The EGI Federated Cloud

Diego Scardaci
diego.scardaci@egi.eu
EGI.eu/INFN

Outline

- **Concept of the EGI Federated Cloud: vision**
  - as a technology to build community-clouds
  - as an infrastructure deployed on NGI resources
- **EGI Federated Cloud: state-of-the-art and technologies**
  - VM Management, Appliances Marketplace, Data management, Information discovery, Accounting, Security
- **High level tools with main capabilities (PaaS, SaaS)**
- **Typical setups and use cases - through examples**
  - Usage Models: deployment of services, heavy computation, heavy memory applications, big data, etc.
- **Plans, next steps**
- **User support**
- **Next steps to become an user**

Concept of the EGI FedCloud Vision

Rationale

Growth of Providers

- **High Throughput Platform:** academic resource providers
- **Federated Cloud Platform:** diversity of resource providers

Growth of Research Communities

Tens of 1000's

Few related use cases
Single application model

Millions

Many diverse use cases
& application models

What is the EGI Federated Cloud

The EGI Federated Cloud is federation of institutional private Clouds, offering Cloud Infrastructure as a Service to scientists in Europe and worldwide.

EGI Federated Cloud is based on:

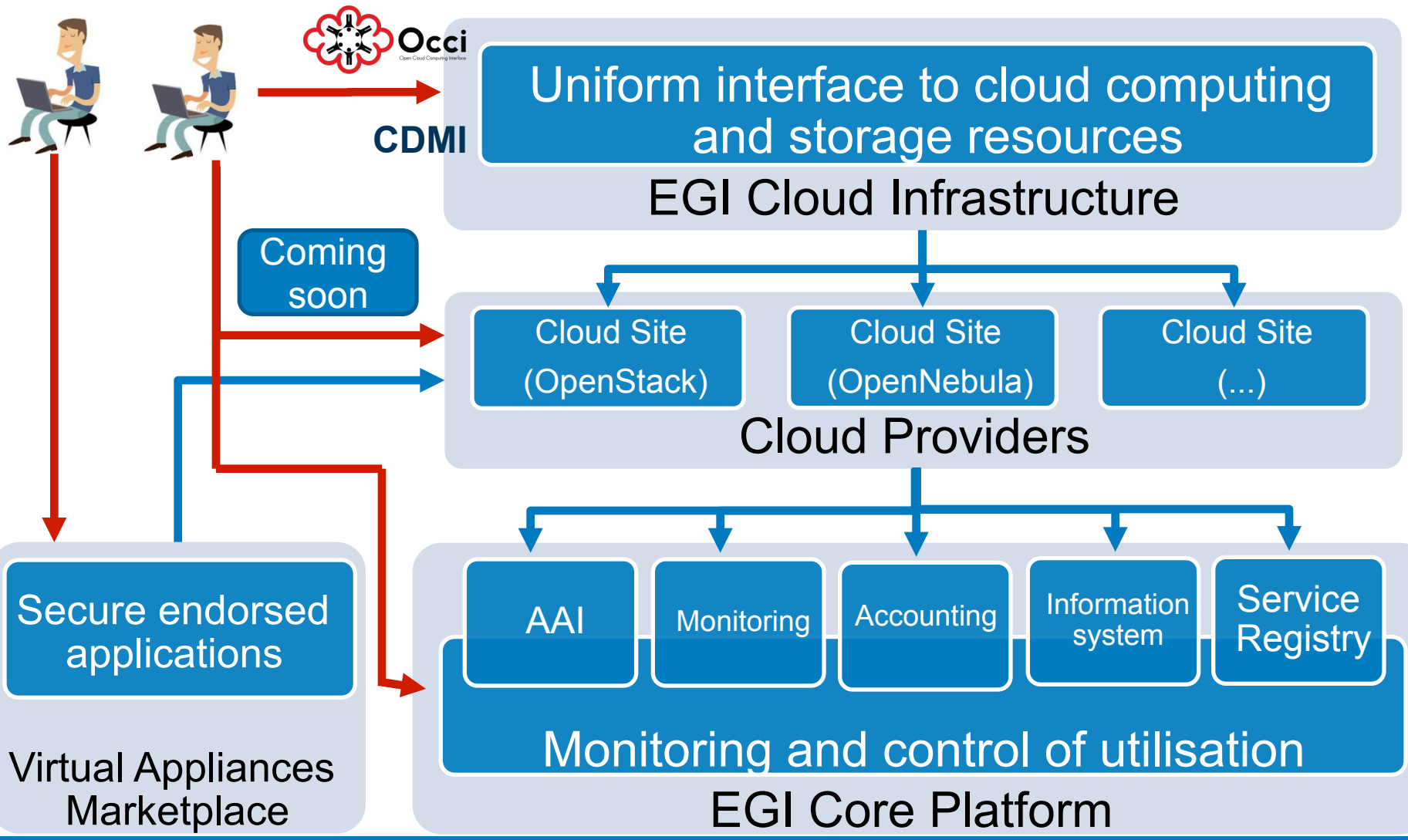
- **Standards and validation**: federation is based on common **Open-Standards** – OCCI, CDMI, OVF, GLUE, etc...
- **Heterogeneous implementation**: no mandate on the cloud technology, the only condition is to expose the chosen interfaces and services.

FedCloud as a technology to build community-clouds

The EGI Federated Cloud is **providing access to digital resources on a flexible environment**.

Cloud resources	• Computing intensive experiments • Data intensive experiments • Hosting services
Integrated in EGI infra	• AAI & monitoring • Accounting & information discovery • Service registry
Standard based	• Heterogeneous implementation • OCCI, CDMI, GLUE2, OVF, etc.
VM Catalogue	• EGI Certified and Secure endorsed VMs • Register your VMs • Re-use public VMs
Single system	• Set of independent cloud services • Uniform interfaces
Target groups	• Individual researchers • Larger research communities or groups

The EGI Federated Cloud



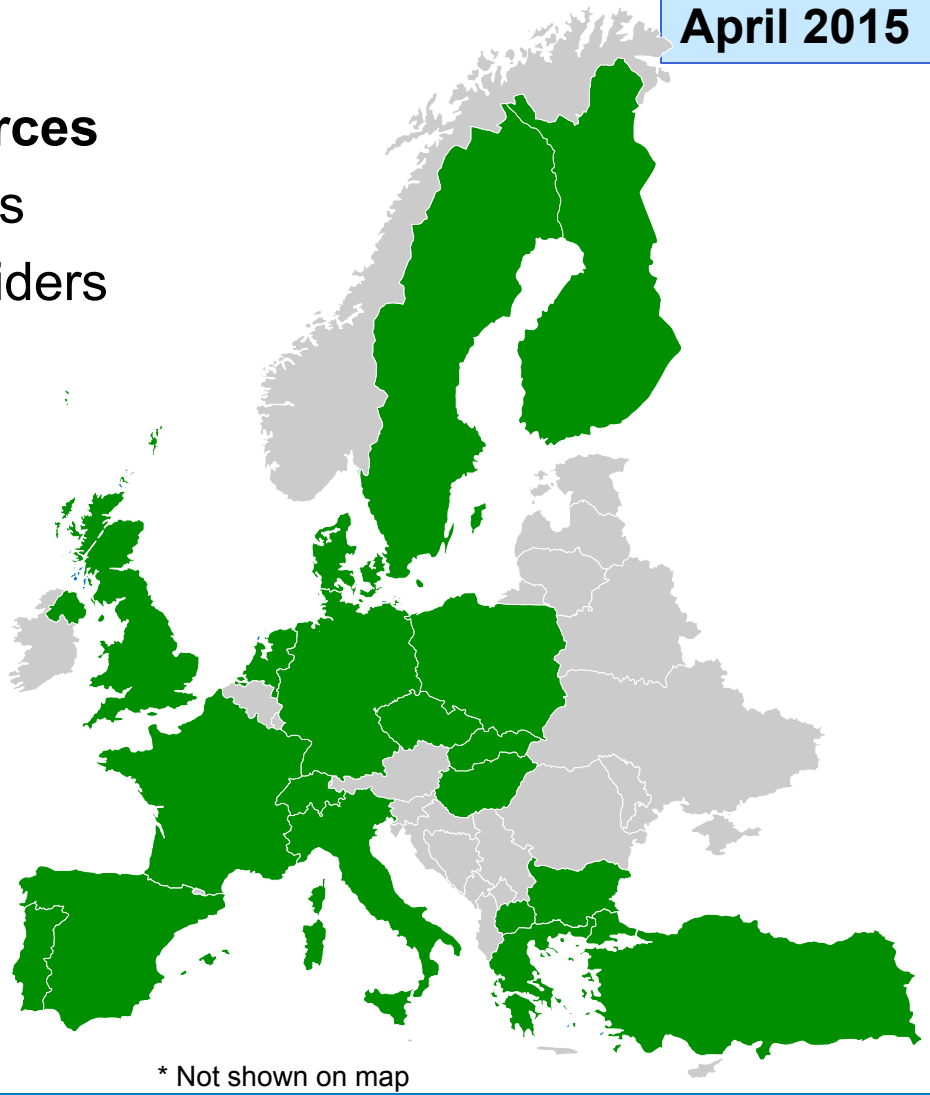
Overview of the cloud services

Features/CMFs	OpenStack	OpenNebula	Synnefo	AWS	MS Azure
OCCI Compute Management	Yes	Yes	Yes	Yes	Coming Soon
OCCI Storage Management	Yes	Yes	Yes	Yes	Coming Soon
OCCI Basic Network Management	Yes	Yes	Yes	Yes	Coming Soon
OCCI FedCloud Contextualisation	Yes	Yes	Yes	Yes	Coming Soon
CMFs Native Interface	Coming Soon	TBD	TBD	No	No
CDMI	Coming Soon	No	Yes	No	No
Cloud-Init	Yes	Yes	Yes	Yes	Coming Soon
Integrated in EGI Core Platform	Yes	Yes	Yes	TBD	TBD

