Royal Netherlands
Meteorological Institute
Ministry of Infrastructure
and Water Management

SATPROC

satellite processing on the European weather cloud

Eelco Verduijn / RDSW Agile Team

EWC User Workshop 22 Sept 2026

Topics

Who are we

Overview of value chain

Our CI/CD setup for EUWC

What does SATPROC solve

Questions

RDSW Agile team: who are we?

Team responsible for running and maintaining RDSW value chains.

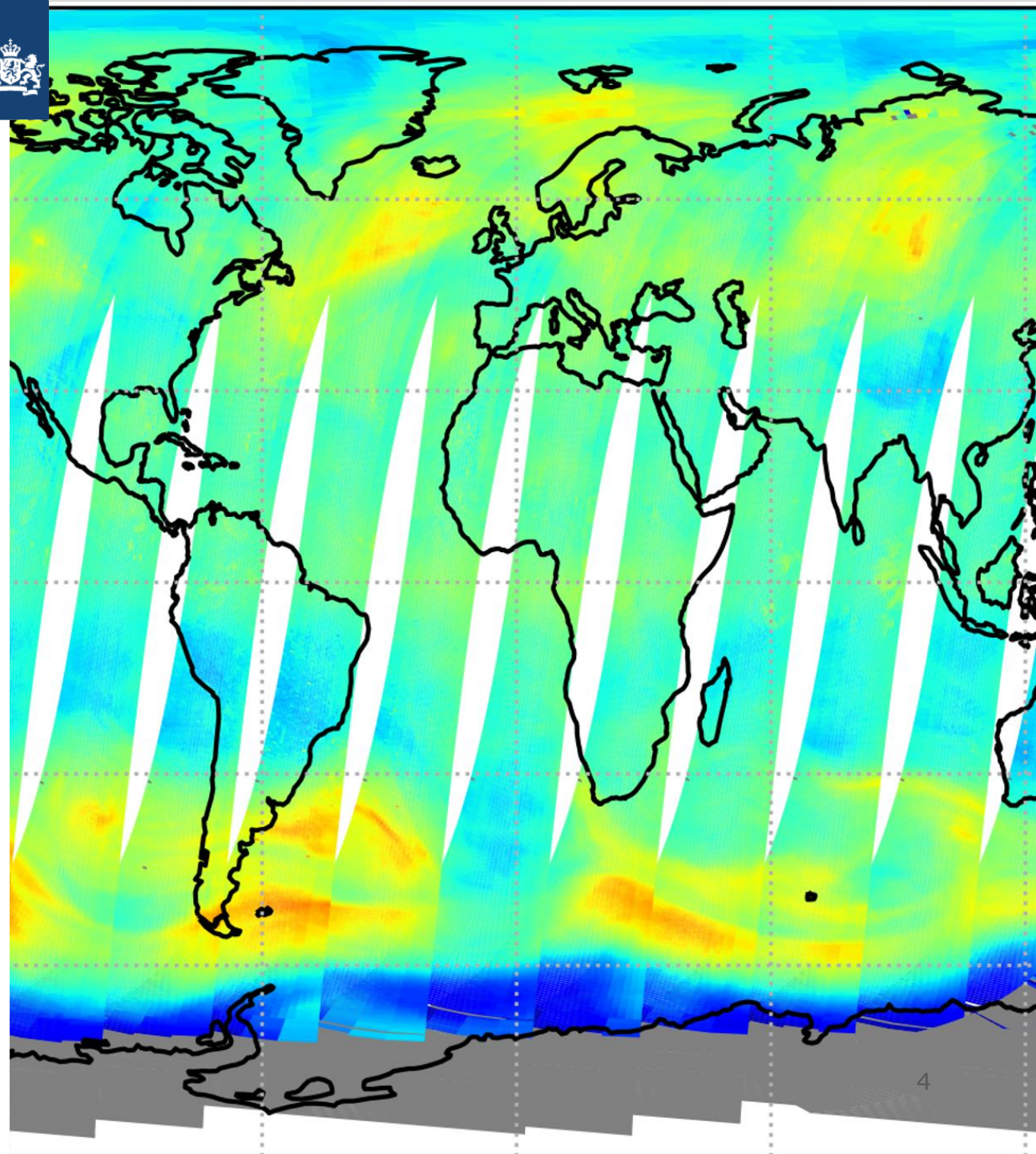
- OSI SAF
- AC SAF
- Satellite Composites
- …

