Clouds in Science: preparatory Steps for LHC Computing

Bob Jones
Head of openlab
IT dept.
CERN

Sverre Jarp
CERN openlab CTO
IT dept.
CERN

ISC12
June 21st 2012, Hamburg

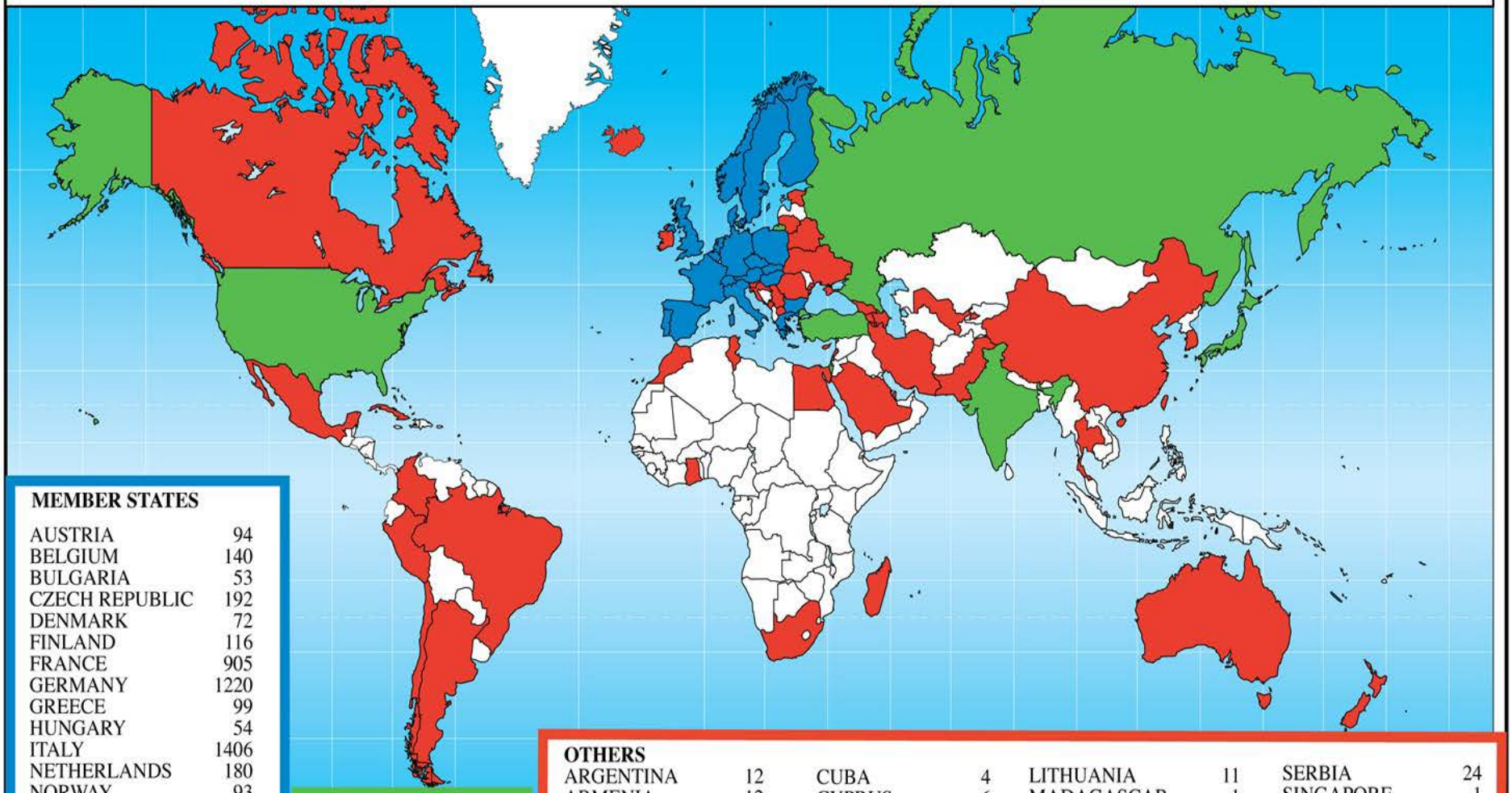
CERN was founded 1954: 12 European States
Today: 20 Member States

~ 2300 staff
~ 790 other paid personnel
> 10’000 users

Budget (2011) ~1000 MCHF

20 Member States: Austria, Belgium, Bulgaria, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland and the United Kingdom

Distribution of All CERN Users by Nation of Institute on 27 June 2011



MEMBER STATES

AUSTRIA	94
BELGIUM	140
BULGARIA	53
CZECH REPUBLIC	192
DENMARK	72
FINLAND	116
FRANCE	905
GERMANY	1220
GREECE	99
HUNGARY	54
ITALY	1406
NETHERLANDS	180
NORWAY	93
POLAND	205
PORTUGAL	141
SLOVAKIA	63
SPAIN	339
SWEDEN	79
SWITZERLAND	359
UNITED KINGDOM	732

