European Weather Cloud: Batch Processing

EUMETSAT

Francesco Murdaca, Mike Grant

21st September 2023, HTCondor Workshop 2023, Orsay (France)

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED
ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF EUMETSAT

About authors

Francesco Murdaca

Cloud Application Developer

EUMETSAT

Mike Grant

Climate Data Processing Engineer

EUMETSAT

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED
ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   EUMETSAT

About authors, who you can trust, as they have no-colour slides

Francesco Murdaca

Cloud Application Developer

EUMETSAT

Mike Grant

Climate Data Processing Engineer

EUMETSAT

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED
ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF EUMETSAT

European Weather Cloud (EWC)

European Weather Cloud (EWC)

What is it?

*The European Weather Cloud (EWC) aims to be **the cloud-based collaboration platform** for **meteorological application development** and **operations** in Europe and enables the digital transformation of the European Meteorological Infrastructure (EMI)*
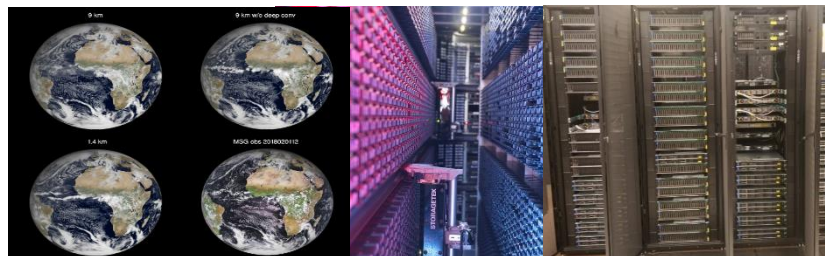
Who maintains it?

- *European Centre for Medium Range Weather Forecast (**ECMWF**)*

- *European Organization for the Exploitation of Meteorological Satellites (**EUMETSAT**)*

What is it for?

*Handling model and observation data on a cloud infra close to the storage, while giving maximum freedom to the users:*
- *Member & Cooperating States*
- *R&D Projects*
- *Partners (e.g. EUMETNET)*
- *Internal Use*

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING–BASED INFRASTRUCTURE, FOCUSED
ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   **EUMETSAT**

European Weather Cloud (EWC): challenges

① New landscape in data ecosystem

② New technical capabilities offered by the cloud
- **Scaling horizontally** as needed
- Abstracting **away the computer's OS** / installation
- Request (or **data**) **driven**
- **Geo-redundancy**

③ Cloud native code needs changes of approach for greatest benefit
- Containers, object storage, stateless code, triggering / notifications, ..
- Most of EO user community isn't far along this path
- Existing code and systems are major investments – **legacy matters**!
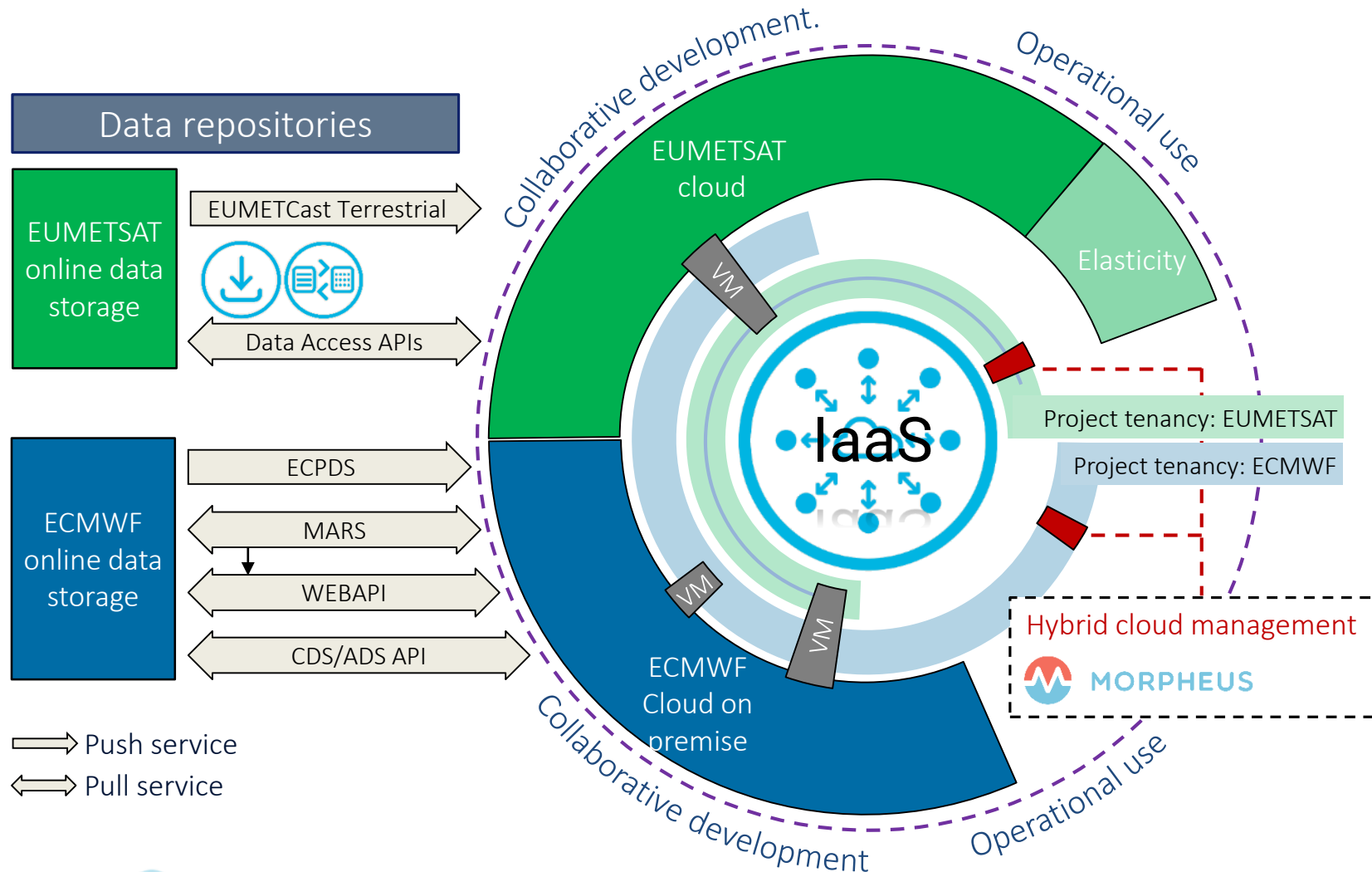
The change is required but is a huge task for individual organizations → **Community is required!**

