CERN's Scientific Programme and the need for computing resources

Bob Jones
Head of openlab
IT dept
CERN

CERN founded 1954: "Science for Peace"

~ 2300 staff
~ 1050 other paid personnel
~ 11000 users

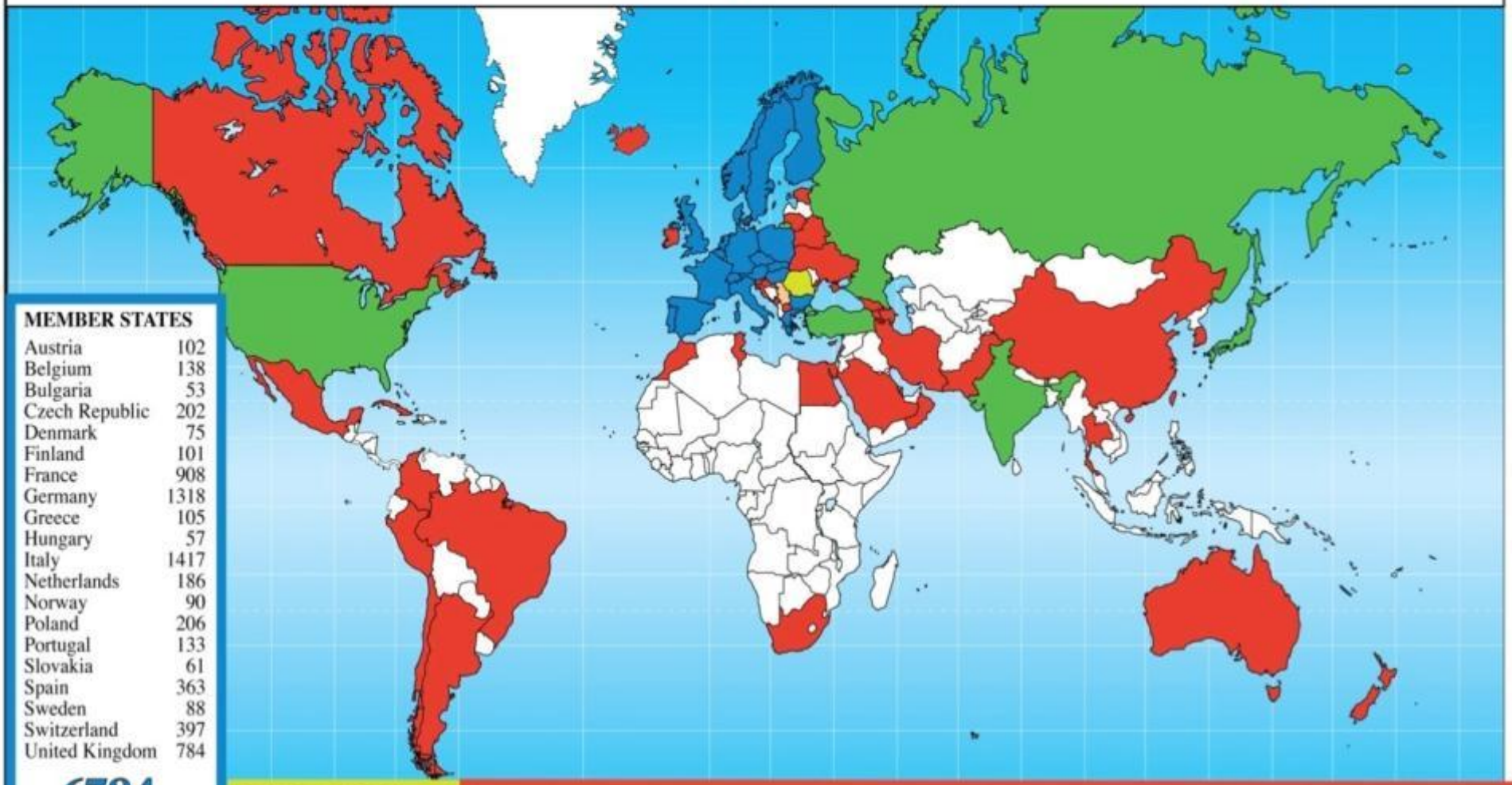
Budget (2012) ~1000 MCHF

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

1 Candidate for Accession: Romania

8 Observers to Council: India, Israel, Japan, the Russian Federation, the United States of America, Turkey, the European Commission and UNESCO

Distribution of All CERN Users by Nation of Institute on 4 April 2012



MEMBER STATES

Austria	102
Belgium	138
Bulgaria	53
Czech Republic	202
Denmark	75
Finland	101
France	908
Germany	1318
Greece	105
Hungary	57
Italy	1417
Netherlands	186
Norway	90
Poland	206
Portugal	133
Slovakia	61
Spain	363
Sweden	88
Switzerland	397
United Kingdom	784