FedCloud as an infrastructure deployed on NGI resources

Resources

- **13 NGIs provide 22 certified resources**
- 4 NGIs currently integrating resources
- 5 NGIs with interested resource providers
- Worldwide interest & integration
  - Australia* (NeCTAR)
  - Africa* (SAGrid)
  - South Korea* (KISTI)
  - United States* (NIST, NSF)



* Not shown on map

Usage – May/Dec 2014

- 244,913 (397,128) VMs certified (uncertified)
- 10.6M (12.3M) CPU hours (wall time)

EGI Federated Cloud
State-of-the-art & Technologies

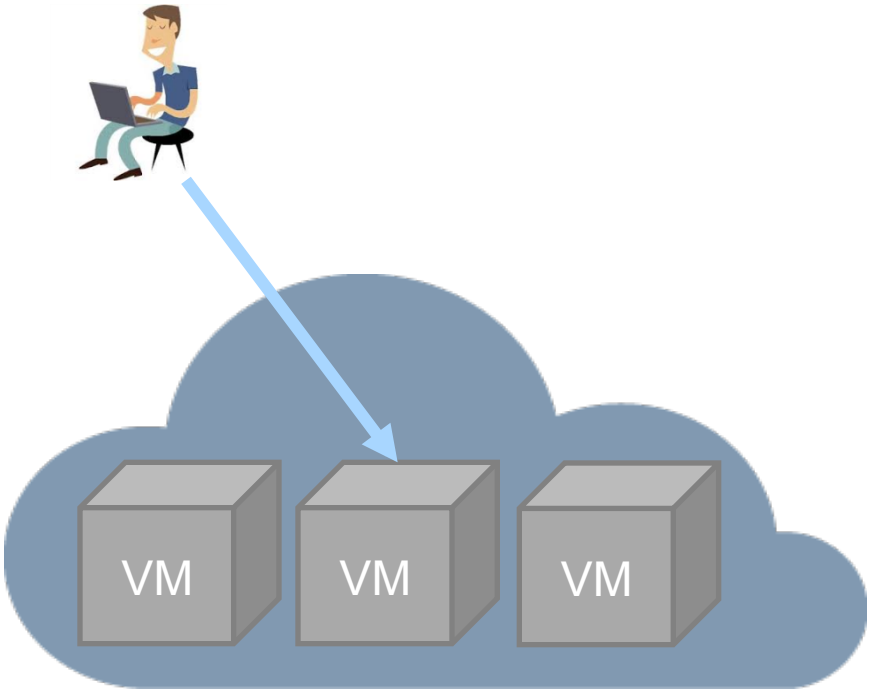
EGI Federated Cloud services

Uniform interface to cloud computing and storage resources

EGI Cloud Infrastructure

OCCI standard tools to easily manage VMs:

- Start, Suspend, Stop, Destroy, etc.



Simple usage

- OCCI standard
- rOCCI CLI client
- jOCCI – JAVA API

Scale to your needs

- Number of CPU cores
- Amount of RAM
- Local Disk

Contextualisation

- Configure your VM at startup
- Install SW at startup
- Complex deployment

Data Management

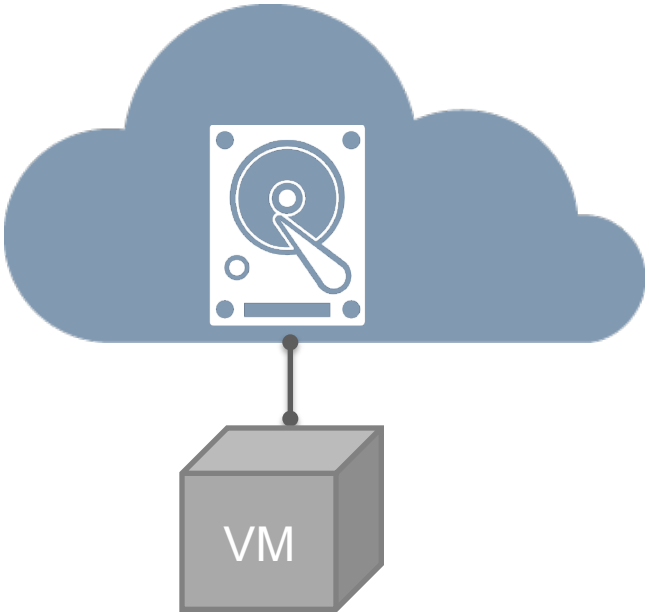
3 solutions available:

- VM Disk Space
- Block Storage (OCCI standard)
- Object Storage (CDMI standard)

VM Disk Space	• Disk stored in the VM image (<100 GB)
Block Storage	• Disk attached to a running VM (any size)
Object Storage	• External storage system • Access through REST API

Data Management - Block Storage

Persistent block level storage to use with VMs



Simple usage

- Use as any other block device from VMs
- Snapshotable

High Performance

- Consistent and low-latency performance
- SSDs (in some sites)

Scale to your needs

- From GB to TB
- Create and attach to VMs on demand

Data Management - Object Storage

Data storage infrastructure for storing and retrieving data from anywhere at any time

API Access

- Simple REST APIs for managing and accessing data

Scalable

- Store as much data as needed.
- Get accounted only for the space used.

Sharing

- Define ACLs on each object, share publicly your data

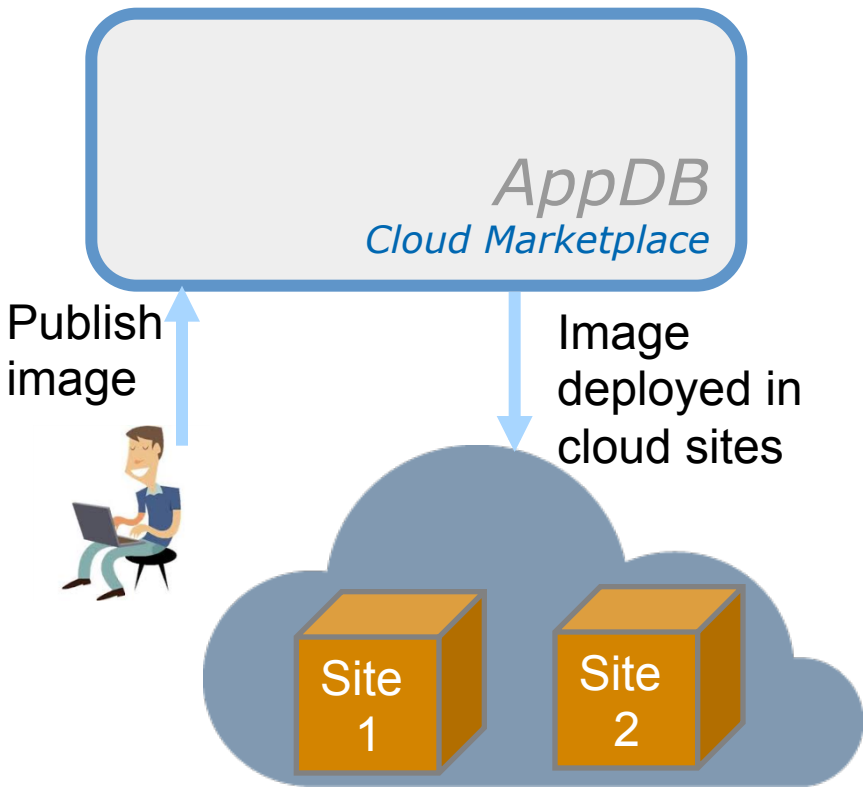
Virtual Appliances Marketplace

Secure endorsed applications

Virtual Appliances Marketplace

Appliances Marketplace

Web-based VM image catalogue with automatic deployment on the EGI Federated Cloud sites



