European Weather Cloud under the hood, Improving cloud offering from IaaS to PaaS

Virtual workshop: Weather and climate in the cloud
February 8-10, 2021

Vasileios A. Baousis (PhD)

ECMWF

Agenda

- European Weather Cloud
  - Building a Community Cloud with OpenStack
  - ECMW's production workflow & European Weather Cloud
- From IaaS to PaaS model
  - Design considerations
  - Building blocks
  - Tailor made virtual resources and alternatives (VM, blueprints and containerisation)
  - Kubernetes (RKE, OKD etc) Templates and Operators
- On going and future activities

ECMWF

European Weather Cloud

- European Weather Cloud is a Community Cloud- EMI (E&E and Member States)

- Three year pilot project started in January 2019 in collaboration between ECMWF and EUMETSAT with operational phase scheduled to start Q1/2022

    "Basic goal is to bring the computation resources (Cloud) closer to our Big data (meteorological archive and satellite data)"

- ECMWF's Pilot infrastructure was built with open source software, OpenStack and Ceph

- Main page www.europeanweather.cloud

ECMWF

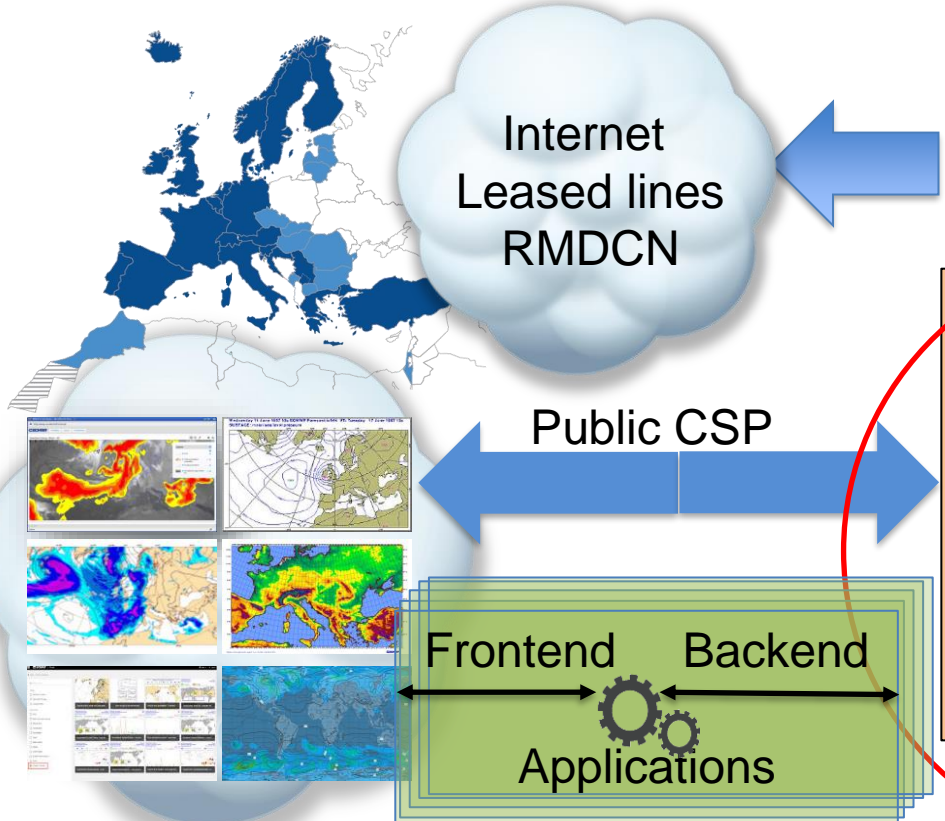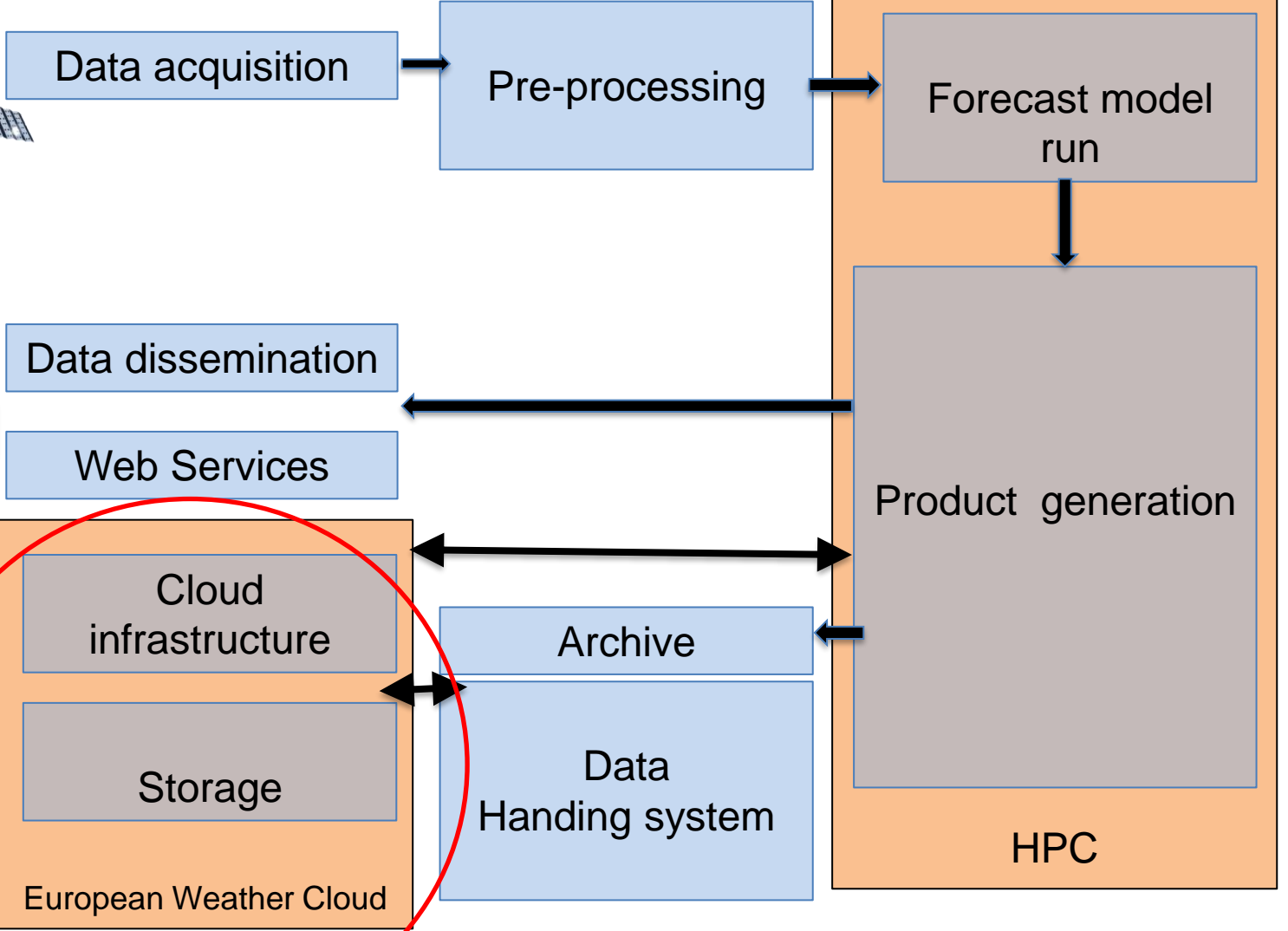# European Weather Cloud pilot infrastructure @ ECMWF
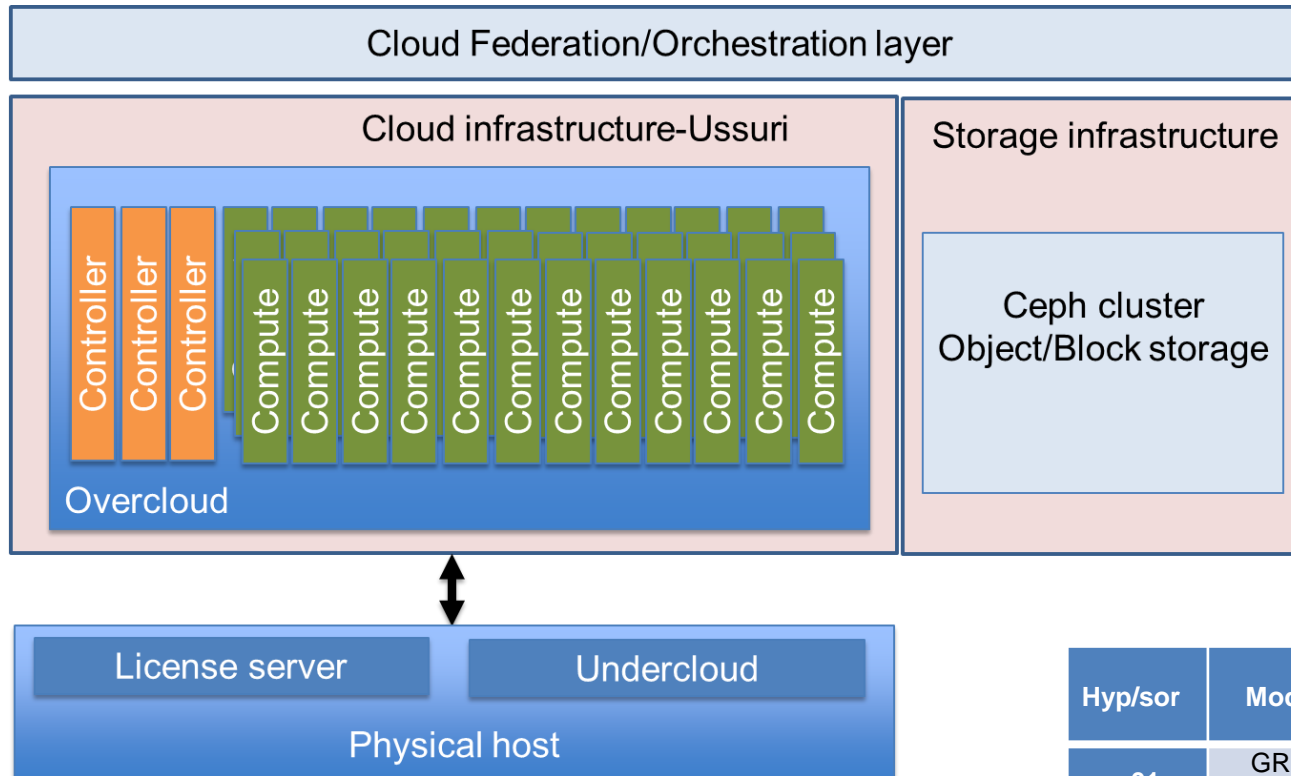
H/W

