Federated Cloud User Communities and Proof of Concepts

Salvatore Pinto
Cloud Technologist
EGI.eu

e-infrastructure

EGI FedCloud Communities

- **Ecology** – BioVeL: Biodiversity Virtual e-Laboratory
- ✔ **Structural biology** – WeNMR: a worldwide e-Infrastructure for NMR and structural biology
- ✔ **Linguistics** – CLARIN: 'British National Corpus' service (BNCWeb)
- ✔ **Earth Observation** – SSEP: European Space Agency's Supersites Exploitation Platform for volcano and earthquakes monitoring
- **Software Engineering** – SCI-BUS: simulated environments for portal testing
- **Software Engineering** – DIRAC: deploying ready-to-use distributed computing systems
- ✔ **Software Engineering** – Catania Science Gateway Framework
- **Musicology** – Peachnote: dynamic analysis of musical scores
- **Earth Observation** – ENVRI: Common Operations of Environmental Research infrastructures
- **Geology** – VERCE: Virtual Earthquake and seismology Research
- **Ecology** – LifeWatch: E-Science European Infrastructure for Biodiversity and Ecosystem Research
- **High Energy Physics** – CERN ATLAS: ATLAS processing cluster via HelixNebula

More info: https://wiki.egi.eu/wiki/Fedcloud-tf:Users

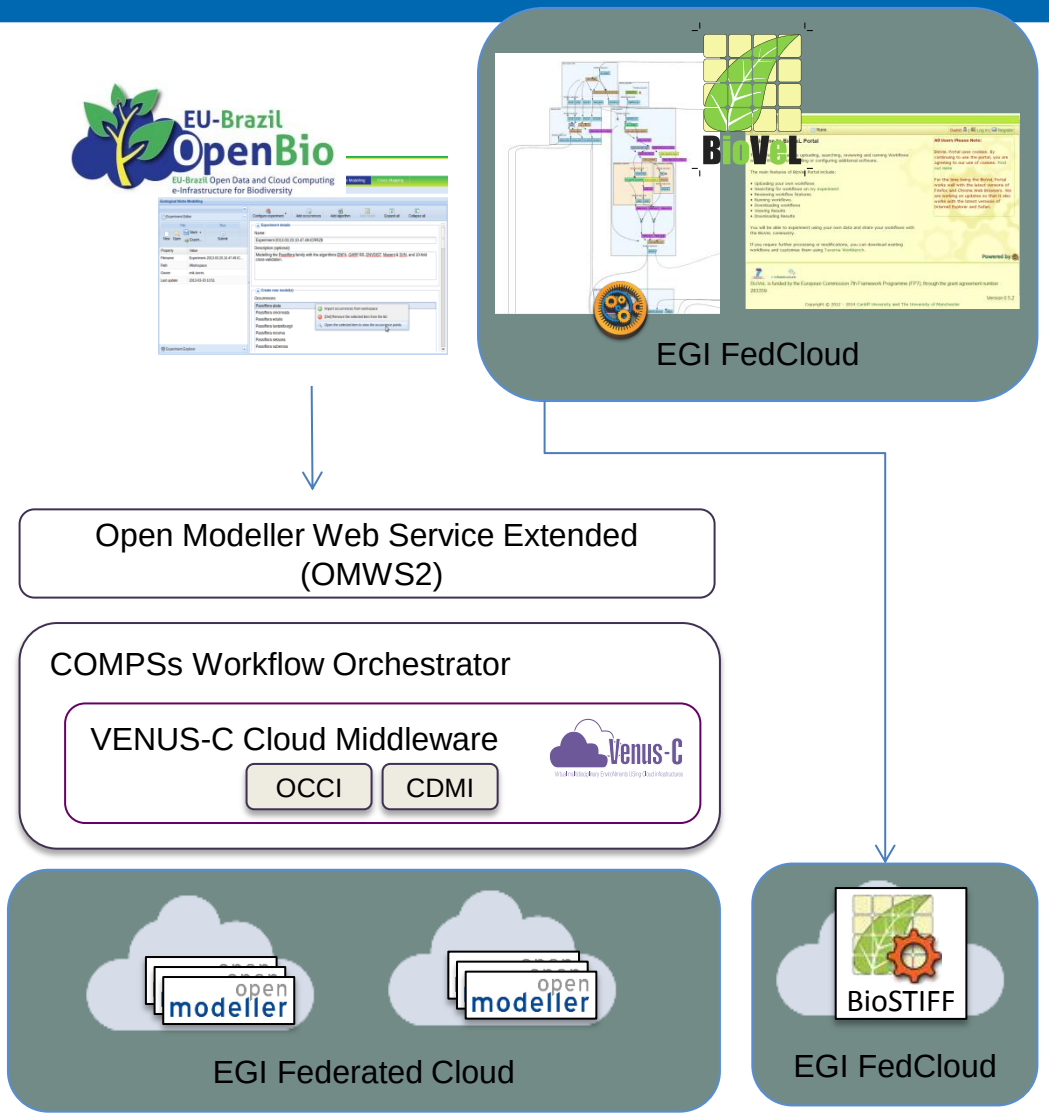
BioVeL

- Use Case #1: OpenModeller 