Cloud VM image repository

- Publish your image
- Re-use images certified by EGI
- Manage versions

Manage your VM images

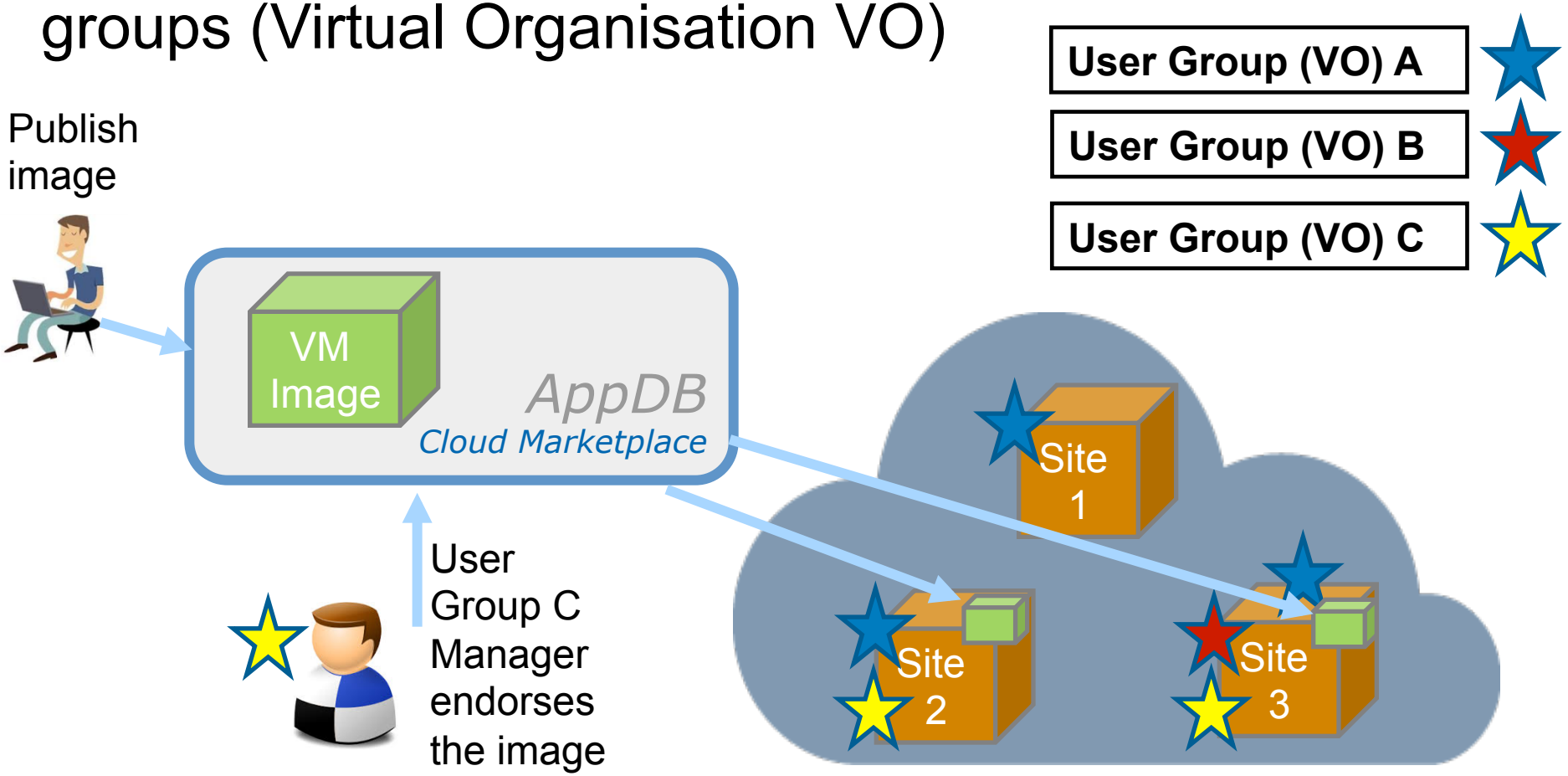
- Sharing
- Create private images

Automatic deployment on the cloud sites

- Images automatically deployed on the cloud sites
- Automatic update

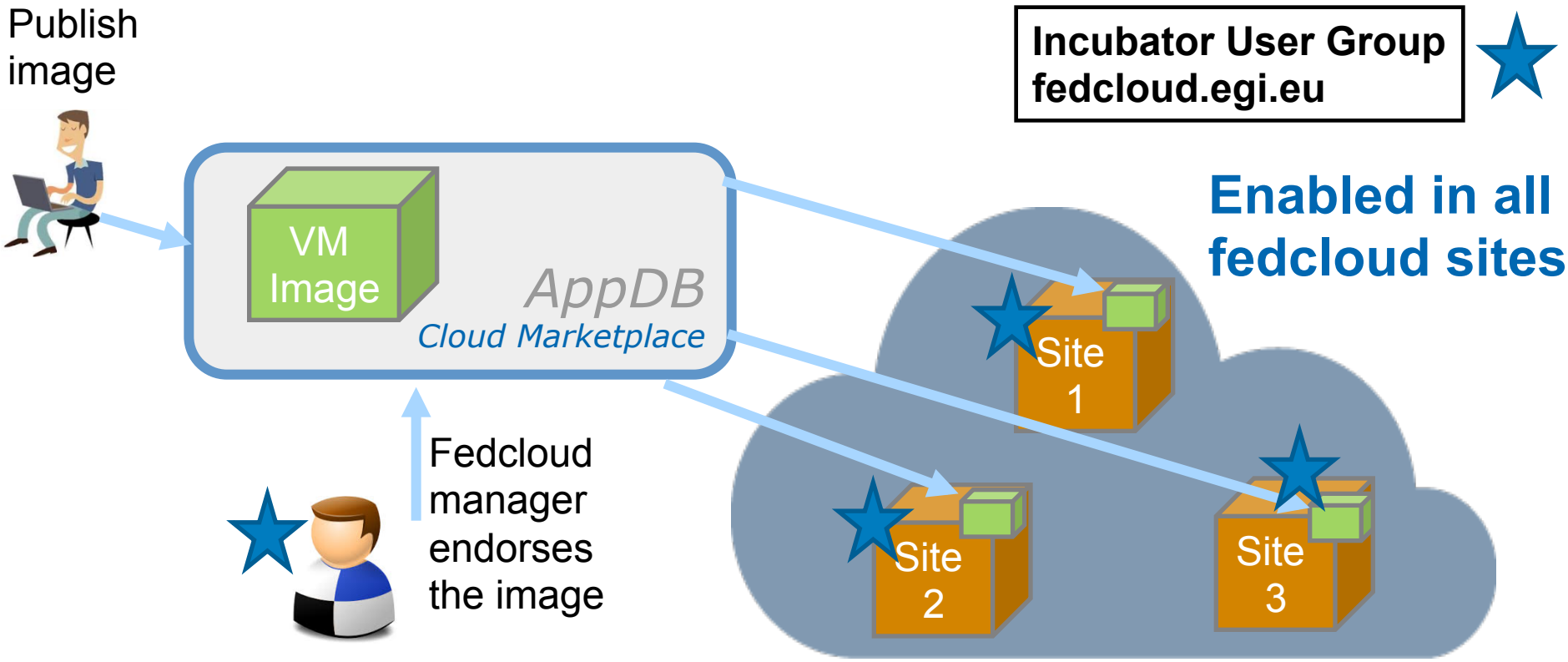
Image distribution in the EGI Federated cloud

Image distribution on the cloud based on user groups (Virtual Organisation VO)

