



Workshop:  
**Weather and  
climate  
in the cloud**



# Data Tailor Web Service: cloud-native data customisation at EUMETSAT

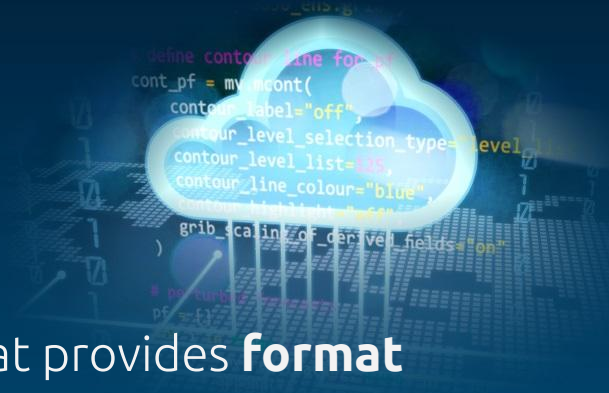
Alessandro Amici - [@alexamici](https://twitter.com/alexamici)  
B-Open - <https://bopen.eu>

# Agenda

- What is the EUMETSAT Data Tailor?
- How does it fit in the pilot EUMETSAT Data Services?
- Integration among cloud services
  - identity federation
  - centralised authorisation and accounting
  - resources scalability
  - service-to-service communication via APIs
  - integrate complex end-user interactions across services



# Data Tailor



The EUMETSAT Data Tailor is an Open Source tool that provides **format conversion** and basic **product customisation** capabilities for a set of **EUMETSAT products**

- Format conversion to well-known scientific and GIS formats
- Aggregation of native products
- Layer / band filtering
- Extraction of a region of interest (ROI)
- Reprojection and resampling

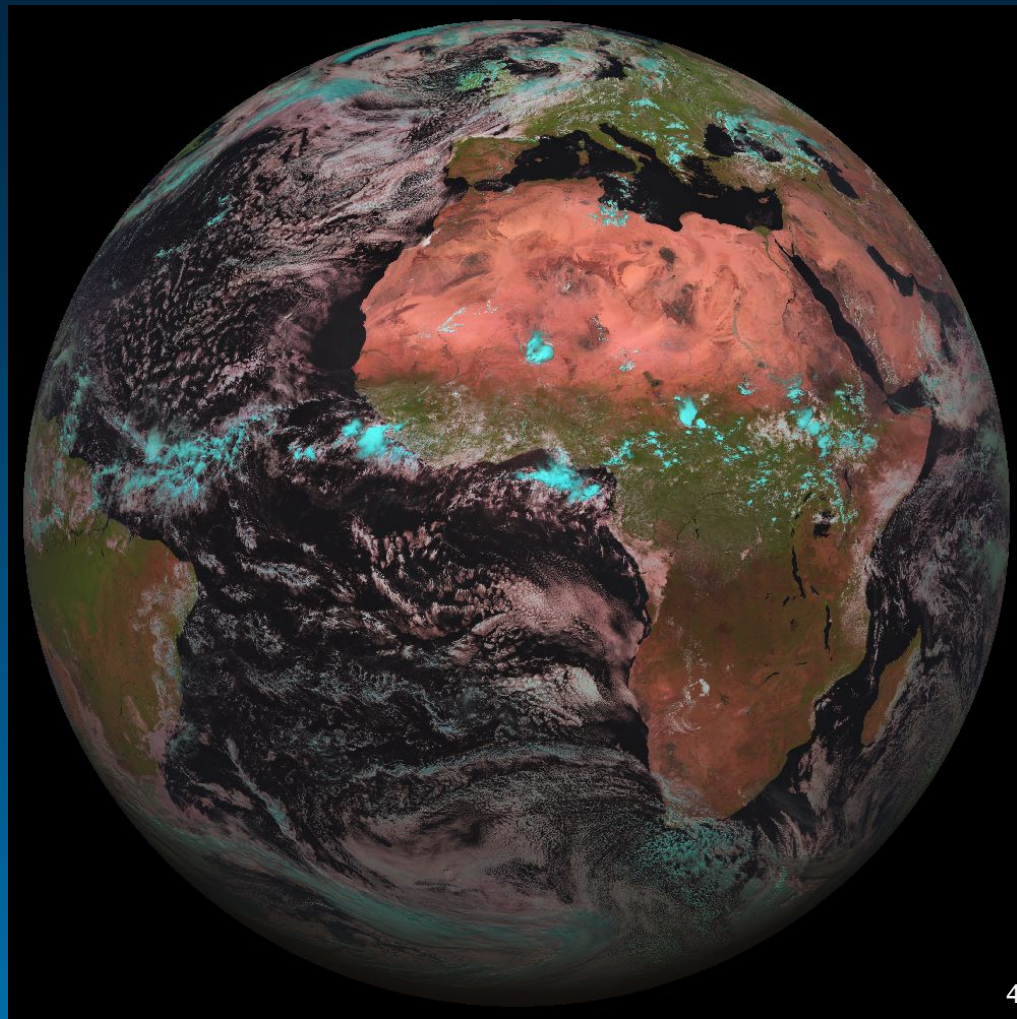
Home: <https://www.eumetsat.int/data-tailor>

```
define contour_line for pf
cont_pf = mv_mcont(
  contour_label="off",
  contour_level_selection_type="level",
  contour_level_list=10,
  contour_line_colour="blue",
  contour_line_style="solid",
  grib_scaling_of_derived_fields="on"
)
# plot contour lines
pf = ff
cont = cont_pf
```

