Implementation of the Production European Weather Cloud

Xavier Abellan – xavier.abellan@ecmwf.int

Stig Telfer – stig@stackhpc.com

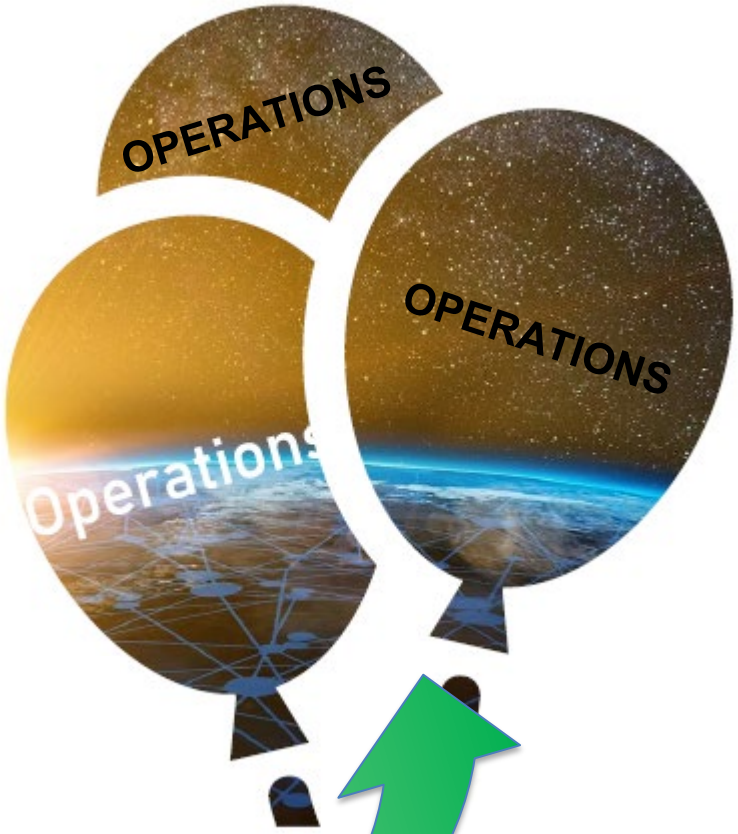
ECMWF StackHPC

What is the European Weather Cloud

"The European Weather Cloud 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.

The European Weather Cloud is dedicated to support the **National Hydro-meteorological Services of the Member States of both ECMWF and EUMETSAT in fulfilling their official duties** to protect life and property from impending meteorological hazards."

"*a community cloud*"



