OAuth Working Group B. Campbell
Internet-Draft J. Bradley
Intended status: Standards Track Ping Identity
Expires: January 9, 2017 M. Jones
Microsoft
July 8, 2016

A Token Binding method for OAuth 2.0 Proof Key for Code Exchange
draft-campbell-oauth-tbpkce-00

Abstract

This specification describes a Proof Key for Code Exchange (PKCE) [RFC7636] method utilizing Token Binding over HTTP [I-D.ietf-tokbind-https] to cryptographically bind the OAuth 2.0 [RFC6749] authorization code to a key pair on the client, which it proves possession of during the access token request with the authorization code.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This specification minimally describes an OAuth 2.0 [RFC6749] PKCE [RFC7636] method based on the Token Binding Protocol [I-D.ietf-tokbind-protocol] and Token Binding over HTTP [I-D.ietf-tokbind-https]. The general details and motivations of PKCE are discussed in that document and this specification defines only the additional pieces needed for a Token Binding PKCE method.

1.1. Requirements Notation and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2. Terminology

This specification uses the terms "authorization code", "authorization endpoint", "authorization server", "authorization request", "access token request", "client", and "token endpoint" defined by OAuth 2.0 [RFC6749], and the terms "Provided", "Token Binding" and "Token Binding ID" defined by Token Binding over HTTP [I-D.ietf-tokbind-https].

2. Code Challenge

As defined in Proof Key for Code Exchange [RFC7636], the client sends the code challenge as part of the OAuth 2.0 Authorization Request with the two additional parameters: "code_challenge" and "code_challenge_method".

For the Token Binding method of PKCE defined herein, "tb2" is used for the value of the "code_challenge_method" parameter.

The value of the "code_challenge" parameter is the base64url encoding (per Section 5 of [RFC4648] with all trailing pad ('=') characters omitted and without the inclusion of any line breaks or whitespace) of the SHA-256 hash [RFC6234] of the Provided Token Binding ID that the client will use when calling the authorization server's token endpoint. Note that, prior to making the authorization request, the client may need to establish a TLS connection between itself and the authorization server's token endpoint in order to obtain the appropriate Token Binding ID.

When the authorization server issues the authorization code in the authorization response, it associates the code challenge and method values with the authorization code so it can be verified later when the code is presented in the access token request.

3. Code Verifier

Upon receipt of the authorization code, the client sends the access token request to the token endpoint. The Token Binding Protocol [I-D.ietf-tokbind-protocol] is negotiated on the TLS connection between the client and the authorization server and the "Sec-Token-Binding" header, as defined in Token Binding over HTTP [I-D.ietf-tokbind-https], is included in the access token request. The authorization server extracts the Provided Token Binding ID from the header value, hashes it with SHA-256, and compares it to the "code_challenge" value previously associated with the authorization code. If the values match, the token endpoint MUST continue processing as normal (as defined by OAuth 2.0 [RFC6749]). If the values do not match, an error response indicating "invalid_grant" MUST be returned.

The "Sec-Token-Binding" header contains sufficient information for verification of the authorization code and its association to the original authorization request. However, PKCE [RFC7636] requires that a "code_verifier" parameter be sent with the access token request, so the static value "provided" is used to meet that requirement and indicate that the Provided Token Binding ID is used for the verification.

4. Security Considerations

TBD

5. IANA Considerations

5.1. PKCE Code Challenge Method Registration

This specification requests registration of the following Code Challenge Method Parameter Name in the IANA "PKCE Code Challenge Methods" registry [IANA.OAuth.Parameters] established by [RFC7636].

5.1.1. Registry Contents

- o Code Challenge Method Parameter Name: tb2
- o Change controller: IESG
- o Specification document(s): Section 2 of [[this specification]]

6. Normative References

[I-D.ietf-tokbind-https]
Popov, A., Nystrom, M., Balfanz, D., Langley, A., and J. Hodges, "Token Binding over HTTP", draft-ietf-tokbind-https-03 (work in progress), March 2016.

[I-D.ietf-tokbind-protocol]
Popov, A., Nystrom, M., Balfanz, D., Langley, A., and J. Hodges, "The Token Binding Protocol Version 1.0", draft-ietf-tokbind-protocol-06 (work in progress), May 2016.

[IANA.OAuth.Parameters]
IANA, "OAuth Parameters", <http://www.iana.org/assignments/oauth-parameters>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, DOI 10.17487/RFC4648, October 2006, <http://www.rfc-editor.org/info/rfc4648>.

[RFC6234] Eastlake 3rd, D. and T. Hansen, "US Secure Hash Algorithms (SHA and SHA-based HMAC and HKDF)", RFC 6234, DOI 10.17487/RFC6234, May 2011, <http://www.rfc-editor.org/info/rfc6234>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <http://www.rfc-editor.org/info/rfc6749>.

[RFC7636] Sakimura, N., Ed., Bradley, J., and N. Agarwal, "Proof Key for Code Exchange by OAuth Public Clients", RFC 7636, DOI 10.17487/RFC7636, September 2015, <http://www.rfc-editor.org/info/rfc7636>.

Appendix A. Acknowledgements

Dirk Balfanz, William Dennis (and others?) also provided input to this specification.

Appendix B. Document History

[[to be removed by the RFC Editor before publication as an RFC]]

draft-campbell-oauth-tbpkce-00

- o Initial version.

Authors' Addresses

Brian Campbell
Ping Identity

Email: brian.d.campbell@gmail.com
URI: https://twitter.com/__b_c

John Bradley
Ping Identity

Email: ve7jtb@ve7jtb.com
URI: http://www.thread-safe.com/

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

OAuth Working Group   M. Jones
Internet-Draft   Microsoft
Intended status: Standards Track   P. Hunt
Expires: September 14, 2017   Oracle
                                                              A. Nadalin
                                                               Microsoft
                                                          March 13, 2017

Authentication Method Reference Values
draft-ietf-oauth-amr-values-08

Abstract

The "amr" (Authentication Methods References) claim is defined and registered in the IANA "JSON Web Token Claims" registry but no standard Authentication Method Reference values are currently defined. This specification establishes a registry for Authentication Method Reference values and defines an initial set of Authentication Method Reference values.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The "amr" (Authentication Methods References) claim is defined and registered in the IANA "JSON Web Token Claims" registry [IANA.JWT.Claims] but no standard Authentication Method Reference values are currently defined. This specification establishes a registry for Authentication Method Reference values and defines an initial set of Authentication Method Reference values.

For context, the "amr" (Authentication Methods References) claim is defined by Section 2 of the OpenID Connect Core 1.0 specification [OpenID.Core] as follows:

amr

OPTIONAL. Authentication Methods References. JSON array of strings that are identifiers for authentication methods used in the authentication. For instance, values might indicate that both password and OTP authentication methods were used. The definition of particular values to be used in the "amr" Claim is beyond the scope of this specification. Parties using this claim will need to agree upon the meanings of the values used, which may be context-specific. The "amr" value is an array of case sensitive strings.

Each "amr" value typically provides an identifier for a family of closely-related authentication methods. For example, the "otp" identifier intentionally covers both time-based and HMAC-based OTPs. Many relying parties will be content to know that an OTP has been used in addition to a password; the distinction between which kind of OTP was used is not useful to them. Thus, there’s a single identifier that can be satisfied in two or more nearly equivalent ways.

Similarly, there’s a whole range of nuances between different fingerprint matching algorithms. They differ in false positive and false negative rates over different population samples and also differ based on the kind and model of fingerprint sensor used. Like the OTP case, many relying parties will be content to know that a fingerprint match mas made, without delving into and differentiating based on every aspect of the implementation of fingerprint capture and match. The "fpt" identifier accomplishes this.

Ultimately, the relying party is depending upon the identity provider to do reasonable things. If it does not trust the identity provider to do so, it has no business using it. The "amr" value lets the identity provider signal to the relying party additional information about what it did, for the cases in which that information is useful to the relying party.

The "amr" values defined by this specification are not intended to be an exhaustive set covering all use cases. Additional values can and will be added to the registry by other specifications. Rather, the values defined herein are an intentionally small set that are already actually being used in practice.

The values defined by this specification only make distinctions that are known to be useful to relying parties. Slicing things more finely than would be used in practice would actually hurt interop, rather than helping it, because it would force relying parties to recognize that several or many different values actually mean the same thing to them.

For context, while the claim values registered pertain to authentication, note that OAuth 2.0 [RFC6749] is designed for resource authorization and cannot be used for authentication without employing appropriate extensions, such as those defined by OpenID Connect Core 1.0 [OpenID.Core]. The existence of the "amr" claim and values for it should not be taken as encouragement to try to use

OAuth 2.0 for authentication without employing extensions enabling secure authentication to be performed.

When used with OpenID Connect, if the identity provider supplies an "amr" claim in the ID Token resulting from a successful authentication, the relying party can inspect the values returned and thereby learn details about how the authentication was performed. For instance, the relying party might learn that only a password was used or it might learn that iris recognition was used in combination with a hardware-secured key. Whether "amr" values are provided and which values are understood by what parties are both beyond the scope of this specification. The OpenID Connect MODRNA Authentication Profile 1.0 [OpenID.MODRNA] is one example of an application context that uses "amr" values defined by this specification.

1.1. Requirements Notation and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2. Terminology

This specification uses the terms defined by JSON Web Token (JWT) [JWT] and OpenID Connect Core 1.0 [OpenID.Core].

2. Authentication Method Reference Values

The following is a list of Authentication Method Reference values defined by this specification:

face
: Biometric authentication [RFC4949] using facial recognition

fpt
: Biometric authentication [RFC4949] using a fingerprint

geo
: Use of geolocation information for authentication, such as that provided by [W3C.REC-geolocation-API-20161108]

hwk
: Proof-of-possession (PoP) of a hardware-secured key. See Appendix C of [RFC4211] for a discussion on PoP.

iris
: Biometric authentication [RFC4949] using an iris scan

kba
   Knowledge-based authentication [NIST.800-63-2] [ISO29115]

mca
   Multiple-channel authentication [MCA]. The authentication involves communication over more than one distinct communication channel. For instance, a multiple-channel authentication might involve both entering information into a workstation's browser and providing information on a telephone call to a pre-registered number.

mfa
   Multiple-factor authentication [NIST.800-63-2] [ISO29115]. When this is present, specific authentication methods used may also be included.

otp
   One-time password [RFC4949]. One-time password specifications that this authentication method applies to include [RFC4226] and [RFC6238].

pin
   Personal Identification Number (PIN) [RFC4949] or pattern (not restricted to containing only numbers) that a user enters to unlock a key on the device. This mechanism should have a way to deter an attacker from obtaining the PIN by trying repeated guesses.

pwd
   Password-based authentication [RFC4949]

rba
   Risk-based authentication [JECM]

retina
   Biometric authentication [RFC4949] using a retina scan

sc
   Smart card [RFC4949]

sms
   Confirmation using SMS [SMS] text message to the user at a registered number

swk
   Proof-of-possession (PoP) of a software-secured key. See Appendix C of [RFC4211] for a discussion on PoP.

tel
Confirmation by telephone call to the user at a registered number. This authentication technique is sometimes also referred to as "call back" [RFC4949].

user
User presence test. Evidence that the end-user is present and interacting with the device. This is sometimes also referred to as "test of user presence" [W3C.WD-webauthn-20170216].

vbm
Biometric authentication [RFC4949] using a voiceprint

wia
Windows integrated authentication [MSDN]

3. Relationship to "acr" (Authentication Context Class Reference)

The "acr" (Authentication Context Class Reference) claim and "acr_values" request parameter are related to the "amr" (Authentication Methods References) claim, but with important differences. An Authentication Context Class specifies a set of business rules that authentications are being requested to satisfy. These rules can often be satisfied by using a number of different specific authentication methods, either singly or in combination. Interactions using "acr_values" request that the specified Authentication Context Classes be used and that the result should contain an "acr" claim saying which Authentication Context Class was satisfied. The "acr" claim in the reply states that the business rules for the class were satisfied -- not how they were satisfied.

In contrast, interactions using the "amr" claim make statements about the particular authentication methods that were used. This tends to be more brittle than using "acr", since the authentication methods that may be appropriate for a given authentication will vary over time, both because of the evolution of attacks on existing methods and the deployment of new authentication methods.

4. Privacy Considerations

The list of "amr" claim values returned in an ID Token reveals information about the way that the end-user authenticated to the identity provider. In some cases, this information may have privacy implications.

While this specification defines identifiers for particular kinds of credentials, it does not define how these credentials are stored or protected. For instance, ensuring the security and privacy of biometric credentials that are referenced by some of the defined Authentication Method Reference values is beyond the scope of this specification.

5. Security Considerations

The security considerations in OpenID Connect Core 1.0 [OpenID.Core] and OAuth 2.0 [RFC6749] and the OAuth 2.0 Threat Model [RFC6819] apply to applications using this specification.

As described in Section 3, taking a dependence upon particular authentication methods may result in brittle systems since the authentication methods that may be appropriate for a given authentication will vary over time.

6. IANA Considerations

6.1. Authentication Method Reference Values Registry

This specification establishes the IANA "Authentication Method Reference Values" registry for "amr" claim array element values. The registry records the Authentication Method Reference value and a reference to the specification that defines it. This specification registers the Authentication Method Reference values defined in Section 2.

Values are registered on an Expert Review [RFC5226] basis after a three-week review period on the jwt-reg-review@ietf.org mailing list, on the advice of one or more Designated Experts. To increase potential interoperability, the experts are requested to encourage registrants to provide the location of a publicly-accessible specification defining the values being registered, so that their intended usage can be more easily understood.

Registration requests sent to the mailing list for review should use an appropriate subject (e.g., "Request to register Authentication Method Reference value: otp").

Within the review period, the Designated Experts will either approve or deny the registration request, communicating this decision to the review list and IANA. Denials should include an explanation and, if applicable, suggestions as to how to make the request successful. Registration requests that are undetermined for a period longer than 21 days can be brought to the IESG’s attention (using the iesg@ietf.org mailing list) for resolution.

IANA must only accept registry updates from the Designated Experts and should direct all requests for registration to the review mailing list.

It is suggested that the same Designated Experts evaluate these registration requests as those who evaluate registration requests for the IANA "JSON Web Token Claims" registry [IANA.JWT.Claims].

Criteria that should be applied by the Designated Experts includes determining whether the proposed registration duplicates existing functionality, whether it is likely to be of general applicability or whether it is useful only for a single application, whether the value is actually being used, and whether the registration description is clear.

6.1.1. Registration Template

Authentication Method Reference Name:
The name requested (e.g., "otp") for the authentication method or family of closely-related authentication methods. Because a core goal of this specification is for the resulting representations to be compact, it is RECOMMENDED that the name be short -- that is, not to exceed 8 characters without a compelling reason to do so. To facilitate interoperability, the name must use only printable ASCII characters excluding double quote ('"') and backslash ('\') (the Unicode characters with code points U+0021, U+0023 through U+005B, and U+005D through U+007E). This name is case sensitive. Names may not match other registered names in a case-insensitive manner unless the Designated Experts state that there is a compelling reason to allow an exception.

Authentication Method Reference Description:
Brief description of the Authentication Method Reference (e.g., "One-time password").

Change Controller:
For Standards Track RFCs, state "IESG". For others, give the name of the responsible party. Other details (e.g., postal address, email address, home page URI) may also be included.

Specification Document(s):
Reference to the document or documents that specify the parameter, preferably including URIs that can be used to retrieve copies of the documents. An indication of the relevant sections may also be included but is not required.

6.1.2. Initial Registry Contents

- Authentication Method Reference Name: "face"
- Authentication Method Reference Description: Facial recognition
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "fpt"
- Authentication Method Reference Description: Fingerprint biometric
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "geo"
- Authentication Method Reference Description: Geolocation
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "hwk"
- Authentication Method Reference Description: Proof-of-possession of a hardware-secured key
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "iris"
- Authentication Method Reference Description: Iris scan biometric
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "kba"
- Authentication Method Reference Description: Knowledge-based authentication
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "mca"
- Authentication Method Reference Description: Multiple-channel authentication
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "mfa"
- Authentication Method Reference Description: Multiple-factor authentication
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "otp"
- Authentication Method Reference Description: One-time password

- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "pin"
- Authentication Method Reference Description: Personal Identification Number or pattern
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "pwd"
- Authentication Method Reference Description: Password-based authentication
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "rba"
- Authentication Method Reference Description: Risk-based authentication
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "retina"
- Authentication Method Reference Description: Retina scan biometric
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "sc"
- Authentication Method Reference Description: Smart card
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "sms"
- Authentication Method Reference Description: Confirmation using SMS
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "swk"
- Authentication Method Reference Description: Proof-of-possession of a software-secured key
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Authentication Method Reference Name: "tel"
- Authentication Method Reference Description: Confirmation by telephone call
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- o Authentication Method Reference Name: "user"
- o Authentication Method Reference Description: User presence test
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Authentication Method Reference Name: "vbm"
- o Authentication Method Reference Description: Voice biometric
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Authentication Method Reference Name: "wia"
- o Authentication Method Reference Description: Windows integrated authentication
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

7. References

7.1. Normative References

[IANA.JWT.Claims]
IANA, "JSON Web Token Claims", <http://www.iana.org/assignments/jwt>.

[JWT] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, May 2015, <http://www.rfc-editor.org/info/rfc7519>.

[OpenID.Core]
Sakimura, N., Bradley, J., Jones, M., de Medeiros, B., and C. Mortimore, "OpenID Connect Core 1.0", November 2014, <http://openid.net/specs/openid-connect-core-1_0.html>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <http://www.rfc-editor.org/info/rfc6749>.

7.2. Informative References

[ISO29115]
International Organization for Standardization, "ISO/IEC 29115:2013 -- Information technology - Security techniques - Entity authentication assurance framework", ISO/IEC 29115:2013, April 2013, <http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45138>.

[JECM] Williamson, G., "Enhanced Authentication In Online Banking", Journal of Economic Crime Management 4.2: 18-19, 2006, <http://utica.edu/academic/institutes/ecii/publications/articles/51D6D996-90F2-F468-AC09C4E8071575AE.pdf>.

[MCA] ldapwiki.com, "Multiple-channel Authentication", August 2016, <https://www.ldapwiki.com/wiki/Multiple-channel%20Authentication>.

[MSDN] Microsoft, "Integrated Windows Authentication with Negotiate", September 2011, <http://blogs.msdn.com/b/benjaminperkins/archive/2011/09/14/iis-integrated-windows-authentication-with-negotiate.aspx>.

[NIST.800-63-2]
National Institute of Standards and Technology (NIST), "Electronic Authentication Guideline", NIST Special Publication 800-63-2, August 2013, <http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-63-2.pdf>.

[OpenID.MODRNA]
Connotte, J. and J. Bradley, "OpenID Connect MODRNA Authentication Profile 1.0", March 2017, <http://openid.net/specs/openid-connect-modrna-authentication-1_0.html>.

[RFC4211] Schaad, J., "Internet X.509 Public Key Infrastructure Certificate Request Message Format (CRMF)", RFC 4211, DOI 10.17487/RFC4211, September 2005, <http://www.rfc-editor.org/info/rfc4211>.

[RFC4226] M'Raihi, D., Bellare, M., Hoornaert, F., Naccache, D., and O. Ranen, "HOTP: An HMAC-Based One-Time Password Algorithm", RFC 4226, DOI 10.17487/RFC4226, December 2005, <http://www.rfc-editor.org/info/rfc4226>.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", FYI 36, RFC 4949, DOI 10.17487/RFC4949, August 2007, <http://www.rfc-editor.org/info/rfc4949>.

[RFC6238] M'Raihi, D., Machani, S., Pei, M., and J. Rydell, "TOTP: Time-Based One-Time Password Algorithm", RFC 6238, DOI 10.17487/RFC6238, May 2011, <http://www.rfc-editor.org/info/rfc6238>.

[RFC6819] Lodderstedt, T., Ed., McGloin, M., and P. Hunt, "OAuth 2.0 Threat Model and Security Considerations", RFC 6819, DOI 10.17487/RFC6819, January 2013, <http://www.rfc-editor.org/info/rfc6819>.

[SMS] 3rd Generation Partnership Project, "Technical realization of the Short Message Service (SMS)", 3GPP Technical Specification (TS) 03.40 V7.5.0 (2001-12), January 2002, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=141>.

[W3C.REC-geolocation-API-20161108]
Popescu, A., "Geolocation API Specification 2nd Edition", World Wide Web Consortium Recommendation REC-geolocation-API-20161108, November 2016, <https://www.w3.org/TR/2016/REC-geolocation-API-20161108>.

[W3C.WD-webauthn-20170216]
Bharadwaj, V., Le Van Gong, H., Balfanz, D., Czeskis, A., Birgisson, A., Hodges, J., Jones, M., Lindemann, R., and J. Jones, "Web Authentication: An API for accessing Scoped Credentials", World Wide Web Consortium Working Draft WD-webauthn-20170216, February 2017, <http://www.w3.org/TR/2017/WD-webauthn-20170216/>.

Appendix A. Examples

In some cases, the "amr" claim value returned may contain a single Authentication Method Reference value. For example, the following "amr" claim value indicates that the authentication performed used an iris scan biometric:

```
"amr": ["iris"]
```

In other cases, the "amr" claim value returned may contain multiple Authentication Method Reference values. For example, the following "amr" claim value indicates that the authentication performed used a password and knowledge-based authentication:

```
"amr": ["pwd", "kba"]
```

Appendix B. Acknowledgements

Caleb Baker participated in specifying the original set of "amr" values. Jari Arkko, John Bradley, Ben Campbell, Brian Campbell, William Denniss, Linda Dunbar, Stephen Farrell, Paul Kyzivat, Elaine Newton, James Manger, Catherine Meadows, Alexey Melnikov, Kathleen Moriarty, Nat Sakimura, and Mike Schwartz provided reviews of the specification.

Appendix C. Document History

[[to be removed by the RFC editor before publication as an RFC]]

-08

- Added text in the IANA Registration Template saying that names can be for families of closely-related authentication methods, as suggested by Stephen Farrell.

-07

- Clarified that the values are intended to provide identifiers for families of closely-related authentication methods.
- Updated the MODRNA Authentication Profile reference.

-06

- Addressed IESG comments. Identifiers are now restricted to using only printable JSON-friendly ASCII characters. Additional references to documentation relevant to specific AMR values were added.

-05

- Specified characters allowed in "amr" values, reusing the IANA Considerations language on this topic from RFC 7638.

-04

- Added examples with single and multiple values.
- Clarified that the actual credentials referenced are not part of this specification to avoid additional privacy concerns for biometric data.
- Clarified that the OAuth 2.0 Threat Model [RFC6819] applies to applications using this specification.

-03

o Addressed shepherd comments.

-02

o Addressed working group last call comments.

-01

o Distinguished between retina and iris biometrics.
o Expanded the introduction to provide additional context to readers.
o Referenced the OpenID Connect MODRNA Authentication Profile 1.0 specification, which uses "amr" values defined by this specification.

-00

o Created the initial working group draft from draft-jones-oauth-amr-values-05 with no normative changes.

Authors' Addresses

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

Phil Hunt
Oracle

Email: phil.hunt@yahoo.com

Anthony Nadalin
Microsoft

Email: tonynad@microsoft.com

OAuth
Internet-Draft
Intended status: Standards Track
Expires: September 12, 2019

W. Denniss
Google
J. Bradley
Ping Identity
M. Jones
Microsoft
H. Tschofenig
ARM Limited
March 11, 2019

OAuth 2.0 Device Authorization Grant
draft-ietf-oauth-device-flow-15

Abstract

The OAuth 2.0 Device Authorization Grant is designed for internet-connected devices that either lack a browser to perform a user-agent based authorization, or are input-constrained to the extent that requiring the user to input text in order to authenticate during the authorization flow is impractical. It enables OAuth clients on such devices (like smart TVs, media consoles, digital picture frames, and printers) to obtain user authorization to access protected resources without using an on-device user-agent.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This OAuth 2.0 [RFC6749] protocol extension, sometimes referred to as "device flow", enables OAuth clients to request user authorization from applications on devices that have limited input capabilities or lack a suitable browser. Such devices include those smart TVs, media console, picture frames and printers which lack an easy input method or suitable browser required for traditional OAuth interactions. The authorization flow defined by this specification instructs the user to review the authorization request on a secondary device, such as a smartphone which does have the requisite input and browser capabilities to complete the user interaction.

The Device Authorization Grant is not intended to replace browser-based OAuth in native apps on capable devices like smartphones. Those apps should follow the practices specified in OAuth 2.0 for Native Apps [RFC8252].

The operating requirements to be able to use this authorization grant type are:

(1) The device is already connected to the Internet.

(2) The device is able to make outbound HTTPS requests.

(3) The device is able to display or otherwise communicate a URI and code sequence to the user.

(4) The user has a secondary device (e.g., personal computer or smartphone) from which they can process the request.

As the device authorization grant does not require two-way communication between the OAuth client and the user-agent (unlike other OAuth 2 grant types such as the Authorization Code and Implicit grant types), it supports several use cases that cannot be served by those other approaches.

Instead of interacting with the end user's user agent, the client instructs the end user to use another computer or device and connect to the authorization server to approve the access request. Since the protocol supports clients that can't receive incoming requests, clients poll the authorization server repeatedly until the end user completes the approval process.

The device typically chooses the set of authorization servers to support (i.e., its own authorization server, or those by providers it has relationships with). It is not uncommon for the device application to support only a single authorization server, such as with a TV application for a specific media provider that supports only that media provider's authorization server. The user may not have an established relationship yet with that authorization provider, though one can potentially be set up during the authorization flow.

```
      +----------+                                +----------------+
      |          |>---(A)-- Client Identifier --->|                |
      |          |                                |                |
      |          |<---(B)-- Device Code,      ---<|                |
      |          |          User Code,            |                |
      |  Device  |          & Verification URI    |                |
      |  Client  |                                |                |
      |          |  [polling]                     |                |
      |          |>---(E)-- Device Code,      --->|                |
      |          |          & Client Identifier   |                |
      |          |                                |  Authorization |
      |          |<---(F)-- Access Token      ---<|     Server     |
      +----------+   (& Optional Refresh Token)   |                |
            v                                     |                |
            :                                     |                |
           (C) User Code & Verification URI       |                |
            :                                     |                |
            v                                     |                |
      +----------+                                |                |
      | End user |                                |                |
      |    at    |<---(D)-- End user reviews  --->|                |
      |  Browser |          authorization request |                |
      +----------+                                +----------------+
```

Figure 1: Device Authorization Flow

The device authorization flow illustrated in Figure 1 includes the following steps:

(A) The client requests access from the authorization server and includes its client identifier in the request.

(B) The authorization server issues a device code, an end-user code, and provides the end-user verification URI.

(C) The client instructs the end user to use its user agent (on another device) and visit the provided end-user verification URI. The client provides the user with the end-user code to enter in order to review the authorization request.

(D) The authorization server authenticates the end user (via the user agent) and prompts the user to grant the client's access request. If the user agrees to the client's access request, the user enters the user code provided by the client. The authorization server validates the user code provided by the user.

(E) While the end user reviews the client's request (step D), the client repeatedly polls the authorization server to find out if the user completed the user authorization step. The client includes the verification code and its client identifier.

(F) The authorization server validates the verification code provided by the client and responds back with the access token if the user granted access, an error if they denied access, or indicates that the client should continue to poll.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Device Authorization Endpoint:
The authorization server's endpoint capable of issuing device verification codes, user codes, and verification URLs.

Device Verification Code:
A short-lived token representing an authorization session.

End-User Verification Code:
A short-lived token which the device displays to the end user, is entered by the user on the authorization server, and is thus used to bind the device to the user.

3. Protocol

3.1. Device Authorization Request

This specification defines a new OAuth endpoint, the device authorization endpoint. This is separate from the OAuth authorization endpoint defined in [RFC6749] with which the user interacts with via a user-agent (i.e., a browser). By comparison, when using the device authorization endpoint, the OAuth client on the device interacts with the authorization server directly without presenting the request in a user-agent, and the end user authorizes the request on a separate device. This interaction is defined as follows.

The client initiates the authorization flow by requesting a set of verification codes from the authorization server by making an HTTP "POST" request to the device authorization endpoint.

The client constructs the request with the following parameters, sent as the body of the request, encoded with the "application/x-www-form-urlencoded" encoding algorithm defined by Section 4.10.22.6 of [HTML5]:

client_id
   REQUIRED, if the client is not authenticating with the authorization server as described in Section 3.2.1. of [RFC6749]. The client identifier as described in Section 2.2 of [RFC6749].

scope
   OPTIONAL. The scope of the access request as described by Section 3.3 of [RFC6749].

For example, the client makes the following HTTPS request:

```
POST /device_authorization HTTP/1.1
Host: server.example.com
Content-Type: application/x-www-form-urlencoded

client_id=459691054427
```

All requests from the device MUST use the Transport Layer Security (TLS) [RFC8446] protocol and implement the best practices of BCP 195 [RFC7525].

Parameters sent without a value MUST be treated as if they were omitted from the request. The authorization server MUST ignore unrecognized request parameters. Request and response parameters MUST NOT be included more than once.

The client authentication requirements of Section 3.2.1 of [RFC6749] apply to requests on this endpoint, which means that confidential clients (those that have established client credentials) authenticate in the same manner as when making requests to the token endpoint, and public clients provide the "client_id" parameter to identify themselves.

Due to the polling nature of this protocol (as specified in Section 3.4), care is needed to avoid overloading the capacity of the token endpoint. To avoid unneeded requests on the token endpoint, the client SHOULD only commence a device authorization request when prompted by the user, and not automatically, such as when the app starts or when the previous authorization session expires or fails.