EUMETSAT archives “native” satellite products:

- data formats
- geometry
- granularity

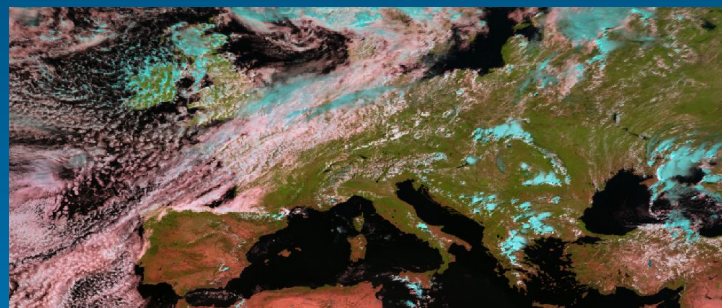
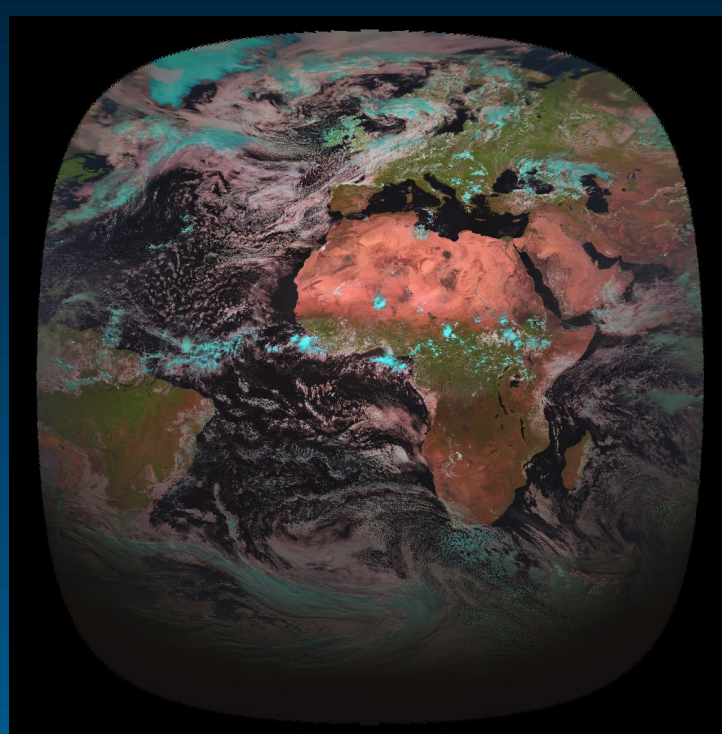
e.g. METEOSAT / SEVIRI 12-bands full-disc.