- Systems : Cloud:43 /Ceph: 23
- Cores   : ~3000
- RAM   : ~21TB
- Storage: ~1PB (HDD+SSD)
- GPUs   : 2x5 NVIDIA Tesla V100

Hyp/sor	Model	Nova profile	VM/GPU /Model	VMs /host	Frame BufferSize	Max. Display Resolution
gpu01	GRID V100-16C	nvidia-301	1	2	16384	4096×21602
gpu02	GRID V100-8C	nvidia-300	2	4	8192	4096×21602
gpu03	GRID V100-8C	nvidia-300	2	4	8192	4096×21602
gpu04	GRID V100-4C	nvidia-299	4	8	4096	4096×21602
gpu05	GRID V100-4C	nvidia-299	4	8	4096	4096×21602

ECMWF's production workflow and European Weather Cloud

800 million observations/day

Data acquisition → Pre-processing → Forecast model run

Data dissemination

Web Services

Internet Leased lines RMDCN

Public CSP

Cloud infrastructure

Storage

European Weather Cloud

Frontend   Backend

Applications

Archive

Data Handing system

Product generation

HPC

ECMWF component of the European Weather Cloud Overall architecture

https://storage.ecmwf.europeanweather.cloud

Public network (136.156.90.*/23)

EUMETSAT

VM

Load Balancer (F5)

OpenStack

VM

136.156.*
192.168.*
10.x.*

Private (e.g. 192.168.*.*/x)

192.168.*

VM

Data network (10.158.20.*)

