

# Federated authentication: overview and status

J. Quinteros, with contributions from: Rob Casey, Jerry Carter, the EIDA Technical Committee, and many others





# Why do we authenticate?

- The first reason is always **because there are restricted data**.
- However, this is important for authorization, not authentication.
- Authentication checks who a person is.
- Authorization checks which specific resources a user has access to.
- To understand usage patterns is fundamental to improve services.
- It could also improve the quality of our statistics.
- It allows our funders to understand the impact of our data and services.





## FDSN current standard

- Only the dataselect web service has the option to be used with authentication. EventWS and StationWS don't have authentication.
- In dataselect there is a "queryauth" method to authenticate and submit a data request.
- We have 2 methods to get data: query and queryauth.
- HTTP Digest Authentication (RFC 2617) should be requested from the client.
- Authentication credentials are data center specific.





# Limitations

- EventWS and StationWS don't understand what a user is.
- To have 2 methods to request data could be confusing.
- Digest Authentication (RFC 2617) and different credentials per data centre seriously limit the scalability of a federation.
- Data centres must keep a list of usernames and passwords.





# Federated authentication at ORFEUS

Challenges:

- Services supporting open/embargoed data.
- Thousands of users/year around the world. Most of them unknown.
- New regulations on privacy (GDPR).
- Avoid the need to manage sensitive data at the data centre.
- Foster user authentication for open data (better statistics).
- Better understanding of how data is being used.
- How to properly manage a user database?





A Commission of

## **Federated authentication at ORFEUS**

- Approach: Complete decoupling from user login and service provisioning.
- User receives a token and presents it to the service providing data.
- The eduGAIN initiative (>8000 institutions) allows users to log in at their home institutions. We don't store any user data!





#### International Federation of Digital Seismograph Networks

# Get a token easily

- As a user, you can get a token from our web page.
- You will be redirected to your institution.
- Log in there.

A Commission of

- You will receive the token as soon as you authenticate.
- This needs to be stored where different tools (or you) can find it.

User documentation with all the details can be found in the following link.

EIDA users requesting Alparray data must complete all the requirements mentioned below for the registration process and later get in contact with the Network PI to be authorized to access the data.

From this page you can request a digitally signed token to be used with all existing EIDA web services (not Arclink) in order to not only retrieve open or restricted data, but also personalize your interaction with the EIDA services.

Please, select a duration for your token. After this amount of time, the token will not be accepted anymore. After clicking on the "Request token" button you will be redirected to B2ACCESS (optionally your home institution) to complete the authentication.

1 day			
2 days			
1 week			
2 weeks			
1 month			

**Request token** 



This work is co-funded by the EOSC-hub project (Horizon 2020) under Grant number 777536. And also by EUDAT2020 / EPOS-IP.



# **EIDA token digitally signed**

```
-----BEGIN PGP SIGNED MESSAGE-----
Hash: SHA1
```

{"valid\_until": "2019-05-08T10:21:26.269027Z", "cn": "Javier Quinteros", "memberof":
 "/epos/alparray;/epos;/", "sn": "Quinteros", "issued": "2019-04-08T10:21:26.269034Z", "mail":
 "javier@gfz-potsdam.de", "givenName": "Javier", "expiration": "1m"}
-----BEGIN PGP SIGNATURE----Version: GnuPG v1

iQEcBAEBAgAGBQJcqyCmAAoJEEFpzp0AlwdXaBQIAL9I7lUriWaoWMDPAnUTLUVEN8XjVN3Kfxa bokCY3jEfl6hRVLmEO8ofaV9iHktrXqvPaC6Ygp3w6raJi9mtmsS1O61FcUcBS1vG1LdWzcpM==8ke x

-----END PGP SIGNATURE-----





# Use of EIDA token with Obspy

165 Trace(s) in Stream: Z3.A022A..HHZ | 2016-03-01 - 2016-03-01T00:02:01 | 100.0 Hz, 12396 samples (163 other traces)... Z3.A216A.00.HHZ | 2016-03-01 - 2016-03-01T00:02:00 | 100.0 Hz, 12001 samples





