A European Cloud federation

EGI's Cloud strategy for 2020

Outline

Brief introduction to EGI

EGI Federated Cloud

From Pilot to Production

EGI's Cloud strategy for H2020

Data, Sharing, Innovation: Safe infrastructures for the ERA

Outline

Brief introduction to EGI

EGI Federated Cloud

From Pilot to Production

EGI's Cloud strategy for H2020

Data, Sharing, Innovation: Safe infrastructures for the ERA

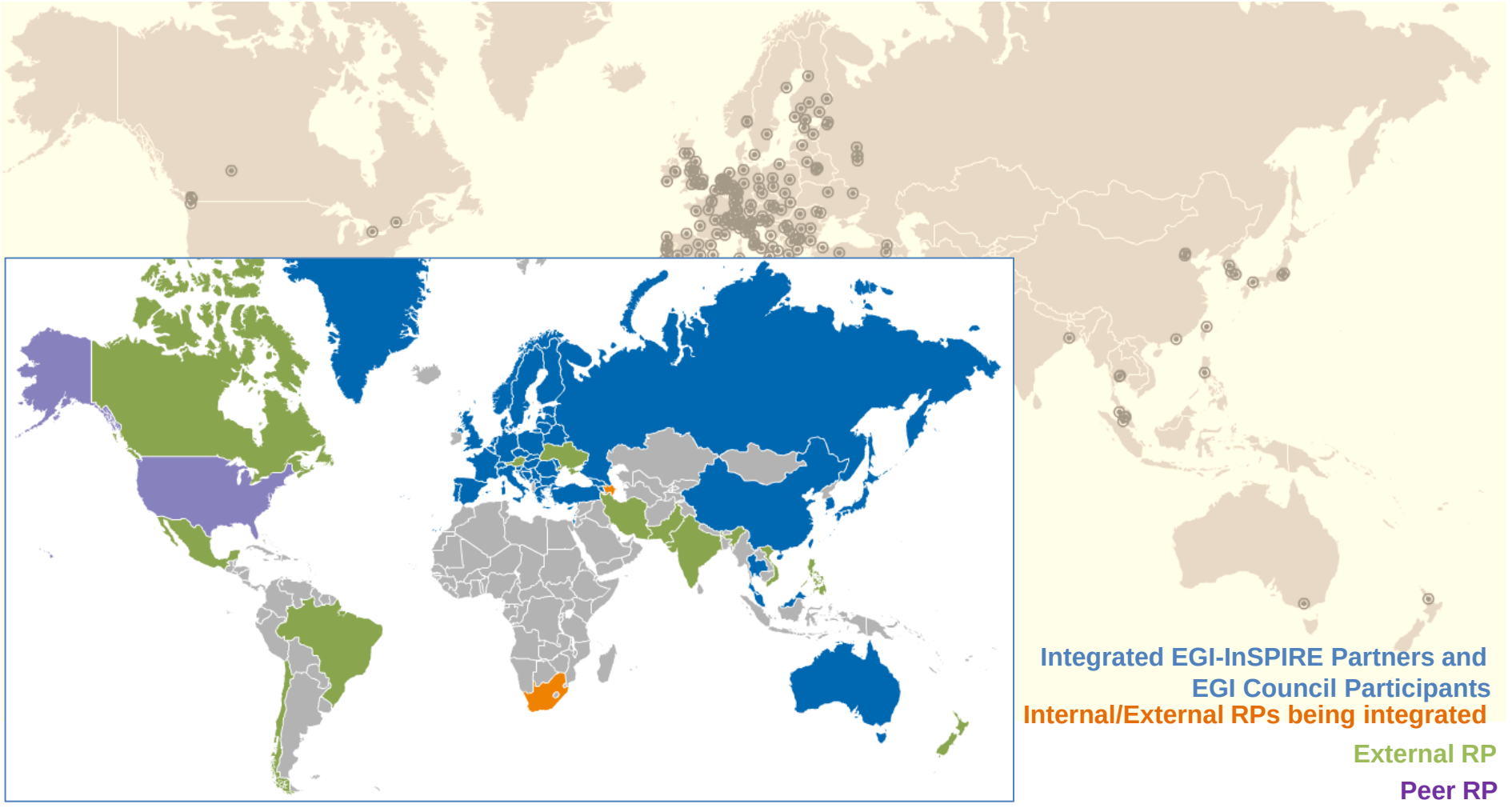
European Grid Infrastructure (EGI)

337 Resource Centres in 34 National Grid Initiatives/EIROs
430,000 logical CPU cores
190 PB disk, 180 PB tape
1.2 M job/day, EGI-InSPIRE PY3: +44.7% increase of CPU wall clock time used (HS-06 h)
53 High activity projects supported by EGI

EGI-InSPIRE EC Project (2010-2014) supporting

- EGI and NGI operations
- Outreach and policy development
- Software validation and verification
- Federated cloud