- **Rianne Giesen**
- **Ruben Ijpelaar**
- **Gijsbert Tilstra**
- **Olaf Tuinder**
- **Paul de Valk**
- **Eelco Verduijn**
- **Anton Verhoef**
- **Robert van Versendaal**

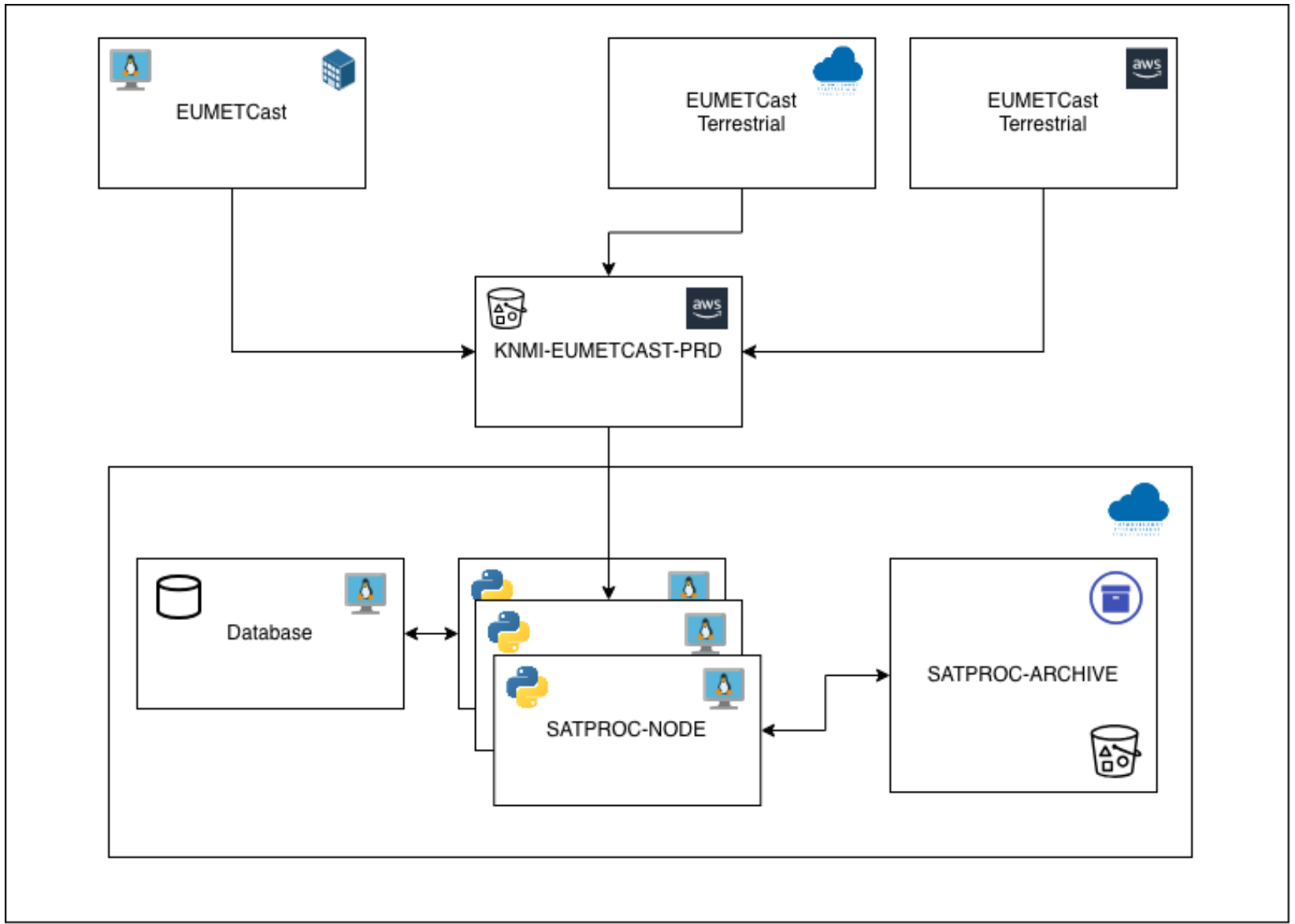
A value chain

- The KNMI is part of the **AC SAF**

- **SAF**: Satellite Application Facilities (SAFs)
- Dedicated centres of excellence for processing satellite data