6784

OBSERVERS

India	134
Japan	225
Russia	859
Turkey	83
USA	1749

3050

CANDIDATE FOR ACCESSION

Romania	78

ASSOCIATE MEMBER IN THE PRE-STAGE TO MEMBERSHIP

Israel	67
Serbia	26

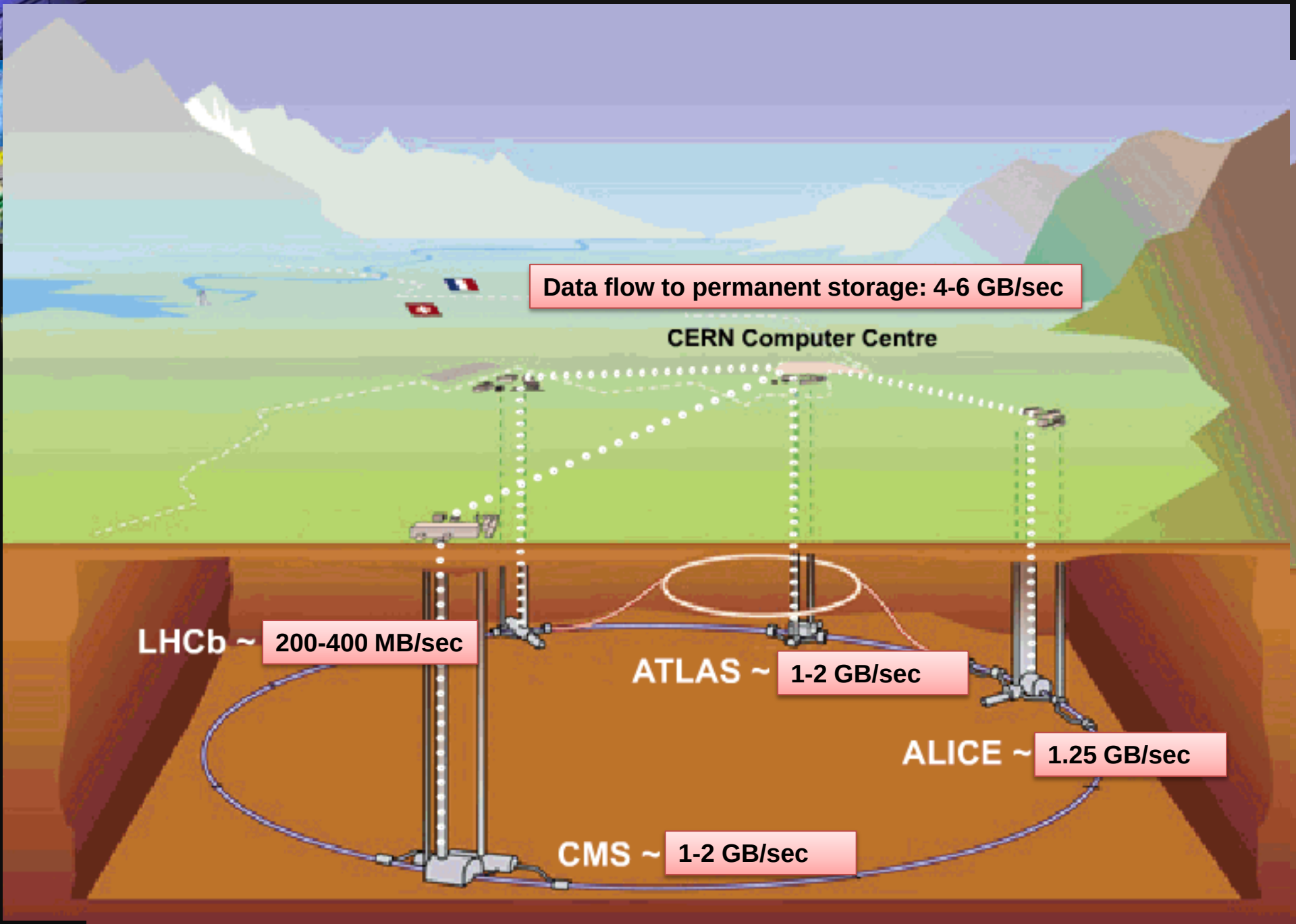
OTHERS

Argentina	18	China	115	Iran	16	Oman	1	Ukraine	21
Armenia	13	China (Taipei)	70	Ireland	10	Pakistan	22	Uzbekistan	1
Australia	28	Colombia	10	Korea	91	Peru	2		
Azerbaijan	1	Croatia	21	Lebanon	1	Qatar	1		
Belarus	22	Cuba	4	Lithuania	13	Saudi Arabia	3		
Brazil	102	Cyprus	9	Malta	1	Slovenia	38		
Canada	170	Egypt	7	Mexico	43	South Africa	21		
Chile	4	Estonia	17	Montenegro	1	Thailand	5		
		Georgia	10	Morocco	6	T.F.Y.R.O.M.	2		
		Iceland	3	New Zealand	11	Tunisia	1		