3.2. Device Authorization Response

In response, the authorization server generates a unique device verification code and an end-user code that are valid for a limited time and includes them in the HTTP response body using the "application/json" format [RFC8259] with a 200 (OK) status code. The response contains the following parameters:

device_code
: REQUIRED. The device verification code.

user_code
: REQUIRED. The end-user verification code.

verification_uri
: REQUIRED. The end-user verification URI on the authorization server. The URI should be short and easy to remember as end users will be asked to manually type it into their user-agent.

verification_uri_complete
: OPTIONAL. A verification URI that includes the "user_code" (or other information with the same function as the "user_code"), designed for non-textual transmission.

expires_in
: REQUIRED. The lifetime in seconds of the "device_code" and "user_code".

interval
: OPTIONAL. The minimum amount of time in seconds that the client SHOULD wait between polling requests to the token endpoint. If no value is provided, clients MUST use 5 as the default.

For example:

```
HTTP/1.1 200 OK
Content-Type: application/json
Cache-Control: no-store

{
  "device_code": "GmRhmhcxhwAzkoEqiMEg_DnyEysNkuNhszIySk9eS",
  "user_code": "WDJB-MJHT",
  "verification_uri": "https://example.com/device",
  "verification_uri_complete":
      "https://example.com/device?user_code=WDJB-MJHT",
  "expires_in": 1800,
  "interval": 5
}
```

In the event of an error (such as an invalidly configured client), the authorization server responds in the same way as the token endpoint specified in Section 5.2 of [RFC6749].

3.3. User Interaction

After receiving a successful Authorization Response, the client displays or otherwise communicates the "user_code" and the "verification_uri" to the end user and instructs them to visit the URI in a user agent on a secondary device (for example, in a browser on their mobile phone), and enter the user code.

```
+-----------------------------------------------+
|                                               |
|  Using a browser on another device, visit:    |
|  https://example.com/device                   |
|                                               |
|  And enter the code:                          |
|  WDJB-MJHT                                    |
|                                               |
+-----------------------------------------------+
```

Figure 2: Example User Instruction

The authorizing user navigates to the "verification_uri" and authenticates with the authorization server in a secure TLS-protected ([RFC8446]) session. The authorization server prompts the end user to identify the device authorization session by entering the "user_code" provided by the client. The authorization server should then inform the user about the action they are undertaking and ask them to approve or deny the request. Once the user interaction is complete, the server MAY inform the user to return to their device.

During the user interaction, the device continuously polls the token endpoint with the "device_code", as detailed in Section 3.4, until the user completes the interaction, the code expires, or another error occurs. The "device_code" is not intended for the end user directly, and thus should not be displayed during the interaction to avoid confusing the end user.

Authorization servers supporting this specification MUST implement a user interaction sequence that starts with the user navigating to "verification_uri" and continues with them supplying the "user_code" at some stage during the interaction. Other than that, the exact sequence and implementation of the user interaction is up to the authorization server, for example, the authorization server may enable new users to sign up for an account during the authorization flow, or add additional security verification steps.

It is NOT RECOMMENDED for authorization servers to include the user code in the verification URI ("verification_uri"), as this increases the length and complexity of the URI that the user must type. While the user must still type the same number of characters with the "user_code" separated, once they successfully navigate to the "verification_uri", any errors in entering the code can be highlighted by the authorization server to improve the user experience. The next section documents user interaction with "verification_uri_complete", which is designed to carry both pieces of information.

3.3.1. Non-textual Verification URI Optimization

When "verification_uri_complete" is included in the Authorization Response (Section 3.2), clients MAY present this URI in a non-textual manner using any method that results in the browser being opened with the URI, such as with QR (Quick Response) codes or NFC (Near Field Communication), to save the user typing the URI.

For usability reasons, it is RECOMMENDED for clients to still display the textual verification URI ("verification_uri") for users not able to use such a shortcut. Clients MUST still display the "user_code", as the authorization server will require the user to confirm it to disambiguate devices, or as a remote phishing mitigation (See Section 5.4).

If the user starts the user interaction by browsing to "verification_uri_complete", then the user interaction described in Section 3.3 is still followed, but with the optimization that the user does not need to type the "user_code". The server SHOULD display the "user_code" to the user and ask them to verify that it matches the "user_code" being displayed on the device, to confirm they are authorizing the correct device. As before, in addition to taking steps to confirm the identity of the device, the user should also be afforded the choice to approve or deny the authorization request.

```
                  +-------------------------------------------------+
                  |                                                 |
                  |  Scan the QR code, or using     +------------+  |
                  |  a browser on another device,   |[_]..  . [_]|  |
                  |  visit:                         | .  ..   . .|  |
                  |  https://example.com/device     | . .  . ....|  |
                  |                                 |.   . . .   |  |
                  |  And enter the code:            |[_]. ... .  |  |
                  |  WDJB-MJHT                      +------------+  |
                  |                                                 |
                  +-------------------------------------------------+
```

Figure 3: Example User Instruction with QR Code Representation of the Complete Verification URI

3.4. Device Access Token Request

After displaying instructions to the user, the client makes an Access Token Request to the token endpoint (as defined by Section 3.2 of [RFC6749]) with a "grant_type" of "urn:ietf:params:oauth:grant-type:device_code". This is an extension grant type (as defined by Section 4.5 of [RFC6749]) created by this specification, with the following parameters:

grant_type
   REQUIRED. Value MUST be set to "urn:ietf:params:oauth:grant-type:device_code".

device_code
   REQUIRED. The device verification code, "device_code" from the Device Authorization Response, defined in Section 3.2.

client_id
   REQUIRED, if the client is not authenticating with the authorization server as described in Section 3.2.1. of [RFC6749]. The client identifier as described in Section 2.2 of [RFC6749].

For example, the client makes the following HTTPS request (line breaks are for display purposes only):

```
POST /token HTTP/1.1
Host: server.example.com
Content-Type: application/x-www-form-urlencoded

grant_type=urn%3Aietf%3Aparams%3Aoauth%3Agrant-type%3Adevice_code
&device_code=GmRhmhcxhwAzkoEqiMEg_DnyEysNkuNhszIySk9eS
&client_id=459691054427
```

If the client was issued client credentials (or assigned other authentication requirements), the client MUST authenticate with the authorization server as described in Section 3.2.1 of [RFC6749]. Note that there are security implications of statically distributed client credentials, see Section 5.6.

The response to this request is defined in Section 3.5. Unlike other OAuth grant types, it is expected for the client to try the Access Token Request repeatedly in a polling fashion, based on the error code in the response.

3.5. Device Access Token Response

If the user has approved the grant, the token endpoint responds with a success response defined in Section 5.1 of [RFC6749]; otherwise it responds with an error, as defined in Section 5.2 of [RFC6749].

In addition to the error codes defined in Section 5.2 of [RFC6749], the following error codes are specified for use with the device authorization grant in token endpoint responses:

authorization_pending
: The authorization request is still pending as the end user hasn't yet completed the user interaction steps (Section 3.3). The client SHOULD repeat the Access Token Request to the token endpoint (a process known as polling). Before each new request the client MUST wait at least the number of seconds specified by the "interval" parameter of the Device Authorization Response (see Section 3.2), or 5 seconds if none was provided, and respect any increase in the polling interval required by the "slow_down" error.

slow_down
: A variant of "authorization_pending", the authorization request is still pending and polling should continue, but the interval MUST be increased by 5 seconds for this and all subsequent requests.

access_denied
: The end user denied the authorization request.

expired_token
: The "device_code" has expired and the device authorization session has concluded. The client MAY commence a new Device Authorization Request but SHOULD wait for user interaction before restarting to avoid unnecessary polling.

The "authorization_pending" and "slow_down" error codes define particularly unique behavior, as they indicate that the OAuth client should continue to poll the token endpoint by repeating the token request (implementing the precise behavior defined above). If the client receives an error response with any other error code, it MUST stop polling and SHOULD react accordingly, for example, by displaying an error to the user.

On encountering a connection timeout, clients MUST unilaterally reduce their polling frequency before retrying. The use of an exponential backoff algorithm to achieve this, such as by doubling the polling interval on each such connection timeout, is RECOMMENDED.

The assumption of this specification is that the separate device the user is authorizing the request on does not have a way to communicate back to device with the OAuth client. This protocol only requires a one-way channel in order to maximise the viability of the protocol in restricted environments, like an application running on a TV that is only capable of outbound requests. If a return channel were to exist for the chosen user interaction interface, then the device MAY wait until notified on that channel that the user has completed the action before initiating the token request (as an alternative to polling). Such behavior is, however, outside the scope of this specification.

4. Discovery Metadata

Support for this specification MAY be declared in the OAuth 2.0 Authorization Server Metadata [RFC8414] by including the value "urn:ietf:params:oauth:grant-type:device_code" in the "grant_types_supported" parameter, and by adding the following new parameter:

device_authorization_endpoint
: OPTIONAL. URL of the authorization server's device authorization endpoint defined in Section 3.1.

5. Security Considerations

5.1. User Code Brute Forcing

Since the user code is typed by the user, shorter codes are more desirable for usability reasons. This means the entropy is typically less than would be used for the device code or other OAuth bearer token types where the code length does not impact usability. It is therefore recommended that the server rate-limit user code attempts.

The user code SHOULD have enough entropy that when combined with rate limiting and other mitigations makes a brute-force attack infeasible. For example, it's generally held that 128-bit symmetric keys for encryption are seen as good enough today because an attacker has to put in 2^96 work to have a 2^-32 chance of guessing correctly via brute force. The rate limiting and finite lifetime on the user code places an artificial limit on the amount of work an attacker can "do", so if, for instance, one uses a 8-character base-20 user code (with roughly 34.5 bits of entropy), the rate-limiting interval and validity period would need to only allow 5 attempts in order to get the same 2^-32 probability of success by random guessing.

A successful brute forcing of the user code would enable the attacker to authenticate with their own credentials and make an authorization grant to the device. This is the opposite scenario to an OAuth bearer token being brute forced, whereby the attacker gains control of the victim's authorization grant. Such attacks may not always make economic sense, for example for a video app the device owner may then be able to purchase movies using the attacker's account, though a privacy risk would still remain and thus is important to protect against. Furthermore, some uses of the device flow give the granting account the ability to perform actions such as controlling the device, which needs to be protected.

The precise length of the user code and the entropy contained within is at the discretion of the authorization server, which needs to consider the sensitivity of their specific protected resources, the practicality of the code length from a usability standpoint, and any mitigations that are in place such as rate-limiting, when determining the user code format.

5.2. Device Code Brute Forcing

An attacker who guesses the device code would be able to potentially obtain the authorization code once the user completes the flow. As the device code is not displayed to the user and thus there are no usability considerations on the length, a very high entropy code SHOULD be used.

5.3. Device Trustworthiness

Unlike other native application OAuth 2.0 flows, the device requesting the authorization is not the same as the device that the user grants access from. Thus, signals from the approving user's session and device are not relevant to the trustworthiness of the client device.

Note that if an authorization server used with this flow is malicious, then it could man-in-the-middle the backchannel flow to another authorization server. In this scenario, the man-in-the-middle is not completely hidden from sight, as the end user would end up on the authorization page of the wrong service, giving them an opportunity to notice that the URL in the browser's address bar is wrong. For this to be possible, the device manufacturer must either directly be the attacker, shipping a device intended to perform the man-in-the-middle attack, or be using an authorization server that is controlled by an attacker, possibly because the attacker compromised the authorization server used by the device. In part, the person purchasing the device is counting on it and its business partners to be trustworthy.

5.4. Remote Phishing

It is possible for the device flow to be initiated on a device in an attacker's possession. For example, an attacker might send an email instructing the target user to visit the verification URL and enter the user code. To mitigate such an attack, it is RECOMMENDED to inform the user that they are authorizing a device during the user interaction step (see Section 3.3), and to confirm that the device is in their possession. The authorization server SHOULD display information about the device so that the person can notice if a software client was attempting to impersonating a hardware device.

For authorization servers that support the option specified in Section 3.3.1 for the client to append the user code to the authorization URI, it is particularly important to confirm that the device is in the user's possession, as the user no longer has to type the code manually. One possibility is to display the code during the authorization flow and asking the user to verify that the same code is being displayed on the device they are setting up.

The user code needs to have a long enough lifetime to be useable (allowing the user to retrieve their secondary device, navigate to the verification URI, login, etc.), but should be sufficiently short to limit the usability of a code obtained for phishing. This doesn't prevent a phisher presenting a fresh token, particularly in the case they are interacting with the user in real time, but it does limit the viability of codes sent over email or SMS.

5.5. Session Spying

While the device is pending authorization, it may be possible for a malicious user to physically spy on the device user interface (by viewing the screen on which it's displayed, for example) and hijack the session by completing the authorization faster than the user that initiated it. Devices SHOULD take into account the operating environment when considering how to communicate the code to the user to reduce the chances it will be observed by a malicious user.

5.6. Non-confidential Clients

Device clients are generally incapable of maintaining the confidentiality of their credentials, as users in possession of the device can reverse engineer it and extract the credentials. Therefore, unless additional measures are taken, they should be treated as public clients (as defined by Section 2.1 of OAuth 2.0) susceptible to impersonation. The security considerations of Section 5.3.1 of [RFC6819] and Sections 8.5 and 8.6 of [RFC8252] apply to such clients.

The user may also be able to obtain the device_code and/or other OAuth bearer tokens issued to their client, which would allow them to use their own authorization grant directly by impersonating the client. Given that the user in possession of the client credentials can already impersonate the client and create a new authorization grant (with a new device_code), this doesn't represent a separate impersonation vector.

5.7. Non-Visual Code Transmission

There is no requirement that the user code be displayed by the device visually. Other methods of one-way communication can potentially be used, such as text-to-speech audio, or Bluetooth Low Energy. To mitigate an attack in which a malicious user can bootstrap their credentials on a device not in their control, it is RECOMMENDED that any chosen communication channel only be accessible by people in close proximity. E.g., users who can see, or hear the device.

6. Usability Considerations

This section is a non-normative discussion of usability considerations.

6.1. User Code Recommendations

For many users, their nearest Internet-connected device will be their mobile phone, and typically these devices offer input methods that are more time consuming than a computer keyboard to change the case or input numbers. To improve usability (improving entry speed, and reducing retries), these limitations should be taken into account when selecting the user-code character set.

One way to improve input speed is to restrict the character set to case-insensitive A-Z characters, with no digits. These characters can typically be entered on a mobile keyboard without using modifier keys. Further removing vowels to avoid randomly creating words results in the base-20 character set: "BCDFGHJKLMNPQRSTVWXZ". Dashes or other punctuation may be included for readability.

An example user code following this guideline containing 8 significant characters and dashes added for end-user readability, with a resulting entropy of 20^8: "WDJB-MJHT".

Pure numeric codes are also a good choice for usability, especially for clients targeting locales where A-Z character keyboards are not used, though their length needs to be longer to maintain a high entropy.

An example numeric user code containing 9 significant digits and dashes added for end-user readability, with an entropy of 10^9: "019-450-730".

When processing the inputted user code, the server should strip dashes and other punctuation it added for readability (making the inclusion of that punctuation by the user optional). For codes using only characters in the A-Z range as with the base-20 charset defined above, the user's input should be upper-cased before comparison to account for the fact that the user may input the equivalent lower-case characters. Further stripping of all characters outside the user_code charset is recommended to reduce instances where an errantly typed character (like a space character) invalidates otherwise valid input.

It is RECOMMENDED to avoid character sets that contain two or more characters that can easily be confused with each other like "0" and "O", or "1", "l" and "I". Furthermore, the extent practical, where a character set contains one character that may be confused with characters outside the character set the character outside the set MAY be substituted with the one in the character set that it is commonly confused with (for example, "O" for "0" when using a numerical 0-9 character set).

6.2. Non-Browser User Interaction

Devices and authorization servers MAY negotiate an alternative code transmission and user interaction method in addition to the one described in Section 3.3. Such an alternative user interaction flow could obviate the need for a browser and manual input of the code, for example, by using Bluetooth to transmit the code to the authorization server's companion app. Such interaction methods can utilize this protocol, as ultimately, the user just needs to identify the authorization session to the authorization server; however, user interaction other than via the verification URI is outside the scope of this specification.

7. IANA Considerations

7.1. OAuth Parameters Registration

This specification registers the following values in the IANA "OAuth Parameters" registry [IANA.OAuth.Parameters] established by [RFC6749].

7.1.1. Registry Contents

- o Parameter name: device_code
- o Parameter usage location: token request
- o Change controller: IESG
- o Specification Document: Section 3.1 of [[this specification]]

7.2. OAuth URI Registration

This specification registers the following values in the IANA "OAuth URI" registry [IANA.OAuth.Parameters] established by [RFC6755].

7.2.1. Registry Contents

- o URN: urn:ietf:params:oauth:grant-type:device_code
- o Common Name: Device flow grant type for OAuth 2.0
- o Change controller: IESG
- o Specification Document: Section 3.1 of [[this specification]]

7.3. OAuth Extensions Error Registration

This specification registers the following values in the IANA "OAuth Extensions Error Registry" registry [IANA.OAuth.Parameters] established by [RFC6749].

7.3.1. Registry Contents

- o Error name: authorization_pending
- o Error usage location: Token endpoint response
- o Related protocol extension: [[this specification]]
- o Change controller: IETF
- o Specification Document: Section 3.5 of [[this specification]]

- o Error name: access_denied
- o Error usage location: Token endpoint response
- o Related protocol extension: [[this specification]]
- o Change controller: IETF
- o Specification Document: Section 3.5 of [[this specification]]

- o Error name: slow_down
- o Error usage location: Token endpoint response
- o Related protocol extension: [[this specification]]
- o Change controller: IETF
- o Specification Document: Section 3.5 of [[this specification]]

- o Error name: expired_token
- o Error usage location: Token endpoint response
- o Related protocol extension: [[this specification]]
- o Change controller: IETF
- o Specification Document: Section 3.5 of [[this specification]]

7.4. OAuth 2.0 Authorization Server Metadata

This specification registers the following values in the IANA "OAuth 2.0 Authorization Server Metadata" registry [IANA.OAuth.Parameters] established by [RFC8414].

7.4.1. Registry Contents

- o Metadata name: device_authorization_endpoint
- o Metadata Description: The Device Authorization Endpoint.
- o Change controller: IESG
- o Specification Document: Section 4 of [[this specification]]

8. Normative References

[HTML5] IANA, "HTML5", <https://www.w3.org/TR/2014/REC-html5-20141028/>.

[IANA.OAuth.Parameters] IANA, "OAuth Parameters", <http://www.iana.org/assignments/oauth-parameters>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <https://www.rfc-editor.org/info/rfc6749>.

[RFC6755] Campbell, B. and H. Tschofenig, "An IETF URN Sub-Namespace for OAuth", RFC 6755, DOI 10.17487/RFC6755, October 2012, <https://www.rfc-editor.org/info/rfc6755>.

[RFC6819] Lodderstedt, T., Ed., McGloin, M., and P. Hunt, "OAuth 2.0 Threat Model and Security Considerations", RFC 6819, DOI 10.17487/RFC6819, January 2013, <https://www.rfc-editor.org/info/rfc6819>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <https://www.rfc-editor.org/info/rfc7525>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8252] Denniss, W. and J. Bradley, "OAuth 2.0 for Native Apps", BCP 212, RFC 8252, DOI 10.17487/RFC8252, October 2017, <https://www.rfc-editor.org/info/rfc8252>.

[RFC8259] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", STD 90, RFC 8259, DOI 10.17487/RFC8259, December 2017, <https://www.rfc-editor.org/info/rfc8259>.

[RFC8414] Jones, M., Sakimura, N., and J. Bradley, "OAuth 2.0 Authorization Server Metadata", RFC 8414, DOI 10.17487/RFC8414, June 2018, <https://www.rfc-editor.org/info/rfc8414>.

[RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/info/rfc8446>.

Appendix A. Acknowledgements

The starting point for this document was the Internet-Draft draft-recordon-oauth-v2-device, authored by David Recordon and Brent Goldman, which itself was based on content in draft versions of the OAuth 2.0 protocol specification removed prior to publication due to a then lack of sufficient deployment expertise. Thank you to the OAuth working group members who contributed to those earlier drafts.

This document was produced in the OAuth working group under the chairpersonship of Rifaat Shekh-Yusef and Hannes Tschofenig with Benjamin Kaduk, Kathleen Moriarty, and Eric Rescorla serving as Security Area Directors.

The following individuals contributed ideas, feedback, and wording that shaped and formed the final specification:

Alissa Cooper, Ben Campbell, Brian Campbell, Roshni Chandrashekhar, Eric Fazendin, Benjamin Kaduk, Jamshid Khosravian, Torsten Lodderstedt, James Manger, Dan McNulty, Breno de Medeiros, Simon Moffatt, Stein Myrseth, Emond Papegaaij, Justin Richer, Adam Roach, Nat Sakimura, Andrew Sciberras, Marius Scurtescu, Filip Skokan, Ken Wang, and Steven E. Wright.

Appendix B. Document History

[[to be removed by the RFC Editor before publication as an RFC]]

-15

- Renamed and dropped most usage of the term "flow"
- Documented error responses on the authorization endpoint
- Documented client authentication for the authorization endpoint

-14

- Added more normative text on polling behavior.
- Added discussion on risk of user retrieving their own device_code.
- Editorial improvements.

-13

- Added a longer discussion about entropy, proposed by Benjamin Kaduk.
- Added device_code to OAuth IANA registry.
- Expanded explanation of "case insensitive".
- Added security section on Device Code Brute Forcing.
- application/x-www-form-urlencoded normativly referenced.
- Editorial improvements.

-12

- Set a default polling interval to 5s explicitly.

- o Defined the slow_down behavior that it should increase the current interval by 5s.
- o expires_in now REQUIRED
- o Other changes in response to review feedback.

-11

- o Updated reference to OAuth 2.0 Authorization Server Metadata.

-10

- o Added a missing definition of access_denied for use on the token endpoint.
- o Corrected text documenting which error code should be returned for expired tokens (it's "expired_token", not "invalid_grant").
- o Corrected section reference to RFC 8252 (the section numbers had changed after the initial reference was made).
- o Fixed line length of one diagram (was causing xml2rfc warnings).
- o Added line breaks so the URN grant_type is presented on an unbroken line.
- o Typos fixed and other stylistic improvements.

-09

- o Addressed review comments by Security Area Director Eric Rescorla about the potential of a confused deputy attack.

-08

- o Expanded the User Code Brute Forcing section to include more detail on this attack.

-07

- o Replaced the "user_code" URI parameter optimization with verification_uri_complete following the IETF99 working group discussion.
- o Added security consideration about spying.
- o Required that device_code not be shown.
- o Added text regarding a minimum polling interval.

-06

- o Clarified usage of the "user_code" URI parameter optimization following the IETF98 working group discussion.

-05

- o response_type parameter removed from authorization request.
- o Added option for clients to include the user_code on the verification URI.
- o Clarified token expiry, and other nits.

-04

- o Security & Usability sections. OAuth Discovery Metadata.

-03

- o device_code is now a URN. Added IANA Considerations

-02

- o Added token request & response specification.

-01

- o Applied spelling and grammar corrections and added the Document History appendix.

-00

- o Initial working group draft based on draft-recordon-oauth-v2-device.

Authors' Addresses

William Denniss
Google
1600 Amphitheatre Pkwy
Mountain View, CA 94043
USA

Email: wdenniss@google.com
URI: http://wdenniss.com/device-flow

John Bradley
Ping Identity

Email: ve7jtb@ve7jtb.com
URI: http://www.thread-safe.com/

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

Hannes Tschofenig
ARM Limited
Austria

Email: Hannes.Tschofenig@gmx.net
URI: http://www.tschofenig.priv.at

OAuth Working Group M. Jones
Internet-Draft Microsoft
Intended status: Standards Track N. Sakimura
Expires: September 5, 2018 NRI
J. Bradley
Ping Identity
March 4, 2018

OAuth 2.0 Authorization Server Metadata
draft-ietf-oauth-discovery-10

Abstract

This specification defines a metadata format that an OAuth 2.0 client can use to obtain the information needed to interact with an OAuth 2.0 authorization server, including its endpoint locations and authorization server capabilities.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This specification generalizes the metadata format defined by "OpenID Connect Discovery 1.0" [OpenID.Discovery] in a way that is compatible with OpenID Connect Discovery, while being applicable to a wider set of OAuth 2.0 use cases. This is intentionally parallel to the way that the "OAuth 2.0 Dynamic Client Registration Protocol" [RFC7591] specification generalized the dynamic client registration mechanisms defined by "OpenID Connect Dynamic Client Registration 1.0" [OpenID.Registration] in a way that was compatible with it.

The metadata for an authorization server is retrieved from a well-known location as a JSON [RFC7159] document, which declares its endpoint locations and authorization server capabilities. This process is described in Section 3.

This metadata can either be communicated in a self-asserted fashion by the server origin via HTTPS or as a set of signed metadata values represented as claims in a JSON Web Token (JWT) [JWT]. In the JWT case, the issuer is vouching for the validity of the data about the authorization server. This is analogous to the role that the Software Statement plays in OAuth Dynamic Client Registration [RFC7591].

The means by which the client chooses an authorization server is out of scope. In some cases, its issuer identifier may be manually configured into the client. In other cases, it may be dynamically discovered, for instance, through the use of WebFinger [RFC7033], as described in Section 2 of "OpenID Connect Discovery 1.0" [OpenID.Discovery].

1.1. Requirements Notation and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

All uses of JSON Web Signature (JWS) [JWS] and JSON Web Encryption (JWE) [JWE] data structures in this specification utilize the JWS Compact Serialization or the JWE Compact Serialization; the JWS JSON Serialization and the JWE JSON Serialization are not used.

1.2. Terminology

This specification uses the terms "Access Token", "Authorization Code", "Authorization Endpoint", "Authorization Grant", "Authorization Server", "Client", "Client Authentication", "Client Identifier", "Client Secret", "Grant Type", "Protected Resource", "Redirection URI", "Refresh Token", "Resource Owner", "Resource Server", "Response Type", and "Token Endpoint" defined by OAuth 2.0 [RFC6749], the terms "Claim Name", "Claim Value", and "JSON Web Token (JWT)" defined by JSON Web Token (JWT) [JWT], and the term "Response Mode" defined by OAuth 2.0 Multiple Response Type Encoding Practices [OAuth.Responses].

2. Authorization Server Metadata

Authorization servers can have metadata describing their configuration. The following authorization server metadata values are used by this specification and are registered in the IANA "OAuth Authorization Server Metadata" registry established in Section 7.1:

issuer
: REQUIRED. The authorization server's issuer identifier, which is a URL that uses the "https" scheme and has no query or fragment components. Authorization server metadata is published at a ".well-known" RFC 5785 [RFC5785] location derived from this issuer identifier, as described in Section 3. The issuer identifier is used to prevent authorization server mix-up attacks, as described in "OAuth 2.0 Mix-Up Mitigation" [I-D.ietf-oauth-mix-up-mitigation].

authorization_endpoint
: URL of the authorization server's authorization endpoint [RFC6749]. This is REQUIRED unless no grant types are supported that use the authorization endpoint.

token_endpoint
: URL of the authorization server's token endpoint [RFC6749]. This is REQUIRED unless only the implicit grant type is supported.

jwks_uri
: OPTIONAL. URL of the authorization server's JWK Set [JWK] document. The referenced document contains the signing key(s) the client uses to validate signatures from the authorization server. This URL MUST use the "https" scheme. The JWK Set MAY also contain the server's encryption key(s), which are used by clients to encrypt requests to the server. When both signing and encryption keys are made available, a "use" (public key use) parameter value is REQUIRED for all keys in the referenced JWK Set to indicate each key's intended usage.

registration_endpoint
: OPTIONAL. URL of the authorization server's OAuth 2.0 Dynamic Client Registration endpoint [RFC7591].

scopes_supported
: RECOMMENDED. JSON array containing a list of the OAuth 2.0 [RFC6749] "scope" values that this authorization server supports. Servers MAY choose not to advertise some supported scope values even when this parameter is used.

response_types_supported

REQUIRED. JSON array containing a list of the OAuth 2.0 "response_type" values that this authorization server supports. The array values used are the same as those used with the "response_types" parameter defined by "OAuth 2.0 Dynamic Client Registration Protocol" [RFC7591].

response_modes_supported
OPTIONAL. JSON array containing a list of the OAuth 2.0 "response_mode" values that this authorization server supports, as specified in OAuth 2.0 Multiple Response Type Encoding Practices [OAuth.Responses]. If omitted, the default is "["query", "fragment"]". The response mode value "form_post" is also defined in OAuth 2.0 Form Post Response Mode [OAuth.Post].

grant_types_supported
OPTIONAL. JSON array containing a list of the OAuth 2.0 grant type values that this authorization server supports. The array values used are the same as those used with the "grant_types" parameter defined by "OAuth 2.0 Dynamic Client Registration Protocol" [RFC7591]. If omitted, the default value is "["authorization_code", "implicit"]".

token_endpoint_auth_methods_supported
OPTIONAL. JSON array containing a list of client authentication methods supported by this token endpoint. Client authentication method values are used in the "token_endpoint_auth_method" parameter defined in Section 2 of [RFC7591]. If omitted, the default is "client_secret_basic" -- the HTTP Basic Authentication Scheme specified in Section 2.3.1 of OAuth 2.0 [RFC6749].

token_endpoint_auth_signing_alg_values_supported
OPTIONAL. JSON array containing a list of the JWS signing algorithms ("alg" values) supported by the token endpoint for the signature on the JWT [JWT] used to authenticate the client at the token endpoint for the "private_key_jwt" and "client_secret_jwt" authentication methods. This metadata entry MUST be present if either of these authentication methods are specified in the "token_endpoint_auth_methods_supported" entry. No default algorithms are implied if this entry is omitted. Servers SHOULD support "RS256". The value "none" MUST NOT be used.

service_documentation
OPTIONAL. URL of a page containing human-readable information that developers might want or need to know when using the authorization server. In particular, if the authorization server does not support Dynamic Client Registration, then information on how to register clients needs to be provided in this documentation.

ui_locales_supported
OPTIONAL. Languages and scripts supported for the user interface, represented as a JSON array of BCP47 [RFC5646] language tag values. If omitted, the set of supported languages and scripts is unspecified.

op_policy_uri
OPTIONAL. URL that the authorization server provides to the person registering the client to read about the authorization server's requirements on how the client can use the data provided by the authorization server. The registration process SHOULD display this URL to the person registering the client if it is given. As described in Section 5, despite the identifier "op_policy_uri", appearing to be OpenID-specific, its usage in this specification is actually referring to a general OAuth 2.0 feature that is not specific to OpenID Connect.

op_tos_uri
OPTIONAL. URL that the authorization server provides to the person registering the client to read about the authorization server's terms of service. The registration process SHOULD display this URL to the person registering the client if it is given. As described in Section 5, despite the identifier "op_tos_uri", appearing to be OpenID-specific, its usage in this specification is actually referring to a general OAuth 2.0 feature that is not specific to OpenID Connect.

revocation_endpoint
OPTIONAL. URL of the authorization server's OAuth 2.0 revocation endpoint [RFC7009].

revocation_endpoint_auth_methods_supported
OPTIONAL. JSON array containing a list of client authentication methods supported by this revocation endpoint. The valid client authentication method values are those registered in the IANA "OAuth Token Endpoint Authentication Methods" registry [IANA.OAuth.Parameters]. If omitted, the default is "client_secret_basic" -- the HTTP Basic Authentication Scheme specified in Section 2.3.1 of OAuth 2.0 [RFC6749].

revocation_endpoint_auth_signing_alg_values_supported
OPTIONAL. JSON array containing a list of the JWS signing algorithms ("alg" values) supported by the revocation endpoint for the signature on the JWT [JWT] used to authenticate the client at the revocation endpoint for the "private_key_jwt" and "client_secret_jwt" authentication methods. This metadata entry MUST be present if either of these authentication methods are specified in the "revocation_endpoint_auth_methods_supported"

entry. No default algorithms are implied if this entry is omitted. The value "none" MUST NOT be used.

introspection_endpoint
OPTIONAL. URL of the authorization server's OAuth 2.0 introspection endpoint [RFC7662].

introspection_endpoint_auth_methods_supported
OPTIONAL. JSON array containing a list of client authentication methods supported by this introspection endpoint. The valid client authentication method values are those registered in the IANA "OAuth Token Endpoint Authentication Methods" registry [IANA.OAuth.Parameters] or those registered in the IANA "OAuth Access Token Types" registry [IANA.OAuth.Parameters]. (These values are and will remain distinct, due to Section 7.2.) If omitted, the set of supported authentication methods MUST be determined by other means.

introspection_endpoint_auth_signing_alg_values_supported
OPTIONAL. JSON array containing a list of the JWS signing algorithms ("alg" values) supported by the introspection endpoint for the signature on the JWT [JWT] used to authenticate the client at the introspection endpoint for the "private_key_jwt" and "client_secret_jwt" authentication methods. This metadata entry MUST be present if either of these authentication methods are specified in the "introspection_endpoint_auth_methods_supported" entry. No default algorithms are implied if this entry is omitted. The value "none" MUST NOT be used.

code_challenge_methods_supported
OPTIONAL. JSON array containing a list of PKCE [RFC7636] code challenge methods supported by this authorization server. Code challenge method values are used in the "code_challenge_method" parameter defined in Section 4.3 of [RFC7636]. The valid code challenge method values are those registered in the IANA "PKCE Code Challenge Methods" registry [IANA.OAuth.Parameters]. If omitted, the authorization server does not support PKCE.