European Grid Infrastructure 2013



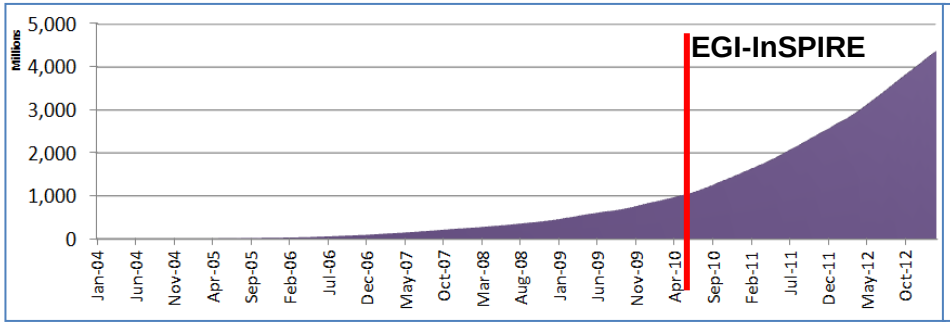
Data management and compute

- Data can be stored on different storage systems
  - Common interface for storage access: SRM, gridFTP, WEBDAV
- Data can be distributed and replicated among different locations
  - File replica catalog
  - Metadata catalog
  - File transfer services
- Computing resources are usually available through interfaces called Computing Elements
  - Data input and output are usually staged-in and staged-out to storage services

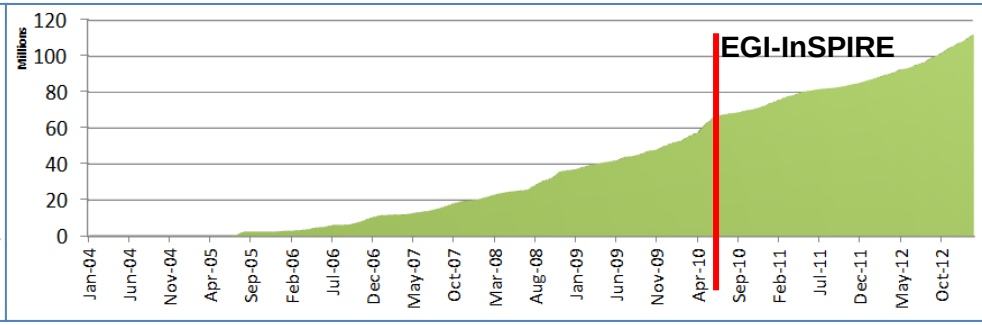
10 years in production

Jan 2004 – Sep 2013		Value
CPU wall time consumed (Jan 2004 – April 2013)	Billion hours	5.7 (CPU wall time) 44.8 (normalized HEP-SPC06)
Yearly estimated overall resource utilization		82%

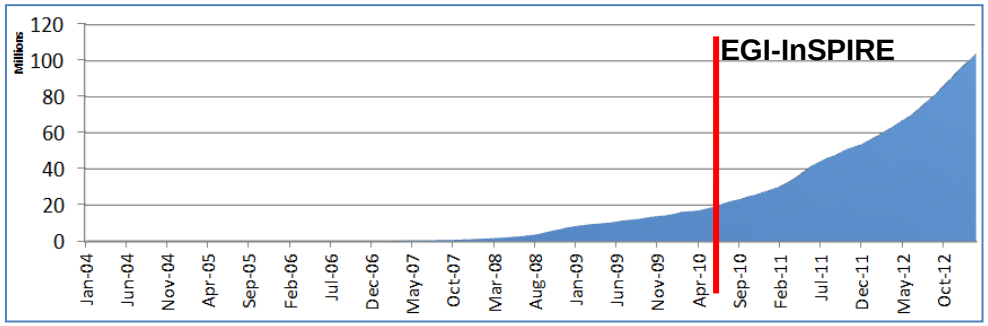
HEP: 4.5 Billion hours



LS: 116.7 Million hours



AA: 110.3 Million hours



Federated Cloud

- New services to meet user demand
  - Use case driven
- Proof of concepts being successfully demonstrated (http://go.egi.eu/PoC)
  - Hosting of services for data dissemination (SaaS) – ENVRI, EISCAT_3D
  - Digital Libraries and Digital Preservation services for memory institutions and human science - DCH-RP
  - Integrated Cloud Iaas-SaaS services to avoid large data transfers (ESA)
  - Virtual laboratories (PaaS and IaaS) – BioVel, LifeWatch
  - Hadoop clusters on demand (PaaS) – Peachnote, BioVel
  - …
- Definition of business models
- Hybrid private-public provisioning - EGI Federated Cloud is part of Helix Nebula

Federated Cloud

- Define the Cloud Federation layer
  - Promote adoption of common interfaces (OCCI,CDMI)
  - Investigate the capabilities for the federation of clouds
- Integrate the cloud services with the EGI core platform
  - auth and authz, accounting and monitoring, service registry, service
- 70 partners from 40 institutions and 13 countries
  - 4 resource centres, 1000 cores, 16 TB Storage
  - Capacity building in 2014
- Production in Spring 2014

Outline

Brief introduction to EGI

EGI Federated Cloud

From Pilot to Production

EGI's Cloud strategy for H2020

Data, Sharing, Innovation: Safe infrastructures for the ERA

EGI Vision

To support the digital European Research Area through a pan-European research infrastructure based on an open federation of reliable services that provide uniform access to national computing, storage and data resources.