26 September 2023

https://www.europeanweather.cloud

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

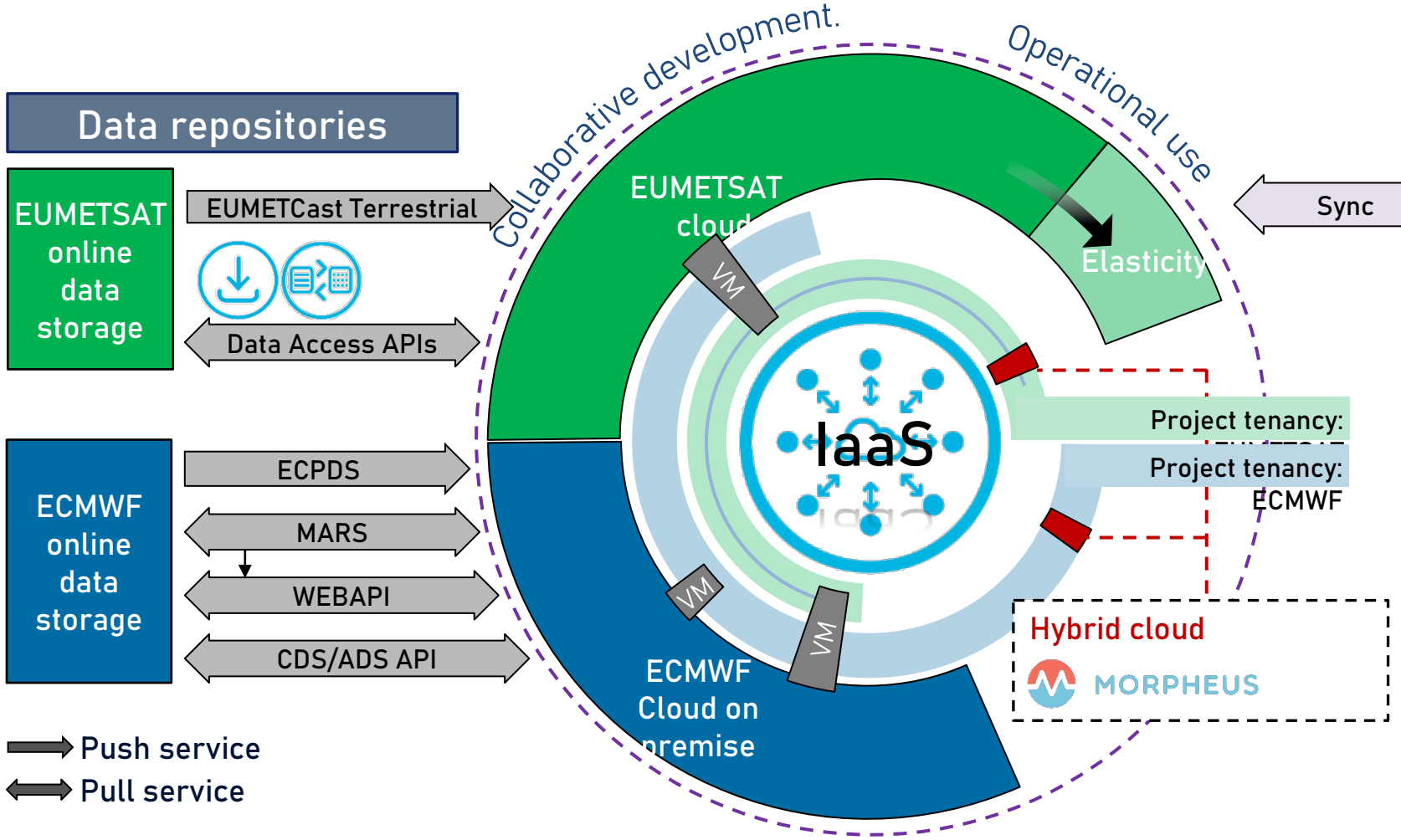
ECMWF EUMETSAT

Why European Weather Cloud

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

ECMWF EUMETSAT

High level architecture



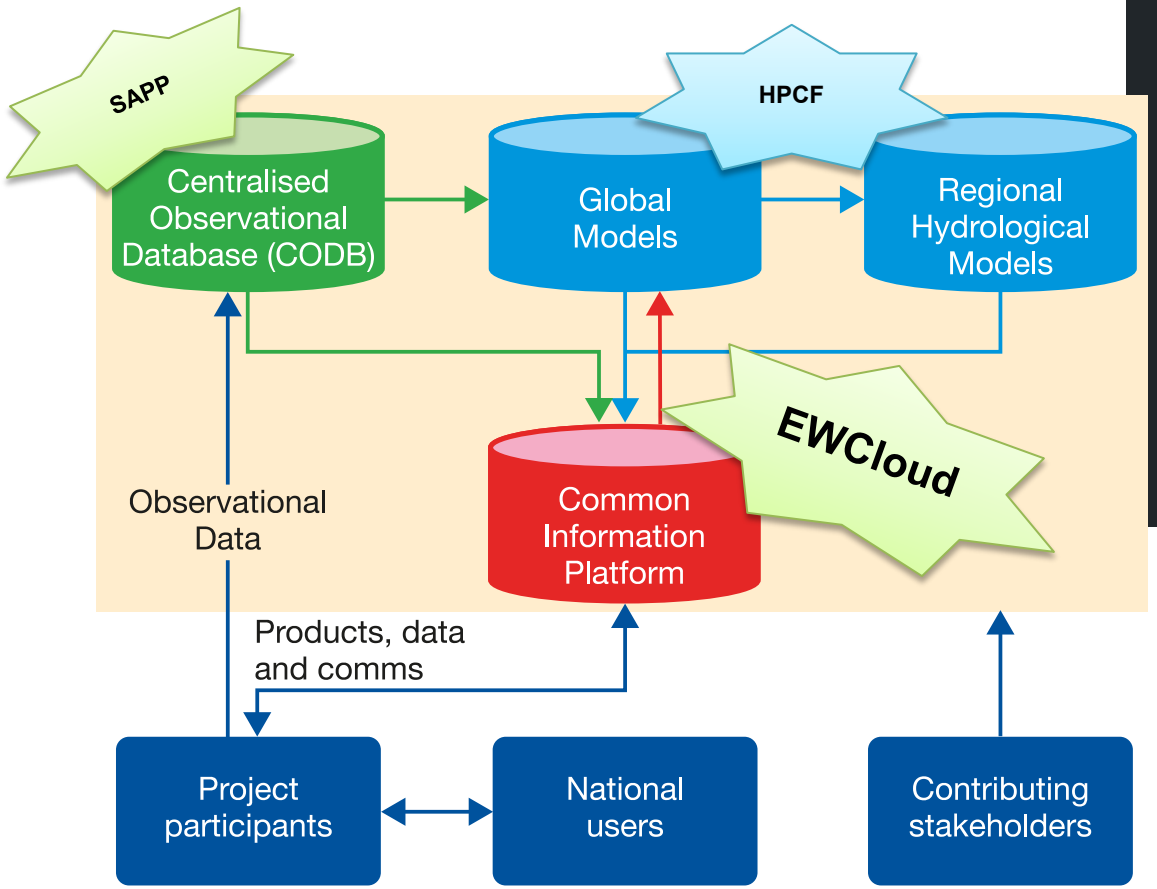
Tenants can provision VMs and host their own applications