Additional authorization server metadata parameters MAY also be used. Some are defined by other specifications, such as OpenID Connect Discovery 1.0 [OpenID.Discovery].

2.1. Signed Authorization Server Metadata

In addition to JSON elements, metadata values MAY also be provided as a "signed_metadata" value, which is a JSON Web Token (JWT) [JWT] that asserts metadata values about the authorization server as a bundle. A set of claims that can be used in signed metadata are defined in

Section 2. The signed metadata MUST be digitally signed or MACed using JSON Web Signature (JWS) [JWS] and MUST contain an "iss" (issuer) claim denoting the party attesting to the claims in the signed metadata. Consumers of the metadata MAY ignore the signed metadata if they do not support this feature. If the consumer of the metadata supports signed metadata, metadata values conveyed in the signed metadata MUST take precedence over the corresponding values conveyed using plain JSON elements.

Signed metadata is included in the authorization server metadata JSON object using this OPTIONAL member:

signed_metadata
: A JWT containing metadata values about the authorization server as claims. This is a string value consisting of the entire signed JWT. A "signed_metadata" metadata value SHOULD NOT appear as a claim in the JWT.

3. Obtaining Authorization Server Metadata

Authorization servers supporting metadata MUST make a JSON document containing metadata as specified in Section 2 available at a path formed by inserting a well-known URI string into the authorization server’s issuer identifier between the host component and the path component, if any. By default, the well-known URI string used is "/.well-known/oauth-authorization-server". This path MUST use the "https" scheme. The syntax and semantics of ".well-known" are defined in RFC 5785 [RFC5785]. The well-known URI suffix used MUST be registered in the IANA "Well-Known URIs" registry [IANA.well-known].

Different applications utilizing OAuth authorization servers in application-specific ways may define and register different well-known URI suffixes used to publish authorization server metadata as used by those applications. For instance, if the Example application uses an OAuth authorization server in an Example-specific way, and there are Example-specific metadata values that it needs to publish, then it might register and use the "example-configuration" URI suffix and publish the metadata document at the path formed by inserting "/.well-known/example-configuration" between the host and path components of the authorization server’s issuer identifier. Alternatively, many such applications will use the default well-known URI string "/.well-known/oauth-authorization-server", which is the right choice for general-purpose OAuth authorization servers, and not register an application-specific one.

An OAuth 2.0 application using this specification MUST specify what well-known URI suffix it will use for this purpose. The same authorization server MAY choose to publish its metadata at multiple well-known locations derived from its issuer identifier, for example, publishing metadata at both "/.well-known/example-configuration" and "/.well-known/oauth-authorization-server".

Some OAuth applications will choose to use the well-known URI suffix "openid-configuration". As described in Section 5, despite the identifier "/.well-known/openid-configuration", appearing to be OpenID-specific, its usage in this specification is actually referring to a general OAuth 2.0 feature that is not specific to OpenID Connect.

3.1. Authorization Server Metadata Request

An authorization server metadata document MUST be queried using an HTTP "GET" request at the previously specified path.

The client would make the following request when the issuer identifier is "https://example.com" and the well-known URI suffix is "oauth-authorization-server" to obtain the metadata, since the issuer identifier contains no path component:

```
GET /.well-known/oauth-authorization-server HTTP/1.1
Host: example.com
```

If the issuer identifier value contains a path component, any terminating "/" MUST be removed before inserting "/.well-known/" and the well-known URI suffix between the host component and the path component. The client would make the following request when the issuer identifier is "https://example.com/issuer1" and the well-known URI suffix is "oauth-authorization-server" to obtain the metadata, since the issuer identifier contains a path component:

```
GET /.well-known/oauth-authorization-server/issuer1 HTTP/1.1
Host: example.com
```

Using path components enables supporting multiple issuers per host. This is required in some multi-tenant hosting configurations. This use of ".well-known" is for supporting multiple issuers per host; unlike its use in RFC 5785 [RFC5785], it does not provide general information about the host.

3.2. Authorization Server Metadata Response

The response is a set of claims about the authorization server’s configuration, including all necessary endpoints and public key location information. A successful response MUST use the 200 OK HTTP status code and return a JSON object using the "application/json" content type that contains a set of claims as its members that are a subset of the metadata values defined in Section 2. Other claims MAY also be returned.

Claims that return multiple values are represented as JSON arrays. Claims with zero elements MUST be omitted from the response.

An error response uses the applicable HTTP status code value.

The following is a non-normative example response:

```
HTTP/1.1 200 OK
Content-Type: application/json

{
 "issuer":
   "https://server.example.com",
 "authorization_endpoint":
   "https://server.example.com/authorize",
 "token_endpoint":
   "https://server.example.com/token",
 "token_endpoint_auth_methods_supported":
   ["client_secret_basic", "private_key_jwt"],
 "token_endpoint_auth_signing_alg_values_supported":
   ["RS256", "ES256"],
 "userinfo_endpoint":
   "https://server.example.com/userinfo",
 "jwks_uri":
   "https://server.example.com/jwks.json",
 "registration_endpoint":
   "https://server.example.com/register",
 "scopes_supported":
   ["openid", "profile", "email", "address",
    "phone", "offline_access"],
 "response_types_supported":
   ["code", "code token"],
 "service_documentation":
   "http://server.example.com/service_documentation.html",
 "ui_locales_supported":
   ["en-US", "en-GB", "en-CA", "fr-FR", "fr-CA"]
}
```

3.3. Authorization Server Metadata Validation

The "issuer" value returned MUST be identical to the authorization server’s issuer identifier value into which the well-known URI string was inserted to create the URL used to retrieve the metadata. If these values are not identical, the data contained in the response MUST NOT be used.

4. String Operations

Processing some OAuth 2.0 messages requires comparing values in the messages to known values. For example, the member names in the metadata response might be compared to specific member names such as "issuer". Comparing Unicode [UNICODE] strings, however, has significant security implications.

Therefore, comparisons between JSON strings and other Unicode strings MUST be performed as specified below:

1. Remove any JSON applied escaping to produce an array of Unicode code points.

2. Unicode Normalization [USA15] MUST NOT be applied at any point to either the JSON string or to the string it is to be compared against.

3. Comparisons between the two strings MUST be performed as a Unicode code point to code point equality comparison.

Note that this is the same equality comparison procedure described in Section 8.3 of [RFC7159].

5. Compatibility Notes

The identifiers "/.well-known/openid-configuration", "op_policy_uri", and "op_tos_uri" contain strings referring to the OpenID Connect [OpenID.Core] family of specifications that were originally defined by "OpenID Connect Discovery 1.0" [OpenID.Discovery]. Despite the reuse of these identifiers that appear to be OpenID-specific, their usage in this specification is actually referring to general OAuth 2.0 features that are not specific to OpenID Connect.

The algorithm for transforming the issuer identifier to an authorization server metadata location defined in Section 3 is equivalent to the corresponding transformation defined in Section 4 of "OpenID Connect Discovery 1.0" [OpenID.Discovery], provided that the issuer identifier contains no path component. However, they are different when there is a path component, because OpenID Connect Discovery 1.0 specifies that the well-known URI string is appended to the issuer identifier (e.g., "https://example.com/issuer1/.well-known/openid-configuration"), whereas this specification specifies that the well-known URI string is inserted before the path component of the issuer identifier (e.g., "https://example.com/.well-known/openid-configuration/issuer1").

Going forward, OAuth authorization server metadata locations should use the transformation defined in this specification. However, when deployed in legacy environments in which the OpenID Connect Discovery 1.0 transformation is already used, it may be necessary during a transition period to publish metadata for issuer identifiers containing a path component at both locations. During this transition period, applications should first apply the transformation defined in this specification and attempt to retrieve the authorization server metadata from the resulting location; only if the retrieval from that location fails should they fall back to attempting to retrive it from the alternate location obtained using the transformation defined by OpenID Connect Discovery 1.0. This backwards-compatibility behavior should only be necessary when the well-known URI suffix employed by the application is "openid-configuration".

6. Security Considerations

6.1. TLS Requirements

Implementations MUST support TLS. Which version(s) ought to be implemented will vary over time and depend on the widespread deployment and known security vulnerabilities at the time of implementation. The authorization server MUST support TLS version 1.2 [RFC5246] and MAY support additional transport-layer security mechanisms meeting its security requirements. When using TLS, the client MUST perform a TLS/SSL server certificate check, per RFC 6125 [RFC6125]. Implementation security considerations can be found in Recommendations for Secure Use of TLS and DTLS [BCP195].

To protect against information disclosure and tampering, confidentiality protection MUST be applied using TLS with a ciphersuite that provides confidentiality and integrity protection.

6.2. Impersonation Attacks

TLS certificate checking MUST be performed by the client, as described in Section 6.1, when making an authorization server metadata request. Checking that the server certificate is valid for the issuer identifier URL prevents man-in-middle and DNS-based attacks. These attacks could cause a client to be tricked into using an attacker’s keys and endpoints, which would enable impersonation of the legitimate authorization server. If an attacker can accomplish this, they can access the resources that the affected client has access to using the authorization server that they are impersonating.

An attacker may also attempt to impersonate an authorization server by publishing a metadata document that contains an "issuer" claim using the issuer identifier URL of the authorization server being impersonated, but with its own endpoints and signing keys. This would enable it to impersonate that authorization server, if accepted by the client. To prevent this, the client MUST ensure that the issuer identifier URL it is using as the prefix for the metadata request exactly matches the value of the "issuer" metadata value in the authorization server metadata document received by the client.

6.3. Publishing Metadata in a Standard Format

Publishing information about the authorization server in a standard format makes it easier for both legitimate clients and attackers to use the authorization server. Whether an authorization server publishes its metadata in an ad-hoc manner or in the standard format defined by this specification, the same defenses against attacks that might be mounted that use this information should be applied.

6.4. Protected Resources

Secure determination of appropriate protected resources to use with an authorization server for all use cases is out of scope of this specification. This specification assumes that the client has a means of determining appropriate protected resources to use with an authorization server and that the client is using the correct metadata for each authorization server. Implementers need to be aware that if an inappropriate protected resource is used by the client, that an attacker may be able to act as a man-in-the-middle proxy to a valid protected resource without it being detected by the authorization server or the client.

The ways to determine the appropriate protected resources to use with an authorization server are in general, application-dependent. For instance, some authorization servers are used with a fixed protected resource or set of protected resources, the locations of which may be well known, or which could be published as metadata values by the authorization server. In other cases, the set of resources that can be used with an authorization server can by dynamically changed by administrative actions. Many other means of determining appropriate associations between authorization servers and protected resources are also possible.

7. IANA Considerations

The following registration procedure is used for the registry established by this specification.

Values are registered on a Specification Required [RFC8126] basis after a two-week review period on the oauth-ext-review@ietf.org mailing list, on the advice of one or more Designated Experts. However, to allow for the allocation of values prior to publication, the Designated Experts may approve registration once they are satisfied that such a specification will be published.

Registration requests sent to the mailing list for review should use an appropriate subject (e.g., "Request to register OAuth Authorization Server Metadata: example").

Within the review period, the Designated Experts will either approve or deny the registration request, communicating this decision to the review list and IANA. Denials should include an explanation and, if applicable, suggestions as to how to make the request successful. Registration requests that are undetermined for a period longer than 21 days can be brought to the IESG's attention (using the iesg@ietf.org mailing list) for resolution.

Criteria that should be applied by the Designated Experts includes determining whether the proposed registration duplicates existing functionality, determining whether it is likely to be of general applicability or whether it is useful only for a single application, and whether the registration makes sense.

IANA must only accept registry updates from the Designated Experts and should direct all requests for registration to the review mailing list.

It is suggested that multiple Designated Experts be appointed who are able to represent the perspectives of different applications using this specification, in order to enable broadly-informed review of registration decisions. In cases where a registration decision could be perceived as creating a conflict of interest for a particular Expert, that Expert should defer to the judgment of the other Experts.

7.1. OAuth Authorization Server Metadata Registry

This specification establishes the IANA "OAuth Authorization Server Metadata" registry for OAuth 2.0 authorization server metadata names. The registry records the authorization server metadata member and a reference to the specification that defines it.

The Designated Experts must either:

(a) require that metadata names and values being registered use only printable ASCII characters excluding double quote ('"') and backslash

('\') (the Unicode characters with code points U+0021, U+0023 through U+005B, and U+005D through U+007E), or

(b) if new metadata members or values are defined that use other code points, require that their definitions specify the exact Unicode code point sequences used to represent them. Furthermore, proposed registrations that use Unicode code points that can only be represented in JSON strings as escaped characters must not be accepted.

7.1.1. Registration Template

Metadata Name:
The name requested (e.g., "issuer"). This name is case-sensitive. Names may not match other registered names in a case-insensitive manner (one that would cause a match if the Unicode toLowerCase() operation were applied to both strings) unless the Designated Experts state that there is a compelling reason to allow an exception.

Metadata Description:
Brief description of the metadata (e.g., "Issuer identifier URL").

Change Controller:
For Standards Track RFCs, list the "IESG". For others, give the name of the responsible party. Other details (e.g., postal address, email address, home page URI) may also be included.

Specification Document(s):
Reference to the document or documents that specify the parameter, preferably including URIs that can be used to retrieve copies of the documents. An indication of the relevant sections may also be included but is not required.

7.1.2. Initial Registry Contents

- Metadata Name: "issuer"
- Metadata Description: Authorization server's issuer identifier URL
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Metadata Name: "authorization_endpoint"
- Metadata Description: URL of the authorization server's authorization endpoint
- Change Controller: IESG
- Specification Document(s): Section 2 of [[this specification]]

- Metadata Name: "token_endpoint"

o Metadata Description: URL of the authorization server’s token endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "jwks_uri"
o Metadata Description: URL of the authorization server’s JWK Set document
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "registration_endpoint"
o Metadata Description: URL of the authorization server’s OAuth 2.0 Dynamic Client Registration Endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "scopes_supported"
o Metadata Description: JSON array containing a list of the OAuth 2.0 "scope" values that this authorization server supports
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "response_types_supported"
o Metadata Description: JSON array containing a list of the OAuth 2.0 "response_type" values that this authorization server supports
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "response_modes_supported"
o Metadata Description: JSON array containing a list of the OAuth 2.0 "response_mode" values that this authorization server supports
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "grant_types_supported"
o Metadata Description: JSON array containing a list of the OAuth 2.0 grant type values that this authorization server supports
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "token_endpoint_auth_methods_supported"
o Metadata Description: JSON array containing a list of client authentication methods supported by this token endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "token_endpoint_auth_signing_alg_values_supported"

- o Metadata Description: JSON array containing a list of the JWS signing algorithms supported by the token endpoint for the signature on the JWT used to authenticate the client at the token endpoint
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Metadata Name: "service_documentation"
- o Metadata Description: URL of a page containing human-readable information that developers might want or need to know when using the authorization server
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Metadata Name: "ui_locales_supported"
- o Metadata Description: Languages and scripts supported for the user interface, represented as a JSON array of BCP47 language tag values
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Metadata Name: "op_policy_uri"
- o Metadata Description: URL that the authorization server provides to the person registering the client to read about the authorization server’s requirements on how the client can use the data provided by the authorization server
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Metadata Name: "op_tos_uri"
- o Metadata Description: URL that the authorization server provides to the person registering the client to read about the authorization server’s terms of service
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Metadata Name: "revocation_endpoint"
- o Metadata Description: URL of the authorization server’s OAuth 2.0 revocation endpoint
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

- o Metadata Name: "revocation_endpoint_auth_methods_supported"
- o Metadata Description: JSON array containing a list of client authentication methods supported by this revocation endpoint
- o Change Controller: IESG
- o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "revocation_endpoint_auth_signing_alg_values_supported"
o Metadata Description: JSON array containing a list of the JWS signing algorithms supported by the revocation endpoint for the signature on the JWT used to authenticate the client at the revocation endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "introspection_endpoint"
o Metadata Description: URL of the authorization server's OAuth 2.0 introspection endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "introspection_endpoint_auth_methods_supported"
o Metadata Description: JSON array containing a list of client authentication methods supported by this introspection endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "introspection_endpoint_auth_signing_alg_values_supported"
o Metadata Description: JSON array containing a list of the JWS signing algorithms supported by the introspection endpoint for the signature on the JWT used to authenticate the client at the introspection endpoint
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

o Metadata Name: "code_challenge_methods_supported"
o Metadata Description: PKCE code challenge methods supported by this authorization server
o Change Controller: IESG
o Specification Document(s): Section 2 of [[this specification]]

7.2. Updated Registration Instructions

This specification adds to the instructions for the Designated Experts of the following IANA registries, both of which are in the "OAuth Parameters" registry [IANA.OAuth.Parameters]:

o OAuth Access Token Types
o OAuth Token Endpoint Authentication Methods

IANA has added a link to this specification in the Reference sections of these registries. [[RFC Editor: The above sentence is written in the past tense as it would appear in the final specification, even though these links won't actually be created until after the IESG has requested publication of the specification. Please delete this note after the links are in place.]]

For these registries, the designated experts must reject registration requests in one registry for values already occurring in the other registry. This is necessary because the "introspection_endpoint_auth_methods_supported" parameter allows for the use of values from either registry. That way, because the values in the two registries will continue to be mutually exclusive, no ambiguities will arise.

7.3. Well-Known URI Registry

This specification registers the well-known URI defined in Section 3 in the IANA "Well-Known URIs" registry [IANA.well-known] established by RFC 5785 [RFC5785].

7.3.1. Registry Contents

- o URI suffix: "oauth-authorization-server"
- o Change controller: IESG
- o Specification document: Section 3 of [[this specification]]
- o Related information: (none)

8. References

8.1. Normative References

[BCP195] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <http://www.rfc-editor.org/info/bcp195>.

[IANA.OAuth.Parameters] IANA, "OAuth Parameters", <http://www.iana.org/assignments/oauth-parameters>.

[JWE] Jones, M. and J. Hildebrand, "JSON Web Encryption (JWE)", RFC 7516, DOI 10.17487/RFC7516, May 2015, <http://tools.ietf.org/html/rfc7516>.

[JWK] Jones, M., "JSON Web Key (JWK)", RFC 7517, DOI 10.17487/RFC7517, May 2015, <http://tools.ietf.org/html/rfc7517>.

[JWS] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Signature (JWS)", RFC 7515, DOI 10.17487/RFC7515, May 2015, <http://tools.ietf.org/html/rfc7515>.

[JWT] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <http://tools.ietf.org/html/rfc7519>.

[OAuth.Post]
Jones, M. and B. Campbell, "OAuth 2.0 Form Post Response Mode", April 2015, <http://openid.net/specs/oauth-v2-form-post-response-mode-1_0.html>.

[OAuth.Responses]
de Medeiros, B., Ed., Scurtescu, M., Tarjan, P., and M. Jones, "OAuth 2.0 Multiple Response Type Encoding Practices", February 2014, <http://openid.net/specs/oauth-v2-multiple-response-types-1_0.html>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <https://www.rfc-editor.org/info/rfc5246>.

[RFC5646] Phillips, A., Ed. and M. Davis, Ed., "Tags for Identifying Languages", BCP 47, RFC 5646, DOI 10.17487/RFC5646, September 2009, <https://www.rfc-editor.org/info/rfc5646>.

[RFC5785] Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known Uniform Resource Identifiers (URIs)", RFC 5785, DOI 10.17487/RFC5785, April 2010, <https://www.rfc-editor.org/info/rfc5785>.

[RFC6125] Saint-Andre, P. and J. Hodges, "Representation and Verification of Domain-Based Application Service Identity within Internet Public Key Infrastructure Using X.509 (PKIX) Certificates in the Context of Transport Layer Security (TLS)", RFC 6125, DOI 10.17487/RFC6125, March 2011, <https://www.rfc-editor.org/info/rfc6125>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <https://www.rfc-editor.org/info/rfc6749>.

[RFC7009] Lodderstedt, T., Ed., Dronia, S., and M. Scurtescu, "OAuth 2.0 Token Revocation", RFC 7009, DOI 10.17487/RFC7009, August 2013, <https://www.rfc-editor.org/info/rfc7009>.

[RFC7033] Jones, P., Salgueiro, G., Jones, M., and J. Smarr, "WebFinger", RFC 7033, DOI 10.17487/RFC7033, September 2013, <https://www.rfc-editor.org/info/rfc7033>.

[RFC7159] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", RFC 7159, DOI 10.17487/RFC7159, March 2014, <https://www.rfc-editor.org/info/rfc7159>.

[RFC7591] Richer, J., Ed., Jones, M., Bradley, J., Machulak, M., and P. Hunt, "OAuth 2.0 Dynamic Client Registration Protocol", RFC 7591, DOI 10.17487/RFC7591, July 2015, <https://www.rfc-editor.org/info/rfc7591>.

[RFC7636] Sakimura, N., Ed., Bradley, J., and N. Agarwal, "Proof Key for Code Exchange by OAuth Public Clients", RFC 7636, DOI 10.17487/RFC7636, September 2015, <https://www.rfc-editor.org/info/rfc7636>.

[RFC7662] Richer, J., Ed., "OAuth 2.0 Token Introspection", RFC 7662, DOI 10.17487/RFC7662, October 2015, <https://www.rfc-editor.org/info/rfc7662>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[UNICODE] The Unicode Consortium, "The Unicode Standard", <http://www.unicode.org/versions/latest/>.

[USA15] Davis, M. and K. Whistler, "Unicode Normalization Forms", Unicode Standard Annex 15, June 2015, <http://www.unicode.org/reports/tr15/>.

8.2. Informative References

[I-D.ietf-oauth-mix-up-mitigation] Jones, M., Bradley, J., and N. Sakimura, "OAuth 2.0 Mix-Up Mitigation", draft-ietf-oauth-mix-up-mitigation-01 (work in progress), July 2016.

[IANA.well-known]
IANA, "Well-Known URIs",
<http://www.iana.org/assignments/well-known-uris>.

[OpenID.Core]
Sakimura, N., Bradley, J., Jones, M., de Medeiros, B., and C. Mortimore, "OpenID Connect Core 1.0", November 2014, <http://openid.net/specs/openid-connect-core-1_0.html>.

[OpenID.Discovery]
Sakimura, N., Bradley, J., Jones, M., and E. Jay, "OpenID Connect Discovery 1.0", November 2014, <http://openid.net/specs/openid-connect-discovery-1_0.html>.

[OpenID.Registration]
Sakimura, N., Bradley, J., and M. Jones, "OpenID Connect Dynamic Client Registration 1.0", November 2014, <http://openid.net/specs/openid-connect-registration-1_0.html>.

Appendix A. Acknowledgements

This specification is based on the OpenID Connect Discovery 1.0 specification, which was produced by the OpenID Connect working group of the OpenID Foundation. This specification standardizes the de facto usage of the metadata format defined by OpenID Connect Discovery to publish OAuth authorization server metadata.

The authors would like to thank the following people for their reviews of this specification: Shwetha Bhandari, Ben Campbell, Brian Campbell, Brian Carpenter, William Denniss, Vladimir Dzhuvinov, Donald Eastlake, Samuel Erdtman, George Fletcher, Dick Hardt, Phil Hunt, Alexey Melnikov, Tony Nadalin, Mark Nottingham, Eric Rescorla, Justin Richer, Adam Roach, Hannes Tschofenig, and Hans Zandbelt.

Appendix B. Document History

[[to be removed by the RFC Editor before publication as an RFC]]

-10

o Clarified the meaning of "case-insensitive", as suggested by Alexey Melnikov.

-09

- o Revised the transformation between the issuer identifier and the authorization server metadata location to conform to BCP 190, as suggested by Adam Roach.

- o Defined the characters allowed in registered metadata names and values, as suggested by Alexey Melnikov.

- o Changed to using the RFC 8174 boilerplate instead of the RFC 2119 boilerplate, as suggested by Ben Campbell.

- o Acknowledged additional reviewers.

-08

- o Changed the "authorization_endpoint" to be REQUIRED only when grant types are supported that use the authorization endpoint.

- o Added the statement, to provide historical context, that this specification standardizes the de facto usage of the metadata format defined by OpenID Connect Discovery to publish OAuth authorization server metadata.

- o Applied clarifications suggested by Mark Nottingham about when application-specific well-known suffixes are and are not appropriate.

- o Acknowledged additional reviewers.

-07

- o Applied clarifications suggested by EKR.

-06

- o Incorporated resolutions to working group last call comments.

-05

- o Removed the "protected_resources" element and the reference to draft-jones-oauth-resource-metadata.

-04

- o Added the ability to list protected resources with the "protected_resources" element.

- o Added ability to provide signed metadata with the "signed_metadata" element.

- Removed "Discovery" from the name, since this is now just about authorization server metadata.

-03

- Changed term "issuer URL" to "issuer identifier" for terminology consistency, paralleling the same terminology consistency change in the mix-up mitigation spec.

-02

- Changed the title to OAuth 2.0 Authorization Server Discovery Metadata.
- Made "jwks_uri" and "registration_endpoint" OPTIONAL.
- Defined the well-known URI string "/.well-known/oauth-authorization-server".
- Added security considerations about publishing authorization server discovery metadata in a standard format.
- Added security considerations about protected resources.
- Added more information to the "grant_types_supported" and "response_types_supported" definitions.
- Referenced the working group Mix-Up Mitigation draft.
- Changed some example metadata values.
- Acknowledged individuals for their contributions to the specification.

-01

- Removed WebFinger discovery.
- Clarified the relationship between the issuer identifier URL and the well-known URI path relative to it at which the discovery metadata document is located.

-00

- Created the initial working group version based on draft-jones-oauth-discovery-01, with no normative changes.

Authors' Addresses

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

Nat Sakimura
Nomura Research Institute, Ltd.

Email: n-sakimura@nri.co.jp
URI: http://nat.sakimura.org/

John Bradley
Ping Identity

Email: ve7jtb@ve7jtb.com
URI: http://www.thread-safe.com/

OAuth Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 7, 2017
M. Jones
Microsoft
J. Bradley
Ping Identity
N. Sakimura
NRI
July 6, 2016

OAuth 2.0 Mix-Up Mitigation
draft-ietf-oauth-mix-up-mitigation-01

Abstract

This specification defines an extension to The OAuth 2.0 Authorization Framework that enables the authorization server to dynamically provide the client using it with additional information about the current protocol interaction that can be validated by the client and that enables the client to dynamically provide the authorization server with additional information about the current protocol interaction that can be validated by the authorization server. This additional information can be used by the client and the authorization server to prevent classes of attacks in which the client might otherwise be tricked into using inconsistent sets of metadata from multiple authorization servers, including potentially using a token endpoint that does not belong to the same authorization server as the authorization endpoint used. Recent research publications refer to these as "IdP Mix-Up" and "Malicious Endpoint" attacks.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

OAuth 2.0 [RFC6749] clients use multiple authorization server endpoints when using some OAuth response types. For instance, when using the "code" response type, the client uses both the authorization endpoint and the token endpoint. It is important that endpoints belonging to the same authorization server always be used together. Otherwise, information produced by one authorization server could mistakenly be sent by the client to different authorization server, resulting in some of the attacks described in Section 7. Recent research publications refer to these specific attacks as "IdP Mix-Up" [arXiv.1601.01229v2] and "Malicious Endpoint" [arXiv.1508.04324v2] attacks.

