



Data Analysis made easy with the ENES Climate Analytics Service (ECAS)

Sofiane Bendoukha, Donatello Elia, Fabrizio Antonio, Sandro Fiore, Tobias Weigel, Alessandro D'Anca

**April 10, 2019 @ EGU 19
SC1.22**

 eosc-hub.eu

 [@EOSC_eu](https://twitter.com/EOSC_eu)

Dissemination level: Public



● Training materials

- <https://github.com/ECAS-Lab/ecas-training>

● ECASLab / JupyterHub

- **ECASLab @ DKRZ** <https://ecaslaboratory.dkrz.de>
- **ECASLab @ CMCC** <https://ecaslaboratory.cmcc.it>

● Ophidia framework documentation

- <http://ophidia.cmcc.it/documentation/users/index.html>

● ECAS is part of the **EOSC-HUB** service catalogue

- ECAS enables scientific end-users to perform data analysis experiments

● Server-based

- Computation @ **CMCC** or **DKRZ**
- Avoid data transfer (download)
- Improved reusability of data and workflows (FAIR approach)

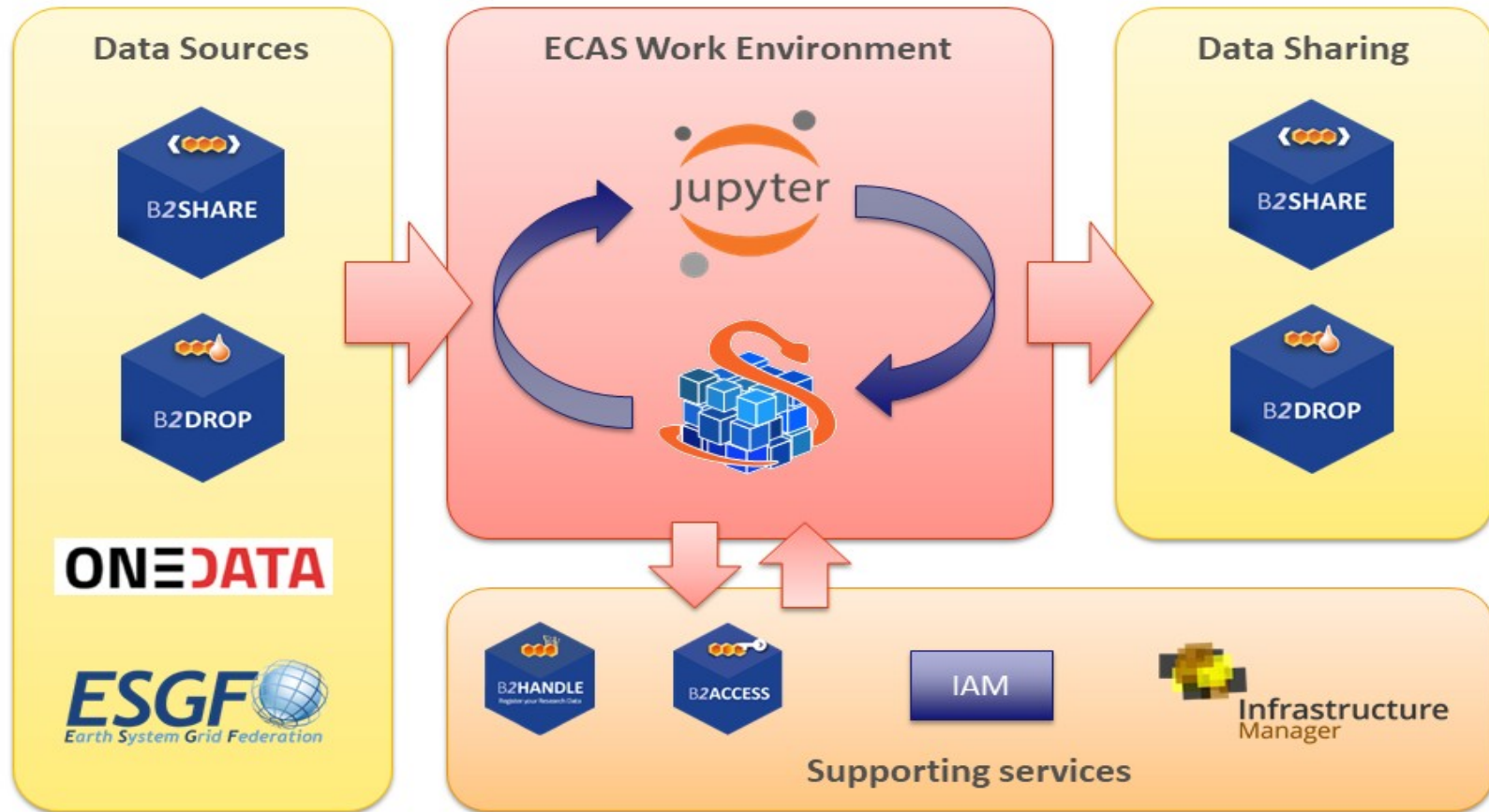


● ECAS supports different Auth* providers

- Local and external AAI providers supported (LDAP, B2ACCESS, EGI Check In)