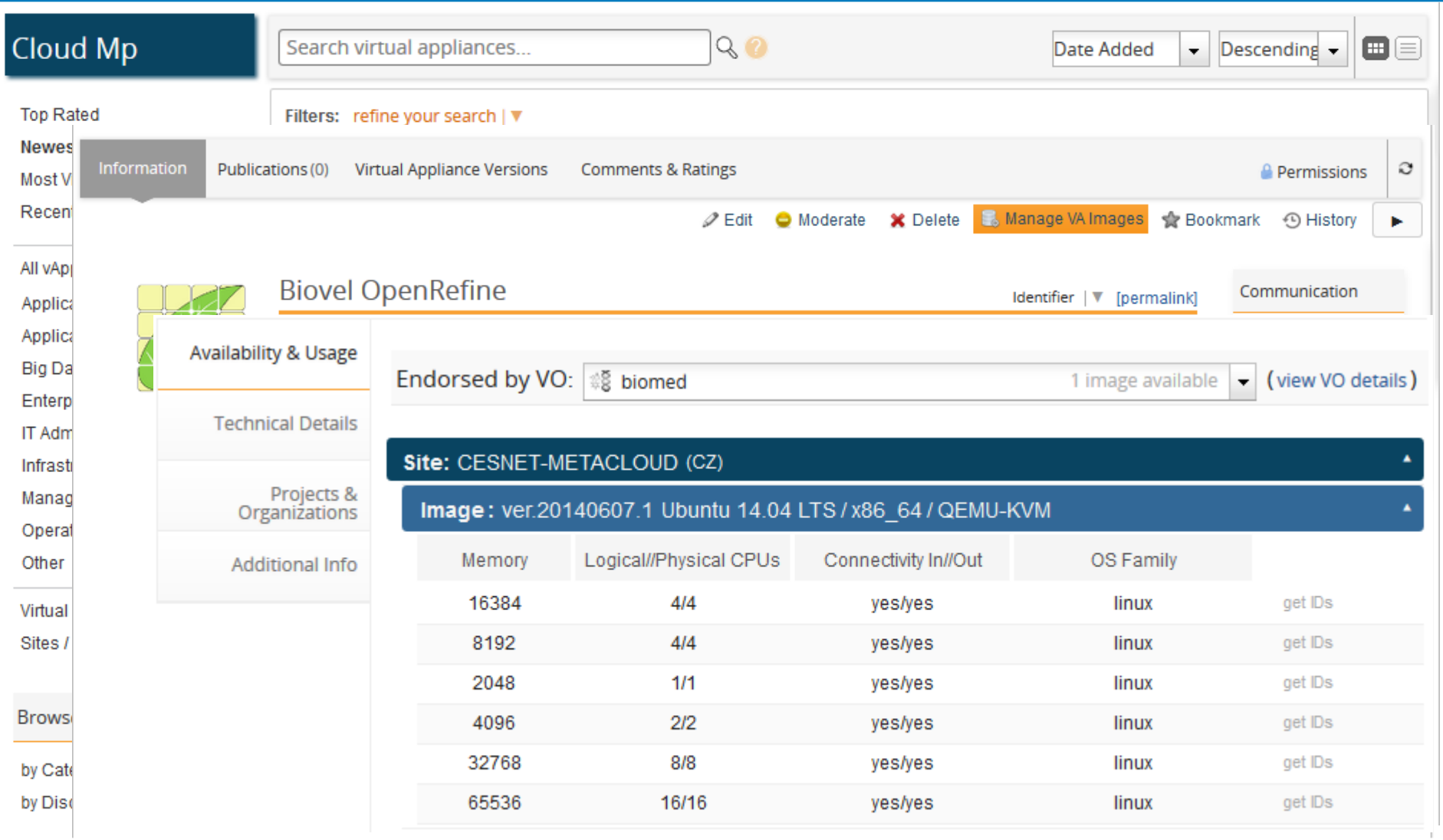


Incubator User Group/ Virtual Organisation

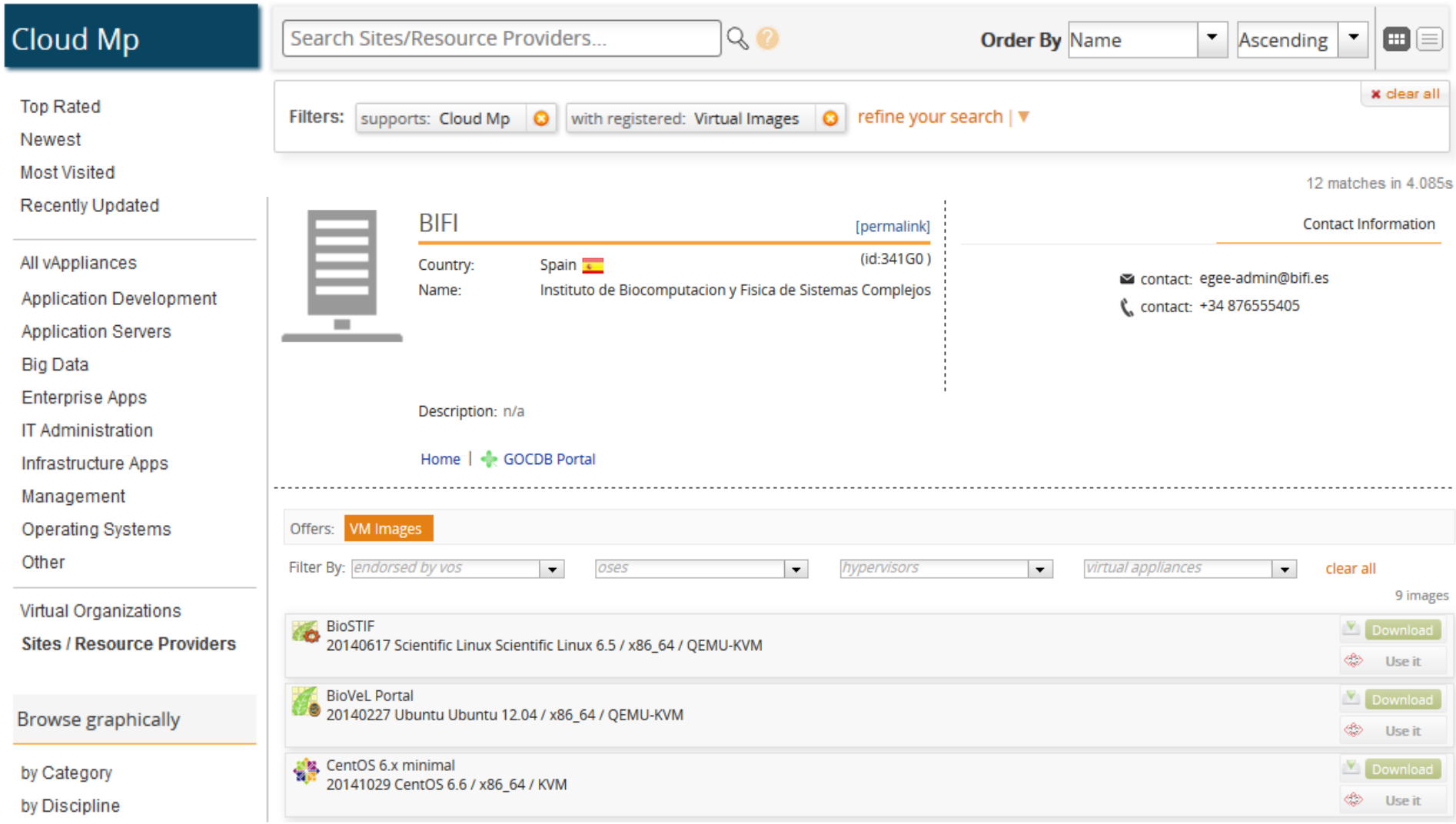
Incubator user group/VO for application prototyping and validation for any new users



