EPJ Web of Conferences **295**, 04054 (2024) https://doi.org/10.1051/epjconf/202429504054
CHEP 2023

WLCG Transition from X.509 to Tokens

Status, Plans, and Timeline

Thomas Dack[1,*], *Federica* Agostini[2], *Jim* Basney[3], *Linda* Cornwall[1], *John Steven* De Stefano Jr[4], *Dave* Dykstra[5], *Francesco* Giacomini[2], *Maarten* Litmaath[6], *Roberta* Miccoli[2], *Mischa* Sallé[7], *Hannah* Short[6], and *Enrico* Vianello[2], representing the WLCG Authorization WG

[1]Science and Technology Facilities Council (UKRI-STFC), United Kingdom
[2]Istituto Nazionale di Fisica Nucleare (INFN-CNAF), Italy
[3]National Center for Supercomputing Applications (NCSA), United States
[4]Brookhaven National Laboratory (BNL), United States
[5]Fermi National Acclerator Laboratory (FNAL), United States
[6]European Organization for Nuclear Research (CERN) Switzerland
[7]Nationaal Instituut voor Subatomaire Fysica (Nikhef), Netherlands

Abstract. Since 2017, the Worldwide LHC Computing Grid (WLCG) has been working towards enabling token-based authentication and authorization throughout its entire middleware stack. Following the initial publication of the WLCG Token Schema v1.0 in 2019, OAuth2.0 token workflows have been integrated across grid middleware. There are many complex challenges to be addressed before the WLCG can be end-to-end token-based, including not just technical hurdles but also interoperability with the wider authentication and authorization landscape.

This paper presents the status of the WLCG coordination and deployment work, and how it relates to software providers and partner communities. The authors also detail how the WLCG token transition timeline has progressed, and how it has changed since its publication.

1 Introduction

Five years ago, at CHEP 2018, in the plenary session *Beyond X.509: Token-based Authentication and Authorization for HEP* [1], the use of the INDIGO Identity and Access Management Service (INDIGO IAM) [2] as a replacement for the existing X.509 and VOMS (Virtual Organization Membership Service) [3] architecture was publicly proposed. This plenary presented the initial work done by the **WLCG Authorization Working Group** (the AuthZ WG), formed the previous year in the summer of 2017, and which had begun the process of defining architecture requirements in the course of that year.

Since that initial phase, significant progress has been made towards this transition, with deployment of IAM instances and integration and testing with experiments and services. The current objective for the AuthZ WG is to build upon the work done thus far and in particular, exercise the use of tokens at scale as part of the 2024 WLCG Data Challenge (DC24).

*e-mail: thomas.dack@stfc.ac.uk

1.1 Contributing Groups

The AuthZ WG was formed in 2017 and brings together experts from several projects and domains – including SciTokens [4], the INDIGO DataCloud project [5] and EGI [6] – in order to ensure that the authorization infrastructure is both full-featured and, to the extent possible, interoperable, and that technical and policy challenges faced are tackled appropriately.

2 WLCG Token Infrastructure

2.1 Motivations for Transition

There are a range of reasons to motivate the transition of the WLCG authentication and authorization infrastructure away from the established VOMS X.509-based system, towards one built on the OAuth 2.0 [7] and OIDC [8] token protocols. First and foremost, whilst providing a functioning solution for authentication and authorization across a distributed infrastructure, user certificates are unfamiliar to new researchers and even keep hampering experienced users. Coupled with this, by construction, authorization with personal certificates poses issues with regards to privacy and protection of user data.

In the years since X.509 was chosen in the early 2000's, the prevalence of token-based authentication and authorization has grown within industry, in particular within the Social Media space, with many web services now allowing users to "Sign in with Google" etc. The majority of users are now very familiar with web workflows which delegate access through the OAuth 2.0/OIDC flows, such that the adoption of these protocols presents users with a more familiar and intuitive process. Many software tools and libraries now support OAuth 2.0/OIDC flows by default thanks to this industry uptake; by enabling tokens for WLCG we are able to benefit from this easier integration.

Transitioning to tokens provides the possibility to introduce finer grained access controls. In fact, tokens can contain specific *capabilities*, limited for the purpose of the token, whereas VOMS proxies are typically more powerful than required. The precise level of authorization granularity (e.g. permissions per file, per experiment, etc.) is still under discussion and will be based on a middle ground being found between security and operational requirements, with different experiments possibly making different choices.

2.2 Token Infrastructure Design

Following extensive discussions and technical pilots [9], the decision was made to use INDIGO IAM (see Figure 1) for the WLCG Token issuer software. As of 2023, this software already follows the AARC Blueprint of best practices for Authentication and Authorisation Infrastructure design [10] [11], but not yet all AEGIS recommendations for integration with other infrastructures.

2.2.1 Deployment Technical Details

For the WLCG deployments of INDIGO IAM, certain configuration settings were chosen:

- Only two authentication sources were enabled - X.509 certificates and CERN Single Sign-On. All WLCG users have a CERN account, which provides additional assurance that the identity of the user has been confirmed.
- A periodic check is made against the CERN Human Resources database to confirm the affiliation of the user with the experiment, and block the user once it has expired.

