Token Transition update

GDB, 11 September 2024

M. Litmaath

Computing state of affairs (1)

- Campaign to have HTCondor CEs upgraded to maintained versions
  - Intermediary version: v9.0.20 – only available for CentOS 7
    - Supports tokens, SSL (no VOMS mapping) and GSI (with VOMS mapping)
  - Versions >= v23.x
    - Support tokens and SSL without VOMS mapping
  - Versions >= v23.5.2
    - Support tokens and SSL **with VOMS** mapping! (release notes)
  - To use SSL mappings with proxies, clients must also run **recent** versions!
  - All versions support *delegation* of VOMS proxies to be used by jobs and APEL
    - Mind this HTCondor (CE) setting for APEL: USE_VOMS_ATTRIBUTES = True
  - 53 tickets, >= 25 solved
  - Upgrading to **EL9** (or 8) at the same time
    - APEL client, parsers and python-argo-ams-library available from the WLCG repository

Computing state of affairs (2)

- APEL support for *tokens* is becoming essential as some VOs have **stopped** equipping their jobs with X509 / VOMS proxies
  - Not yet on the horizon for the LHC experiments
    - ALICE **HTCondor CE** jobs still come with VOMS proxies **only** for APEL
  - Stopgap approaches were suggested for the time being
    - Map token issuers / subjects / … to **pseudo** VOMS FQANs
    - The rest of the machinery can stay unchanged
  - Medium-term solution expected from the GUT Profile WG

- ARC 7 continues supporting X509 / VOMS besides tokens
  - Some VOs may prefer switching to tokens in the near future to make their job submissions more uniform

IAM service developments (1)

- All production instances at CERN are on v1.10.1 since Sep 9
  - A bugfix release on top of v1.10.0 that mainly addresses most of the remaining high-priority issues in the area of VO management
  - Other improvements are planned for this autumn

- The “dteam” instance is usable for service testing and monitoring with tokens as well as VOMS proxies
  - VO membership managed by EGI Operations and WLCG Ops Coordination

IAM service developments (2)

- New instances for the LHC experiments are available on **Kubernetes**, sharing their DBs with the current production instances on *OpenShift*
  - For better **load-balancing**, **logging**, **monitoring**, **GitOps** and **HA** options
  - They will replace the current production instances in the **next weeks**
    - Dates to be decided per experiment
  - Sites were ticketed to add support for the future VOMS endpoints and token issuers by May 31 → new deadline: Sep 16
    - 45 tickets still open, all have been reminded on Sep 2
    - **ETF** (SAM) **preprod** will only use the new configurations to check all services

- Transition to Kubernetes **HA setups** **has started**
  - Each IAM service will be load-balanced over 3 clusters in 3 different OpenStack availability zones, with a shared DB per VO
  - The HA configuration is **non-trivial** due to current requirements of the application
  - Already in **production** for most small VOs as of Aug 14
  - Remaining small VOs + ALICE will be done on Sep 11
  - ATLAS, CMS and LHCb are planned for Oct 2

Data Challenge 2024 followup (1)

- Several ideas for more **sustainable** use of tokens in large-scale data management have been **discussed between experts** of the services involved
  - Focusing on FTS workflows for now

More details on today's agenda of the XRootD & FTS workshop

- A **new model** was proposed for testing
  1. Tokens have scopes per individual file and longish lifetimes
     - A stolen token thus could be used for some time, but with only little potential damage
  2. The FTS just uses those tokens without any exchanges or refreshing
     - Thus avoiding a big load on itself as well as IAM
  3. If a token runs out, its corresponding transfer just fails, passing the ball back to Rucio / DIRAC

- The FTS code now supports this **new model** alongside the model used in DC24, which will remain needed for other communities
  - Further enhancements were also discussed and will be considered later

Data Challenge 2024 followup (2)

- ATLAS have started using this **in production** as of late August
  - 15 sites, all served by the CERN FTS, which has the new code
  - Starting with SCRATCH_DISK transfers, followed by DATA_DISK
  - No use of tokens during weekends for the time being
  - Typical token **rates** are 1-2 Hz, which occasional spikes of 5 Hz
  - **Lifetimes** currently are 2 weeks, to be reduced with more experience
  - The removal of tokens is left to the background cleanup job in IAM
    - Avoid Rucio complexity & interference that may affect IAM performance (see below)
  - The max number of concurrent tokens stored so far has been **1.7 M**
    - Already more than the overall maximum seen in DC24, no problems so far
  - A few SEs were found not to implement the WLCG token profile correctly
    - To be followed up separately

- The best would be to **stop storing access tokens** altogether
  - That implies moving DB handling code out of the third-party framework that is currently used under the hood, which in turn implies other changes first
  - The desired feature is planned to become available before the end of the year

Data Challenge 2024 followup (3)

- **High-rate stress tests** are desirable, but currently not an option while the same IAM instances are serving all other use cases
  - During DC24 we could have tolerated downtimes of 1 or 2 days, because IAM use cases were much less time-critical then

- An opportunity has been recognized: the migration from the IAM instances on *OpenShift* to the ones on **Kubernetes**
  - Instead of decommissioning the old services shortly, we can first reuse them for **stress tests** – when reconfigured with their own DB instances

- This would allow token usage to be steadily ramped up until we run into instabilities
  - Possibly due to **DB limitations** encountered by the current IAM code

AuthZ WG items

- Various IAM improvements are still desirable in the short term
  - Fixes for the last of the current high-priority issues
  - A similar dashboard is planned for the autumn releases
  - The next IAM hackathon will be held Nov 27-28 at **IJCLab**, Orsay

- Version 2.0 of the **WLCG token profile** is under preparation
  - Fixing a number of issues encountered with v1.0
  - A few **need to be agreed** in AuthZ or DOMA BDT WG meetings

- The **Token Trust & Traceability WG** met on July 23 and Aug 27
  - Aiming to equip site admins, VO experts, developers, … with best practices for token usage, which will also provide **input for policy documents**
    - Recipes, tools, log mining, testing, debugging, monitoring, banning, ...
    - Example: guidelines for **large-scale data transfers**, based on what works!

Conclusions and outlook

- **Collaborative** efforts will involve many of us in the next months
  - **IAM usability** for VO administration by LHC experiments and others
    - High-priority issues mostly done
  - **HA options** for LHC experiment IAM instances – almost there now!
  - **Data management:** a **new model** for large-scale data transfers
    - Aiming to reach the **next level** of token usage this autumn
  - HTCondor CE versions that are fully maintained
  - **APEL** adjustments for tokens – short vs. medium term
  - GUT Profile WG progress toward a new VO attribute – for accounting etc.
  - Version 2.0 of the WLCG token profile – to signal where we intend to go
  - More **deployment and operations** know-how – also input for policies
  - More use of auxiliary services – for robustness and hiding complexity