EPJ Web of Conferences **214**, 09002 (2019) https://doi.org/10.1051/epjconf/201921409002
CHEP 2018

Beyond X.509: token-based authentication and authorization for HEP

Andrea Ceccanti[1,*], *Enrico* Vianello[1], *Marco* Caberletti[1], and *Francesco* Giacomini[1]

[1]INFN-CNAF, via Berti Pichat 6/2 40137 Bologna

Abstract. X.509 certificates and VOMS have proved to be a secure and reliable solution for authentication and authorization on the Grid, but also showed usability issues and required the development of ad-hoc services and libraries to support VO-based authorization schemes in Grid middleware and experiment computing frameworks. The need to move beyond X.509 certificates is recognized as an important objective in the HEP R&D roadmap for software and computing, to overcome the usability issues of the current AAI and embrace recent advancement in web technologies widely adopted in industry, but also to enable the secure composition of computing and storage resources provisioned across heterogeneous providers in order to meet the computing needs of HL-LHC. A flexible and usable AAI based on modern web technologies is a key enabler of such secure composition and has been a major topic of research of the recently concluded INDIGO-DataCloud project. In this contribution, we present an integrated solution, based on the INDIGO-DataCloud Identity and Access Management service that demonstrates how a next generation, token-based VO-aware AAI can be built in support of HEP computing use cases, while maintaining compatibility with the existing, VOMS-based AAI used by the Grid.

Introduction

The current WLCG Authentication and Authorization Infrastructure (AAI), shown in Figure 1, is composed of the following main building blocks:

- the trust fabric: provided by IGTF [1], it basically tells services which are the certificate authorities (CAs) that can be trusted;
- X.509 certificates [2]: issued by trusted CAs to users and services for mutual authentication purposes;
- Proxy certificates [3]: a mechanism used to implement single sign-on and delegation starting from X.509 certificates;
- Virtual Organization Membership Service (VOMS) attribute certificates [4]: X.509 attribute certificates [5] embedded in proxy certificates and used to augment identity information with VO-issued authorization attributes that drive the authorization at services.

During the past years this AAI has proved to work quite well, providing a secure infrastructure that has scaled to millions of jobs and hundreds of sites, has supported important scientific discoveries and has been adopted by several research communities besides HEP.

*e-mail: andrea.ceccanti@cnaf.infn.it

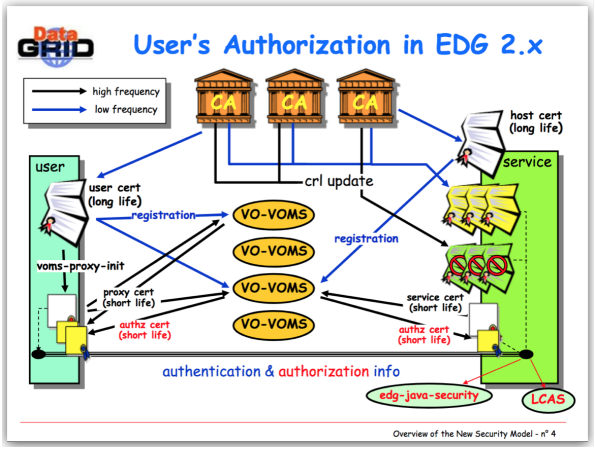

Figure 1. The current WLCG AAI as appeared in an ancient summary slide from the European DataGrid project, 2002.

Experience has also exposed the main limitations of this AAI, the biggest one being probably poor usability: handling certificates is a convoluted process that annoys most scientific users and leads to errors, resulting in user support requests and complaints. VOMS proxy certificates do not work in browsers, so complex workarounds had to be put in place in order to integrate VOMS with Science Gateways and experiment frameworks.

It is also hard to integrate support for identity federations like EduGAIN [6], since the authorization model at Grid services, being based on VOMS, is tightly bound to X.509 certificates.

Finally, X.509 has often represented a barrier when integrating computing and storage resources from external partners, such as commercial providers, hybrid clouds and HPC centers, whose importance is becoming crucial to address the expected computing needs of HL-LHC [7].

In this contribution we describe a novel AAI, conceived in the context of the INDIGO DataCloud project [8], which is based on industry standard technologies and represents a possible solution for the aforementioned problems.

1 A token-based AAI for HEP

What are the main requirements that a novel AAI for WLCG should satisfy?