The Enterprise Journey to Cloud

Startups Journey to Cloud
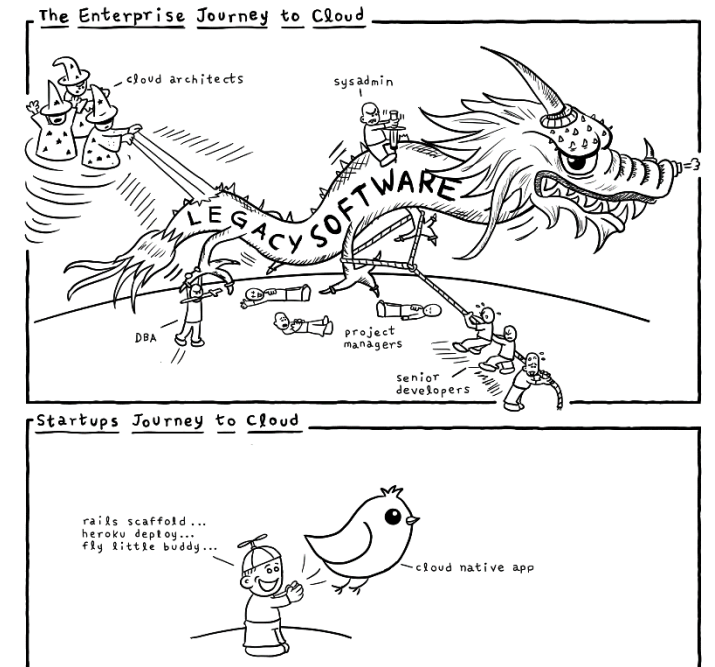
Daniel Stori {turnoff.us}
Thanks to Michael Tharrington
Image taken from: http://turnoff.us/geek/enterprise-vs-startup-journey-to-cloud/ (License CC4BY)

EWC is an efficient way to focus on the key competences of NMHSs

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING–BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   EUMETSAT

European Weather Cloud (EWC): Based on OpenStack and managed by Morpheus

Users can provision VMs and host their own applications

- Key advantages are data proximity and community
- Data accessed from pull data services (data APIs)
- Some library items (e.g. Kubernetes templates) for users' convenience
- Managed services
- Not tight to any cloud
- Federation is possible

Data repositories

EUMETSAT online data storage
- EUMETCast Terrestrial
- Data Access APIs

ECMWF online data storage
- ECPDS
- MARS
- WEBAPI
- CDS/ADS API

Push service
Pull service

Collaborative development.
Operational use
EUMETSAT cloud
Elasticity
Collaborative development
ECMWF Cloud on premise
Operational use

IaaS
VM

Project tenancy: EUMETSAT
Project tenancy: ECMWF

Hybrid cloud management
MORPHEUS

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   EUMETSAT

Batch Processing on EWC

Batch Processing on EWC: use cases

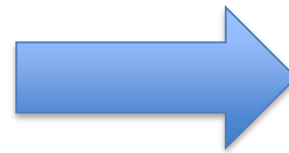| Use case	Example
Analysis job	Create one job from each file in a list. Data stored at the end
Processing job	Using different inputs with timestamps and generate new output
MPI job	Cloud mask generation breaking up the Earth disk into tiles
ICDR job	Run a satellite processing of some kind, one every day as time goes by

Batch Processing on EWC: use cases

Use case	Example
Analysis job	Create one job from each file in a list. Data stored at the end
Processing job	Using different inputs with timestamps and generate new output
MPI job	Cloud mask generation breaking up the Earth disk into tiles
ICDR job	Run a satellite processing of some kind, one every day as time goes by