Appliances Marketplace



Appliances Marketplace



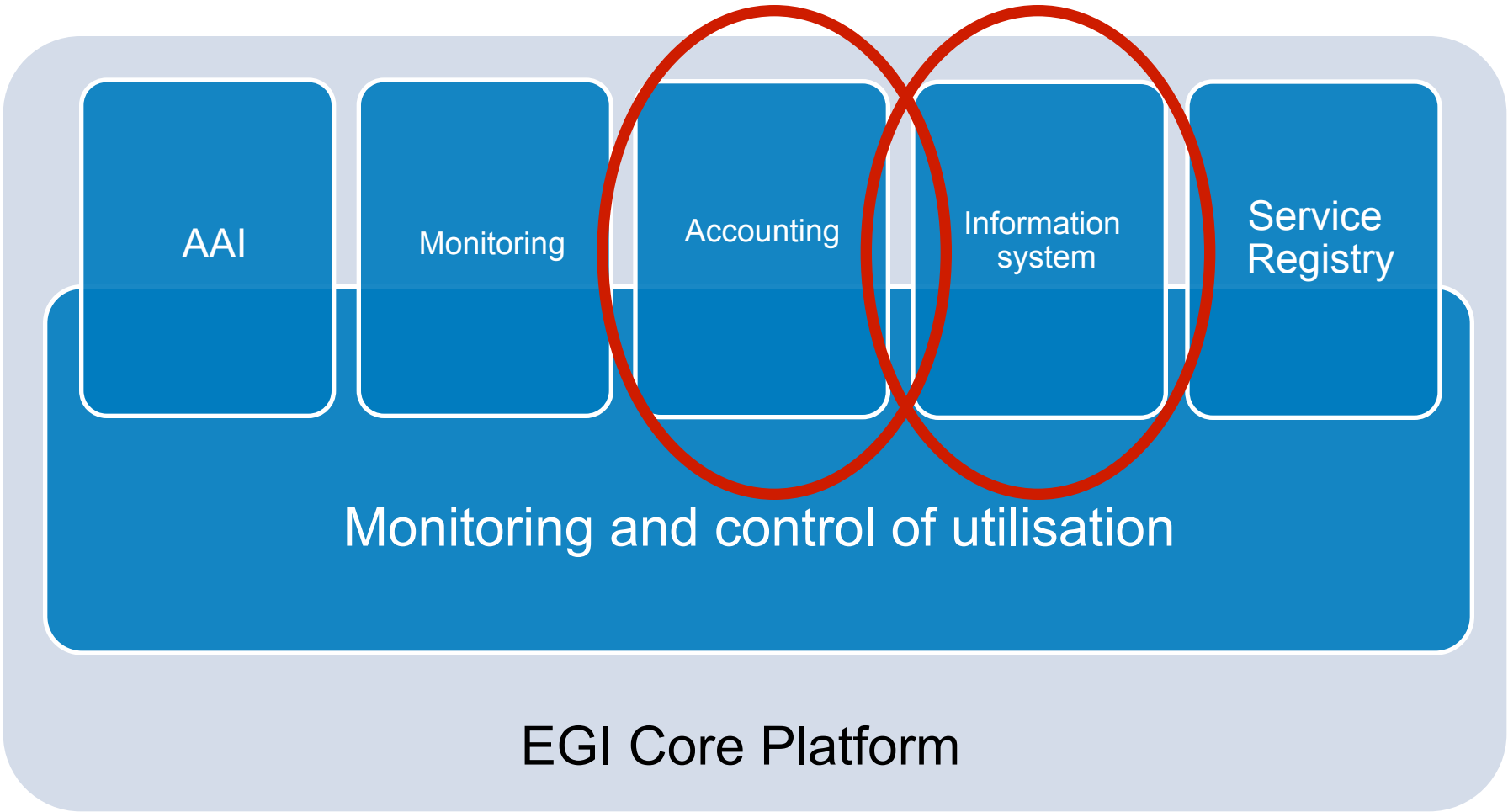
Security

Secure endorsed application and service deployment

EGI CSIRT
Computer Security Incident Response Team
Web page

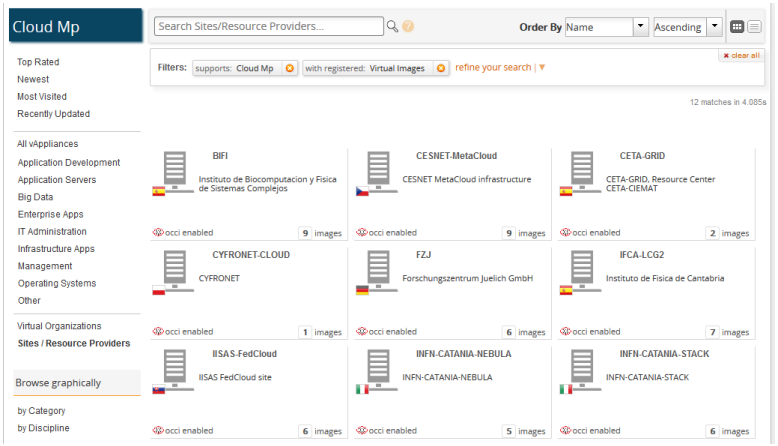
EGI VM Images	• Endorsed by the EGI security team • Periodically updated
User VM Images	• Guidelines to create secure images • Each user can ask for an EGI endorsement
Security monitoring	• Continuous monitoring of the running VMs • Alarms and immediate actions in case of security issues

EGI Core Platform



Information discovery

Information about the available resources and their current status



Sites

- List of the certified EGI Federated Cloud sites

VM images

- List of images available on each site
- Get IDs to be used with the rOCCI client

Views

- High level view for final users (VA marketplace)
- Advanced view for expert users through LDAP

Data about usage of the FedCloud resources

The following table shows the distribution of Total number of VM run grouped by SITE and DATE.

Total number of VM run by SITE and DATE

SITE	Nov 2014	Dec 2014	Jan 2015	Feb 2015	Mar 2015	Apr 2015	Total	%
100IT	716	709	748	675	743	48	3,639	1.08%
BIFI	11,113	15,293	18,209	42,974	47,682	0	135,271	40.00%
CERN-PROD	13,310	185	437	8,085	171	34	22,222	6.57%
CESGA	1,726	929	0	1	556	52	3,264	0.97%
CESNET-MetaCloud	1,258	1,407	792	837	770	49	5,113	1.51%
CETA-GRID	153	99	2,307	227	454	45	3,285	0.97%
CYFRONET-CLOUD	456	0	0	235	751	47	1,489	0.44%
FZJ	5,530	17,979	10,385	21,145	3,088	47	58,174	17.20%
GoeGrid	1,504	798	751	681	752	49	4,535	1.34%
HG-09-Okeanos-Cloud	493	229	241	185	179	0	1,327	0.39%
IFCA-LCG2	4,426	650	621	539	537	35	6,808	2.01%
IISAS-FedCloud	2,970	6,014	5,357	2,766	3,435	207	20,749	6.13%
IN2P3-IRES	0	0	0	188	729	44	961	0.28%
INFN-CATANIA-NEBULA	663	705	779	672	767	49	3,635	1.07%
INFN-CATANIA-STACK	115	202	847	697	774	47	2,682	0.79%
INFN-PADOVA-STACK	5,482	8,652	13,131	11,162	7,683	1,075	47,185	13.95%
MK-04-FINKICLOUD	284	344	749	688	227	0	2,292	0.68%
NCG-INGRID-PT	667	674	712	593	714	30	3,390	1.00%
PRISMA-INFN-BARI	2,460	753	791	208	212	1	4,425	1.31%
SZTAKI	551	572	487	707	31	0	2,348	0.69%
TR-FC1-ULAKBIM	157	147	752	707	499	28	2,290	0.68%
UPV-GRYCAP	756	574	257	888	600	49	3,124	0.92%
Total	54,790	56,915	58,353	94,860	71,354	1,936	338,208	
Percentage	16.20%	16.83%	17.25%	28.05%	21.10%	0.57%		

VO Group/Role:	fedcloud.egi.eu - ALL Groups and Roles
ROC/Country:	ALL ROCs
Infrastructure	Cloud
Order by:	Number of VMs — Total number of VMs run
Period:	Start year: 2014 — Start month: 3 — End year: 2015 — End month: 2

Refresh

VO (ALL Groups and Roles) Users information

Mar 2014 - Feb 2015

The following table shows the Usage of the Users ordered by Total number of jobs and the Total Usage of the Other Users. A detailed view can be obtained by selecting an individual user

VO Users ordered by Total number of jobs

#	User ID	Jobs #	Jobs %	CPU time Hrs	CPU time %	Norm. CPU time Hrs	Norm. CPU time %	WCT Hrs	WCT %	Norm. WCT Hrs	Norm. WCT %	Monetary Cost €	Monetary Cost %	CPU Efficiency %	Avg. CPU time Hrs	Avg. WCT Hrs
1	User 1	34,158	76.8%	102,855,939	5.4%	0	0.0%	4,359,851,264	64.5%	0	0.0%	166,977,133	15.5%	2.4	3011.18	127637.78
2	User 2	2,040	4.6%	79,340,745	4.2%	0	0.0%	79,340,745	1.2%	0	0.0%	80,398,622	7.4%	100.0	38892.52	38892.52
3	User 3	1,826	4.1%	2,154,330	0.1%	0	0.0%	2,154,330	0.0%	0	0.0%	0	0.0%	100.0	1179.81	1179.81
4	User 4	1,138	2.6%	324,439,069	17.1%	0	0.0%	324,439,069	4.8%	0	0.0%	258,438,790	23.9%	100.0	285095.84	285095.84
5	User 5	545	1.2%	64,898,295	3.4%	0	0.0%	137,729,786	2.0%	0	0.0%	54,738,970	5.1%	47.1	119079.44	252715.20
6	User 6	539	1.2%	126,928,755	6.7%	0	0.0%	146,890,024	2.2%	0	0.0%	998,313	0.1%	86.4	235489.34	272523.24
7	User 7	403	0.9%	151,702,037	8.0%	0	0.0%	268,360,390	4.0%	0	0.0%	105,699,879	9.8%	56.5	376431.85	665906.67
8	User 8	293	0.7%	22,456,541	1.2%	0	0.0%	22,456,541	0.3%	0	0.0%	22,273,052	2.1%	100.0	76643.48	76643.48
9	User 9	227	0.5%	122,210,566	6.4%	0	0.0%	167,553,207	2.5%	0	0.0%	95,657,093	8.9%	72.9	538372.54	738119.85
10	User 10	193	0.4%	28,527,866	1.5%	0	0.0%	116,496,393	1.7%	0	0.0%	24,203,168	2.2%	24.4	147812.78	603608.25
11	User 11	177	0.4%	2,859,069	0.2%	0	0.0%	2,859,069	0.0%	0	0.0%	2,897,190	0.3%	100.0	16152.93	16152.93
12	User 12	120	0.3%	8,908,708	0.5%	0	0.0%	9,296,073	0.1%	0	0.0%	2,742,567	0.3%	95.8	74239.23	77467.28
13	User 13	113	0.3%	4,134,284	0.2%	0	0.0%	4,150,007	0.1%	0	0.0%	538	0.0%	99.6	36586.58	36725.73

Groups (VO) and users

- N. of VMs created for each group (VO) or user
- Total CPU times
- Total RAM, etc.

Sites

- N. of VMs
- Total CPU times
- Total RAM
- Disk size, etc.

VMs

- N. of cores
- RAM
- Disk size
- VM image, etc.