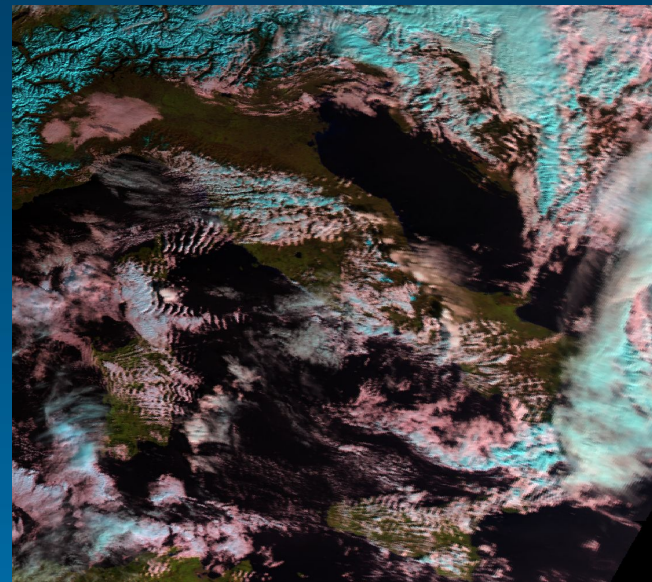
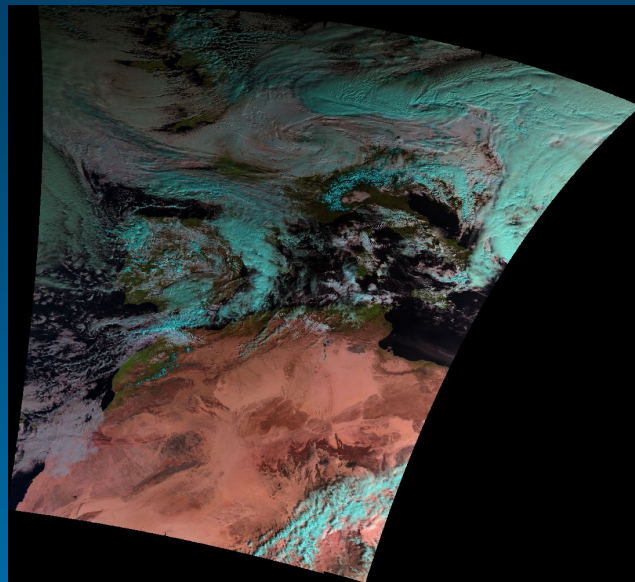
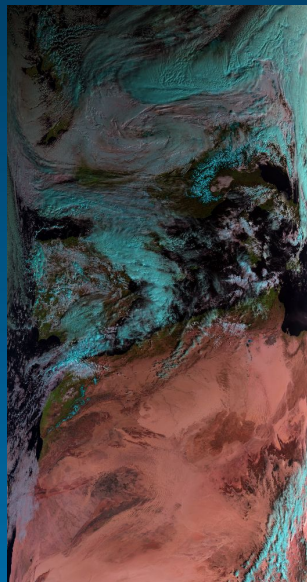
```
def define_contour_line_for_pf
  cont_pf = my_mcont(
    contour_label="off",
    contour_level_selection_type="level",
    contour_level_list=[10],
    contour_line_colour="blue",
    contour_line_width=2,
    grib_scaling_of_derived_fields="on"
  )
  # print(contour_line)
  pf = pf
end
```

With the Data Tailor a user can:

- select bands
- project to lat-lon
- extract a geographic region of interest
- convert to a scientific or GIS data format



The Data Tailor can customise products of geostationary and polar satellites and higher level SAF products, e.g. MetOp / AVHRR 6-channels



# Technologies



The Data Tailor is built on top of the **Python** scientific and geospatial ecosystem:

- core: **Dask**, **Xarray**, netcdf4, NumPy, Falcon, PyJWT
- GUI: React / JavaScript
- plugins: **GDAL**, SciPy, Pandas, cfrib / ecCodes, h5py

The Data Tailor is also one of the first EUMETSAT project that went full Open Source: <https://gitlab.eumetsat.int/open-source/data-tailor>

