Helix Nebula, the Science Cloud

e-IRG strategy workshop

11 & 12 June 2012

Maryline Lengert, ESA

From Requirement Collection to Strategic Plan & Proof of Concept

HELIX NEBULA THESCIENCECLOUD

- **End 2010**: ESA started collecting **Cloud Computing requirements** for Space/Science activities
- **January 2011**: CERN and EMBL joined
- **Spring 2011**: ESA started collecting **European Industry requirements** from Space/Science users
- **June 2011**: **Strategic Plan** endorsed at ESRIN Workshop
- **October 2011**: 3 **Pilot projects presented to industry and selected** at Heidelberg Workshop
- **January 2012**: pilot projects started with Proof of Concept, governance in place
- **February 2012**: **NDA** signed by all (18+5) members of the consortium

http://cdsweb.cern.ch/record/1374172/files/CERN-OPEN-2011-036.pdf

Strategic Plan for a Scientific Cloud Computing infrastructure for Europe

HELIX NEBULA THESCIENCECLOUD

CERN-OPEN-2011-036
08/08/2011

1.1

8th August 2011

- **Establish a sustainable multi-tenant cloud computing infrastructure in Europe**
- **Initially based on the needs for the European Research Area & space agencies**
- **Based on commercial services from multiple IT industry providers**
- **Adhere to internationally recognised policies and quality standards**
- **Governance structure involving all stakeholders**

Dr. Maryline Lengert
ESA - European Space Agency
Senior Advisor
Maryline.Lengert@esa.int
Tel +39 06 941 80430

Dr. Bob Jones
CERN – European Organization for Nuclear Research
IT department
Bob.Jones@cern.ch
Tel. +41 22 767 14 82