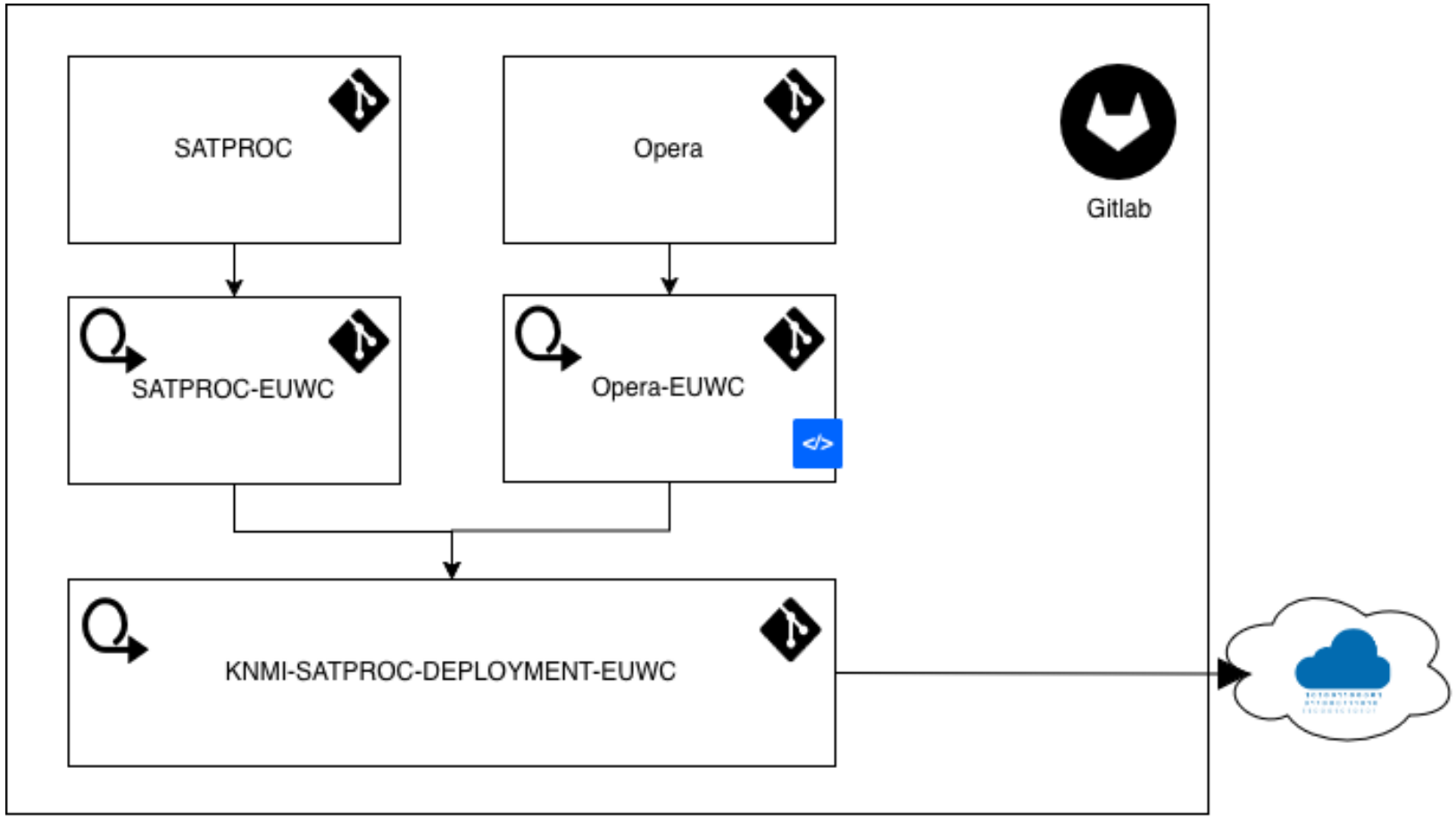
High Level Design



deployment

- Gitlab-opentofu-component
- Gitlab for terraform state