- *Support a VO-centric authorization model*: today, computing and storage resources are shared in the context of a VO. Access to resources is granted, with varying levels of authorization, only to those who can prove membership in the Virtual Organisation (VO) by presenting an authorization token issued by a trusted central VO service;
- *Flexible authentication*: the AAI should support many authentication mechanisms, from identity federations (e.g., eduGAIN [6]) and social logins (e.g., Google, Facebook, or LinkedIn single sign-on) to X.509 certificates. Different authentication mechanisms would be mapped to different level-of-assurance profiles, so that authorization policies can be applied accordingly at services;

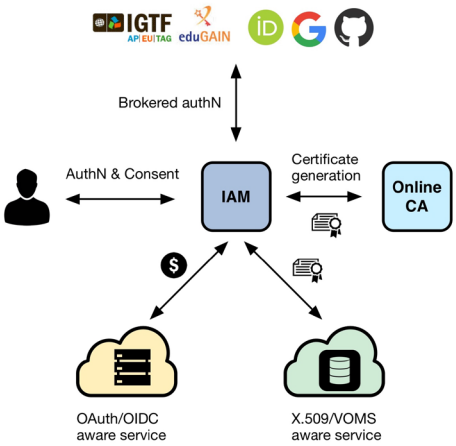

Figure 2. A token-based AAI for WLCG.

- *Authentication and authorization orthogonality*: authorization is decoupled from authentication mechanisms by introducing an intermediate credential, the authorization token, which provides all the information needed to grant access to shared resources;
- *Account linking and identity harmonization*: users should have the ability to link their multiple identities (e.g., their institutional login, social accounts, ssh keys, etc.) to their VO account. This account is assigned a persistent, VO-scoped opaque identifier which is then used across the infrastructure for traceability and accounting purposes;
- *Support for delegation and long-running jobs*: the AAI should support constrained delegation, where a user can delegate part of her rights to an agent or a service that acts on user's behalf. The service must be able to further delegate a subset of these rights to other services down the line, with the explicit or implicit approval of the user. The AAI could provide provide mechanisms and tools to safely define trusted delegation chains, i.e., which services can take part in a delegation chain and which set of privileges can be delegated across the chain;
- *Provisioning APIs*: the AAI must provide the ability to provision, manage and de-provision identity and authorization information to relying services, to enable, for instance, local or service-specific account management;
- *Integration and token translation*: the AAI must support integration with legacy or external services that cannot be modified to natively support tokens through controlled token translation.

1.1 A VO-centric token-based AAI

We propose the same model that was successful for VOMS, shown in Figure 2. We envision a central, VO-scoped Identity and Access Management (IAM) service that deals with user authentication supporting various mechanisms and that exposes information about user's identity and authorization attributes to services in two ways:

- via JSON Web Tokens (JWTs) [9] and standard OAuth [10] and OpenID Connect [11] protocol flows;
- via VOMS X.509 Attribute Certificates (ACs), as is the case for the current WLCG AAI.

Supporting both token-based and VOMS-based authorization will allow a gradual transition from the current X.509 and VOMS-based AAI to the new, token-based authorization model. Legacy services will continue to work without changes, while new services could be integrated leveraging the token-based paradigm.

In order to hide certificate management complexity from users, an online Certificate Authority such as RCAuth.eu [12] can be integrated with the central IAM service, to provide user certificates on demand. The generated proxy certificates can then be provided to agents or services acting on behalf of the users via provisioning APIs, as part of the information returned in an OpenID Connect authentication flow or leveraging other means like SSH authentication [13].

1.2 OAuth, OpenID Connect and JWTs

OAuth 2.0 [10] is the standard framework for delegated authorization in industry for HTTP services and is the main building block of the proposed AAI. It defines authorization flows targeted at service, desktop and mobile applications that describe how access tokens can be obtained from authorization servers and presented at services to be granted access to resources.

OpenID Connect [11] extends OAuth with an identity layer. It defines how authentication information is actually provided to services, so that authorization, accounting and other user identity-related functionality can be implemented.

JSON Web Tokens (JWTs) provide a mechanism to express in a secure way claims meant to be exchanged between services. Claims typically describe user identity, authentication properties, attributes and capabilities. Like VOMS, JWTs support distributed verification of token signatures.