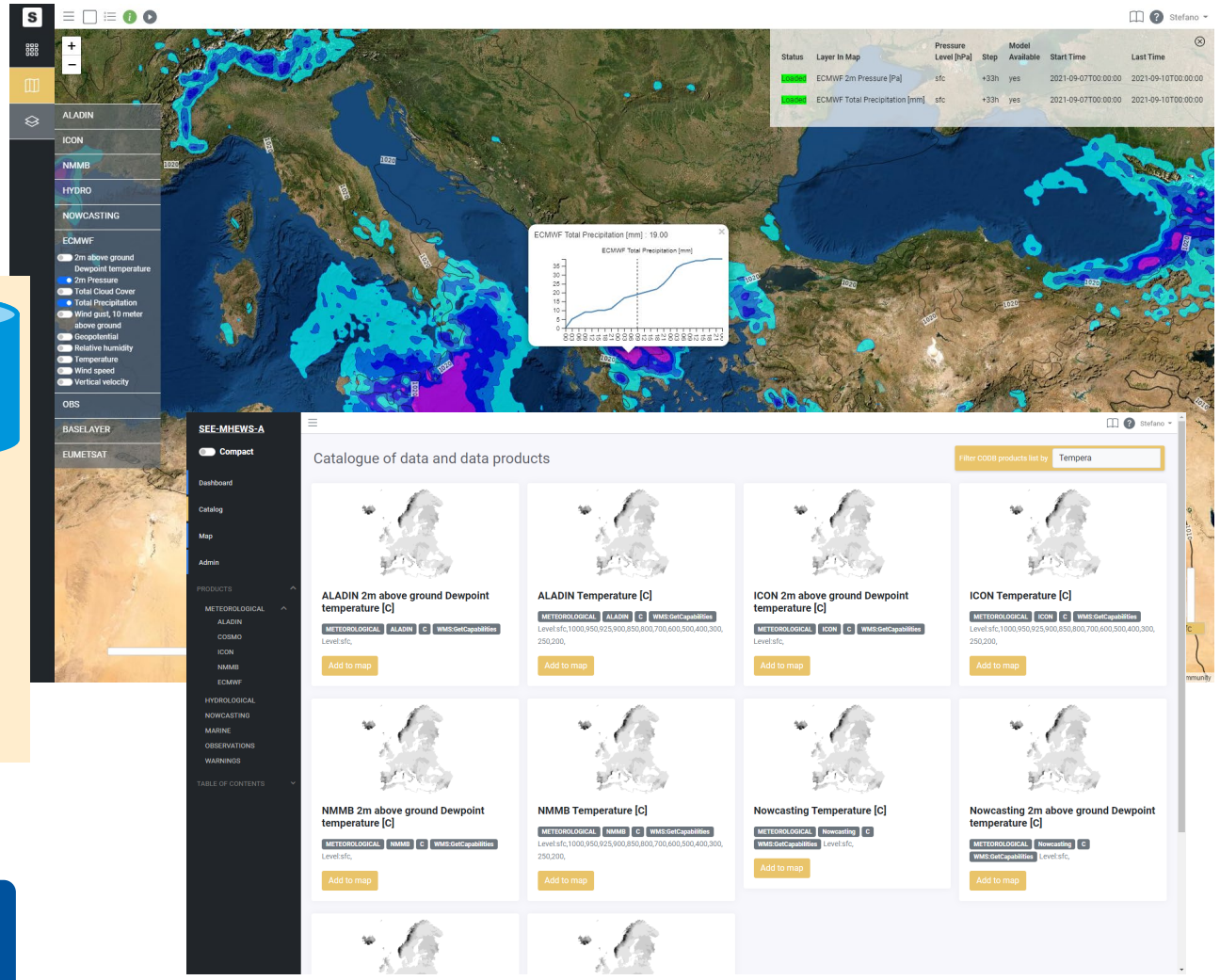
- All deployed services inside the tenancy are maintained by the tenants
- Tenants can sync third-party data and collaborate with data access using e.g. object storage

