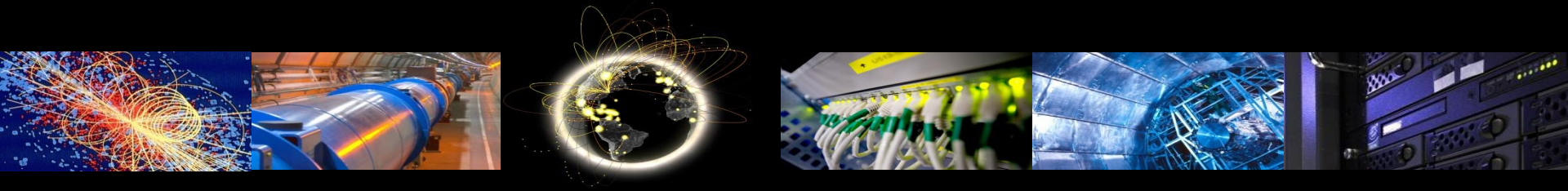
WLCG AuthZ Update

Grid Deployment Board

June 9th 2021



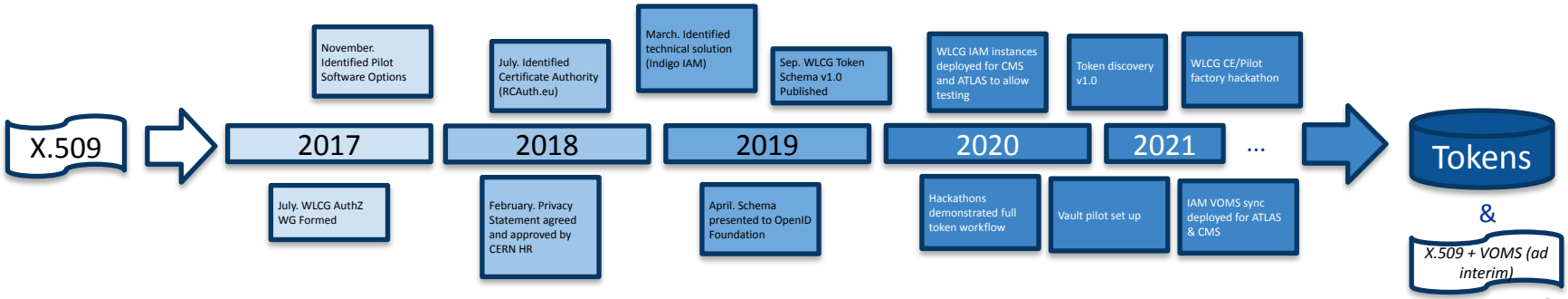
WLCG Authorization (AuthZ) WG

- Membership includes current major users of tokens in High Energy Physics
  - INDIGO IAM
  - EGI Check-in
  - SciTokens
  - dCache
  - ALICE
- Development work of pilot projects supported by:

- Priority to stick to industry and R&E standards wherever possible
- Collaboration with DOMA Working Group - essential for accelerating token support in workflows

Towards Tokens for WLCG

- Documentation
  - ✓ Token schema v1.0
  - ✓ Token discoverability specification v1.0
- Hackathons
  - ✓ Focus on Data management (Jan & Sept. 2020) and **CE/Pilot factories (May 2021)**
- Infrastructure
  - Token issuers per VO (Indigo IAM) - **ATLAS and CMS ready for testing**
    - Operational support and maintenance (CERN IT)
  - Command line solution (Hashicorp Vault) - *Pilot set up*
  - Token support throughout WLCG stack - *Timeline TBC*