6542

OBSERVER STATES

INDIA	109
ISRAEL	60
JAPAN	190
RUSSIA	822
TURKEY	79
USA	1786

3046

OTHERS

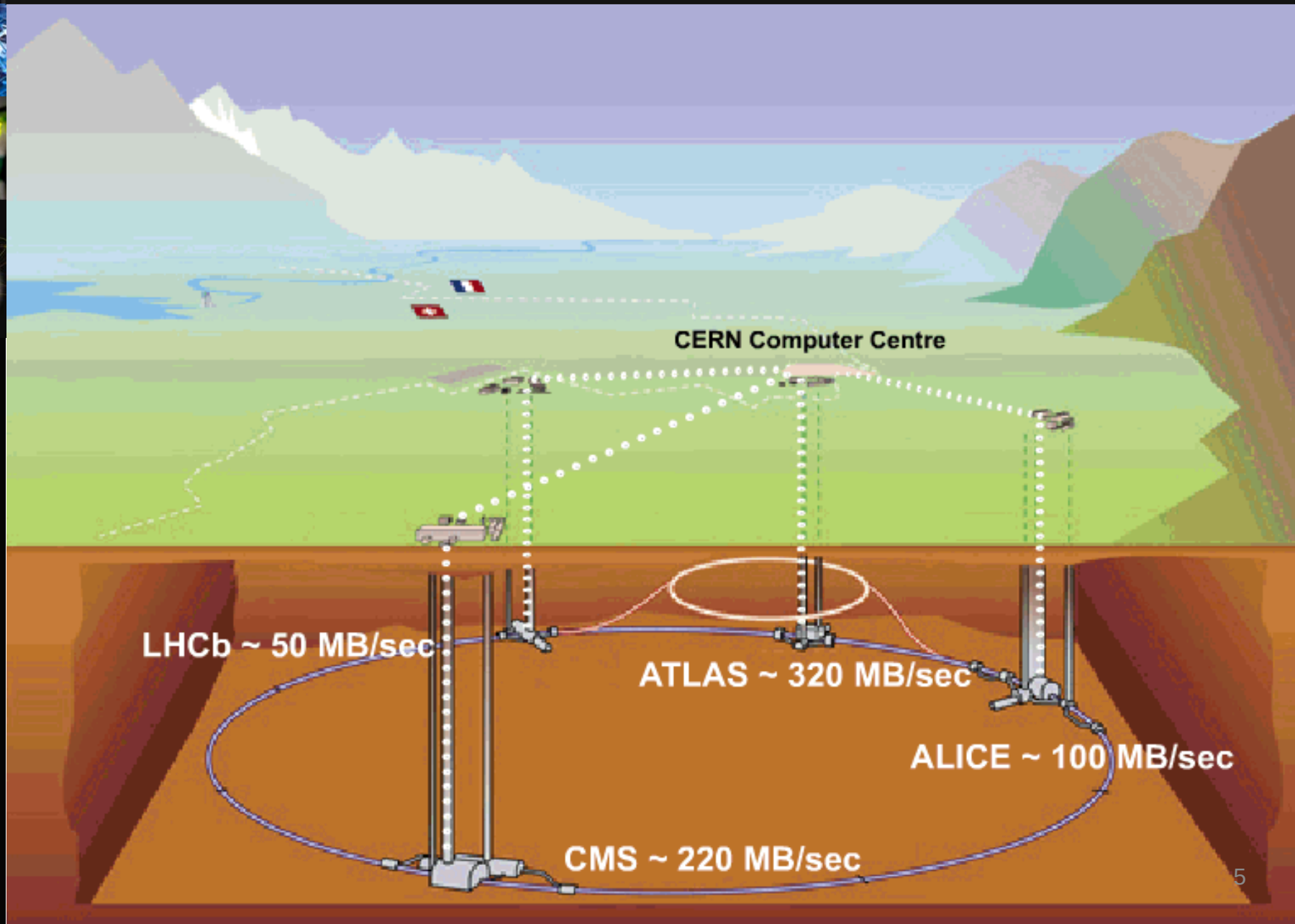
ARGENTINA	12	CUBA	4	LITHUANIA	11	SERBIA	24
ARMENIA	12	CYPRUS	6	MADAGASCAR	1	SINGAPORE	1
AUSTRALIA	22	EGYPT	6	MALTA	1	SLOVENIA	31
AZERBAIJAN	1	ESTONIA	18	MEXICO	39	SOUTH AFRICA	15
BELARUS	19	GEORGIA	10	MONTENEGRO	1	THAILAND	1
BRAZIL	79	GHANA	1	MOROCCO	7	F.Y.R.O.M.	3
CANADA	160	HONG KONG	1	NEW ZEALAND	9	TUNISIA	1
CHILE	3	ICELAND	3	PAKISTAN	19	UKRAINE	19
CHINA	87	IRAN	15	PERU	2	UZBEKISTAN	1
CHINA (TAIPEI)	53	IRELAND	13	QATAR	1		
COLOMBIA	13	KOREA	85	ROMANIA	66		
CROATIA	15	LEBANON	1	SAUDI ARABIA	2		