  - Integration of the OpenModeller web service with the COMPSs framework and VENUS-C middleware to run OpenModeller jobs on the EGI Federated Cloud
  - Requirements:
    - IaaS service with custom images
    - Storage for input layers (currently stored directly on the VMs)

- Use Case #2: BioSTIF ✔
  - BioSTIF service is a web service for data visualization
  - The activity is to integrate the
  - Requirements:
    - IaaS service with basic RHEL image and contextualization
    - Cloud Storage for input layers and user custom data
    - High Availability

BioVeL & EUBrazilOpenBio

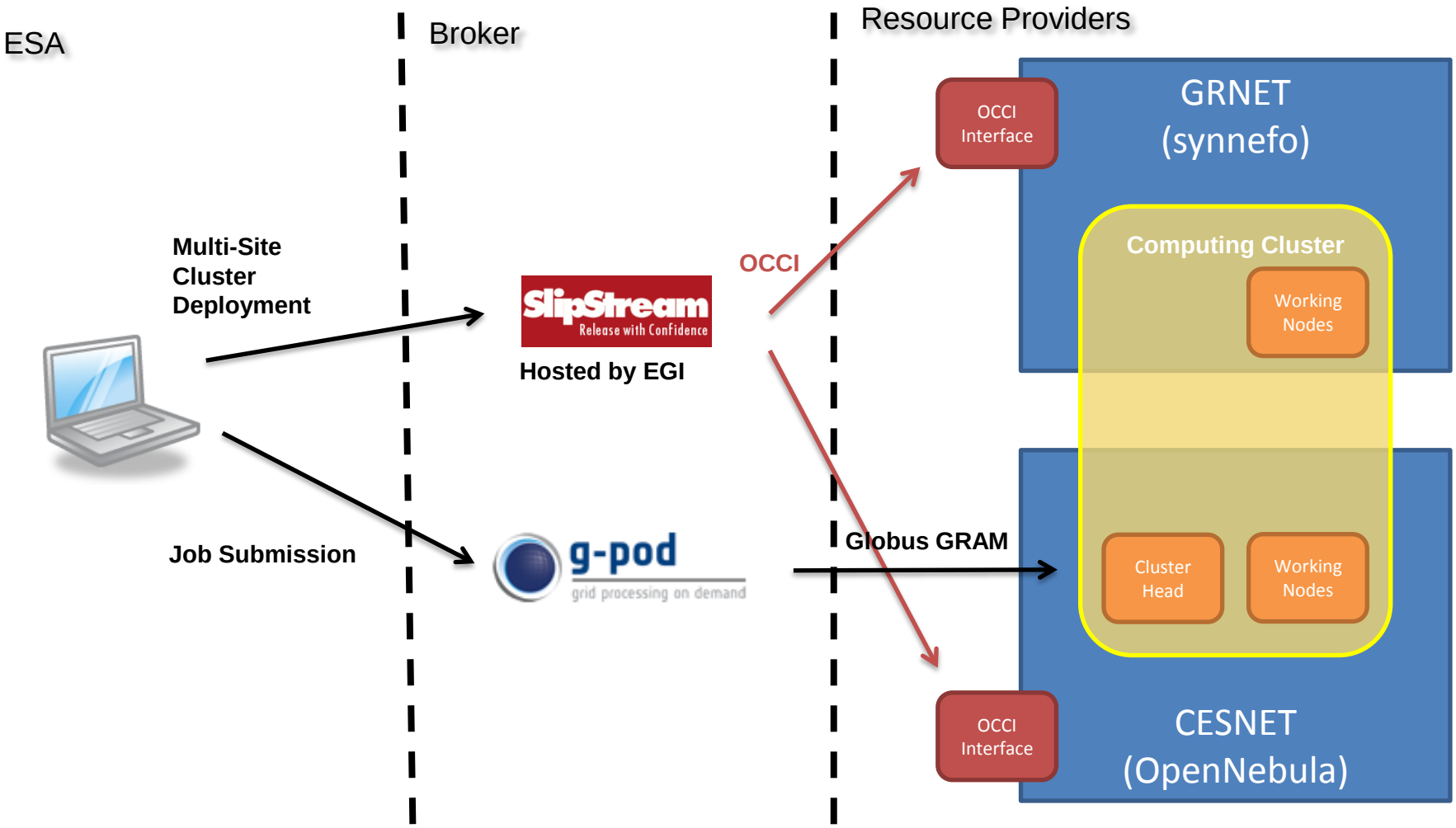
- Services are accessed from the **BioVeL Portal**, **Taverna Workbench** and **EU-Brazil OpenBio VRE**.
- The openModeller Web Service interface (**OMWS2** in the picture) is integrated into the **COMPs Workflow Orchestrator**, which, using the **VENUS-C** middleware, dynamically instantiates openModeller workflows on **cloud resources** from different providers.
- **BioSTIFF** services are also integrated in the EGI Federated Cloud, to be used in the ENM workflow.
- **BioVeL Portal** is also running on the EGI Federated Cloud
- Deployment of BioSTIFF and BioVeL Portal services on the EGI Federated Cloud is performed via contextualization and Ansible scripts



