Understanding OAuth2

Friday January 22nd, 2016

* [What is OAuth2?](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#What_is_OAuth2)
* [Basic knowledge](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Basic_knowledge)
  + [Roles](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Roles)
  + [Tokens](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Tokens)
  + [Access token scope](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Access_token_scope)
  + [HTTPS](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#HTTPS)
* [Register as a client](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Register_as_a_client)
  + [Client registration](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Client_registration)
  + [Authorization server response](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Authorization_server_response)
* [Authorization grant types](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Authorization_grant_types)
  + [Authorization Code Grant](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Authorization_Code_Grant)
  + [Implicit Grant](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Implicit_Grant)
  + [Resource Owner Password Credentials Grant](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Resource_Owner_Password_Credentials_Grant)
  + [Client Credentials Grant](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Client_Credentials_Grant)
* [Access token usage](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Access_token_usage)
  + [Request parameter (GET or POST)](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Request_parameter_GET_or_POST)
  + [Authorization header](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Authorization_header)
* [Security](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Security)
  + [Vulnerability in Authorization Code Grant](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Vulnerability_in_Authorization_Code_Grant)
  + [Vulnerability in Implicit Grant](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Vulnerability_in_Implicit_Grant)
  + [Clickjacking](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Clickjacking)
* [Conclusion](http://www.bubblecode.net/en/2016/01/22/understanding-oauth2/#Conclusion)

If OAuth2 is still a vague concept for you or you simply want to be sure you understand its behaviours, this article should interest you.

What is OAuth2?



[OAuth2](https://oauth.net/2/) is, you guessed it, the version 2 of the OAuth protocol (also called framework).

This protocol allows third-party applications to grant limited access to an HTTP service, either on behalf of a resource owner or by allowing the third-party application to obtain access on its own behalf. Access is requested by a client, it can be a website or a mobile application for example.

Version 2 is expected to simplify the previous version of the protocol and to facilitate interoperability between different applications.

[Specifications](https://tools.ietf.org/html/rfc6749) are still being drafted and the protocol is constantly evolving but that does not prevent it from being implemented and acclaimed by several internet giants such as [Google](https://developers.google.com/accounts/docs/OAuth2) or [Facebook](https://developers.facebook.com/docs/reference/dialogs/oauth/).

Basic knowledge

Roles

OAuth2 defines 4 roles :

* **Resource Owner**: generally yourself.
* **Resource Server**: server hosting protected data (for example Google hosting your profile and personal information).
* **Client**: application requesting access to a resource server (it can be your PHP website, a Javascript application or a mobile application).
* **Authorization Server**: server issuing access token to the client. This token will be used for the client to request the resource server. This server can be the same as the authorization server (same physical server and same application), and it is often the case.

Tokens

Tokens are random strings generated by the authorization server and are issued when the client requests them.

There are 2 types of token:

* **Access Token**: this is the most important because it allows the user data from being accessed by a third-party application. This token is sent by the client as a parameter or as a header in the request to the resource server. It has a limited lifetime, which is defined by the authorization server. **It must be kept confidential as soon as possible** but we will see that this is not always possible, especially when the client is a web browser that sends requests to the resource server via Javascript.
* **Refresh Token**: this token is issued with the access token but unlike the latter, it is not sent in each request from the client to the resource server. It merely serves to be sent to the authorization server for renewing the access token when it has expired. For security reasons, it is not always possible to obtain this token. We will see later in what circumstances.

Access token scope

The scope is a parameter used to limit the rights of the access token. This is the authorization server that defines the list of the available scopes. The client must then send the scopes he wants to use for his application during the request to the authorization server. More the scope is reduced, the greater the chance that the resource owner authorizes access.

More information: [http://tools.ietf.org/html/rfc6749#section-3.3](https://tools.ietf.org/html/rfc6749#section-3.3).