934

CERN

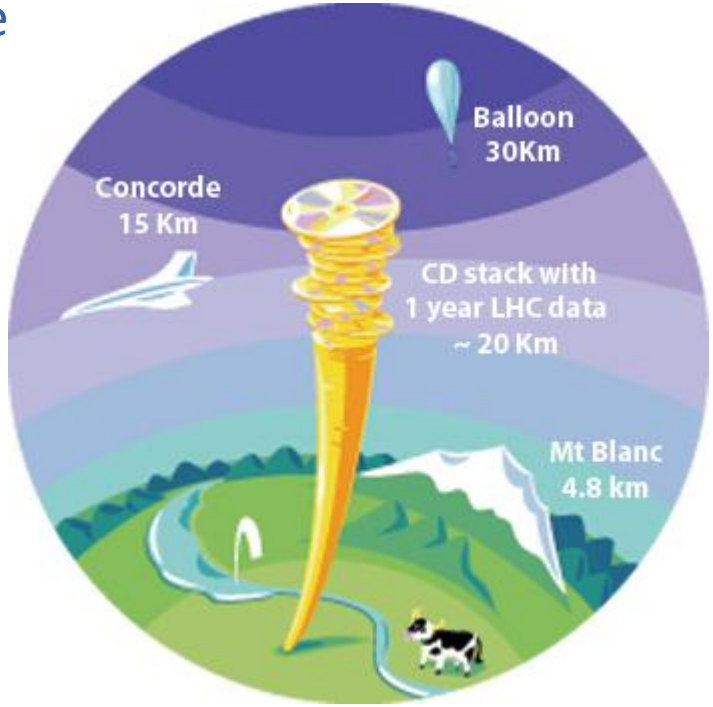
Accelerating Science and Innovation


Data flow to permanent storage: 4-6 GB/sec
CERN Computer Centre
LHCb ~ 200-400 MB/sec
ATLAS ~ 1-2 GB/sec
ALICE ~ 1.25 GB/sec
CMS ~ 1-2 GB/sec

The LHC Data Challenge

CERN IT Department

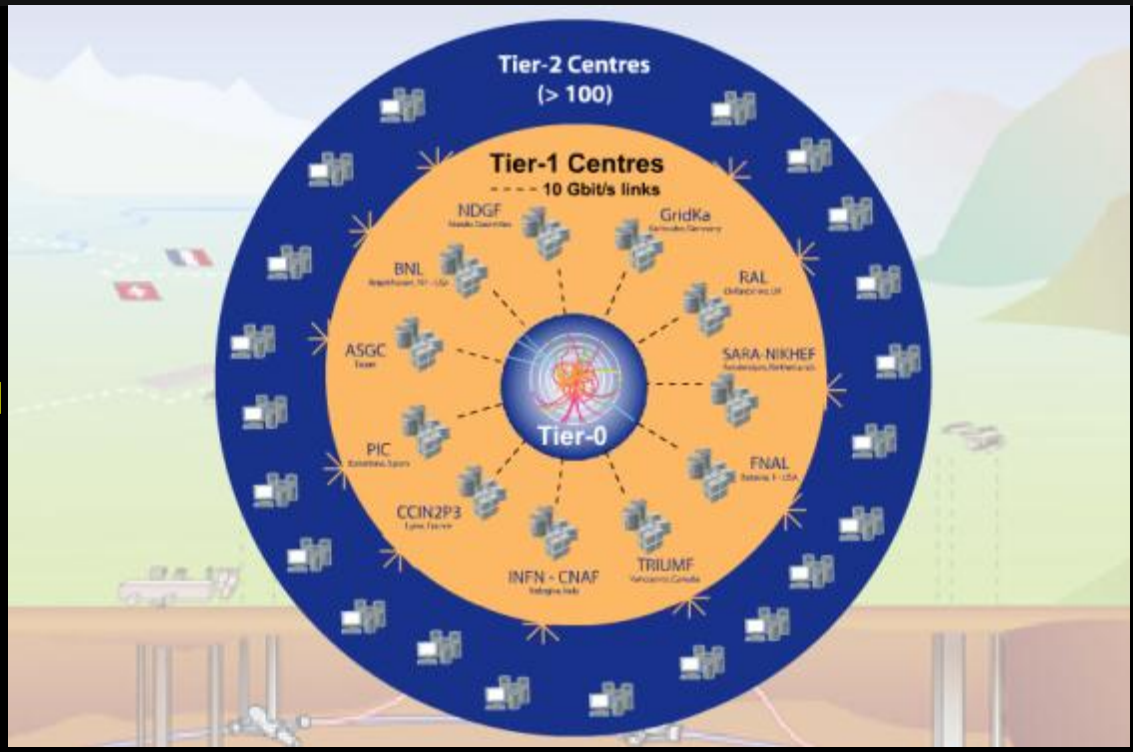
- The accelerator will run for 20 years
- Experiments ***are*** producing more than **20 Million Gigabytes** of data each year
- LHC data analysis requires a computing power equivalent to **~100,000 of today's fastest PC processors**
- Requires many cooperating computer centres, as CERN can only provide <20% of the capacity