SSEP

- Use Case #1: Data Processing Services ✔
  - Enable users to invoke processing of data from the ESA archive on the Cloud. A custom Globus cluster is started on demand and connected to the ESA Grid system to absorb peaks in the processing.
  - Requirements:
    - IaaS service with basic RHEL images and contextualization
    - Infrastructure Broker for automated cluster deployment on multiple sites
    - Block storage attached to the VM
    - Fast network connection between the sites

ESA PoC Scenario 2: Multi site Processing Cluster



SSEP

- Use Case #3: Virtual Laboratory
  - Run the SSEP Virtual Laboratory (CloudToolbox) VMI for accessing SSEP processing services and elaborate the results
  - Requirements:
    - IaaS service with custom images
    - Personal storage for user data
    - Persistent images
    - GUI interface for starting the VMs

SSEP (Virtual Laboratory)

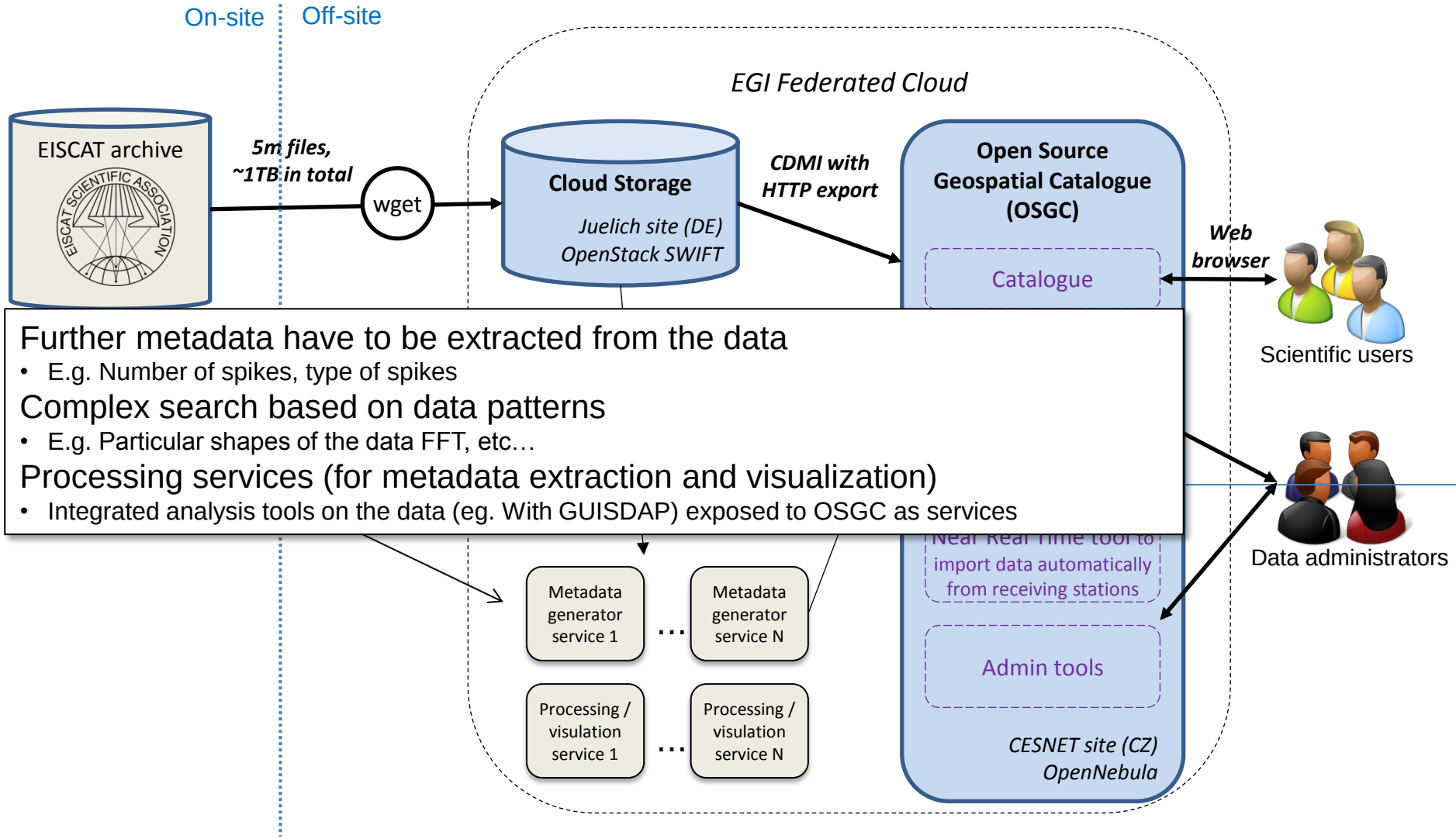
CloudToolbox provides an on-demand Virtual Laboratory on the cloud (with Remote Shell and Virtual Desktop support), equipped with all the basic tools for SAR processing (full GAMMA licensed software, NEST, BEAM) and linked to the VA4 data archive.

ENVRI

- Use Case #1: Data access and dissemination
  - Use the EGI Federated Cloud services to perform hosting, cataloguing (according to custom metadata) and dissemination
  - Requirements:
    - OpenSearch Catalogue PaaS
    - STorage-as-a-Service (integrated within the PaaS solution to provide hosting)