894

CERN

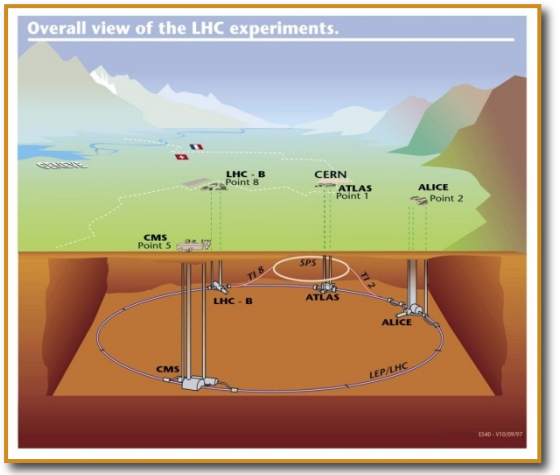
Accelerating Science and Innovation

Data acquisition and storage for LHC @ CERN

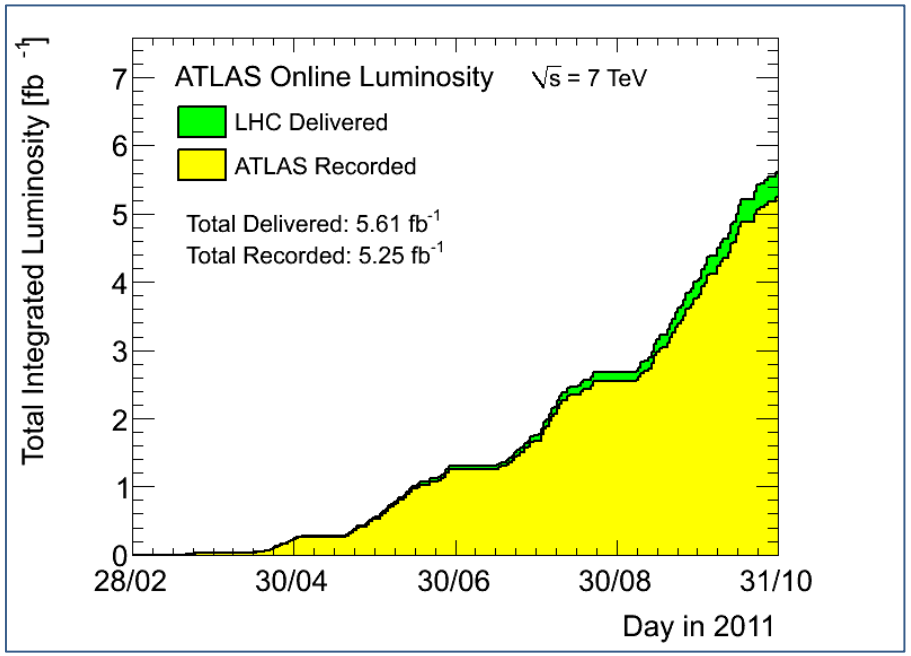


Large Hadron Collider (LHC)

- The biggest machine ever built
  - 27 km, 100 meters below ground
- Activities started in 2009
  - Highest energy in an accelerator
  - Large data sample of recorded collisions (events) available for high energy physics (HEP) measurements



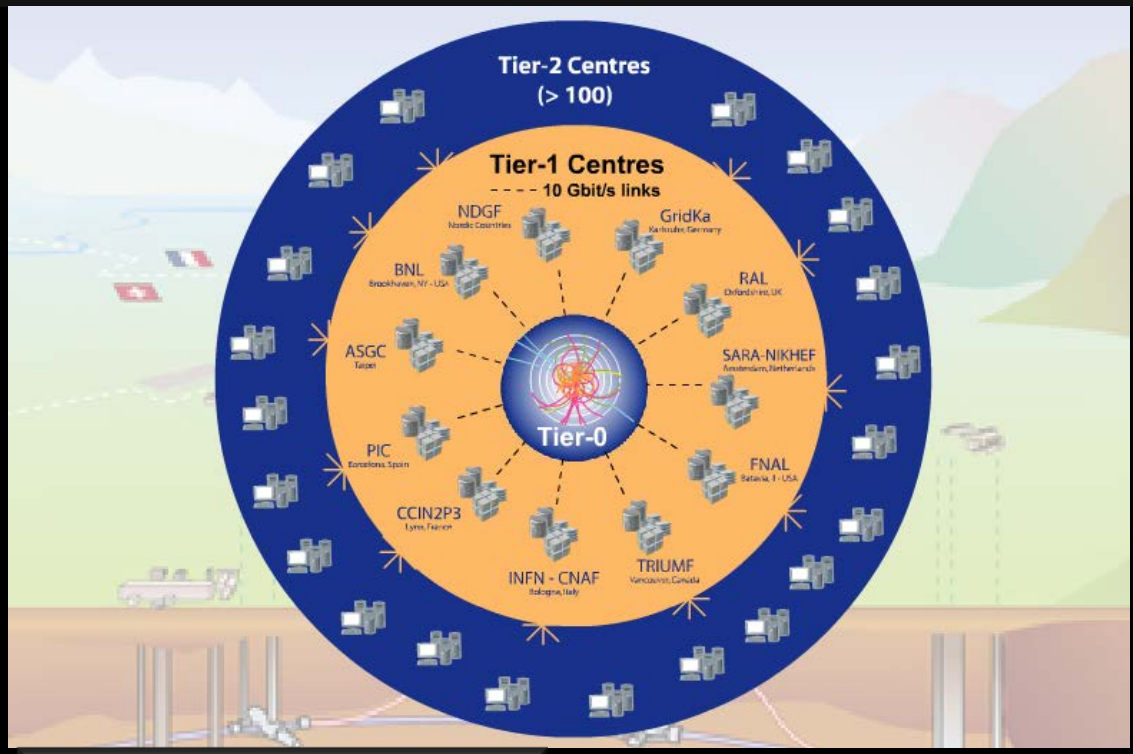
- **10^7 collisions per second**
- **Fortunately most collisions are uninteresting !**



CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

WLCG – what and why?