Rados-GW (S3) Rados-RBD

librados

CEPH

APIs | DHS (MARS) (ecfs) | ECPDS | HPC (FDB) (IFS) (pgen) | Notifications

Internal systems

Local area network
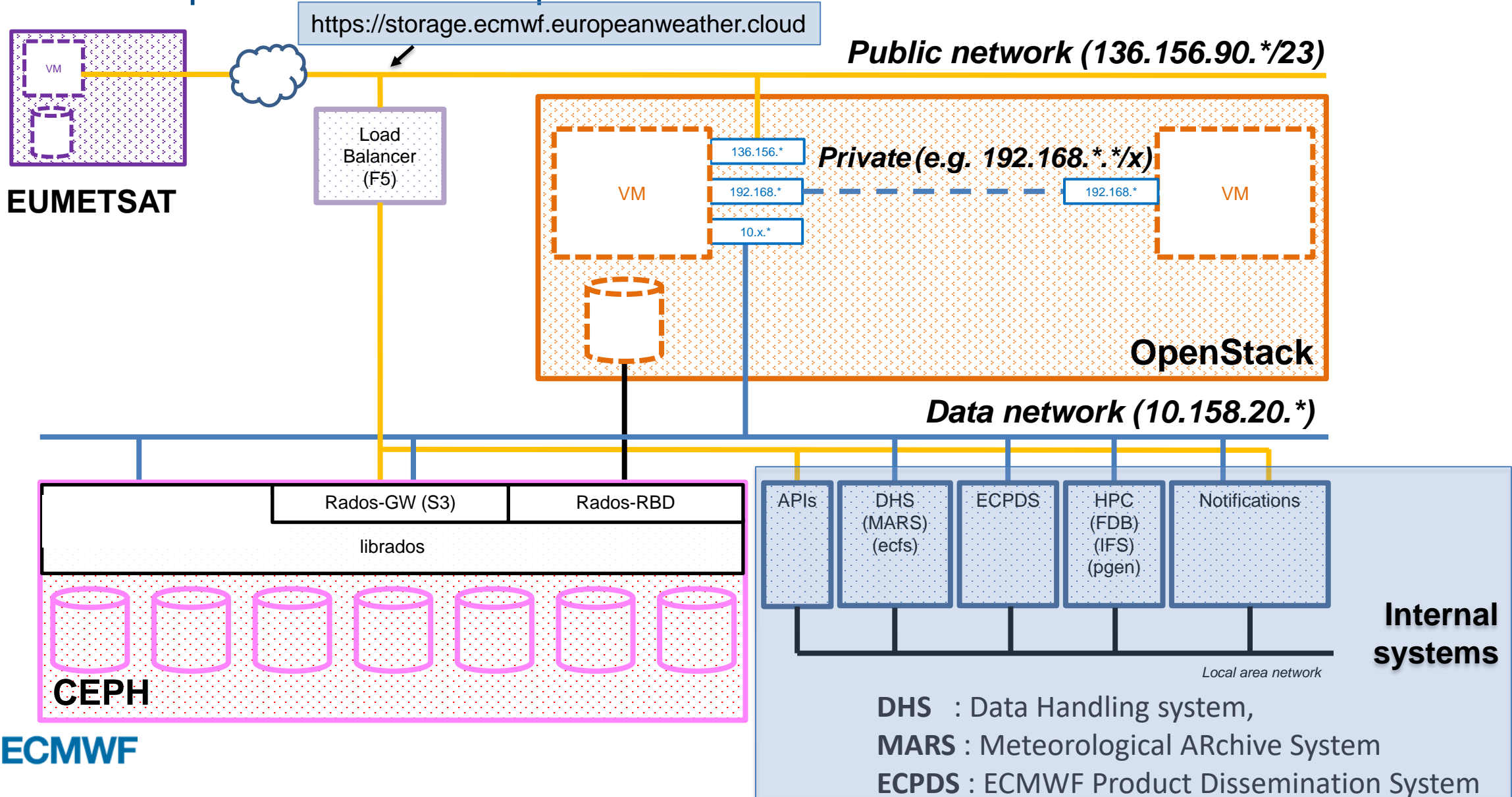
DHS : Data Handling system,
MARS : Meteorological ARchive System
ECPDS : ECMWF Product Dissemination System

European Weather Cloud: From IaaS to PaaS model

Platform as a Service (PaaS)

The capability provided to the consumer to deploy onto the cloud infrastructure **consumer-created or acquired applications** created using programming languages, libraries, services, and tools supported by the provider.