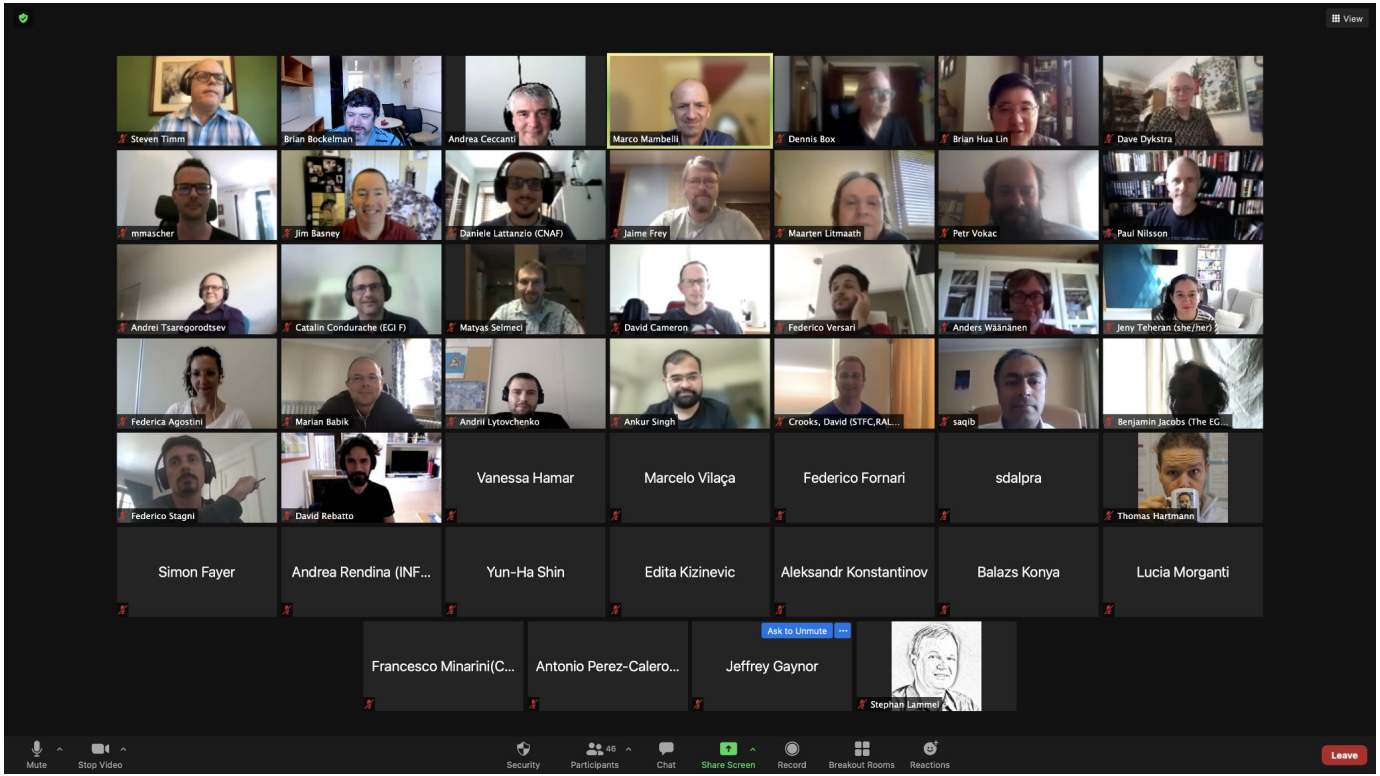


WLCG CE/Factory Hackathon

- Much of the activity around tokens has involved storage and DOMA software providers.
- Identified the need to better organize the CE/pilot factory communities.
- Idea: host a 2-day virtual workshop to start bringing together the independent threads of work.
  - And thus the CE/Factory hackathon was born!

https://indico.cern.ch/event/1032742

Hackathon



~70 registered, ~50 participants.

Goals of the Event

Three basic ideas:

1. Start the community-building process necessary for a major transition.
2. Ensure all participants have a basic understanding of the relevant technology (boot camp!).
3. Start identifying the biggest gaps between what we have and what we need.

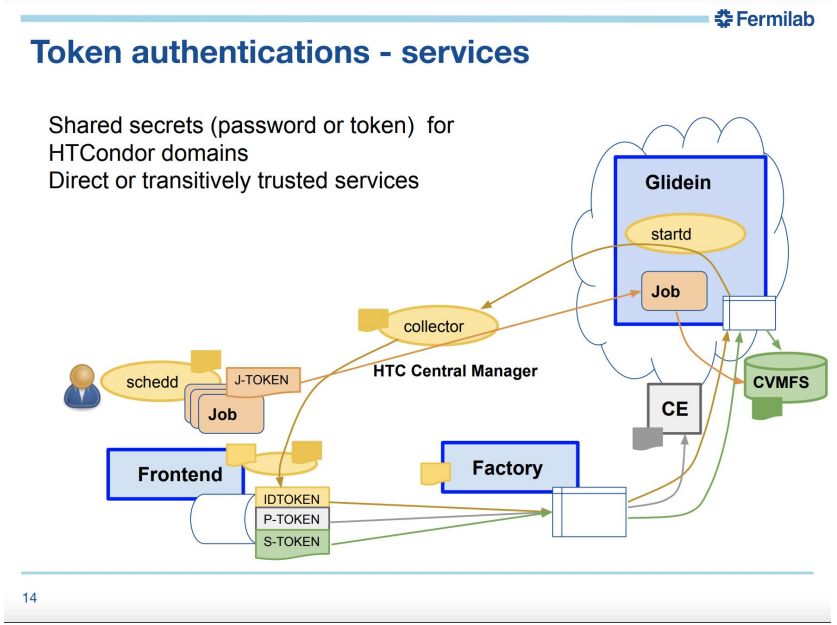
Day 1 - Presentations & Discussions

Morning presentations:

- A tour-de-force overview of token technologies, OAuth, IAM, and the WLCG profile by Andrea.
- An explanation of how GlideinWMS does token auth to submit pilot jobs and connect back to the resource pool by Marco.
- A hands-on demo from Brian Lin on using a condor submit file to submit jobs to a HTCondor-CE using a WLCG token.

Discussions:

- What services does a CE need to offer for token lifetime management? What is needed beyond initial encrypted transfer?
- What's the best way to automate renewal of tokens for the pilot systems?

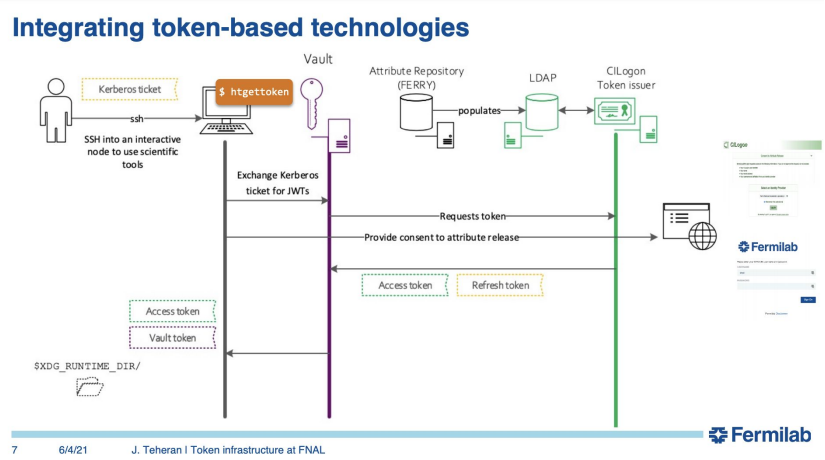


Slide from "GlideinWMS approach to token auth" presented by Marco Mambelli and Dennis Box.

Day 2 - Presentations

Presentations:

- ARC-CE token support status and plans by Aleksandr.
  - Short version: it works in the existing version but there are a number of “gotchas” and refinements in the pipeline.
- FNAL status of rolling out tokens for the site AAI by Jeny.
  - FNAL is an interesting case as they’re taken a holistic view to the transition - covering far more site-level services.



FNAL is leveraging Hashicorp’s Vault to manage and maintain its tokens. Tokens are sent to vault via typical OAuth (browser-based) flows but can be acquired by clients using configurable policies and different authentication methods (like Kerberos). Potentially ideal for automated renewal.

Day 2: Discussions

- Tokens are far more prescriptive about authorization than X.509.
  - However, there are still tough "mapping" issues from token to a local identity. What data goes into the map? How complex should it be?
  - HTCondor-CE's existing mapping mechanism probably needs additional functionality. *However* strong pushback from devs against making it arbitrarily complex. Identified need for callout for complex decisions.
- What's the best way to automate token acquisition and renewal?
  - Particularly, monitoring systems (ETF) currently have more complexity than other use cases as the VO's pilot credentials are delegated to a non-VO-member.
  - Vault appears to fill the role but there's a need for a production instance at CERN to really make progress. Follow-up with CERN IT.
- What services should a CE provide for tokens beyond just "copy like a file"?
  - Does the CE need to understand token expiration? Token renewal?
  - ARC-CE may need ability to stage in data for ARC control tower functionality.
  - Should factory push renewed tokens versus pilots pulling the tokens they need?
  - What's the tradeoff for a long-lived but low-power credential?

Hackathon - Conclusions

Note there wasn't much "hacking"!

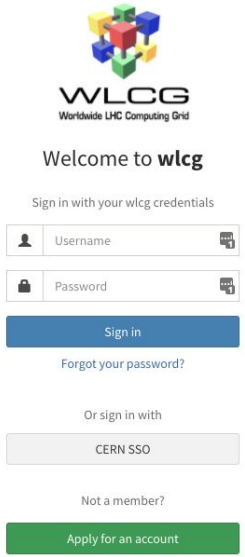
- Award goes to Petr who was able to confirm some ARC-CE functionality during the discussions!