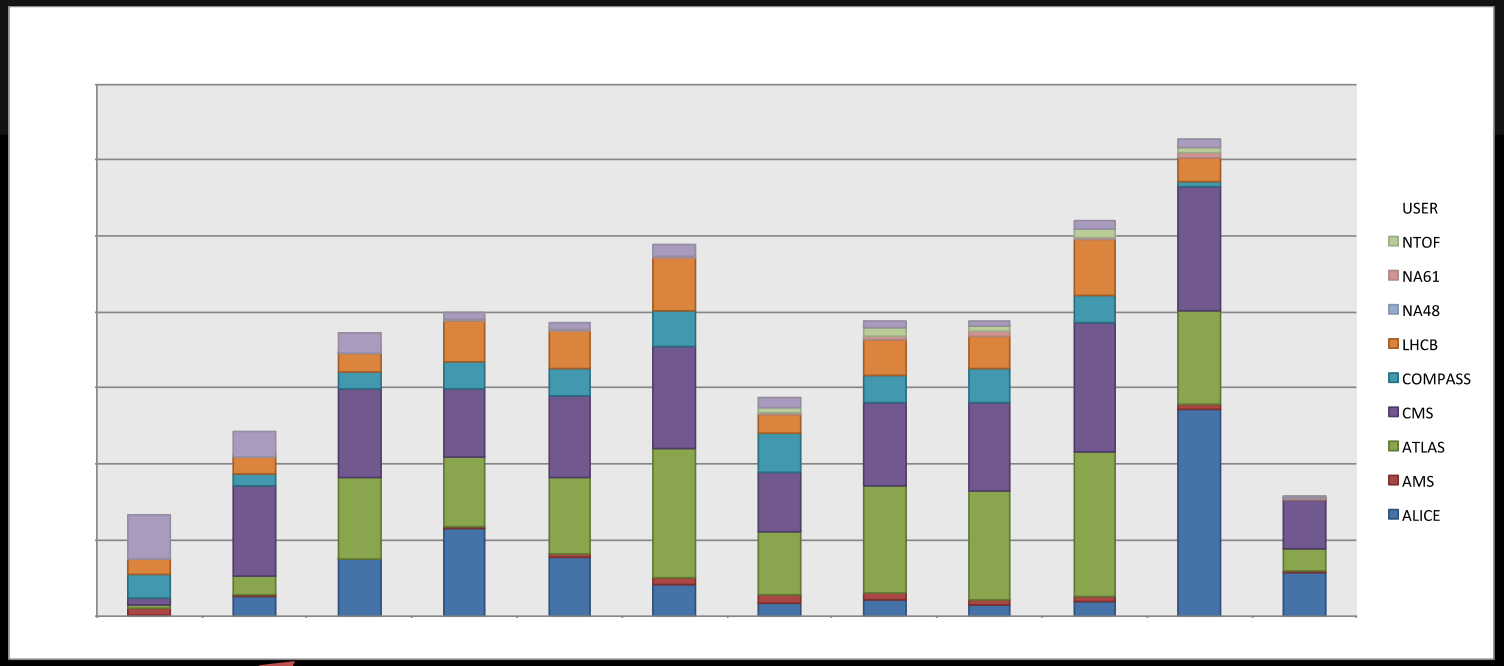
- A distributed computing infrastructure to provide the production and analysis environments for the LHC experiments
- Managed and operated by a worldwide collaboration between the experiments and the participating computer centres
- The resources are distributed – for funding and sociological reasons
- Our task was to make use of the resources available to us – no matter where they are located



Tier-0 (CERN):
- Data recording
- Initial data reconstruction
- Data distribution