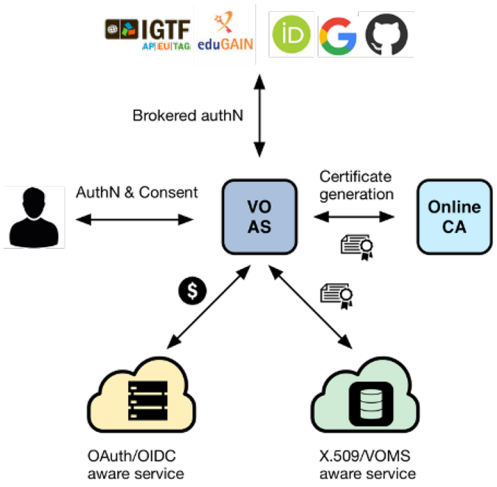

Figure 1. The INDIGO IAM software supports multiple Authentication (AuthN) sources and provides access control to both OAuth 2.0/OIDC and X.509/VOMS aware services. Integration with the RCauth online Certificate Authority (CA) is optionally supported.

The WLCG IAM instances have been deployed on the CERN OpenShift infrastructure, with the configuration stored in the CERN GitLab service. The MySQL database instances are hosted by the CERN Database-on-demand service and the monitoring uses the CERN central monitoring infrastructure comprising Fluentbit, Grafana, OpenSearch, Prometheus etc. Efforts have been made wherever possible to leverage standard, supported tools to promote quality of service and decrease support overhead on the IAM service operators. There are currently 4 WLCG IAM deployments at CERN (ALICE, ATLAS, CMS, and LHCb) and an additional testing instance hosted by INFN-CNAF.

3 Token Transition Status

The current aim for the WLCG AuthZ WG is to facilitate the uptake of tokens to support workflows across the main LHC experiments, and within the wider WLCG context. In the areas of data storage, access and transfers, progress is mainly driven through activities in the **WLCG Bulk Data Transfers WG**, comprising grid workflow experts from LHC experiments and partner communities (e.g. Belle II and DUNE), developers of data management middleware products, service managers and site representatives. The transition process is underway, though not without adjustments to initially conceived time-frames - due to both implementation and technology challenges.

3.1 Token Uptake to Date

3.1.1 LHC Experiments

The LHC experiments use tokens at least for job submissions to HTCondor CE instances, while for ARC CE instances the transition to tokens is not urgent and for certain sites may also depend on the ARC CE data management functionality to be enhanced with token support. A deployment campaign is soon foreseen to be launched for EGI sites with HTCondor CEs

to upgrade to *supported* versions that permit the continued use of VOMS proxies for job submissions, albeit with a much less flexible mapping scheme, which will allow those sites to keep working with communities that cannot switch yet to a supported type of tokens. The most recent HTCondor CE and ARC CE versions also bring support for the use of EGI Check-in tokens that many such communities are expected to make use of eventually.

The *sustained token issuance rates* of IAM test instances at CERN have increased to very promising levels thanks to deployment tuning, with further increases expected after DB handling improvements in IAM. Token workflow details for DC24 have been agreed between the FTS [12] and Rucio [13] teams, with further optimizations already foreseen to be considered afterwards. As more Storage Elements will continue to get configured with token support, large numbers of DC24 transfers are expected to be done with tokens, thereby allowing all token handling aspects to be tested *at scale* against realistic loads.

3.1.2 Wider WLCG

Whilst the core focus of the AuthZ WG has been the deployment of token-based workflows for the WLCG, this infrastructure must be able to support and interoperate with the authentication and authorization systems in place for other experiments. Examples include:

- **Belle II**
  The Belle II collaboration already have an IAM service and expect to upgrade their DIRAC [14] service to v8.0 or higher in the coming months and subsequently start using WLCG tokens initially for job submissions, while Belle II data transfers in DC24 may still have to rely on VOMS proxies instead.
- **DUNE**
  The authentication and authorization infrastructure for DUNE is hosted at FNAL, which is much further advanced in the transition to tokens. They are currently using WLCG-profile tokens in their production infrastructure. Their infrastructure is based on CILogon as the token issuer, HashiCorp Vault for storing refresh tokens, *htvault-config* for configuring Vault, *htgettoken* as the client of Vault, and an integration with HTCondor for keeping short-lived bearer tokens refreshed in batch jobs [15]. It is likely that at least some of the LHC experiments will use some of the same software components with their INDIGO IAM token issuers.

3.2 Token Transition Timeline

The WLCG Token Transition Timeline v1.0 [16] defines targeted milestones for the technical process of transitioning from X.509 to tokens.

The current focus of *development* is on facilitating the use of tokens in DC24.

M.2 (Dec 2022): DIRAC versions supporting job submission tokens deployed for concerned VOs

Delayed. LHCb have upgraded to v8.0 and validated job submission to HTCondor and ARC CEs with tokens. Belle II expect to upgrade in the next months.