HTTPS

OAuth2 requires the use of HTTPS for communication between the client and the authorization server because of sensitive data passing between the two (tokens and possibly resource owner credentials). In fact you are not forced to do so if you implement your own authorization server but you must know that you are opening a big security hole by doing this.

Register as a client

Since you want to retrieve data from a resource server using OAuth2, you have to register as a client of the authorization server.

Each provider is free to allow this by the method of his choice. The protocol only defines the parameters that must be specified by the client and those to be returned by the authorization server.

Here are the parameters (they may differ depending of the providers):

Client registration

* Application Name: the application name
* Redirect URLs: URLs of the client for receiving authorization code and access token
* Grant Type(s): authorization types that will be used by the client
* Javascript Origin (optional): the hostname that will be allowed to request the resource server via XMLHttpRequest

Authorization server response

* Client Id: unique random string
* Client Secret: secret key that must be kept confidential

More information: [RFC 6749 — Client Registration](https://tools.ietf.org/html/rfc6749#section-2).

Authorization grant types

OAuth2 defines 4 grant types depending on the location and the nature of the client involved in obtaining an access token.

Authorization Code Grant

When it should be used?

It should be used as soon as the client is a web server. It allows you to obtain a long-lived access token since it can be renewed with a refresh token (if the authorization server enables it).

Example:

* Resource Owner: you
* Resource Server: a Google server
* Client: any website
* Authorization Server: a Google server

Scenario:

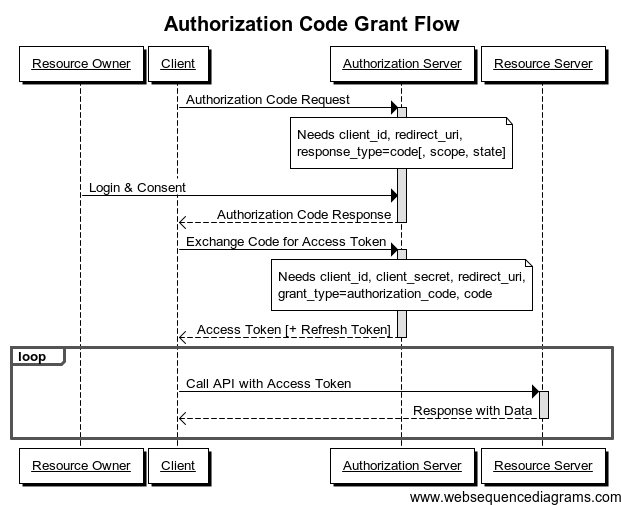
1. A website wants to obtain information about your Google profile.
2. You are redirected by the client (the website) to the authorization server (Google).
3. If you authorize access, the authorization server sends an authorization code to the client (the website) in the callback response.
4. Then, this code is exchanged against an access token between the client and the authorization server.
5. The website is now able to use this access token to query the resource server (Google again) and retrieve your profile data.

You never see the access token, it will be stored by the website (in session for example). Google also sends other information with the access token, such as the token lifetime and eventually a refresh token.

This is the ideal scenario and the safer one because the access token is not passed on the client side (web browser in our example).

More information: [RFC 6749 — Authorization Code Grant](https://tools.ietf.org/html/rfc6749#section-4.1).

Sequence diagram:

[](https://goo.gl/hbl1B)

Implicit Grant

When it should be used?

It is typically used when the client is running in a browser using a scripting language such as Javascript. This grant type does not allow the issuance of a refresh token.

Example:

* Resource Owner: you
* Resource Server: a Facebook server
* Client: a website using [AngularJS](http://angularjs.org/) for example
* Authorization Server: a Facebook server

Scenario:

1. The client (AngularJS) wants to obtain information about your Facebook profile.
2. You are redirected by the browser to the authorization server (Facebook).
3. If you authorize access, the authorization server redirects you to the website with the access token in the URI fragment (not sent to the web server). Example of callback: http://example.com/oauthcallback#access\_token=MzJmNDc3M2VjMmQzN.
4. This access token can now be retrieved and used by the client (AngularJS) to query the resource server (Facebook). Example of query: https://graph.facebook.com/me?access\_token=MzJmNDc3M2VjMmQzN.