- Use Case #2: Data Processing Services
  - Integrate the ENVRI processing services developed by CNR-ISTI into the EGI Fedearated Cloud. Services can runs as a set of workers on custom VM images or on an Hadoop cluster.
  - Requirements:
    - IaaS service with custom image
    - Hadoop cluster for Big Data processing

Data Access and Dissemination PaaS (concept with EGI solutions)



Further metadata have to be extracted from the data
- E.g. Number of spikes, type of spikes

Complex search based on data patterns
- E.g. Particular shapes of the data FFT, etc…

Processing services (for metadata extraction and visualization)
- Integrated analysis tools on the data (eg. With GUISDAP) exposed to OSGC as services

Hybrid Cloud Federation

- Use Case #1: Hybrid cloud federation
  - Use the Helix Nebula broker and common core services (ex. common AAI) to federate private cloud resources coming from different institutions, keeping compatibility with public resources for burst processing.
  - Requirements:
    - IaaS & STaaS service, with basic and custom images support
    - Fast network for data sharing
    - Automatic VM deployment close to the data (avoid large data transfer)
    - Common AAI, enable easy resource sharing between the sites

Thank you

EGI Federated Cloud resources

- Wiki site: http://go.egi.eu/fedcloud
- Mailing List: fedcloud-tf@mailman.egi.eu
- User support: ucst@egi.eu

e-infrastructure