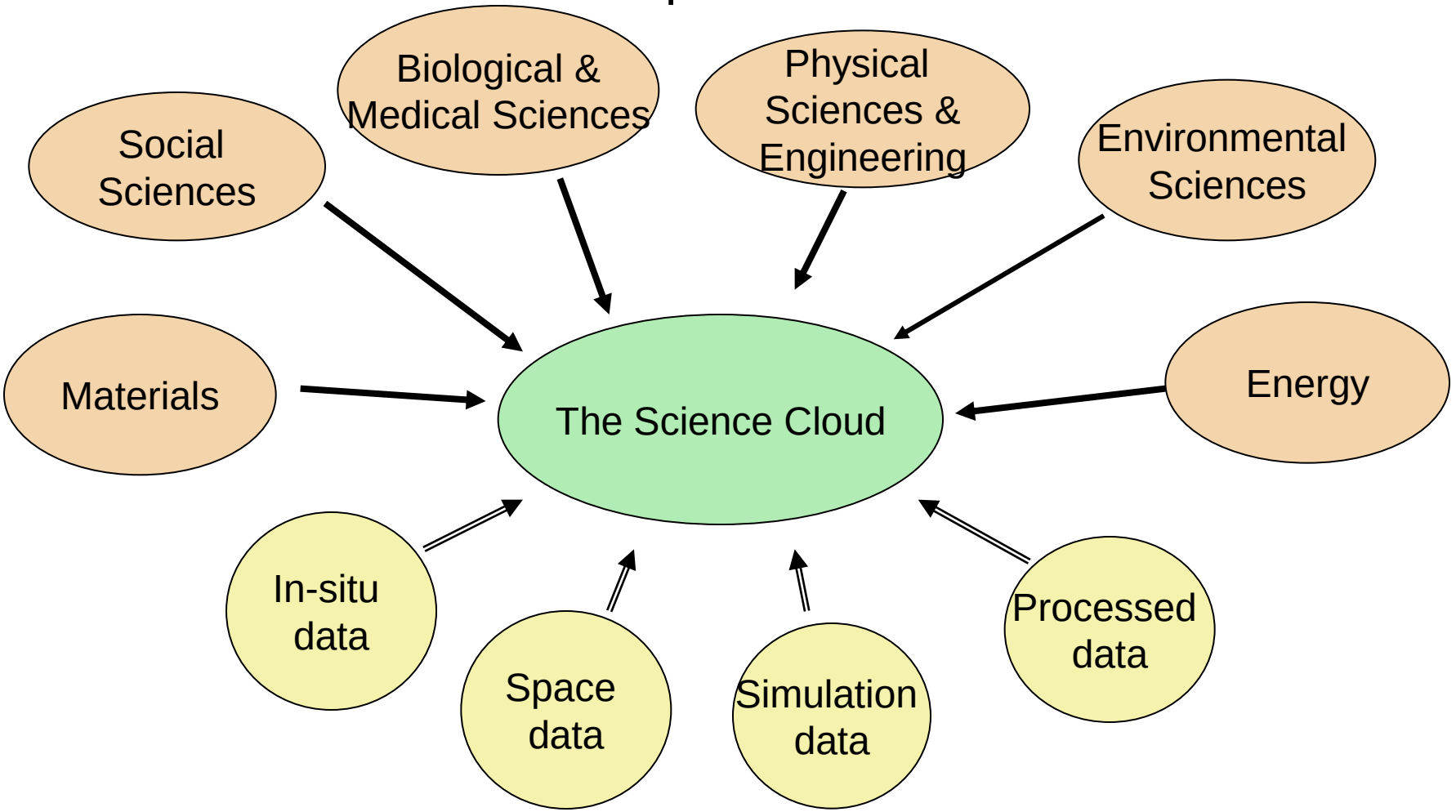
4 Goals outlined in the Strategic Plan

HELIX NEBULA THESCIENCECLOUD

1. Set up a **cloud computing infrastructure** for European Research Area
2. Identify and adopt **policies** for trust, security and privacy on a European-level
3. Create a light-weight **governance** structure involving all stakeholders
4. Define a short and medium term **funding** scheme