Use cases commonly require
- Analysis on large, long-timescale datasets (e.g. 50 years)
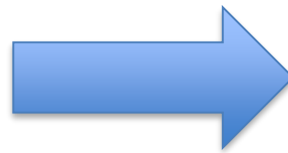- Millions of independent, relatively small jobs

Batch Processing on EWC: use cases

Use case	Example
Analysis job	Create one job from each file in a list. Data stored at the end
Processing job	Using different inputs with timestamps and generate new output
MPI job	Cloud mask generation breaking up the Earth disk into tiles
ICDR job	Run a satellite processing of some kind, one every day as time goes by

Use cases commonly require
- Analysis on large, long-timescale datasets (e.g. 50 years)
- Millions of independent, relatively small jobs
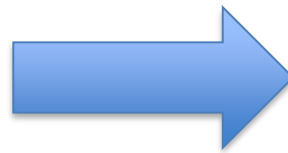
High Throughput Computing

Batch Processing on EWC: use cases

Use case	Example
Analysis job	Create one job from each file in a list. Data stored at the end
Processing job	Using different inputs with timestamps and generate new output
MPI job	Cloud mask generation breaking up the Earth disk into tiles
ICDR job	Run a satellite processing of some kind, one every day as time goes by

Use cases commonly require
- Analysis on large, long-timescale datasets (e.g. 50 years)
- Millions of independent, relatively small jobs

High Throughput Computing

HTCondor already used internally at EUMETSAT for full mission dataset reprocessing activities (Climate Team)

Batch Processing on EWC: use cases

Use case	Example
Analysis job	Create one job from each file in a list. Data stored at the end
Processing job	Using different inputs with timestamps and generate new output
MPI job	Cloud mask generation breaking up the Earth disk into tiles
ICDR job	Run a satellite processing of some kind, one every day as time goes by

Use cases commonly require
- Analysis on large, long-timescale datasets (e.g. 50 years)
- Millions of independent, relatively small jobs

High Throughput Computing

HTCondor already used internally at EUMETSAT for full mission dataset reprocessing activities (Climate Team)

What solution for EWC users?