The consumer **does not manage or control the underlying cloud infrastructure** including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment
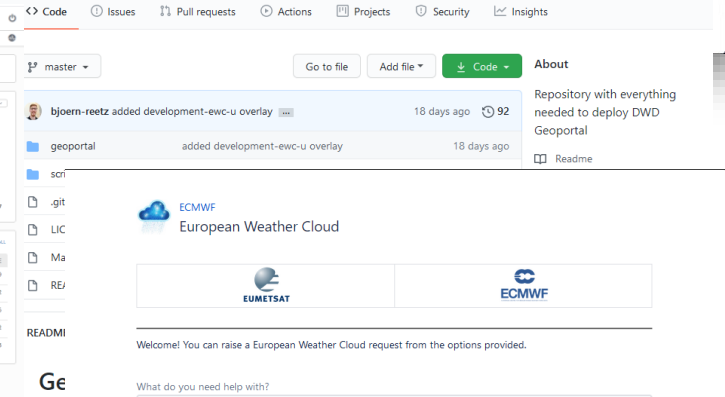
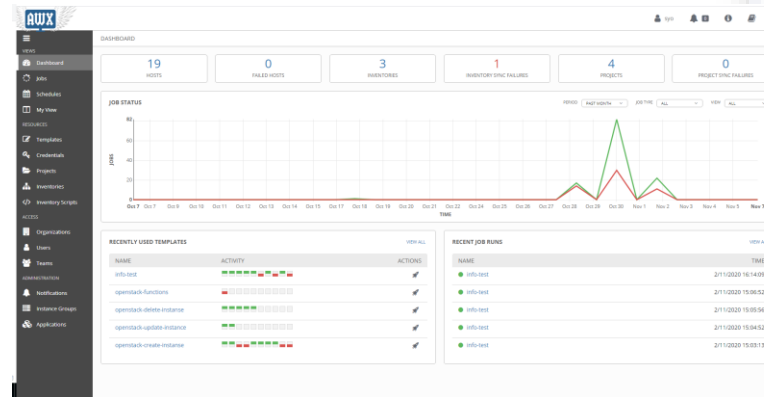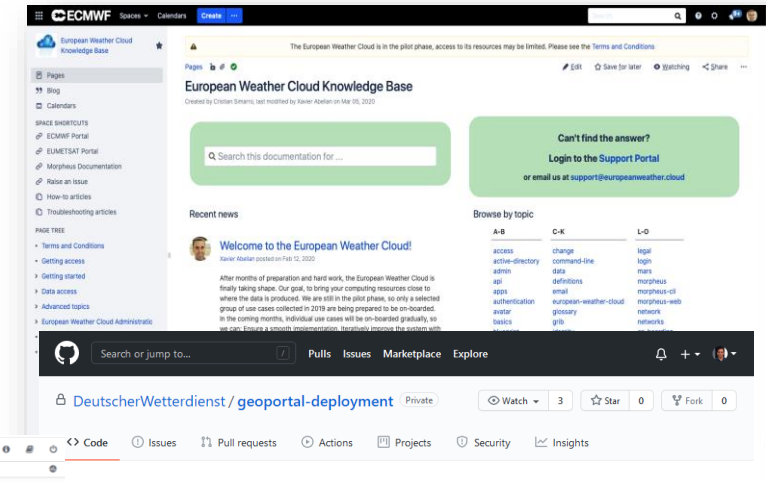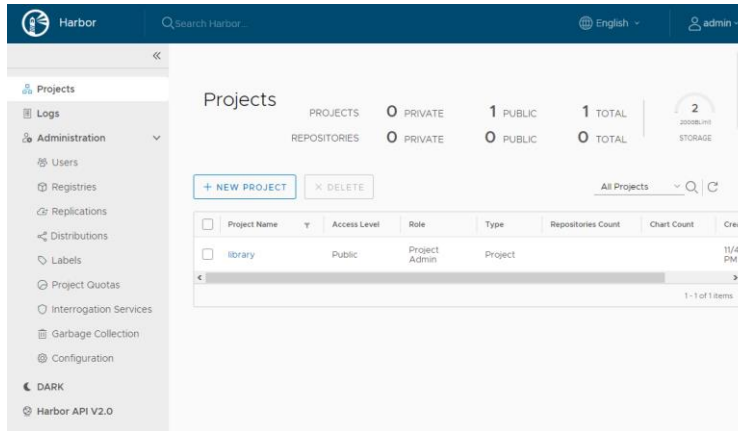[1] NIST 800-145 The NIST Definition of Cloud Computing

European Weather Cloud: From IaaS to PaaS model

Design considerations

- User base diversity

- Mainly meteorological applications

  – but necessary general purpose platforms and applications

- Multi-tier applications

- Resource demanding workloads

  – Support ML/AI

- Fast access to data

- Containerization and orchestration

- Quick/easy creation/disposal of platforms (e.g. Jupyterlab supported with GPUs)

- User and application security, privacy and isolation (certificate creation etc)

- User accounting and reporting

- Data privacy

ECMWF   EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Building blocks

- git*(lab, hub etc) repositories
- Private Container registries
- Common Configuration management tools like Ansible (awx), Puppet etc
- Support and guidance
  - European Weather Cloud Knowledge base
- Public & private (git, docker registries etc)