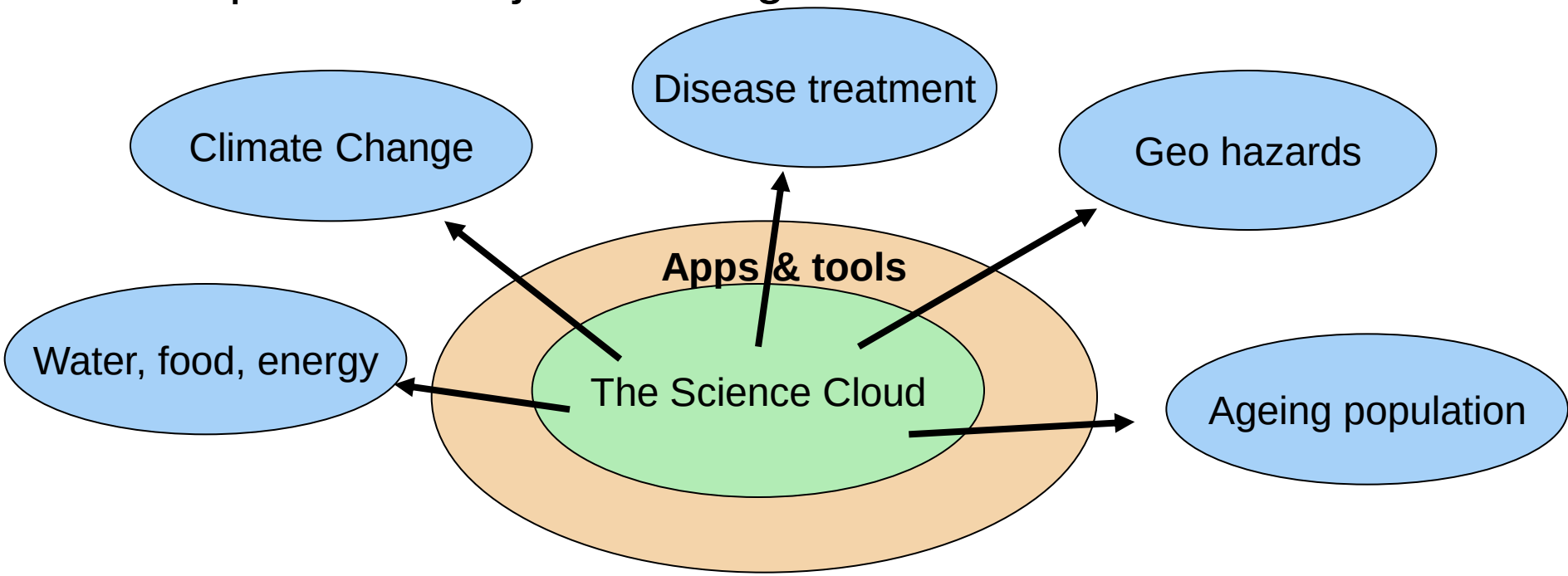
The Science Cloud : INPUT

The Science Cloud : a unique mine of scientific data



The Science Cloud: OUTPUT

The Science Cloud : a unique opportunity for Scientists to comprehend major challenges



Breakthrough in Science
Increased number of New Patents
Publications cross communities/domains

Addressing actions of the "Digital Agenda for Europe"

- Supporting the **single digital market**: building such a European

Digital Agenda **@DigitalAgendaEU** is now following you (**@HelixNebulaSC**).

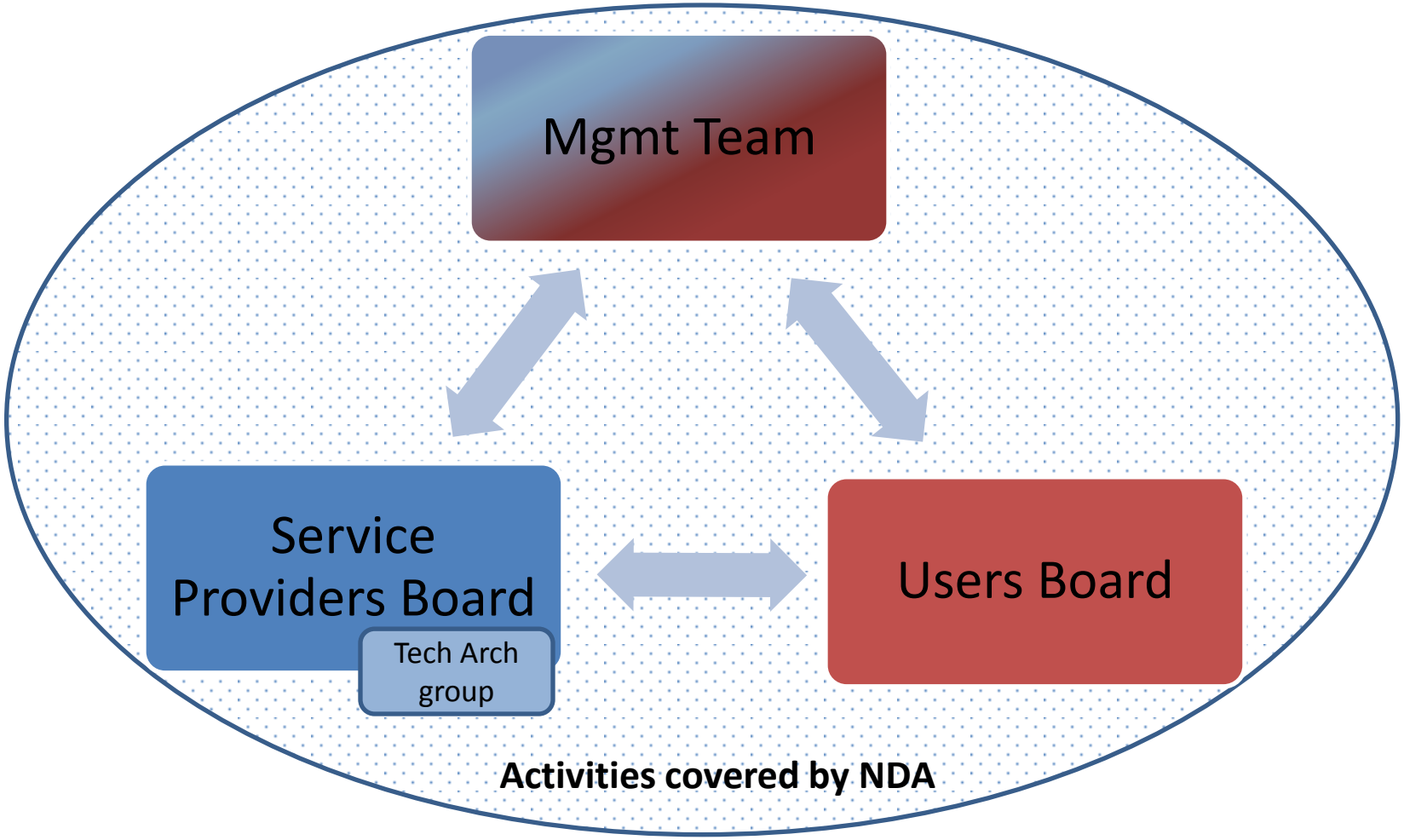
DigitalAgendaEU Digital Agenda
This is the official account of the EU's Digital Agenda policy flagship - providing all the news you need about maximising the potential of ICT in Europe.