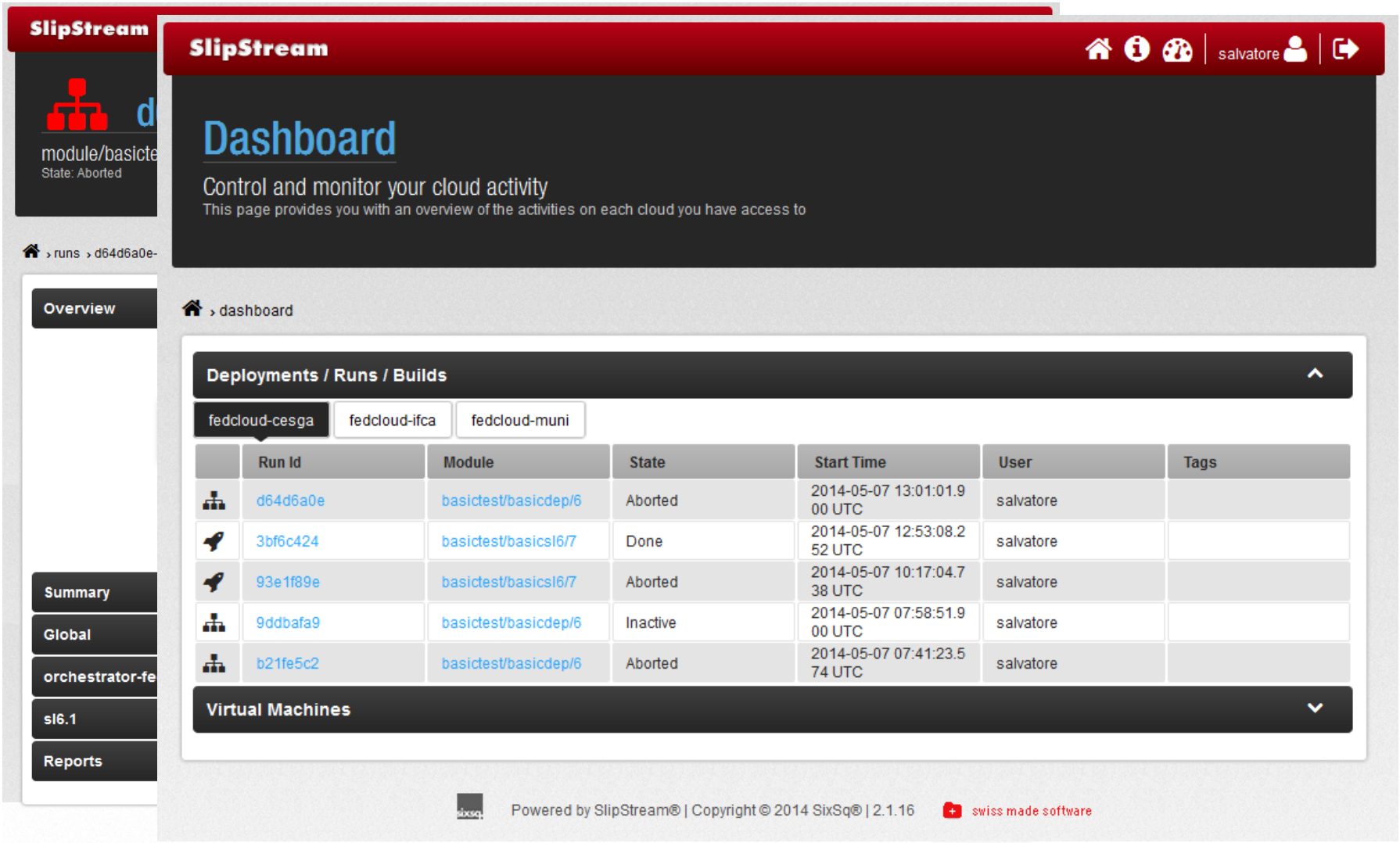
High level tools with main capabilities (PaaS, SaaS)

EGI FedCloud
High Level Tools (PaaS & SaaS)

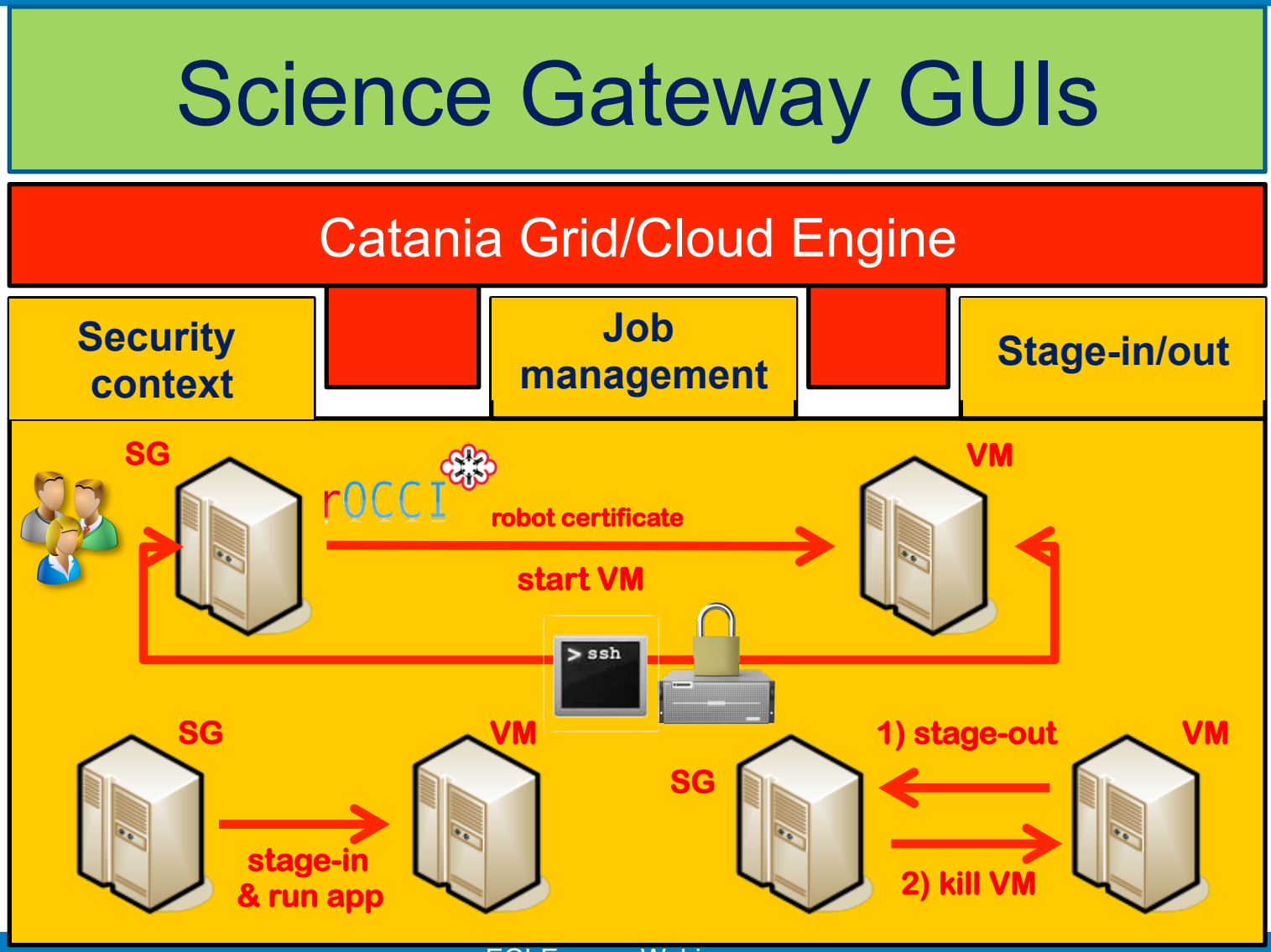
- Extend the IaaS capabilities of the EGI cloud
  - 'Alternatives' of the OCCI client and API
  - More than OCCI
- External contributions (→ support many other clouds too)

Catania Science Gateway	• SaaS • Identity Federation
Slipstream	• PaaS for automating deployments • Helix Nebula
COMPSs	• Programming framework for auto-parallelisation
VMDIRAC	• Abstraction on top of various HPC/HTC/cloud
WS-PGRADE	• Workflow development and enactment
Vcycle	• VM lifecycle manager

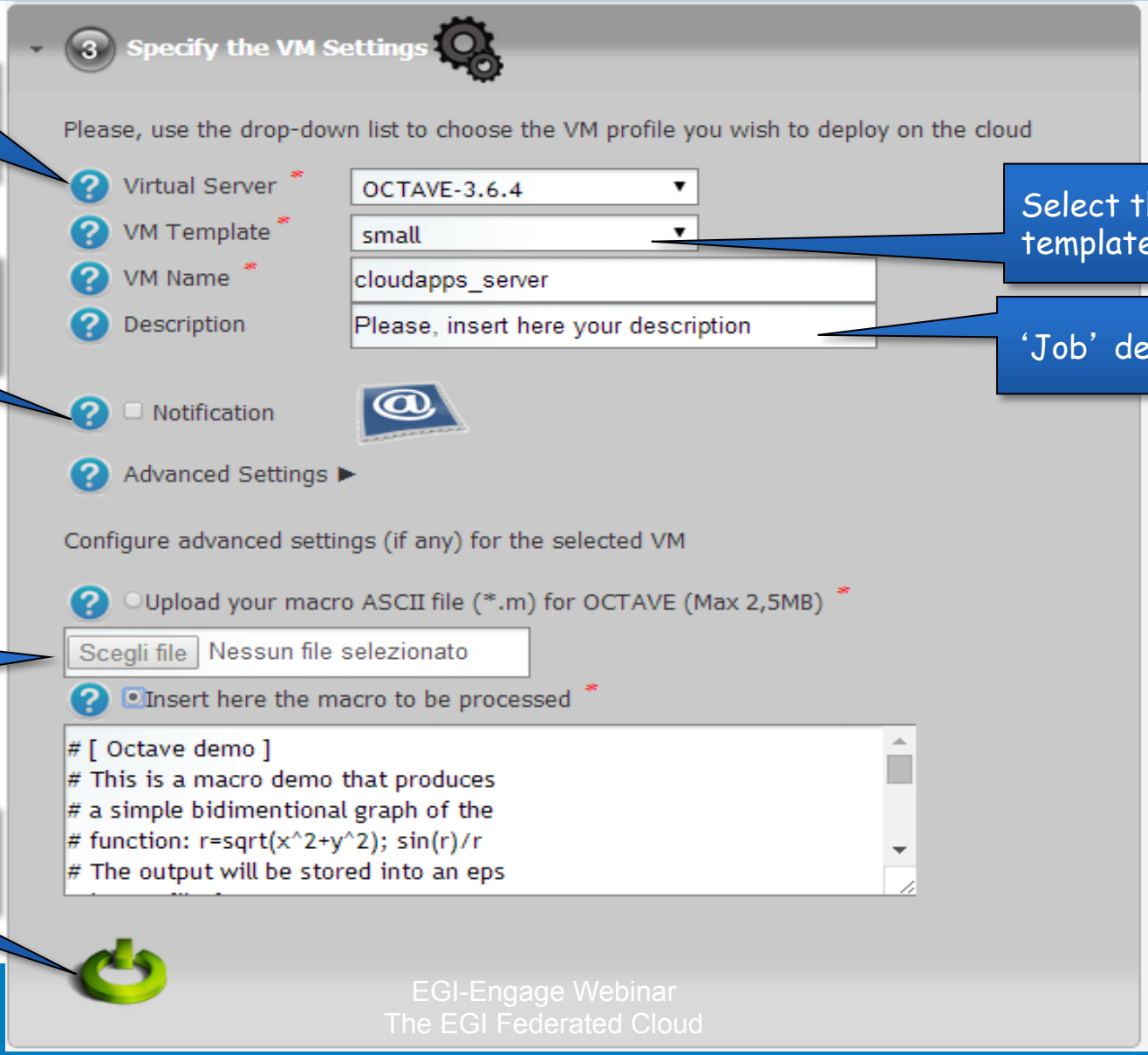
Slipstream
PaaS for automatic deployment