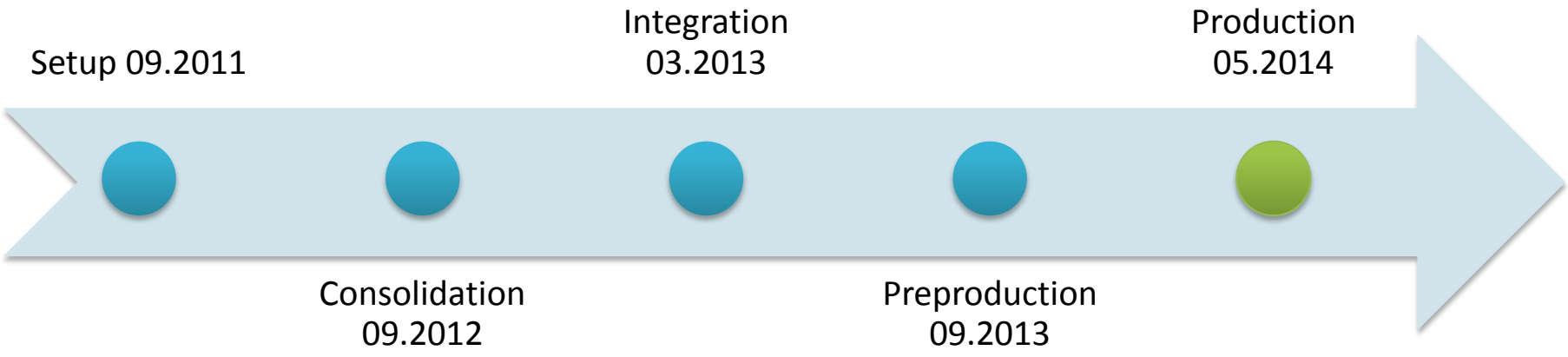
EGI federated Cloud vision for H2020:

10M cores Cloud compute
1 EB Cloud storage

From Pilot to Production

- **Objectives:**
  - Identify and investigate the capabilities needed to federate private clouds
  - Identify the technical solutions, deploy proof of concepts in a pre production testbed, test the solutions with real use cases

- **Capabilities**
  1. Manage VM instances
  2. Data access/transfer interface
  3. Cloud service information federation
  4. Resource consumption management
  5. Cloud service availability
  6. Notification & Automation
  7. Federated AAI
  8. VM Image Management
  9. Brokering
  10. Contextualisation



Participation – Nov 2013

Cyfronet | FZJ | OeRC | EGI.eu | CESNET | GWDG | INFN-BARI

CNRS | IN2P3

KTH | Masaryk

FCTSG | INFN-CNAF

CETA | CESGA

IGI | SARA

RADICAL | IFCA

STFC | SZTAKI

BSC | GRNET

Imperial | DANTE

LMU | IPHC | IISAS | SixSq | 100%IT | CSC | IFAE | SRCE

esa

we-nmr

Members

- ~70 individuals
- ~35 institutions
- >13 countries

Stakeholders

- 23 Resource Providers
  - 10 production
- 10 Technology Providers
- 8 User Communities
- 4 Liaisons

Technologies

- OpenNebula.
- StratusLab.
- OpenStack.
- Synnefo.
- WNoDeS.
- PERUN
- SlipStream