coordination and pooling of resources;

- **Improving trust and security**: it will also provide a coordinated European approach to security for cloud computing and adhere to rules on data protection.


A Collaboration Initiative
European Commission
& relevant projects
User organisations
Demand-side
European
Cloud Computing
Strategy
Commercial Service
Providers
Supply-side
Bringing together all the stakeholders to establish a public-private partnership

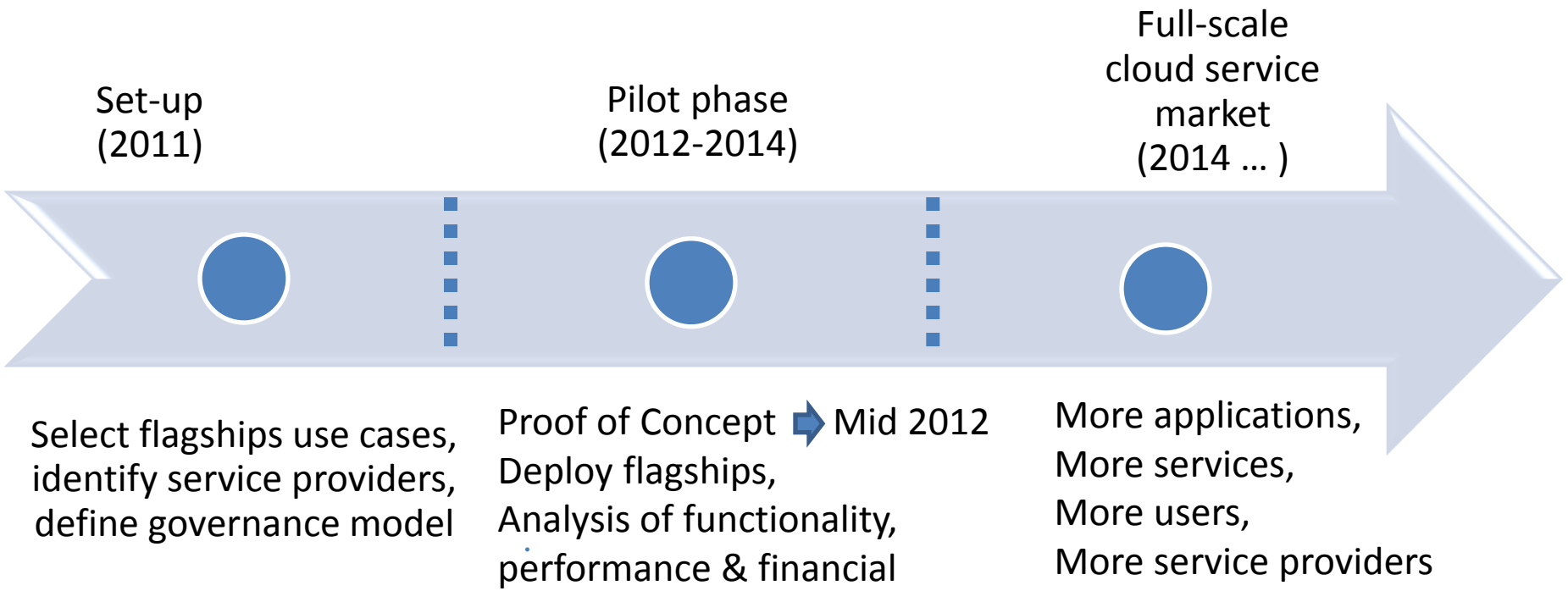
Governance Model during Proof of Concept in the Pilot Phase