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

ECMWF EUMETSAT

International collaboration: SEE-MHEWS

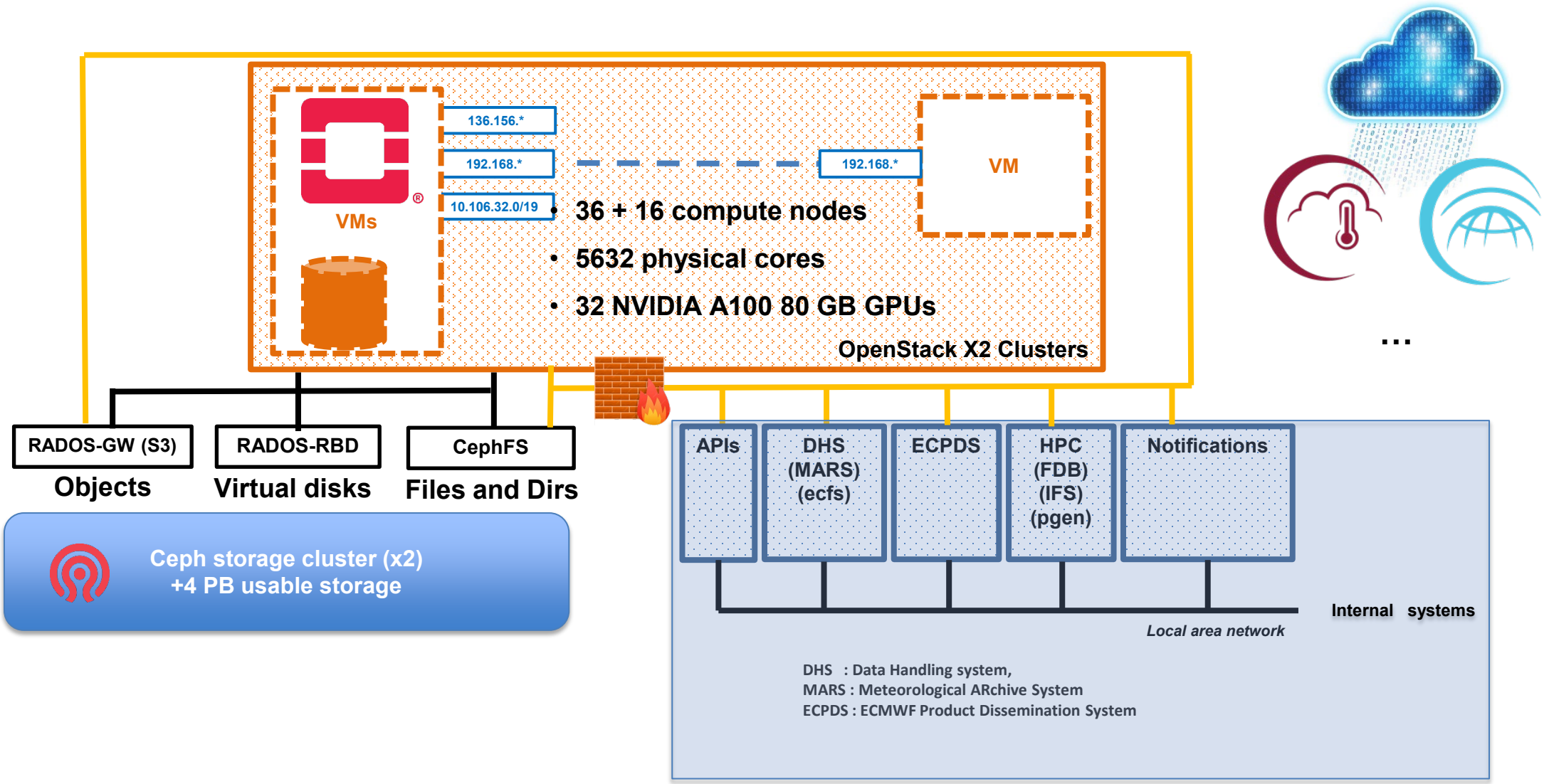




EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF EUMETSAT

The Common Cloud Infrastructure - CCI

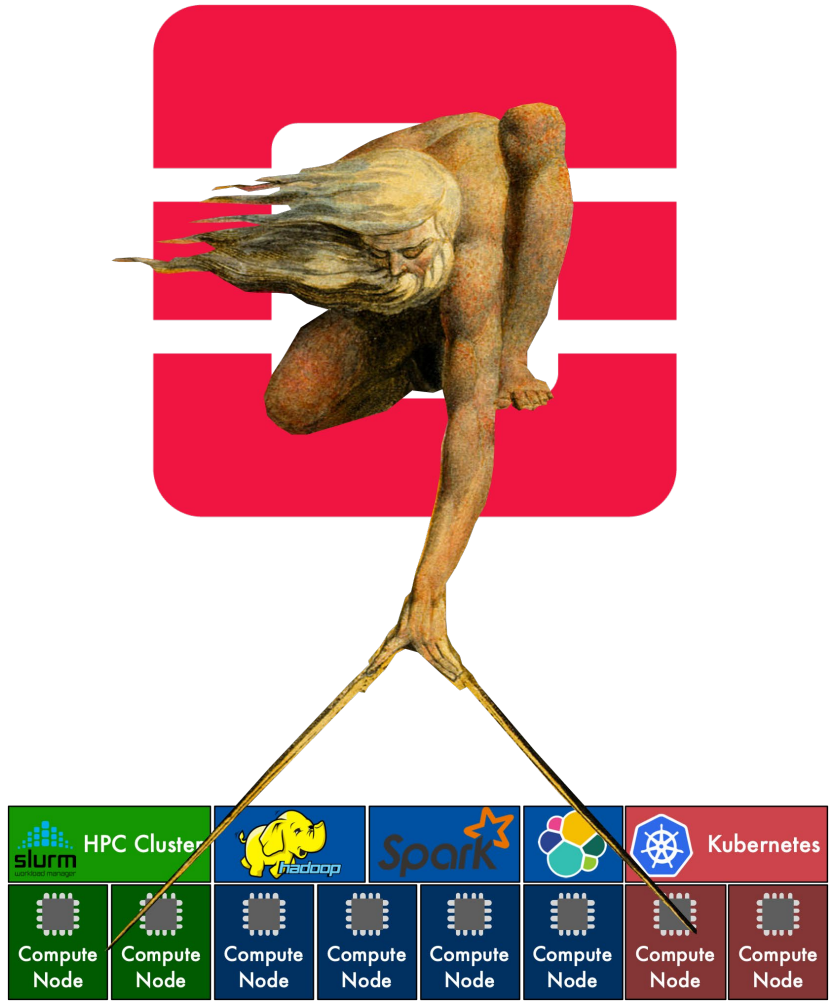


EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

ECMWF EUMETSAT