M.3 (Feb 2023): VOMS-Admin is switched off for one or more experiments

Delayed. Supporting work underway, but pushed back to allow for further IAM development to improve on VO use-cases and concerns.

M.4 (Mar 2023): HTCondor installations at EGI sites have been upgraded to supported versions > 9.0.x

Delayed. Work in progress, but milestone postponed to the spring of 2024.

M.5 (Mar 2023): End of HTCondor support for GSI Auth

Achieved. Was postponed to May. Officially there is no supported version featuring GSI as of that month. The HTCondor team have provided newer 9.0.x versions as stepping stones for EGI sites toward supported versions >= 10.x. EGI and WLCG Operations will run campaigns to help sites get there in the next months.

M.6 (Mar 2023): Some storage endpoints provide support for tokens

Achieved. A steadily increasing number of CMS production storage services already pass token tests. ATLAS also have a number of early adopters and aim for all their big sites to be ready by DC24.

M.7 (Feb 2024): Rucio, DIRAC, and FTS have sufficient token support in released versions to perform DC24 using token authorization.

On track. Currently work in progress.

M.8 (Mar 2024): Sufficient storage endpoints support tokens to allow DC24 to be done using only tokens.

On track. Having WLCG token support ready in time for DC24 is a major current emphasis. It remains to be seen if, for a given experiment, *all* data transfers can be done with tokens. It looks much more likely and in fact would be sufficient that *large fractions* will be making use of tokens.

M.9 (Mar 2025): Grid jobs use tokens for reading and stageout.

Requiring changes also inside the job pilots.

M.10 (Mar 2026): Users no longer need X.509 certificates

There is a longer gap between M.9 (Mar 25) and M.10 (Mar 26), so as to allow the development of utilities, workflows, and onboarding to ensure a smooth user transition. Users should need to know nothing about tokens.

4 Conclusions and Outlook

The WLCG has positive momentum for its token transition, with progress being made across the grid. Though the token transition timeline has seen some delays, progress has been steady in many areas concerned.

The main milestones at this time are:

- **The switch to HTCondor CE versions that no longer support GSI**
  Several scenarios are required in order to smooth the transition for legacy use cases.
- **Integration and deployment ready to run DC24 using tokens to a large extent**
  Aim for production-quality infrastructure at the largest sites (preferably all T0/1 plus the big T2 sites).

Notable work is underway to meet these objectives, including the implementation of the required level of token support across the Rucio, DIRAC, and FTS services, as well as a sufficient level of storage endpoints. Demonstration of data transfers using tokens during DC24 is a key step in achieving the ultimate transition goal, where the current user-intensive certificate authorisation flows have been replaced with new token workflows. This work lays the technical foundation for user migration and onboarding activities, planned for completion with Milestone 10 in March 2026.

References

[1] A. Ceccanti, E. Vianello, M. Caberletti, F. Giacomini, *Beyond X.509: token-based authentication and authorization for HEP* (2019), https://doi.org/10.1051/epjconf/201921409002

[2] *INDIGO Identity and Access Management Service (IAM)*, https://indigo-iam.github.io/, https://doi.org/10.5281/zenodo.8366226

[3] *Virtual Organisation Membership Service (VOMS)*, https://italiangrid.github.io/voms, https://doi.org/10.5281/zenodo.1875371

[4] A. Withers, B. Bockelman, D. Weitzel, D. Brown, J. Gaynor, J. Basney, T. Tannenbaum, Z. Miller, *SciTokens: Capability-Based Secure Access to Remote Scientific Data*, in *Proceedings of the Practice and Experience on Advanced Research Computing* (Association for Computing Machinery, New York, NY, USA, 2018), PEARC '18, ISBN 9781450364461, https://doi.org/10.1145/3219104.3219135

[5] *Indigo Data Cloud*, https://www.indigo-datacloud.eu

[6] *EGI*, https://www.egi.eu

[7] *RFC6749: The OAuth 2.0 Authorization Framework*, https://www.ietf.org/rfc/rfc6749.txt

[8] *The OpenID Connect Authentication Protocol: OpenID Connect Core 1.0 incorporating errata set 1*, https://openid.net/specs/openid-connect-core-1_0.html

[9] *WLCG AAI pilots update*, https://indico.cern.ch/event/805596/#2-wlcg-aai-pilots-update

[10] *Authentication and Authorisation for Research and Collaboration (AARC)*, https://aarc-project.eu

[11] *Authentication and Authorisation for Research and Collaboration (AARC) Community Guidelines*, https://aarc-community.org/guidelines/

[12] *The CERN File Transfer Service*, https://fts.web.cern.ch

[13] *RUCIO: Scientific Data Management*, http://rucio.cern.ch/

[14] *DIRAC: the Interware*, https://github.com/DIRACGrid

[15] D. Dykstra, M. Altunay, J. Teheran, *Secure Command Line Solution for Token-based Authentication* (EPJ Web of Conferences, 2021), CHEP '21, https://doi.org/10.1051/epjconf/202125102036

[16] WLCG Authorization Working Group, *WLCG Token Transition Timeline (1.0)* (2022), https://doi.org/10.5281/zenodo.7014668