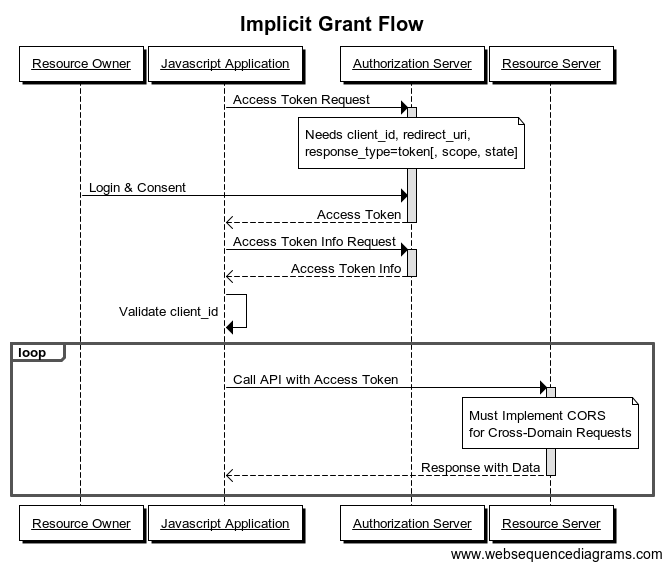
Maybe you wonder how the client can make a call to the Facebook API with Javascript without being blocked because of the [Same Origin Policy](https://en.wikipedia.org/wiki/Same_origin_policy)? Well, this cross-domain request is possible because Facebook authorizes it thanks to a header called Access-Control-Allow-Origin present in the response.

More information about Cross-Origin Resource Sharing (CORS): <https://developer.mozilla.org/en-US/docs/HTTP/Access_control_CORS#The_HTTP_response_headers>.

Attention! This type of authorization should only be used if no other type of authorization is available. Indeed, it is the least secure because the access token is exposed (and therefore vulnerable) on the client side.

More information: [RFC 6749 — Implicit Grant](https://tools.ietf.org/html/rfc6749#section-4.2).

Sequence diagram:

[](https://goo.gl/zfeof)

Resource Owner Password Credentials Grant

When it should be used?

With this type of authorization, the credentials (and thus the password) are sent to the client and then to the authorization server. It is therefore imperative that there is absolute trust between these two entities. It is mainly used when the client has been developed by the same authority as the authorization server. For example, we could imagine a website named example.com seeking access to protected resources of its own subdomain api.example.com. The user would not be surprised to type his login/password on the site example.com since his account was created on it.

Example:

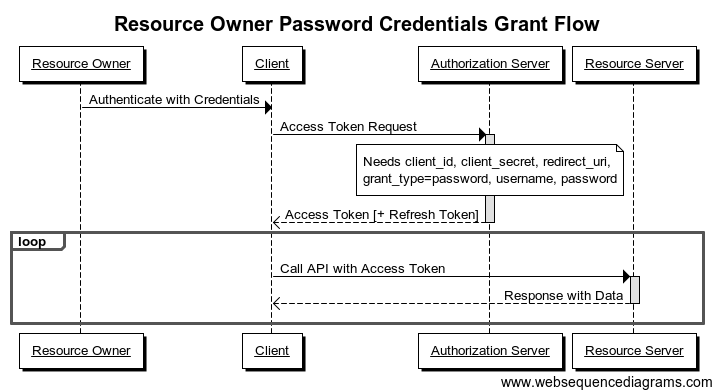
* Resource Owner: you having an account on acme.com website of the Acme company
* Resource Server: Acme company exposing its API at api.acme.com
* Client: acme.com website from Acme company
* Authorization Server: an Acme server

Scenario:

1. Acme company, doing things well, thought to make available a RESTful API to third-party applications.
2. This company thinks it would be convenient to use its own API to avoid reinventing the wheel.
3. Company needs an access token to call the methods of its own API.
4. For this, company asks you to enter your login credentials via a standard HTML form as you normally would.
5. The server-side application (website acme.com) will exchange your credentials against an access token from the authorization server (if your credentials are valid, of course).
6. This application can now use the access token to query its own resource server (api.acme.com).