The client obviously cannot be confused into using endpoints from multiple authorization servers in an authorization flow if the client is configured to use only a single authorization server. However, the client can potentially be tricked into mixing endpoints if it is configured to use more than one authorization server, whether the configuration is dynamic or static. The client may be confused if it has no way to determine whether the set of endpoints belongs to the same authorization server. Or, a client may be confused simply because it is receiving authorization responses from more than one authorization server at the same redirection endpoint and the client is insufficiently able to determine that the response received is associated with the correct authorization server.

This specification enables the authorization server to dynamically provide the client using it with additional information about the current protocol interaction that can be validated by the client and that enables the client to dynamically provide the authorization server with additional information about the current protocol interaction that can be validated by the authorization server. This enables them to abort interactions in which endpoints from multiple authorization servers would otherwise be used.

The mitigation data provided by the authorization server to the client is an issuer identifier, which is used to identify the authorization server, and a client ID, which is used to verify that the response is from the correct authorization server and is intended for this client. The issuer identifier is defined in Section 2 of [OAuth.Discovery]. If supported by the authorization server, the issuer identifier can also be used to obtain a consistent set of metadata describing the authorization server configuration, as also described in [OAuth.Discovery].

This mitigation data is returned to the client in the authorization response. The syntax for returning the mitigation data from the authorization server is dependent upon the OAuth response type being used. The syntax used with the existing response types registered in the IANA "OAuth Authorization Endpoint Response Types" registry [IANA.OAuth.Parameters] as of the time of this writing is defined by this specification. Two of these response types are defined by RFC 6749 [RFC6749]; the rest are defined by [OAuth.Responses].

The mitigation data provided by the client to the authorization server is the existing "state" value defined by RFC 6749 [RFC6749], but adding also sending it from the client to the token endpoint. This is used by the authorization server to verify that the authorization code and state both belong to the same protocol interaction.

1.1. Requirements Notation and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2. Terminology

This specification uses the terms "Access Token", "Authorization Code", "Authorization Endpoint", "Authorization Grant", "Authorization Server", "Client", "Client Authentication", "Client Identifier", "Client Secret", "Grant Type", "Protected Resource", "Redirection URI", "Refresh Token", "Resource Owner", "Resource Server", "Response Type", and "Token Endpoint" defined by OAuth 2.0 [RFC6749], the terms "Claim Name", "Claim Value", and "JSON Web Token (JWT)" defined by JSON Web Token (JWT) [JWT].

2. The OAuth Issuer Identifier

The OAuth issuer identifier serves as a concrete identifier for the authorization server. As defined in [OAuth.Discovery], the issuer identifier is a URL that uses the "https" scheme and has no query or fragment components. Also as specified there, this is the location where ".well-known" RFC 5785 [RFC5785] resources containing information about the authorization server are published. In particular, when discovery is supported, the authorization server’s metadata is retrieved as a JSON document [RFC7159] from a path derived from this URL. This metadata document contains a consistent set of metadata describing the authorization server configuration.

Implementations supporting this specification MAY also support discovery or they MAY simply use the issuer identifier as a concrete identifier for the authorization server. This specification does not rely upon the authorization server publishing or the client retrieving a discovery metadata document.

3. Mitigation Data Returned in Authorization Response

Mitigating the attacks relies on the authorization server returning additional data about the interaction and the client checking that data. The mitigation data returned is the client ID and the issuer identifier. The syntax for returning the mitigation data from the authorization server is dependent upon the OAuth response type being used.

3.1. Mitigation Data Returned in Authorization Response Parameters

Some OAuth response types do not already return the issuer identifier and client ID in the authorization response. When this is the case, the mitigation data is returned as additional OAuth response parameters.

These new response parameters are defined for this purpose:

client_id
: Client that this response is intended for. It MUST contain the OAuth 2.0 client ID of the client as its value.

iss
: Issuer identifier for the authorization server issuing the response. The "iss" value is a case-sensitive URL using the "https" scheme that contains scheme, host, and optionally, port number and path components and no query or fragment components.

As of the time of this writing, these are the existing response types that are registered in the IANA "OAuth Authorization Endpoint Response Types" registry [IANA.OAuth.Parameters] that do not already return the issuer identifier and client ID in the authorization response: "code", "code token", "none", and "token". Therefore, the client ID and issuer are returned using the new authorization response parameters when using these response types. To avoid duplication, as discussed in Section 7.2, it is NOT RECOMMENDED to also return them in this manner when the response type already returns these values in the authorization response.

3.1.1. Example Authorization Response using Response Parameters

The following example authorization response is to a request that used the "code" response type. It uses the "iss" and "client_id" response parameters to return the mitigation information to the client.

The example successful authorization response follows (with line breaks within lines for display purposes only):

```
HTTP/1.1 302 Found
Location: https://client.example.org/cb?
  code=Qcb0Orv1zh30vL1MPRsbm-diHiMwcLyZvn1arpZv-Jxf_11jnpEX3Tgfvk
  &state=nrsz6AnHzPSVVBYRVTXV6ZTXQeg_eih7hdpewHNXmZ8
  &iss=https://server.example.com
  &client_id=5d9e8a36-569d-4c40-8d6b-6e279ac1c5f1
```

3.2. Mitigation Data Returned in JWT

As of the time of this writing, these are the existing response types that are registered in the IANA "OAuth Authorization Endpoint Response Types" registry [IANA.OAuth.Parameters] that already return the issuer identifier and client ID in the authorization response: "code id_token", "code id_token token", "id_token", and "id_token token". All of these return these values as the "iss" (issuer) claim value and as an "aud" (audience) claim value in a signed ID Token, which is a JSON Web Token [JWT], as specified in "OpenID Connect Core 1.0" [OpenID.Core]. When using these response types, the client MUST use the client ID and issuer values returned in the ID Token for validating the mitigation data.

3.2.1. Example Authorization Response using JWT

The following example authorization response is to a request that used the "id_token token" response type. It uses the "iss" and "aud" claims in the ID Token to return the mitigation information to the client.

The example successful authorization response follows (with line breaks within lines for display purposes only):

```
HTTP/1.1 302 Found
Location: https://client.example.org/cb#
  access_token=jHkWEdUXMU1BwAsC4vtUsZwnNvTIxEl0z9K3vx5KF0Y
  &token_type=Bearer
  &id_token=eyJraWQiOiIxZTlnZGs3IiwiYWxnIjoiUlMyNTYifQ.
  ewogImlzcyI6ICJodHRwczovL3NlcnZlci5leGFtcGxlLmNvbSIsCiAic3ViIjog
  IjI0ODI4OTc2MTAwMSIsCiAiYXVkIjogInM2QmhkUmtxdDMiLAogIm5vbmNlIjog
  Im4tMFM2X1d6QTJNaiIsCiAiZXhwIjogMTMxMTI4MTk3MCwKICJpYXQiOiAxMzEx
  MjgwOTcwLAogImF0X2hhc2giOiAiNzdRbVVQdGpQZnpXdEYyQW5wSzlSUSIKfQ.
  kdqTmftlaXg5WBYBr1wkxhkqCGZPc0k8vTiV5g2jj67jQ7XkrDamYx2bOkZLdZrp
  MPIzkdYB1nZI_G8vQGQuamRhJcEIt21kblGPZ-yhEhdkAiZIZLu38rChalDS2Mh0
  glE_rke5XXRhmqqoEFFdziFdnO3p61-7y51co84OEAZvARSINQaOWIzvioRfs4zw
  IFOaT33Vpxfqr8HDyh31zo9eBW2dSQuCa071z0ENWChWoPliK1JCo_Bk9eDg2uwo
  2ZwhsvHzj6TMQ0lYOTzufSlSmXIKfjlOsb3nftQeR697_hA-nMZyAdL8_NRfaC37
  XnAbW8WB9wCfECp7cuNuOg
  &state=af0ifjsldkj
```

Decoding the ID Token in the response will yield the following claims, which includes the mitigation information in the "iss" and "aud" claims:

```
{
 "iss": "https://server.example.com",
 "sub": "248289761001",
 "aud": "s6BhdRkqt3",
 "nonce": "n-0S6_WzA2Mj",
 "exp": 1311281970,
 "iat": 1311280970,
 "at_hash": "77QmUPtjPfzWtF2AnpK9RQ"
}
```

4. Validating the Authorization Response

Upon receiving the mitigation data in an authorization response, the client MUST validate that the response was intended for it and that the authorization server metadata that it obtained at client registration time is consistent with the authorization server metadata contained in the metadata referenced by the issuer identifier.

The client MUST validate the authorization server configuration as follows:

1. Compare the issuer identifier for the authorization server that the client received when it registered at the authorization

server that it made the request to with the issuer value returned in the "iss" response parameter or the "iss" claim in the ID Token, depending upon the response type being used. If they do not exactly match, the client MUST NOT proceed with the authorization.

2. Verify that the response is intended for this client by confirming that the client's client identifier for the authorization server the request was made to matches the value of the "client_id" response parameter or that the client's client identifier is an audience value of the ID Token, depending upon the response type being used. If not, the client MUST NOT proceed with the authorization.

5. Mitigation Data Sent to the Token Endpoint

Mitigating the attacks also relies on the client sending additional data about the interaction to the token endpoint, for response types that use it, and the authorization server checking that data. The mitigation data sent is the same state value that is sent in the authorization request and returned in the authorization response. This specification defines the new "state" token request parameter for passing this additional information.

As of the time of this writing, these are the existing response types that are registered in the IANA "OAuth Authorization Endpoint Response Types" registry [IANA.OAuth.Parameters] that use the token endpoint: "code", "code id_token", "code id_token token", and "code token". The state value is to be sent in the "state" token request parameter when using these response types, and any new response types registered that use the token endpoint.

5.1. Example Token Request

The following example token request is part of a protocol interaction that used the "code" response type. It uses the "state" request parameter to send mitigation information to the authorization server.

The example of token request follows (with line breaks within lines for display purposes only):

```
POST /token HTTP/1.1
Host: server.example.com
Content-Type: application/x-www-form-urlencoded
Authorization: Basic czZCaGRSa3F0MzpnWDFmQmF0M2JW

grant_type=authorization_code
  &code=SplxlOBeZQQYbYS6WxSbIA
  &redirect_uri=https%3A%2F%2Fclient.example.org%2Fcb
  &state=ZSGXNBavNc-B3kU3DeJnZoWWOzYxsbvj7jp-S0x_z8U
```

6. Validating the Token Request

When the authorization server receives a token request at the token endpoint that contains a value in the "state" parameter, it MUST validate that the state value received exactly matches the state value previously received in the corresponding authorization request. If the recorded state value and the state value received do not exactly match, the authorization server MUST NOT proceed with the authorization.

7. Security Considerations

7.1. IdP Mix-Up and Malicious Endpoint Attacks

The attacks mitigated by this extension are described in detail in "A Comprehensive Formal Security Analysis of OAuth 2.0" [arXiv.1601.01229v2] and "On the security of modern Single Sign-On Protocols: Second-Order Vulnerabilities in OpenID Connect" [arXiv.1508.04324v2]. To mitigate these attacks, clients configured to use more than one authorization server should use authorization servers that return issuer and client ID information and should validate that a consistent set of authorization server endpoints are being used when using response types that utilize multiple endpoints.

When registering, clients SHOULD NOT allow multiple authorization servers to return the same issuer value, and MUST NOT allow multiple authorization servers to return the same issuer and client ID value pair.

7.2. Duplicate Information Attacks

If a protocol is defined to return the same information in multiple locations, this can create an additional attack surface. Knowing that the information is supposed to be the same, recipients will often be lazy and use the information from only one of the locations, not validating that all the supposedly duplicate instances are the same. This can enable attackers to create illegal protocol messages that have different values in the multiple locations and those illegal messages will not be detected or rejected by these lazy recipients.

For this reason, if an OAuth profile is being used that returns the mitigation information defined by this specification in one location, it SHOULD NOT also be returned in another. In particular, if a JWT containing the client ID and issuer values is being returned in the authorization response, they SHOULD NOT also be returned as individual authorization response parameters.

7.3. Cut-and-Paste Attacks

OAuth authorization responses are sent as redirects to redirection URIs, with the response parameters typically passed as URI query parameters or fragment values. A "cut-and-paste" attack is performed by the attacker creating what appears to be a legitimate authorization response, but that substitutes some of the response parameter values with values of the attacker's choosing. Sometimes this is done by copying or "cutting" some values out of a legitimate response and replacing or "pasting" some of these values into a different response, the original version of which may have also been legitimate, creating a combination of response values that are not legitimate and that may cause behaviors sought by the attacker. The Code Substitution threat described in Section 4.4.1.13 of [RFC6819] is one example of the use of a cut-and-paste attack.

A concern with returning the mitigation information as new individual authorization response parameters whose values are not cryptographically bound together is that cut-and-paste attacks against their values will not be detected. A security analysis has not been done of the effects of the new attacks that the use of cut-and-paste against these new values will enable.

To prevent replay of the state in another browser instance by an attacker, the state value MUST be tied to the browser instance in a way that cannot be forged by an attacker. Section 4 of [I-D.bradley-oauth-jwt-encoded-state] provides several examples of how a client can accomplish this.

In the replay attack, the attacker can set cookies in the browser. Using an unsigned cookie to bind state to the browser is not sufficient.

8. IANA Considerations

8.1. OAuth Parameters Registration

This specification registers the following parameters in the IANA "OAuth Parameters" registry [IANA.OAuth.Parameters] established by RFC 6749 [RFC6749].

8.1.1. Registry Contents

- o Parameter name: "client_id"
- o Parameter usage location: Authorization Response
- o Change controller: IESG
- o Specification document(s): Section 3.1 of [[this specification]]
- o Related information: None

- o Parameter name: "iss"
- o Parameter usage location: Authorization Response
- o Change controller: IESG
- o Specification document(s): Section 3.1 of [[this specification]]
- o Related information: None

- o Parameter name: "state"
- o Parameter usage location: Token Request
- o Change controller: IESG
- o Specification document(s): Section 5 of [[this specification]]
- o Related information: None

9. References

9.1. Normative References

[IANA.OAuth.Parameters]
IANA, "OAuth Parameters", <http://www.iana.org/assignments/oauth-parameters>.

[JWT] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <http://tools.ietf.org/html/rfc7519>.

[OAuth.Discovery]
Jones, M., Sakimura, N., and J. Bradley, "OAuth 2.0 Discovery", draft-ietf-oauth-discovery-02 (work in progress), March 2016, <http://tools.ietf.org/html/draft-ietf-oauth-discovery-02>.

[OAuth.Responses]
de Medeiros, B., Ed., Scurtescu, M., Tarjan, P., and M. Jones, "OAuth 2.0 Multiple Response Type Encoding Practices", February 2014, <http://openid.net/specs/oauth-v2-multiple-response-types-1_0.html>.

[OpenID.Core]
Sakimura, N., Bradley, J., Jones, M., de Medeiros, B., and C. Mortimore, "OpenID Connect Core 1.0", November 2014, <http://openid.net/specs/openid-connect-core-1_0.html>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, DOI 10.17487/RFC3629, November 2003, <http://www.rfc-editor.org/info/rfc3629>.

[RFC5785] Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known Uniform Resource Identifiers (URIs)", RFC 5785, DOI 10.17487/RFC5785, April 2010, <http://www.rfc-editor.org/info/rfc5785>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <http://www.rfc-editor.org/info/rfc6749>.

[RFC6819] Lodderstedt, T., Ed., McGloin, M., and P. Hunt, "OAuth 2.0 Threat Model and Security Considerations", RFC 6819, DOI 10.17487/RFC6819, January 2013, <http://www.rfc-editor.org/info/rfc6819>.

[RFC7159] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", RFC 7159, DOI 10.17487/RFC7159, March 2014, <http://www.rfc-editor.org/info/rfc7159>.

[RFC7662] Richer, J., Ed., "OAuth 2.0 Token Introspection", RFC 7662, DOI 10.17487/RFC7662, October 2015, <http://www.rfc-editor.org/info/rfc7662>.

9.2. Informative References

[arXiv.1508.04324v2]
Mladenov, V., Mainka, C., and J. Schwenk, "On the security of modern Single Sign-On Protocols: Second-Order Vulnerabilities in OpenID Connect", arXiv 1508.04324v2, January 2016, <http://arxiv.org/abs/1508.04324v2/>.

[arXiv.1601.01229v2]
Fett, D., Kuesters, R., and G. Schmitz, "A Comprehensive Formal Security Analysis of OAuth 2.0", arXiv 1601.01229v2, January 2016, <http://arxiv.org/abs/1601.01229v2/>.

[I-D.bradley-oauth-jwt-encoded-state]
Bradley, J., Lodderstedt, T., and H. Zandbelt, "Encoding claims in the OAuth 2 state parameter using a JWT", draft-bradley-oauth-jwt-encoded-state-05 (work in progress), December 2015.

[RFC7591] Richer, J., Ed., Jones, M., Bradley, J., Machulak, M., and P. Hunt, "OAuth 2.0 Dynamic Client Registration Protocol", RFC 7591, DOI 10.17487/RFC7591, July 2015, <http://www.rfc-editor.org/info/rfc7591>.

Appendix A. Implementation Notes

The authorization server can compare the two state values either by recording the complete state value between the authorization request and the token request, possibly in the same data structure in which the authorization code issued was recorded, or by recording only a cryptographic hash of the state value, possibly resulting in substantial size savings.

Appendix B. Acknowledgements

Alfred Albrecht, John Bradley, Brian Campbell, Joerg Connotte, William Denniss, Sebastian Ebling, Florian Feldmann, Daniel Fett, Roland Hedberg, Phil Hunt, Ralf Kuesters, Torsten Lodderstedt, Christian Mainka, Vladislav Mladenov, Anthony Nadalin, Justin Richer, Nat Sakimura, Antonio Sanso, Guido Schmitz, Joerg Schwenk, Hannes Tschofenig, and Hans Zandbelt all contributed to the discussions that led to the creation of this specification.

This specification is partially based on the OpenID Connect Core 1.0 specification, which was produced by the OpenID Connect working group of the OpenID Foundation.

Appendix C. Open Issues

- We need to do a security analysis of the cut-and-paste attacks that may be enabled when mitigation information is returned to the client using individual authorization response parameters.

Appendix D. Document History

[[to be removed by the RFC Editor before publication as an RFC]]

-01

- Changed terms "issuer URL" and "configuration information location" to "issuer identifier" so that consistent terminology is used for this.

-00

- Created the initial working group draft from draft-jones-oauth-mix-up-mitigation-01 with no normative changes and adding Nat Sakimura as an editor.

Authors' Addresses

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

John Bradley
Ping Identity

Email: ve7jtb@ve7jtb.com
URI: http://www.thread-safe.com/

Nat Sakimura
Nomura Research Institute, Ltd.

Email: n-sakimura@nri.co.jp
URI: http://nat.sakimura.org/

OAuth Working Group W. Denniss
Internet-Draft Google
Updates: 6749 (if approved) J. Bradley
Intended status: Best Current Practice Ping Identity
Expires: December 11, 2017 June 9, 2017

OAuth 2.0 for Native Apps
draft-ietf-oauth-native-apps-12

Abstract

OAuth 2.0 authorization requests from native apps should only be made through external user-agents, primarily the user's browser. This specification details the security and usability reasons why this is the case, and how native apps and authorization servers can implement this best practice.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The OAuth 2.0 [RFC6749] authorization framework documents two approaches in Section 9 for native apps to interact with the authorization endpoint: an embedded user-agent, and an external user-agent.

This best current practice requires that only external user-agents like the browser are used for OAuth by native apps. It documents how native apps can implement authorization flows using the browser as the preferred external user-agent, and the requirements for authorization servers to support such usage.

This practice is also known as the AppAuth pattern, in reference to open source libraries [AppAuth] that implement it.

2. Notational Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in Key words for use in RFCs to Indicate Requirement Levels [RFC2119]. If these words are used without being spelled in uppercase then they are to be interpreted with their normal natural language meanings.

3. Terminology

In addition to the terms defined in referenced specifications, this document uses the following terms:

"native app" An app or application that is installed by the user to their device, as distinct from a web app that runs in the browser context only. Apps implemented using web-based technology but distributed as a native app, so-called hybrid apps, are considered equivalent to native apps for the purpose of this specification.

"app" In this document, "app" means a "native app" unless further specified.

"app store" An ecommerce store where users can download and purchase apps.

"OAuth" In this document, OAuth refers to the OAuth 2.0 Authorization Framework [RFC6749].

"external user-agent" A user-agent capable of handling the authorization request that is a separate entity or security domain to the native app making the request (such as a browser), such that the app cannot access the cookie storage, nor inspect or modify page content.

"embedded user-agent" A user-agent hosted inside the native app itself (such as via a web-view), with which the app has control over to the extent it is capable of accessing the cookie storage and/or modifying the page content.

"browser" The default application launched by the operating system to handle "http" and "https" scheme URI content.

"in-app browser tab" A programmatic instantiation of the browser that is displayed inside a host app, but retains the full security properties and authentication state of the browser. Has different platform-specific product names, such as SFSafariViewController on iOS, and Custom Tabs on Android.

"inter-app communication" Communication between two apps on a device.

"claimed HTTPS URI" Some platforms allow apps to claim a HTTPS scheme URI after proving ownership of the domain name. URIs claimed in such a way are then opened in the app instead of the browser.

"private-use URI scheme" A private-use URI scheme defined by the app and registered with the operating system. URI requests to such schemes trigger the app which registered it to be launched to handle the request.

"web-view" A web browser UI (user interface) component that can be embedded in apps to render web pages, used to create embedded user-agents.

"reverse domain name notation" A naming convention based on the domain name system, but where the domain components are reversed, for example "app.example.com" becomes "com.example.app".

4. Overview

The best current practice for authorizing users in native apps is to perform the OAuth authorization request in an external user-agent (typically the browser), rather than an embedded user-agent (such as one implemented with web-views).

Previously it was common for native apps to use embedded user-agents (commonly implemented with web-views) for OAuth authorization requests. That approach has many drawbacks, including the host app being able to copy user credentials and cookies, and the user needing to authenticate from scratch in each app. See Section 8.12 for a deeper analysis of using embedded user-agents for OAuth.

Native app authorization requests that use the browser are more secure and can take advantage of the user's authentication state. Being able to use the existing authentication session in the browser enables single sign-on, as users don't need to authenticate to the authorization server each time they use a new app (unless required by authorization server policy).

Supporting authorization flows between a native app and the browser is possible without changing the OAuth protocol itself, as the authorization request and response are already defined in terms of URIs, which encompasses URIs that can be used for inter-app communication. Some OAuth server implementations that assume all clients are confidential web-clients will need to add an understanding of public native app clients and the types of redirect URIs they use to support this best practice.

4.1. Authorization Flow for Native Apps Using the Browser

```
  +~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~+
  |          User Device           |
  |                                |
  | +--------------------------+   | (5) Authorization  +---------------+
  | |                          |   |     Code           |               |
  | |        Client App        |---------------------->|     Token     |
  | |                          |<----------------------|    Endpoint   |
  | +--------------------------+   | (6) Access Token,  |               |
  |   |             ^              |     Refresh Token  +---------------+
  |   |             |              |
  |   |             |              |
  |   | (1)         | (4)          |
  |   | Authorizat- | Authoriza-   |
  |   | ion Request | tion Code    |
  |   |             |              |
  |   |             |              |
  |   v             |              |
  | +---------------------------+  | (2) Authorization  +---------------+
  | |                           |  |     Request        |               |
  | |          Browser          |--------------------->| Authorization |
  | |                           |<---------------------|    Endpoint   |
  | +---------------------------+  | (3) Authorization  |               |
  |                                |     Code           +---------------+
  +~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~+
```

Figure 1: Native App Authorization via External User-agent

Figure 1 illustrates the interaction of the native app with a browser external user-agent to authorize the user.

(1) The client app opens a browser tab with the authorization request.

(2) Authorization endpoint receives the authorization request, authenticates the user and obtains authorization. Authenticating the user may involve chaining to other authentication systems.

(3) Authorization server issues an authorization code to the redirect URI.

(4) Client receives the authorization code from the redirect URI.

(5) Client app presents the authorization code at the token endpoint.

(6) Token endpoint validates the authorization code and issues the tokens requested.

5. Using Inter-app URI Communication for OAuth

Just as URIs are used for OAuth 2.0 [RFC6749] on the web to initiate the authorization request and return the authorization response to the requesting website, URIs can be used by native apps to initiate the authorization request in the device's browser and return the response to the requesting native app.

By adopting the same methods used on the web for OAuth, benefits seen in the web context like the usability of a single sign-on session and the security of a separate authentication context are likewise gained in the native app context. Re-using the same approach also reduces the implementation complexity and increases interoperability by relying on standards-based web flows that are not specific to a particular platform.

To conform to this best practice, native apps MUST use an external user-agent to perform OAuth authentication requests. This is achieved by opening the authorization request in the browser (detailed in Section 6), and using a redirect URI that will return the authorization response back to the native app, as defined in Section 7.

6. Initiating the Authorization Request from a Native App

Native apps needing user authorization create an authorization request URI with the authorization code grant type per Section 4.1 of OAuth 2.0 [RFC6749], using a redirect URI capable of being received by the native app.

The function of the redirect URI for a native app authorization request is similar to that of a web-based authorization request. Rather than returning the authorization response to the OAuth client's server, the redirect URI used by a native app returns the response to the app. Several options for a redirect URI that will return the authorization response to the native app in different platforms are documented in Section 7. Any redirect URI that allows the app to receive the URI and inspect its parameters is viable.

Public native app clients MUST implement the Proof Key for Code Exchange (PKCE [RFC7636]) extension to OAuth, and authorization servers MUST support PKCE for such clients, for the reasons detailed in Section 8.1.

After constructing the authorization request URI, the app uses platform-specific APIs to open the URI in an external user-agent. Typically the external user-agent used is the default browser, that is, the application configured for handling "http" and "https" scheme URIs on the system, but different browser selection criteria and other categories of external user-agents MAY be used.

This best practice focuses on the browser as the RECOMMENDED external user-agent for native apps. An external user-agent designed specifically for processing authorization requests capable of processing the request and redirect URIs in the same way MAY also be used. Other external user-agents, such as a native app provided by the authorization server may meet the criteria set out in this best practice, including using the same redirection URI properties, but their use is out of scope for this specification.

Some platforms support a browser feature known as in-app browser tabs, where an app can present a tab of the browser within the app context without switching apps, but still retain key benefits of the browser such as a shared authentication state and security context. On platforms where they are supported, it is RECOMMENDED for usability reasons that apps use in-app browser tabs for the authorization request.

7. Receiving the Authorization Response in a Native App

There are several redirect URI options available to native apps for receiving the authorization response from the browser, the availability and user experience of which varies by platform.

To fully support this best practice, authorization servers MUST offer at least the following three redirect URI options to native apps. Native apps MAY use whichever redirect option suits their needs best, taking into account platform specific implementation details.

7.1. Private-use URI Scheme Redirection

Many mobile and desktop computing platforms support inter-app communication via URIs by allowing apps to register private-use URI schemes (sometimes colloquially referred to as custom URL schemes) like "com.example.app". When the browser or another app attempts to load a URI with a custom scheme, the app that registered it is launched to handle the request.

To perform an OAuth 2.0 authorization request with a private-use URI scheme redirect, the native app launches the browser with a standard authorization request, but one where the redirection URI utilizes a custom URI scheme it registered with the operating system.

When choosing a URI scheme to associate with the app, apps MUST use a URI scheme based on a domain name under their control, expressed in reverse order, as recommended by Section 3.8 of [RFC7595] for private-use URI schemes.

For example, an app that controls the domain name "app.example.com" can use "com.example.app" as their scheme. Some authorization servers assign client identifiers based on domain names, for example "client1234.usercontent.example.net", which can also be used as the domain name for the scheme when reversed in the same manner. A scheme such as "myapp" however would not meet this requirement, as it is not based on a domain name.

Care must be taken when there are multiple apps by the same publisher that each scheme is unique within that group. On platforms that use app identifiers that are also based on reverse order domain names, those can be reused as the private-use URI scheme for the OAuth redirect to help avoid this problem.

Following the requirements of [RFC3986] Section 3.2, as there is no naming authority for private-use URI scheme redirects, only a single slash ("/") appears after the scheme component. A complete example of a redirect URI utilizing a private-use URI scheme:

```
com.example.app:/oauth2redirect/example-provider
```

When the authentication server completes the request, it redirects to the client’s redirection URI as it would normally. As the redirection URI uses a custom scheme it results in the operating system launching the native app, passing in the URI as a launch parameter. The native app then processes the authorization response like normal.

7.2. Claimed HTTPS URI Redirection

Some operating systems allow apps to claim HTTPS scheme [RFC7230] URIs in domains they control. When the browser encounters a claimed URI, instead of the page being loaded in the browser, the native app is launched with the URI supplied as a launch parameter.

Such URIs can be used as redirect URIs by native apps. They are indistinguishable to the authorization server from a regular web-based client redirect URI. An example is:

```
https://app.example.com/oauth2redirect/example-provider
```

As the redirect URI alone is not enough to distinguish public native app clients from confidential web clients, it is REQUIRED in Section 8.4 that the client type be recorded during client registration to enable the server to determine the client type and act accordingly.

App-claimed HTTPS redirect URIs have some advantages compared to other native app redirect options in that the identity of the destination app is guaranteed to the authorization server by the operating system. For this reason, native apps SHOULD use them over the other options where possible.

7.3. Loopback Interface Redirection

Native apps that are able to open a port on the loopback network interface without needing special permissions (typically, those on desktop operating systems) can use the loopback interface to receive the OAuth redirect.

