EPJ Web of Conferences **245**, 03001 (2020) https://doi.org/10.1051/epjconf/202024503001
CHEP 2019

WLCG Authorisation

from X.509 to Tokens

Brian Bockelman[6], *Andrea* Ceccanti[3], *Ian* Collier[4], *Linda* Cornwall[4], *Thomas* Dack[4], *Jaroslav* Guenther[1], *Mario* Lassnig[1], *Maarten* Litmaath[1], *Paul* Millar[2], *Mischa* Sallé[5], *Hannah* Short[1,*], *Jeny* Teheran[7], and *Romain* Wartel[1]

[1]European Organization for Nuclear Research (CERN), Switzerland
[2]Deutsches Elektronen-Synchrotron (DESY), Germany
[3]Istituto Nazionale di Fisica Nucleare (INFN), Italy
[4]Science and Technology Facilities Council (UKRI-STFC), United Kingdom
[5]Nationaal Instituut voor Subatomaire Fysica (Nikhef), Netherlands
[6]Morgridge Institute for Research, United States
[7]Fermi National Accelerator Laboratory, United States

Abstract. The WLCG Authorisation Working Group was formed in July 2017 with the objective to understand and meet the needs of a future-looking Authentication and Authorisation Infrastructure (AAI) for WLCG experiments. Much has changed since the early 2000s when X.509 certificates presented the most suitable choice for authorisation within the grid; progress in token based authorisation and identity federation has provided an interesting alternative with notable advantages in usability and compatibility with external (commercial) partners. The need for interoperability in this new model is paramount as infrastructures and research communities become increasingly interdependent. Over the past two years, the working group has made significant steps towards identifying a system to meet the technical needs highlighted by the community during staged requirements gathering activities. Enhancement work has been possible thanks to externally funded projects, allowing existing AAI solutions to be adapted to our needs. A cornerstone of the infrastructure is the reliance on a common token schema in line with evolving standards and best practices, allowing for maximum compatibility and easy cooperation with peer infrastructures and services. We present the work of the group and an analysis of the anticipated changes in authorisation model by moving from X.509 to token based authorisation. A concrete example of token integration in Rucio is presented.

1 Introduction

This paper describes ongoing work by the WLCG Authorisation Working Group [1] to transition the Worldwide LHC Computing Grid's authorisation model from X.509 Certificates [2] and certificate proxies [3] to tokens. The tokens are defined as JSON Web Tokens (JWT) [4], to be provisioned over OpenID Connect (OIDC) [5] and OAuth2 [6] workflows.

*e-mail: hannah.short@cern.ch

1.1 Motivation

When X.509 was chosen in the early 2000s, there was no mature alternative. The Globus Toolkit [7], chosen by WLCG, supported X.509 and provided a functional solution for distributed authentication and authorisation, when coupled with policies controlled by the Interoperable Global Trust Federation (IGTF) [8]. Despite providing a solution, certificate management proved a steep learning curve for researchers and the authorisation model adopted (although sufficient at the time) is outdated by today's standards regarding the protection of privacy of user data.

In the meantime, alternative authentication and authorisation technologies have emerged and X.509 is no longer considered the most suitable option. It is increasingly common that many off-the-shelf software choices and commercially hosted services offer inbuilt options for federated authentication, using standards such as SAML [9] or OIDC. Integrating grid software with such third party systems has therefore become progressively more difficult, since a lot of X.509 developments are deeply integrated in the grid software stack. Users are now well accustomed to web-based authorisation workflows that use OAuth2 or OIDC to delegate access rights, meaning that adoption of such technologies would present a more user-friendly experience. These factors came together to build a strong motivation for redesigning the WLCG authorisation model based on JWTs over OIDC and OAuth2.