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

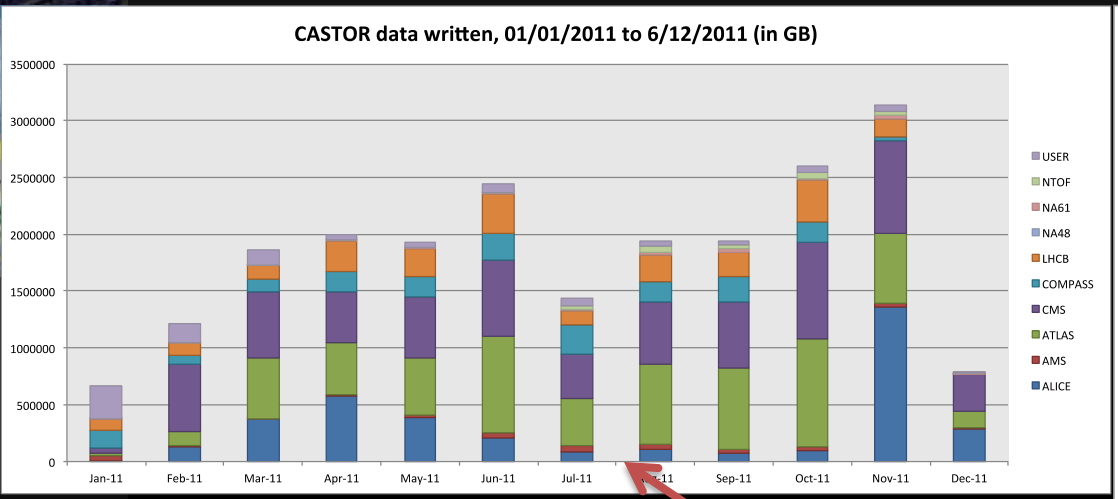


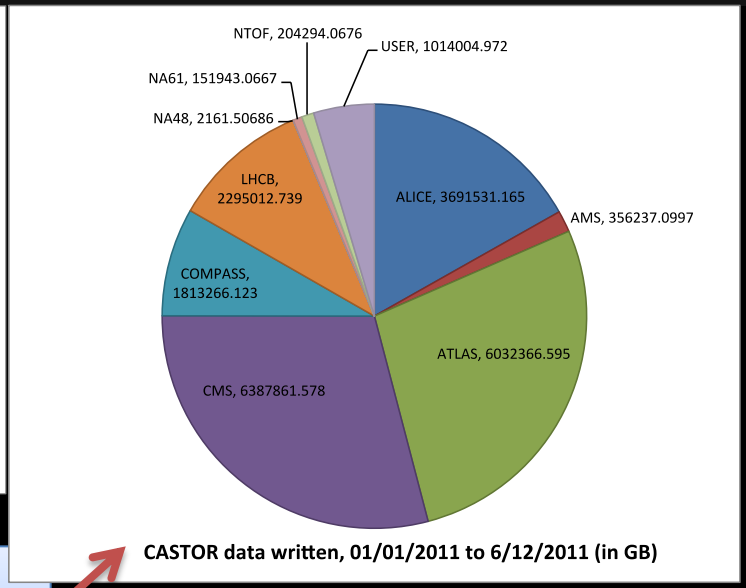
Tier-0 (CERN):
- Data recording
- Initial data reconstruction
- Data distribution