Consortium membership

- Consortium includes all participating supply-side and demand-side companies/organisations
  - Member status and adopter status
  - All sign a non-disclosure agreement
  - “interested parties” can also register
- Initial membership is defined
  - More members and adopters will be added following the Proof of Concept stage within the Pilot Phase (end of 2012)

Timeline


Set-up
(2011)
Pilot phase
(2012-2014)
Full-scale
cloud service
market
(2014 …)
Select flagships use cases,
identify service providers,
define governance model
Proof of Concept
Mid 2012
Deploy flagships,
Analysis of functionality,
performance & financial
model
More applications,
More services,
More users,
More service providers

Helix Nebula Pilot Phase

Flagship use cases

Pilot Phase Goals

- Through the pilot phase we expect to explore/push a series of perceived barriers to Cloud adoption:
  - Security: Unknown or low compliance and security standards
  - Reliability: Availability of service for business critical tasks
  - Data privacy: Moving sensitive data to the Cloud
  - Scalability/Elasticity: Will the Cloud scale-up to our needs
  - Network performance: Data transfer bottleneck; QoS
  - Integration: Hybrid systems with in-house/legacy systems
  - Vendor lock-in: Dependency on vendors once data & applications have been transferred to the Cloud
  - Legal concerns: Such as who has legal liability
  - Transparency: Clarity of conditions, terms and pricing

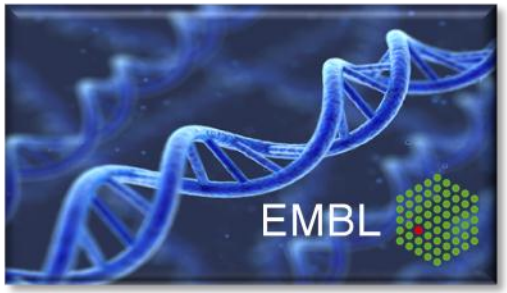
Flagship use cases

- **Proposed by demand-side user organisations addressing scientific challenges with societal impact**
  - High-profile applications that catch the public imagination and encourage others to use the services
  - Show need for significant scale of resources, federation/aggregation of data sets, long-term archiving and on-demand processing
  - Bring people and data together : Building communities and stimulating innovation

Use the Cloud for what it is good for !