# Products supported by the Data Tailor

- AMSU-A GDS Level 1B - Metop
- ASCAT Level 1 Sigma0 at Full Sensor Resolution - Metop - Global Data Service
- ASCAT GDS Level 1 Sigma0 Swath Grid - Metop
- ASCAT Soil Moisture Swath Grid in NRT - Metop
- AVHRR GDS Level 1B - Metop
- Surface Solar Radiation Data Set - Heliosat (SARAH) - Edition 2.1
- ECMWF - Near-surface air temperature
- Global L3C AVHRR Sea Surface Temperature - Metop
- GOME-2 GDS Level 1B - Metop
- Polar Multi-Sensor Aerosol Optical Properties - Metop
- GPCC - Thickness of the liquid water equivalent total precipitation
- HIRS GDS Level 1B - Metop
- High Rate SEVIRI Level 1.5 Image Data - MSG
- Scatterometer Root Zone Soil Moisture (RZSM) Data Record 10km resolution - Multimission
- IASI GDS Level 1C - All Spectral Samples - Metop
- IASI Combined Sounding Products - Metop
- LSA SAF - Climate Data Records - Fraction of absorbed photosynthetically active radiation
- LSA SAF - Climate Data Records - Fractional vegetation cover
- LSA SAF - Climate Data Records - Leaf area index
- Land Surface Temperature Climate Data Record - MSG
- LSA SAF - Climate Data Records - Reference evapotranspiration
- LSA SAF - Climate Data Records - Normalized difference vegetation index
- LSA SAF - Near-Real Time - Land surface temperature
- LSA SAF - Near-Real Time - Reference evapotranspiration
- Daily Fraction of Absorbed Photosynthetic Active Radiation - MSG
- Evapotranspiration - MSG - 0 degree
- Reference Evapotranspiration - MSG
- MHS GDS Level 1B - Metop
- Land Surface Temperature Climate Data Record - MSG
- Atmospheric Motion Vectors - MSG
- Cloud Analysis - MSG
- Cloud Mask - MSG
- Active Fire Monitoring (GRIB) - MSG
- Multi-Sensor Precipitation Estimate (GRIB) - MSG
- Optimal Cloud Analysis - MSG
- Flexible Combined Imager (FCI) Level 1C - MTG
- Flexible Combined Imager (FCI) Level 2 Global Instability Index (GII) - MTG
- IRS Level 1 Principal Component - MTG
- HRI Level 1.5 Image Data - MFG
- ASCAT Winds and Soil Moisture at 25 km Swath Grid - Metop
- ASCAT Winds at 25 km Swath Grid - Metop
- ASCAT Coastal Winds at 12.5 km Swath Grid - Metop
- ASCAT L2 25 km Winds Data Record Release 1 - Metop
- ASCAT L2 12.5 km Winds Data Record Release 1 - Metop
- ERS Scatterometer L2 25 km Winds Data Record Release 1 - ERS
- SeaWinds L2 25 km Winds Data Record Release 1 - QuikSCAT





# DTWS in EUMETSAT Data Services

## Pilot EUMETSAT Data Services:

- EUMETSAT Data Store - <https://data.eumetsat.int>
  - catalogue and archive
- EUMETSAT Data Tailor - <https://tailor.eumetsat.int>
  - product customisation service
- EUMETView - <https://view.eumetsat.int>
  - online map service



# Data Tailor Web Service



The Data Tailor Web Service forms part of the pilot EUMETSAT Data Services and is the adapter between the Data Store and the end-users:

- download only the data they need (bands and spatial area)
- download ready-to-use data (data type, projection and format)

# Pilot Data Services: user journey 1/6

```
define contour_line for set  
cont_pf = my_mcont(  
  contour_label="off",  
  contour_level_selection_type="level",  
  contour_level_list=135,  
  contour_line_colour="blue",  
  grib_section_list=644,  
  grib_section_list_of_derived_fields="on"
```

data.eumetsat.int/search?query=&results=19

EUMETSAT DATA SERVICES

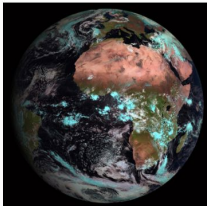
Cart 0 items alexamici

Data Access / Search results

## Data Store

e.g. "Atlantic Sea"

Perform an advanced search >

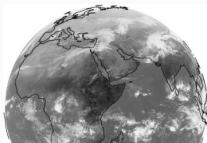


### High Rate SEVIRI Level 1.5 Image Data - MSG - 0 degree

Rectified (level 1.5) Meteosat SEVIRI image data. The data is transmitted as High Rate transmissions in 12 spectral channels. Level 1.5 image data corresponds to the geolocated and radiometrically pre-processed image data, ready for further processing, e.g. the extraction of meteorological products. Any spacecraft specific ...