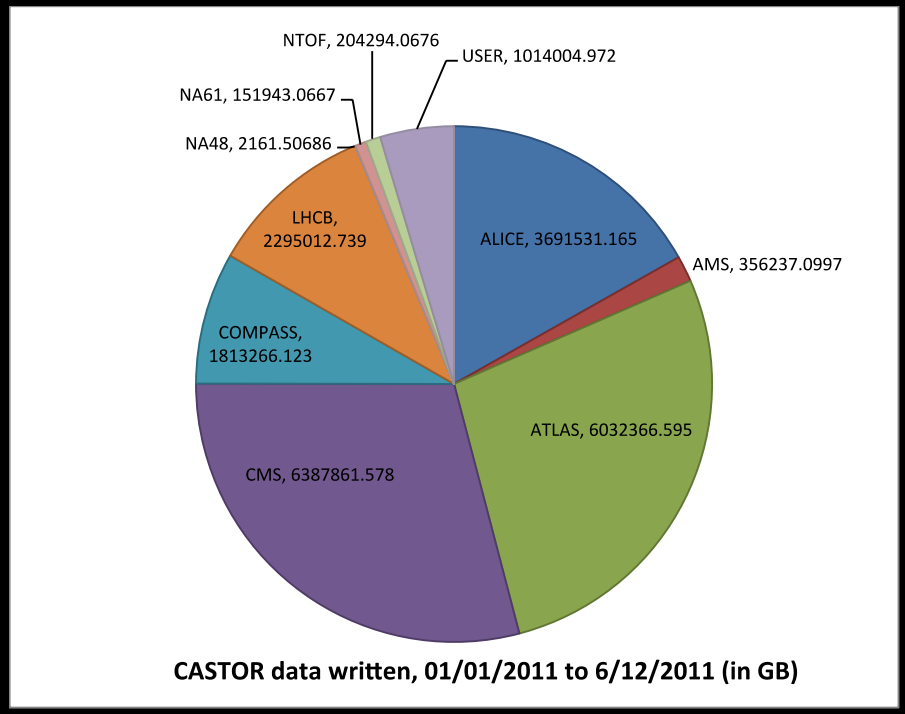
Tier-1 (11 centres):
- Permanent storage
- Re-processing
- Analysis

Tier-2 (~130 centres):
- Simulation
- End-user analysis



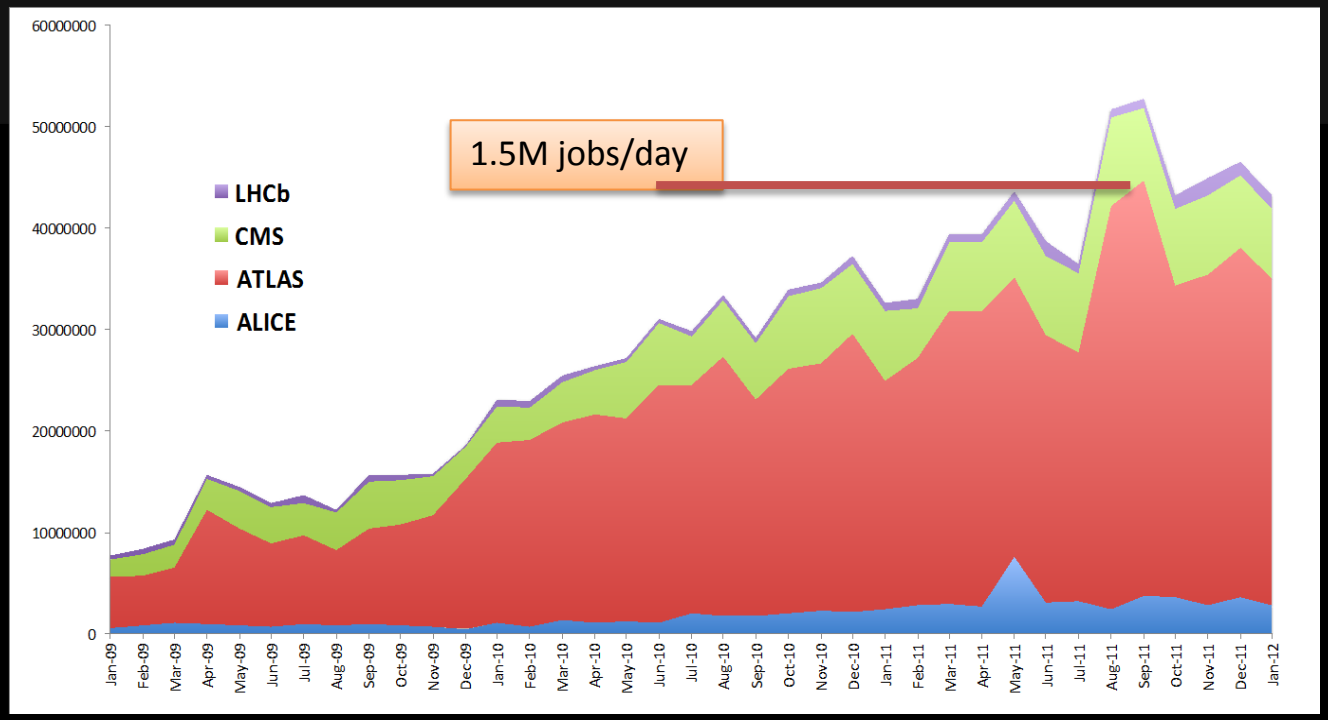
22 PB data written in 2011
More than 6 GB/s to tape during HI run