- Additional AAI providers can be integrated (e.g. INDIGO IAM)

- ECAS provides data access via ESGF
- Coordinated Regional Climate Downscaling Experiment
 - ~ 100 Tbyte Cordex
- Coupled Model Intercomparison Project 5
 - ~ 1.2 Pbyte CMIP5 Data
- Coupled Model Intercomparison Project 6
 - ~ 250 Tbyte CMIP6 Data from the 1PByte published
- Other Data pools can be mounted on demand
 - MPI Grand ensemble (**MPI - GE**)
 - data collection exposed in the Federated Data Archive (e.g. through **OneData**)

Service architecture and interfaces

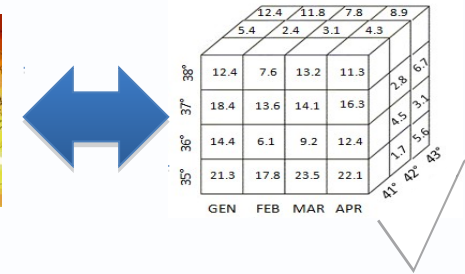
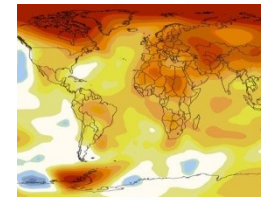
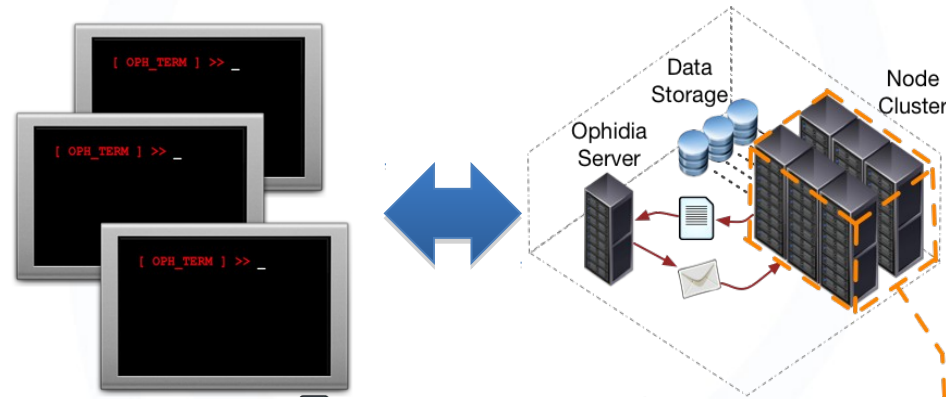


● The Ophidia framework addresses big data challenges for eScience

- support for declarative, parallel, server-side data analysis exploiting parallel computing techniques
- end-to-end mechanisms to support complex experiments and large processing workflows on scientific multi-dimensional datacubes