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

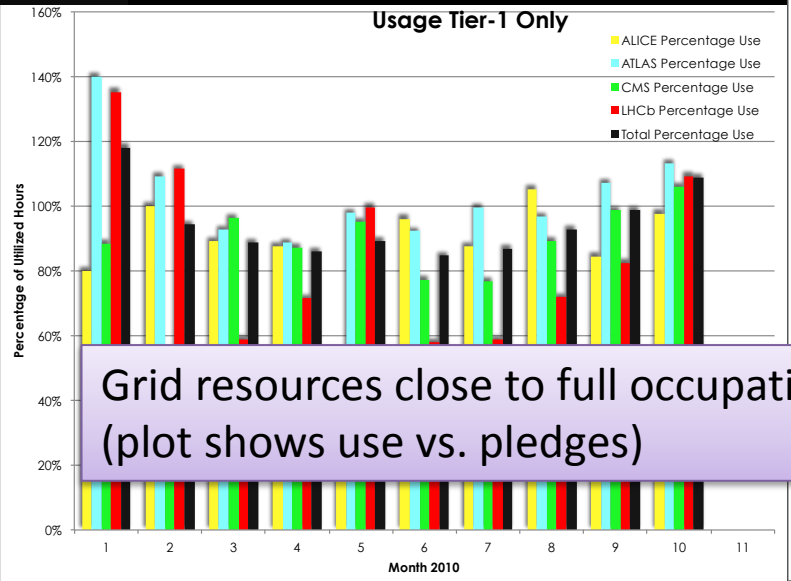
Tier-2 (~130 centres):
- Simulation
- End-user analysis

WLCG in 2011

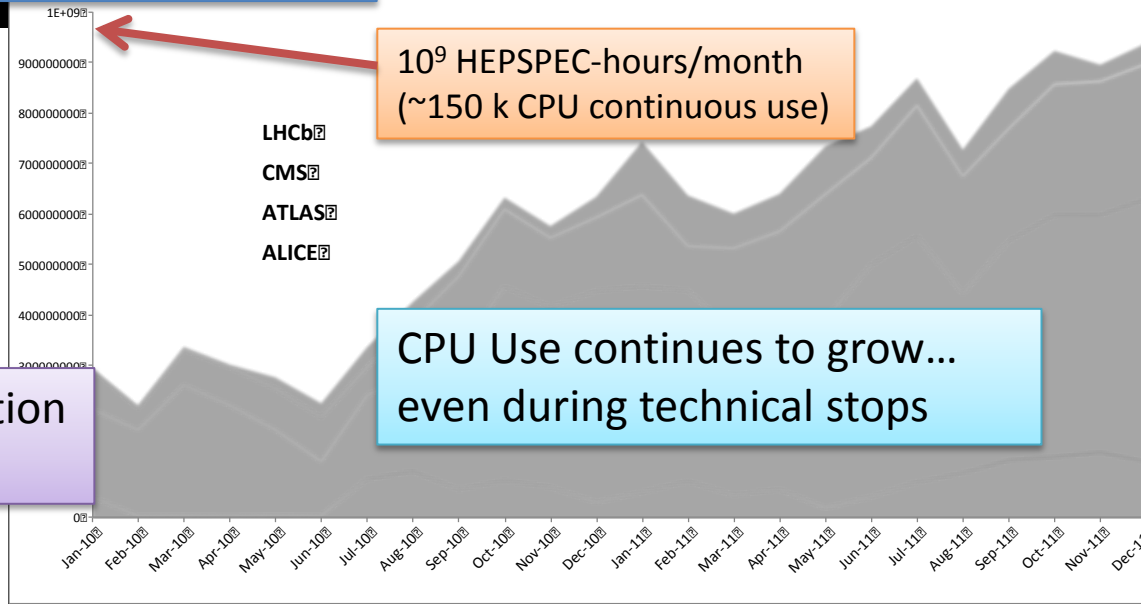




22 PB data written in 2011
More than 6 GB/s to tape during HI run



Grid resources close to full occupation
(plot shows use vs. pledges)



10^9 HEPSPEC-hours/month
(~150 k CPU continuous use)

CPU Use continues to grow…
even during technical stops

How to evolve WLCG?

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

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

Evolve the **Implementation** while preserving the **collaboration** & **service**

CERN openlab in a nutshell

- A science – industry partnership to drive R&D and innovation with over a decade of success
- Evaluate state-of-the-art technologies in a challenging environment and improve them
- Test in a research environment today what will be used in many business sectors tomorrow
- Train next generation of engineers/employees
- Disseminate results and outreach to new audiences