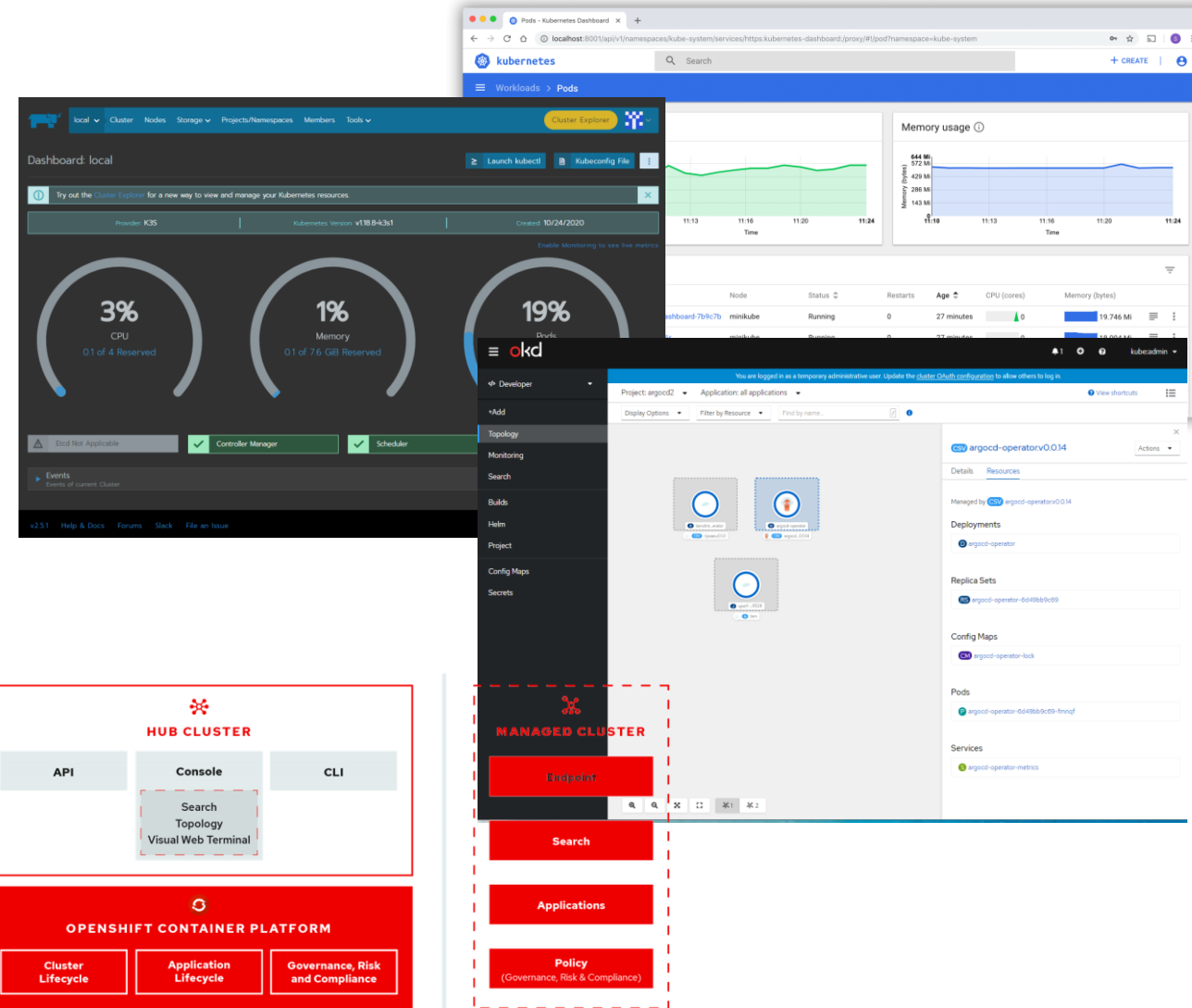|	
Container registry	https://registry.europeanweather.cloud
CM tools/Ansible tower	https://config.europeanweather.cloud
Gitlab repository	https://repository.europeanweather.cloud
Knowledge base	https://confluence.ecmwf.int/display/EWCLOUDKB
Support portal	https://jira.ecmwf.int/servicedesk/customer/portal/9

Tailor-made Virtual resources and alternatives

- Prebuild VMs with meteorological applications/tools

  – Not difficult to produce (CI/CD Pipelines can deliver images)

  – Appropriate solution for a small number of applications (Not scalable for n app and z versions =>n*z images.

  – Cumbersome to maintain

- Cloud orchestrator blueprints & workflows

  – Easy to deploy and maintain, Flexible to maintain n applications & z versions

  – Increased provisioning time of the VM, depending on the application complexity and size

- Containerisation

  – Provides solutions to many requirements

  – Different dimension of abstraction/flexibility/additional complexity from the admin point of view?

  – In/de-creased user isolation/security?

   - Proper RBAC implementations ? Admin overhead?

➔A combination of these

Kubernetes, Rancher OpenShift-OKD

- Security and dynamic SSL certificates

  – https://github.com/tnozicka/openshift-acme

- Higher storage abstraction (in addition to the natural integration with Ceph/Cinder)

