AAI and InCommon Federation

OSG Council Meeting, July 2024

Brian Bockelman

MORGRIDGE
INSTITUTE FOR RESEARCH

Authorization and Authentication Infrastructure for OSG Services

- Most here are familiar with the traditional AAI used by OSG services:
  - Host certificates (with TLS) for server authentication.
  - X.509 user certificates + proxies + “GSI” for client authentication; VOMS extensions for group or role information.
    - The authentication information was mapped to a Unix account at the service-side; the client was authorized to do anything the Unix account could.
    - GSI is a set of extensions + rules on top of traditional web PKI. Some pieces (set of accepted CAs, DN namespaces) are managed by the IGTF
- We are transitioning to a new AAI, where the credentials used are JWT (tokens), signed by “issuers”, and containing authorizations for a specific operation.
  - OSG services only need to interpret the JWT; no mapping needed.
  - Authorization is based on the issuer’s assertion, *not* based on the user.
    - User information is kept for the .

X.509 -> tokens

So, where are we?

Server Authentication

- Starting in 2018, OSG Security began accepting a limited number of “CAB CAs” (the CAs used by your browser and operating systems).
  - These CAs verify domain ownership, not organizational ownership. This difference was viewed as a manageable acceptable risk (hint: same level of verification of www.wisc.edu).
  - This risk has worked out well in practice. The estimated cost to run a CA was $250k/year at the time.
- The WLCG still mandates use of IGTF host certificates.
  - The year-over-year number of CAs participating in the IGTF have been decreasing.
  - The current “CA of last resort” are commercial CAs; Sectigo (a company) holds the contract for both InCommon members and GEANT TCS.
    - These contracts are up for renewal. I’d guess outside the UK, Netherlands, and CERN, Sectigo does >90% of the IGTF host certificates.
    - There are other options. Probably 1-5 host certificates total are purchased through the DigiCert IGTF CA.
- Generally, the other GSI components (namespaces, CRLs) are not used; rather, TLS and the Internet PKI are.
  - GSI mostly ended with Globus Toolkit. XRootD reimplements some of this functionality.

Client authorization

- OSG 3.6 (February 2021) was our first release series where each major service could use capability-based authz.
  - I.e., “token support”
  - No dependency on Globus Toolkit
  - You could perform transfers with tokens, submit glideins with tokens, run a HTCondor pool with tokens.
  - Notably, XRootD still supported VOMS and had its own GSI-like protocol.
  - OSG 3.6 support ended June 2024.
- The addition of tokens doesn’t automatically move users off X.509:
  - In the past 3 years, we’ve added tooling around token acquisition and management.
  - HTCondor’s working on its second generation of token support. Simplified trust model, less moving parts, more reliable.
  - Pelican has token-management built in. Only legacy (LIGO) use cases for OSDF use X.509.
  - And then there’s selecting a token issuer…

Issuing Tokens

- There are three common implementations of token issuers used: “local issuer”, OA4MP (used by CILogon), and IAM.
- The “local issuer” is simplest and good for simple cases: have the private signing key on the service and generate a token directly.
- OA4MP: Integrated with Pelican, CILogon will operate as a contract-based service.
  - OA4MP is not a “full solution” but rather plugs in to your existing user/group management and authentication/SSO. No web UI.
  - Authorization policies are written in a scripting language called “QDL”. New/different policies & ideas are an update of a script, not a new software release.
  - In production at CHTC, JLab, Fermilab.
- IAM: Used by the WLCG community.
  - “Full stack”: can only issue tokens for users/groups in its database (API available for sync’ing between systems).
  - Token issuing component implemented in Java – evolving the system requires new software releases.
  - Only option that *also* can issue VOMS extensions.