BioVeL

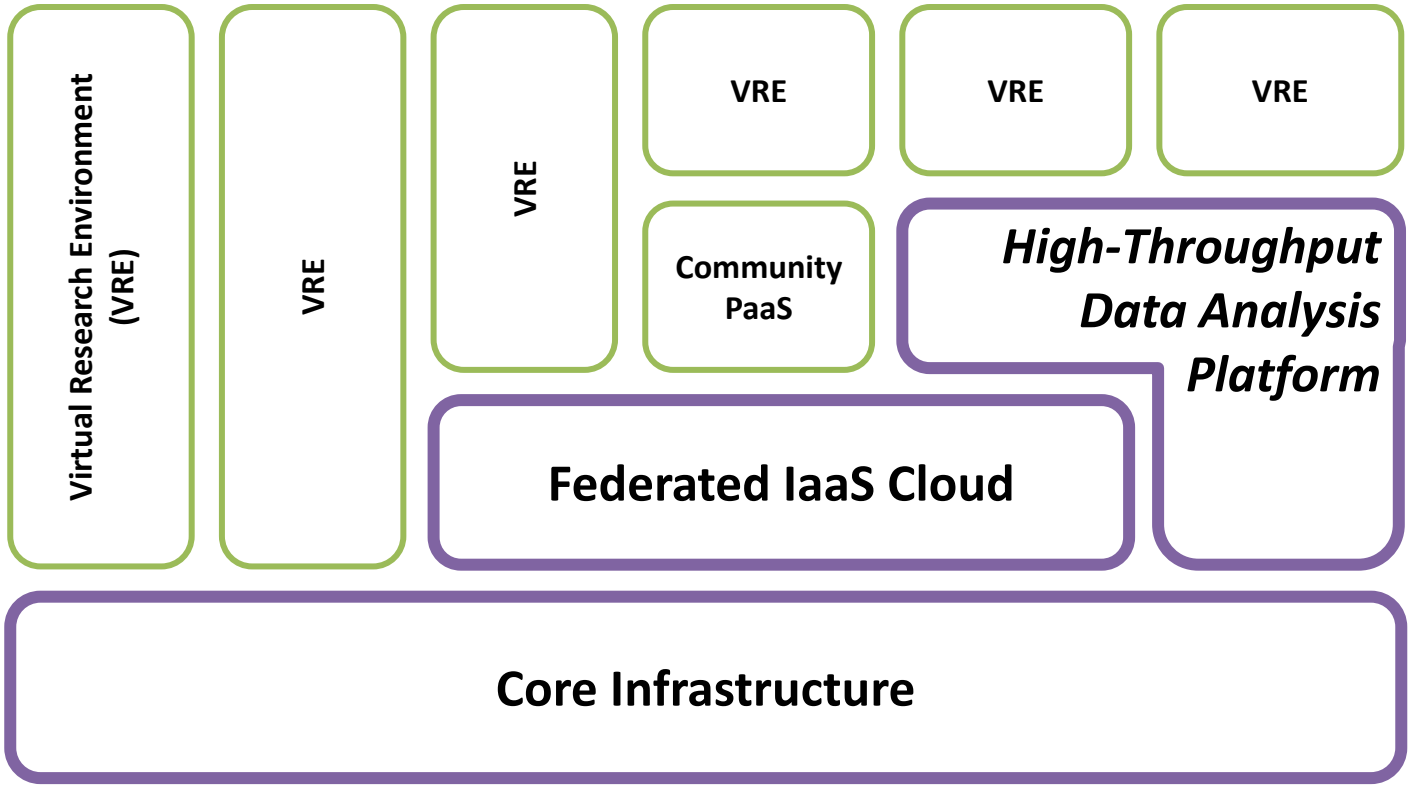
SCI-BUS

Catania Science Gateway Framework

Gaia

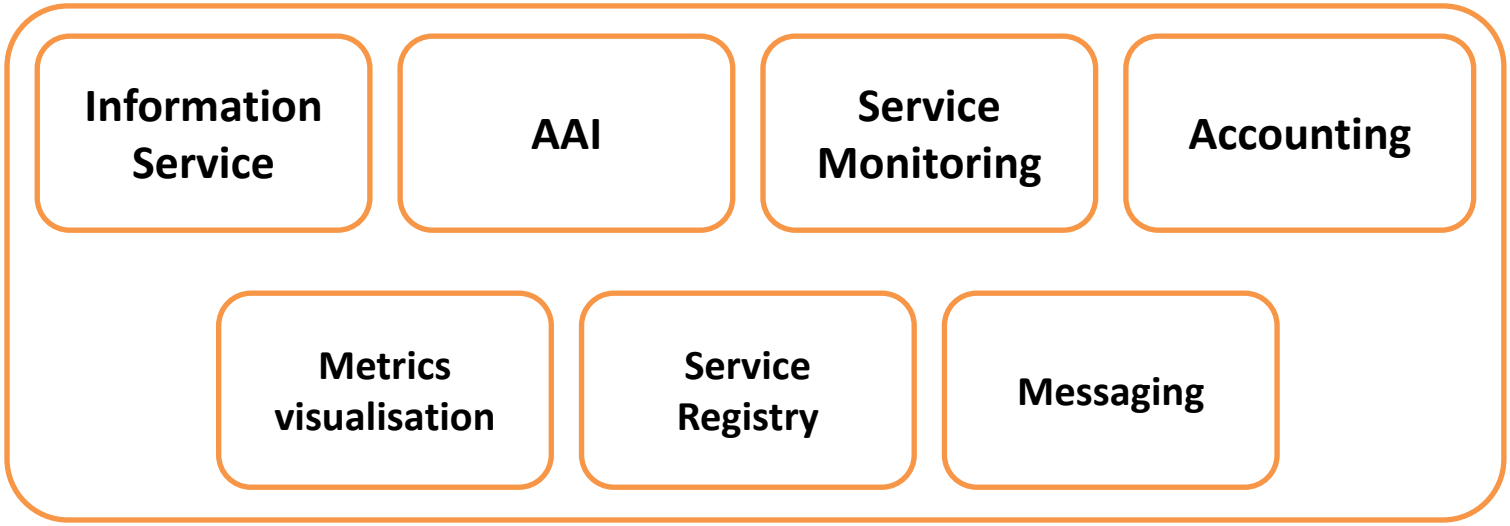
CLARIN
Common Language Resources and Technology Infrastructure