  – https://www.noobaa.io

- Machine learning (ML) workflows creation

  – https://github.com/kubeflow/

- Kubernetes federation

  – https://github.com/open-cluster-management

ECMWF

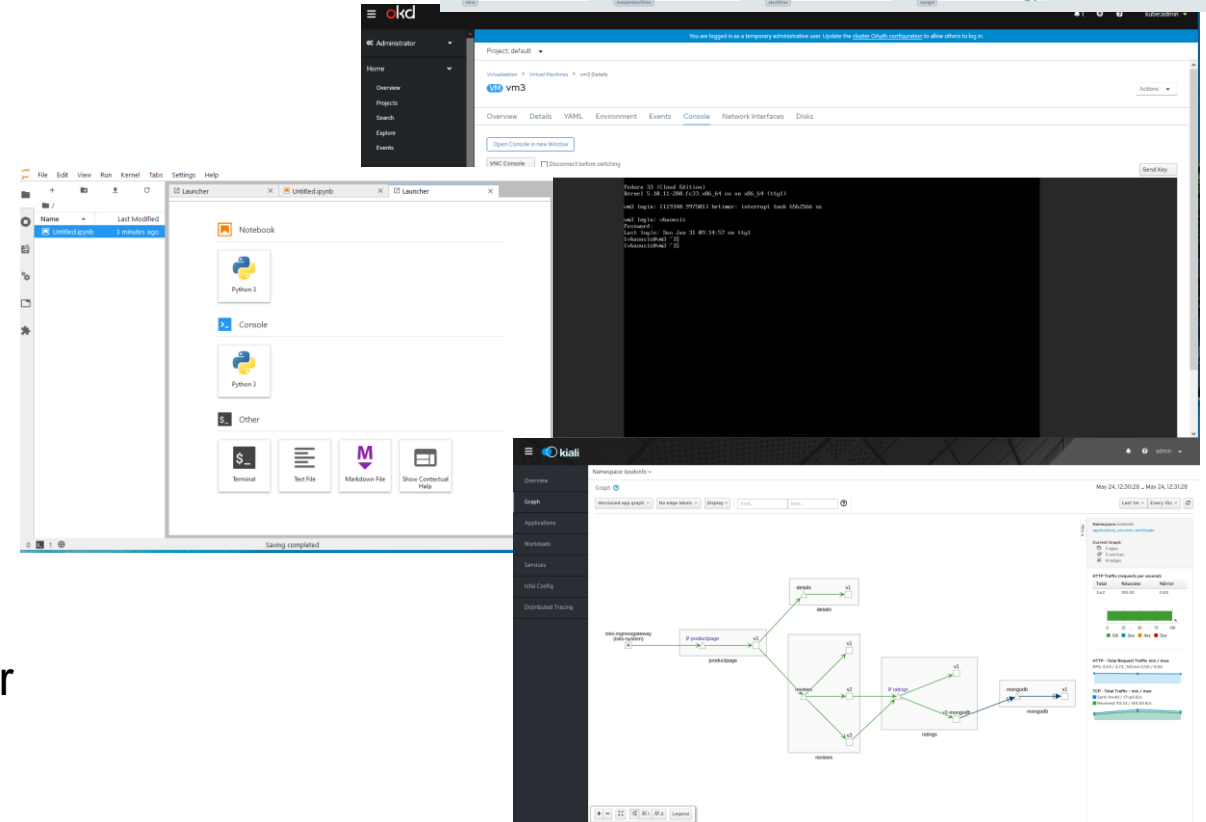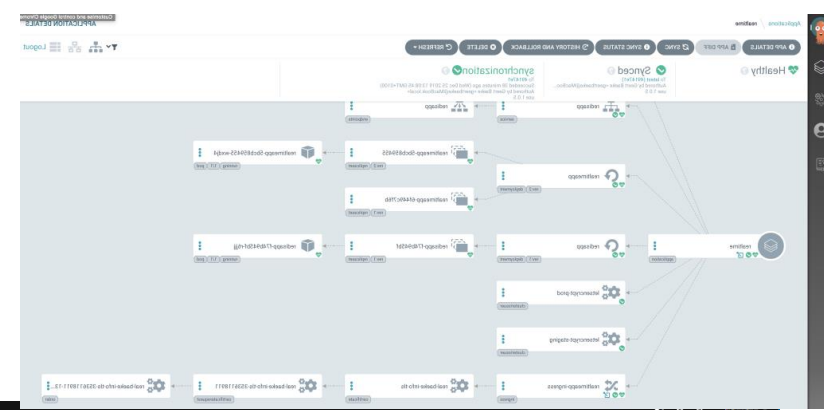**EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS**

Kubernetes operators

- *Operators* take human operational knowledge and encode it into software that is more easily shared with consumers

- A method of packaging, deploying, and managing a Kubernetes application.

- Repeatable, health checks, easily updated, encapsulate knowledge

- Can be created using Ansible, Helm

- Operator SDK

- Openshift templates.