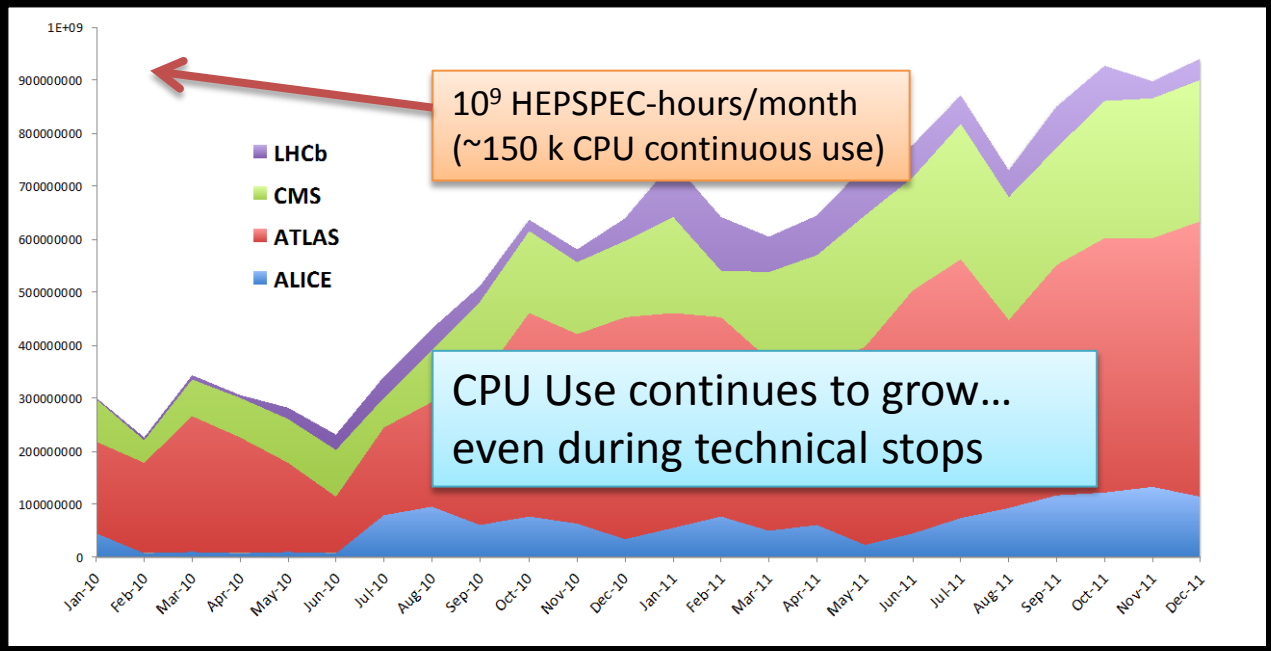
WLCG in 2011



CASTOR data written, 01/01/2011 to 6/12/2011 (in GB)

WLCG in 2011





WLCG Worldwide LHC Computing Grid

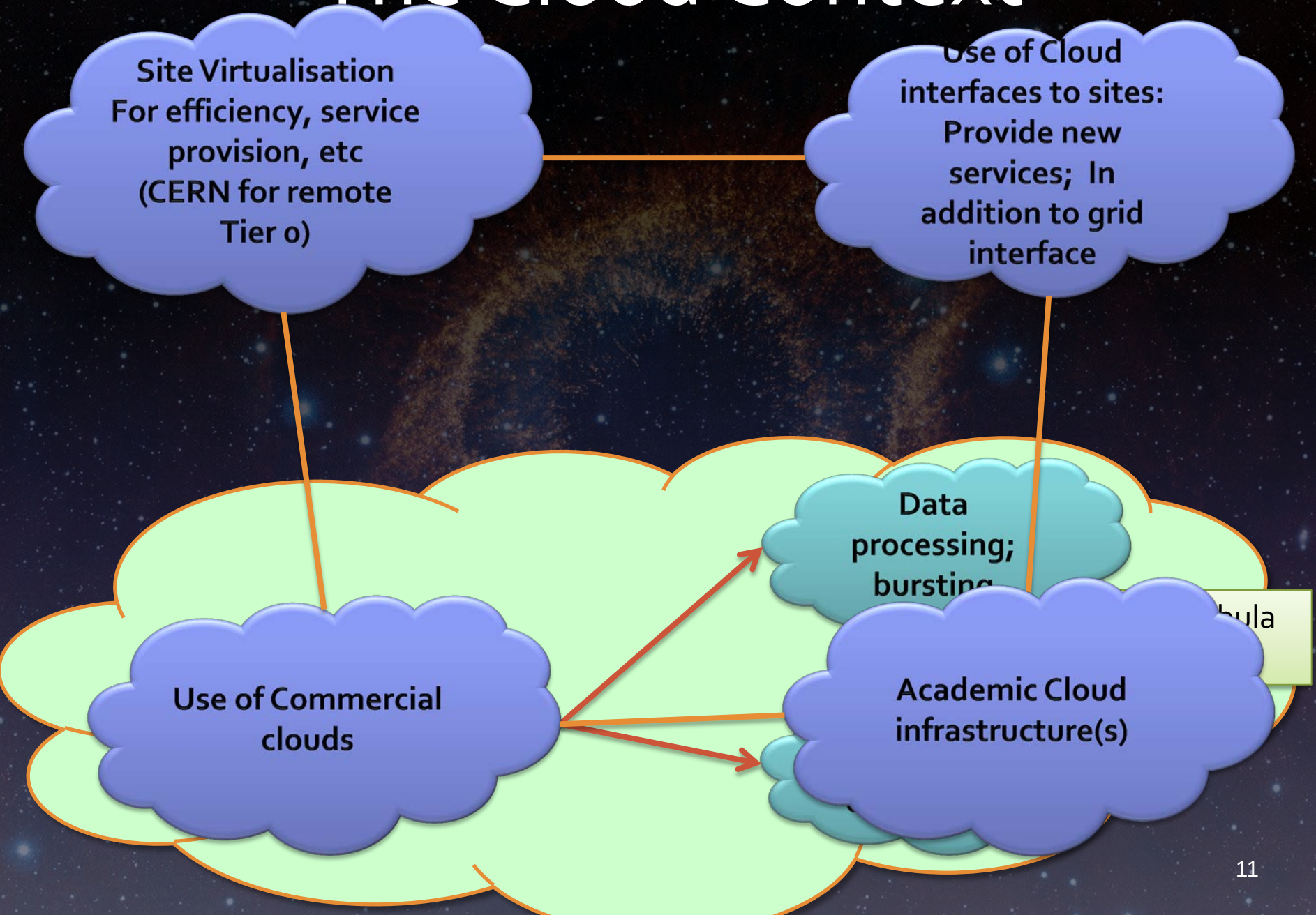
What is WLCG?