Loopback redirect URIs use the HTTP scheme and are constructed with the loopback IP literal and whatever port the client is listening on. That is, "http://127.0.0.1:{port}/{path}" for IPv4, and "http://[::1]:{port}/{path}" for IPv6. An example redirect using the IPv4 loopback interface with a randomly assigned port:

```
http://127.0.0.1:50719/oauth2redirect/example-provider
```

An example redirect using the IPv6 loopback interface with a randomly assigned port:

```
http://[::1]:61023/oauth2redirect/example-provider
```

The authorization server MUST allow any port to be specified at the time of the request for loopback IP redirect URIs, to accommodate clients that obtain an available ephemeral port from the operating system at the time of the request.

Clients SHOULD NOT assume the device supports a particular version of the Internet Protocol. It is RECOMMENDED that clients attempt to bind to the loopback interface using both IPv4 and IPv6, and use whichever is available.

8. Security Considerations

8.1. Protecting the Authorization Code

The redirect URI options documented in Section 7 share the benefit that only a native app on the same device can receive the authorization code which limits the attack surface, however code interception by a different native app running on the same device may be possible.

A limitation of using private-use URI schemes for redirect URIs is that multiple apps can typically register the same scheme, which makes it indeterminate as to which app will receive the Authorization Code. Section 1 of PKCE [RFC7636] details how this limitation can be used to execute a code interception attack.

Loopback IP based redirect URIs may be susceptible to interception by other apps accessing the same loopback interface on some operating systems.

App-claimed HTTPS redirects are less susceptible to URI interception due to the presence of the URI authority, but they are still public clients and the URI is sent using the operating system's URI dispatch handler with unknown security properties.

The Proof Key for Code Exchange by OAuth Public Clients (PKCE [RFC7636]) standard was created specifically to mitigate against this attack. It is a proof of possession extension to OAuth 2.0 that protects the code grant from being used if it is intercepted. It achieves this by having the client generate a secret verifier, a hash of which it passes in the initial authorization request, and which it must present in full when redeeming the authorization code grant. An app that intercepted the authorization code would not be in possession of this secret, rendering the code useless.

Section 6 requires that both clients and servers use PKCE for public native app clients. Authorization servers SHOULD reject authorization requests from native apps that don't use PKCE by returning an error message as defined in Section 4.4.1 of PKCE [RFC7636].

8.2. OAuth Implicit Grant Authorization Flow

The OAuth 2.0 implicit grant authorization flow as defined in Section 4.2 of OAuth 2.0 [RFC6749] generally works with the practice of performing the authorization request in the browser, and receiving the authorization response via URI-based inter-app communication. However, as the implicit flow cannot be protected by PKCE [RFC7636] (which is a required in Section 8.1), the use of the Implicit Flow with native apps is NOT RECOMMENDED.

Tokens granted via the implicit flow also cannot be refreshed without user interaction, making the authorization code grant flow - which can issue refresh tokens - the more practical option for native app authorizations that require refreshing.

8.3. Loopback Redirect Considerations

Loopback interface redirect URIs use the "http" scheme (i.e., without TLS). This is acceptable for loopback interface redirect URIs as the HTTP request never leaves the device.

Clients should open the network port only when starting the authorization request, and close it once the response is returned.

Clients should listen on the loopback network interface only, to avoid interference by other network actors.

While redirect URIs using localhost (i.e., "http://localhost:{port}/") function similarly to loopback IP redirects described in Section 7.3, the use of "localhost" is NOT RECOMMENDED. Specifying a redirect URI with the loopback IP literal rather than localhost avoids inadvertently listening on network interfaces other than the loopback interface. It is also less susceptible to client side firewalls, and misconfigured host name resolution on the user’s device.

8.4. Registration of Native App Clients

Native apps, except when using a mechanism like Dynamic Client Registration [RFC7591] to provision per-instance secrets, are classified as public clients, as defined by Section 2.1 of OAuth 2.0 [RFC6749] and MUST be registered with the authorization server as such. Authorization servers MUST record the client type in the client registration details in order to identify and process requests accordingly.

Authorization servers MUST require clients to register their complete redirect URI (including the path component), and reject authorization requests that specify a redirect URI that doesn't exactly match the one that was registered, with the exception of loopback redirects, where an exact match is required except for the port URI component.

For private-use URI scheme based redirects, authorization servers SHOULD enforce the requirement in Section 7.1 that clients use reverse domain name based schemes. At a minimum, any scheme that doesn't contain a period character ("."), SHOULD be rejected.

In addition to the collision resistant properties, requiring a URI scheme based on a domain name that is under the control of the app can help to prove ownership in the event of a dispute where two apps claim the same private-use URI scheme (where one app is acting maliciously). For example, if two apps claimed "com.example.app", the owner of "example.com" could petition the app store operator to remove the counterfeit app. Such a petition is harder to prove if a generic URI scheme was used.

Authorization servers MAY request the inclusion of other platform-specific information, such as the app package or bundle name, or other information used to associate the app that may be useful for verifying the calling app's identity, on operating systems that support such functions.

8.5. Client Authentication

Secrets that are statically included as part of an app distributed to multiple users should not be treated as confidential secrets, as one user may inspect their copy and learn the shared secret. For this reason, and those stated in Section 5.3.1 of [RFC6819], it is NOT RECOMMENDED for authorization servers to require client authentication of public native apps clients using a shared secret, as this serves little value beyond client identification which is already provided by the "client_id" request parameter.

Authorization servers that still require a statically included shared secret for native app clients MUST treat the client as a public client (as defined by Section 2.1 of OAuth 2.0 [RFC6749]), and not accept the secret as proof of the client's identity. Without additional measures, such clients are subject to client impersonation (see Section 8.6).

8.6. Client Impersonation

As stated in Section 10.2 of OAuth 2.0 [RFC6749], the authorization server SHOULD NOT process authorization requests automatically without user consent or interaction, except when the identity of the client can be assured. This includes the case where the user has previously approved an authorization request for a given client id - unless the identity of the client can be proven, the request SHOULD be processed as if no previous request had been approved.

Measures such as claimed HTTPS redirects MAY be accepted by authorization servers as identity proof. Some operating systems may offer alternative platform-specific identity features which MAY be accepted, as appropriate.

8.7. Fake External User-Agent

The native app which is initiating the authorization request has a large degree of control over the user interface and can potentially present a fake external user-agent, that is, an embedded user-agent made to appear as an external user agent.

The advantage when all good actors are using external user-agents is that it is possible for security experts to detect bad actors, as anyone faking an external user-agent is provably bad. If good and bad actors alike are using embedded user-agents, bad actors don’t need to fake anything, making them harder to detect. Once malicious apps are detected, it may be possible to use this knowledge to blacklist the apps signatures in malware scanning software, take removal action in the case of apps distributed by app stores, and other steps to reduce the impact and spread of the malicious app.

Authorization servers can also directly protect against fake external user-agents by requiring an authentication factor only available to true external user-agents.

Users who are particularly concerned about their security when using in-app browser tabs may also take the additional step of opening the request in the full browser from the in-app browser tab, and complete the authorization there, as most implementations of the in-app browser tab pattern offer such functionality.

8.8. Malicious External User-Agent

If a malicious app is able to configure itself as the default handler for "https" scheme URIs in the operating system, it will be able to intercept authorization requests that use the default browser and abuse this position of trust for malicious ends such as phishing the user.

Many operating systems mitigate this issue by requiring an explicit user action to change the default handler for HTTP URIs. This attack is not confined to OAuth for Native Apps, a malicious app configured in this way would present a general and ongoing risk to the user beyond OAuth usage.

8.9. Cross-App Request Forgery Protections

Section 5.3.5 of [RFC6819] recommends using the "state" parameter to link client requests and responses to prevent CSRF (Cross Site Request Forgery) attacks.

To mitigate CSRF style attacks using inter-app URI communication, it is similarly RECOMMENDED that native apps include a high entropy secure random number in the "state" parameter of the authorization request, and reject any incoming authorization responses without a state value that matches a pending outgoing authorization request.

8.10. Authorization Server Mix-Up Mitigation

To protect against a compromised or malicious authorization server attacking another authorization server used by the same app, it is REQUIRED that a unique redirect URI is used for each authorization server used by the app (for example, by varying the path component), and that authorization responses are rejected if the redirect URI they were received on doesn't match the redirect URI in a outgoing authorization request.

The native app MUST store the redirect URI used in the authorization request with the authorization session data (i.e., along with "state" and other related data), and MUST verify that the URI on which the authorization response was received exactly matches it.

The requirements of Section 8.4 that authorization servers reject requests with URIs that don't match what was registered are also required to prevent such attacks.

8.11. Non-Browser External User-Agents

This best practice recommends a particular type of external user-agent, the user's browser. Other external user-agent patterns may also be viable for secure and usable OAuth. This document makes no comment on those patterns.

8.12. Embedded User-Agents

OAuth 2.0 [RFC6749] Section 9 documents two approaches for native apps to interact with the authorization endpoint. This best current practice requires that native apps MUST NOT use embedded user-agents to perform authorization requests, and allows that authorization endpoints MAY take steps to detect and block authorization requests in embedded user-agents. The security considerations for these requirements are detailed herein.

Embedded user-agents are an alternative method for authorizing native apps. These embedded user agents are unsafe for use by third-parties to the authorization server by definition, as the app that hosts the embedded user-agent can access the user’s full authentication credential, not just the OAuth authorization grant that was intended for the app.

In typical web-view based implementations of embedded user-agents, the host application can: log every keystroke entered in the form to capture usernames and passwords; automatically submit forms and bypass user-consent; copy session cookies and use them to perform authenticated actions as the user.

Even when used by trusted apps belonging to the same party as the authorization server, embedded user-agents violate the principle of least privilege by having access to more powerful credentials than they need, potentially increasing the attack surface.

Encouraging users to enter credentials in an embedded user-agent without the usual address bar and visible certificate validation features that browsers have makes it impossible for the user to know if they are signing in to the legitimate site, and even when they are, it trains them that it’s OK to enter credentials without validating the site first.

Aside from the security concerns, embedded user-agents do not share the authentication state with other apps or the browser, requiring the user to login for every authorization request which is often considered an inferior user experience.

9. IANA Considerations

[RFC Editor: please do NOT remove this section.]

This document has no IANA actions.

Section 7.1 specifies how private-use URI schemes are used for inter-app communication in OAuth protocol flows. This document requires in Section 7.1 that such schemes are based on domain names owned or assigned to the app, as recommended in Section 3.8 of [RFC7595]. Per Section 6 of [RFC7595], registration of domain based URI schemes with IANA is not required.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, DOI 10.17487/RFC3986, January 2005, <http://www.rfc-editor.org/info/rfc3986>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <http://www.rfc-editor.org/info/rfc6749>.

[RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, DOI 10.17487/RFC7230, June 2014, <http://www.rfc-editor.org/info/rfc7230>.

[RFC7595] Thaler, D., Ed., Hansen, T., and T. Hardie, "Guidelines and Registration Procedures for URI Schemes", BCP 35, RFC 7595, DOI 10.17487/RFC7595, June 2015, <http://www.rfc-editor.org/info/rfc7595>.

[RFC7636] Sakimura, N., Ed., Bradley, J., and N. Agarwal, "Proof Key for Code Exchange by OAuth Public Clients", RFC 7636, DOI 10.17487/RFC7636, September 2015, <http://www.rfc-editor.org/info/rfc7636>.

10.2. Informative References

[RFC6819] Lodderstedt, T., Ed., McGloin, M., and P. Hunt, "OAuth 2.0 Threat Model and Security Considerations", RFC 6819, DOI 10.17487/RFC6819, January 2013, <http://www.rfc-editor.org/info/rfc6819>.

[RFC7591] Richer, J., Ed., Jones, M., Bradley, J., Machulak, M., and P. Hunt, "OAuth 2.0 Dynamic Client Registration Protocol", RFC 7591, DOI 10.17487/RFC7591, July 2015, <http://www.rfc-editor.org/info/rfc7591>.

[AppAuth] Denniss, W., Wright, S., McGinniss, I., Ravikumar, R., and others, "AppAuth", May 22, <https://appauth.io>.

[AppAuth.iOSmacOS]
Wright, S., Denniss, W., and others, "AppAuth for iOS and macOS", February 2016, <https://github.com/openid/AppAuth-iOS>.

[AppAuth.Android]
McGinniss, I., Denniss, W., and others, "AppAuth for Android", February 2016, <https://github.com/openid/AppAuth-Android>.

[SamplesForWindows]
Denniss, W., "OAuth for Apps: Samples for Windows", July 2016, <https://github.com/googlesamples/oauth-apps-for-windows>.

Appendix A. Server Support Checklist

OAuth servers that support native apps must:

1. Support private-use URI scheme redirect URIs. This is required to support mobile operating systems. See Section 7.1.

2. Support HTTPS scheme redirect URIs for use with public native app clients. This is used by apps on advanced mobile operating systems that allow app-claimed URIs. See Section 7.2.

3. Support loopback IP redirect URIs. This is required to support desktop operating systems. See Section 7.3.

4. Not assume native app clients can keep a secret. If secrets are distributed to multiple installs of the same native app, they should not be treated as confidential. See Section 8.5.

5. Support PKCE [RFC7636]. Required to protect authorization code grants sent to public clients over inter-app communication channels. See Section 8.1

Appendix B. Operating System Specific Implementation Details

This document primarily defines best practices in a generic manner, referencing techniques commonly available in a variety of environments. This non-normative section documents operating system specific implementation details of the best practice.

The implementation details herein are considered accurate at the time of publishing but will likely change over time. It is hoped that such change won’t invalidate the generic principles in the rest of the document, and those principles should take precedence in the event of a conflict.

B.1. iOS Implementation Details

Apps can initiate an authorization request in the browser without the user leaving the app, through the SFSafariViewController class which implements the in-app browser tab pattern. Safari can be used to handle requests on old versions of iOS without SFSafariViewController.

To receive the authorization response, both private-use URI scheme redirects (referred to as Custom URL Schemes) and claimed HTTPS links (known as Universal Links) are viable choices, and function the same whether the request is loaded in SFSafariViewController or the Safari app. Apps can claim Custom URI schemes with the "CFBundleURLTypes" key in the application's property list file "Info.plist", and HTTPS links using the Universal Links feature with an entitlement file and an association file on the domain.

Universal Links are the preferred choice on iOS 9 and above due to the ownership proof that is provided by the operating system.

A complete open source sample is included in the AppAuth for iOS and macOS [AppAuth.iOSmacOS] library.

B.2. Android Implementation Details

Apps can initiate an authorization request in the browser without the user leaving the app, through the Android Custom Tab feature which implements the in-app browser tab pattern. The user's default browser can be used to handle requests when no browser supports Custom Tabs.

Android browser vendors should support the Custom Tabs protocol (by providing an implementation of the "CustomTabsService" class), to provide the in-app browser tab user experience optimization to their users. Chrome is one such browser that implements Custom Tabs.

To receive the authorization response, private-use URI schemes are broadly supported through Android Implicit Intents. Claimed HTTPS redirect URIs through Android App Links are available on Android 6.0 and above. Both types of redirect URIs are registered in the application's manifest.

A complete open source sample is included in the AppAuth for Android [AppAuth.Android] library.

B.3. Windows Implementation Details

Both traditional and Universal Windows Platform (UWP) apps can perform authorization requests in the user's browser. Traditional apps typically use a loopback redirect to receive the authorization response, and listening on the loopback interface is allowed by default firewall rules. When creating the loopback network socket, apps SHOULD set the "SO_EXCLUSIVEADDRUSE" socket option to prevent other apps binding to the same socket.

UWP apps can use private-use URI scheme redirects to receive the authorization response from the browser, which will bring the app to the foreground. Known on the platform as "URI Activation", the URI scheme is limited to 39 characters in length, and may include the "." character, making short reverse domain name based schemes (as recommended in Section 7.1) possible.

UWP apps can alternatively use the Web Authentication Broker API in SSO (Single Sign-on) mode, which is an external user agent designed for authorization flows. Cookies are shared between invocations of the broker but not the user's preferred browser, meaning the user will need to sign-in again even if they have an active session in their browser, but the session created in the broker will be available to subsequent apps that use the broker. Personalisations the user has made to their browser, such as configuring a password manager may not available in the broker. To qualify as an external user-agent, the broker MUST be used in SSO mode.

To use the Web Authentication Broker in SSO mode, the redirect URI must be of the form "msapp://{appSID}" where "appSID" is the app's SID, which can be found in the app's registration information. While Windows enforces the URI authority on such redirects, ensuring only the app with the matching SID can receive the response on Windows, the URI scheme could be claimed by apps on other platforms without the same authority present, thus this redirect type should be treated similar to private-use URI scheme redirects for security purposes.

An open source sample demonstrating these patterns is available [SamplesForWindows].

B.4. macOS Implementation Details

Apps can initiate an authorization request in the user's default browser using platform APIs for opening URIs in the browser.

To receive the authorization response, private-use URI schemes are a good redirect URI choice on macOS, as the user is returned right back to the app they launched the request from. These are registered in the application's bundle information property list using the "CFBundleURLSchemes" key. Loopback IP redirects are another viable option, and listening on the loopback interface is allowed by default firewall rules.

A complete open source sample is included in the AppAuth for iOS and macOS [AppAuth.iOSmacOS] library.

B.5. Linux Implementation Details

Opening the Authorization Request in the user's default browser requires a distro-specific command, "xdg-open" is one such tool.

The loopback redirect is the recommended redirect choice for desktop apps on Linux to receive the authorization response. Apps SHOULD NOT set the "SO_REUSEPORT" or "SO_REUSEADDR" socket options, to prevent other apps binding to the same socket.

Appendix C. Acknowledgements

The author would like to acknowledge the work of Marius Scurtescu, and Ben Wiley Sittler whose design for using private-use URI schemes in native OAuth 2.0 clients at Google formed the basis of Section 7.1.

The following individuals contributed ideas, feedback, and wording that shaped and formed the final specification:

Andy Zmolek, Steven E Wright, Brian Campbell, Nat Sakimura, Eric Sachs, Paul Madsen, Iain McGinniss, Rahul Ravikumar, Breno de Medeiros, Hannes Tschofenig, Ashish Jain, Erik Wahlstrom, Bill Fisher, Sudhi Umarji, Michael B. Jones, Vittorio Bertocci, Dick Hardt, David Waite, Ignacio Fiorentino, Kathleen Moriarty, and Elwyn Davies.

Authors' Addresses

William Denniss
Google
1600 Amphitheatre Pkwy
Mountain View, CA 94043
USA

Email: wdenniss@google.com
URI: http://wdenniss.com/appauth

John Bradley
Ping Identity

Phone: +1 202-630-5272
Email: ve7jtb@ve7jtb.com
URI: http://www.thread-safe.com/p/appauth.html

OAuth
Internet-Draft
Intended status: Informational
Expires: January 9, 2017

P. Hunt, Ed.
Oracle Corporation
J. Richer

W. Mills

P. Mishra
Oracle Corporation
H. Tschofenig
ARM Limited
July 8, 2016

OAuth 2.0 Proof-of-Possession (PoP) Security Architecture
draft-ietf-oauth-pop-architecture-08.txt

Abstract

The OAuth 2.0 bearer token specification, as defined in RFC 6750, allows any party in possession of a bearer token (a "bearer") to get access to the associated resources (without demonstrating possession of a cryptographic key). To prevent misuse, bearer tokens must be protected from disclosure in transit and at rest.

Some scenarios demand additional security protection whereby a client needs to demonstrate possession of cryptographic keying material when accessing a protected resource. This document motivates the development of the OAuth 2.0 proof-of-possession security mechanism.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The OAuth 2.0 protocol family ([RFC6749], [RFC6750], and [RFC6819]) offer a single token type known as the "bearer" token to access protected resources. RFC 6750 [RFC6750] specifies the bearer token mechanism and defines it as follows:

> "A security token with the property that any party in possession of the token (a "bearer") can use the token in any way that any other party in possession of it can. Using a bearer token does not require a bearer to prove possession of cryptographic key material."

The bearer token meets the security needs of a number of use cases the OAuth 2.0 protocol had originally been designed for. There are, however, other scenarios that require stronger security properties and ask for active participation of the OAuth client in form of cryptographic computations when presenting an access token to a resource server.

This document outlines additional use cases requiring stronger security protection in Section 3, identifies threats in Section 4, proposes different ways to mitigate those threats in Section 6, outlines an architecture for a solution that builds on top of the existing OAuth 2.0 framework in Section 7, and concludes with a requirements list in Section 5.

2. Terminology

The key words 'MUST', 'MUST NOT', 'REQUIRED', 'SHALL', 'SHALL NOT', 'SHOULD', 'SHOULD NOT', 'RECOMMENDED', 'MAY', and 'OPTIONAL' in this specification are to be interpreted as described in [RFC2119], with the important qualification that, unless otherwise stated, these terms apply to the design of the protocol, not its implementation or application.

3. Use Cases

The main use case that motivates improvement upon "bearer" token security is the desire of resource servers to obtain additional assurance that the client is indeed authorized to present an access token. The expectation is that the use of additional credentials (symmetric or asymmetric keying material) will encourage developers to take additional precautions when transferring and storing access token in combination with these credentials.

Additional use cases listed below provide further requirements for the solution development. Note that a single solution does not necessarily need to offer support for all use cases.

3.1. Preventing Access Token Re-Use by the Resource Server

In a scenario where a resource server receives a valid access token, the resource server then re-uses it with other resource server. The reason for re-use may be malicious or may well be legitimate. In a legitimate case, the intent is to support chaining of computations whereby a resource server needs to consult other third party resource servers to complete a requested operation. In both cases it may be assumed that the scope and audience of the access token is sufficiently defined that to allow such a re-use. For example, imagine a case where a company operates email services as well as picture sharing services and that company had decided to issue access tokens with a scope and audience that allows access to both services.

With this use case the desire is to prevent such access token re-use. This also implies that the legitimate use cases require additional enhancements for request chaining.

3.2. TLS and DTLS Channel Binding Support

In this use case we consider the scenario where an OAuth 2.0 request to a protected resource is secured using TLS or DTLS (see [RFC4347]), but the client and the resource server demand that the underlying TLS/DTLS exchange is bound to additional application layer security to prevent cases where the TLS/DTLS connection is terminated at a TLS/DTLS intermediary, which splits the TLS/DTLS connection into two separate connections.

In this use case additional information should be conveyed to the resource server to ensure that no entity entity has tampered with the TLS/DTLS connection.

3.3. Access to a Non-TLS Protected Resource

This use case is for a web client that needs to access a resource that makes data available (such as videos) without offering integrity and confidentiality protection using TLS. Still, the initial resource request using OAuth, which includes the access token, must be protected against various threats (e.g., token replay, token modification).

While it is possible to utilize bearer tokens in this scenario with TLS protection when the request to the protected resource is made, as described in [RFC6750], there may be the desire to avoid using TLS between the client and the resource server at all. In such a case the bearer token approach is not possible since it relies on TLS for ensuring integrity and confidentiality protection of the access token exchange since otherwise replay attacks are possible: First, an eavesdropper may steal an access token and present it at a different resource server. Second, an eavesdropper may steal an access token and replay it against the same resource server at a later point in time. In both cases, if the attack is successful, the adversary gets access to the resource owners data or may perform an operation selected by the adversary (e.g., sending a message). Note that the adversary may obtain the access token (if the recommendations in [RFC6749] and [RFC6750] are not followed) using a number of ways, including eavesdropping the communication on the wireless link.

Consequently, the important assumption in this use case is that a resource server does not have TLS support and the security solution should work in such a scenario. Furthermore, it may not be necessary to provide authentication of the resource server towards the client.

3.4. Offering Application Layer End-to-End Security

In Web deployments resource servers are often placed behind load balancers, which are deployed by the same organization that operates the resource servers. These load balancers may terminate the TLS connection setup and HTTP traffic is transmitted without TLS protection from the load balancer to the resource server. With application layer security in addition to the underlying TLS security it is possible to allow application servers to perform cryptographic verification on an end-to-end basis.

The key aspect in this use case is therefore to offer end-to-end security in the presence of load balancers via application layer security. Enterprise networks also deploy proxies that inspect traffic and thereby break TLS.

4. Security and Privacy Threats

The following list presents several common threats against protocols utilizing some form of token. This list of threats is based on NIST Special Publication 800-63 [NIST800-63]. We exclude a discussion of threats related to any form of identity proofing and authentication of the resource owner to the authorization server since these procedures are not part of the OAuth 2.0 protocol specification itself.

Token manufacture/modification:

An attacker may generate a bogus token or modify the token content (such as authentication or attribute statements) of an existing token, causing resource server to grant inappropriate access to the client. For example, an attacker may modify the token to extend the validity period. A client, which MAY be a normal client or MAY be assumed to be constrained (see [RFC7252]), may modify the token to have access to information that they should not be able to view.

Token disclosure:

Tokens may contain personal data, such as real name, age or birthday, payment information, etc.

Token redirect:

An attacker uses the token generated for consumption by the resource server to obtain access to another resource server.

Token reuse:

An attacker attempts to use a token that has already been used once with a resource server. The attacker may be an eavesdropper who observes the communication exchange or, worse, one of the communication end points. A client may, for example, leak access tokens because it cannot keep secrets confidential. A client may also reuse access tokens for some other resource servers. Finally, a resource server may use a token it had obtained from a client and use it with another resource server that the client interacts with. A resource server, offering relatively unimportant application services, may attempt to use an access token obtained from a client to access a high-value service, such as a payment service, on behalf of the client using the same access token.

Token repudiation:

Token repudiation refers to a property whereby a resource server is given an assurance that the authorization server cannot deny to have created a token for the client.

5. Requirements

RFC 4962 [RFC4962] gives useful guidelines for designers of authentication and key management protocols. While RFC 4962 was written with the AAA framework used for network access authentication in mind the offered suggestions are useful for the design of other key management systems as well. The following requirements list applies OAuth 2.0 terminology to the requirements outlined in RFC 4962.

These requirements include

Cryptographic Algorithm Independent:

The key management protocol MUST be cryptographic algorithm independent.

Strong, fresh session keys:

Session keys MUST be strong and fresh. Each session deserves an independent session key, i.e., one that is generated specifically for the intended use. In context of OAuth this means that keying material is created in such a way that can only be used by the combination of a client instance, protected resource, and authorization scope.

Limit Key Scope:

Following the principle of least privilege, parties MUST NOT have access to keying material that is not needed to perform their role. Any protocol that is used to establish session keys MUST specify the scope for session keys, clearly identifying the parties to whom the session key is available.

Replay Detection Mechanism:

The key management protocol exchanges MUST be replay protected. Replay protection allows a protocol message recipient to discard any message that was recorded during a previous legitimate dialogue and presented as though it belonged to the current dialogue.

Authenticate All Parties:

Each party in the key management protocol MUST be authenticated to the other parties with whom they communicate. Authentication mechanisms MUST maintain the confidentiality of any secret values used in the authentication process. Secrets MUST NOT be sent to another party without confidentiality protection.

Authorization:

Client and resource server authorization MUST be performed. These entities MUST demonstrate possession of the appropriate keying material, without disclosing it. Authorization is REQUIRED whenever a client interacts with an authorization server. Authorization checking prevents an elevation of privilege attack.

Keying Material Confidentiality and Integrity:

While preserving algorithm independence, confidentiality and integrity of all keying material MUST be maintained.

Confirm Cryptographic Algorithm Selection:

The selection of the "best" cryptographic algorithms SHOULD be securely confirmed. The mechanism SHOULD detect attempted roll-back attacks.

Uniquely Named Keys:

Key management proposals require a robust key naming scheme, particularly where key caching is supported. The key name provides a way to refer to a key in a protocol so that it is clear to all parties which key is being referenced. Objects that cannot be named cannot be managed. All keys MUST be uniquely named, and the key name MUST NOT directly or indirectly disclose the keying material.

Prevent the Domino Effect:

Compromise of a single client MUST NOT compromise keying material held by any other client within the system, including session keys and long-term keys. Likewise, compromise of a single resource server MUST NOT compromise keying material held by any other Resource Server within the system. In the context of a key hierarchy, this means that the compromise of one node in the key hierarchy must not disclose the information necessary to compromise other branches in the key hierarchy. Obviously, the compromise of the root of the key hierarchy will compromise all of the keys; however, a compromise in one branch MUST NOT result in the compromise of other branches. There are many implications of this requirement; however, two implications deserve highlighting. First, the scope of the keying material must be defined and understood by all parties that communicate with a party that holds that keying material. Second, a party that holds keying material in a key hierarchy must not share that keying material with parties that are associated with other branches in the key hierarchy.

Bind Key to its Context:

Keying material MUST be bound to the appropriate context. The context includes the following.

* The manner in which the keying material is expected to be used.

* The other parties that are expected to have access to the keying material.

* The expected lifetime of the keying material. Lifetime of a child key SHOULD NOT be greater than the lifetime of its parent in the key hierarchy.

Any party with legitimate access to keying material can determine its context. In addition, the protocol MUST ensure that all parties with legitimate access to keying material have the same context for the keying material. This requires that the parties are properly identified and authenticated, so that all of the parties that have access to the keying material can be determined. The context will include the client and the resource server identities in more than one form.

Authorization Restriction:

If client authorization is restricted, then the client SHOULD be made aware of the restriction.

Client Identity Confidentiality:

A client has identity confidentiality when any party other than the resource server and the authorization server cannot sufficiently identify the client within the anonymity set. In comparison to anonymity and pseudonymity, identity confidentiality is concerned with eavesdroppers and intermediaries. A key management protocol SHOULD provide this property.

Resource Owner Identity Confidentiality:

Resource servers SHOULD be prevented from knowing the real or pseudonymous identity of the resource owner, since the authorization server is the only entity involved in verifying the resource owner’s identity.

Collusion:

Resource servers that collude can be prevented from using information related to the resource owner to track the individual. That is, two different resource servers can be prevented from determining that the same resource owner has authenticated to both of them. Authorization servers MUST bind different keying material to access tokens used for resource servers from different origins (or similar concepts in the app world).