1.2 WLCG Authorisation Working Group

The WLCG Authorisation Working Group was formed in the summer of 2017, at a time when multiple activities were independently beginning to seriously consider token based authorisation. SciTokens [10] had produced a working system in the US focusing on capability-based authorization and an increasing number of efforts were ongoing in Europe to enable token based workflows, in particular in the context of the INDIGO DataCloud project [11] and EGI [12]. Experts from multiple domains came together in this group to chart a path towards token based authorisation for WLCG. Work to enhance software was supported by several European Commission Projects: EOSC-Hub [13], EOSC Pilot [14] and AARC2 [15]. The group's objective was to understand and meet the requirements of a future-looking authorisation service for WLCG experiments. In particular, effort focused on addressing usability, i.e. removing the need for users to manage certificates themselves and enable more intuitive authentication methods. A second objective was to simplify integration with third party services and software by adopting industry standards. A significant effort has been made to be as inclusive as possible, by including resource and service providers during the construction of the schema and the testing phase, as well as making all outputs available to the wider community such as through the FIM4R [16] workshop at Fermilab in September 2019. Documentation and a list of meetings are maintained on the WLCG Authorisation Working Group Twiki [1].

2 Milestones

Several key milestones have been achieved since the WLCG Authorisation Working Group was formed, with major progress included in Figure 1. In particular, a competitive study of two pilot solutions was made, resulting in an identified technical solution in March 2019. A schema (see Section 4) was negotiated between multiple stakeholders over the course of a year and was published in September 2019.

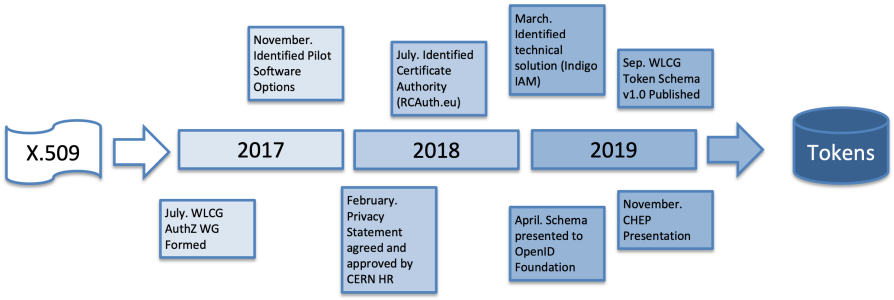

Figure 1. Key milestones achieved between 2017 and 2019 towards a token based authorisation scheme. Note, it is not implied that the final goal has been reached; much additional effort is required to fully transition from X.509 to tokens.

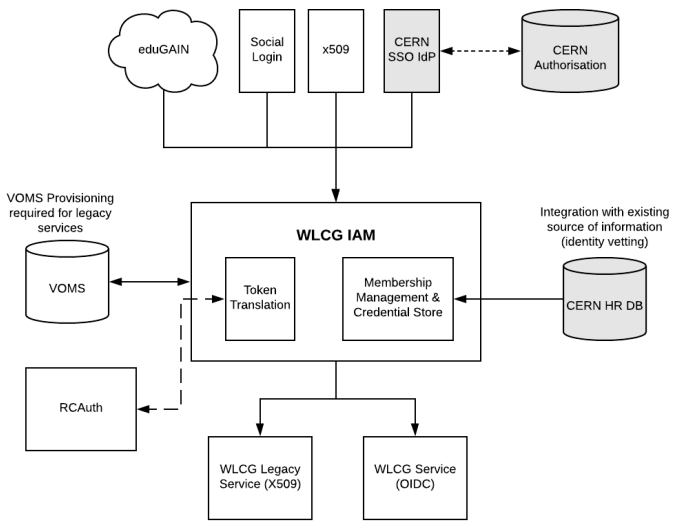

Figure 2. Design of the WLCG Token Authentication and Authorisation Infrastructure. Grey boxes represent configurable components, in this case tuned for the CERN environment.