1.3 Token-based authentication and authorization

In the authorization model we envision, shown in Figure 3, services expose functionality through OAuth protected APIs: only agents presenting a valid and trusted access token are granted access. Access tokens, which are signed JWTs, can be obtained by client applications (browsers, command line interfaces (CLIs), or other services) from the central IAM service. Access tokens, depending on VO configuration and security requirements, can provide identity information (e.g., an opaque user identifier, groups, and other attributes) and other authorization information (e.g., capabilities). Authorization is then performed at services based on the token contents, after a token verification step that assesses token integrity and validity.

Some services may require the exchange of the VO issued access token with a local, service-specific one issued by the service itself that is then used to drive authorization decisions. In both scenarios, the agent is authorized based on the information asserted by the central VO service in the initial access token.

The token-based authorization model described above is very similar to the VOMS-based one used in production today, but has two main advantages:

- it is not bound to X.509 certificates or any other specific authentication mechanism
- it is already widely adopted in industry and can be implemented with existing libraries and off-the-shelf components

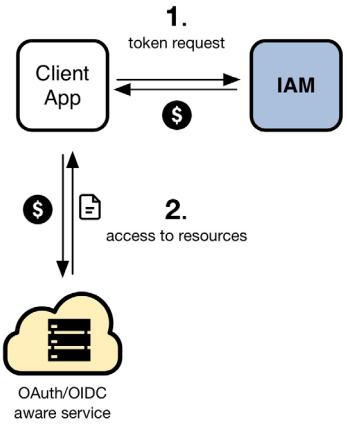

Figure 3. Token-based authentication and authorization for WLCG.

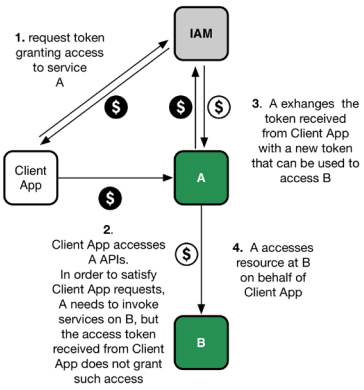

Figure 4. Token exchange for controlled delegation of privileges.

1.4 Trust and discovery

Trust is managed is a way similar to the one used today for VOMS. The list of trusted VO authorization servers is statically configured for a service by the service administrator. This configuration provides the minimum information required to obtain key material from the central server, using standard protocols [14, 15].

Services that will drive an authentication flow will need to be explicitly registered at the central VO authorization service, as required by the OpenID connect and OAuth standards. This registration process could require privileged access, depending on the security policies agreed upon by the infrastructure and the VO.

1.5 Controlled delegation of privileges

OAuth has native support for delegation, and was in fact designed to enable controlled third-party application access to user resources like Facebook or Twitter timelines.

The basic delegation mechanism is implemented via authorization flows that result in a client application receiving an access token which grants access to a specific set of resources. This token is typically a *bearer* token, a token whose access rights are granted to anyone presenting it, no other proof-of-possession steps are needed. This is the main difference with the WLCG authorization model, where an authentication step based on an X.509 certificate chain is always required when interacting with a service.

In the bearer-token model, a simple delegation scheme is implemented by passing access tokens across services participating in a delegation chain. This simple approach, while suitable for some scenarios, has the drawback of going against the principle of least privilege in requiring a token that could work for several possibly heterogeneous services and that, if stolen, could grant significant access rights to an attacker (in a way similar to how today an attacker stealing a VOMS proxy certificate has basically unlimited access to VO resources for the lifetime of the proxy).

OAuth provides finer-grained control of privileges (and audience) with scopes. JWTs provide a dedicated claim, `audience`, to restrict the token scope only to selected services/resources.

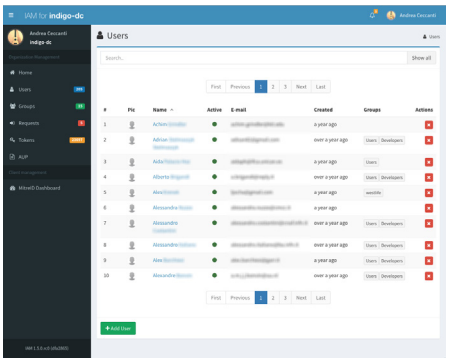

Figure 5. The INDIGO IAM management dashboard.

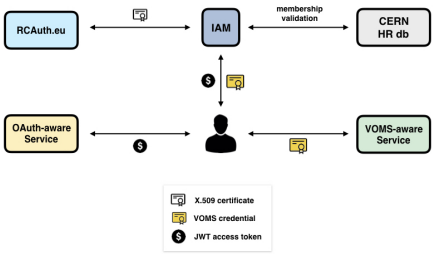

Figure 6. IAM WLCG integration work.