AS-to-RS Relationship Anonymity:

For solutions using asymmetric key cryptography the client MAY conceal information about the resource server it wants to interact with. The authorization server MAY reject such an attempt since it may not be able to enforce access control decisions.

Channel Binding:

A solution MUST enable support for channel bindings. The concept of channel binding, as defined in [RFC5056], allows applications to establish that the two end-points of a secure channel at one network layer are the same as at a higher layer by binding authentication at the higher layer to the channel at the lower layer.

There are performance concerns with the use of asymmetric cryptography. Although symmetric key cryptography offers better performance asymmetric cryptography offers additional security properties. A solution MUST therefore offer the capability to support both symmetric as well as asymmetric keys.

There are threats that relate to the experience of the software developer as well as operational practices. Verifying the servers identity in TLS is discussed at length in [RFC6125].

A number of the threats listed in Section 4 demand protection of the access token content and a standardized solution, for example, in the form of a JSON-based format, is available with the JWT [RFC7519].

6. Threat Mitigation

A large range of threats can be mitigated by protecting the content of the token, for example using a digital signature or a keyed message digest. Alternatively, the content of the token could be passed by reference rather than by value (requiring a separate message exchange to resolve the reference to the token content).

To simplify discussion in the following example we assume that the token itself cannot be modified by the client, either due to cryptographic protection (such as signature or encryption) or use of a reference value with sufficient entropy and associated secure lookup. The token remains opaque to the client. These are characteristics shared with bearer tokens and more information on best practices can be found in [RFC6819] and in the security considerations section of [RFC6750].

To deal with token redirect it is important for the authorization server to include the identifier of the intended recipient - the resource server. A resource server must not be allowed to accept access tokens that are not meant for its consumption.

To provide protection against token disclosure two approaches are possible, namely (a) not to include sensitive information inside the token or (b) to ensure confidentiality protection. The latter approach requires at least the communication interaction between the client and the authorization server as well as the interaction between the client and the resource server to experience confidentiality protection. As an example, TLS with a ciphersuite that offers confidentiality protection has to be applied as per [RFC7525]. Encrypting the token content itself is another alternative. In our scenario the authorization server would, for example, encrypt the token content with a symmetric key shared with the resource server.

To deal with token reuse more choices are available.

6.1. Confidentiality Protection

In this approach confidentiality protection of the exchange is provided on the communication interfaces between the client and the resource server, and between the client and the authorization server. No eavesdropper on the wire is able to observe the token exchange. Consequently, a replay by a third party is not possible. An authorization server wants to ensure that it only hands out tokens to clients it has authenticated first and who are authorized. For this purpose, authentication of the client to the authorization server will be a requirement to ensure adequate protection against a range of attacks. This is, however, true for the description in Section 6.2 and Section 6.3 as well. Furthermore, the client has to make sure it does not distribute (or leak) the access token to entities other than the intended the resource server. For that purpose the client will have to authenticate the resource server before transmitting the access token.

6.2. Sender Constraint

Instead of providing confidentiality protection, the authorization server could also put the identifier of the client into the protected token with the following semantic: 'This token is only valid when presented by a client with the following identifier.' When the access token is then presented to the resource server how does it know that it was provided by the client? It has to authenticate the client! There are many choices for authenticating the client to the resource server, for example by using client certificates in TLS [RFC5246], or pre-shared secrets within TLS [RFC4279]. The choice of the preferred authentication mechanism and credential type may depend on a number of factors, including

- o security properties
- o available infrastructure
- o library support
- o credential cost (financial)
- o performance
- o integration into the existing IT infrastructure
- o operational overhead for configuration and distribution of credentials

This long list hints to the challenge of selecting at least one mandatory-to-implement client authentication mechanism.

6.3. Key Confirmation

A variation of the mechanism of sender authentication, described in Section 6.2, is to replace authentication with the proof-of-possession of a specific (session) key, i.e., key confirmation. In this model the resource server would not authenticate the client itself but would rather verify whether the client knows the session key associated with a specific access token. Examples of this approach can be found with the OAuth 1.0 MAC token [RFC5849], and Kerberos [RFC4120] when utilizing the AP_REQ/AP_REP exchange (see also [I-D.hardjono-oauth-kerberos] for a comparison between Kerberos and OAuth).

To illustrate key confirmation, the first example is borrowed from Kerberos and use symmetric key cryptography. Assume that the authorization server shares a long-term secret with the resource server, called K(Authorization Server-Resource Server). This secret would be established between them out-of-band. When the client requests an access token the authorization server creates a fresh and unique session key Ks and places it into the token encrypted with the long term key K(Authorization Server-Resource Server). Additionally, the authorization server attaches Ks to the response message to the client (in addition to the access token itself) over a confidentiality protected channel. When the client sends a request to the resource server it has to use Ks to compute a keyed message digest for the request (in whatever form or whatever layer). The resource server, when receiving the message, retrieves the access token, verifies it and extracts K(Authorization Server-Resource Server) to obtain Ks. This key Ks is then used to verify the keyed message digest of the request message.

Note that in this example one could imagine that the mechanism to protect the token itself is based on a symmetric key based mechanism to avoid any form of public key infrastructure but this aspect is not further elaborated in the scenario.

A similar mechanism can also be designed using asymmetric cryptography. When the client requests an access token the authorization server creates an ephemeral public / privacy key pair (PK/SK) and places the public key PK into the protected token. When the authorization server returns the access token to the client it also provides the PK/SK key pair over a confidentiality protected channel. When the client sends a request to the resource server it has to use the privacy key SK to sign the request. The resource server, when receiving the message, retrieves the access token, verifies it and extracts the public key PK. It uses this ephemeral public key to verify the attached signature.

6.4. Summary

As a high level message, there are various ways the threats can be mitigated. While the details of each solution are somewhat different, they all accomplish the goal of mitigating the threats.

The three approaches are:

Confidentiality Protection:

The weak point with this approach, which is briefly described in Section 6.1, is that the client has to be careful to whom it discloses the access token. What can be done with the token entirely depends on what rights the token entitles the presenter and what constraints it contains. A token could encode the identifier of the client but there are scenarios where the client is not authenticated to the resource server or where the identifier of the client rather represents an application class rather than a single application instance. As such, it is possible that certain deployments choose a rather liberal approach to security and that everyone who is in possession of the access token is granted access to the data.

Sender Constraint:

The weak point with this approach, which is briefly described in Section 6.2, is to setup the authentication infrastructure such that clients can be authenticated towards resource servers. Additionally, the authorization server must encode the identifier of the client in the token for later verification by the resource server. Depending on the chosen layer for providing client-side authentication there may be additional challenges due to Web server load balancing, lack of API access to identity information, etc.

Key Confirmation:

The weak point with this approach, see Section 6.3, is the increased complexity: a complete key distribution protocol has to be defined.

In all cases above it has to be ensured that the client is able to keep the credentials secret.

7. Architecture

The proof-of-possession security concept assumes that the authorization server acts as a trusted third party that binds keys to access tokens. These keys are then used by the client to demonstrate the possession of the secret to the resource server when accessing the resource. The resource server, when receiving an access token, needs to verify that the key used by the client matches the one included in the access token.

There are slight differences between the use of symmetric keys and asymmetric keys when they are bound to the access token and the subsequent interaction between the client and the authorization server when demonstrating possession of these keys. Figure 1 shows the symmetric key procedure and Figure 2 illustrates how asymmetric keys are used. While symmetric cryptography provides better performance properties the use of asymmetric cryptography allows the client to keep the private key locally and never expose it to any other party.

For example, with the JSON Web Token (JWT) [RFC7519] a standardized format for access tokens is available. The necessary elements to bind symmetric or asymmetric keys to a JWT are described in [I-D.ietf-oauth-proof-of-possession].

Note: The negotiation of cryptographic algorithms between the client and the authorization server is not shown in the examples below and assumed to be present in a protocol solution to meet the requirements for crypto-agility.

7.1. Client and Authorization Server Interaction

7.1.1. Symmetric Keys

```
                        +---------------+
                       ^|               |
                     // | Authorization |
                    /   | Server        |
                  //    |               |
                 /      |               |
          (I)  //      /+---------------+
   Access     /      //
   Token     /      /
   Request //     //  (II) Access Token
   +Params /      /        +Symmetric Key
        //     //
       /      v
     +-----------+                       +------------+
     |           |                       |            |
     |           |                       | Resource   |
     | Client    |                       | Server     |
     |           |                       |            |
     |           |                       |            |
     +-----------+                       +------------+
```

Figure 1: Interaction between the Client and the Authorization Server (Symmetric Keys).

In order to request an access token the client interacts with the authorization server as part of the a normal grant exchange, as shown in Figure 1. However, it needs to include additional information elements for use with the PoP security mechanism, as depicted in message (I). In message (II) the authorization server then returns the requested access token. In addition to the access token itself, the symmetric key is communicated to the client. This symmetric key is a unique and fresh session key with sufficient entropy for the given lifetime. Furthermore, information within the access token ties it to this specific symmetric key.

Note: For this security mechanism to work the client as well as the resource server need to have access to the session key. While the key transport mechanism from the authorization server to the client has been explained in the previous paragraph there are three ways for communicating this session key from the authorization server to the resource server, namely

Embedding the symmetric key inside the access token itself. This requires that the symmetric key is confidentiality protected.

The resource server queries the authorization server for the symmetric key. This is an approach envisioned by the token introspection endpoint [RFC7662].

The authorization server and the resource server both have access to the same back-end database. Smaller, tightly coupled systems might prefer such a deployment strategy.

7.1.2. Asymmetric Keys

```
                        +---------------+
                       ^|               |
   Access Token Req. // | Authorization |
   +Parameters      /   | Server        |
   +[Fingerprint] //    |               |
                 /      |               |
       (I)     //      /+---------------+
              /      //
             /      /     (II)
           //     //  Access Token
           /      /   +[ephemeral
        //     //      asymmetric key pair]
       /      v
     +-----------+                     +------------+
     |           |                     |            |
     |           |                     | Resource   |
     | Client    |                     | Server     |
     |           |                     |            |
     |           |                     |            |
     +-----------+                     +------------+
```

Figure 2: Interaction between the Client and the Authorization Server (Asymmetric Keys).

The use of asymmetric keys is slightly different since the client or the server could be involved in the generation of the ephemeral key pair. This exchange is shown in Figure 1. If the client generates the key pair it either includes a fingerprint of the public key or the public key in the request to the authorization server. The authorization server would include this fingerprint or public key in the confirmation claim inside the access token and thereby bind the asymmetric key pair to the token. If the client did not provide a fingerprint or a public key in the request then the authorization server is asked to create an ephemeral asymmetric key pair, binds the fingerprint of the public key to the access token, and returns the asymmetric key pair (public and private key) to the client. Note that there is a strong preference for generating the private/public key pair locally at the client rather than at the server.

7.2. Client and Resource Server Interaction

The specification describing the interaction between the client and the authorization server, as shown in Figure 1 and in Figure 2, can be found in [I-D.ietf-oauth-pop-key-distribution].

Once the client has obtained the necessary access token and keying material it can start to interact with the resource server. To demonstrate possession of the key bound to the access token it needs to apply this key to the request by computing a keyed message digest (i.e., a symmetric key-based cryptographic primitive) or a digital signature (i.e., an asymmetric cryptographic computation). When the resource server receives the request it verifies it and decides whether access to the protected resource can be granted. This exchange is shown in Figure 3.

```
                             +--------------+
                             |              |
                             | Authorization |
                             | Server       |
                             |              |
                             |              |
                             +--------------+




                                 Request
      +-----------+  + Signature/MAC (a)  +------------+
      |           |---------------------->|            |
      |           |  [+Access Token]      | Resource   |
      | Client    |                       | Server     |
      |           |    Response (b)       |            |
      |           |<----------------------|            |
      +-----------+  [+ Signature/MAC]    +------------+

           ^                                     ^
           |                                     |
           |                                     |
       Symmetric Key                      Symmetric Key
          or                                  or
       Asymmetric Key Pair              Public Key (Client)
          +                                   +
        Parameters                         Parameters
```

Figure 3: Client Demonstrates PoP.

The specification describing the ability to sign the HTTP request from the client to the resource server can be found in [I-D.ietf-oauth-signed-http-request].

7.3. Resource and Authorization Server Interaction (Token Introspection)

So far the examples talked about access tokens that are passed by value and allow the resource server to make authorization decisions immediately after verifying the request from the client. In some deployments a real-time interaction between the authorization server and the resource server is envisioned that lowers the need to pass self-contained access tokens around. In that case the access token merely serves as a handle or a reference to state stored at the authorization server. As a consequence, the resource server cannot autonomously make an authorization decision when receiving a request from a client but has to consult the authorization server. This can, for example, be done using the token introspection endpoint (see [RFC7662]). Figure 4 shows the protocol interaction graphically. Despite the additional token exchange previous descriptions about associating symmetric and asymmetric keys to the access token are still applicable to this scenario.

```
                      +---------------+
        Access       ^|               |
        Token Req. // | Authorization |^
          (I)     /   | Server        | \  (IV) Token
                //    |               |  \ Introspection Req.
               /      |               |   \     +Access
             //      /+---------------+    \     Token
            /      // (II)             \    \\
           /      /   Access            \     \
         //     //    Token              \ (V) \
         /      /                         \Resp.\
      //     //                            \     \
     /      v                               V     \
   +-----------+ Request +Signature/MAC+------------+
   |           |  (III)  +Access Token |            |
   |           |----------------------->| Resource   |
   | Client    |   (VI) Success or     | Server     |
   |           |        Failure        |            |
   |           |<----------------------|            |
   +-----------+                       +------------+
```

Figure 4: Token Introspection and Access Token Handles.

8. Security Considerations

The purpose of this document is to provide use cases, requirements, and motivation for developing an OAuth security solution extending Bearer Tokens. As such, this document is only about security.

9. IANA Considerations

This document does not require actions by IANA.

10. Acknowledgments

This document is the result of conference calls late 2012/early 2013 and in design team conference calls February 2013 of the IETF OAuth working group. The following persons (in addition to the OAuth WG chairs, Hannes Tschofenig, and Derek Atkins) provided their input during these calls: Bill Mills, Justin Richer, Phil Hunt, Prateek Mishra, Mike Jones, George Fletcher, Leif Johansson, Lucy Lynch, John

Bradley, Tony Nadalin, Klaas Wierenga, Thomas Hardjono, Brian Campbell

In the appendix of this document we reuse content from [RFC4962] and the authors would like thank Russ Housely and Bernard Aboba for their work on RFC 4962.

We would like to thank Reddy Tirumaleswar for his review.

11. References

11.1. Normative References

[I-D.ietf-oauth-pop-key-distribution]
Bradley, J., Hunt, P., Jones, M., and H. Tschofenig, "OAuth 2.0 Proof-of-Possession: Authorization Server to Client Key Distribution", draft-ietf-oauth-pop-key-distribution-02 (work in progress), October 2015.

[I-D.ietf-oauth-proof-of-possession]
Jones, M., Bradley, J., and H. Tschofenig, "Proof-of-Possession Key Semantics for JSON Web Tokens (JWTs)", draft-ietf-oauth-proof-of-possession-11 (work in progress), December 2015.

[I-D.ietf-oauth-signed-http-request]
Richer, J., Bradley, J., and H. Tschofenig, "A Method for Signing HTTP Requests for OAuth", draft-ietf-oauth-signed-http-request-02 (work in progress), February 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <http://www.rfc-editor.org/info/rfc5246>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <http://www.rfc-editor.org/info/rfc6749>.

[RFC7519] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <http://www.rfc-editor.org/info/rfc7519>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <http://www.rfc-editor.org/info/rfc7525>.

[RFC7662] Richer, J., Ed., "OAuth 2.0 Token Introspection", RFC 7662, DOI 10.17487/RFC7662, October 2015, <http://www.rfc-editor.org/info/rfc7662>.

11.2. Informative References

[I-D.hardjono-oauth-kerberos]
Hardjono, T., "OAuth 2.0 support for the Kerberos V5 Authentication Protocol", draft-hardjono-oauth-kerberos-01 (work in progress), December 2010.

[NIST800-63]
Burr, W., Dodson, D., Perlner, R., Polk, T., Gupta, S., and E. Nabbus, "NIST Special Publication 800-63-1, INFORMATION SECURITY", December 2008.

[RFC4120] Neuman, C., Yu, T., Hartman, S., and K. Raeburn, "The Kerberos Network Authentication Service (V5)", RFC 4120, DOI 10.17487/RFC4120, July 2005, <http://www.rfc-editor.org/info/rfc4120>.

[RFC4279] Eronen, P., Ed. and H. Tschofenig, Ed., "Pre-Shared Key Ciphersuites for Transport Layer Security (TLS)", RFC 4279, DOI 10.17487/RFC4279, December 2005, <http://www.rfc-editor.org/info/rfc4279>.

[RFC4347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security", RFC 4347, DOI 10.17487/RFC4347, April 2006, <http://www.rfc-editor.org/info/rfc4347>.

[RFC4962] Housley, R. and B. Aboba, "Guidance for Authentication, Authorization, and Accounting (AAA) Key Management", BCP 132, RFC 4962, DOI 10.17487/RFC4962, July 2007, <http://www.rfc-editor.org/info/rfc4962>.

[RFC5056] Williams, N., "On the Use of Channel Bindings to Secure Channels", RFC 5056, DOI 10.17487/RFC5056, November 2007, <http://www.rfc-editor.org/info/rfc5056>.

[RFC5849] Hammer-Lahav, E., Ed., "The OAuth 1.0 Protocol", RFC 5849, DOI 10.17487/RFC5849, April 2010, <http://www.rfc-editor.org/info/rfc5849>.

[RFC6125] Saint-Andre, P. and J. Hodges, "Representation and Verification of Domain-Based Application Service Identity within Internet Public Key Infrastructure Using X.509 (PKIX) Certificates in the Context of Transport Layer Security (TLS)", RFC 6125, DOI 10.17487/RFC6125, March 2011, <http://www.rfc-editor.org/info/rfc6125>.

[RFC6750] Jones, M. and D. Hardt, "The OAuth 2.0 Authorization Framework: Bearer Token Usage", RFC 6750, DOI 10.17487/RFC6750, October 2012, <http://www.rfc-editor.org/info/rfc6750>.

[RFC6819] Lodderstedt, T., Ed., McGloin, M., and P. Hunt, "OAuth 2.0 Threat Model and Security Considerations", RFC 6819, DOI 10.17487/RFC6819, January 2013, <http://www.rfc-editor.org/info/rfc6819>.

[RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, DOI 10.17487/RFC7252, June 2014, <http://www.rfc-editor.org/info/rfc7252>.

Authors’ Addresses

Phil Hunt (editor)
Oracle Corporation

Email: phil.hunt@yahoo.com

Justin Richer

Email: ietf@justin.richer.org

William Mills

Email: wmills@yahoo-inc.com

Prateek Mishra
Oracle Corporation

Email: prateek.mishra@oracle.com

Hannes Tschofenig
ARM Limited
Hall in Tirol 6060
Austria

Email: Hannes.Tschofenig@gmx.net
URI: http://www.tschofenig.priv.at

OAuth Working Group M. Jones
Internet-Draft A. Nadalin
Intended status: Standards Track Microsoft
Expires: January 21, 2020 B. Campbell, Ed.
Ping Identity
J. Bradley
Yubico
C. Mortimore
Salesforce
July 20, 2019

OAuth 2.0 Token Exchange
draft-ietf-oauth-token-exchange-19

Abstract

This specification defines a protocol for an HTTP- and JSON- based Security Token Service (STS) by defining how to request and obtain security tokens from OAuth 2.0 authorization servers, including security tokens employing impersonation and delegation.

Table of Contents

1. Introduction

A security token is a set of information that facilitates the sharing of identity and security information in heterogeneous environments or across security domains. Examples of security tokens include JSON Web Tokens (JWTs) [JWT] and SAML 2.0 Assertions [OASIS.saml-core-2.0-os]. Security tokens are typically signed to achieve integrity and sometimes also encrypted to achieve confidentiality. Security tokens are also sometimes described as Assertions, such as in [RFC7521].

A Security Token Service (STS) is a service capable of validating security tokens provided to it and issuing new security tokens in response, which enables clients to obtain appropriate access credentials for resources in heterogeneous environments or across security domains. Web Service clients have used WS-Trust [WS-Trust] as the protocol to interact with an STS for token exchange. While WS-Trust uses XML and SOAP, the trend in modern Web development has been towards RESTful patterns and JSON. The OAuth 2.0 Authorization Framework [RFC6749] and OAuth 2.0 Bearer Tokens [RFC6750] have emerged as popular standards for authorizing third-party applications' access to HTTP and RESTful resources. The conventional OAuth 2.0 interaction involves the exchange of some representation of resource owner authorization for an access token, which has proven to be an extremely useful pattern in practice. However, its input and output are somewhat too constrained as is to fully accommodate a security token exchange framework.

This specification defines a protocol extending OAuth 2.0 that enables clients to request and obtain security tokens from authorization servers acting in the role of an STS. Similar to OAuth 2.0, this specification focuses on client developer simplicity and requires only an HTTP client and JSON parser, which are nearly universally available in modern development environments. The STS protocol defined in this specification is not itself RESTful (an STS doesn't lend itself particularly well to a REST approach) but does utilize communication patterns and data formats that should be familiar to developers accustomed to working with RESTful systems.

A new grant type for a token exchange request and the associated specific parameters for such a request to the token endpoint are defined by this specification. A token exchange response is a normal OAuth 2.0 response from the token endpoint with a few additional parameters defined herein to provide information to the client.

The entity that makes the request to exchange tokens is considered the client in the context of the token exchange interaction. However, that does not restrict usage of this profile to traditional OAuth clients. An OAuth resource server, for example, might assume the role of the client during token exchange in order to trade an access token that it received in a protected resource request for a new token that is appropriate to include in a call to a backend service. The new token might be an access token that is more narrowly scoped for the downstream service or it could be an entirely different kind of token.

The scope of this specification is limited to the definition of a basic request-and-response protocol for an STS-style token exchange utilizing OAuth 2.0. Although a few new JWT claims are defined that enable delegation semantics to be expressed, the specific syntax, semantics and security characteristics of the tokens themselves (both those presented to the authorization server and those obtained by the client) are explicitly out of scope and no requirements are placed on the trust model in which an implementation might be deployed. Additional profiles may provide more detailed requirements around the specific nature of the parties and trust involved, such as whether signing and/or encryption of tokens is needed or if proof-of-possession style tokens will be required or issued; however, such details will often be policy decisions made with respect to the specific needs of individual deployments and will be configured or implemented accordingly.

The security tokens obtained may be used in a number of contexts, the specifics of which are also beyond the scope of this specification.

1.1. Delegation vs. Impersonation Semantics

One common use case for an STS (as alluded to in the previous section) is to allow a resource server A to make calls to a backend service C on behalf of the requesting user B. Depending on the local site policy and authorization infrastructure, it may be desirable for A to use its own credentials to access C along with an annotation of some form that A is acting on behalf of B ("delegation"), or for A to be granted a limited access credential to C but that continues to identify B as the authorized entity ("impersonation"). Delegation and impersonation can be useful concepts in other scenarios involving multiple participants as well.

When principal A impersonates principal B, A is given all the rights that B has within some defined rights context and is indistinguishable from B in that context. Thus, when principal A impersonates principal B, then insofar as any entity receiving such a token is concerned, they are actually dealing with B. It is true that some members of the identity system might have awareness that impersonation is going on, but it is not a requirement. For all intents and purposes, when A is impersonating B, A is B within the context of the rights authorized by the token. A's ability to impersonate B could be limited in scope or time, or even with a one-time-use restriction, whether via the contents of the token or an out-of-band mechanism.

Delegation semantics are different than impersonation semantics, though the two are closely related. With delegation semantics, principal A still has its own identity separate from B and it is explicitly understood that while B may have delegated some of its rights to A, any actions taken are being taken by A representing B. In a sense, A is an agent for B.

Delegation and impersonation are not inclusive of all situations. When a principal is acting directly on its own behalf, for example, neither delegation nor impersonation are in play. They are, however, the more common semantics operating for token exchange and, as such, are given more direct treatment in this specification.

Delegation semantics are typically expressed in a token by including information about both the primary subject of the token as well as the actor to whom that subject has delegated some of its rights. Such a token is sometimes referred to as a composite token because it is composed of information about multiple subjects. Typically, in the request, the "subject_token" represents the identity of the party on behalf of whom the token is being requested while the "actor_token" represents the identity of the party to whom the access rights of the issued token are being delegated. A composite token issued by the authorization server will contain information about both parties. When and if a composite token is issued is at the discretion of the authorization server and applicable policy and configuration.

The specifics of representing a composite token and even whether or not such a token will be issued depend on the details of the implementation and the kind of token. The representations of composite tokens that are not JWTs are beyond the scope of this specification. The "actor_token" request parameter, however, does provide a means for providing information about the desired actor and the JWT "act" claim can provide a representation of a chain of delegation.

1.2. Requirements Notation and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

1.3. Terminology

This specification uses the terms "access token type", "authorization server", "client", "client identifier", "resource server", "token endpoint", "token request", and "token response" defined by OAuth 2.0 [RFC6749], and the terms "Base64url Encoding", "Claim", and "JWT Claims Set" defined by JSON Web Token (JWT) [JWT].

2. Token Exchange Request and Response

2.1. Request

A client requests a security token by making a token request to the authorization server's token endpoint using the extension grant type mechanism defined in Section 4.5 of [RFC6749].

Client authentication to the authorization server is done using the normal mechanisms provided by OAuth 2.0. Section 2.3.1 of [RFC6749] defines password-based authentication of the client, however, client authentication is extensible and other mechanisms are possible. For example, [RFC7523] defines client authentication using bearer JSON Web Tokens (JWTs) [JWT]. The supported methods of client authentication and whether or not to allow unauthenticated or unidentified clients are deployment decisions that are at the discretion of the authorization server. Note that omitting client authentication allows for a compromised token to be leveraged via an STS into other tokens by anyone possessing the compromised token. Thus client authentication allows for additional authorization checks by the STS as to which entities are permitted to impersonate or receive delegations from other entities.

The client makes a token exchange request to the token endpoint with an extension grant type using the HTTP "POST" method. The following parameters are included in the HTTP request entity-body using the "application/x-www-form-urlencoded" format with a character encoding of UTF-8 as described in Appendix B of RFC6749 [RFC6749].

grant_type
REQUIRED. The value "urn:ietf:params:oauth:grant-type:token-exchange" indicates that a token exchange is being performed.

resource
OPTIONAL. A URI that indicates the target service or resource where the client intends to use the requested security token. This enables the authorization server to apply policy as appropriate for the target, such as determining the type and content of the token to be issued or if and how the token is to be encrypted. In many cases, a client will not have knowledge of the logical organization of the systems with which it interacts and will only know a URI of the service where it intends to use the token. The "resource" parameter allows the client to indicate to the authorization server where it intends to use the issued token by providing the location, typically as an https URL, in the token exchange request in the same form that will be used to access that resource. The authorization server will typically have the capability to map from a resource URI value to an appropriate policy. The value of the "resource" parameter MUST be an absolute URI, as specified by Section 4.3 of [RFC3986], which MAY include a query component and MUST NOT include a fragment component. Multiple "resource" parameters may be used to indicate that the issued token is intended to be used at the multiple resources listed. See [I-D.ietf-oauth-resource-indicators] for additional background and uses of the "resource" parameter.

audience
OPTIONAL. The logical name of the target service where the client intends to use the requested security token. This serves a purpose similar to the "resource" parameter, but with the client providing a logical name for the target service. Interpretation of the name requires that the value be something that both the client and the authorization server understand. An OAuth client identifier, a SAML entity identifier [OASIS.saml-core-2.0-os], an OpenID Connect Issuer Identifier [OpenID.Core], are examples of things that might be used as "audience" parameter values. However, "audience" values used with a given authorization server must be unique within that server, to ensure that they are properly interpreted as the intended type of value. Multiple "audience" parameters may be used to indicate that the issued token is intended to be used at the multiple audiences listed. The "audience" and "resource" parameters may be used together to indicate multiple target services with a mix of logical names and resource URIs.

scope
OPTIONAL. A list of space-delimited, case-sensitive strings, as defined in Section 3.3 of [RFC6749], that allow the client to specify the desired scope of the requested security token in the context of the service or resource where the token will be used. The values and associated semantics of scope are service specific and expected to be described in the relevant service documentation.

requested_token_type
OPTIONAL. An identifier, as described in Section 3, for the type of the requested security token. If the requested type is unspecified, the issued token type is at the discretion of the authorization server and may be dictated by knowledge of the requirements of the service or resource indicated by the "resource" or "audience" parameter.

subject_token
REQUIRED. A security token that represents the identity of the party on behalf of whom the request is being made. Typically, the subject of this token will be the subject of the security token issued in response to the request.

subject_token_type
REQUIRED. An identifier, as described in Section 3, that indicates the type of the security token in the "subject_token" parameter.

actor_token
OPTIONAL. A security token that represents the identity of the acting party. Typically, this will be the party that is authorized to use the requested security token and act on behalf of the subject.

actor_token_type
An identifier, as described in Section 3, that indicates the type of the security token in the "actor_token" parameter. This is REQUIRED when the "actor_token" parameter is present in the request but MUST NOT be included otherwise.

In processing the request, the authorization server MUST perform the appropriate validation procedures for the indicated token type and, if the actor token is present, also perform the appropriate validation procedures for its indicated token type. The validity criteria and details of any particular token are beyond the scope of this document and are specific to the respective type of token and its content.

In the absence of one-time-use or other semantics specific to the token type, the act of performing a token exchange has no impact on the validity of the subject token or actor token. Furthermore, the exchange is a one-time event and does not create a tight linkage between the input and output tokens, so that (for example) while the expiration time of the output token may be influenced by that of the input token, renewal or extension of the input token is not expected to be reflected in the output token's properties. It may still be appropriate or desirable to propagate token revocation events. However, doing so is not a general property of the STS protocol and would be specific to a particular implementation, token type or deployment.

