. **IDTOKEN-based EP authentication:** About 95% of EP's come via IDTOKENS (rest are believed to be misconfigurations).
3. **Storage access via tokens:** Enabled for all users for over 3 years.

CHTC shares many of the technologies as the OSPool and has taken another step: upgraded to a HTCondor 9.3.0 RC. None of the daemons use GSI!

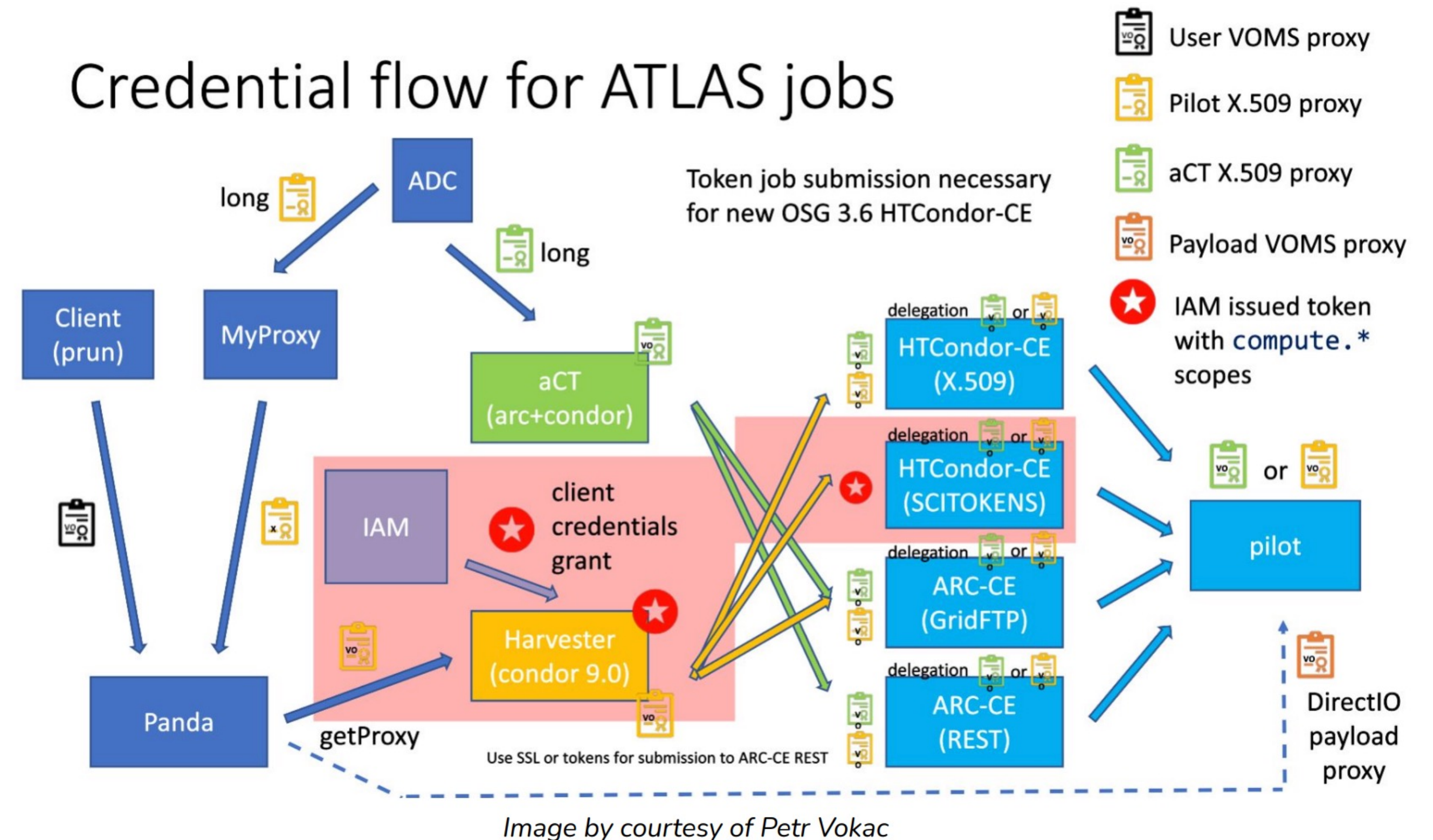
- This week CHTC will start an upgrade drive for the sites where SciTokens isn't working.
- OSPool is about a month behind CHTC.