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   EUMETSAT

Batch Processing on EWC: initial state

- Each tenancy have different allocated resources

- No shared network possible from the cloud provider

- Scheduling on the internet with public IPs not possible
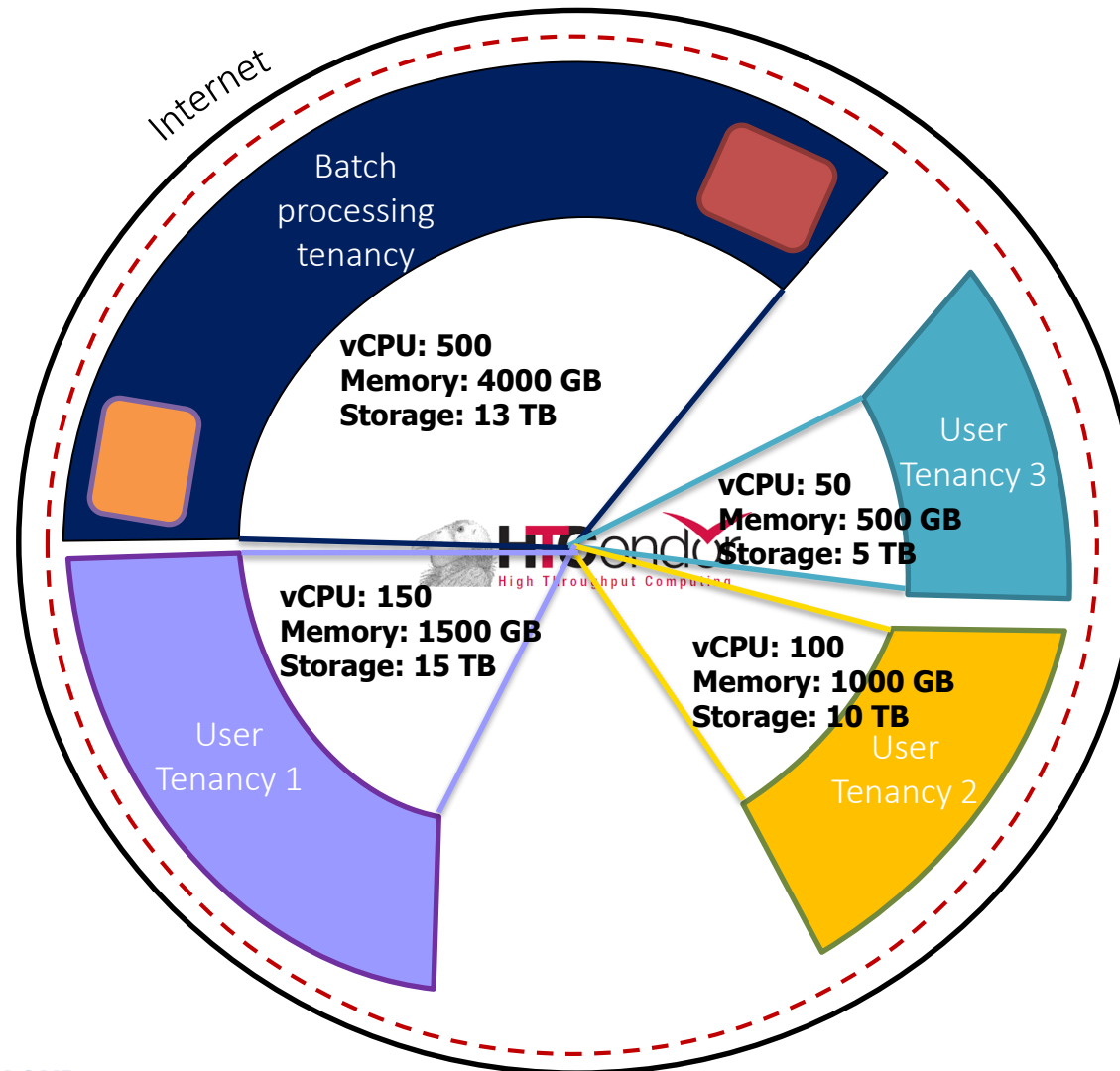
- Not all tenancies have resources to maintain a service
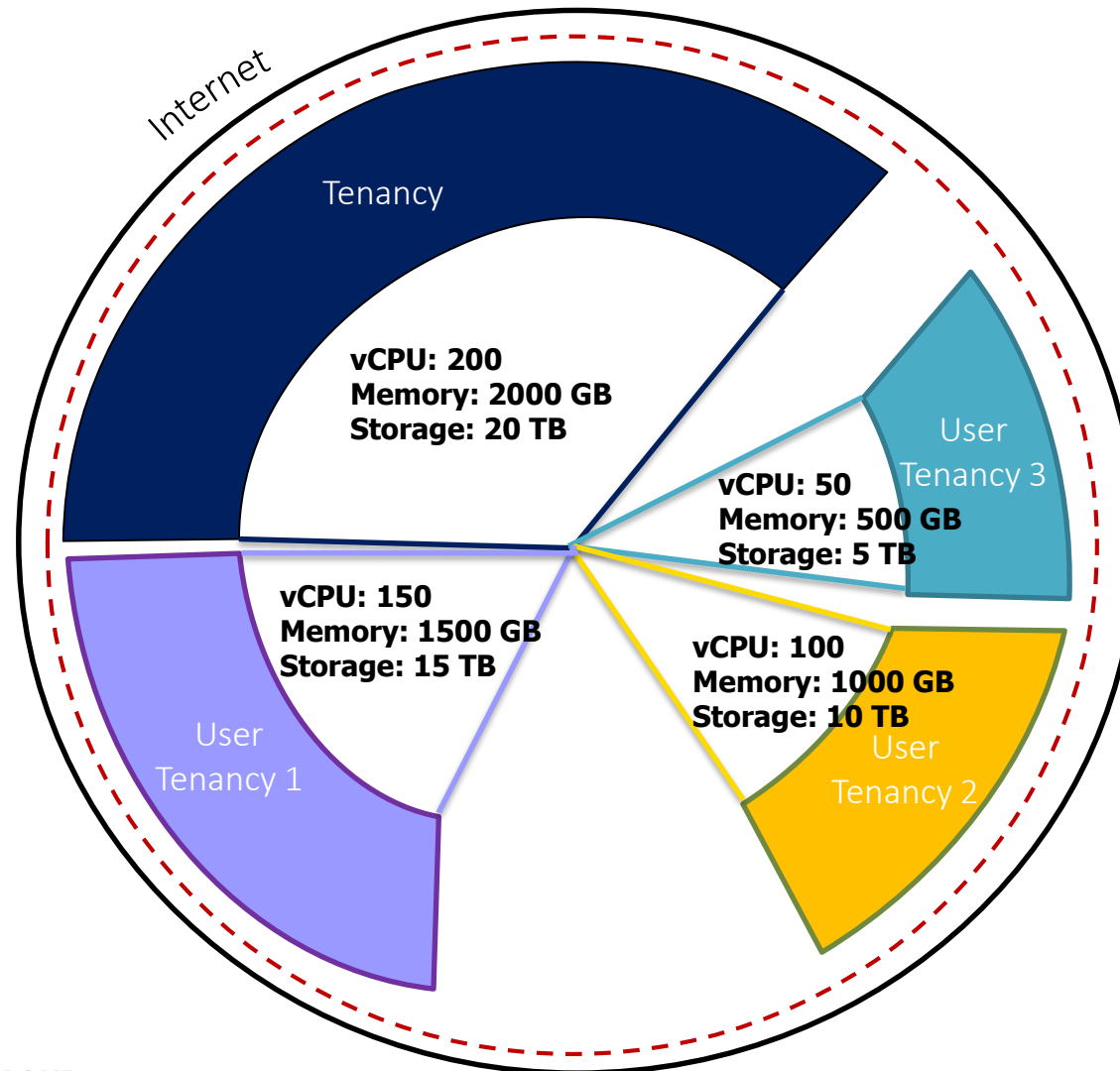
Tenancy definition:
Isolated environment with unique users, resources, private networks and workloads, with no access or visibility to other tenants in the same infrastructure.