OSPool, LIGO

- OSPool, for the most part, uses the local issuer. Tokens are issued completely transparently to the user – users only ‘see’ them when bugs disrupt the jobs.
- LIGO uses a combination of the local issuer and CILogon. Challenges in integrating DAGMan + CILogon – focus has been making things simpler and more reliable. For now, that means using local issuer.
  - Closely collaborating with the HTCSS team to evolve the token interaction.
  - When the current “O4” run started, token integration was not ready; some users on X.509. To not ‘pull the rug’ from production cases, they have asked OSG to commit to supporting X.509 use cases of OSDF until end of O4 (June 2025).
  - Lots of access went through CVMFS - wonderful when it works, impossible to debug when it doesn’t. Mostly converting to invoking the Pelican client directly.
  - LIGO is the big user of CILogon’s CA for ‘user certificates’. When O4 ends, CILogon plans to shutdown its CA component.

JLab, Fermilab

- JLab:
  - JLab has a CILogon contract for a token issuer. Integrated, online, in use.
  - JLab has not been a ‘squeaky wheel’. Current projects include moving them from a “role your own” XRootD setup to OSDF/Pelican. We can more easily support/change things in the Pelican layer to be responsive to their needs.
  - Some of the observed pain points have been at the HTCondor/Vault boundary layer.
- Fermilab:
  - Another CILogon issuer; also developed the integration with Vault.
    - Vault allows one to manage and distribute tokens via simpler workflows (easier to automate from CLI-esque environments) than OAuth2.
  - Tokens are heavily used for users/experiments on the “JobSub” framework.
  - Experiments moving data between sites (DUNE) are still doing this via X.509.
  - Fermilab has ~dozen experiments and a legacy systems to deal with. There will be a “long tail” of conversion.
    - Under the same CILogon user certificate CA shutdown deadline. <12 months to go!

WLCG

- The token transition is definitely a “LS3” project.
- IAM rollout is done: lingered for years, immensely accelerated once folks realized VOMS-Admin was never going to be upgraded past EL7.
  - Worrying symptom: Can WLCG make progress outside “crisis mode”?
  - Good sign: Starting to make progress on the “high priority” issues. New releases are happening.
  - Still has open security issues affecting tokens.
- The “middleware” component where token integration is struggling is FTS.
  - Success: A pre-release version of FTS was used in DC24 for token-based transfers. At its peak, half of the CMS traffic (by volume) was transferred via tokens.
  - Challenge: the code used in DC24 is not in an official release (DC24 was ~5 months ago!).
    - Comparison: the production Rucio code for tokens was released in November 2023.
- Beyond middleware, each experiment has its own TODO list to convert over their own systems.
  - Each experiment is proceeding on this on their own pace; not clear there’s a lot of coordination (coordination is likely not that important on experiment-specific software).
  - Example: CMS has a list of all tasks for the effort, their dependencies, and the assigned team for each task. Complete project is ~12 months.

Other Token Profiles

- A token is simply a JSON object (key-value pairs) where a few keys have standardized definitions.
  - Very flexible, one can design any AAI you want with tokens. You can re-implement GSI!
- Beyond WLCG, the other profile “in the community” is is the EGI Check-In / AARC profile.
  - The WLCG profile based on the idea that each experiment / entity signs assertions on its behalf.
  - The EGI Check-In model is a single, central issuer (the EGI Check-In issuer) is the single authority for the planet.
- Wildly different AAI models, not clear how the two will co-exist.
  - There’s a group, the “Grand Unified Token” (GUT) profile, trying to devise a common schema for both EGI and WLCG.
  - The profile is the easy part, a common AAI model is the hard part.
  - All OSG services go through a multi-profile abstraction layer. Anything the “GUT profile” group devises will map through this layer and OSG services will be able to support such tokens.
    - Whether the AAI models mesh – that remains to be seen.

Conclusions

- Transition to the new authorization model is uneven.
  - For PATh-related services and some organizations, this is done.
  - Fermilab and JLab experiments are on excellent path with their CILogon partnership.
  - For WLCG, this is firmly a “LS3” activity. Good recent progress – although some possibly only came due to external deadlines
- The most significant risk is FTS:
  - Small core team, difficult for external folks to contribute.
  - No/minimal contributors outside the CERN IT group.
  - Many priorities, many requests going to the team, no coordinated stakeholder meetings.
  - Critical for WLCG and Fermilab experiments.

Questions?

This project is supported by the National Science Foundation under Cooperative Agreements PHY-2323298 and OAC-2030508. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.