More information: [RFC 6749 — Resource Owner Password Credentials Grant](https://tools.ietf.org/html/rfc6749#section-4.3).

Sequence diagram:

[](https://goo.gl/f1Tu7)

Client Credentials Grant

When it should be used?

This type of authorization is used when the client is himself the resource owner. There is no authorization to obtain from the end-user.

Example:

* Resource Owner: any website
* Resource Server: Google Cloud Storage
* Client: the resource owner
* Authorization Server: a Google server

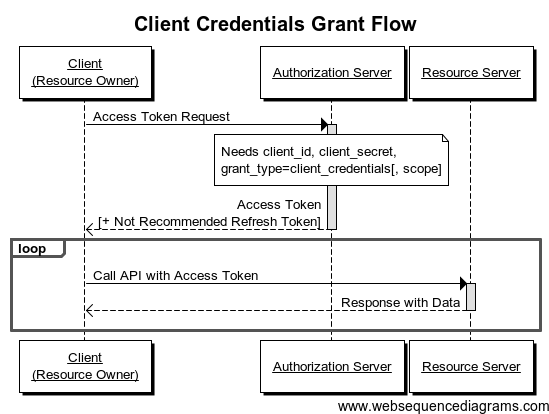
Scenario:

1. A website stores its files of any kind on Google Cloud Storage.
2. The website must go through the Google API to retrieve or modify files and must authenticate with the authorization server.
3. Once authenticated, the website obtains an access token that can now be used for querying the resource server (Google Cloud Storage).

Here, the end-user does not have to give its authorization for accessing the resource server.

More information: [RFC 6749 — Client Credentials Grant](https://tools.ietf.org/html/rfc6749#section-4.4).

Sequence diagram:

[](https://goo.gl/EpGVN)

Access token usage

The access token can be sent in several ways to the resource server.

Request parameter (GET or POST)

Example using GET: https://api.example.com/profile**?access\_token=MzJmNDc3M2VjMmQzN**

This is not ideal because the token can be found in the access logs of the web server.

Authorization header

GET /profile HTTP/1.1  
Host: api.example.com  
**Authorization: Bearer MzJmNDc3M2VjMmQzN**

It is elegant but all resource servers do not allow this.

Security

OAuth2 is sometimes criticized for its permeability, but it is often due to bad implementations of the protocol. There are big mistakes to avoid when using it, here are some examples.

Vulnerability in Authorization Code Grant

There is a vulnerability in this flow that allows an attacker to steal a user’s account under certain conditions. This hole is often encountered and also in many known websites (such as Pinterest, SoundCloud, Digg, …) that have not properly implemented the flow.

Example:

* Your victim has a valid account on a website called A.
* The A website allows a user to login or register with Facebook and is previously registered as a client in Facebook OAuth2 authorization server.
* You click on the Facebook Connect button of website A but do not follow the redirection thanks to Firefox [NoRedirect](https://addons.mozilla.org/fr/firefox/addon/noredirect/) addon or by using [Burp](http://portswigger.net/burp/) for example (callback looks like this: http://site-internet-a.com/facebook/login?code=OGI2NmY2NjYxN2Y4YzE3).
* You get the url (containing the authorization code) to which you would be redirected (visible in Firebug).
* Now you have to force your victim to visit this url via a hidden iframe on a website or an image in an email for example.
* If the victim is logged in website A, jackpot! Now you have access to the victim’s account in website A with your Facebook account. You just have to click on the Facebook Connect button and you will be connected with the victim’s account.