Initial Flagship Use Cases

HELIX NEBULA THESCIENCECLOUD

ATLAS High Energy Physics Cloud Use	Genomic Assembly in the Cloud	SuperSites Exploitation Platform
		
To support the computing capacity needs for the ATLAS experiment	A new service to simplify large scale genome analysis; for a deeper insight into evolution and biodiversity	To create an Earth Observation platform, focusing on earthquake and volcano research

Call for proposals

- Template agreed by demand and supply side
- Eligibility review and analysis with cloud service suppliers

CERN-ATLAS Flagship Use Case

- Real time processing of ATLAS experiment
- Goals:
  - Evaluate cloud technologies for ATLAS use cases
  - Design model to integrate cloud resources with ATLAS distributed computing
  - Implementation in ATLAS software framework

EMBL Flagship Use Case

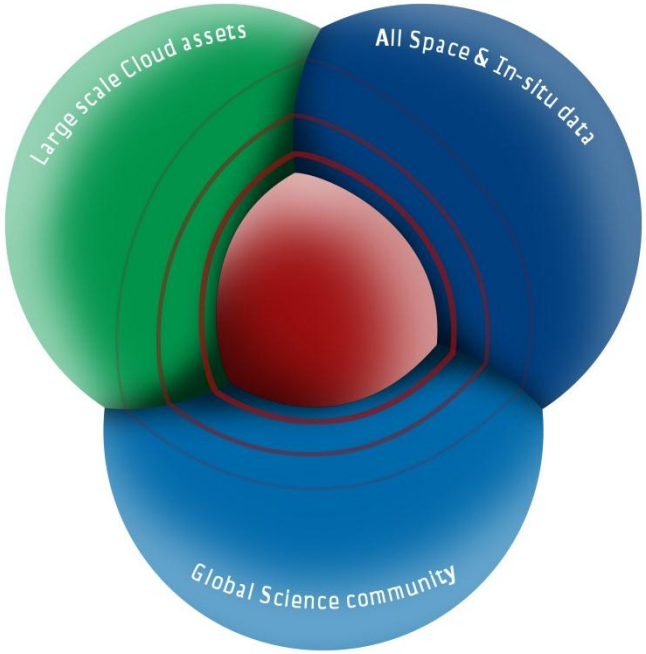
- Genomic Assembly in the Cloud
  - Genomic Sequencing is now an Affordable Solution and of interest for Academic Research Groups, Medical Research, Pharmaceutical industry and Agricultural Research
  - Sequencing technologies produce vast amount of data which need to be analysed
- Objectives
  - To provide a Genomic Assembly and Annotation Service to a broad range of researchers in various communities
  - To remove computational infrastructure hurdles that may prevent Genomic Assembly projects taking place

SuperSites Exploitation Platform (SSEP)