Adoption of tokens – ATLAS

ATLAS:

- Have jobs submitted by their pilot factory (Harvester) to BNL & Chicago.
- On schedule to have OSG sites move by February 2022.
- ARC-CE sites worldwide will need to switch to the REST interface by June 2022.
- User jobs will use X.509 for some time...

Credential flow for ATLAS jobs



Adoption of tokens - CMS

CMS:

- Has successfully tested tokens to OSG sites. Not in production.
- Using some CERN-local services (Argus) to map user proxies internally; need to figure out how to replace this.
- Switched to HTTP-TPC (necessary to retire Globus GridFTP) at CERN, all T1s, and 38/47 T2s.
 - N.b.: ATLAS is in a similar state.



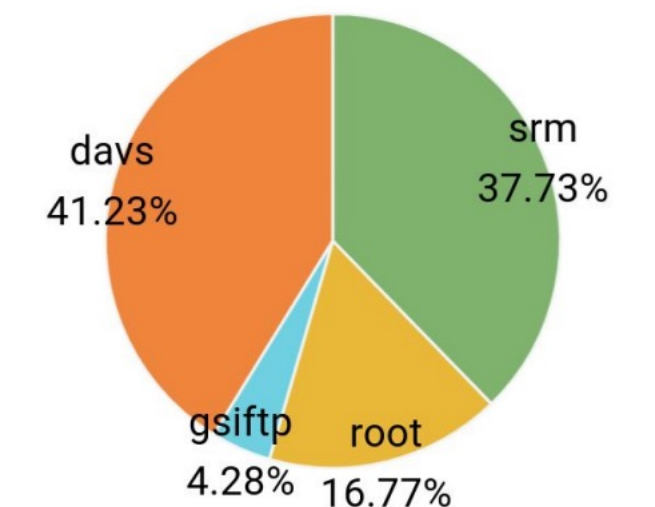
GridFTP to WebDav-TPC transition status

- **Long process** that requires all sites to enable WebDAV endpoint
 - Conversation initiated by CMS opening a ticket to site
 - **Manual tests** + manual **verification** of storage.json
 - **LoadTests** on a _Test instance has to **succeed** before **production** instance is **configured**

Results:

- **WebDAV** has been deployed at **all T1** Disk endpoints
- **38 out of 47 T2s** has completed the transition
- **Few T3s** out of ~27 has started working on it
- **SAM** tests **waiting** for **fixes** on gfal python interface