2.1.1. Relationship Between Resource, Audience and Scope

When requesting a token, the client can indicate the desired target service(s) where it intends to use that token by way of the "audience" and "resource" parameters, as well as indicating the desired scope of the requested token using the "scope" parameter. The semantics of such a request are that the client is asking for a token with the requested scope that is usable at all the requested target services. Effectively, the requested access rights of the token are the cartesian product of all the scopes at all the target services.

An authorization server may be unwilling or unable to fulfill any token request but the likelihood of an unfulfillable request is significantly higher when very broad access rights are being solicited. As such, in the absence of specific knowledge about the relationship of systems in a deployment, clients should exercise discretion in the breadth of the access requested, particularly the number of target services. An authorization server can use the "invalid_target" error code, defined in Section 2.2.2, to inform a client that it requested access to too many target services simultaneously.

2.2. Response

The authorization server responds to a token exchange request with a normal OAuth 2.0 response from the token endpoint, as specified in Section 5 of [RFC6749]. Additional details and explanation are provided in the following subsections.

2.2.1. Successful Response

If the request is valid and meets all policy and other criteria of the authorization server, a successful token response is constructed by adding the following parameters to the entity-body of the HTTP response using the "application/json" media type, as specified by [RFC8259], and an HTTP 200 status code. The parameters are serialized into a JavaScript Object Notation (JSON) structure by adding each parameter at the top level. Parameter names and string values are included as JSON strings. Numerical values are included as JSON numbers. The order of parameters does not matter and can vary.

access_token
REQUIRED. The security token issued by the authorization server in response to the token exchange request. The "access_token" parameter from Section 5.1 of [RFC6749] is used here to carry the requested token, which allows this token exchange protocol to use the existing OAuth 2.0 request and response constructs defined for the token endpoint. The identifier "access_token" is used for historical reasons and the issued token need not be an OAuth access token.

issued_token_type
REQUIRED. An identifier, as described in Section 3, for the representation of the issued security token.

token_type
REQUIRED. A case-insensitive value specifying the method of using the access token issued, as specified in Section 7.1 of [RFC6749]. It provides the client with information about how to utilize the access token to access protected resources. For example, a value of "Bearer", as specified in [RFC6750], indicates that the issued security token is a bearer token and the client can simply present it as is without any additional proof of eligibility beyond the contents of the token itself. Note that the meaning of this parameter is different from the meaning of the "issued_token_type" parameter, which declares the representation of the issued security token; the term "token type" is more typically used with the aforementioned meaning as the structural or syntactical representation of the security token, as it is in all "*_token_type" parameters in this specification. If the issued token is not an access token or usable as an access token, then the "token_type" value "N_A" is used to indicate that an OAuth 2.0 "token_type" identifier is not applicable in that context.

expires_in
RECOMMENDED. The validity lifetime, in seconds, of the token issued by the authorization server. Oftentimes the client will not have the inclination or capability to inspect the content of the token and this parameter provides a consistent and token-type-agnostic indication of how long the token can be expected to be

valid. For example, the value 1800 denotes that the token will expire in thirty minutes from the time the response was generated.

scope
OPTIONAL, if the scope of the issued security token is identical to the scope requested by the client; otherwise, REQUIRED.

refresh_token
OPTIONAL. A refresh token will typically not be issued when the exchange is of one temporary credential (the subject_token) for a different temporary credential (the issued token) for use in some other context. A refresh token can be issued in cases where the client of the token exchange needs the ability to access a resource even when the original credential is no longer valid (e.g., user-not-present or offline scenarios where there is no longer any user entertaining an active session with the client). Profiles or deployments of this specification should clearly document the conditions under which a client should expect a refresh token in response to "urn:ietf:params:oauth:grant-type:token-exchange" grant type requests.

2.2.2. Error Response

If the request itself is not valid or if either the "subject_token" or "actor_token" are invalid for any reason, or are unacceptable based on policy, the authorization server MUST construct an error response, as specified in Section 5.2 of [RFC6749]. The value of the "error" parameter MUST be the "invalid_request" error code.

If the authorization server is unwilling or unable to issue a token for any target service indicated by the "resource" or "audience" parameters, the "invalid_target" error code SHOULD be used in the error response.

The authorization server MAY include additional information regarding the reasons for the error using the "error_description" as discussed in Section 5.2 of [RFC6749].

Other error codes may also be used, as appropriate.

2.3. Example Token Exchange

The following example demonstrates a hypothetical token exchange in which an OAuth resource server assumes the role of the client during the exchange. It trades an access token, which it received in a protected resource request, for a new token that it will use to call to a backend service (extra line breaks and indentation in the examples are for display purposes only).

Figure 1 shows the resource server receiving a protected resource request containing an OAuth access token in the Authorization header, as specified in Section 2.1 of [RFC6750].

```
GET /resource HTTP/1.1
Host: frontend.example.com
Authorization: Bearer accVkjcJyb4BWCxGsndESCJQbdFMogUC5PbRDqceLTC
```

Figure 1: Protected Resource Request

In Figure 2, the resource server assumes the role of client for the token exchange and the access token from the request in Figure 1 is sent to the authorization server using a request as specified in Section 2.1. The value of the "subject_token" parameter carries the access token and the value of the "subject_token_type" parameter indicates that it is an OAuth 2.0 access token. The resource server, acting in the role of the client, uses its identifier and secret to authenticate to the authorization server using the HTTP Basic authentication scheme. The "resource" parameter indicates the location of the backend service, https://backend.example.com/api, where the issued token will be used.

```
POST /as/token.oauth2 HTTP/1.1
Host: as.example.com
Authorization: Basic cnMwODpsb25nLXNlY3VyZS1yYW5kb20tc2VjcmV0
Content-Type: application/x-www-form-urlencoded

grant_type=urn%3Aietf%3Aparams%3Aoauth%3Agrant-type%3Atoken-exchange
&resource=https%3A%2F%2Fbackend.example.com%2Fapi
&subject_token=accVkjcJyb4BWCxGsndESCJQbdFMogUC5PbRDqceLTC
&subject_token_type=
 urn%3Aietf%3Aparams%3Aoauth%3Atoken-type%3Aaccess_token
```

Figure 2: Token Exchange Request

The authorization server validates the client credentials and the "subject_token" presented in the token exchange request. From the "resource" parameter, the authorization server is able to determine the appropriate policy to apply to the request and issues a token suitable for use at https://backend.example.com. The "access_token" parameter of the response shown in Figure 3 contains the new token, which is itself a bearer OAuth access token that is valid for one minute. The token happens to be a JWT; however, its structure and format are opaque to the client so the "issued_token_type" indicates only that it is an access token.

```
HTTP/1.1 200 OK
Content-Type: application/json
Cache-Control: no-cache, no-store

{
 "access_token":"eyJhbGciOiJFUzI1NiIsImtpZCI6IjllciJ9.eyJhdWQiOiJo
   dHRwczovL2JhY2tlbmQuZXhhbXBsZS5jb20iLCJpc3MiOiJodHRwczovL2FzLmV
   4YW1wbGUuY29tIiwiZXhwIjoxNDQxOTE3NTkzLCJpYXQiOjE0NDE5MTc1MzMsIn
   N1YiI6ImJkY0BleGFtcGxlLmNvbSIsInNjb3BlIjoiYXBpIn0.40y3ZgQedw6rx
   f59WlwHDD9jryFOr0_Wh3CGozQBihNBhnXEQgU85AI9x3KmsPottVMLPIWvmDCM
   y5-kdXjwhw",
 "issued_token_type":
     "urn:ietf:params:oauth:token-type:access_token",
 "token_type":"Bearer",
 "expires_in":60
}
```

Figure 3: Token Exchange Response

The resource server can then use the newly acquired access token in making a request to the backend server as illustrated in Figure 4.

```
GET /api HTTP/1.1
Host: backend.example.com
Authorization: Bearer eyJhbGciOiJFUzI1NiIsImtpZCI6IjllciJ9.eyJhdWQ
   iOiJodHRwczovL2JhY2tlbmQuZXhhbXBsZS5jb20iLCJpc3MiOiJodHRwczovL2
   FzLmV4YW1wbGUuY29tIiwiZXhwIjoxNDQxOTE3NTkzLCJpYXQiOjE0NDE5MTc1M
   zMsInN1YiI6ImJkY0BleGFtcGxlLmNvbSIsInNjb3BlIjoiYXBpIn0.40y3ZgQe
   dw6rxf59WlwHDD9jryFOr0_Wh3CGozQBihNBhnXEQgU85AI9x3KmsPottVMLPIW
   vmDCMy5-kdXjwhw
```

Figure 4: Backend Protected Resource Request

Additional examples can be found in Appendix A.

3. Token Type Identifiers

Several parameters in this specification utilize an identifier as the value to describe the token in question. Specifically, they are the "requested_token_type", "subject_token_type", "actor_token_type" parameters of the request and the "issued_token_type" member of the response. Token type identifiers are URIs. Token Exchange can work with both tokens issued by other parties and tokens from the given authorization server. For the former the token type identifier indicates the syntax (e.g., JWT or SAML 2.0) so the authorization server can parse it; for the latter it indicates what the given authorization server issued it for (e.g., access_token or refresh_token).

The following token type identifiers are defined by this specification. Other URIs MAY be used to indicate other token types.

urn:ietf:params:oauth:token-type:access_token
: Indicates that the token is an OAuth 2.0 access token issued by the given authorization server.

urn:ietf:params:oauth:token-type:refresh_token
: Indicates that the token is an OAuth 2.0 refresh token issued by the given authorization server.

urn:ietf:params:oauth:token-type:id_token
: Indicates that the token is an ID Token, as defined in Section 2 of [OpenID.Core].

urn:ietf:params:oauth:token-type:saml1
: Indicates that the token is a base64url-encoded SAML 1.1 [OASIS.saml-core-1.1] assertion.

urn:ietf:params:oauth:token-type:saml2
: Indicates that the token is a base64url-encoded SAML 2.0 [OASIS.saml-core-2.0-os] assertion.

The value "urn:ietf:params:oauth:token-type:jwt", which is defined in Section 9 of [JWT], indicates that the token is a JWT.

The distinction between an access token and a JWT is subtle. An access token represents a delegated authorization decision, whereas JWT is a token format. An access token can be formatted as a JWT but doesn't necessarily have to be. And a JWT might well be an access token but not all JWTs are access tokens. The intent of this specification is that "urn:ietf:params:oauth:token-type:access_token" be an indicator that the token is a typical OAuth access token issued by the authorization server in question, opaque to the client, and usable the same manner as any other access token obtained from that authorization server. (It could well be a JWT, but the client isn't and needn't be aware of that fact.) Whereas, "urn:ietf:params:oauth:token-type:jwt" is to indicate specifically that a JWT is being requested or sent (perhaps in a cross-domain use-case where the JWT is used as an authorization grant to obtain an access token from a different authorization server as is facilitated by [RFC7523]).

Note that for tokens which are binary in nature, the URI used for conveying them needs to be associated with the semantics of a base64 or other encoding suitable for usage with HTTP and OAuth.

4. JSON Web Token Claims and Introspection Response Parameters

It is useful to have defined mechanisms to express delegation within a token as well as to express authorization to delegate or impersonate. Although the token exchange protocol described herein can be used with any type of token, this section defines claims to express such semantics specifically for JWTs and in an OAuth 2.0 Token Introspection [RFC7662] response. Similar definitions for other types of tokens are possible but beyond the scope of this specification.

Note that the claims not established herein but used in examples and descriptions, such as "iss", "sub", "exp", etc., are defined by [JWT].

4.1. "act" (Actor) Claim

The "act" (actor) claim provides a means within a JWT to express that delegation has occurred and identify the acting party to whom authority has been delegated. The "act" claim value is a JSON object and members in the JSON object are claims that identify the actor. The claims that make up the "act" claim identify and possibly provide additional information about the actor. For example, the combination of the two claims "iss" and "sub" might be necessary to uniquely identify an actor.

However, claims within the "act" claim pertain only to the identity of the actor and are not relevant to the validity of the containing JWT in the same manner as the top-level claims. Consequently, non-identity claims (e.g., "exp", "nbf", and "aud") are not meaningful when used within an "act" claim, and therefore are not used.

Figure 5 illustrates the "act" (actor) claim within a JWT Claims Set. The claims of the token itself are about user@example.com while the "act" claim indicates that admin@example.com is the current actor.

```
{
  "aud":"https://consumer.example.com",
  "iss":"https://issuer.example.com",
  "exp":1443904177,
  "nbf":1443904077,
  "sub":"user@example.com",
  "act":
  {
    "sub":"admin@example.com"
  }
}
```

Figure 5: Actor Claim

A chain of delegation can be expressed by nesting one "act" claim within another. The outermost "act" claim represents the current actor while nested "act" claims represent prior actors. The least recent actor is the most deeply nested. The nested "act" claims serve as a history trail that connects the initial request and subject through the various delegation steps undertaken before reaching the current actor. In this sense, the current actor is considered to include the entire authorization/delegation history, leading naturally to the nested structure described here.

For the purpose of applying access control policy, the consumer of a token MUST only consider the token's top-level claims and the party identified as the current actor by the "act" claim. Prior actors identified by any nested "act" claims are informational only and are not to be considered in access control decisions.

The following example in Figure 6 illustrates nested "act" (actor) claims within a JWT Claims Set. The claims of the token itself are about user@example.com while the "act" claim indicates that the system https://service16.example.com is the current actor and https://service77.example.com was a prior actor. Such a token might come about as the result of service16 receiving a token in a call from service77 and exchanging it for a token suitable to call service26 while the authorization server notes the situation in the newly issued token.


```
{
  "aud":"https://service26.example.com",
  "iss":"https://issuer.example.com",
  "exp":1443904100,
  "nbf":1443904000,
  "sub":"user@example.com",
  "act":
  {
    "sub":"https://service16.example.com",
    "act":
    {
      "sub":"https://service77.example.com"
    }
  }
}
```


Figure 6: Nested Actor Claim

When included as a top-level member of an OAuth token introspection response, "act" has the same semantics and format as the claim of the same name.

4.2. "scope" (Scopes) Claim

The value of the "scope" claim is a JSON string containing a space-separated list of scopes associated with the token, in the format described in Section 3.3 of [RFC6749].

Figure 7 illustrates the "scope" claim within a JWT Claims Set.

```
{
  "aud":"https://consumer.example.com",
  "iss":"https://issuer.example.com",
  "exp":1443904177,
  "nbf":1443904077,
  "sub":"dgaf4mvfs75Fci_FL3heQA",
  "scope":"email profile phone address"
}
```

Figure 7: Scopes Claim

OAuth 2.0 Token Introspection [RFC7662] already defines the "scope" parameter to convey the scopes associated with the token.

4.3. "client_id" (Client Identifier) Claim

The "client_id" claim carries the client identifier of the OAuth 2.0 [RFC6749] client that requested the token.

The following example in Figure 8 illustrates the "client_id" claim within a JWT Claims Set indicating an OAuth 2.0 client with "s6BhdRkqt3" as its identifier.

```
{
  "aud":"https://consumer.example.com",
  "iss":"https://issuer.example.com",
  "exp":1443904177,
  "sub":"user@example.com",
  "client_id":"s6BhdRkqt3"
}
```

Figure 8: Client Identifier Claim

OAuth 2.0 Token Introspection [RFC7662] already defines the "client_id" parameter as the client identifier for the OAuth 2.0 client that requested the token.

4.4. "may_act" (Authorized Actor) Claim

The "may_act" claim makes a statement that one party is authorized to become the actor and act on behalf of another party. The claim might be used, for example, when a "subject_token" is presented to the token endpoint in a token exchange request and "may_act" claim in the subject token can be used by the authorization server to determine whether the client (or party identified in the "actor_token") is authorized to engage in the requested delegation or impersonation.

The claim value is a JSON object and members in the JSON object are claims that identify the party that is asserted as being eligible to act for the party identified by the JWT containing the claim. The claims that make up the "may_act" claim identify and possibly provide additional information about the authorized actor. For example, the combination of the two claims "iss" and "sub" are sometimes necessary to uniquely identify an authorized actor, while the "email" claim might be used to provide additional useful information about that party.

However, claims within the "may_act" claim pertain only to the identity of that party and are not relevant to the validity of the containing JWT in the same manner as top-level claims. Consequently, claims such as "exp", "nbf", and "aud" are not meaningful when used within a "may_act" claim, and therefore are not used.

Figure 9 illustrates the "may_act" claim within a JWT Claims Set. The claims of the token itself are about user@example.com while the "may_act" claim indicates that admin@example.com is authorized to act on behalf of user@example.com.

```
{
  "aud":"https://consumer.example.com",
  "iss":"https://issuer.example.com",
  "exp":1443904177,
  "nbf":1443904077,
  "sub":"user@example.com",
  "may_act":
  {
    "sub":"admin@example.com"
  }
}
```

Figure 9: Authorized Actor Claim

When included as a top-level member of an OAuth token introspection response, "may_act" has the same semantics and format as the claim of the same name.

5. Security Considerations

Much of the guidance from Section 10 of [RFC6749], the Security Considerations in The OAuth 2.0 Authorization Framework, is also applicable here. Furthermore, [RFC6819] provides additional security considerations for OAuth and [I-D.ietf-oauth-security-topics] has updated security guidance based on deployment experience and new threats that have emerged since OAuth 2.0 was originally published.

All of the normal security issues that are discussed in [JWT], especially in relationship to comparing URIs and dealing with unrecognized values, also apply here.

In addition, both delegation and impersonation introduce unique security issues. Any time one principal is delegated the rights of another principal, the potential for abuse is a concern. The use of the "scope" claim (in addition to other typical constraints such as a limited token lifetime) is suggested to mitigate potential for such abuse, as it restricts the contexts in which the delegated rights can be exercised.

6. Privacy Considerations

Tokens employed in the context of the functionality described herein may contain privacy-sensitive information and, to prevent disclosure of such information to unintended parties, MUST only be transmitted over encrypted channels, such as Transport Layer Security (TLS). In cases where it is desirable to prevent disclosure of certain information to the client, the token MUST be encrypted to its intended recipient. Deployments SHOULD determine the minimally necessary amount of data and only include such information in issued tokens. In some cases, data minimization may include representing only an anonymous or pseudonymous user.

7. IANA Considerations

7.1. OAuth URI Registration

This specification registers the following values in the IANA "OAuth URI" registry [IANA.OAuth.Parameters] established by [RFC6755].

7.1.1. Registry Contents

- URN: urn:ietf:params:oauth:grant-type:token-exchange
- Common Name: Token exchange grant type for OAuth 2.0
- Change controller: IESG
- Specification Document: Section 2.1 of [[this specification]]

- URN: urn:ietf:params:oauth:token-type:access_token
- Common Name: Token type URI for an OAuth 2.0 access token
- Change controller: IESG
- Specification Document: Section 3 of [[this specification]]

- URN: urn:ietf:params:oauth:token-type:refresh_token
- Common Name: Token type URI for an OAuth 2.0 refresh token
- Change controller: IESG
- Specification Document: Section 3 of [[this specification]]

- URN: urn:ietf:params:oauth:token-type:id_token
- Common Name: Token type URI for an ID Token
- Change controller: IESG
- Specification Document: Section 3 of [[this specification]]

- URN: urn:ietf:params:oauth:token-type:saml1
- Common Name: Token type URI for a base64url-encoded SAML 1.1 assertion
- Change Controller: IESG
- Specification Document: Section 3 of [[this specification]]

- URN: urn:ietf:params:oauth:token-type:saml2
- Common Name: Token type URI for a base64url-encoded SAML 2.0 assertion
- Change Controller: IESG
- Specification Document: Section 3 of [[this specification]]

7.2. OAuth Parameters Registration

This specification registers the following values in the IANA "OAuth Parameters" registry [IANA.OAuth.Parameters] established by [RFC6749].

7.2.1. Registry Contents

- Parameter name: resource
- Parameter usage location: token request
- Change controller: IESG
- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: audience
- Parameter usage location: token request
- Change controller: IESG
- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: requested_token_type
- Parameter usage location: token request
- Change controller: IESG
- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: subject_token
- Parameter usage location: token request
- Change controller: IESG
- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: subject_token_type
- Parameter usage location: token request
- Change controller: IESG

- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: actor_token
- Parameter usage location: token request
- Change controller: IESG
- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: actor_token_type
- Parameter usage location: token request
- Change controller: IESG
- Specification document(s): Section 2.1 of [[this specification]]

- Parameter name: issued_token_type
- Parameter usage location: token response
- Change controller: IESG
- Specification document(s): Section 2.2.1 of [[this specification]]

7.3. OAuth Access Token Type Registration

This specification registers the following access token type in the IANA "OAuth Access Token Types" registry [IANA.OAuth.Parameters] established by [RFC6749].

7.3.1. Registry Contents

- Type name: N_A
- Additional Token Endpoint Response Parameters: (none)
- HTTP Authentication Scheme(s): (none)
- Change controller: IESG
- Specification document(s): Section 2.2.1 of [[this specification]]

7.4. JSON Web Token Claims Registration

This specification registers the following Claims in the IANA "JSON Web Token Claims" registry [IANA.JWT.Claims] established by [JWT].

7.4.1. Registry Contents

- Claim Name: "act"
- Claim Description: Actor
- Change Controller: IESG
- Specification Document(s): Section 4.1 of [[this specification]]

- Claim Name: "scope"
- Claim Description: Scope Values
- Change Controller: IESG

- Specification Document(s): Section 4.2 of [[this specification]]

- Claim Name: "client_id"
- Claim Description: Client Identifier
- Change Controller: IESG
- Specification Document(s): Section 4.3 of [[this specification]]

- Claim Name: "may_act"
- Claim Description: Authorized Actor - the party that is authorized to become the actor
- Change Controller: IESG
- Specification Document(s): Section 4.4 of [[this specification]]

7.5. OAuth Token Introspection Response Registration

This specification registers the following values in the IANA "OAuth Token Introspection Response" registry [IANA.OAuth.Parameters] established by [RFC7662].

7.5.1. Registry Contents

- Claim Name: "act"
- Claim Description: Actor
- Change Controller: IESG
- Specification Document(s): Section 4.1 of [[this specification]]

- Claim Name: "may_act"
- Claim Description: Authorized Actor - the party that is authorized to become the actor
- Change Controller: IESG
- Specification Document(s): Section 4.4 of [[this specification]]

7.6. OAuth Extensions Error Registration

This specification registers the following values in the IANA "OAuth Extensions Error" registry [IANA.OAuth.Parameters] established by [RFC6749].

7.6.1. Registry Contents

- Error Name: "invalid_target"
- Error Usage Location: token error response
- Related Protocol Extension: OAuth 2.0 Token Exchange
- Change Controller: IETF
- Specification Document(s): Section 2.2.2 of [[this specification]]

8. References

8.1. Normative References

[IANA.JWT.Claims]
IANA, "JSON Web Token Claims", <http://www.iana.org/assignments/jwt>.

[IANA.OAuth.Parameters]
IANA, "OAuth Parameters", <http://www.iana.org/assignments/oauth-parameters>.

[JWT] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <http://tools.ietf.org/html/rfc7519>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, DOI 10.17487/RFC3986, January 2005, <https://www.rfc-editor.org/info/rfc3986>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <https://www.rfc-editor.org/info/rfc6749>.

[RFC7662] Richer, J., Ed., "OAuth 2.0 Token Introspection", RFC 7662, DOI 10.17487/RFC7662, October 2015, <https://www.rfc-editor.org/info/rfc7662>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8259] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", STD 90, RFC 8259, DOI 10.17487/RFC8259, December 2017, <https://www.rfc-editor.org/info/rfc8259>.

8.2. Informative References

[I-D.ietf-oauth-resource-indicators]
Campbell, B., Bradley, J., and H. Tschofenig, "Resource Indicators for OAuth 2.0", draft-ietf-oauth-resource-indicators-02 (work in progress), January 2019.

[I-D.ietf-oauth-security-topics]
Lodderstedt, T., Bradley, J., Labunets, A., and D. Fett, "OAuth 2.0 Security Best Current Practice", draft-ietf-oauth-security-topics-13 (work in progress), July 2019.

[OASIS.saml-core-1.1]
Maler, E., Mishra, P., and R. Philpott, "Assertions and Protocol for the OASIS Security Assertion Markup Language (SAML) V1.1", OASIS Standard oasis-sstc-saml-core-1.1, September 2003.

[OASIS.saml-core-2.0-os]
Cantor, S., Kemp, J., Philpott, R., and E. Maler, "Assertions and Protocol for the OASIS Security Assertion Markup Language (SAML) V2.0", OASIS Standard saml-core-2.0-os, March 2005.

[OpenID.Core]
Sakimura, N., Bradley, J., Jones, M., de Medeiros, B., and C. Mortimore, "OpenID Connect Core 1.0", November 2014, <http://openid.net/specs/openid-connect-core-1_0.html>.

[RFC6750] Jones, M. and D. Hardt, "The OAuth 2.0 Authorization Framework: Bearer Token Usage", RFC 6750, DOI 10.17487/RFC6750, October 2012, <https://www.rfc-editor.org/info/rfc6750>.

[RFC6755] Campbell, B. and H. Tschofenig, "An IETF URN Sub-Namespace for OAuth", RFC 6755, DOI 10.17487/RFC6755, October 2012, <https://www.rfc-editor.org/info/rfc6755>.

[RFC6819] Lodderstedt, T., Ed., McGloin, M., and P. Hunt, "OAuth 2.0 Threat Model and Security Considerations", RFC 6819, DOI 10.17487/RFC6819, January 2013, <https://www.rfc-editor.org/info/rfc6819>.

[RFC7521] Campbell, B., Mortimore, C., Jones, M., and Y. Goland, "Assertion Framework for OAuth 2.0 Client Authentication and Authorization Grants", RFC 7521, DOI 10.17487/RFC7521, May 2015, <https://www.rfc-editor.org/info/rfc7521>.

[RFC7523] Jones, M., Campbell, B., and C. Mortimore, "JSON Web Token (JWT) Profile for OAuth 2.0 Client Authentication and Authorization Grants", RFC 7523, DOI 10.17487/RFC7523, May 2015, <https://www.rfc-editor.org/info/rfc7523>.

[WS-Trust] Nadalin, A., Goodner, M., Gudgin, M., Barbir, A., and H. Granqvist, "WS-Trust 1.4", February 2012, <http://docs.oasis-open.org/ws-sx/ws-trust/v1.4/ws-trust.html>.

Appendix A. Additional Token Exchange Examples

Two example token exchanges are provided in the following sections illustrating impersonation and delegation, respectively (with extra line breaks and indentation for display purposes only).