3 Technical Design

The INDIGO IAM [17] software was chosen as the core of WLCG's future, token based authentication and authorisation infrastructure. In Figure 2, we see the WLCG IAM (a WLCG Instance of INDIGO IAM) in the centre, proxying multiple authentication channels, centralising authorisation, providing token translation to and from X.509, and serving as a uniform token issuer for downstream WLCG Services. The Token Translation is performed using the RCauth.eu [18] online Certificate Authority and users can request a certificate from the WLCG IAM User Interface. This architecture is in line with guidelines from the AARC Project, including the AARC Blueprint Architecture.

3.1 Usage Flow

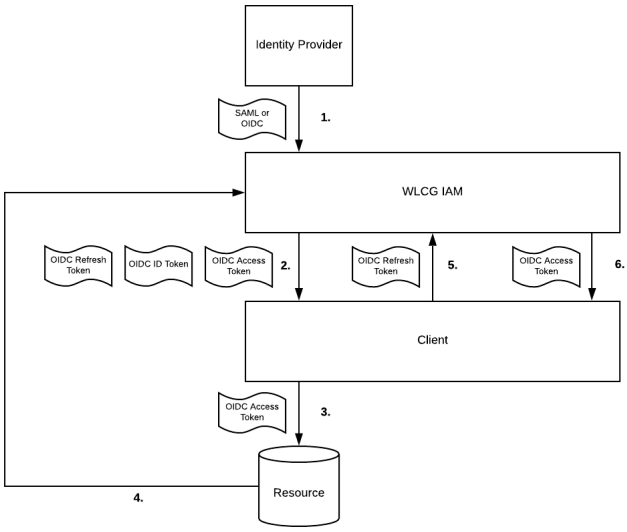

Figure 3. Anticipated token flow, starting from a SAML or OIDC token being issued from the Identity provider to WLCG IAM. WLCG IAM then issues multiple OIDC tokens to downstream clients, granting access to protected resources.

It is envisaged that there will be a small number of registered clients for WLCG and a much larger number of unregistered Resources. The user's authentication from their Identity Provider is technically decoupled from the tokens issued by the WLCG IAM. A user may authenticate with SAML, OAuth2, OIDC, X.509 or potentially any future protocol that is accepted by WLCG IAM. A more detailed description of IAM [19] and the work done to address WLCG requirements is given in a dedicated contribution to this journal [20]. Figure 3 shows the anticipated token workflow, the numbered steps are explained below:

1. A SAML or OIDC credential is sent from an Identity Provider to WLCG IAM

2. WLCG IAM sends three OIDC tokens to a registered Client, the ID Token, Access Token and Refresh Token

3. The Access Token is used to authorise access to downstream Resources (such as storage)

4. The Resource validates the token against known trust roots. The trust roots may be cached in advance to avoid a high number of round trips to the token issuer.

5. When the Access Token expires, the Client may use the Refresh Token to request a newly valid Access Token from WLCG IAM

6. WLCG IAM returns an Access Token to the Client, which may repeat steps 3 to 6

4 Schema

The WLCG Common JWT Profiles document was published on September 25th, 2019 [21]. This document describes how WLCG users may use the available geographically distributed resources without X.509 credentials. In this model, clients are issued with bearer tokens; these tokens are subsequently used to interact with resources. The tokens may contain authorization groups and/or capabilities, according to the preference of the Virtual Organisation (VO), applications and relying parties.

Wherever possible, the document builds on existing standards when describing profiles to support current and anticipated WLCG usage. In particular, three major technologies are identified as providing the basis for this system: OAuth2 (RFC 6749 and RFC 6750), OIDC and JWTs (RFC 7519). Additionally, trust roots are established via, among others, the OpenID Discovery [22] mechanism or OAuth2 Authorization Server Metadata (RFC 8414 [23]). The document provides a profile for OAuth2 Access Tokens and OIDC ID Tokens, the claims for which are shown in Table 1. Although many WLCG requirements were covered by standard claims defined in RFC7519 [4] or OIDC core [5], a claim to convey assurance was taken from the Research and Education Federations (REFEDS) [24] specification. Claims to convey version (*wlcg.ver*) and groups (*wlcg.groups*) have been defined within the WLCG schema document itself.