- Cloud-init for initialization

Repository Overview



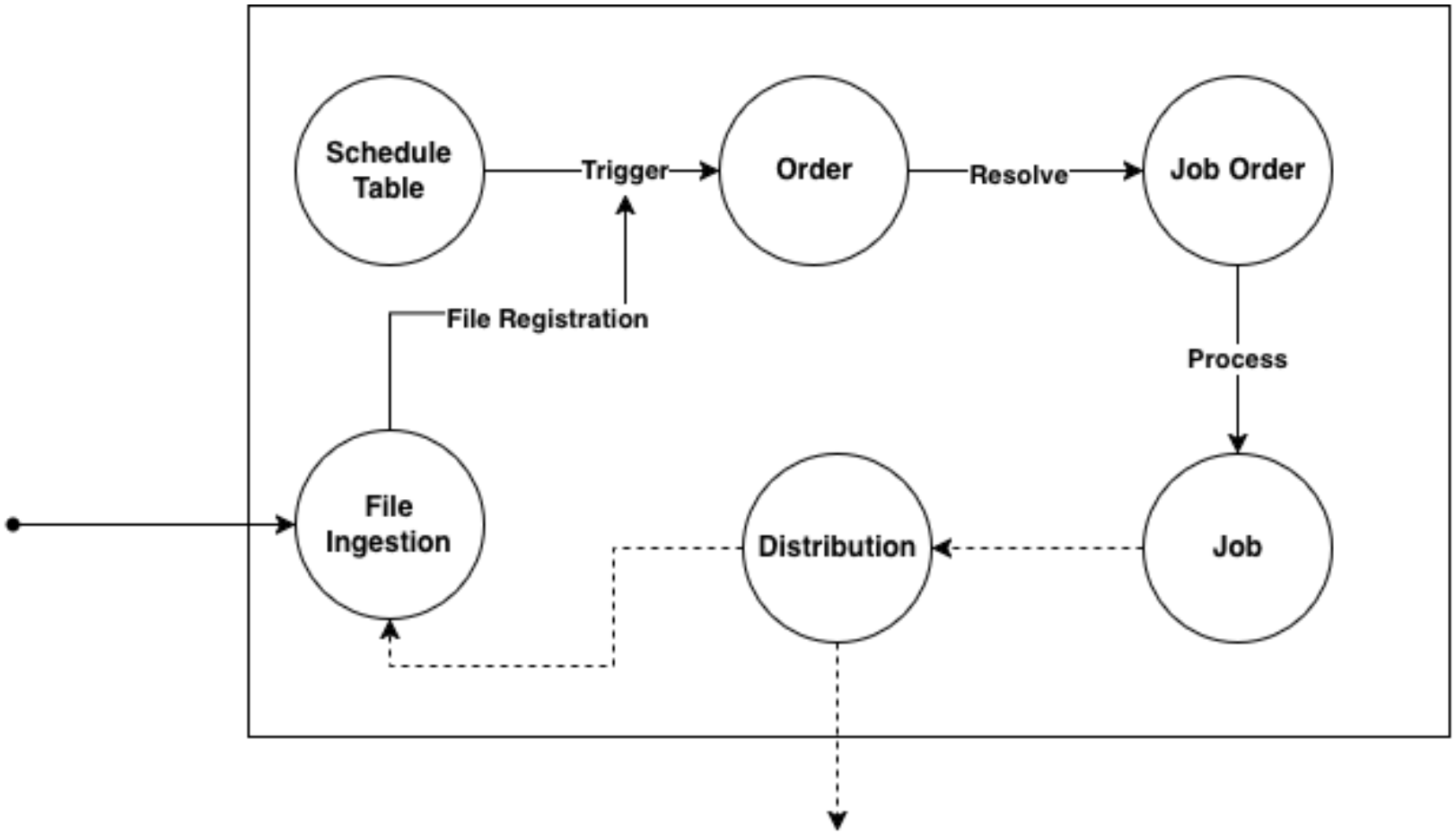
What is **SATPROC**?