Building the Production European Weather Cloud

- Potential and Motivations
  - Agility and Flexibility
  - Isolation and Sharing
  - Self-service compute platforms
- Method
  - HPC network fabrics
  - High-performance virtualisation
  - High-speed access to storage
  - Self-service compute platforms



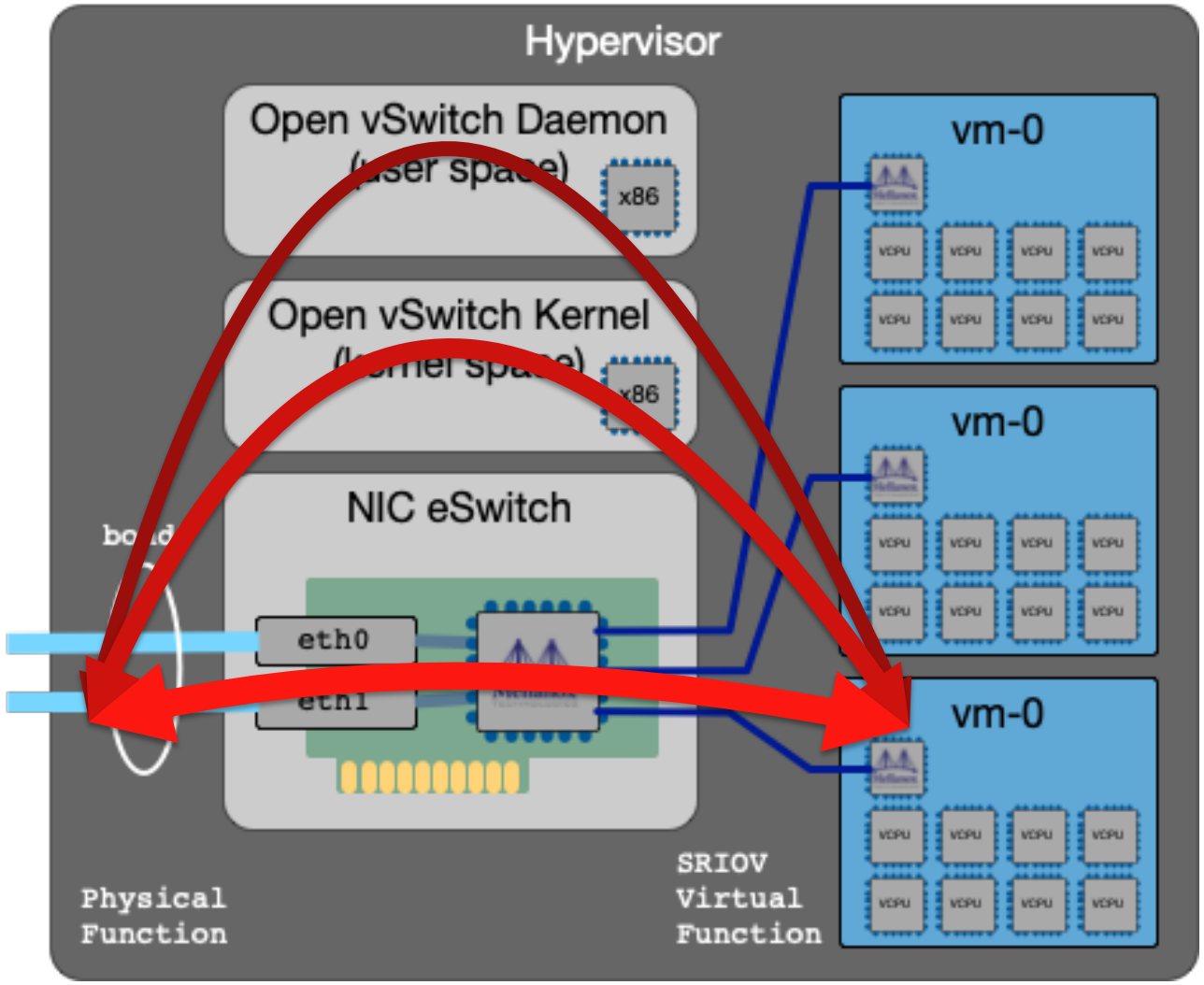
EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