Production Data transferred by protocol

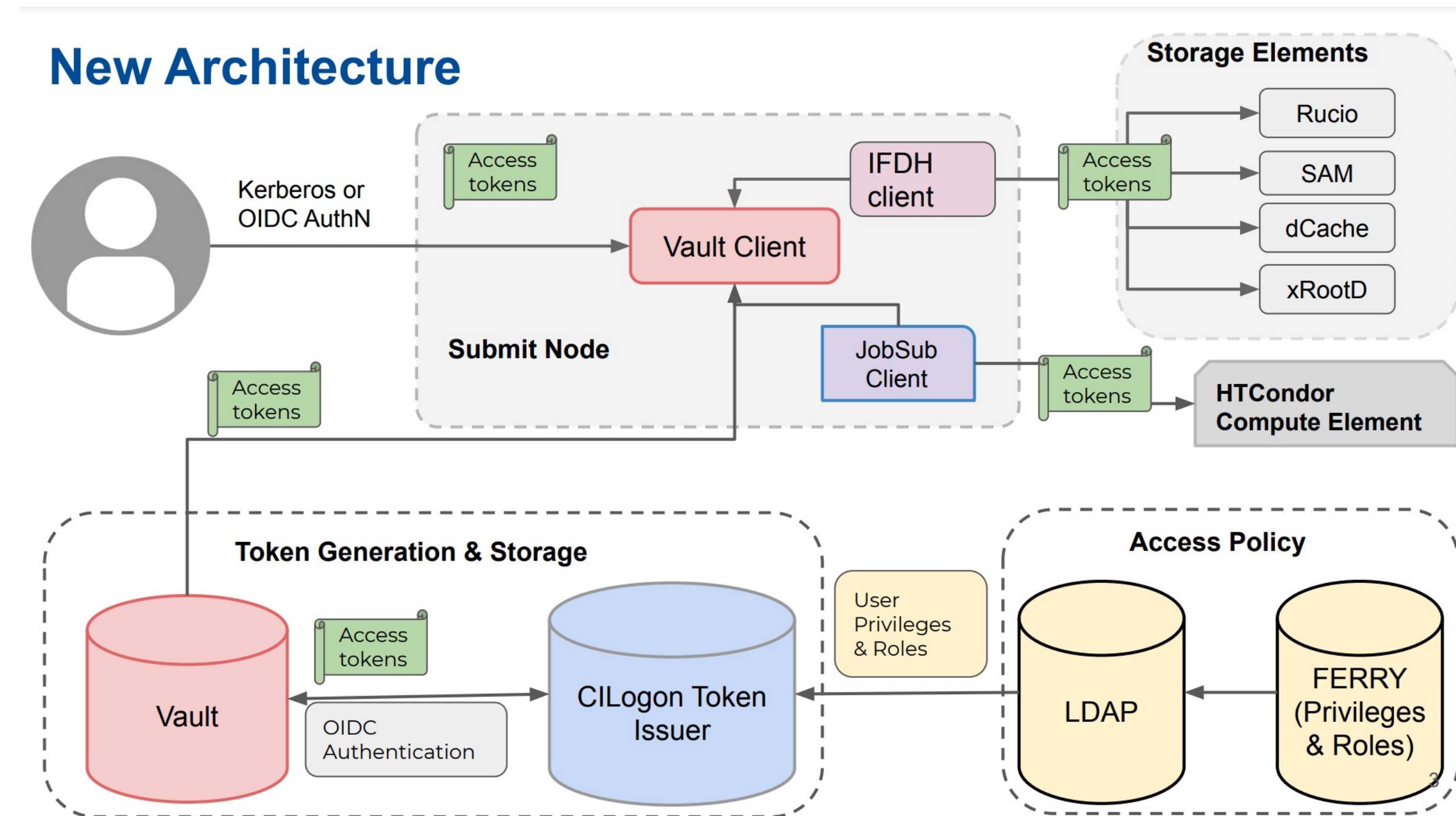


12

Adoption of tokens – FNAL & BNL

[FNAL](#) and [BNL](#) had site reports on their use of tokens and federated identity:

- FNAL is transitioning many of its experiments to an architecture heavily leveraging Hashicorp's Vault for token acquisition and management.
 - Hope to start moving over users of local experiments next spring.
- BNL's outlined progress across 6 different experiments – noting many timelines and plans are driven by the host lab (e.g., KEK for Belle II).
- Both highlighted the need to coordinate closely with DOE cybersecurity.



Closing Thoughts



First steps toward a token-based future!

Even though we've been at this for many years – and accelerated in the last 4 – it feels like we're still taking the first 'production' steps this year.

- The OSG's end-of-support for the Grid Community Toolkit focused the community on a retirement date of February 2022. Looks like we'll even make it!
- However, this is not the end-goal: many workflows still rely on GSI for user identities. WLCG has outlined a plan lasting through 2025 (which means the entire transition would take ~8 years).
- These efforts should look at the entire approach to authorization – especially the use of capabilities – and not just 'translate' the existing system to new technologies.
 - This makes efforts like [SciAuth](#) so important.

You can write FORTRAN in any language.

Corollary: You can implement X.509 in any technology

(our goal is to not!)



morgridge.org

This material is based upon work supported by the National Science Foundation under Grant No. 1836650, 2030508, and 2114989. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

FEARLESS SCIENCE