CERN openlab

www.cern.ch/openlab

PARTNERS

hp invent

intel

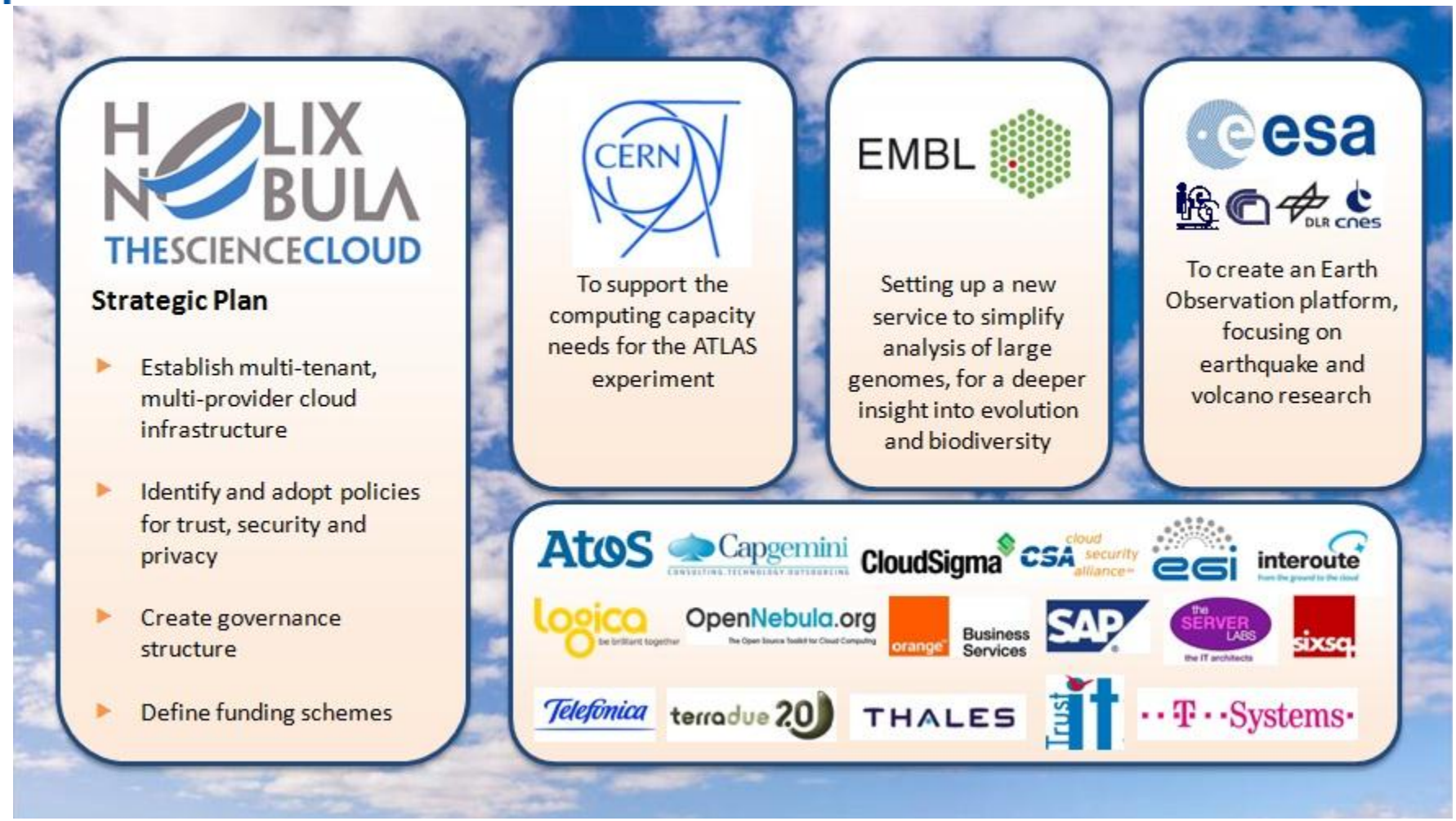
ORACLE

SIEMENS

CONTRIBUTOR (2012)

HUAWEI

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



Email:contact@helix-nebula.eu Twitter: HelixNebulaSC Website: http://www.helix-nebula.eu/

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 for the European Research Area & space agencies
- Integrate commercial services from multiple IT industry providers