EGI Platform architecture



EGI Core Infrastructure

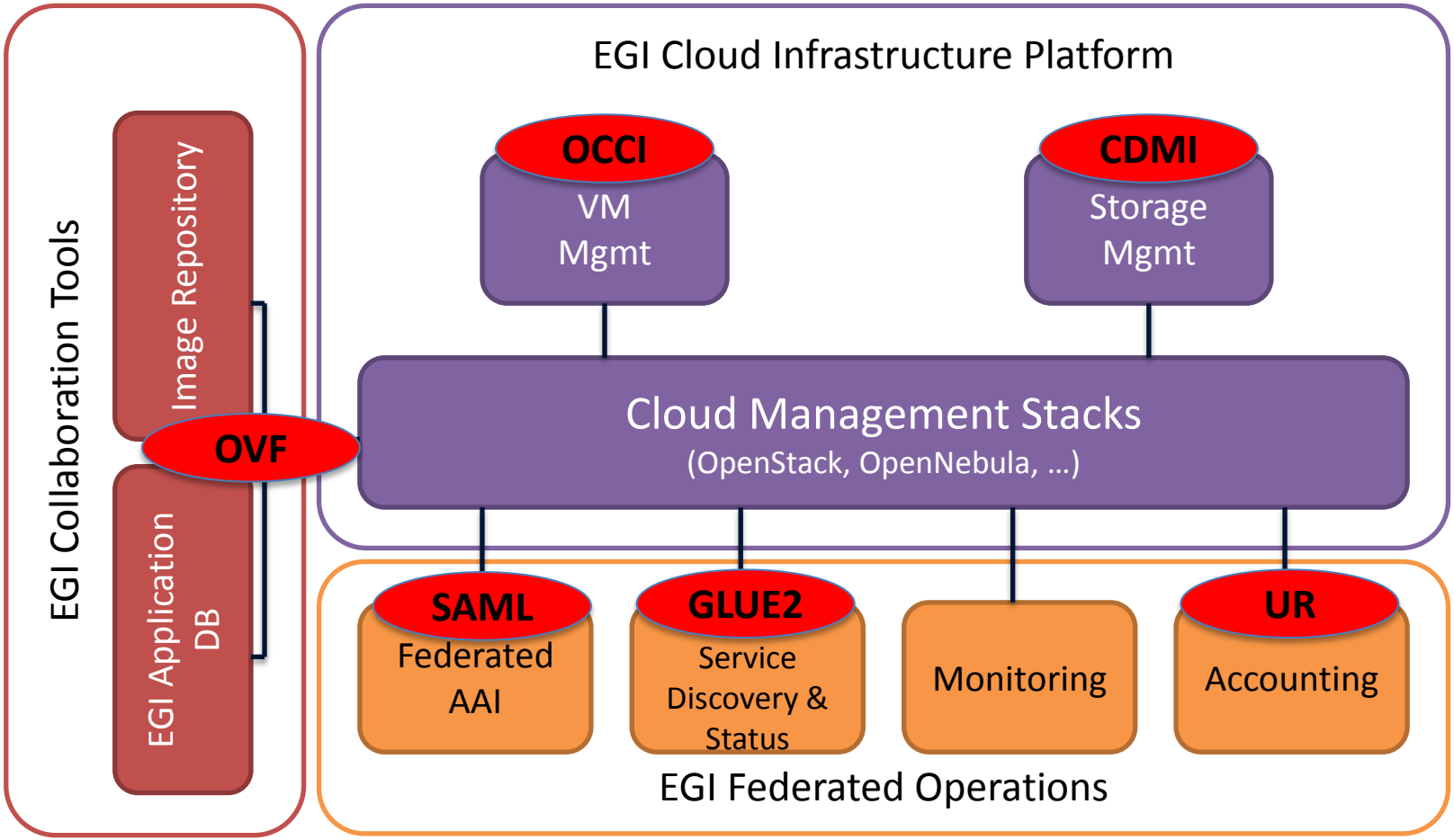
Services that federate and integrate the functional services deployed in the **production infrastructure**



For e-Infrastructures & Research Infrastructures

EGI's Cloud Infrastructure

Enable an open ecosystem of services



Open Standards approach

- **OCCI:** Open Cloud Computing Interface
  - RESTFul API to manage virtual machine in the Cloud
  - From OGF
- **CDMI:** Cloud Data Management Interface
  - RESTFul API to create, retrieve, update and delete data elements from the Cloud
  - From SNIA
- **OVF:** Open Virtualization Format
  - Standard format for packaging and distributing virtual machines
  - From DMTF
- **SAML:** Security Assertion Markup Language
  - Standard for conveying identity tokens and attributes
  - From OASIS
- **GLUE2:**
  - Standard to describe and publish information on structured distributed infrastructures
  - From OGF
- **UR:** Usage Records, v2
  - Standard to express, collect and aggregate usage accounting records
  - From OGF