● Ophidia supports both **batch** & **interactive** data analytics

- More than 50 datacube-oriented **operators** are available, including: data reduction and subsetting, data intercomparison, metadata and provenance management, time series analysis with array-based primitives
- A wide set of (low-level) array-based primitives (over 100) to perform, e.g. data summarization, algebraic expressions, predicates evaluation, statistical analysis
- Support for complex workflows and Python applications execution

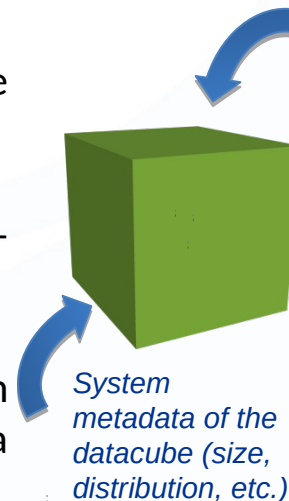


Oph_Term: a terminal-like commands interpreter serving as a client for the Ophidia framework

PyOphidia: a Python interface for datacube management & analytics with Ophidia

Ophidia framework: declarative, parallel server-side processing

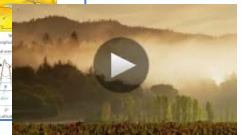
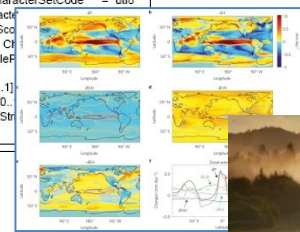
Through **oph_term/PyOphidia** the user run ("send") commands ("operators") to the Ophidia framework to manipulate datasets ("datacubes")



<<Abstract>>
MD_Metadata

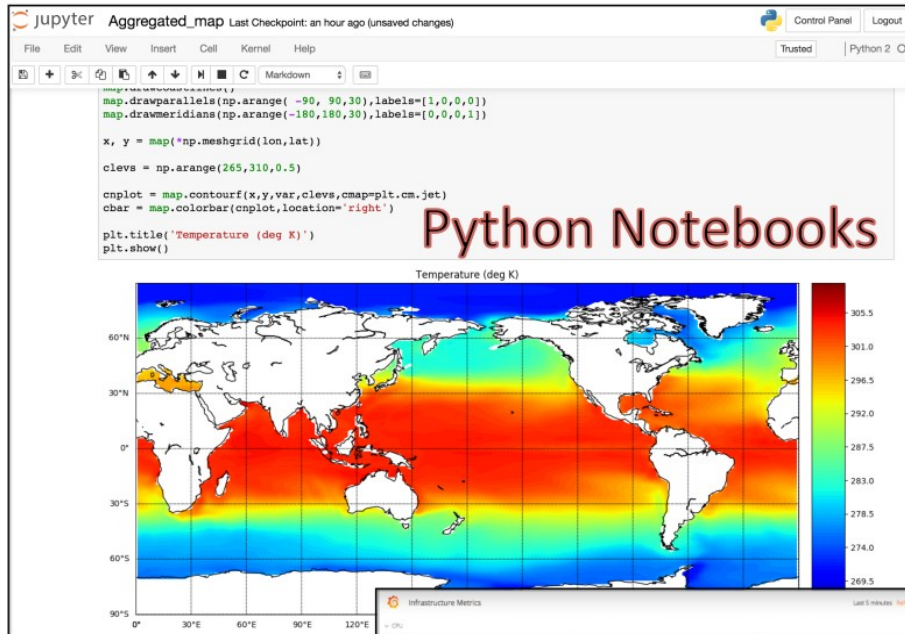
- + fieldIdentifier [0..1]: CharacterString
- + language [0..1]: CharacterString
- + characterSet [0..1]: MD_CharacterSetCode = "utf8"
- + parentIdentifier [0..1]: CharacterString
- + hierarchyLevel [0..*]: MD_Scope
- + hierarchyLevelName [0..*]: CharacterString
- + contact [1..*]: CI_Contact
- + dateStamp : Date
- + metadataStandardName [0..1]: CharacterString
- + metadataStandardVersion [0..1]: CharacterString
- + datasetURI [0..1]: CharacterString
- + locale [0..*]: PT_Locale