http://cdsweb.cern.ch/record/1374172/

A Collaboration Initiative

HELIX NEBULA
THESCIENCECLOUD

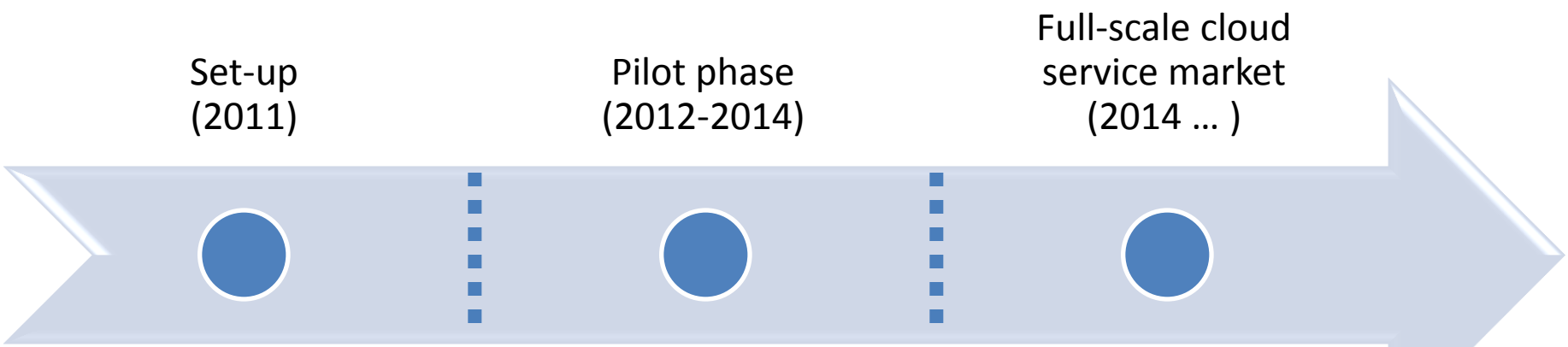
European Commission
& relevant projects

User organisations
Demand-side

Commercial Service Providers
Supply-side

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

Timeline



- Agree strategic plan
- Select flagships use cases
- Identify service providers
- Define governance model

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

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

co-funded by EC under grant 312301 with 1.8M€

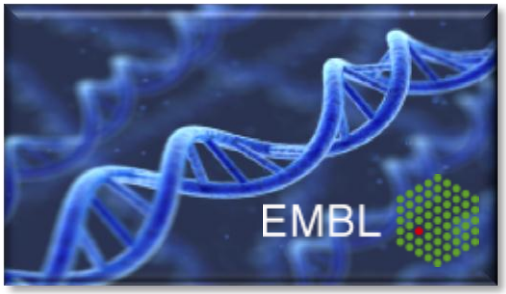
Pilot Phase

Explore / push a series of perceived barriers to Cloud adoption via deployment of selected flagship applications:

- **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
- **Business Models**: Finding a win-win model for the suppliers and users
- **Legal concerns**: liability, jurisdiction, intellectual property
- **Transparency**: Clarity of conditions, terms and pricing