Software-Defined Infrastructure: Agility and Flexibility, Isolation and Sharing

- Implemented using OpenStack and Ceph
  - Multi-tenancy model supports isolation and sharing of resources
  - Fine-grained access control
  - Role-based access control
- Managed using a new generation of tools:
  - Terraform
  - Ansible
  - Helm (Kubernetes)
  - REST APIs
- Self-service compute platforms build on these components

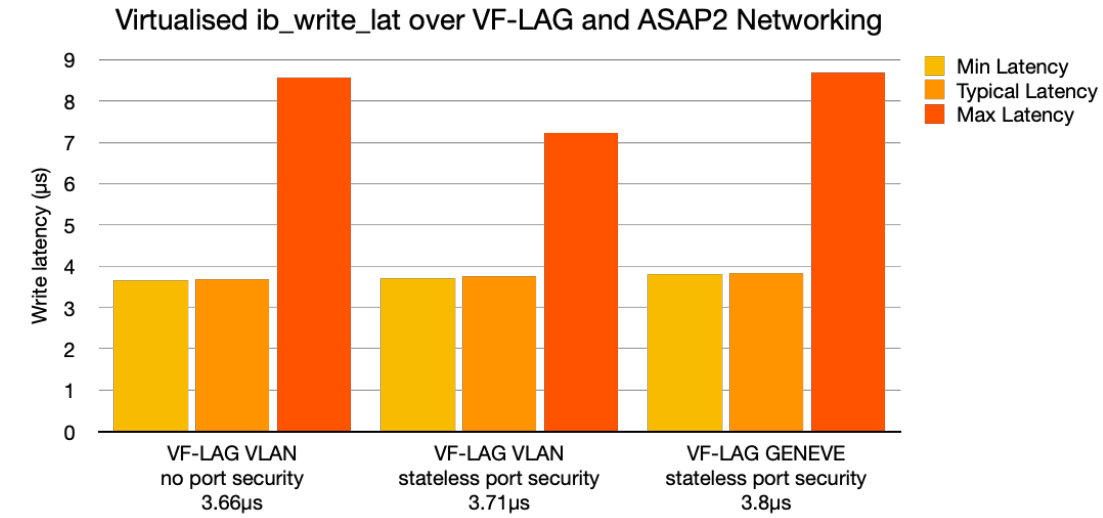
HPC network fabrics

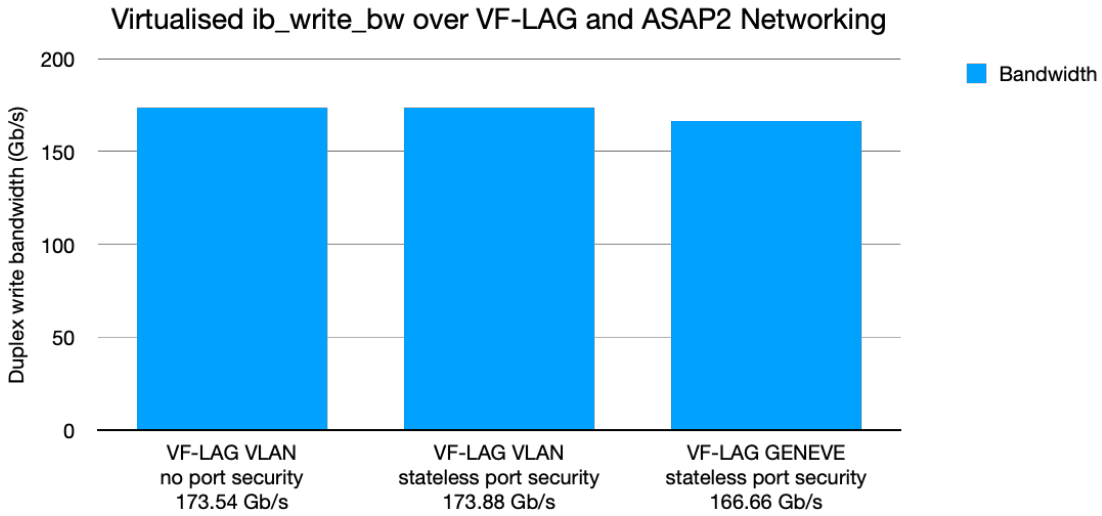
- SR-IOV
  - Present NIC virtual functions to VMs
- VF-LAG
  - Active-active bonded network ports
- ASAP^2
  - Hardware offload of SDN flow rules
- RoCE
  - RDMA over Converged Ethernet
- Live migration supported
- Security groups supported