Claim	Origin	WLCG OIDC ID Token	WLCG OAuth2 Access Token
sub	RFC7519	X	X
exp	RFC7519	X	X
iss	RFC7519	X	X
acr	OIDC core	X	X
aud	RFC7519	X	X
iat	RFC7519	X	X
nbf	RFC7519	X	X
jti	RFC7519	X	X
eduperson_assurance	REFEDS	X	X
wlcg.ver	WLCG	X	X
wlcg.groups	WLCG	X	X
auth_time	OIDC core	X	
standard OIDC claims	OIDC core	X	
scope	*Inspired by Auth token exchange draft*		X

Table 1. Token Claims used in the WLCG Schema, including their origin in either RFC7519 [4], OIDC core [5], the Research and Education Federations (REFEDS) [24] or the WLCG Schema itself [21].“X” indicates their use in ID and Access tokens. Claim definitions can be found at the origin and are not included here for brevity.

4.1 Authorisation

Two forms of Authorisation exist in parallel in the WLCG Common JWT Profile Schema; Group based Authorisation and Capability based Authorisation. Groups are semantically equivalent to existing VOMS groups, i.e. a group contains a list of users. A capability is

the ability to perform an action, optionally at a specific path, e.g. *"the bearer of this token is authorised to write to /data"*. Extended examples are included in the Schema document [21].

5 WLCG Token Issuer

In late 2019, a WLCG token issuer was set up at the Istituto Nazionale di Fisica Nucleare (INFN) to provide a stable platform against which software enhancement could be tested. The token issuer was made available at https://wlcg.cloud.cnaf.infn.it/, primarily to facilitate tests planned by the WLCG DOMA Working Group. Clients are able to register to receive authentication tokens from the issuer over OIDC, and a web based platform is available to visualise tokens and their content. A VOMS attribute authority linked to this IAM instance has also been deployed, to allow users to get VOMS attribute certificates for the `wlcg` VO. The objective is to demonstrate IAM interoperability with the current X.509 WLCG AAI and enable a gradual transition to tokens.

6 WLCG Tokens working example: Rucio

Rucio [25] is one of the main Data Management tools used by WLCG experiments and is a key service to be enhanced to accept WLCG tokens. During 2019, the Rucio authentication and authorization mechanism was extended to support JWTs using the Open ID Connect protocol (following the OAuth 2.0 specifications). The implementation is based on the OIDC certified [26] *pyoidc* [27] library and follows the WLCG specification. During this process several improvements to the schema were identified, which will be incorporated into the next version of the WLCG token profile specification. To perform operations with Rucio *1.22.0.dev3* a user can now log in via the authorization code flow, which has been implemented in both the Rucio WebUI and Rucio command line interface (CLI). Several CLI login workflows are supported - including username and password, or session validation using a browser - to fulfil a variety of researcher use cases. The Rucio REST API has been configured as an OAuth2 protected resource and accepts tokens from trusted OIDC providers (including the WLCG token issuer discussed in Section 5).

6.1 Rucio and Downstream Storage

Tests have been successfully performed to demonstrate token based authorization from Rucio to the File Transfer Service (FTS) [28]. When a Rucio user requests a file transfer, Rucio can use either a JWT issued for the user or a JWT issued to the Rucio Admin user itself, which submits the job on the user's behalf. In order to allow fine grained access control downstream (Rucio → FTS3), token exchange and token refresh grant flows were also implemented. The first functional tests of a third party copy were performed (Rucio → FTS3 → dCache) and a version *1.22.0.dev3* of Rucio supporting all these new features is currently deployed on the DOMA instance [29]. Additional tests with other downstream storage elements are planned.