Software for ingesting, scheduling, processing and distributing data

Inspired by IPF (ESA standard) and SQS

- Schedules containers to process jobs
- Cloud agnostic (custom worker implementations)
- Use S&T’s muninn for file metadata (open-source)
- Yet more YAML

What is **SATPROC**?



What problems does **SATPROC** solve?

GOME_xxx_1B_M03_20260909111758Z_20260909112058Z_N_O_20260909115814Z

Filename contains:

- satellite
- observation start
- observation end
- creation date

Not sufficient to make decisions.

What problems does **SATPROC** solve?

Are there valid observation in this file?

What orbit does it belong to?

Do we have a complete orbit?

What is the nearest forecast to this observation?

What if input isn’t available?

What problems does **SATPROC** solve?

python script that ingests data

Identify(filename: str): -> bool

Analyze(filename: str, filename_only: bool=False) -> muninn.struct

Product_type() -> str

Default schema:

- Validity start
- Validity stop
- Foot print
- Size
- Hash
- …

Custom namespace:

- Satellite
- Version
- Orbit
- Integrity
- …

Example: schedule table trigger

Are there valid observation in this file?

```
trigger:
     - archive:
              name: archive
             namespace: gome2l1b
     conditions:
        - if_all:
          - function: ValueEquals
            returns: True
            parameters:
              - 'GOOD'
              - ${trigger.gome2l1b.integrity}
```

Example: schedule table input

What is the nearest forecast to this observation?

```
- name: ecmwf_o3_forecast
      is_required: true
      resolver:
          - archive_name: archive
              product_type: ecmwfozoneforecast
     function: ValidityIntersect
     parameters:
      - ${trigger.core.validity_start - timedelta(minutes=1)}
      - ${trigger.core.validity_stop + timedelta(minutes=1)}
      sort_by:
        - property: ecmwfforecast.forecast_total_minutes
          mode: ascending
```

Example: schedule table input

Do we have a complete orbit?

```
inputs:
  - name: complete_orbit
    is_required: true
    resolver:
      - archive_name: archive
        product_type: gome2l2
        timeout: 6d
        function: gome2_resolvers.ResolveL2Orbit
        parameters:
          - ${trigger.gome2product.sub_product}
          - ${trigger.gome2product.orbit}
          - ${trigger.gome2product.satellite}
```

How much has it ran so far?