#### **International Federation of Digital Seismograph Networks**

## Also for WebDC3 (GUI)

			Event and Station Map		
	Time Window selection:		<b>∲</b>		
	Relative Mode	Absolute Mode			
	Use an absolute	time window.	Authentication	):	
	Start	End			
	2020-12-14	2020-12-14			
	00:00:00	23:59:59	Current ID: javier	@gfz-potsdam.de	
			Valid until: Sat Ap	or 18 2020	
			17:35:54 GMT+02	00 (Central Eu-	
			ropean Summer	Time)	Le
	Request Type:				
	• Waveform (	Mini-SEED)			
	<ul> <li>Metadata (S</li> <li>Metadata (T</li> </ul>	Fext)	Load Token	Remove Toke	
	Authentication				
	Current ID: Anon				
	Valid until: N/A	ymous			
A Commission of 🛟 🌗	IASPI	Remove Toke			



## Problems with the current standard

- Our current specification forces us to use digest authentication (username and password).
- To avoid this, we use of a token, which is a "certificate" to recognize the user without username and password.
- To respect the standard we create temporary username/passwords, which are transparent for the user, and use them with "queryauth".
- Only clients take care of these steps (e.g. obspy, fdsnwsscripts, pyrocko).
- The client has to do these **two steps to request data**.
- The data centre has no reliable statistics from the user perspective.





# On-going activities on the IRIS side

- IRIS (EarthScope) implemented a similar AAI system.
- The basic, philosophical approach is basically the same.
- They implemented their token as a standard JSON Web Token (JWT).
- This one of the formats that everyone is using all the time. Even if we don't see them.





## International Federation of Digital Seismograph Networks

# JSON Web Token (JWT)

- Tokens can be signed.
- They cannot be edited.
- A token consumer can verify the issuer of the token and that hasn't been changed.

 They can also be encrypted (not interesting for us).

A Commission of 🛟



Debugger Libraries

Introduction Ask

k Get a T-shirt!

eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJ zdWIiOiIxMjM0NTY3ODkwIiwibmFtZSI6IkpvaG4 gRG9lIiwiZ3JvdXBzIjoiL04xOy9OMiIsImlhdCI 6MTUxNjIzOTAyMn0.x2GobD7EAkP\_NBAQMRtjw-Gw9KUj4HDg0H3BNBmbi6E

HEADER: ALGORITHM & TOKEN TYPE
{ "alg": "HS256", "typ": "JWT" }
PAYLOAD: DATA
{ "sub": "1234567890", "name": "John Doe", "groups": "/N1;/N2", "iat": 1516239022 }
VERIFY SIGNATURE
HMACSHA256( base64UrlEncode(header) + "." + base64UrlEncode(payload), your-256-bit-secret ) _ secret base64 encoded



## Improvements for the community – A roadmap?

- Adopt JWT as our standard format for tokens.
  - IRIS has done it. EIDA could do it transparently soon.
- Discuss a minimum number of fields we must have in the token, taking into account GDPR. Any data centres can always add more fields, if needed.
  - Some are: "issued at", "expiration", "email", "name", "groups" (for the future authorization)
- Use the HTTP header to transmit a token to the query method, if needed. Authorization: Bearer <token>





## Improvements for the community – A roadmap?

- Get rid of the queryauth method. (Finally!)
- Provide a simple, minimalistic way to issue tokens by small data centres.
  - Small data centres do not have the capacity to manage complex IT solutions. Implementation and operational procedures should be very simple.
- Discuss how to trust tokens from other data centres. It could be as easy as exchange a data centre key.
  - Data centres to agree on the algorithm to use and how to exchange the key.





# Conclusion

- We have a unique oportunity to make a step forward and simplify our data provision system.
  - Adapt a standard as **JWT** and **avoid the 2 steps**.
  - Simplify in a unique "query" method to request data.
  - All services could support authentication to improve statistics and analysis.
  - No more passwords for our users.
- This, plus the addition of the FDSN Data Centre Registry, would allow the user to **detach completely from where the data is hosted** (a unique, global seismological data centre).