At the end of day 2, we discussed "what's next?". Ideas:

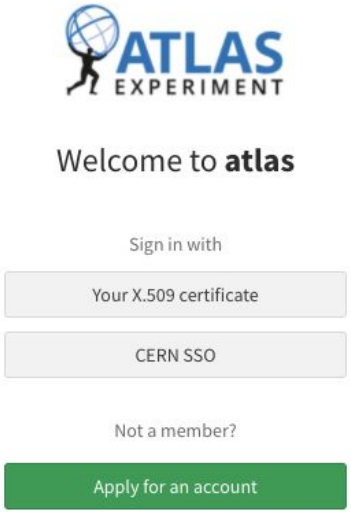
1. Start setting up a token interoperability testbed, analogous to what was done with HTTP-TPC.
2. Have a biweekly, 2-hour, 'open mic' for people to work on projects and have discussions about tokens in CEs.
3. As we gather significant discussion topics - such as missing functionality - the community will use the existing AuthZ working group to discuss with a broader audience.

IAM Deployments status

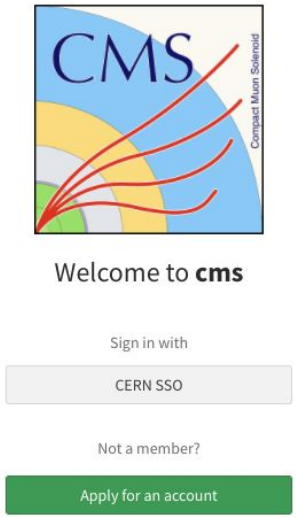
The following token issuers have been deployed. The ATLAS and CMS instances are available for testing and integration, with the expectation that they will become the future production token issuers.