35 schedule tables, running in shadow

knmi-satproc-deployment-aws

1.000.000+ orders processed (shut-down)

Knmi-satproc-deployment-euwc

400.000+ orders processed since june

Summary

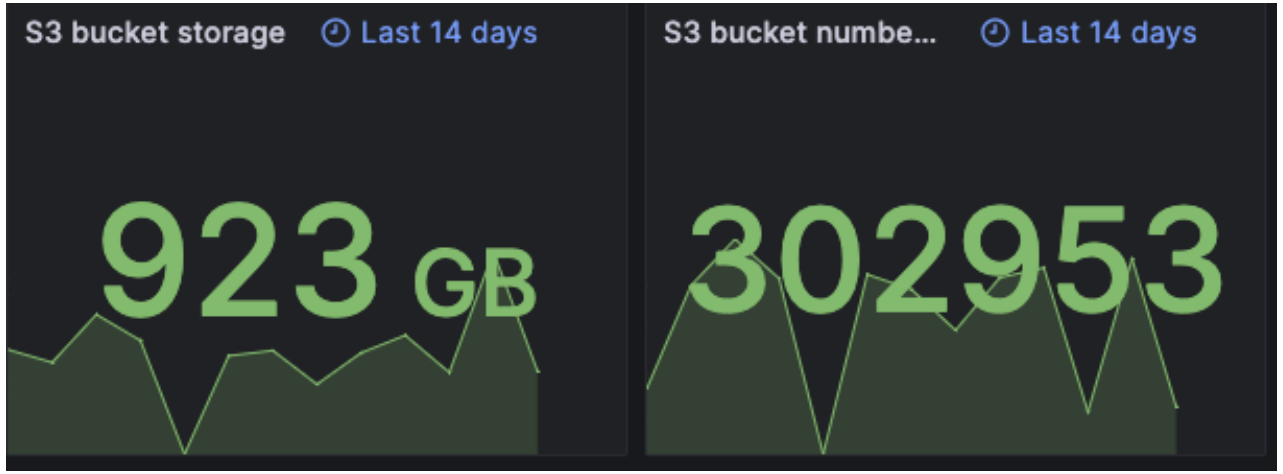
We receive satellite data and process it on the european weather cloud

To process our data we made satproc, to solve our specific challenges

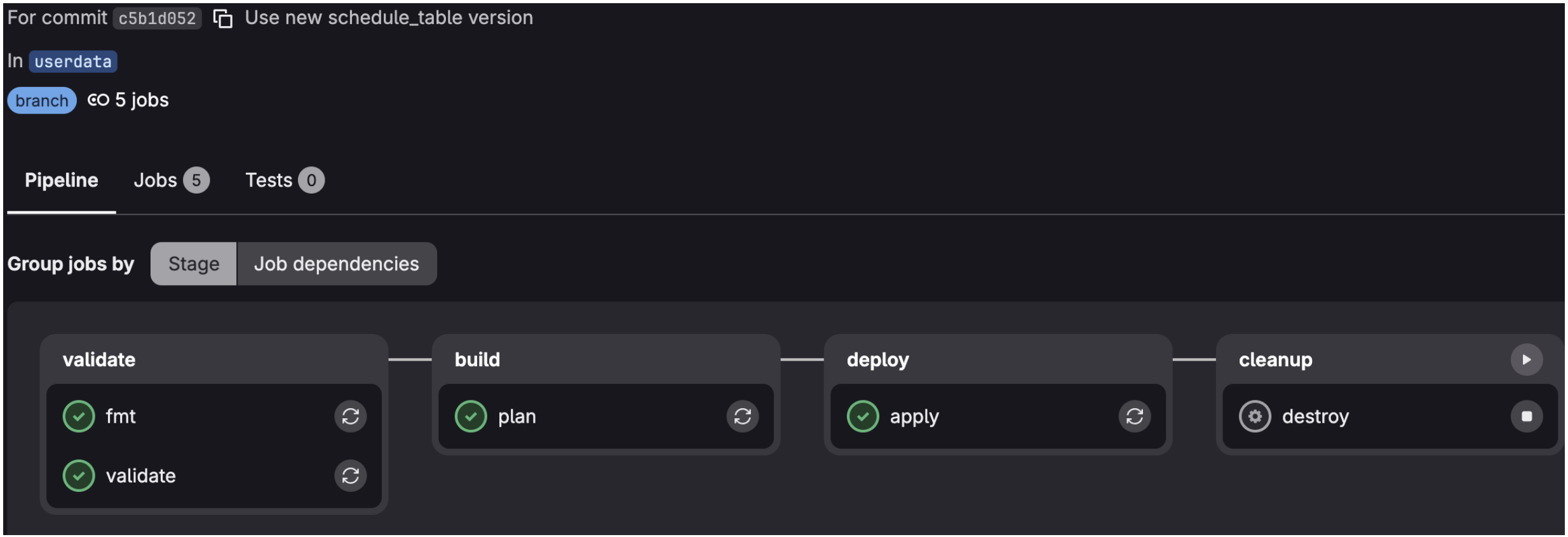
Satproc is deployed to the european weather cloud using CI/CD