To support the use of fine-grained and scoped access tokens, a more secure approach for delegation is achieved by providing a token exchange endpoint at the central VO authorization service. In this scenario, shown in Figure 4, a service that needs to interact with a downstream service exchanges a local token with one suitable for the downstream call by interacting with the central VO server. At the central VO server, a VO policy determines which services are authorized to exchange tokens and which privileges can be delegated in a given chain. The protocol used in this exchange [16] is being standardized by the IETF.

Long-running computations are supported thanks to OAuth refresh tokens which are long lasting credentials used to obtain a new access token from an authorization server when the current token expires. One example of a successful application of this approach is the CMS Dynamic On-Demand Analysis Service (DODAS) [17].

2 The INDIGO Identity and Access Management Service

The INDIGO Identity and Access Management (IAM) service is the core of the AAI and is designed to satisfy all the requirements and guidelines described in Section 1. It provides a central VO-scoped authorization server, dealing with user authentication, registration and high-level authorization for a VO. This centralization of the authentication responsibility in a single service is an emerging architectural pattern (the *IdP-SP-Proxy* pattern in AARC blueprint architecture terminology [18]), which is gaining adoption for integrating identity federations with research computing infrastructures. Conceptually and practically, IAM replaces VOMS as the VO attribute authority, without being limited to a single authentication mechanism.

IAM provides a registration service that implements a moderated enrollment flow similar to the one used in production by WLCG, with support for periodic Acceptable Usage Policy (AUP) enforcement.

IAM supports account linking, allowing a user to link multiple identities to his VO account. Identities are then mapped to different, configurable level-of-assurance labels, which are exposed to relying services via standard OpenID Connect claims.

The IAM management dashboard, shown in Figure 5, provides an intuitive tool for common administrative tasks, such as group membership and OAuth/OpenID Connect client management.

In order to fully address WLCG use cases and enable a seamless transition from an X.509-based AAI to a token-based one, IAM already implements:

- *VOMS provisioning*: a VOMS endpoint encodes user attributes as a standard VOMS attribute certificate (AC) [5];
- *CERN SSO integration*: this improves the user experience in enrollment and AUP signature management flows;
- *CERN HR database integration*: in order to support LHC VOs identity vetting process, the CERN HR database code currently used in VOMS Admin [19] has been extracted and refactored as a REST API. This REST API is used by a synchronization microservice to keep the VO membership information in IAM consistent with the data coming from the CERN HR database.

3 Related work and initiatives

Token-based authentication and authorization have recently gained much attention in scientific computing.

The Scitokens [20] project proposes a profile for OAuth/JWT to enable capability-based authorization for WLCG computing and is working on integrating support for this profile in existing Grid middleware (e.g., HTCondor, XRootD). The Scitokens approach is based on the same technologies proposed in this paper, but with a stronger focus towards capability-based authorization.

The EOSC-Hub [21] project is promoting harmonization across the main European identity solutions (EGI CheckIn, B2Access, INDIGO IAM, EduTEAMS) also covering token-based authentication and authorization.

AARC [22] and FIM4R [23] are working on recommendations and policy frameworks to support the adoption of federated identity management (FIM) by research communities.

The OpenID Research and Education Working Group [24] is developing a set of profiles to ease the adoption of OpenID Connect in the Research and Education sector.

Finally, the WLCG authorization working group [25] is bringing together all WLCG AAI stakeholders to collect requirements and provide recommendations towards a token-based AAI.

4 Conclusions and future work

Moving beyond X.509 certificates is recognized as a key challenge for HEP computing to improve usability, simplify the middleware stack and enable interoperability with heterogeneous computing and storage resource providers.

In this contribution we have described a token-based AAI that relies on standard authentication and authorization technologies widely adopted in industry, such as OAuth and OpenID Connect.

INDIGO IAM is the identity and authorization hub at the core of this future AAI; yet, it provides support also for existing VOMS-aware services, thus enabling a smooth transition path. It has already been successfully integrated with existing Grid services (e.g., dCache [26], StoRM [27], FTS [28]) and other off-the-shelf components (e.g., Openstack [29], Kubernetes [30]).

The next major step is the integration of IAM with the RCAuth online CA, in order to provide on-demand X.509 user certificates. RCAuth has been chosen since it represents an interesting accredited IOTA CA capable of provisioning X.509 certificates to end-entities based on a successful authentication to a trusted Federated Identity Management System (FIMS) [12], and provides a simple OAuth-based protocol to client applications [31].