Virtualised HPC Network Performance

- Ethernet RoCE latency typically 2.6µs
- Virtualisation and SDN overhead is up to 1.0µs over bare metal Ethernet RDMA performance
- Additional functions / overheads incurred are:
  - LAG / VF-LAG
  - ASAP^2
  - SRIOV / virtualisation
  - Virtual tenant network encapsulation
- Bandwidth achieved (depending on VM and hypervisor configuration) can saturate 2x100Gb/s Ethernet links
  - Here benchmarked at 174 Gb/s into a VM

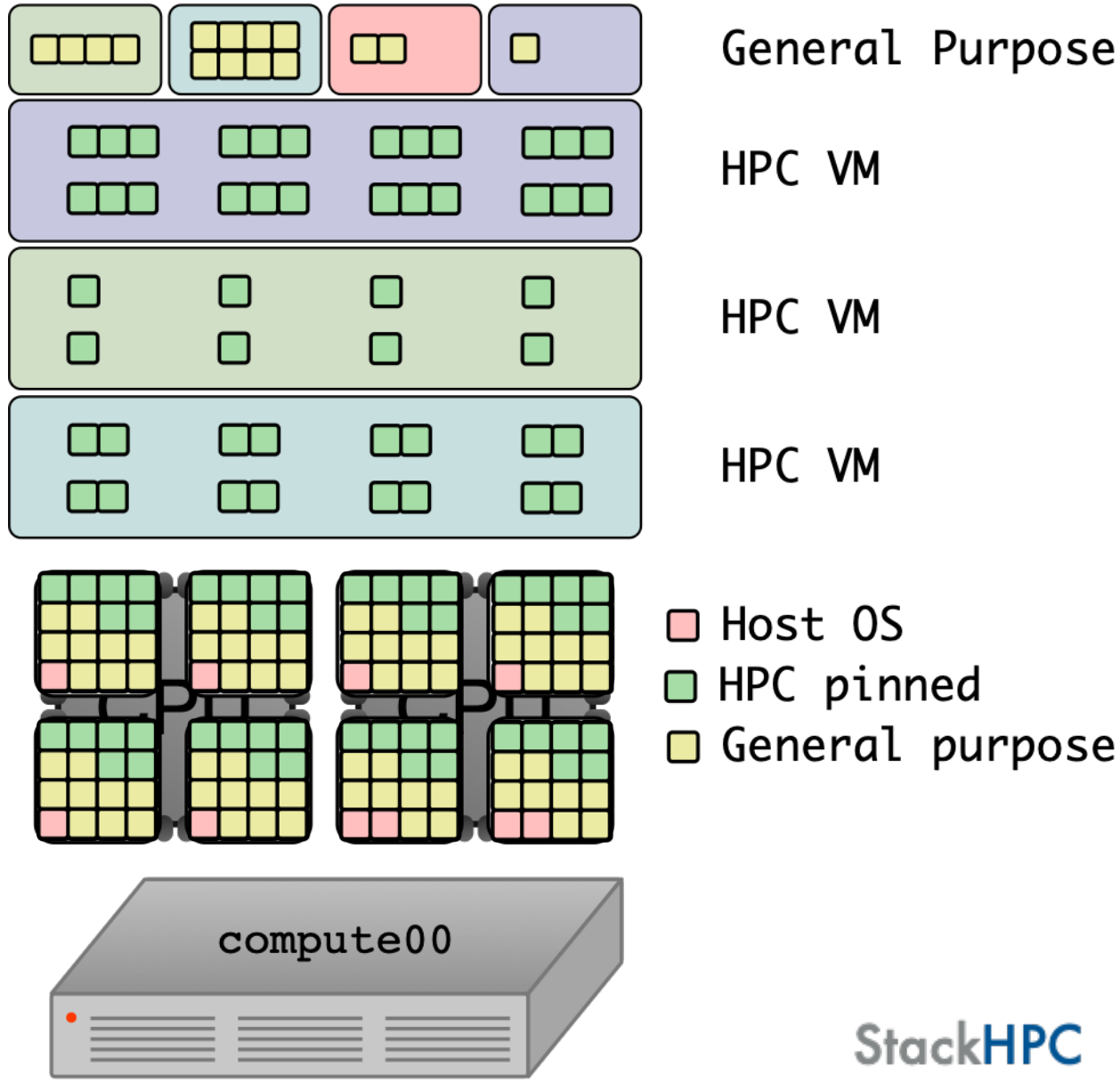




EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

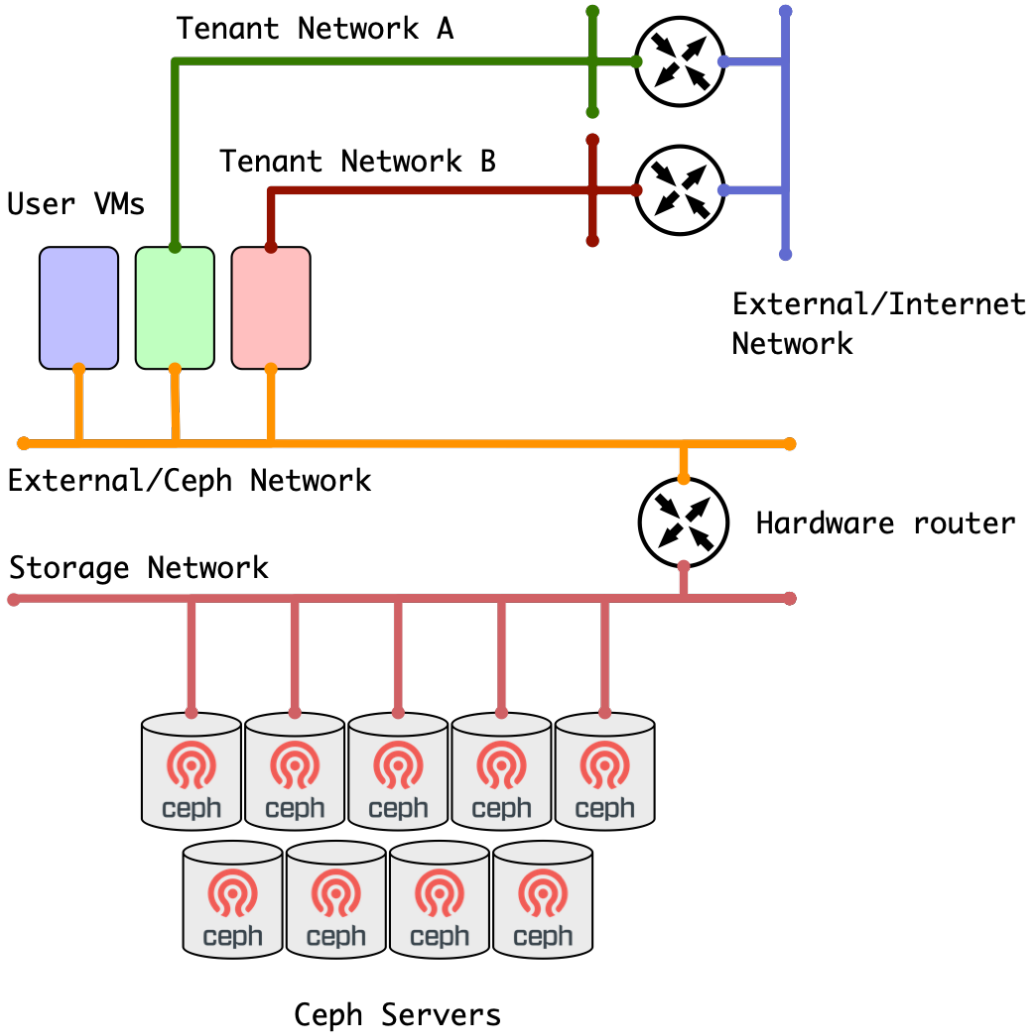
High-performance virtualisation

- Host CPU model passthrough
- Processor core pinning
- Pass through NUMA topology
- Static huge pages
- Also general-purpose flavours



European Weather Cloud Storage

- European Weather Cloud includes significant internal Ceph storage resources.
- Object storage - S3 and Swift APIs
- Block storage - for VM root disks and data volumes
- Ceph filesystem - POSIX-compliant filesystem
  - Shared access to storage network - retaining isolation from other project resources