Catania Science Gateway Framework SaaS & Identity Federation



Catania Science Gateway Framework



Select the VM profile

Enable e-mail notification

Octave-specific settings

Start VM

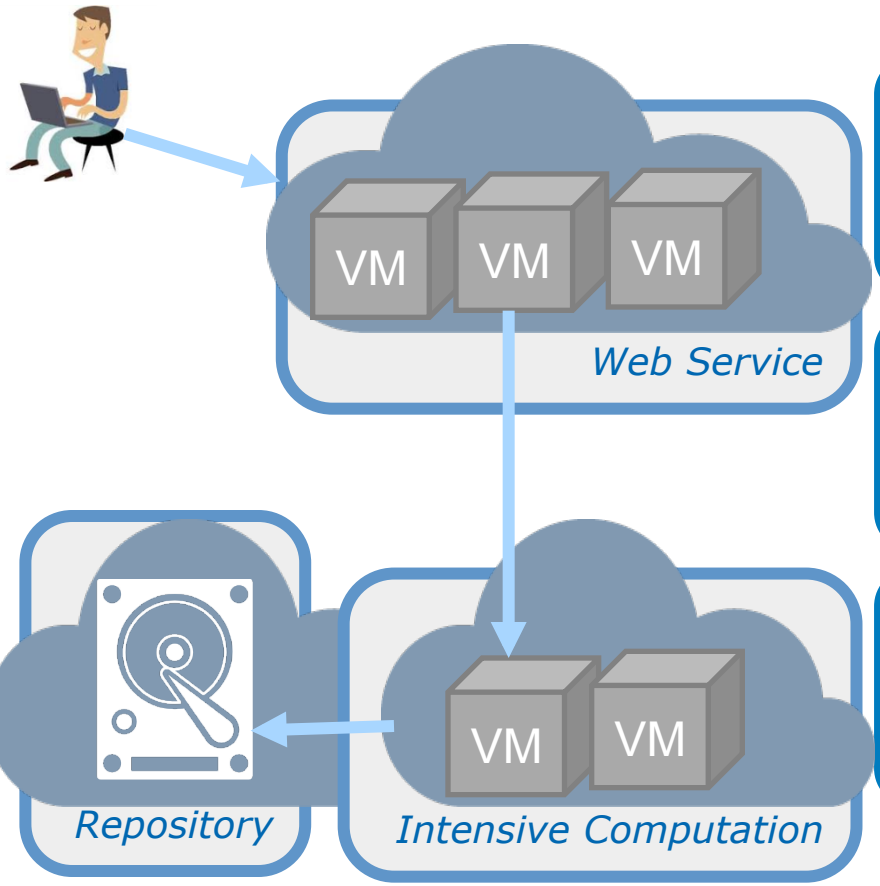
Select the VM template

'Job' description

Typical setups and use cases Examples

Usage Models

The EGI Federated Cloud is able to support different use case models

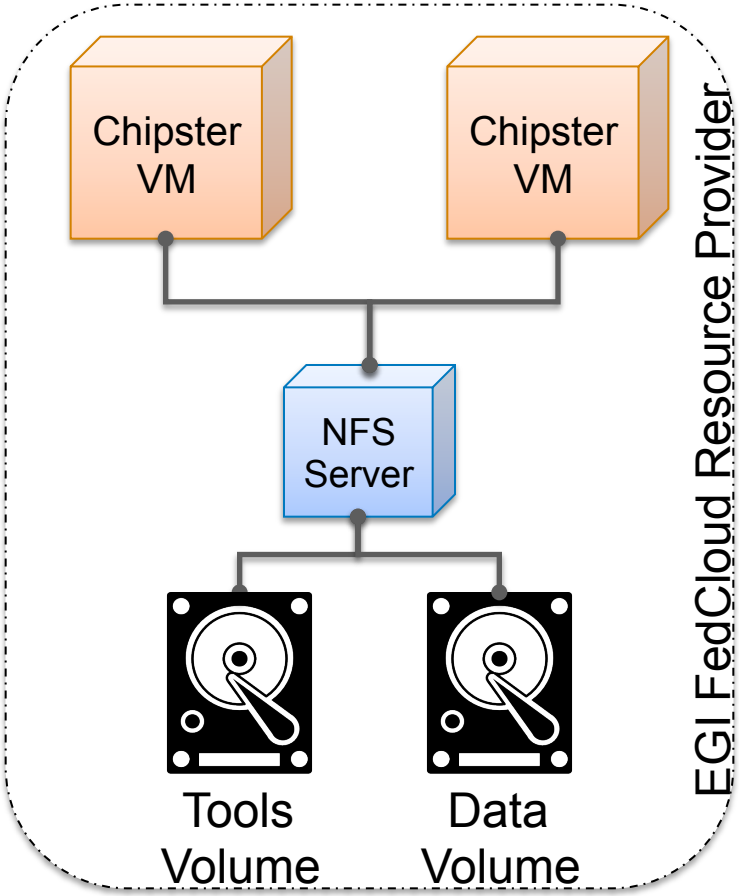


You can combine them!

Compute & Data intensive	• Heavy computation • Large Memory
Datasets Repository	• Store & manage large datasets
Web Services	• Hosting web-sites, web services, portal, etc. • Exploit cloud features as elasticity

Chipster

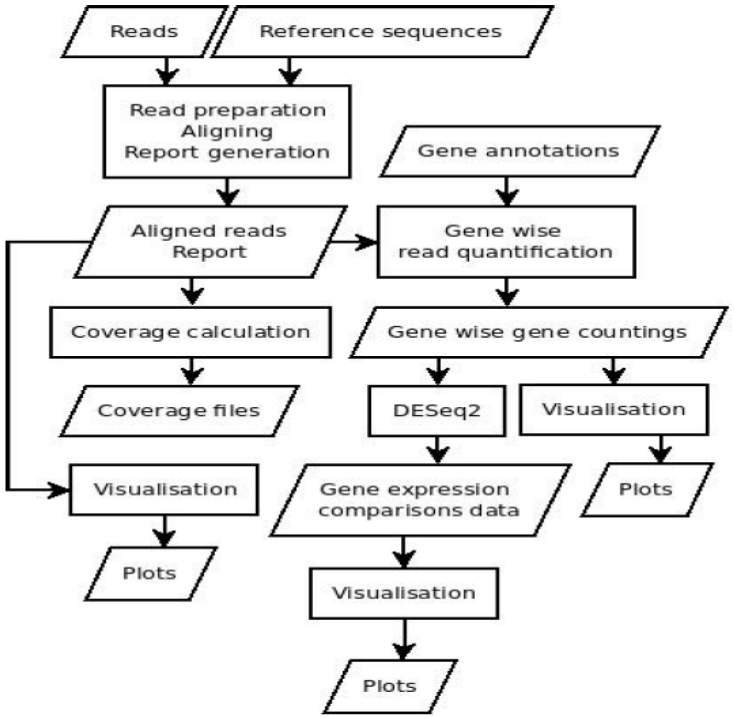
Analysis software contains over 300 analysis tools for NGS, microarray, proteomics and sequence data.