  - Instant App and Quickstart templates to quickly get started creating a new application for different languages.
  - Rails (Ruby), Django (Python), Node.js, CakePHP (PHP), and Dancer (Perl).
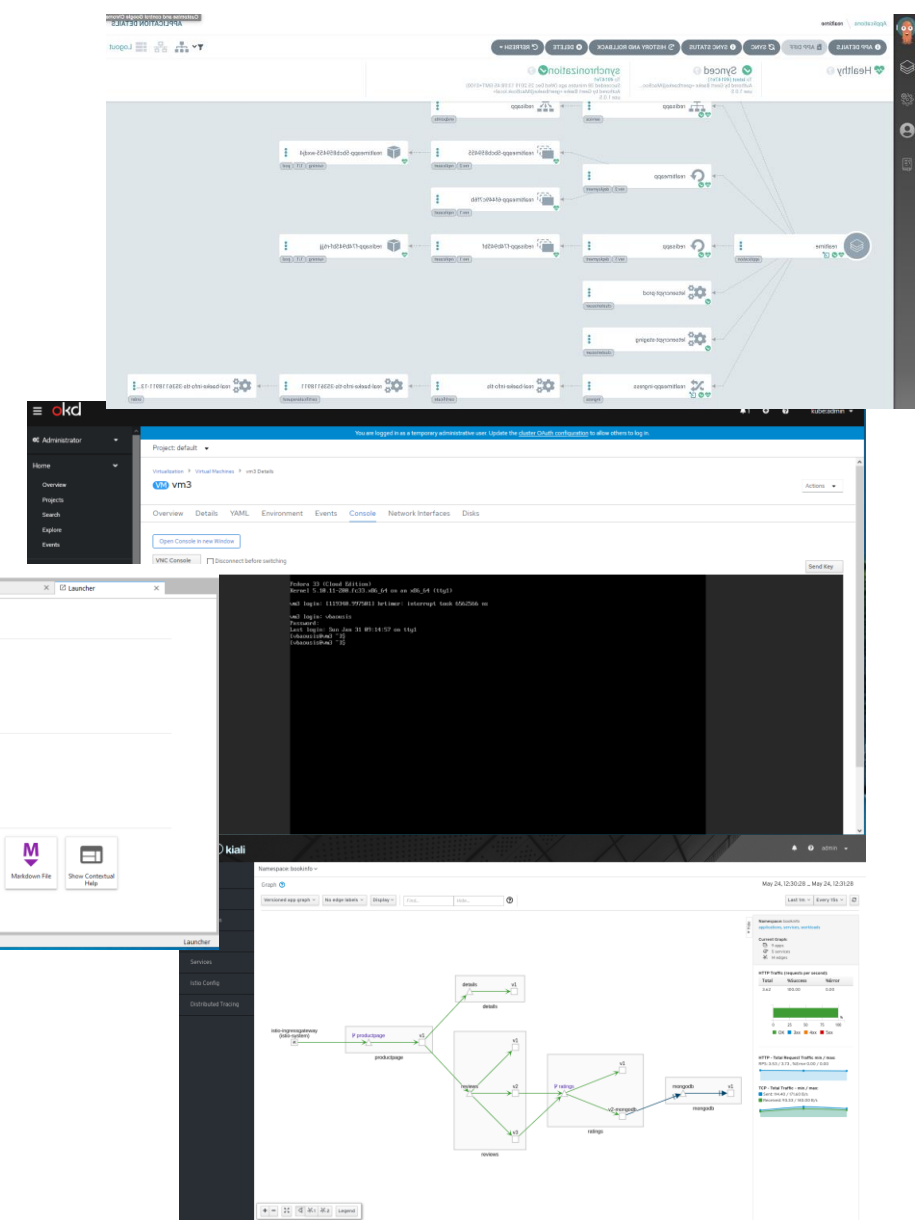
Kubernetes operators

- Opportunistic VMs with Kibevirt within K8 : https://kubevirt.io/

- GitOps continuous delivery tool for Kubernetes.

  – https://argoproj.github.io/argo-cd/

- Serverless with knative: https://knative.dev/

- JupyterHub https://github.com/jupyter

- Microservice connectivity with Istio service mesh

  – https://github.com/kiali/kiali

- SSO with Keycloack

  – https://operatorhub.io/operator/keycloak-operator

- Apache Kafka

  – https://github.com/strimzi/strimzi-kafka-operator

Meteorological applications delivered through K8 Operators?

ECMWF

EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Ongoing and future activities

- IaaS to PaaS offering

- Integration/interface with other projects
  - European Weather Cloud will be interfacing with the Digital Twin Earth/ Part of the Destination Earth Program
  - There are also other European Projects to be considered like European Open Science Cloud (EOSC) etc

- Contributing to the OpenStack, Ceph, okd community
  - Code
  - Troubleshooting problems
  - Presentations

ECMWF

European Weather Cloud under the hood,

Improving cloud offering from IaaS to PaaS

Virtual workshop: Weather and climate in the cloud
February 8-10, 2021

Q/A

Vasileios A. Baousis (PhD)
vasileios.baousis.at.ecmwf.int

ECMWF