Batch Processing on EWC: challenges

Challenge	
Maintenance	Not all users have resources/skills to maintain a new service
Deployment	Run workloads across different tenancies/clouds
Resource	Not enough resources to run a realistic workload in reasonable time for single tenancy
Usage	There are idle resources from tenancies that can be given to the community for a certain time
Network	Connect machines in different isolated environments
Scheduling	Users from different tenancies require access to scheduling

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED
ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   EUMETSAT

Batch Processing on EWC: centrally managed tenancy

HTCondor manager

HTCondor submit node

HTCondor execute node

Internet

Batch processing tenancy

vCPU: 500
Memory: 4000 GB
Storage: 13 TB

User Tenancy 3

vCPU: 50
Memory: 500 GB
Storage: 5 TB

vCPU: 150
Memory: 1500 GB
Storage: 15 TB

vCPU: 100
Memory: 1000 GB
Storage: 10 TB

User Tenancy 2

User Tenancy 1

Advantages
- No maintenance from the users
- Maintainers update the system, users can simply redeploy a node

Batch Processing on EWC: VPN
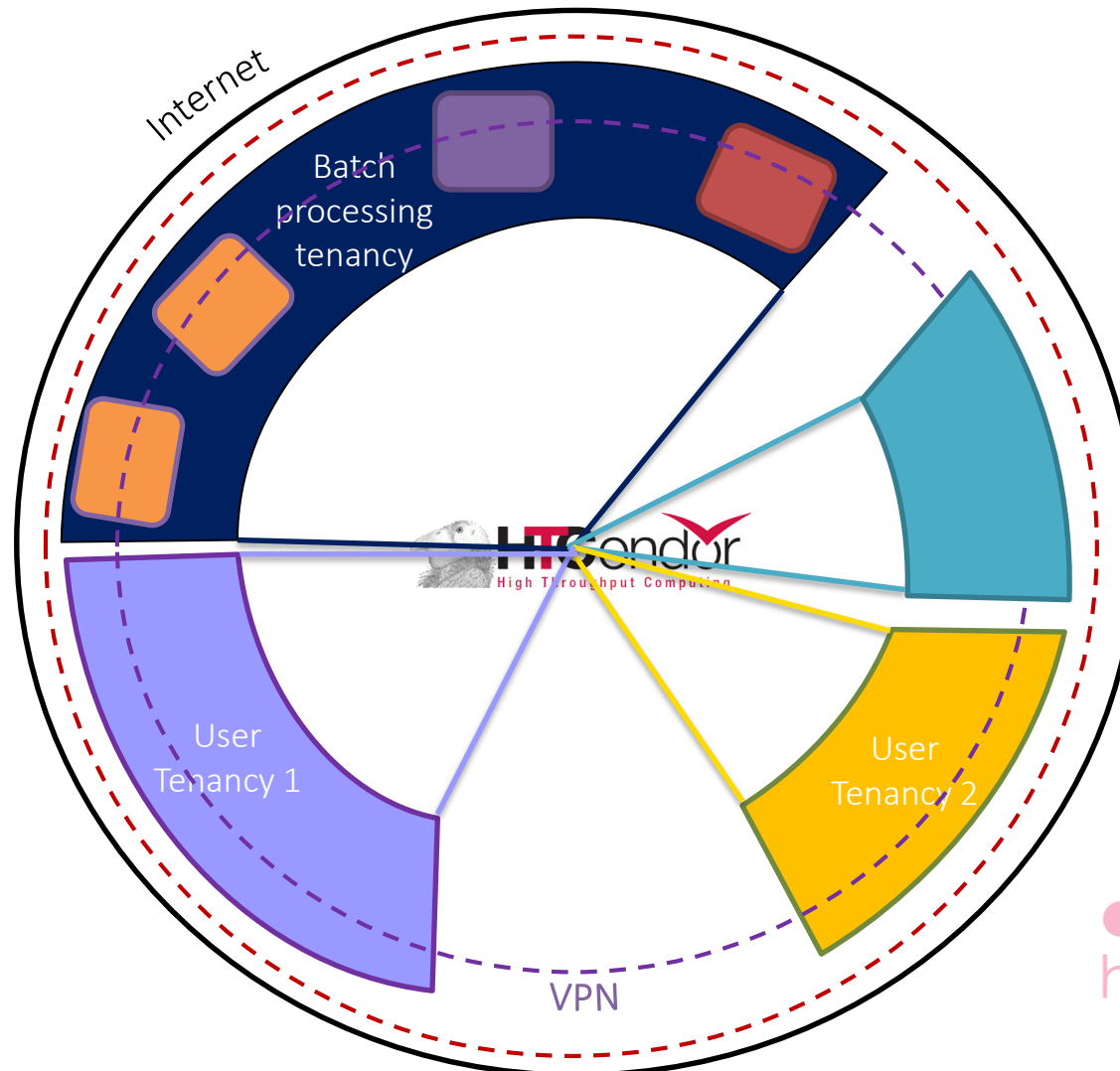
VPN:

- Headscale VPN
- Tailscale clients on the nodes
- RBAC/ACL policies (e.g. Block port below 1000, Restricted user access)
- Automate offline nodes removal
- Automate backup