[learn more](#)

Access Data



### High Rate SEVIRI Level 1.5 Image Data - MSG - Indian Ocean 41.5 degrees E

Rectified (level 1.5) Meteosat SEVIRI image data. The data is transmitted as High Rate transmissions in 12 spectral channels. Level 1.5 image data corresponds to the geolocated and radiometrically pre-processed image data, ready for further processing, e.g. the extraction of meteorological products. Any spacecraft specific ...

# Pilot Data Services: user journey 2/6

The screenshot shows the EUMETSAT Data Services interface. The browser address bar displays the URL: `data.eumetsat.int/data/map/EO:EUM:DAT:MSG:HRSEVIRI?start=2021-02-08T14:27:43.895Z&end=2021-0...`. The page header includes the EUMETSAT logo, 'DATA SERVICES', a shopping cart icon with '0 Items', and a user profile for 'alexamici'. The breadcrumb trail reads: `Data Access / Search results / High Rate SEVIRI Level 1.5 Image Data - MSG - 0 degree`.

The main content area is divided into a left sidebar and a right main panel. The sidebar contains the following elements:

- AVAILABLE TIME RANGE:** 2017-01-01 - 2021-02-09
- Two date range selectors, both set to 2021-02-08 and 2021-02-09, with time filters set to 14:27.
- Sort by:** Sensing time (Descending)
- Filters:** Mission / Satellite
- Buttons:** Reset Filters, Show Results

The right main panel displays search results:

- 1-20 of 97 products found
- Navigation: 1 2 ... 5 >
- Actions:  Select all in page, Add selected to Cart
- Table of products:

<input type="checkbox"/>	Time Range	MSG4	Size	Download	Cart
<input type="checkbox"/>	2021-02-09 (12:00) - 2021-02-09 (12:12)	MSG4	184.64 MB		
<input type="checkbox"/>	2021-02-09 (11:45) - 2021-02-09 (11:57)	MSG4	184.75 MB		
<input checked="" type="checkbox"/>	2021-02-09 (11:30) - 2021-02-09 (11:42)	MSG4	184.79 MB		
<input checked="" type="checkbox"/>	2021-02-09 (11:15) - 2021-02-09 (11:27)	MSG4	184.79 MB		
<input type="checkbox"/>	2021-02-09 (11:00) - 2021-02-09 (11:12)	MSG4	184.70 MB		
<input type="checkbox"/>	2021-02-09 (10:45) - 2021-02-09 (10:57)	MSG4	184.50 MB		
<input type="checkbox"/>	2021-02-09 (10:30) - 2021-02-09 (10:42)	MSG4	184.19 MB		
<input type="checkbox"/>	2021-02-09 (10:15) - 2021-02-09 (10:27)	MSG4	183.75 MB		

The footer contains: CONTACT, LEGAL INFORMATION, Follow us, and social media icons for Twitter, Facebook, YouTube, Instagram, RSS, and LinkedIn.

# Pilot Data Services: user journey 3/6

```
define contour_line for sat  
cont_pf = my_mcont(  
  contour_label="off",  
  contour_level_selection_type="level",  
  contour_level_list=105,  
  contour_line_colour="blue",  
  grib_cal_in/of_derived_fields="on"
```

data.eumetsat.int/data/map/EO:EUM:DAT:MSG:HRSEVIRI/cart

EUMETSAT DATA SERVICES

Cart 2 Items alexamici

Data Access / Search results / High Rate SEVIRI Level 1.5 Image Data - MSG - 0 degree

### Cart (2 Items)

[Return to Map view](#)