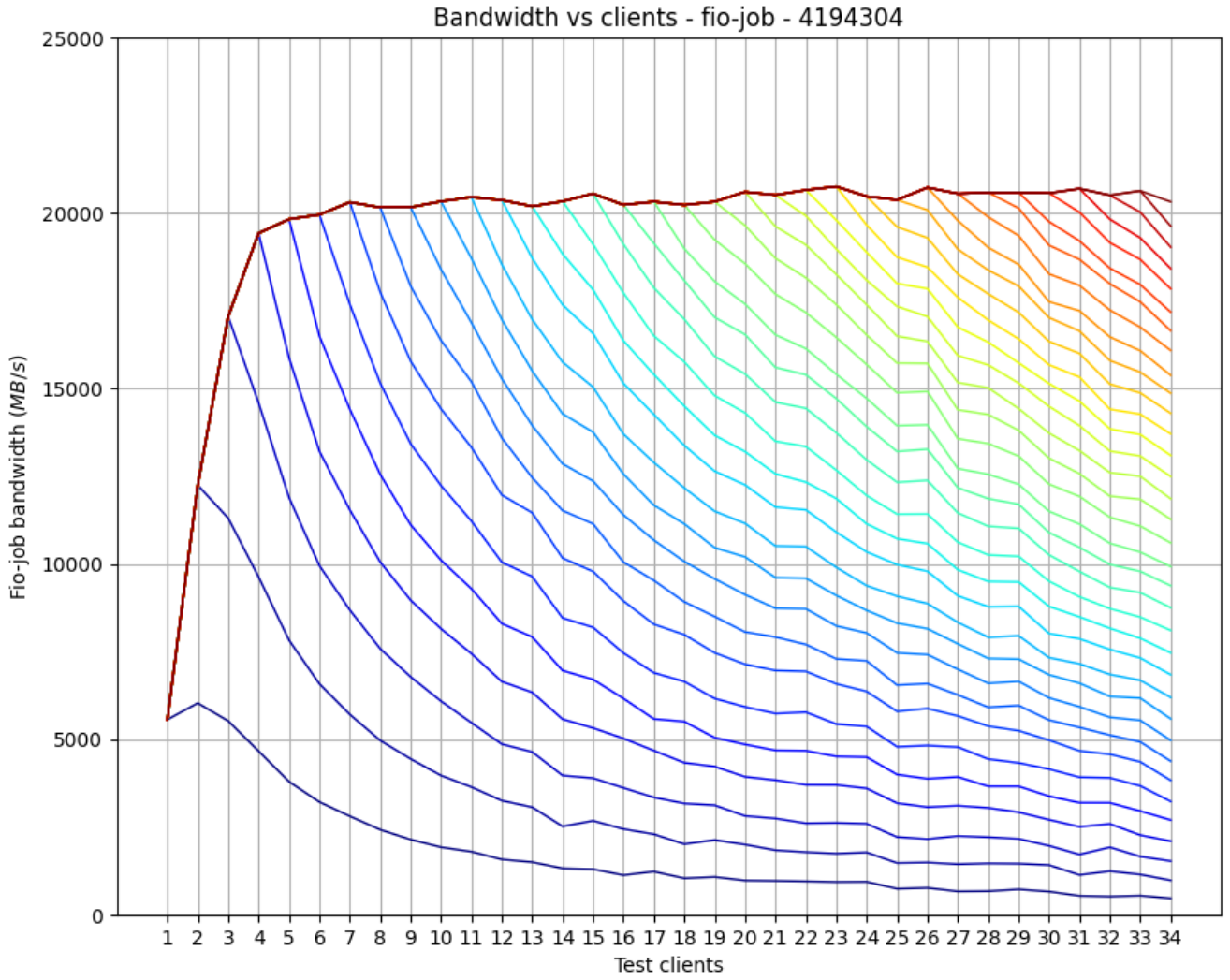


openstack ceph

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

Scale-Out Performance of Ceph Storage

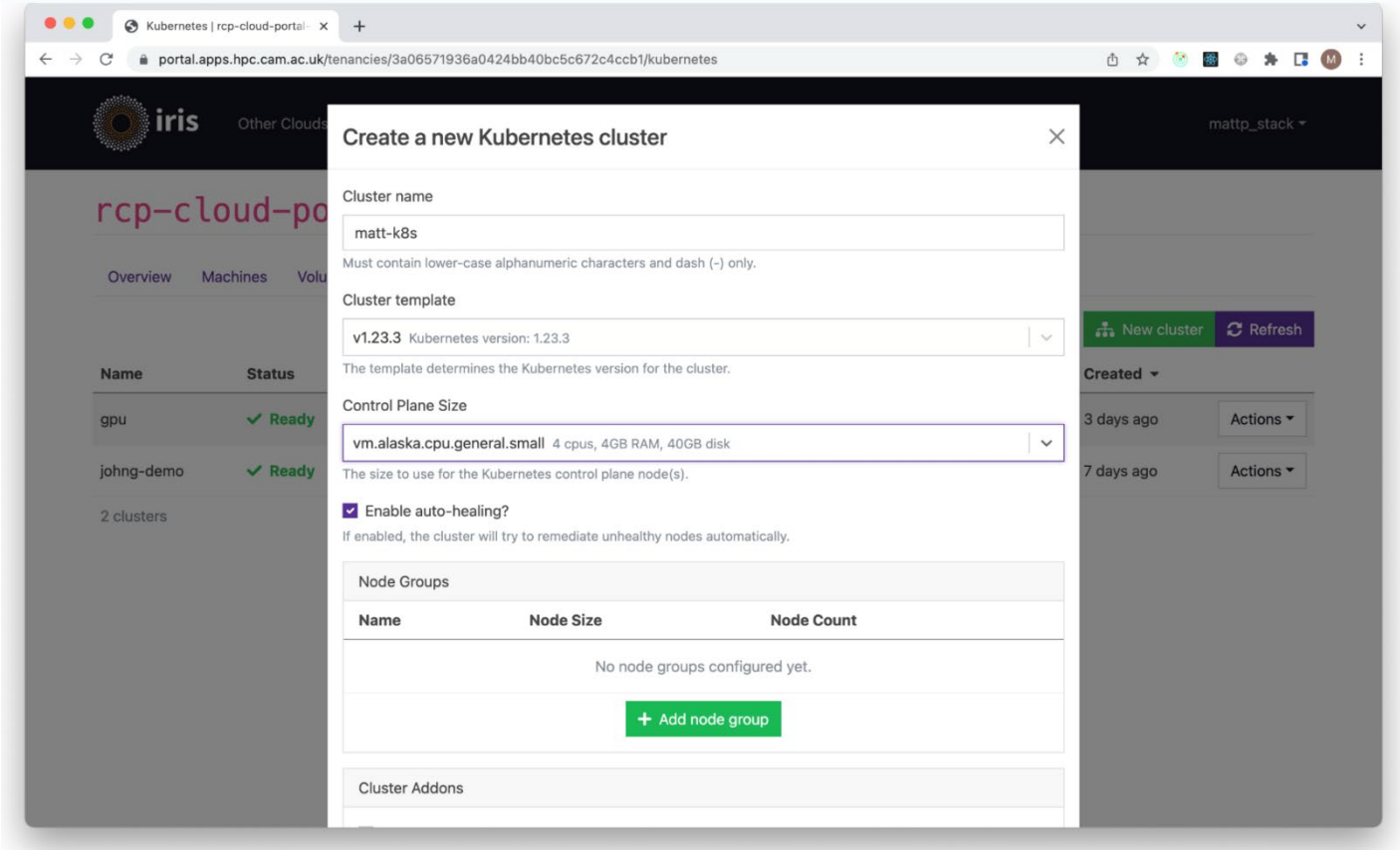
- Individual client VMs sustain ~6 GB/s (4MB IO reads)
- Performance scales to ~21 GB/s aggregate read bandwidth on each cloud
- Performance sustained once storage hardware saturated
- Similar performance levels measured for block and file protocols
- Write bandwidth lower due to data replication factors
- Aggregate bandwidth will scale with hardware extension



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

Self-service Cloud Application Portals: Morpheus, Kubernetes, Azimuth

- Azimuth is a Cloud Portal designed for researchers
- A catalogue of compute platforms is available for self-service
- The catalogue is maintained by cloud admins and configured for optimal HPC integrations
- Compute platforms are defined using cloud-native automation - Terraform, Ansible, Kubernetes, Helm



Azimuth

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

The Software-Defined Supercomputer

- Bringing it all together:
  - The performance of HPC
  - The flexibility of Cloud
- Bare metal, Virtualisation, Containerisation
- Consolidated in a common infrastructure
- Self-service and multi-tenancy
- Driven by open source and open standard APIs

Thank You

Stig Telfer
stig@stackhpc.com
https://www.stackhpc.com

Xavier Abellan
xavier.abellan@ecmwf.int

With thanks to the StackHPC team:

- Pierre Riteau
- Grzegorz Koper
- Will Szumski
- Jakub Darmach
- Justin Coquillon

StackHPC