

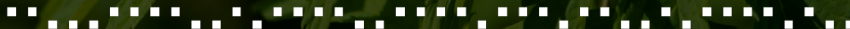
cesnet
"...."

CZECH COLLABORATIVE GROUND SEGMENT

Zdeněk Šustr, Tomáš Stibor

CESNET

11 October 2022



1. Introduction, status & news
2. DHuS Technical Issues
3. Alternative Interfaces to Sentinel data
4. Project C-SCALE

Czech CollGS Structure

■ Ministry of Transport

- NCP
- Management
- Representation towards ESA, the state, businesses, ...

■ CESNET

- Technology partner
- Operator of CollGS digital services
- User support
- Non-commercial/academic processing



Recycled
slide

Two distinct services

■ National Mirror

- Since late 2016
- National archive
- Hub for Czech++ user groups

■ DataHub Relay

- Since Spring 2018
- Demanding technical service in partnership with ESA
- Hidden to end users
- **DataHub mirrors welcome!**

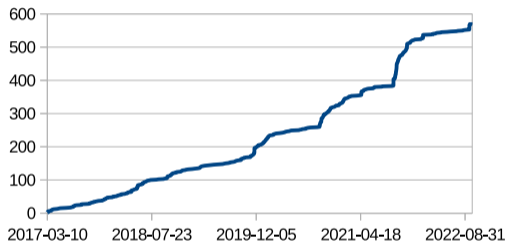


Recycled
slide

Numbers for 2022 so far

- ↗ **574** registered users
 - ↘ **98** interacting
 - ↘ **72** actually downloading
- ≡ **2.5 million** established connections
- ↘ **1.4 million** SOLR queries
 - ↘ 13.5 thousand Intersects
- ↘ 257 thousand OData queries
- ↗ 337 thousand products in archive
 - > 85 % cover Czech territory

No. of registered users



↗ 329 thousand products served

- **Non-trivial computing resources available to CollGS users at CESNET**
 - ~39 thousand CPU cores + a few hundred GPUs
 - Grid (HTC/HPC), Cloud, Kubernetes, Jupyter Notebooks
 - National access: **MetaCentrum**
 - International access: **EGI**, EOSC
 - ~20 % of product downloads go straight to Computing Elements
- **Technical partner in scientific projects, lead in technical projects**
 - **C-SCALE (to be introduced in a few slides)**
 - InterTwin
 - **GREAT (Green Deal Dataspace Alignment)**
 - EOSC Future
 - EGI-ACE

Page size excession

- Client site occasionally exceeds maximum page size setting
- SERCO recommendation: "set page size to half"

Skipped products accumulating

- DHuS keeps track of products it wants to retry
 - but asks for them all at once (CRDR currently at 70 k)

⇒ ODada synchronizers could **benefit from a total redesign!**

- replace *query/response* with a *feed-based* model

or at least allow setting ID on import and we will implement our own download

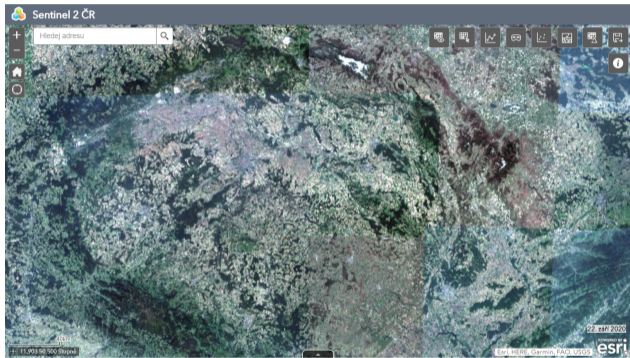
Airavata (as presented on Monday) – is it the answer?

Product revocation

- Automated mechanism to advertise products **removed from SciHub** as faulty, incomplete, republished, repackaged, ...
i.e., products that **need to be removed from national mirrors**
- Human-readable announcements are not enough
 - imprecise, error-prone, requiring interpretation and simultaneous action
- Could go hand in hand with OData synchronizer replacement

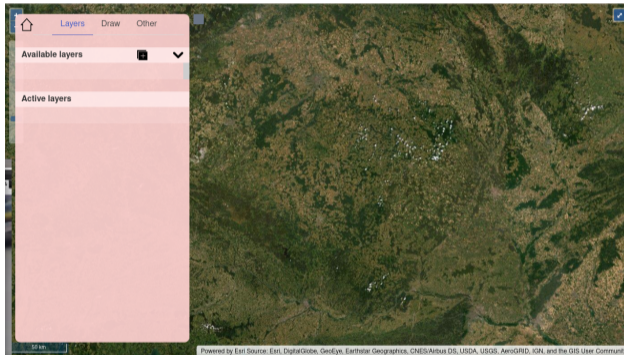
“Alternative”, as in: other than *select region* → *download*