Customize Cart Download Cart Remove Selected X Clear Cart X

<input type="checkbox"/>	Product Name	Satellite	Product Type	Sensing Start (UTC)	Sensing Stop (UTC)	Download	Remove
<input type="checkbox"/>	MSG4-SEVI-MSG15... <i>i</i>	MSG4	MSG15	2021-02-09 11:30	2021-02-09 11:42		X
<input type="checkbox"/>	MSG4-SEVI-MSG15... <i>i</i>	MSG4	MSG15	2021-02-09 11:15	2021-02-09 11:27		X

CONTACT LEGAL INFORMATION Follow us

# Pilot Data Services: user journey 4/6

The screenshot shows the EUMETSAT Launchpad web interface. The browser address bar displays `tailor.eumetsat.int/launchpad`. The page header includes the EUMETSAT logo and navigation tabs: LAUNCHPAD, AGGREGATION, LAYER FILTER, REPROJECTION, ROI, QUICK LOOK, and OUTPUT OPTIONS. The main content area is titled "Launchpad" and contains three sections:

- Product type:** High Rate SEVIRI Level 1.5 Image Data - MSG
- Output format:** GeoTiff
- Configuration:** Natural color disc

A configuration popup window is open, displaying the following details:

```
product: HRSEVIRI
name: Natural color disc
filter: hrseviri_natural_color
format: geotiff
```

The **Input products** section lists two files:

- MSG4-SEVI-MSG15-0100-NA-20210209114243.404000000Z-NA.zip
- MSG4-SEVI-MSG15-0100-NA-20210209112743.140000000Z-NA.zip

At the bottom, there is a vertical sidebar with expandable sections: Aggregation, Layer Filter, Reprojection, ROI, and Quick look.

```
define contour_line for pt
cont_pf = my_cont(
  contour_label="off",
  contour_level_selection_type="level",
  contour_level_list=100,
  contour_line_colour="blue",
  contour_line_width=2,
  grib_cal_in/of_derived_fields="on"
```

# Pilot Data Services: user journey 5/6

tailor.eumetsat.int/reprojection

LAUNCHPAD AGGREGATION LAYER FILTER **REPROJECTION** ROI QUICK LOOK OUTPUT OPTIONS

Launchpad

Aggregation

Layer Filter

Reprojection

Target projection  
Geographic / Plate-Carrée

- Gall Peters
- Gall Stereographic
- Geographic / Plate-Carrée**
- Goode Homolosine
- Lambert cylindrical equal-area
- Mercator

Target resolution (degree)  
Xres Yes  
1 1

Resampling algorithm  
Bilinear resampling

```
define contour_line for ...  
cont_pf = my_contour(  
  contour_label="off",  
  contour_level_selection_type="level",  
  contour_level_list=100,  
  contour_line_colour="blue",  
  grib_subfields=64,  
  grib_calin_of_derived_fields="on"
```

# Pilot Data Services: user journey 6/6

```
define contour_line for set
cont_pf = my_cont(
  contour_label="off",
  contour_level_selection_type="level",
  contour_level_list=100,
  contour_line_colour="blue",
  contour_line_width=2,
  grib_cal/in/of_derived_fields="on"
```

tailor.eumetsat.int/output-options

EUMETSAT LAUNCHPAD AGGREGATION LAYER FILTER REPROJECTION ROI QUICK LOOK OUTPUT OPTIONS

Launchpad

Aggregation

Layer Filter

Reprojection

ROI

Quick look

Output Options

Customisation ID	Size	Status	Progress/Time
<input type="checkbox"/> 7e290dd0	272.8KB	✓ Completed	↓ 1m 59s
<input type="checkbox"/> 8aa7467f	272.5KB	✓ Completed	↓ 10s

460.5MB of 100.0GB 0%

Delete selected

Log

```
2021-02-09 14:50:20 - PROCESSING.submission_helpers[87] - INFO - Sub
2021-02-09 14:50:20 - PROCESSING.api[393] - INFO - FUTURE: run_chain
```