Cloud service provisions

Resource Providers

- IaaS Cloud compute
- IaaS Cloud storage

EGI.eu

- Service availability monitoring
- Resource usage accounting
- Federated AAI coordination
- Service discovery (GOCDB & BDII)
- Application/VM database
- VM Image distribution network

Other providers

- Meta service-brokering
- User community management

Outline

Brief introduction to EGI

EGI Federated Cloud

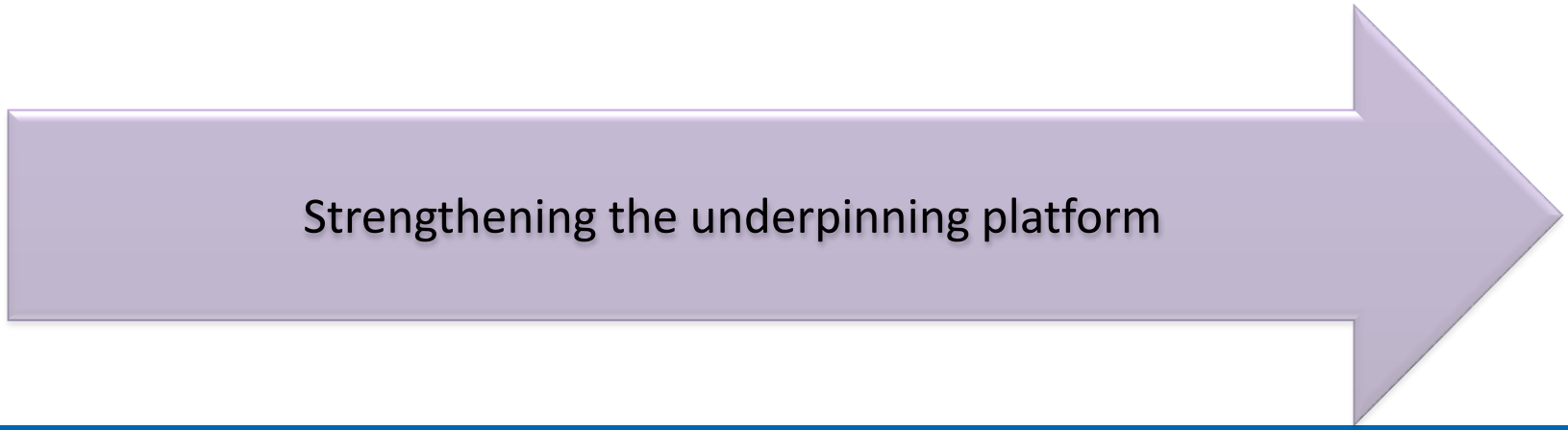
From Pilot to Production

EGI's Cloud strategy for H2020

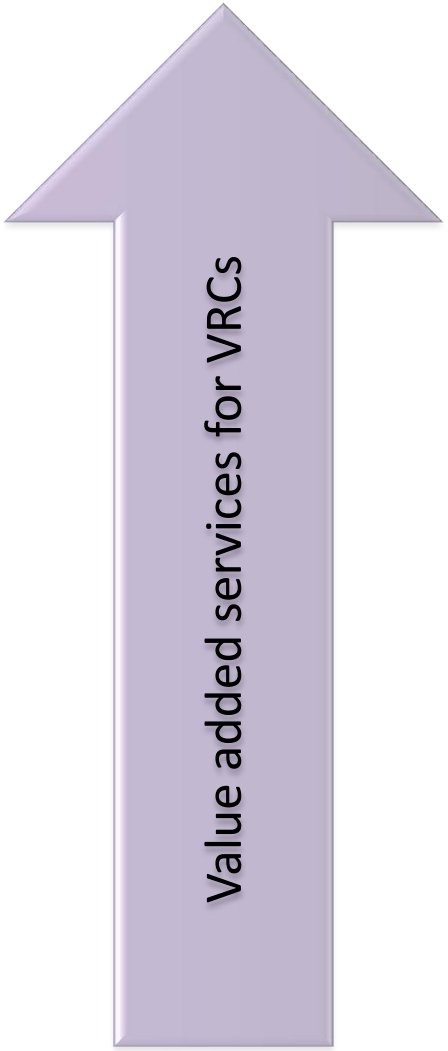
Data, Sharing, Innovation: Safe infrastructures for the ERA

Technology Evolution

- Broader support for open standards in Cloud management frameworks
- Federated Network as a Service
- Messaging network as a Service
- Improve Application Database to App Marketplace
  - Stable market for cloud services, cloud products and applications between providers and consumers from the academia, public sector and enterprise



Technology Innovation



- Use federated IaaS Cloud as EGIs backbone
- Open standards for open platforms
- Innovate and deploy PaaS & SaaS on top
- Engaging in PoCs & EC projects with EGI customers