A distributed computing infrastructure to provide the production and analysis environments for the LHC experiments

- A **collaboration**
  - The resources are distributed and provided "in-kind"
- A **service**
  - Managed and operated by a worldwide collaboration between the experiments and the participating computer centres
- An **implementation**
  - Today general grid technology with high-energy physics specific higher-level services

Need to evolve the implementation while preserving the collaboration and the service

The Cloud Context



Strategic Plan

for a scientific Cloud Computing Infrastructure in Europe

- Establish a sustainable multi-tenant cloud computing infrastructure in Europe
- Initially based on the needs of the European Research Area & space agencies
- Integrate commercial services from multiple IT industry providers

Lengert, Maryline, Jones, Robert (2011) CERN-OPEN-2011-036
http://cdsweb.cern.ch/record/1374172/

A Collaboration Initiative

HELIX NEBULA
THESCIENCECLOUD

European Commission
& relevant projects

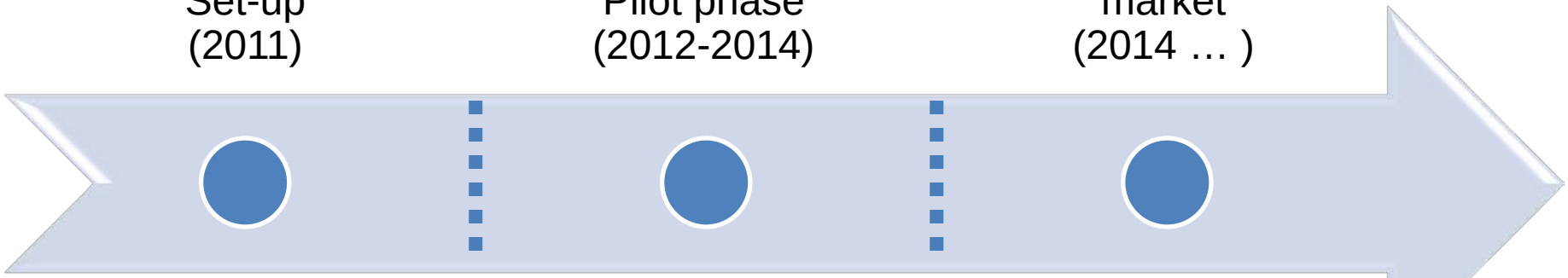
User organisations
Demand-side

European Cloud Computing Strategy

Commercial Service Providers
Supply-side

Bringing together all the stakeholders to establish a public-private partnership

Timeline



Set-up (2011)

- Select flagship use cases
- Identify service providers
- Define governance model

Pilot phase (2012-2014)

- Deploy flagships
- Analysis of functionality, performance & financial model
- Success Stories

Full-scale cloud service market (2014 …)