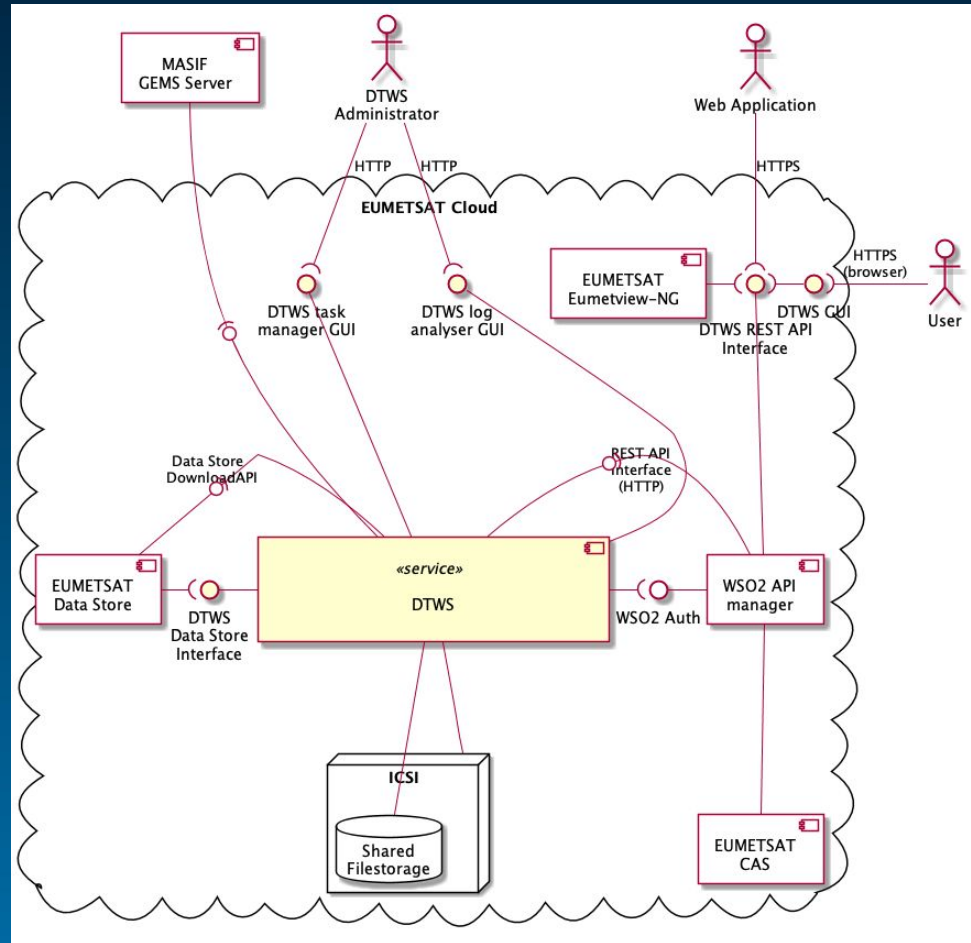
```

define contour_line_for_pf
cont_pf = my_contour(
  contour_label="off",
  contour_level_selection_type="level",
  contour_level_list=10,
  contour_line_colour="blue",
  contour_line_style="solid",
  grib_scaling_of_derived_fields="on"
)
# print(contour_line_for_pf)
pf = pf

```

## Internal cloud services:

- scalable computing and storage (ICSI)
- monitoring service (MASIF)
- identity service (CAS)
- API management service (WSO2)
  - authorisation and accounting
  - OpenAPI