User metadata information

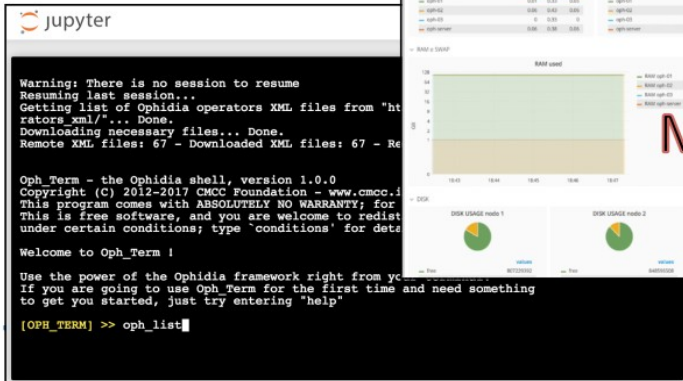


Metadata provenance

```
--> https://ophidia.cmcc.it:8443/162/169 (ROOT)
  https://ophidia.cmcc.it:8443/162/170 (oph_reduce)
    https://ophidia.cmcc.it:8443/162/171 (oph_merge)
      https://ophidia.cmcc.it:8443/162/172 (oph_aggregate2)
        https://ophidia.cmcc.it:8443/162/173 (oph_rollUp)
          https://ophidia.cmcc.it:8443/162/174 (oph_reduce)
            https://ophidia.cmcc.it:8443/162/175 (oph_reduce)
              https://ophidia.cmcc.it:8443/162/176 (oph_aggregate)
                https://ophidia.cmcc.it:8443/162/177 (oph_aggregate)
```

ECAS Terminal



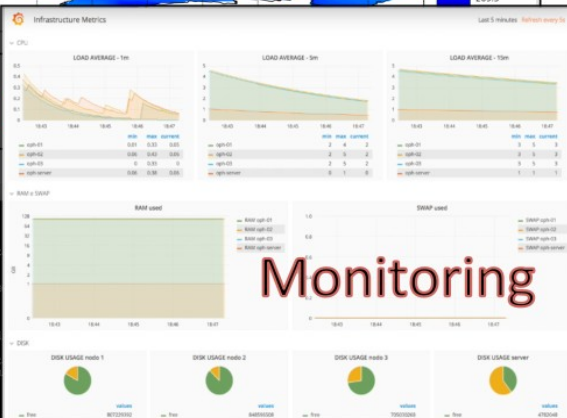
```

Warning: There is no session to resume
Resuming last session...
Getting list of Ophidia operators XML files from "ht
rators.xml"... Done.
Downloading necessary files... Done.
Remote XML files: 67 - Downloaded XML files: 67 - Re
Oph_Term - the Ophidia shell, version 1.0.0
Copyright (C) 2012-2017 CMCC Foundation - www.cmcc.it
This program comes with ABSOLUTELY NO WARRANTY; for
This is free software, and you are welcome to redist
under certain conditions; type 'conditions' for deta
Welcome to Oph_Term !
Use the power of the Ophidia framework right from yo
If you are going to use Oph_Term for the first time and need something
to get you started, just try entering "help"

[OPH_TERM] >>> oph_list

```

Monitoring



Quick Start

OphidiaLab provides two different ways to get access to its scientific eco-system: JupyterHub and Ophidia client.

Jupyter supports interactive data science and scientific computing. OphidiaLab includes a JupyterHub installation and, thanks to the Jupyter Notebooks, scientists can create and share documents that contain live code, equations, visualizations and explanatory text.

The JupyterHub interface is available [here](#).*

After you login, open "Quick Start.ipynb" notebook available under the *quickstart/* folder in your home to get started with OphidiaLab environment capabilities.