A.1. Impersonation Token Exchange Example

A.1.1. Token Exchange Request

In the following token exchange request, a client is requesting a token with impersonation semantics (with only a "subject_token" and no "actor_token", delegation is impossible). The client tells the authorization server that it needs a token for use at the target service with the logical name "urn:example:cooperation-context".

```
POST /as/token.oauth2 HTTP/1.1
Host: as.example.com
Content-Type: application/x-www-form-urlencoded

grant_type=urn%3Aietf%3Aparams%3Aoauth%3Agrant-type%3Atoken-exchange
&audience=urn%3Aexample%3Acooperation-context
&subject_token=eyJhbGciOiJFUzI1NiIsImtpZCI6IjE2In0.eyJhdWQiOiJodHRwc
  zovL2FzLmV4YW1wbGUuY29tIiwiaXNzIjoiaHR0cHM6Ly9vcmlnaW5hbC1pc3N1ZXI
  uZXhhbXBsZS5uZXQiLCJleHAiOjE0NDE5MTA2MDAsIm5iZiI6MTQ0MTkwOTAwMCwic
  3ViIjoiYmRjQGV4YW1wbGUubmV0Iiwic2NvcGUiOiJvcmRlcnMgcHJvZmlsZSBoaXN
  0b3J5In0.PRBg-jXn4cJuj1gmYXFiGkZzRuzbXZ_sDxdE98ddW44ufsbWLKd3JJ1VZ
  hF64pbTtfjy4VXFVBDaQpKjn5JzAw
&subject_token_type=urn%3Aietf%3Aparams%3Aoauth%3Atoken-type%3Ajwt
```

Figure 10: Token Exchange Request

A.1.2. Subject Token Claims

The "subject_token" in the prior request is a JWT and the decoded JWT Claims Set is shown here. The JWT is intended for consumption by the authorization server within a specific time window. The subject of the JWT ("bdc@example.net") is the party on behalf of whom the new token is being requested.

```
{
  "aud":"https://as.example.com",
  "iss":"https://original-issuer.example.net",
  "exp":1441910600,
  "nbf":1441909000,
  "sub":"bdc@example.net",
  "scope":"orders profile history"
}
```

Figure 11: Subject Token Claims

A.1.3. Token Exchange Response

The "access_token" parameter of the token exchange response shown below contains the new token that the client requested. The other parameters of the response indicate that the token is a bearer access token that expires in an hour.

```
HTTP/1.1 200 OK
Content-Type: application/json
Cache-Control: no-cache, no-store

{
 "access_token":"eyJhbGciOiJFUzI1NiIsImtpZCI6IjcyIn0.eyJhdWQiOiJ1cm4
   6ZXhhbXBsZTpjb29wZXJhdGlvbi1jb250ZXh0IiwiaXNzIjoiaHR0cHM6Ly9hcy5l
   eGFtcGxlLmNvbSIsImV4cCI6MTQ0MTkxMzYxMCwic3ViIjoiYmRjQGV4YW1wbGUub
   mV0Iiwic2NvcGUiOiJvcmRlcnMgcHJvZmlsZSBoaXN0b3J5In0.rMdWpSGNACTvnF
   uOL74sYZ6MVuld2Z2WkGLmQeR9ztj6w2OXraQlkJmGjyiCq24kcB7AI2VqVxl3wSW
   nVKh85A",
 "issued_token_type":
   "urn:ietf:params:oauth:token-type:access_token",
 "token_type":"Bearer",
 "expires_in":3600
}
```

Figure 12: Token Exchange Response

A.1.4. Issued Token Claims

The decoded JWT Claims Set of the issued token is shown below. The new JWT is issued by the authorization server and intended for consumption by a system entity known by the logical name "urn:example:cooperation-context" any time before its expiration. The subject ("sub") of the JWT is the same as the subject the token used to make the request, which effectively enables the client to impersonate that subject at the system entity known by the logical name of "urn:example:cooperation-context" by using the token.

```
{
  "aud":"urn:example:cooperation-context",
  "iss":"https://as.example.com",
  "exp":1441913610,
  "sub":"bdc@example.net",
  "scope":"orders profile history"
}
```

Figure 13: Issued Token Claims

A.2. Delegation Token Exchange Example

A.2.1. Token Exchange Request

In the following token exchange request, a client is requesting a token and providing both a "subject_token" and an "actor_token". The client tells the authorization server that it needs a token for use at the target service with the logical name "urn:example:cooperation-context". Policy at the authorization server dictates that the issued token be a composite.

```
POST /as/token.oauth2 HTTP/1.1
Host: as.example.com
Content-Type: application/x-www-form-urlencoded

grant_type=urn%3Aietf%3Aparams%3Aoauth%3Agrant-type%3Atoken-exchange
&audience=urn%3Aexample%3Acooperation-context
&subject_token=eyJhbGciOiJFUzI1NiIsImtpZCI6IjE2In0.eyJhdWQiOiJodHRwc
  zovL2FzLmV4YW1wbGUuY29tIiwiaXNzIjoiaHR0cHM6Ly9vcmlnaW5hbC1pc3N1ZXI
  uZXhhbXBsZS5uZXQiLCJleHAiOjE0NDE5MTAwNjAsInNjb3BlIjoic3RhdHVzIGZlZ
  WQiLCJzdWIiOiJ1c2VyQGV4YW1wbGUubmV0IiwibWF5X2FjdCI6eyJzdWIiOiJhZG1
  pbkBleGFtcGxlLm5ldCJ9fQ.4rPRSWihQbpMIgAmAoqaJojAxj-p2X8_fAtAGTXrvM
  xU-eEZHnXqY0_AOZgLdxw5DyLzua8H_I10MCcckF-Q_g
&subject_token_type=urn%3Aietf%3Aparams%3Aoauth%3Atoken-type%3Ajwt
&actor_token=eyJhbGciOiJFUzI1NiIsImtpZCI6IjE2In0.eyJhdWQiOiJodHRwczo
  vL2FzLmV4YW1wbGUuY29tIiwiaXNzIjoiaHR0cHM6Ly9vcmlnaW5hbC1pc3N1ZXIuZ
  XhhbXBsZS5uZXQiLCJleHAiOjE0NDE5MTAwNjAsInN1YiI6ImFkbWluQGV4YW1wbGU
  ubmV0In0.7YQ-3zPfhUvzje5oqw8COCvN5uP6NsKik9CVV6cAOf4QKgM-tKfiOwcgZ
  oUuDL2tEs6tqPlcBlMjiSzEjm3yBg
&actor_token_type=urn%3Aietf%3Aparams%3Aoauth%3Atoken-type%3Ajwt
```

Figure 14: Token Exchange Request

A.2.2. Subject Token Claims

The "subject_token" in the prior request is a JWT and the decoded JWT Claims Set is shown here. The JWT is intended for consumption by the authorization server before a specific expiration time. The subject of the JWT ("user@example.net") is the party on behalf of whom the new token is being requested.

```
{
  "aud":"https://as.example.com",
  "iss":"https://original-issuer.example.net",
  "exp":1441910060,
  "scope":"status feed",
  "sub":"user@example.net",
  "may_act":
  {
    "sub":"admin@example.net"
  }
}
```

Figure 15: Subject Token Claims

A.2.3. Actor Token Claims

The "actor_token" in the prior request is a JWT and the decoded JWT Claims Set is shown here. This JWT is also intended for consumption by the authorization server before a specific expiration time. The subject of the JWT ("admin@example.net") is the actor that will wield the security token being requested.

```
{
  "aud":"https://as.example.com",
  "iss":"https://original-issuer.example.net",
  "exp":1441910060,
  "sub":"admin@example.net"
}
```

Figure 16: Actor Token Claims

A.2.4. Token Exchange Response

The "access_token" parameter of the token exchange response shown below contains the new token that the client requested. The other parameters of the response indicate that the token is a JWT that expires in an hour and that the access token type is not applicable since the issued token is not an access token.

```
HTTP/1.1 200 OK
Content-Type: application/json
Cache-Control: no-cache, no-store
```


```
{
 "access_token":"eyJhbGciOiJFUzI1NiIsImtpZCI6IjcyIn0.eyJhdWQiOiJ1cm4
   6ZXhhbXBsZTpjb29wZXJhdGlvbi1jb250ZXh0IiwiaXNzIjoiaHR0cHM6Ly9hcy5l
   eGFtcGxlLmNvbSIsImV4cCI6MTQ0MTkxMzYxMCwic2NvcGUiOiJzdGF0dXMgZmVlZ
   CIsInN1YiI6InVzZXJAZXhhbXBsZS5uZXQiLCJhY3QiOnsic3ViIjoiYWRtaW5AZX
   hhbXBsZS5uZXQifX0.3paKl9UySKYB5ng6_cUtQ2qlO8Rc_y7Mea7IwEXTcYbNdwG
   9-G1EKCFe5fW3H0hwX-MSZ49Wpcb1SiAZaOQBtw",
 "issued_token_type":"urn:ietf:params:oauth:token-type:jwt",
 "token_type":"N_A",
 "expires_in":3600
}
```


Figure 17: Token Exchange Response

A.2.5. Issued Token Claims

The decoded JWT Claims Set of the issued token is shown below. The new JWT is issued by the authorization server and intended for consumption by a system entity known by the logical name "urn:example:cooperation-context" any time before its expiration. The subject ("sub") of the JWT is the same as the subject of the "subject_token" used to make the request. The actor ("act") of the JWT is the same as the subject of the "actor_token" used to make the request. This indicates delegation and identifies "admin@example.net" as the current actor to whom authority has been delegated to act on behalf of "user@example.net".


```
{
  "aud":"urn:example:cooperation-context",
  "iss":"https://as.example.com",
  "exp":1441913610,
  "scope":"status feed",
  "sub":"user@example.net",
  "act":
  {
    "sub":"admin@example.net"
  }
}
```


Figure 18: Issued Token Claims

Appendix B. Acknowledgements

This specification was developed within the OAuth Working Group, which includes dozens of active and dedicated participants. It was produced under the chairmanship of Hannes Tschofenig, Derek Atkins, and Rifaat Shekh-Yusef with Kathleen Moriarty, Stephen Farrell, Eric Rescorla, Roman Danyliw, and Benjamin Kaduk serving as Security Area Directors. The following individuals contributed ideas, feedback, and wording to this specification:

Caleb Baker, Vittorio Bertocci, Mike Brown, Thomas Broyer, Roman Danyliw, William Denniss, Vladimir Dzhuvinov, Eric Fazendin, Phil Hunt, Benjamin Kaduk, Jason Keglovitz, Torsten Lodderstedt, Barry Leiba, Adam Lewis, James Manger, Nov Matake, Matt Miller, Hilarie Orman, Matthew Perry, Eric Rescorla, Justin Richer, Adam Roach, Rifaat Shekh-Yusef, Scott Tomilson, and Hannes Tschofenig.

Appendix C. Document History

[[to be removed by the RFC Editor before publication as an RFC]]

-19

- o Fix-up changes introduced in -18.
- o Fix invalid JSON in the Nested Actor Claim example.
- o Reference figure numbers in text when introducing the examples in Section 2 and 4.
- o Editorial updates from additional IESG evaluation comments.
- o Add an informational reference to ietf-oauth-resource-indicators
- o Update ietf-oauth-security-topics ref to 13

-18

- o Editorial updates based on a few more IESG evaluation comments.

-17

- o Editorial improvements and example fixes resulting from IESG evaluation comments.
- o Added a pointer to RFC6749's Appendix B. on the "Use of application/x-www-form-urlencoded Media Type" as a way of providing a normative citation (by reference) for the media type.
- o Strengthened some of the wording in the privacy considerations to bring it inline with RFC 7519 Sec. 12 and RFC 6749 Sec. 10.8.

-16

- o Fixed typo and added an AD to Acknowledgements.

-15

- o Updated the nested actor claim example to (hopefully) be more straightforward.
- o Reworked Privacy Considerations to say to use TLS in transit, minimize the amount of information in the token, and encrypt the token if disclosure of its information to the client is a concern per https://mailarchive.ietf.org/arch/msg/secdir/KJhx4aq_U5uk3k6zpYP-CEHbpVM
- o Moved the Security and Privacy Considerations sections to before the IANA Considerations.

-14

- o Added text in Section 4.1 about the "act" claim stating that only the top-level claims and the current actor are to be considered in applying access control decisions.

-13

- o Updated the claim name and value syntax for scope to be consistent with the treatment of scope in RFC 7662 OAuth 2.0 Token Introspection.
- o Updated the client identifier claim name to be consistent with the treatment of client id in RFC 7662 OAuth 2.0 Token Introspection.

-12

- o Updated to use the boilerplate from RFC 8174.

-11

- o Added new WG chair and AD to the Acknowledgements.
- o Applied clarifications suggested during AD review by EKR.

-10

- o Defined token type URIs for base64url-encoded SAML 1.1 and SAML 2.0 assertions.
- o Applied editorial fixes.

-09

- o Changed "security tokens obtained could be used in a number of contexts" to "security tokens obtained may be used in a number of contexts" per a WGLC suggestion.

- o Clarified that the validity of the subject or actor token have no impact on the validity of the issued token after the exchange has occurred per a WGLC comment.
- o Changed use of invalid_target error code to a SHOULD per a WGLC comment.
- o Clarified text about non-identity claims within the "act" claim being meaningless per a WGLC comment.
- o Added brief Privacy Considerations section per WGLC comments.

-08

- o Use the bibxml reference for OpenID.Core rather than defining it inline.
- o Added editor role for Campbell.
- o Minor clarification of the text for actor_token.

-07

- o Fixed typo (desecration -> discretion).
- o Added an explanation of the relationship between scope, audience and resource in the request and added an "invalid_target" error code enabling the AS to tell the client that the requested audiences/resources were too broad.

-06

- o Drop "An STS for the REST of Us" from the title.
- o Drop "heavyweight" and "lightweight" from the abstract and introduction.
- o Clarifications on the language around xxxxxx_token_type.
- o Remove the want_composite parameter.
- o Add a short mention of proof-of-possession style tokens to the introduction and remove the respective open issue.

-05

- o Defined the JWT claim "cid" to express the OAuth 2.0 client identifier of the client that requested the token.
- o Defined and requested registration for "act" and "may_act" as Token introspection response parameters (in addition to being JWT claims).
- o Loosen up the language about refresh_token in the response to OPTIONAL from NOT RECOMMENDED based on feedback form real world deployment experience.
- o Add clarifying text about the distinction between JWT and access token URIs.
- o Close out (remove) some of the Open Issues bullets that have been resolved.

-04

- o Clarified that the "resource" and "audience" request parameters can be used at the same time (via http://www.ietf.org/mail-archive/web/oauth/current/msg15335.html).
- o Clarified subject/actor token validity after token exchange and explained a bit more about the recommendation to not issue refresh tokens (via http://www.ietf.org/mail-archive/web/oauth/current/msg15318.html).
- o Updated the examples appendix to use an issuer value that doesn't imply that the client issued and signed the tokens and used "Bearer" and "urn:ietf:params:oauth:token-type:access_token" in one of the responses (via http://www.ietf.org/mail-archive/web/oauth/current/msg15335.html).
- o Defined and registered urn:ietf:params:oauth:token-type:id_token, since some use cases perform token exchanges for ID Tokens and no URI to indicate that a token is an ID Token had previously been defined.

-03

- o Updated the document editors (adding Campbell, Bradley, and Mortimore).
- o Added to the title.
- o Added to the abstract and introduction.
- o Updated the format of the request to use application/x-www-form-urlencoded request parameters and the response to use the existing token endpoint JSON parameters defined in OAuth 2.0.
- o Changed the grant type identifier to urn:ietf:params:oauth:grant-type:token-exchange.
- o Added RFC 6755 registration requests for urn:ietf:params:oauth:token-type:refresh_token, urn:ietf:params:oauth:token-type:access_token, and urn:ietf:params:oauth:grant-type:token-exchange.
- o Added RFC 6749 registration requests for request/response parameters.
- o Removed the Implementation Considerations and the requirement to support JWTs.
- o Clarified many aspects of the text.
- o Changed "on_behalf_of" to "subject_token", "on_behalf_of_token_type" to "subject_token_type", "act_as" to "actor_token", and "act_as_token_type" to "actor_token_type".
- o Added an "audience" request parameter used to indicate the logical names of the target services at which the client intends to use the requested security token.
- o Added a "want_composite" request parameter used to indicate the desire for a composite token rather than trying to infer it from the presence/absence of token(s) in the request.

- o Added a "resource" request parameter used to indicate the URLs of resources at which the client intends to use the requested security token.
- o Specified that multiple "audience" and "resource" request parameter values may be used.
- o Defined the JWT claim "act" (actor) to express the current actor or delegation principal.
- o Defined the JWT claim "may_act" to express that one party is authorized to act on behalf of another party.
- o Defined the JWT claim "scp" (scopes) to express OAuth 2.0 scope-token values.
- o Added the "N_A" (not applicable) OAuth Access Token Type definition for use in contexts in which the token exchange syntax requires a "token_type" value, but in which the token being issued is not an access token.
- o Added examples.

-02

- o Enabled use of Security Token types other than JWTs for "act_as" and "on_behalf_of" request values.
- o Referenced the JWT and OAuth Assertions RFCs.

-01

- o Updated references.

-00

- o Created initial working group draft from draft-jones-oauth-token-exchange-01.

Authors' Addresses

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

Anthony Nadalin
Microsoft

Email: tonynad@microsoft.com

Brian Campbell (editor)
Ping Identity

Email: brian.d.campbell@gmail.com

John Bradley
Yubico

Email: ve7jtb@ve7jtb.com

Chuck Mortimore
Salesforce

Email: cmortimore@salesforce.com

OAuth Working Group M. Jones
Internet-Draft Microsoft
Intended status: Standards Track J. Bradley
Expires: January 5, 2017 B. Campbell
Ping Identity
July 4, 2016

OAuth 2.0 Token Binding
draft-jones-oauth-token-binding-00

Abstract

This specification enables OAuth 2.0 implementations to apply Token Binding to Access Tokens and Refresh Tokens. This cryptographically binds these tokens to the TLS connections over which they are intended to be used. This use of Token Binding protects these tokens from man-in-the-middle and token export and replay attacks.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This specification enables OAuth 2.0 [RFC6749] implementations to apply Token Binding The Token Binding Protocol Version 1.0 [I-D.ietf-tokbind-protocol] Token Binding over HTTP [I-D.ietf-tokbind-https] to Access Tokens and Refresh Tokens. This cryptographically binds these tokens to the TLS connections over which they are intended to be used. This use of Token Binding protects these tokens from man-in-the-middle and token export and replay attacks.

1.1. Requirements Notation and Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

1.2. Terminology

This specification uses the terms "Access Token", "Authorization Code", "Authorization Endpoint", "Authorization Grant", "Authorization Server", "Client", "Client Authentication", "Client Identifier", "Client Secret", "Grant Type", "Protected Resource", "Redirection URI", "Refresh Token", "Resource Owner", "Resource Server", "Response Type", and "Token Endpoint" defined by OAuth 2.0 [RFC6749], the terms "Claim", "Claim Name", "Claim Value", and "JSON Web Token (JWT)" defined by JSON Web Token (JWT) [JWT], the term "User Agent" defined by RFC 7230 [RFC7230], and the terms "Provided", "Referred", "Token Binding" and "Token Binding ID" defined by Token Binding over HTTP [I-D.ietf-tokbind-https].

2. Token Binding for Refresh Tokens

Token Binding of refresh tokens is a straightforward first-party scenario, applying term "first-party" as used in Token Binding over HTTP [I-D.ietf-tokbind-https]. It cryptographically binds the refresh token to the TLS connection between the client and the token endpoint. This case is straightforward because the refresh token is both retrieved by the client from the token endpoint and sent by the client to the token endpoint. Unlike the federated scenarios described in Section 3 (Federation Use Cases) of Token Binding over HTTP [I-D.ietf-tokbind-https] and the access token case described in the next section, only a single TLS connection is involved in the refresh token case.

Token Binding a refresh token requires that the authorization server do two things. First, when refresh token is sent to the client, the authorization server needs to remember the Provided Token Binding ID and remember its association with the issued refresh token. Second, when a token request containing a refresh token is received at the token endpoint, the authorization server needs to verify that the Provided Token Binding ID for the request matches the remembered Token Binding ID associated with the refresh token. If the Token Binding IDs do not match, the authorization server should return an error in response to the request.

The means by which the authorization server remembers the association between the refresh token and the Token Binding ID is an implementation detail that beyond the scope of this specification. Some authorization servers will choose to store the Token Binding ID (or a cryptographic hash of it, such a SHA-256 hash [SHS]) in the refresh token itself, thus reducing the amount of state to be kept by the server. Other authorization servers will add the Token Binding ID value (or a hash of it) to an internal data structure also containing other information about the refresh token, such as grant type information. These choices make no difference to the client, since the refresh token is opaque to it.

3. Token Binding for Access Tokens

Token Binding for access tokens cryptographically binds the access token to the TLS connection between the client and the resource server. Token Binding is applied to access tokens in a similar manner to that described in Section 3 (Federation Use Cases) of Token Binding over HTTP [I-D.ietf-tokbind-https]. It is also builds upon the mechanisms for Token Binding of ID Tokens defined in OpenID Connect Token Bound Authentication 1.0 [OpenID.TokenBinding].

In the OpenID Connect [OpenID.Core] use case, HTTP redirects are used to pass information between the identity provider and the relying party; this HTTP redirect makes the Token Binding ID of the relying party available to the identity provider as the Referred Token Binding ID, information about which is then added to the ID Token. No such redirect occurs between the authorization server and the resource server in the access token case; therefore, information about the Token Binding ID for the TLS connection between the client and the resource server needs to be explicitly communicated by the client to the authorization server to achieve Token Binding of the access token. This information is passed to the authorization server using this request parameter:

resource_tbh
: Base64url encoding of the SHA-256 hash [SHS] of the Token Binding ID for the TLS connection between the client and the resource server.

Note that to obtain this Token Binding ID, the client needs to establish a TLS connection between itself and the resource server prior to making the authorization request so that the Provided Token Binding ID for the TLS connection to the resource server can be obtained. The means by which the client retrieves this Token Binding ID from the underlying Token Binding API is implementation and operating system specific. An alternative, if supported, is for the client to generate a Token Binding key to use for the resource server, use the Token Binding ID for that key, and then later use that key when the TLS connection to the resource server is established.

The authorization server MUST ignore the "resource_tbh" parameter if it does not support Token Binding for the access token.

3.1. Initial Access Tokens

Upon receiving the hash of the Token Binding ID in an authorization request containing the "resource_tbh" (resource token binding hash) authorization request parameter, the authorization server then records it in the issued access token. Alternatively, in some implementations, the resource’s Token Binding ID hash might be communicated to the resource server by other means, such as by introspecting [RFC7662] the access token.

3.2. Refreshed Access Tokens

Access tokens obtained from refresh requests can also be token bound. In this case, the hash of the Token Binding ID of the TLS connection between the client and the resource server is sent to the authorization server at the token endpoint using the "resource_tbh" (resource token binding hash) token request parameter; its syntax is exactly the same as the corresponding authorization request parameter. The authorization server then records it in the issued access token or communicates it to the resource server by other means, just as in the previous case.

3.3. Resource Server Token Binding Validation

Upon receiving a token bound access token, the resource server validates the binding by computing a SHA-256 hash of the Provided Token Binding ID and comparing it to the token binding hash value for the access token. If these values do not match, the resource access attempt MUST be rejected with an error.

3.4. Representing Token Binding in JWT Access Tokens

If the access token is represented as a JWT, the token binding information SHOULD be represented in the same way that it is in token bound OpenID Connect ID Tokens [OpenID.TokenBinding]. That specification defines the new JWT Confirmation Method RFC 7800 [RFC7800] member "tbh" (token binding hash) to represent the SHA-256 hash of a Token Binding ID in an ID Token. The value of the "tbh" member is the base64url encoding of the SHA-256 hash of the Token Binding ID.

The following example demonstrates the JWT Claims Set of an access token containing the base64url encoding of the SHA-256 hash of a Token Binding ID as the value of the "tbh" (token binding hash) element in the "cnf" (confirmation) claim:


```
  {
   "iss": "https://server.example.com",
   "aud": "https://resource.example.com",
   "iat": 1467324320,
   "exp": 1467324920,
   "cnf":{
     "tbh": "n0jI3trBK6_Gp2qiLOf48ZEZTjpBnhm-QOyzJxhBeAk"
    }
  }
```


4. Phasing in Token Binding and Preventing Downgrade Attacks

Many OAuth implementations will be deployed in situations in which not all participants support Token Binding. Any of combination of the client, the authorization server, the resource server, and the User Agent may not yet support Token Binding, in which case it will not work end-to-end.

It is a context-dependent deployment choice whether to allow interactions to proceed in which Token Binding is not supported or whether to treat Token Binding failures at any step as fatal errors. Particularly in dynamic deployment environments in which End Users have choices of clients, authorization servers, resource servers, and/or User Agents, it is RECOMMENDED that authorizations using one or more components that do not implement Token Binding be allowed to successfully proceed. This enables different components to be upgraded to supporting Token Binding at different times, providing a smooth transition path for phasing in Token Binding. However, when Token Binding has been performed, any Token Binding key mismatches MUST be treated as fatal errors.

If all the participants in an authorization interaction support Token Binding and yet one or more of them does not use it, this is likely evidence of a downgrade attack. In this case, the authorization SHOULD be aborted with an error. For instance, if the resource server knows that the authorization server and the User Agent both support Token Binding and yet the access token received does not contain Token Binding information, this is almost certainly a sign of an attack.

The authorization server and client can determine whether the other supports Token Binding using the metadata values defined in the next section. They can determine whether the User Agent supports Token

Binding by whether it negotiated Token Binding for the TLS connection. At present, there is no defined mechanism for determining whether the resource server supports Token Binding or not. However, it always safe to proceed as if it does; at worst, the resource server simply won't verify the Token Binding.

5. Token Binding Metadata

5.1. Token Binding Client Metadata

Clients supporting Token Binding that also support the OAuth 2.0 Dynamic Client Registration Protocol [RFC7591] use these metadata values to register their support for Token Binding of Access Tokens and Refresh Tokens:

client_access_token_token_binding_supported
: OPTIONAL. Boolean value specifying whether the Client supports Token Binding of Access Tokens. If omitted, the default value is "false".

client_refresh_token_token_binding_supported
: OPTIONAL. Boolean value specifying whether the Client supports Token Binding of Refresh Tokens. If omitted, the default value is "false".

5.2. Token Binding Authorization Server Metadata

Authorization Servers supporting Token Binding that also support OAuth 2.0 Discovery Metadata [OAuth.Discovery] use this metadata values to register their support for Token Binding of Access Tokens and Refresh Tokens:

as_access_token_token_binding_supported
: OPTIONAL. Boolean value specifying whether the Authorization Server supports Token Binding of Access Tokens. If omitted, the default value is "false".

as_refresh_token_token_binding_supported
: OPTIONAL. Boolean value specifying whether the Authorization Server supports Token Binding of Refresh Tokens. If omitted, the default value is "false".

6. Security Considerations

If a refresh request is received by the authorization server containing a "resource_tbh" (resource token binding hash) value requesting a token bound access token and the refresh token in the request is not itself token bound, then it is not clear that token binding the access token adds significant value. This situation should be considered an open issue for discussion by the working group.

7. IANA Considerations

7.1. OAuth Parameters Registration

This specification registers the following parameter in the IANA "OAuth Parameters" registry [IANA.OAuth.Parameters] established by RFC 6749 [RFC6749]:

7.1.1. Registry Contents

- Parameter name: "resource_tbh"
- Parameter usage location: Authorization Request, Token Request
- Change controller: IESG
- Specification document(s): Section 3 of this document
- Related information: None

7.2. OAuth Dynamic Client Registration Metadata Registration

This specification registers the following client metadata definitions in the IANA "OAuth Dynamic Client Registration Metadata" registry [IANA.OAuth.Parameters] established by [RFC7591]:

7.2.1. Registry Contents

- Client Metadata Name: "client_access_token_token_binding_supported"
- Client Metadata Description: Boolean value specifying whether the Client supports Token Binding of Access Tokens
- Change Controller: IESG
- Specification Document(s): Section 5.1 of [[this specification]]

- Client Metadata Name: "client_refresh_token_token_binding_supported"
- Client Metadata Description: Boolean value specifying whether the Client supports Token Binding of Refresh Tokens
- Change Controller: IESG
- Specification Document(s): Section 5.1 of [[this specification]]

7.3. OAuth Authorization Server Discovery Metadata Registration

This specification registers the following discovery metadata definitions in the IANA "OAuth Authorization Server Discovery Metadata" registry established by [OAuth.Discovery]:

7.3.1. Registry Contents

- Discovery Metadata Name: "as_access_token_token_binding_supported"
- Discovery Metadata Description: Boolean value specifying whether the Authorization Server supports Token Binding of Access Tokens
- Change Controller: IESG
- Specification Document(s): Section 5.2 of [[this specification]]

- Discovery Metadata Name: "as_refresh_token_token_binding_supported"
- Discovery Metadata Description: Boolean value specifying whether the Authorization Server supports Token Binding of Refresh Tokens
- Change Controller: IESG
- Specification Document(s): Section 5.2 of [[this specification]]

8. References

8.1. Normative References

[I-D.ietf-tokbind-https]
Popov, A., Nystrom, M., Balfanz, D., Langley, A., and J. Hodges, "Token Binding over HTTP", draft-ietf-tokbind-https-03 (work in progress), March 2016.

[I-D.ietf-tokbind-protocol]
Popov, A., Nystrom, M., Balfanz, D., Langley, A., and J. Hodges, "The Token Binding Protocol Version 1.0", draft-ietf-tokbind-protocol-06 (work in progress), May 2016.

[IANA.OAuth.Parameters]
IANA, "OAuth Parameters", <http://www.iana.org/assignments/oauth-parameters>.

[JWT] Jones, M., Bradley, J., and N. Sakimura, "JSON Web Token (JWT)", RFC 7519, DOI 10.17487/RFC7519, May 2015, <http://tools.ietf.org/html/rfc7519>.

[OpenID.TokenBinding]
Jones, M., Bradley, J., and B. Campbell, "OpenID Connect Token Bound Authentication 1.0", July 2016, <http://self-issued.info/docs/openid-connect-token-bound-authentication-1_0.html>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC6749] Hardt, D., Ed., "The OAuth 2.0 Authorization Framework", RFC 6749, DOI 10.17487/RFC6749, October 2012, <http://www.rfc-editor.org/info/rfc6749>.

[RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, DOI 10.17487/RFC7230, June 2014, <http://www.rfc-editor.org/info/rfc7230>.

[RFC7662] Richer, J., Ed., "OAuth 2.0 Token Introspection", RFC 7662, DOI 10.17487/RFC7662, October 2015, <http://www.rfc-editor.org/info/rfc7662>.

[RFC7800] Jones, M., Bradley, J., and H. Tschofenig, "Proof-of-Possession Key Semantics for JSON Web Tokens (JWTs)", RFC 7800, DOI 10.17487/RFC7800, April 2016, <http://www.rfc-editor.org/info/rfc7800>.

[SHS] National Institute of Standards and Technology, "Secure Hash Standard (SHS)", FIPS PUB 180-4, March 2012, <http://csrc.nist.gov/publications/fips/fips180-4/fips-180-4.pdf>.

8.2. Informative References

[OAuth.Discovery] Jones, M., Sakimura, N., and J. Bradley, "OAuth 2.0 Discovery", draft-ietf-oauth-discovery-02 (work in progress), March 2016, <http://tools.ietf.org/html/draft-ietf-oauth-discovery-02>.

[OpenID.Core] Sakimura, N., Bradley, J., Jones, M., de Medeiros, B., and C. Mortimore, "OpenID Connect Core 1.0", November 2014, <http://openid.net/specs/openid-connect-core-1_0.html>.

[RFC7591] Richer, J., Ed., Jones, M., Bradley, J., Machulak, M., and P. Hunt, "OAuth 2.0 Dynamic Client Registration Protocol", RFC 7591, DOI 10.17487/RFC7591, July 2015, <http://www.rfc-editor.org/info/rfc7591>.

Appendix A. Acknowledgements

The authors would like to thank the following people for their contributions to the specification: Dirk Balfanz, William Denniss, Andrei Popov, and Nat Sakimura.

Appendix B. Open Issues

- o Some token binding implementations apparently provide APIs that enable native applications to provide Referred Token Bindings, just as the federation support in the HTTPS Token Binding spec does. Can we count on these APIs being supported on all platforms, and if so, does this enable us to somehow do without the "resource_tbh" parameter by mandating that the client send both a Provided and a Referred Token Binding to the authorization server? If this isn't the case, is "resource_tbh" actually secure or does this open a cross-channel validation hole? This area probably needs more attention from both the Token Binding and OAuth working groups.
- o How should we support crypto agility for the hash function?

Appendix C. Document History

[[to be removed by the RFC Editor before publication as an RFC]]

-00

- o Created the initial version.

Authors' Addresses

Michael B. Jones
Microsoft

Email: mbj@microsoft.com
URI: http://self-issued.info/

John Bradley
Ping Identity

Email: ve7jtb@ve7jtb.com
URI: http://www.thread-safe.com/

Brian Campbell
Ping Identity

Email: brian.d.campbell@gmail.com
URI: https://twitter.com/__b_c