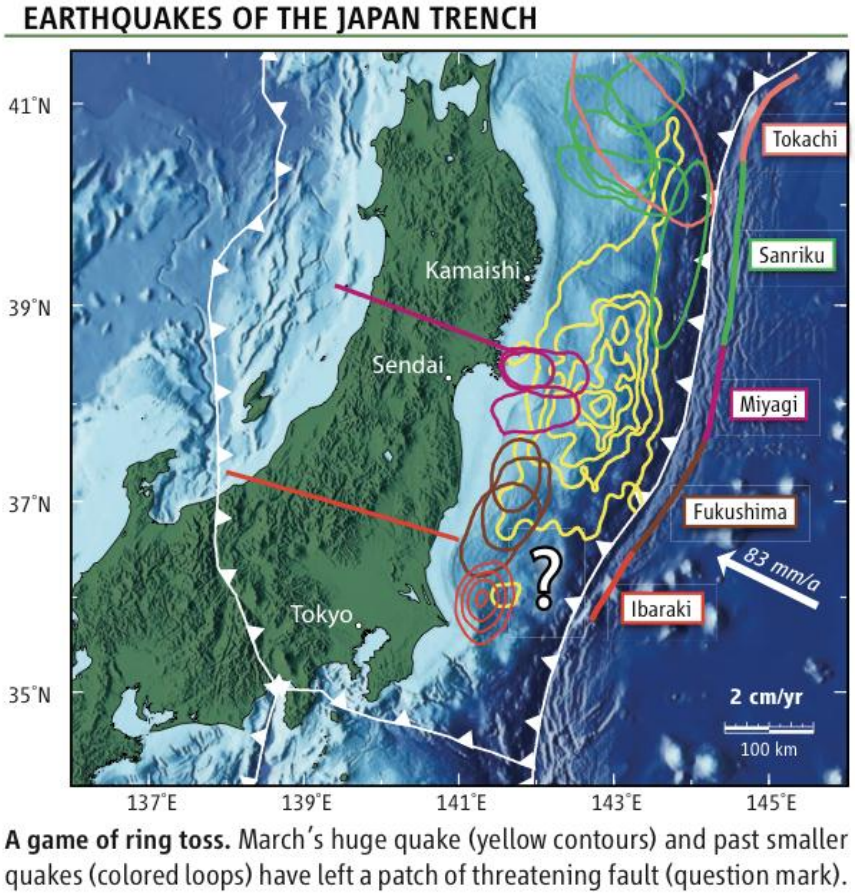
"Science Cloud" flagship proposal

by CNES, DLR, ESA with the

participation of Italian CNR



Geo Hazard : Japan earthquake



A game of ring toss. March's huge quake (yellow contours) and past smaller quakes (colored loops) have left a patch of threatening fault (question mark).

- Tohoku-oki: unprecedented >50 m slip in places (Simons et al., Science 2011, NASA-funded study).

- Will another magnitude 9 occur further south?

- It is unknown whether this fault segment has been accumulating slip.

- Need all InSAR, GPS, Seismic, Petrology, Geochemistry , … !

(-2004 magnitude 9.2 Sumatra earthquake was followed by magnitude 8.7 half-a-year later)

→ The Science Cloud with its "unlimited" resources on data, processing capacity and tools will allow cross-domain science and ease data sharing. The easy usage of this infrastructure will pull "intelligence" to apprehend the challenges.

SSEP Expected Results

1. Science: better scientific understanding of geohazards with the aim of providing sound information about the risks and the potential mitigation measures

2. Data sharing: information extracted from different sources (satellite & in-situ) will open a wide range of new approaches: Cross-domain research

3. Building Communities : In return to SSEP access, scientists will be asked either:
    - to provide results into the information repository on the Science Cloud,
    - to provide their data processing open source code,
    - to provide application tools (“Apps”)

Flagship use cases

	ATLAS H.E.P. Cloud Use (CERN)	Genomic Assembly in the Cloud (EMBL)	SuperSites Exploitation Platform (ESA/CNES/DLR)
Scientific goal/society impact/photogenic	•	•	•
Scale of resources used	•	•	
Federation/Aggregation of datasets		•	•
Long-term archiving of data			•
On-demand processing	•	•	•
Impact on community & benefits	•	•	•
Potential increase of users	•	•	•
Interoperability	•	•	•
Data security	•	•	•
Maturity	•	•	•
Access to license-controlled sw			•

Flagship deployments: Current status

- Proof of Concept stage within the Pilot Phase started January 2012
  - ✓ Acting committees are in place
  - ✓ Pilot now active
  - ✓ NDA signed by all members of the consortium
- Each flagship will be deployed with a series of providers independently
- Expect to have completed initial proof of concept by end of 2012 -> open to new entrants
- Intermediate results of PoC will be published early July 2012

Flagship use cases
Participating Suppliers in Proof of Concept stage

HELIX NEBULA
THESCIENCECLOUD

CloudSigma

Atos

interoute
from the ground to the cloud

terradue 2.0

logica
be brilliant together

sixsq

T-Systems

the SERVER LABS
the IT architects

What's next ?