*Please note that for security reasons, the access to our JupyterHub instance is restricted to authorised users only and needs an additional step after the registration process.



QuickStart

The Ophidia Terminal is a robust, comprehensive, and user-friendly Ophidia client, developed with characteristics similar to the bash shell present in almost all Unix-like environments. Please have a look at the online available documentation to learn more about the basic functionalities of the **Ophidia terminal** as well as some **advanced features** useful for more skilled users.

Two short guides (**basic**, **advanced**) in pdf format are also available.

Several examples of real-world usage of the terminal are also available on the Ophidia website tutorial section.

The latest client RPM for CentOS is available [here](#).

The related DEB package can be downloaded from [here](#).

Once installed you can simply run:

```

/usr/local/ophidia/oph-terminal/bin/oph_term -H
ophidiolab.cmcc.it -u <username> -p <password> -P 11732

```


Thank you for your attention!

Questions?

Sofiane Bendoukha bendoukha@dkrz.de
Donatello Elia donatello.elia@cmcc.it
Fabrizio Antonio fabrizio.antonio@cmcc.it



EOSC-hub

 eosc-hub.eu  [@EOSC_eu](https://twitter.com/EOSC_eu)



This material by Parties of the EOSC-hub Consortium is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

● Training materials

- <https://github.com/ECAS-Lab/ecas-training>

● ECASLab / JupyterHUB

- **ECASLab @ DKRZ** <https://ecaslaboratory.dkrz.de>
- **ECASLab @ CMCC** <https://ecaslaboratory.cmcc.it>

● Ophidia framework documentation

- <http://ophidia.cmcc.it/documentation/users/index.html>

- Share your results with your team or with researchers from a broader community
 - Post-processing Datasets
 - Notebooks
- Cloud-based storage services integrated within ECAS
 - **B2DROP** online storage for external collaborators as well as for collaborators who uses B2DROP too. Keep data synchronized and up-to-date.
 - **B2SHARE** store and publish research data from diverse contexts. Data assigned a **Persistent Identifier** for better findability.
 - **DataHub (OneData)** a global data access solution for eScience.



ONE≡DATA

<https://b2drop.eudat.eu/s/gDyJjMeJ2Xiapwi>



Files Running Clusters

Select items to perform actions on them.

0 /

b2drop-private

b2drop-shared

b2share

conf

data

notebook



All files

Recent

Favorites

Shares

Tags

ecas

	Name	Size	Modified
<input type="checkbox"/>	TestNotebook.ipynb	< 1 KB	5 months ago
<input type="checkbox"/>	Time_series_... .ipynb	2 KB	4 months ago
<input type="checkbox"/>	Tropical+Nights.ipynb	4 KB	11 minutes ago

Share a Notebook

You want to share this notebook?

Jupyter extension

Share

Cancel

Connect to the remote ECAS instance

```
In [ ]: from PyOphidia import cube
cube.Cube.setclient(read_env=True)
```

Import source data (minimum temperature °K)