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

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

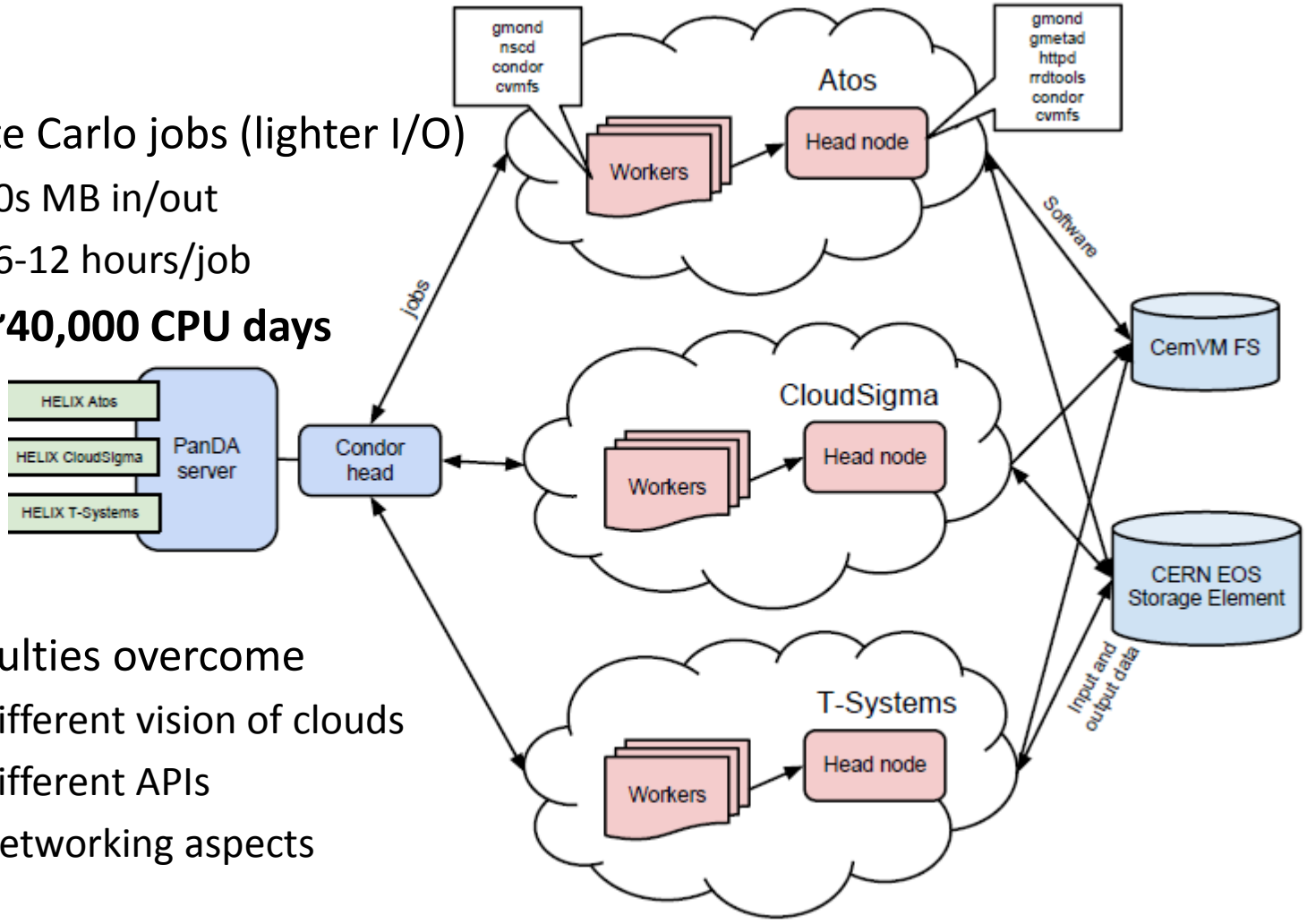
Flagship deployments
First results

- Proof of Concept deployments within the Pilot Phase started January 2012
- Each flagship has been deployed with a series of providers independently

CERN, EMBL and ESA succeeded in deploying scientific applications each involving tens of thousands of jobs running at data centres operated by Atos, CloudSigma and T-Systems

CERN-ATLAS flagship configuration

- Monte Carlo jobs (lighter I/O)
  - 10s MB in/out
  - ~6-12 hours/job
- **Ran ~40,000 CPU days**



- Difficulties overcome
  - Different vision of clouds
  - Different APIs
  - Networking aspects

Ramón Medrano Llamas,Fernando Barreiro, Dan van der Ster (CERN IT), Rodney Walker (LMU Munich)

Open to new members

Users

Service Providers

Adopters

Interested Parties

Become a new member

Over the next two years the Helix Nebula Consortium is expected to involve an increasing number of members.

The potential members are likely to be stemming from the categories below:

Users: An organisation that applies to become a user member should be a science or space organisation and commit to provide at least one flagship use case for cloud computing that can be verified and validated through a Proof of Concept with multiple service providers. Users will name a representative to participate in regular meetings of the Helix Nebula Users Board.

Service Providers: An organisation that applies to become a service provider member should commit to support a minimum set of cloud computing services and perform at least one Proof of Concept of a flagship use case with a user. Service Providers can apply to one or more of the following categories: Connectivity Provider, Infrastructure-as-a-Service Provider (IaaS) Platform-as-a-Service Provider (PaaS), Software-as-a-Service Provider (SaaS), Integrator, Consultant or Broker. Service Providers will name a representative to participate in regular meetings of the Helix Nebula Service Providers Board.

Adopters: An organisation may apply to become an adopter, if they initially do not want to be directly involved in the flagship use cases but wish to make use of the Helix Nebula's products and services on a pay-per-use basis and be able to provide feedback. Adopters may be invited to Boards (either Service Providers or Users) meetings for information purposes. Adopters will name a representative to be informed or invited by the Helix Nebula Consortium on a regular basis.

Interested Parties: An organisation may apply to become an Interested Party, if they initially do not want to be directly involved in the flagship use cases but wish to be kept informed of the work of or use information provided by the Consortium.

For all types of members a formal acceptance procedure will apply (specified hereafter), to be implemented by the Helix Nebula Management Team following the receipt of the membership request. **Membership application implies acceptance of the vision of Helix Nebula** as outlined in the strategic plan Strategic Plan for a Scientific Cloud Computing infrastructure for Europe, CERN-OPEN-2011-036, August 2011, **and willingness to collaborate with the other partners** in order to achieve the vision. All decisions must be adopted by consensus of both the user organisations and service provider companies. In case of lack of consensus, decision will be taken by a qualified majority of all members of the two Boards, which must include the positive vote of the public user organisations.

All users and service providers applying to become a member of the Helix Nebula Consortium (except as an Interested Party) agree to sign a multi-lateral NDA prior to becoming an active member.

For more details and updates about how to join, write to us at contact@helix-nebula.eu

Events

Helix Nebula @ DCI Workshop, 18 September 2012, Prague

Bob Jones (CERN) and Michael Symonds (Atos) will participate to the DCI Workshop *"Distributed Computing Infrastructures for e-Science: Future Perspectives"*.

More

Participants

http://www.helix-nebula.eu

Thank you for your attention

Accelerating Science and Innovation