- After PoC, new Suppliers and new Users can join
  - New comers can either submit flagships that propose some **innovation** in terms of functionality, performance, scope, business opportunities or impact of the European Cloud Computing infrastructure, or use the HN platform as is.
  - The flagship use cases must be **sponsored by user organisations** and **Service Providers**. Will be selected so as to be complementary and maximise coverage of the objectives outlined in the Strategic Plan

→ **Templates will be provided**

Relevance to e-IRG

- Address specific recommendations of **e-IRG White Paper 2011**:
  - Develop commercial offerings of e-Infrastructure services leading to their **sustainability**
  - Introduce governance models based on **Public-Private-Partnership**
  - Increase **collaboration within research area** (ESFRI, GEANT, Terena, PRACE, …)
  - Federated **Identity Management** will be addressed within Helix Nebula
  - e-Infrastructure services promoting **standards** will boost innovation and enlarge commercial market
  - Helix Nebula has the potential to become the major **European data infrastructure** allowing all scientific users to work cross-domain in order to tackle major **societal challenges** (Climate change, aging population, geo-hazards)

Possible collaboration with e-IRG

- We could propose input for e-IRG whitepaper on any policy aspects which we think will help Helix Nebula reach sustainability as an infrastructure
  - For example, we need the ability to have hybrid clouds that will allow us to integrate publicly funded infrastructures (such as our own data centres as well as GEANT/EGI/PRACE) within Helix Nebula

- e-IRG could help us in defining an appropriate Governance model for Helix Nebula (answering interviews or questionnaires)

A European Cloud Computing Partnership: big science teams up with big business

HELIX NEBULA THESCIENCECLOUD

Strategic Plan

- Establish multi-tenant, multi-provider cloud infrastructure
- Identify and adopt policies for trust, security and privacy
- Create governance structure
- Define funding schemes

CERN

To support the computing capacity needs for the ATLAS experiment

EMBL

Setting up a new service to simplify analysis of large genomes, for a deeper insight into evolution and biodiversity

esa

DLR cnes

To create an Earth Observation platform, focusing on earthquake and volcano research

Atos · Capgemini CONSULTING.TECHNOLOGY.OUTSOURCING · CloudSigma · CSA cloud security alliance · egi · interoute from the ground to the cloud · logica be brilliant together · OpenNebula.org The Open Source Toolkit for Cloud Computing · orange Business Services · SAP · the SERVER LABS the IT architects · sixsq · Telefónica · terradue 2.0 · THALES · Trust IT · T-Systems

Email:contact@helix-nebula.eu Twitter: HelixNebulaSC Facebook: HelixNebula.TheScienceCloud

Follow us …

- Helix Nebula on **twitter :** @HelixNebulaSC
- On **Facebook** : https://www.facebook.com/HelixNebula.TheScienceCloud
- On **LinkedIn** : Open Group Helix Nebula the Science Cloud

Big science teams up with big business to kick-start European cloud computing

PR03.12
01.03.2012

Geneva, 1 March 2012. Today a consortium of leading IT providers and three of Europe's biggest research centres (CERN[1], EMBL[2] and ESA[3]) announced a partnership to launch a European cloud computing platform. "Helix Nebula - the Science Cloud", will support the massive IT requirements of European scientists, and become available to governmental organisations and industry after an initial pilot phase.

CERN to tap new Europe cloud computing project

Three key research centers and a
European cloud computing platform

A statement from CERN, the Geneva-ba would be made available to government European Space Agency (ESA) and the vast volumes of information that cannot based Atos, Paris-based Cap Gemini, U based Thales, the statement said. "This and create new commercial markets," s bodies will initially use Helix Nebula to bc which is believed to have given mass to

Cloud computing centralizes data and ap companies and homes, and allows the s

A big cloud for big science | iSGTW

www.isgtw.org

Last week, a major European collaboration called Helix Nebula – the Science Cloud was announced. Its goal is to provide a Pan-European cloud computing platform for scientific research. Commercial partners are bringing the latest computing advancements to enable cheaper science and better collaborati…

Press Release 2012

Heidelberg, 1 March 2012

Flying high in Europe

Big science teams up with big business to kick-start European cloud computing

Today a consortium of leading IT providers and three of Europe's biggest research centres (CERN, EMBL and ESA) announced a partnership to launch a European cloud computing platform. 'Helix Nebula - the Science Cloud', will support the massive IT requirements of European scientists, and become available to governmental organisations and industry after an initial pilot phase.