Questions?

EUMETCast


S3 bucket storage
Last 14 days
923 GB
S3 bucket numbe...
Last 14 days
302953

deduplication warnings
400
300
200
100
0
08/27
08/30
09/02
09/05
09/08
warnings


For commit c5b1d052 Use new schedule_table version
In userdata
branch 5 jobs
Pipeline
Jobs 5
Tests 0
Group jobs by
Stage
Job dependencies
validate
fmt
validate
build
plan
deploy
apply
cleanup
destroy

Where does the data come from?

Copyright: ESA /AOES Medialab

Copyright: EUMETSAT

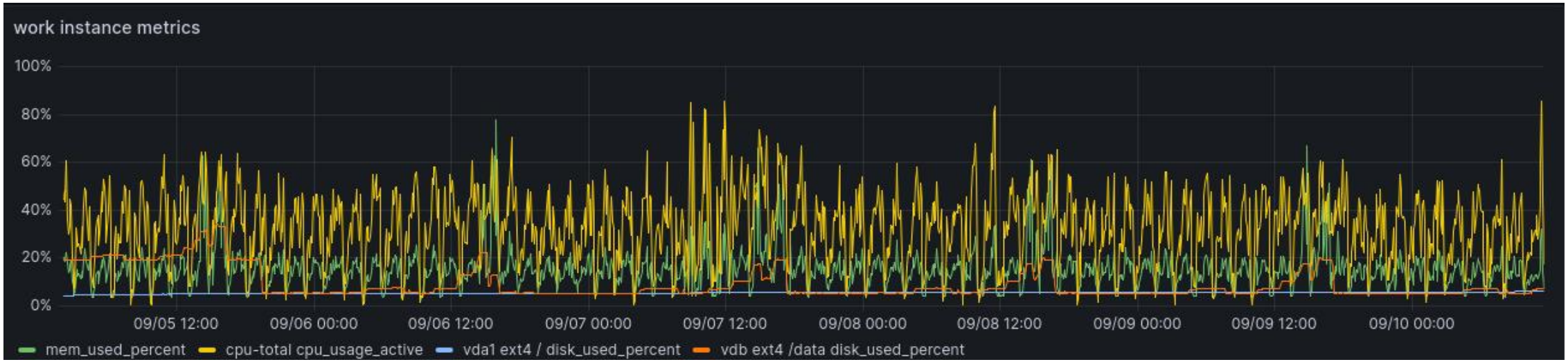
MetOp A: 2006-2021
MetOp B: 2012-2030
MetOp C: 2018-2033