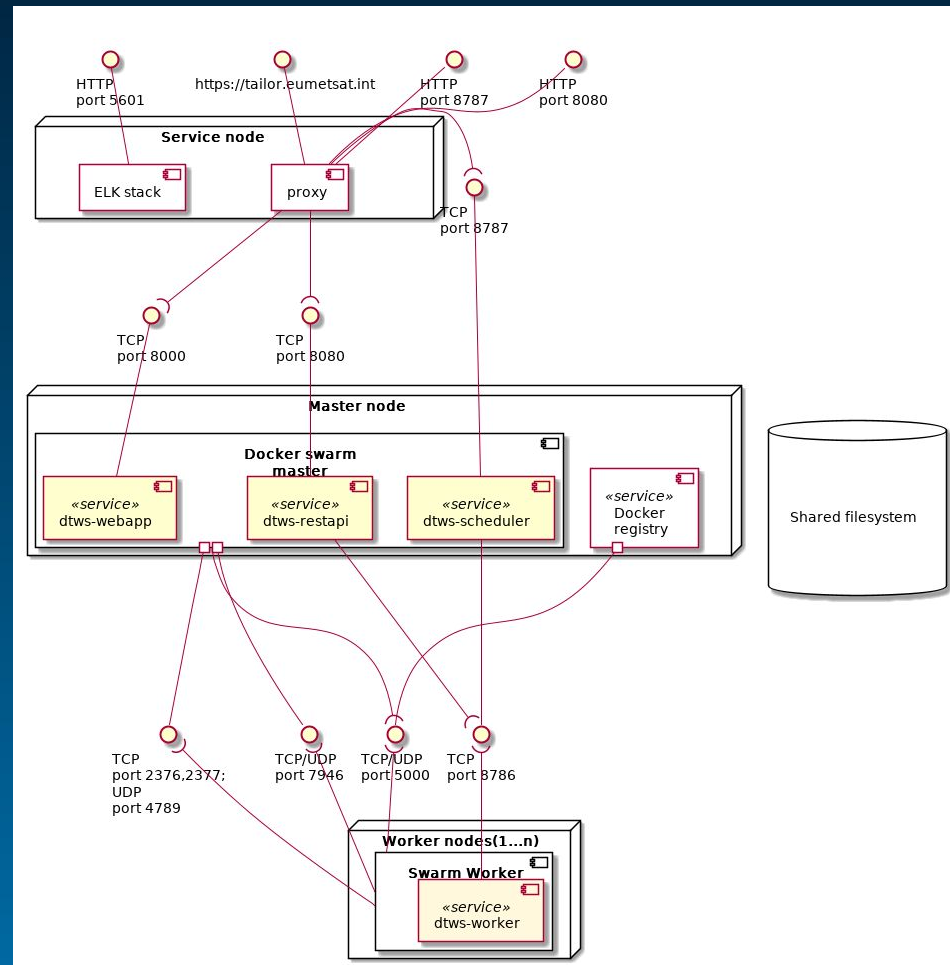
```

define contour_line for pf
cont_pf = my_contour(
  contour_label="off",
  contour_level_selection_type="level",
  contour_level_list=100,
  contour_line_colour="blue",
  contour_line_style="solid",
  grib_scaling_of_derived_fields="on"
)
# plot contour lines
pf = pf

```

## Data Tailor Web Service architecture:

- dtws-webapp (JS, React)
- dtws-restapi (Python, Falcon)
- dtws-scheduler (Dask)
- dtws-worker (Docker, Dask, Data Tailor)



# Cloud integration: AAA

OAuth unify Authentication, authorisation and accounting:

- user logs into the Data Store GUI and search the catalogue
  - authentication via CAS, API authorisation and accounting from WSO2
- Data Store directs the user to the Data Tailor GUI
  - identity passes transparently to the DTWS (CAS)
  - DTWS accesses the native products as the user (WSO2)
  - user configuration, processing and storage quotas and access to customised products are authorised and accounted by the DTWS\* (room for improvement)



# Cloud integration: computing

- Scalable pool of VMs (ICSI)
- Orchestration via dedicated Docker Swarm\* (will migrate to ICSI Kubernetes Service):
  - dedicated Dask scheduler
  - scalable pool of Dask workers
  - DTWS maintains per-user processing quotas
- Customisation are neither cached nor shared between users\*



# Cloud integration: storage

- Scalable NFS share (ICSI)
- DTWS manages users workspaces\*:
  - DTWS maintains per-user product storage quotas



# Cloud integration: GUI



The Data Store and the Data Tailor Web Service have dedicated GUIs

- Single sign on experience
- Passing user configuration between services
  - Data Store selection
- Crossing GUI boundaries\*
  - look and feel, widget set
  - cart / download experience not integrated with the Data Store



# Thank you



B-Open - <https://bopen.eu>

Alessandro Amici - [@alexamici](https://twitter.com/alexamici)

Key persons: Daniel Lee (EUMETSAT) and Maurizio Bottaccio (B-Open)

EUMETSAT Data Tailor: <https://www.eumetsat.int/data-tailor>

EUMETSAT Data Services to go operational in 2021!