Usage Model	• **Web service** • **Heavy computation and large memory** • **Manage large datasets**
Scientific Disciplines	• **Bioinformatics**
Deployment in the FedCloud	• **Complex deployment through contextualisation** • **shared block storage exported as NFS up to 1 TB**

READemption

Pipeline for the computational evaluation of RNA-Seq. data



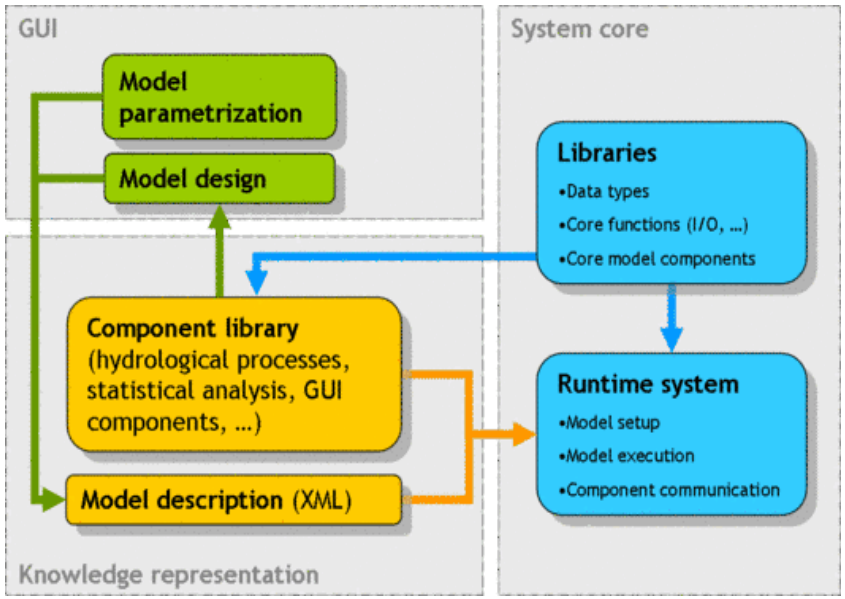
Source: Konrad U. Förstner

UNIVERSITÄT WÜRZBURG

Usage Model	• **Heavy computation** • **Large Memory**
Scientific Disciplines	• **Bioinformatics**
Deployment in the FedCloud	• **VMs with 24 cores, 128 GB of RAM** • **Block storage up to 3 TB**

JAMS
Jena Adaptable Modelling System

Platform for process-based hydrological model development

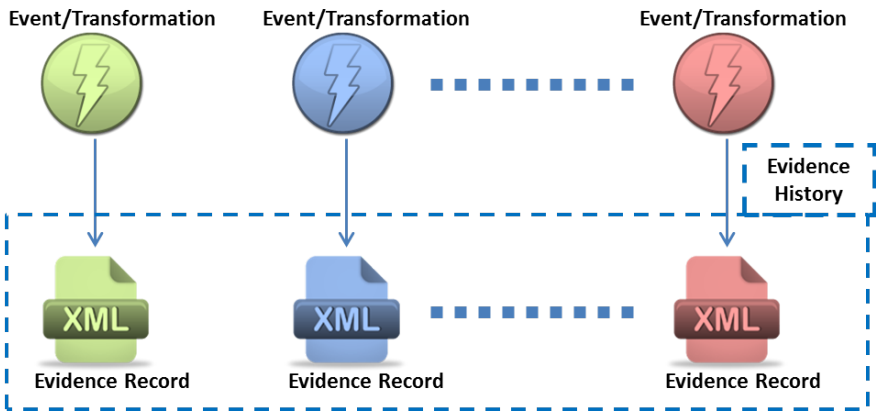


Usage Model	• **Web Service** • **Heavy computation** • **Large Memory**
Scientific Disciplines	• **Hydrology**
Deployment in the FedCloud	• **VMs with > 8 cores, > 16 GB of RAM** • **Image available in the marketplace**

HAPPI

Supports the archive manager and curator to capture and manage part of the Preservation Descriptive Information

SCIDIP-ES
HAPPI Toolkit



Usage Model	• **Web Service**
Scientific Disciplines	• **Digital Archives**
Deployment in the FedCloud	• **VMs with 2 cores, 4 GB of RAM** • **Block Storage**

Plans & next steps

Plans and next steps

The EGI Federated Cloud Task Force continue to work to evolve and improve the Cloud Infrastructure

Compute	• OCCI 1.2 • MS Azure • Native CMF interfaces
Storage	• OCCI advanced for Block Storage • CDMI for OpenStack • Native CMF interfaces
High level tools	• More PaaS and SaaS integrated in the FedCloud
Network	• Advanced network management

User Support

Technical Consultancy

Dedicated technical consultancy for each community (support@egi.eu)

F2F/Web Meetings	• Initial requirements • Allocation of technical experts • Milestones
Doc	• Step by step guides • Tutorials • Examples
Continuous tracking and support	• Technical integration support from EGI/NGI team • Periodic meetings
Basic VM Images	• Main OS available • Secure endorsed • Contextualisation
Incubator VO	• Resources for application prototyping • Enabled in all sites • Up to 6 months (renew)
Migration into production	• Resource providers commitment • Support to create a production VO

EGI Federated Cloud User Support doc. entry page:

- https://wiki.egi.eu/wiki/Federated_Cloud_user_support

Federated Cloud user support

Main | Roadmap and Innovation | Technology | **For Users** | For Resource Providers | Media

Contents [hide]

Technical support is available via the EGI.eu Support Team

Users of the EGI Federated Cloud are scientists working in many fields, who can benefit of a flexible environment for running their workloads. Also, the EGI cloud is suitable to projects aiming to provide services and platforms to the scientific community.

Concept [edit]

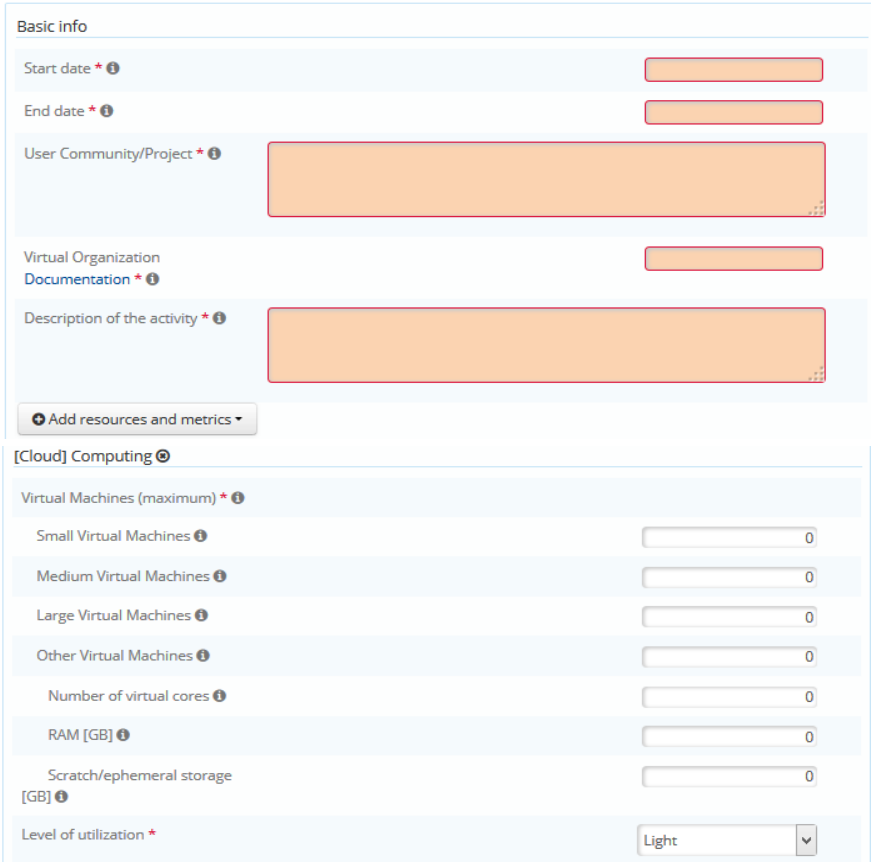
The EGI Federated Cloud is a seamless grid of academic private clouds and virtualised resources built around open standards and focusing on the requirements of the scientific community. The result is a new type of research e-infrastructure based on the mature federated operations services that make EGI a reliable resource for science. When using EGI Federated Cloud resources, researchers and research communities can count on:

- Total control over deployed applications
- Elastic resource consumption based on real needs
- Immediately processed workloads – no more waiting time
- An extended e-Infrastructure across resource providers in Europe
- Service performance scaled with elastic resource consumption
- Single sign-on at multiple, independent providers

Next steps to become an user

Next steps to become a user

Different way to access the EGI Federated Cloud



e-Grant Resource Allocation

User want to access resources

- Join one of the existing groups/VOs
- Find relevant VO in the Ops Portal

Allocation on the EGI RPs

- Dedicated resources
- Submit request via e-Grant: https://e-grant.egi.eu

Join FedCloud as RPs

- Turn your site into a community federate cloud
- Contact EGI Operations: operations@egi.eu

References

EGI Federated Cloud resources

- Wiki site: http://go.egi.eu/fedcloud
- User support: https://wiki.egi.eu/wiki/Federated_Cloud_user_support
- User support e-mail: support@egi.eu
- Federated Cloud Communities: https://wiki.egi.eu/wiki/Federated_Cloud_Communities

Related Standards:

- OCCI: http://occi-wg.org
- CDMI: http://cdmi.sniacloud.com/

Questions ?