- Arc-GIS
- Can integrate with users' GIS instances
- Web-based browser



<https://collgs.czechspaceportal.cz/mapova-aplikace/>

- Visualisation/distribution of products not compatible with DHuS
- Simple inclusion of custom layers



<https://envision.cerit-sc.cz>

Spatio-Temporal Asset Catalogue

- Machine-to-machine discovery interface
- Much richer than OpenSearch
 - down to band level
- Required by other **ESA-sanctioned products**

e.g. OpenEO platform

```

    ],
    platform: "Sentinel-2A",
    constellation: "Sentinel 2",
    instruments: [
      "msi"
    ]
  },
  "eo:cloud_cover": 2.08287124494024,
  "sat:orbit_state": "descending",
  "sat:relative_orbit": 22,
  "proj:epsg": 32633,
  "s2:product_uri": "S2A_MSIL1C_20220816T100611_N0400_R022_T33UVP_20220816T135125.SAFE",
  "s2:generation_time": "2022-08-16T13:51:25.000000Z",
  "s2:processing_baseline": "04.00",
  "s2:product_type": "S2MSI1C",
  "s2:datatake_id": "GS2A_20220816T100611_037342_N04.00",
  "s2:datatake_type": "INS-NOBS",
  "s2:datastrip_id": "S2A_OPER_MSI_L1C_DS_ATOS_20220816T135125_S20220816T100922_N04.00",
  "s2:granule_id": "S2A_OPER_MSI_L1C_TL_ATOS_20220816T135125_A037342_T33UVP_N04.00",
  "s2:mgrs_tile": "33UVP",
  "s2:reflectance_conversion_factor": 0.974147148598951,
  "s2:degraded_msi_data_percentage": 0.0262,
  "s2:mean_solar_zenith": 36.0603030370621,
  "s2:mean_solar_azimuth": 159.420648714476,
  updated: "2022-08-17T13:35:56.715623Z",
  created: "2022-08-17T13:35:56.715623Z",
  "resto:keywords": [
    {
      id: "collection:S2-experimental",
      name: "S2-experimental"
    }
  ]
}

```

<https://stac.cesnet.cz>

C-SCALE

Copernicus – eoSC AnaLytics Engine

Goals of the C-SCALE Project ...and how does it relate to CollGS

Zdeněk Šustr, CESNET
sustr4@cesnet.cz

Prague | 11 October 2022

Copernicus Analytics Engine



Data and compute federation under **EOSC**

c-scale.eu

- Look around! Other C-SCALE partners among you!
- *Each partner covers part of the globe*
 - can a federation cover all of it?
- Avoid creating yet another metadata catalogue, rather learn to query existing sources and federate identities
 - Ground segments (CESNET, GRNET)
 - DIASes (CreoDIAS)
 - National institutes (EODC, VITO)
- Make EO data and analytic workflows accessible in EOSC
- Involve users, bring in new communities

Two Main Data Access Topics

Accessibility

- “Analytics Engine” \implies putting **compute** and **data** resources “under the **same roof**”
 - Seamless access by robots (automated workflows)
 - Federated identities recognized by compute and data providers
 - Still obtaining statistical information **ESA** loves so much
 - \implies attributes carried with OIDC tokens
- Reusing ^(integrating)_{with} work done by **EGI** over the past 10 years
 - Federated cloud (+ additional HTC/HPC resources)
 - OIDC-based identity federation
- DHuS v. 3 makes great progress towards that goal
 - Searching for a **new home for our user base**

Two Main Data Access Topics

Discoverability

1. Common protocol

- Currently supported? (Open Search, CWS, OData, ...)
- Ambitious Choice – **STAC** (Spatio-temporal asset catalogue)
 - + Active community, rich ecosystem
 - Metadata richer than existing DBs
 - Repopulating metadata catalogues from **files (plural!)** – is that a candidate for the transformation framework? Or the successor tho the SAFE format?

2. Common endpoint

[↗ mqs.eodc.eu/stac/v1](https://mqs.eodc.eu/stac/v1)

- Redistribute incoming queries
 - ▶ Avoid querying unlikely candidates if possible
- Collate responses

Verdict: Simple contribution to DHuS is not enough
...but providing interfaces relevant to other ESA products would be nice

Thank you for your attention!

Zdeněk Šustr, CESNET
sustr4@cesnet.cz