Use cases 1

Platform	Description	Key services	Use cases
Data dissemination/Open data	Cloud storage provides an infrastructure to collect and disseminate scientific data. Data intake is curated, data access is inherently open (anonymous) or tracked (social identity?)	• SaaS catalogues tailored to user communities (ex. EO, Biology, etc…) • Custom Data ACL • Federated AAI for data access • Monitoring of data usage	ENVRI GeoCatalogue EISCAT 3D Peachnote
Secure Storage	Stored data is protected even from RP access (e.g. through encryption). Safe for storing (personal) confidential data.	• Keys management • Server-side encryption • Secure data deletion • Data sharing with ACL	Arvados
Location-based computing	Input data is stored on the cloud, processing is on the same cloud or in federated data-centers to have high data access throughput with own applications; “	• Integrated Cloud Iaas-SaaS services to avoid large data transfer • IaaS to provide processing flexibility • Input data stored openly to attract different communities	ESA G-POD Arvados
Virtual Laboratories	Tools to customize and manage virtual laboratories for different communities. Laboratories have shared tools to access data from different sources. User communities manages the instruments, EGI operates the underlying infrastructure and provides the generic tools to access storage (cloud storage, etc...)	• Common shared tools to access data from different resources • Tools to ease laboratories setup • Underlying IaaS and SaaS to support running of the laboratories and store user data storage (with easy scalability) • Simple user interface to request laboratories access	BioVeL virtual e-laboratory LifeWatch virtual laboratory

Use cases 2

Platform	Description	Key services	Use cases
Data preservation	Data long-term preservation. To provide consolidation, persistency, integrity, redundancy, and usability over long periods of time.	• Data consolidation (ensure all the data is harmonized in terms of format, nomenclature, access, etc…) • Data integrity and redundancy (ensure no loss of data) • Data access preservation (ensuring software to read and analyze the data is maintained)	EISCAT 3D DCH-RP HEP
Platform-as-a-Service	Pre configured processing facilities with integrated access to data, running on top of cloud IaaS and SaaS solutions. Possible PaaS services are: • Grid Computing • Hadoop Clusters on demand • Generic High Avaliability service	• Scalability (adapt to the workload) • High Availability (resources are always available) • Resource sharing (different services share the same underlying physical resources)	Peachnote BioVeL HA

Models for Sustainability

- **Pricing**
  - **Free at the point of use**
    - accounted resource consumption, best effort support, no direct reimbursement of accrued costs
  - **Try before you buy**
    - Limited access to free resources, expires after defined time.
    - Facilitates conversion rate from user to customer
  - **Pay-as-you-go**
    - Fixed price per consumption unit, fees directly correlate with consumed resources
  - **Wholesale resource guarantee**
    - Soft-quota, 2-tier cost plan: Resources within quota paid as you go on wholesale price, overdraft within limits costed at pay-as-you-go plan
  - **Reserved resources**
    - Exclusive resource reservation up to agreed limits. Overdraft not allowed, fixed payments

- **Service Levels Agreements**
  - **Best effort**
    - Available at all Cloud service providers
    - “Free at the point of use” billing plan and higher
  - **Basic**
    - Available on all Cloud service providers
    - “Pay-as-you-go” plan or higher
  - **Tailored to community requirements**
    - Advanced SLAs subject to research and innovation
    - Applicable to subset of providers