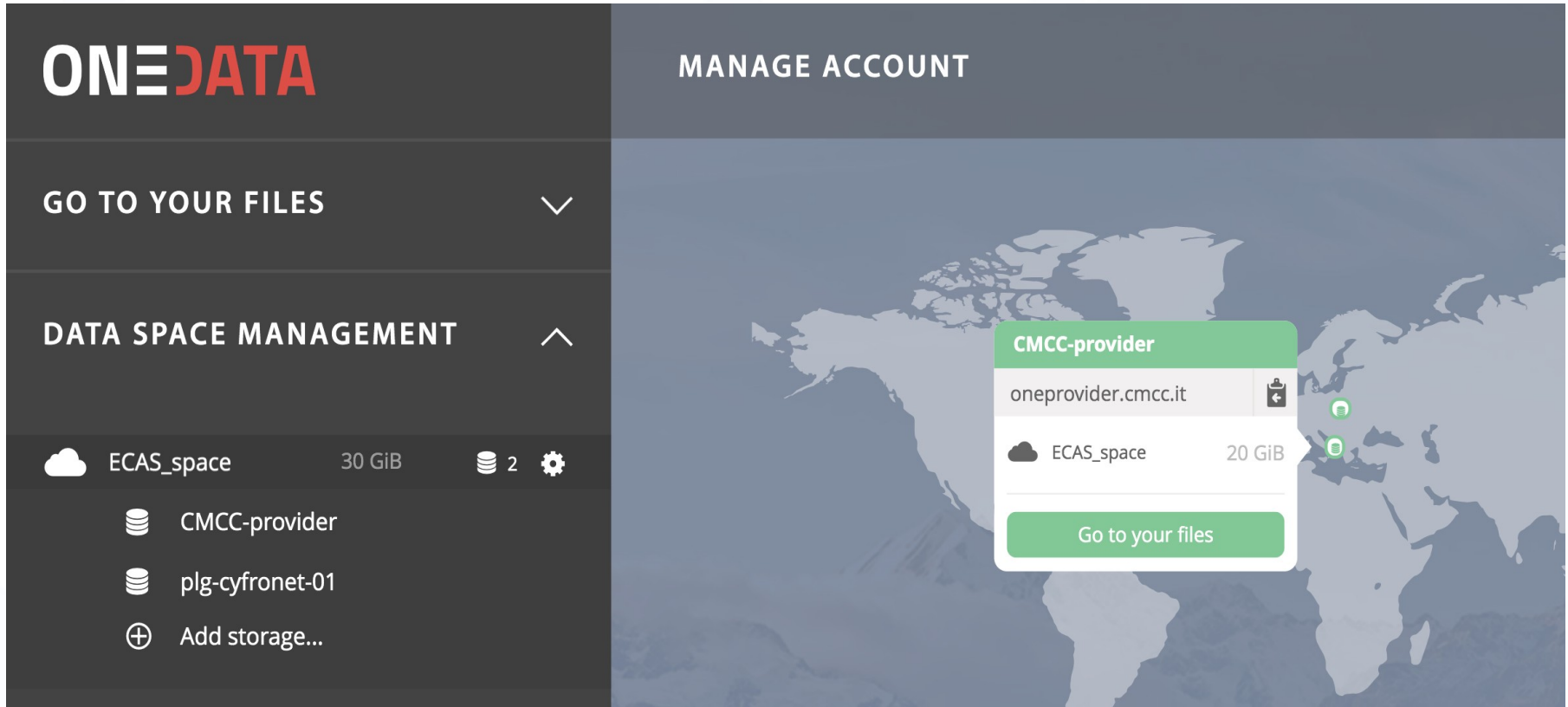
a month ago

hs ago

hs ago

ys ago

B2DROP Web site



The screenshot displays the ONE DATA web interface. On the left is a dark sidebar with the ONE DATA logo at the top. Below the logo are two main menu items: 'GO TO YOUR FILES' with a downward arrow, and 'DATA SPACE MANAGEMENT' with an upward arrow. Under 'DATA SPACE MANAGEMENT', there is a list of data spaces. The first entry is 'ECAS_space' with a cloud icon, '30 GiB' storage, and '2' datasets. Below it are 'CMCC-provider' and 'plg-cyfronet-01', each with a database icon. At the bottom of the list is an 'Add storage...' option with a plus icon. The main content area on the right is titled 'MANAGE ACCOUNT' and features a world map background. A modal window is open over the map, titled 'CMCC-provider'. It contains the URL 'oneprovider.cmcc.it' with a clipboard icon, a cloud icon next to 'ECAS_space' with '20 GiB' storage, and a green button labeled 'Go to your files'.

ONE DATA

MANAGE ACCOUNT

GO TO YOUR FILES

DATA SPACE MANAGEMENT

ECAS_space 30 GiB 2

CMCC-provider

plg-cyfronet-01

Add storage...

CMCC-provider

oneprovider.cmcc.it

ECAS_space 20 GiB

Go to your files

Check the notebook [here](#) for more details