Batch Processing on EWC: multitenancy

HTCondor manager

HTCondor submit node

HTCondor execute node

VPN

Internet

Batch processing tenancy

User Tenancy 3

User Tenancy 1

User Tenancy 2

VPN
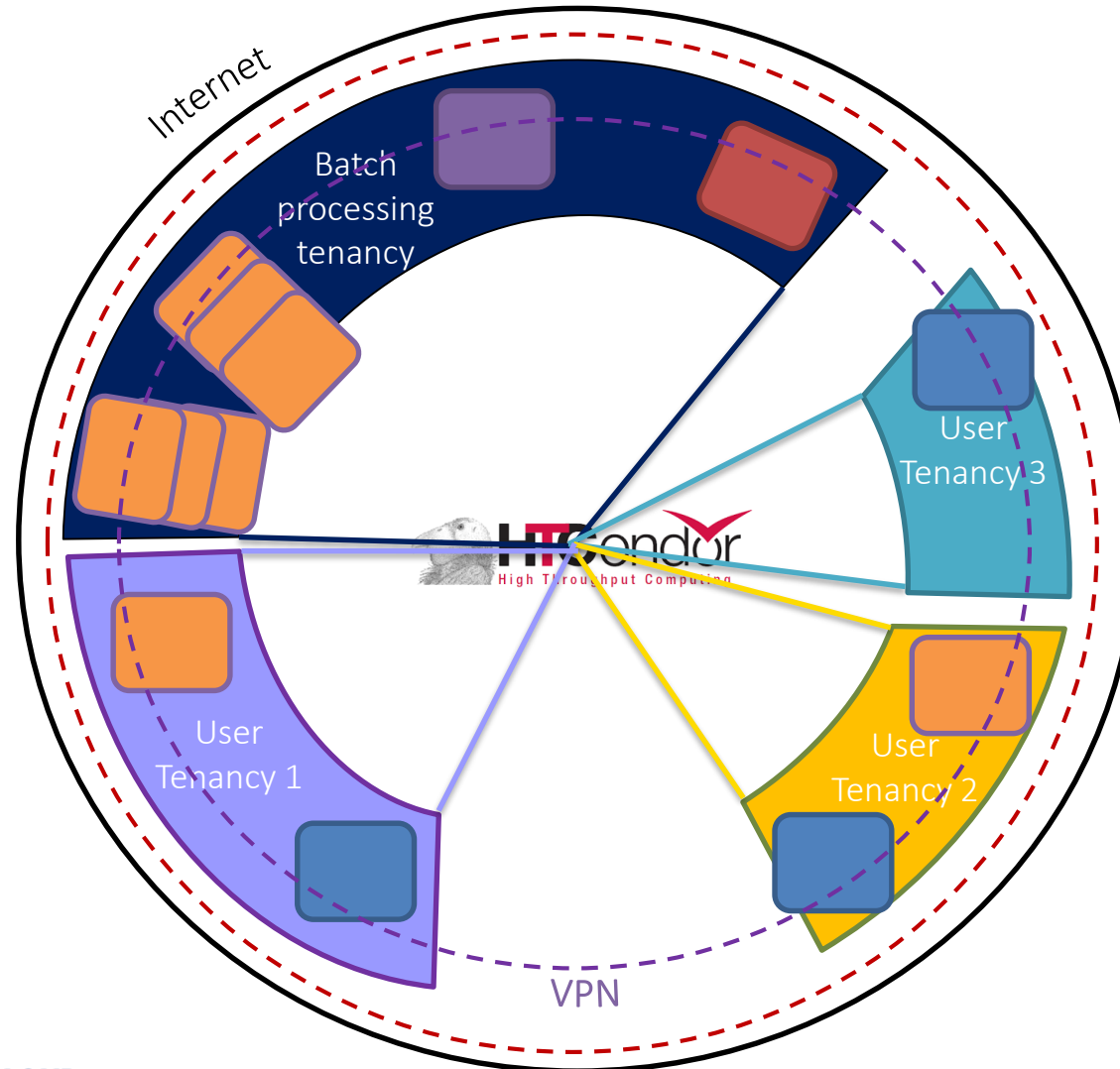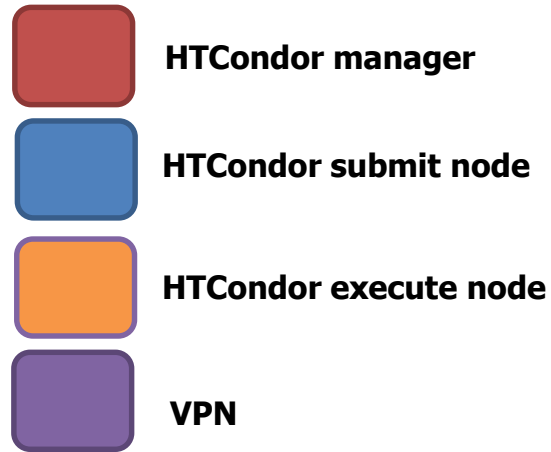
Advantages:

- Use Idle machines

- Small tenancies can benefit from external large amount of resources

- User can submit job from their tenancy with their submit node

- Easy to deploy/maintain/scale in each tenancy

- Retain individual privacy
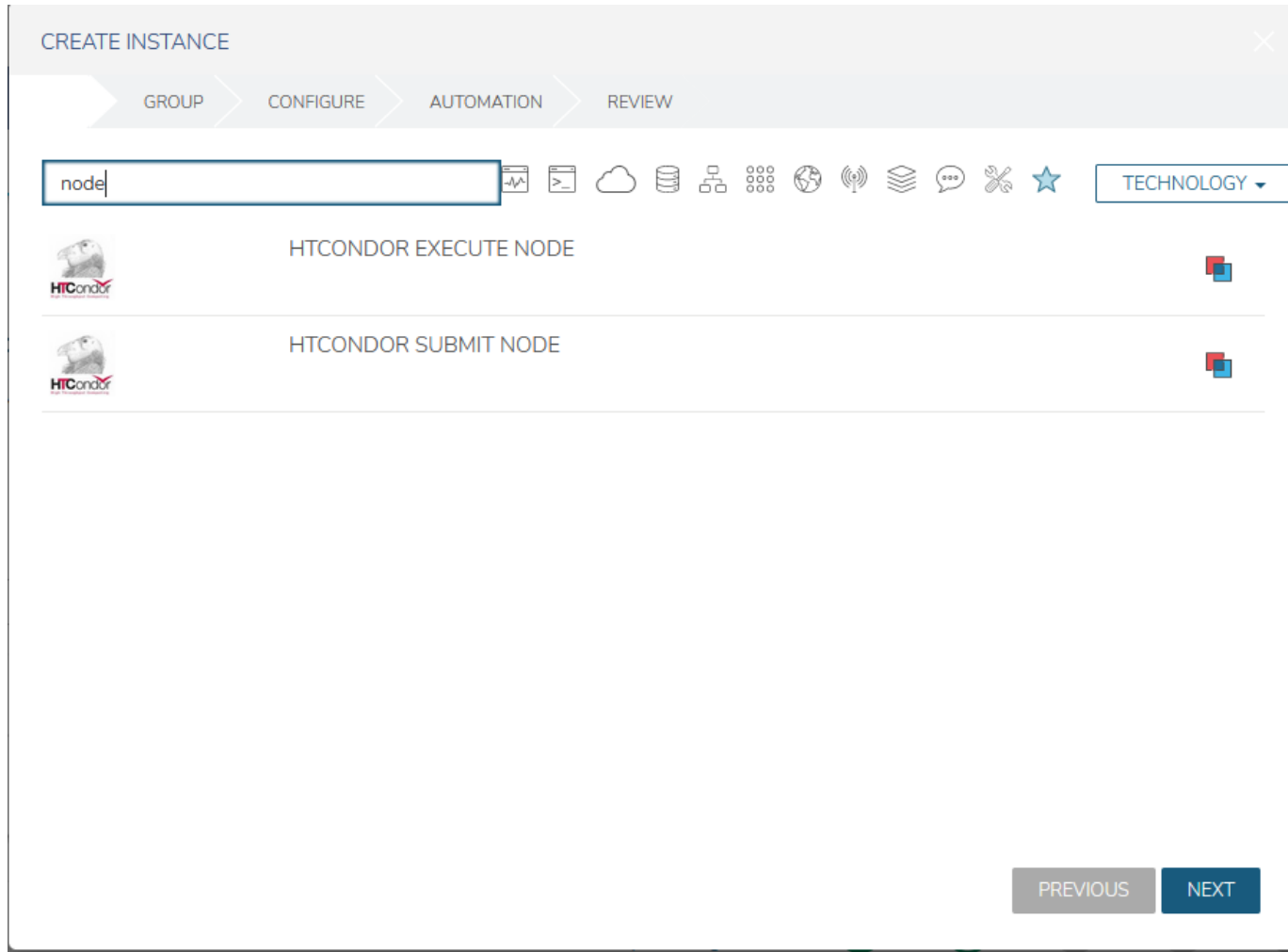
Batch Processing on EWC: monitoring