EGI Cloud service tiers

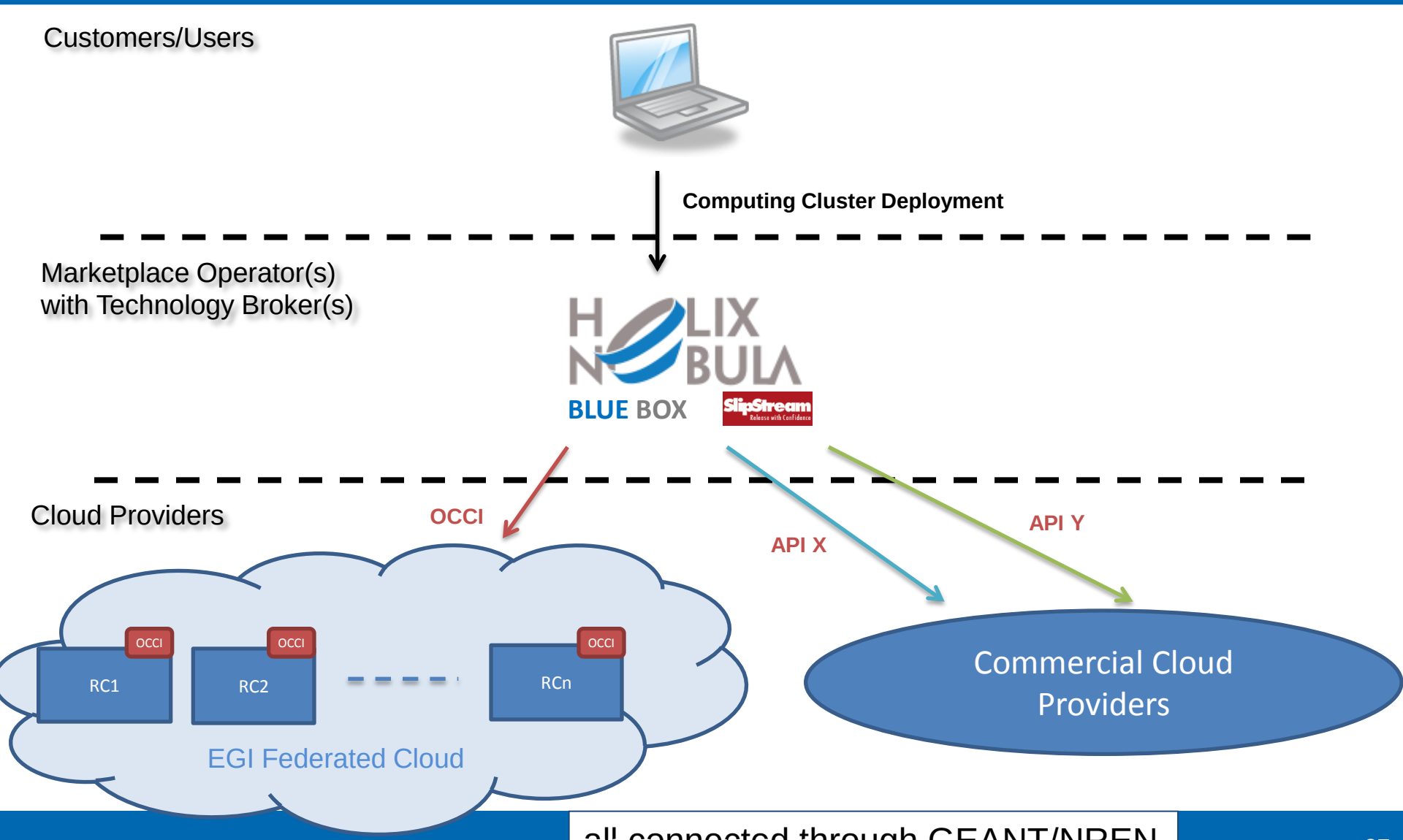
1. **Reliable IaaS services**
   - Expose federated IaaS Clouds to customers
   - User driven choice of provider depending on high level criteria
2. **General purpose platform services**
   - Expose platform services as individual service offerings
   - Customer mixes and matches according to need
3. **Platform as a Service**
   - Consistent platforms comprising of individual platform services
   - Platform services integrated/configured in meaningful way
   - Targets customer segments (e.g. CH) or usage scenarios (data preservation)
4. **Zero ICT infrastructures**
   - Specific customer infrastructures delivered as SaaS
   - Extends beyond ICT into complete VRE supply (e.g. lab facility management)

Marketplace(s) for Cloud Services

Standardisation of Cloud interfaces empowers freedom of choice for customers, better competition among providers and emergence of cloud marketplaces

- **EGI Marketplace**
  - Rich and diverse set of individual offerings
  - Application/VM Image repository
  - User community choice of provider at all levels
  - Academic and commercial resource & service providers
- **Helix Nebula Marketplace**
  - EGI Federated Cloud as service provider
  - “Hidden” behind Slipstream/CGI broker
- **EU Cloud for Europe marketplace**
  - Integrate EGI’s solution and service portfolio
  - Provide services for academia, government and business

Helix Nebula Marketplace



all connected through GEANT/NREN

Value proposition

- **Open standards, open technologies**
  - Use of Open Standards is key to the establishment of an effective, fair and transparent cloud market in Europe
  - Open Source components raise the barrier for hidden backdoors, thus lead to more trusted services
- **Firmly rooted in Europe**
  - Strong public sector involvement through NRENs, NGIs, EIROs are EGI's members
  - European commercial Cloud resource providers, including SME
- **A single cross-border market**
  - Reaching out for research, government & business sectors
  - Level playing field for innovation and services on multi-service tiers

Fit for the Digital Agenda

Action 1: Interoperable, federated IaaS Cloud infrastructure

Action 2: Public sector Cloud federation for Cloud for Europe marketplace

Action 3: Engage in and lead H2020 e-Infrastructure, ICT LEIT & CEF projects to boost Cloud service market

Action 4: Transparent, accounting, billing & SLAs; common T&Cs

Action 5: EGI maintains close relationships with policy makers through strategic partnerships (e.g. SIENA, CloudWATCH projects, e-IRG) and concertation meetings (e.g. ICT, Cloudscape series)

Horizon 2020 – LEIT

- **ICT 7 – 2014**: Advanced cloud infrastructures and services
  - Research & Innovation action
    - High performance heterogeneous cloud infrastructures
    - Federated Cloud networking – SDN collaboration with GEANT
    - Dynamic configuration, automated provisioning and orchestration of cloud resources
    - Automated discovery and composition of services
    - Cloud security

Conclusion

- **EGI:** Community of resource providers with long-term tradition in providing federated ICT services for research
- **EGI Federated Cloud:** Paving the way for a federated cloud in Europe
  - Open standards, open technology
  - Open membership, open processes
- Driver for innovation in Europe

Thank you!

Questions?