SATPROC HOME SCHEDULE TABLE ORDER JOB ORDER JOB DISTRIBUTION ORDER SCHEDULE TABLE ORDER STATUS MUNINN CORE

METRICS

SCHEDULE TABLE ORDER STATUS

SCHEDULE TABLE NAME	RESOLVING (TOTAL)	NEW (24H)	RESOLVED (24H)	DONE (24H)	FAILED (24H)
gome2_l1b_bzip	0	943	0	943	0
gome2_cron_daily_l3_nap_gridder	0	2	0	2	0
gome2_l2_nar_nap_nfo	0	554	0	554	0
gome2_bufr_nhp_gts_converter	0	556	0	550	6
gome2_cron_hourly_nap_vaac	0	96	0	96	0
gome2_cron_daily_nap_l3_monthly_gridder	0	2	0	2	0
gome2_cron_daily_l3_nhp_gridder	0	2	0	2	0
gome2_cron_daily_nar_l3_monthly_gridder	0	2	0	2	0
gome2_cron_daily_nar_vaac	0	2	0	2	0
gome2_cron_daily_l3_nhp_monthly_gridder	0	2	0	2	0
gome2_spectral_drift	0	2	0	2	0
gome2_cron_daily_nhp_globalfield_plot	0	2	0	2	0
gome2_ohp_catter	19	28	0	25	0

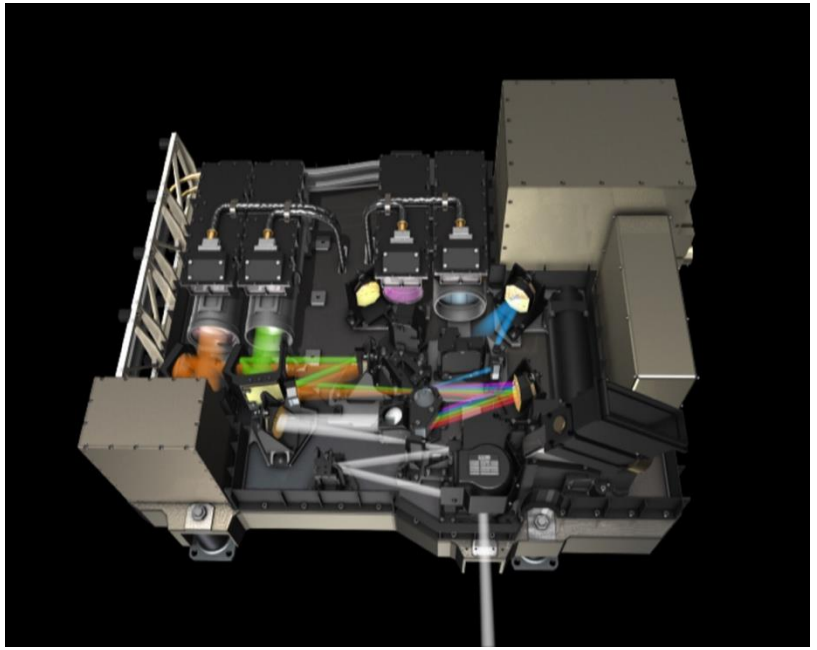

work instance metrics
100%
80%
60%
40%
20%
0%
09/05 12:00
09/06 00:00
09/06 12:00
09/07 00:00
09/07 12:00
09/08 00:00
09/08 12:00
09/09 00:00
09/09 12:00
09/10 00:00
mem_used_percent
cpu-total cpu_usage_active
vda1 ext4 / disk_used_percent
vdb ext4 /data disk_used_percent

SATPROC HOME SCHEDULE TABLE ORDER JOB ORDER JOB DISTRIBUTION ORDER SCHEDULE TABLE ORDER STATUS MUNINN CORE

MUNINN FILE

PROPERTY	VALUE
core.uuid	32e97b26-5c49-44dd-9c75-c75f08f1d631
core.active	True
core.hash	md5:17c1e6fe3c9e9d3d3275f1b9a883f2cd
core.size	1029855
core.metadata_date	2026-07-28 05:08:23.043082
core.archive_date	2026-07-28 05:08:22.930823
core.archive_path	metop/metopb/gome2/level1/gome2lerpmdl1_hdf5/2026/07/27
core.product_type	gome2lerpmdl1
core.product_name	GOME_xxx_1B_M01_20260727010556Z_20260727010856Z_N_O_20260727015920Z_LER_PMD
core.physical_name	GOME_xxx_1B_M01_20260727010556Z_20260727010856Z_N_O_20260727015920Z_LER_PMD.hdf5

Where does the data come from?



Copyright: ESA /AOES Medialab