Workaround:

There is a way to prevent this by adding a “state” parameter. The latter is only recommended and not required in the specifications. If the client sends this parameter when requesting an authorization code, it will be returned unchanged by the authorization server in the response and will be compared by the client before the exchange of the authorization code against the access token. This parameter generally matches to a unique hash of a random number that is stored in the user session. For example in PHP: sha1(uniqid(mt\_rand(), true)).

In our example, if the website A was using the parameter “state”, he would have realized in the callback that the hash does not match the one stored in the session of the victim and would therefore prevented the theft of victim’s account.

More information: [RFC 6749 — Cross-Site Request Forgery](https://tools.ietf.org/html/rfc6749#section-10.12).

Vulnerability in Implicit Grant

This type of authorization is the least secure of all because it exposes the access token to client-side (Javascript most of the time). There is a widespread hole that stems from the fact that the client does not know if the access token was generated for him or not ([Confused Deputy Problem](https://en.wikipedia.org/wiki/Confused_deputy_problem)). This allows an attacker to steal a user account.

Example:

* An attacker aims to steal a victim’s account on a website A. This website allows you to connect via your Facebook account and uses implicit authorization.
* The attacker creates a website B allowing login via Facebook too.
* The victim logs in to the website B with his Facebook account and therefore implicitly authorized the generation of an access token for this.
* The attacker gets the access token via his website B and uses it on website A by modifying the access token in the URI fragment. If website A is not protected against this attack, the victim’s account is compromised and the attacker has now access to it.

Workaround:

To avoid this, the authorization server must provide in its API a way to retrieve access token information. Thus, website A would be able to compare the client\_id of the access token of the attacker against its own client\_id. As the stolen access token was generated for the website B, client\_id would have been different from client\_id of website A and the connection would have been refused.

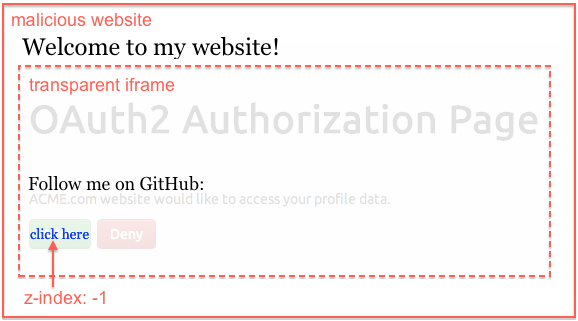
Google describes this in its API documentation: <https://developers.google.com/accounts/docs/OAuth2Login#validatingtoken>.

More information in RFC: [http://tools.ietf.org/html/rfc6819#section-4.4.2.6](https://tools.ietf.org/html/rfc6819#section-4.4.2.6)

Clickjacking

This technique allows the attacker to cheat by hiding the authorization page in a transparent iframe and getting the victim to click a link that is visually over the “Allow” button of the authorization page.

Example:



Workaround:

To avoid this, it is necessary that the authorization server returns a header named X-Frame-Options on the authorization page with the value DENY or SAMEORIGIN. This prevents the authorization page to be displayed in an iframe (DENY) or requires consistency between the domain name of the main page and the domain name specified in the iframe “src” attribute (SAMEORIGIN).

This header is not standard but is supported in the following browsers: IE8+, Firefox3.6.9+, Opera10.5+, Safari4+, Chrome 4.1.249.1042+.

More information: <https://developer.mozilla.org/en-US/docs/HTTP/X-Frame-Options>.

Here is the RFC that lists the potential vulnerabilities in the protocol implementations and the countermeasures: [http://tools.ietf.org/html/rfc6819](https://tools.ietf.org/html/rfc6819).

Conclusion

Like it or not, ([troll from a former contributor of the procotol](http://hueniverse.com/2012/07/oauth-2-0-and-the-road-to-hell/)), OAuth2 seems to impose itself for several years as a standard solution for the delegation of authority between different applications.

In any case, I hope I have helped you understand its workings. We will see in another article how to create your own OAuth2 authorization server with Symfony2.