https://wlcg.cloud.cnaf.infn.it

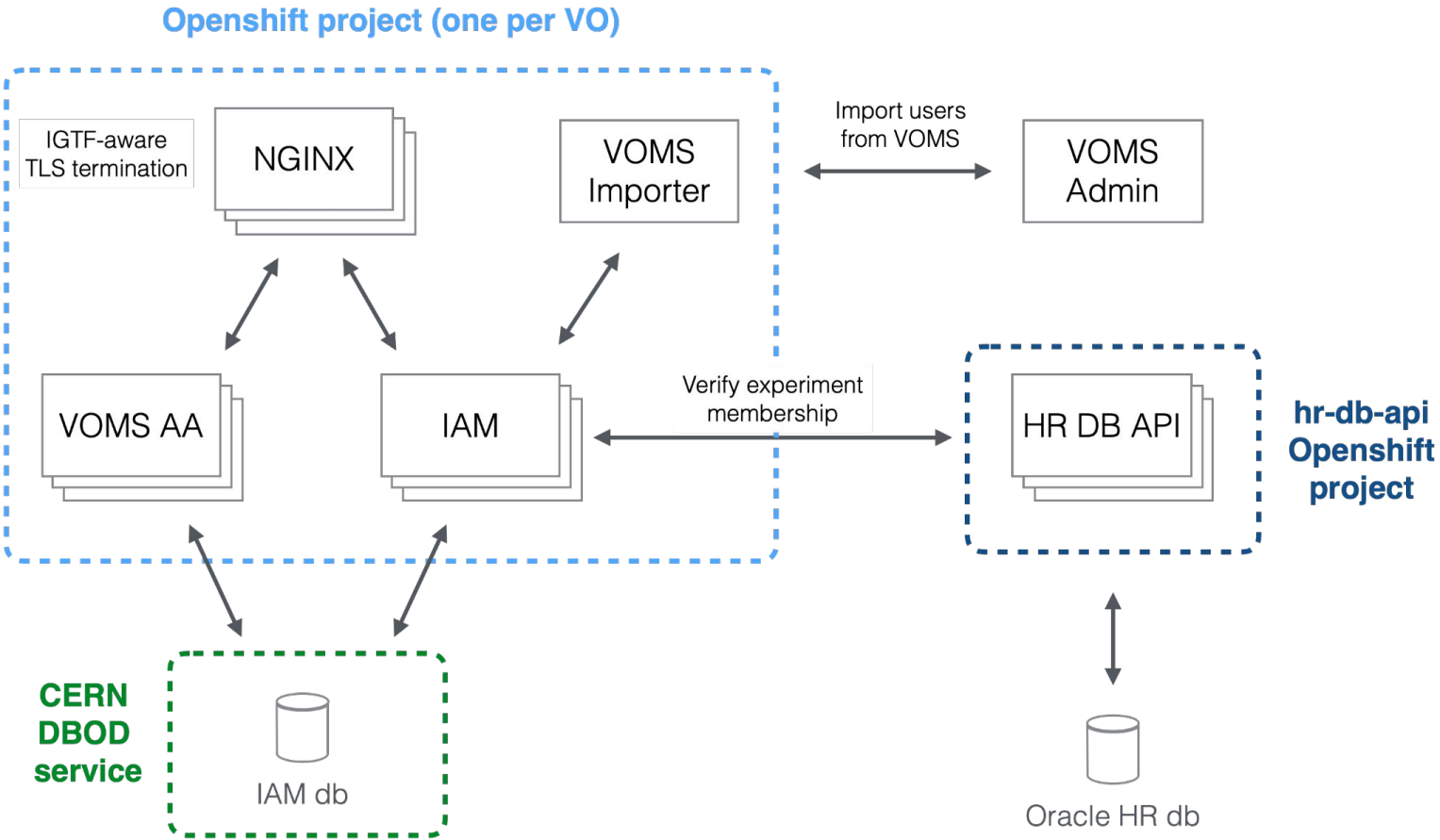


https://atlas-auth.web.cern.ch



https://cms-auth.web.cern.ch

IAM deployment overview



VOMS -> IAM

The transition from X.509 to tokens will take time so IAM **was designed to be backward-compatible with our existing infrastructure**

IAM provides a VOMS endpoint that can issue VOMS credentials understood by existing clients and libraries

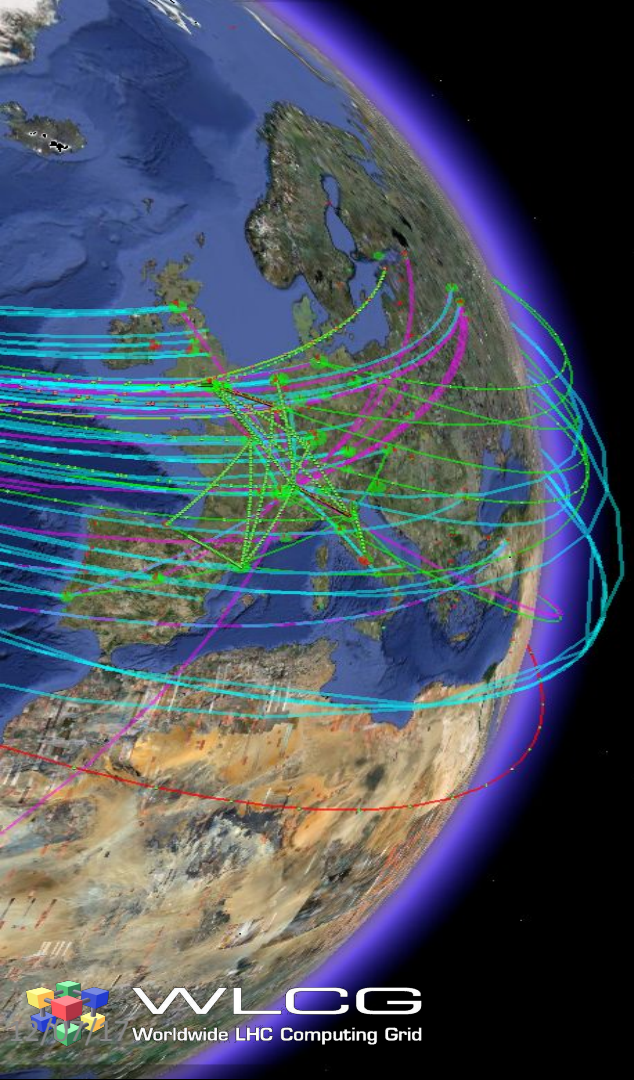
- VOMS clients >= 2.0.16
- VOMS configuration RPMs **already available** in WLCG and OSG repositories for CMS and ATLAS (thanks to Maarten and Brian)

A migration script has been developed and deployed to import users from VOMS to IAM

- users will NOT have to re-register en masse to IAM, and their IAM account will be automagically linked to their CERN account
- the script will keep IAM in sync with the VOMS instances until the VO registration process is migrated to IAM

Next steps

- Test equivalence between new and old VOMS servers
- Setup monitoring for IAM instances
  - metrics
  - centralized logging
  - alerts
- Load testing
- Look into a Vault service in CERN-IT



Questions?