



Copernicus – eoSC AnaLytics Engine

WP2 Status Report

Did we do anything over the Summer?

Zdeněk Šustr, CESNET
sustr4@cesnet.cz

C-SCALE | 14 June 2022

Goals Reminder

Main goals for the WP

- **Discoverability**
 - **find** data in a uniform way \implies STAC (T2.2)
- **Accessibility**
 - **get** data in a uniform way \implies HTTPS/OIDC (T2.1)

- **STAC** (Spatio-Temporal Asset Catalogue) selected
 - Currently **STAC Core**, discussing extensions
 - Greater challenge, no support at any partner site, originally
 - Greater detail than other protocols (Open Search, CSW)
 - STAC `item` includes multiple `assets` (individual image files), which is typically not covered by current metadata databases
- ⇒ additional database under the STAC interface
- ⇒ **fun challenge** in populating it
- **Reconstruct from various metadata files**
 - Predict from subset (e.g., OpenSearch response)
 - ▶ Would make sense for high-turnover repositories (Relays)

Common Collections Structure

- Intention formulated over the Summer, delaying **M06**
- Partners who are only setting up their STACs could benefit from a shared structure of STAC collections.
 - e.g. do all Sentinel 2 products belong into a single collection, or do we separate L1C and L2A?
 - Similar for other platforms
 - Simplifies provider onboarding in the future
- Work in progress (disrupted by holidays)

User-level Catalogue

User-level STAC Catalogue, a.k.a. “central” catalogue

- Preferably owned by WP4, but supported by WP2
 - User-driven design
- <https://resto.c-scale.zcu.cz/>
- Place to store STAC metadata **at user level**
 - *“I have produced, downloaded or found something and I need to keep track of it...”*
 - Also useful for temporary datasets, e.g., for demos
- Only for associated use cases
 - Ask for access, tread carefully
 - Might integrate with Check-in in *Upscale*
- Based on <https://github.com/jjrom/resto>

Implementation Status

- MQS online: <https://mqs.eodc.eu/stac/v1/>
- VITO
 - Done, without extensions
 - <https://services.terrascope.be/stac/>
- EODC
 - Partly Done: <https://stac.eodc.eu/api/v1>
 - Only Sentinel 1 at the moment
- CESNET/GRNET
 - <https://stac.cesnet.cz/>
 - Metadata production cron job available for S1, S2, S3, S5p
 - Only problem with S3 SRAL L1 – not supported by *stactools*
- CloudFerro
 - ?

Why HTTP?

- simple *many:1*
- remote-friendly, latency agnostic
 - range specification is there for partial downloads
- default for many providers, easy to implement for others
- other “object-like” alternatives (S3?) permissible
- Random access (mounted file system) is complicated
 - Difficult to establish and track
 - + Tempts users to reuse what they know from local workflows

```
find /mnt/c-scale-federation -name "*_20220922*"
```

Authentication

Why can't it just be open (Why OIDC?)

- Open Data vs. service funding
 - Statistics required by funders to “pay” for the resources

How do we address that

- VOs set up, attributes collected on registration
 - attributes included in OIDC tokens
 - <https://confluence.egi.eu/x/f4YcC>
- Problem of coexistence with “national” users
 - How to make sure Check-in users and local federation (Terragrid, e-Infra.CZ, ...) users can access the same resources.
 - **Still under discussion three months later**
 - Integrating all under Check-in is not feasible.

Implementation Status – Central



- Support in Perun
 - Done
- Support in Check-in
 - Done

- Entitlements in token:

```
"urn:mace:egi.eu:res:c-scale:user-category:education#aai.egi.eu",  
"urn:mace:egi.eu:res:c-scale:company:CESNET#aai.egi.eu",  
"urn:mace:egi.eu:res:c-scale:accept-email:true#aai.egi.eu",  
"urn:mace:egi.eu:res:c-scale:country:CZ#aai.egi.eu",  
"urn:mace:egi.eu:res:c-scale:research-field:land#aai.egi.eu",  
"urn:mace:egi.eu:res:c-scale:user-function:researcher#aai.egi.eu",
```

... hence every user carries their aggregation details with them

Implementation Status – Partners

- VITO
 - integration in IdP done and tested
 - integration in download service in progress
- EODC
 - Early progress
- CESNET/GRNET
 - Functionality for free with DHuS 3.0, integration pending
 - Need to move the national userbase to an IdP as well
 - ▶ which simplifies things
- CloudFerro
 - ?

Summary: Status of the Federation



- Uneven coverage
 - Expected
 - Partially overcome by specific provisions for specific use cases
- Not growing
 - Understandable, no proper procedure to join – **TODO**
 - Attracting sites to participate
 - ▶ C-SCALE “compute” partners?
 - ▶ CollGS sites? We have candidates but no takers.
- Heavily Sentinel-centric
 - We say “Copernicus”, but we almost always mean “Sentinel”
- Delayed goal: **M06 Enhanced Federation-wide metadata**

Out

Zdeněk Šustr, CESNET
sustr4@cesnet.cz

C-SCALE | 14 June 2022