- More applications
- More services
- More users,
- More service providers

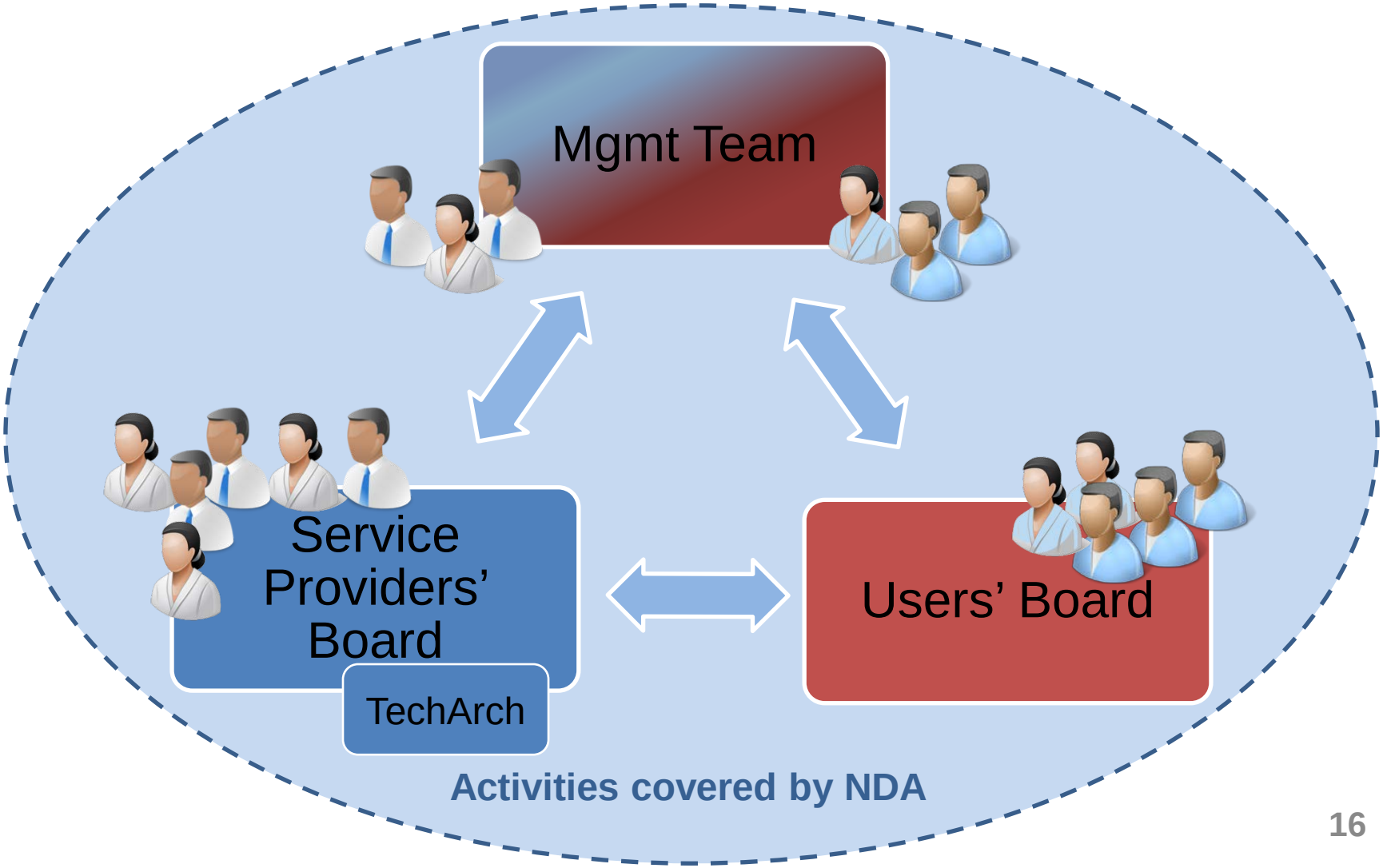
Pilot Phase

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**: Vendor dependency once data & applications are transferred to the Cloud
- **Legal concerns**: liability, jurisdiction, intellectual property
- **Transparency**: Clarity of conditions, terms and pricing

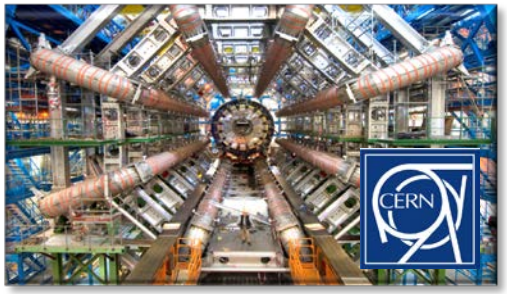
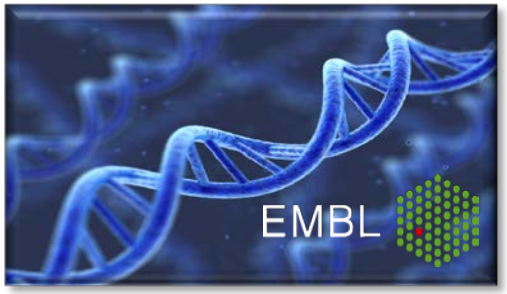
Governance Model

Proof of Concept stage



Initial flagships 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

- **Scientific challenges with societal impact**
- **Sponsored by user organisations**
- ***Stretch* what is possible with the cloud today**

Flagship use cases. Participating Suppliers in Proof of Concept stage

HELIX NEBULA THESCIENCECLOUD

CloudSigma

Atos

interoute
from the ground to the cloud

logica
be brilliant together

terradue 2.0

sixsq

T-Systems

the SERVER LABS
the IT architects

Initial results July'12

Addressing actions of the "Digital Agenda for Europe"

- Supporting the **single digital market**: building such a European Infrastructure will help create a single digital market for cloud computing;
- Enhancing **interoperability and standards**: the European Cloud Computing Infrastructure will permit geographically dispersed and separately managed devices, applications and services to interact seamlessly;
- **Stimulating research and innovation**: implementing this strategic plan will generate more private investment for IT research to develop a new generation of applications and services as well as reinforce the 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.

Helix Nebula funded by EC under grant 312301 with 1.8M€

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

There will be more news in September

Thank you!