References

[1] *The Interoperable Global Trust Federation*, https://www.igtf.net/
[2] *X.509*, https://en.wikipedia.org/wiki/X.509
[3] V. Welch, F. Siebenlist, I.T. Foster, J. Bresnahan, K. Czajkowski, J. Gawor, C. Kesselman, S. Meder, L. Pearlman, S. Tuecke, *Security for grid services*, CoRR **cs.CR/0306129** (2003)
[4] V. Ciaschini, V. Venturi, A. Ceccanti, *The Virtual Organisation Membership Service*, https://doi.org/10.5281/zenodo.1875371
[5] V. Ciaschini, V. Venturi, A. Ceccanti, *The VOMS Attribute Certificate format* , Tech. rep., Open Grid Forum (2011), https://www.ogf.org/documents/GFD.182.pdf
[6] *eduGAIN interfederation website*, http://www.geant.org/Services/Trust_identity_and_security/eduGAIN
[7] HEP Software Foundation, J. Albrecht, A.A.A. Jr, G. Amadio, G. Andronico, N. Anh-Ky, L. Aphecetche, J. Apostolakis, M. Asai, L. Atzori et al., *A Roadmap for HEP Software and Computing R&D for the 2020s* (2017), arXiv:1712.06982
[8] I.D. Collaboration, :, D. Salomoni, I. Campos, L. Gaido, J.M. de Lucas, P. Solagna, J. Gomes, L. Matyska, P. Fuhrman et al., *INDIGO-DataCloud:A data and computing platform to facilitate seamless access to e-infrastructures* (2017), arXiv:1711.01981
[9] M.B. Jones, J. Bradley, N. Sakimura, *The JSON Web Token RFC*, RFC 7519, IETF Tools (2015), https://tools.ietf.org/rfc/rfc7519.txt
[10] D. Hardt, *The OAuth 2.0 Authorization Framework*, RFC 6749, IETF Tools (2012), https://tools.ietf.org/rfc/rfc6749.txt
[11] OpenID Foundation, *The OpenID Connect identity layer* (2018), https://openid.net/connect/
[12] *The RCAuth online CA*, https://rcauth.eu
[13] M. Sallé, *RCAuth.eu: getting proxies using SSH key AuthN*, https://indico.cern.ch/event/669715/contributions/2739035/attachments/1532101/2398499/RCauth_SSH_wlcg_authz_wg.pdf
[14] M.B. Jones, N. Sakimura, J. Bradley, *OAuth 2.0 Authorization Server Metadata*, RFC 8414, IETF Tools (2018), https://tools.ietf.org/rfc/rfc8414.txt
[15] Nat Sakimura and John Bradley and Michael B. Jones and Edmund Jay, *The OpenID Connect discovery specification* (2014), https://openid.net/specs/openid-connect-discovery-1_0.html
[16] M.B. Jones, A. Nadalin, B. Campbell, J. Bradley, C. Mortimore, *OAuth 2.0 Token Exchange*, Internet-Draft draft-ietf-oauth-token-exchange-16.txt (2019)
[17] *Dynamic On Demand Analysis Service: DODAS*, https://dodas-ts.github.io/dodas-doc
[18] *The AARC Blueprint Architecture*, https://aarc-project.eu/architecture
[19] A. Ceccanti, *The VOMS administration service*, https://doi.org/10.5281/zenodo.1875616
[20] *The SciTokens project*, https://scitokens.org
[21] *The EOSC-Hub project*, https://www.eosc-hub.eu
[22] *The AARC project*, https://aarc-project.eu
[23] *Federated Identity Management for Research*, https://fim4r.org
[24] *The OpenID Research & Education working group*, https://openid.net/wg/rande
[25] *The WLCG Authorization Working Group*, https://twiki.cern.ch/twiki/bin/view/LCG/WLCGAuthorizationWG
[26] *The dCache storage solution*, https://dcache.org
[27] *The StoRM storage element*, https://italiangrid.github.io/storm
[28] *The CERN File Transfer Service*, https://fts.web.cern.ch
[29] *The Openstack IAAS framework*, https://www.openstack.org
[30] *The Kubernetes container orchestrator*, https://kubernetes.io
[31] *OAuth for MyProxy*, http://grid.ncsa.illinois.edu/myproxy/oauth/