Legend:
- HTCondor manager
- HTCondor submit node
- HTCondor execute node
- VPN
- Monitoring
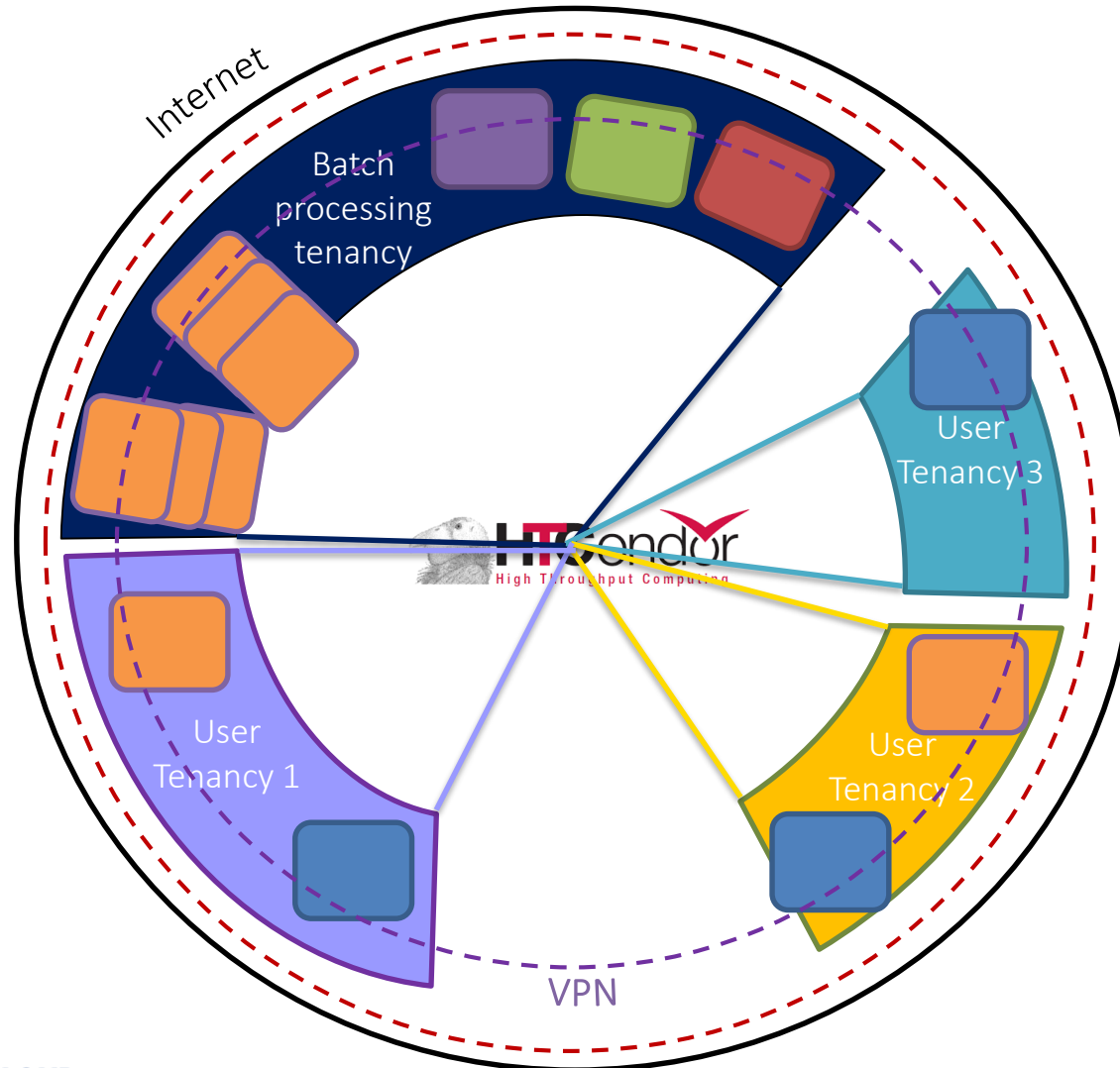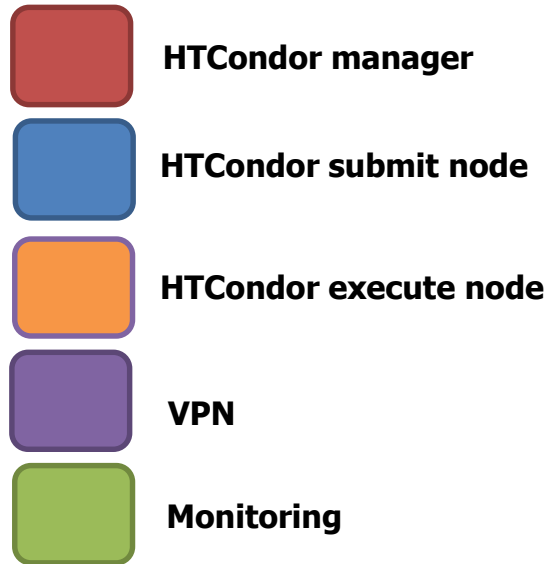
Components:
- Metrics exporter (Python)
- PostgreSQL database
- Grafana dashboard

Batch Processing on EWC: join EWC HTCondor pool

- Few click from the Morpheus Interface to have a new node joining existing pool

- No need for any configuration except selection of the plan for the node (e.g. 32 CPU, 400 RAM, 5 TB)

Batch Processing on EWC: current solution

Challenge	Solution
Maintenance	Centrally Managed Tenancy, easy 'one click' deployment
Deployment	Multi tenancy
Resource	Join automatically the main HTCondor Pool
Usage	Easy 'one click' deployment
Network	VPN
Scheduling	Single schedulers in each tenancy

Submit node:
- Only docker universe allowed
- Only condor_submit command allowed
- Private network in the tenancy enabled to allow access to tenancy-internal resources/files
- Condor transfer mechanism allowed

Execute node:
- No access to execute host for containers
- No access to other containers running on execute node
- Isolated environment for containers
- No autoscaling
- No NFS

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING–BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF   EUMETSAT

Batch Processing on EWC: next steps

- Pilot tests run with users' workload (we are at alpha -> early beta stage)

- User examples (e.g. use of S3 bucket, container)

- Python wrapper for job submission?

- Learn more (at this workshop!) about similar use cases and/or other approaches with pros and cons

Thank you for your attention!