7 Conclusion

The WLCG Authorisation Working Group is well on the way to enabling token based authorisation for WLCG workflows. This work has been possible thanks to the participation of multiple individuals and projects, who together have been able to identify a technical solution and define a token schema. Requirements from Virtual Organizations and other research infrastructures outside of WLCG have been included as a continuous cooperative process to

guarantee interoperability. Preliminary token based workflows have been tested with many more planned for 2020. Further efforts are required to complete the enhancement of middleware, and to ensure that the existing WLCG policies are adhered to in the years to come.

This work was achieved thanks to support from the following European Commission funded projects; EOSC-Hub, EOSC Pilot and AARC2. EOSC-hub receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777536. EOSCPilot.eu has received funding from the European Commission's Horizon 2020 research and innovation programme under the Grant Agreement no 739563. AARC2 is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement number 730941. This material is based upon work supported by the National Science Foundation under Grant No. 1836650.

References

[1] WLCG Authorisation Working Group Twiki, https://twiki.cern.ch/twiki/bin/view/LCG/WLCGAuthorizationWG

[2] Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile, RFC 5280, https://tools.ietf.org/html/rfc5280

[3] Internet X.509 Public Key Infrastructure (PKI) Proxy Certificate Profile, RFC 3820, https://tools.ietf.org/html/rfc3820

[4] JSON Web Tokens, RFC 7519, https://tools.ietf.org/html/rfc7519

[5] OpenID Connect, https://openid.net/specs/openid-connect-core-1_0.html

[6] OAuth2.0, RFC 6749, https://tools.ietf.org/html/rfc6749

[7] Globus Toolkit, http://toolkit.globus.org

[8] Interoperable Global Trust Federation (IGTF), https://www.igtf.net

[9] Security Assertion Markup Language, SAML, RFC 7522, https://tools.ietf.org/html/rfc7522

[10] A. Withers, B. Bockelman, D. Weitzel, D. Brown, J. Gaynor, J. Basney, T. Tannenbaum, and Z. Miller *SciTokens: Capability-Based Secure Access to Remote Scientific Data.* Proceedings of PEARC '18. July 2018. https://doi.org/10.1145/3219104.3219135

[11] INDIGO DataCloud, https://www.indigo-datacloud.eu

[12] EGI, https://www.egi.eu

[13] EOSC Hub, https://www.eosc-hub.eu

[14] EOSC Pilot, https://eoscpilot.eu

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

[16] Federated Identity Management for Research and Collaboration, https://fim4r.org

[17] INDIGO Identity and Access Management (IAM), https://indigo-iam.github.io/docs

[18] RCauth, https://rcauth.eu

[19] A. Ceccanti, E. Vianello, M. Caberletti, F. Giacomini. *Beyond X.509: token-based authentication and authorization for HEP.* EPJ Web Conf., **214** 09002, (2019) https://doi.org/10.1051/epjconf/201921409002

[20] A. Ceccanti, E. Vianello, M. Caberletti, F. Giacomini. *Beyond X.509: token-based authentication and authorization in practice.* Submitted for publication in this journal

[21] M. Altunay, et al. *WLCG Common JWT Profiles (Version 1.0).*, September 2019, Zenodo. http://doi.org/10.5281/zenodo.3460258

[22] OpenID Connect Discovery, https://openid.net/specs/openid-connect-discovery-1_0.html

[23] OAuth 2.0 Authorization Server Metadata, RFC 8414, https://tools.ietf.org/html/rfc8414

[24] Research and Education Federations Group, REFEDS, https://refeds.org

[25] M. Barisits, T. Beermann, F. Berghaus et al., *Rucio - Scientific Data Management*, Comput Softw Big Sci **3**: 11, (2019) https://doi.org/10.1007/s41781-019-0026-3

[26] Certified OpenID Connect Implementations https://openid.net/developers/certified/

[27] Python OIDC library https://github.com/OpenIDC/pyoidc

[28] File Transfer Service http://information-technology.web.cern.ch/services/file-transfer

[29] Rucio DOMA instance https://rucio-doma.cern.ch