
CERN

Agile Infrastructure at CERN - Moving 9'000 Servers into a Private Cloud

Helge Meinhard
Leader, Platform and Engineering Services Group, IT Department
04 April 2014

CERN

CERN

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954

1954: 12 Member States

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954
- Facilities for fundamental research in particle physics

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954
- Facilities for fundamental research in particle physics
- 21 member states, 1 B CHF budget

1954: 12 Member States

Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954
- Facilities for fundamental research in particle physics
- 21 member states, 1 B CHF budget
- 3’360 staff, fellows, students, apprentices, …

1954: 12 Member States

Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

2’512 staff members, 540 fellows, 315 students, 24 apprentices

CERN

- International organisation close to Geneva, straddling Swiss-French border, founded 1954
- Facilities for fundamental research in particle physics
- 21 member states, 1 B CHF budget
- 3’360 staff, fellows, students, apprentices, …
- 11’000 users

1954: 12 Member States

Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

2’512 staff members, 540 fellows, 315 students, 24 apprentices

6’700 member states, 1’800 USA, 900 Russia, 236 Japan, …

CERN

"Science for peace"

- International organisation close to Geneva, straddling Swiss-French border, founded 1954
- Facilities for fundamental research in particle physics
- 21 member states, 1 B CHF budget
- 3'360 staff, fellows, students, apprentices, …
- 11'000 users

1954: 12 Member States

Members: Austria, Belgium, Bulgaria, Czech republic, Denmark, Finland, France, Germany, Greece, Hungary, Israel, Italy, Netherlands, Norway, Poland, Portugal, Slovak republic, Spain, Sweden, Switzerland, United Kingdom
Candidate for membership: Romania
Associate member: Serbia
Observers: European Commission, India, Japan, Russia, Turkey, UNESCO, United States of America
Numerous **non-member states with collaboration agreements**

2'512 staff members, 540 fellows, 315 students, 24 apprentices

6'700 member states, 1'800 USA, 900 Russia, 236 Japan, …

CERN – where the Web was born

CERN – where the Web was born

CERN – where the Web was born


LHCb
ATLAS
CERN Meyrin
CERN Prévessin
SPS 7 km
SUISSE
FRANCE
CMS
ALICE
LHC 27 km

LHC – Large Hadron Collider

- Proton-proton collider

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator
- Run 1 until early 2013

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator
- Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator
- Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade
- Restart early 2015 at 6.5 + 6.5 TeV

LHC – Large Hadron Collider

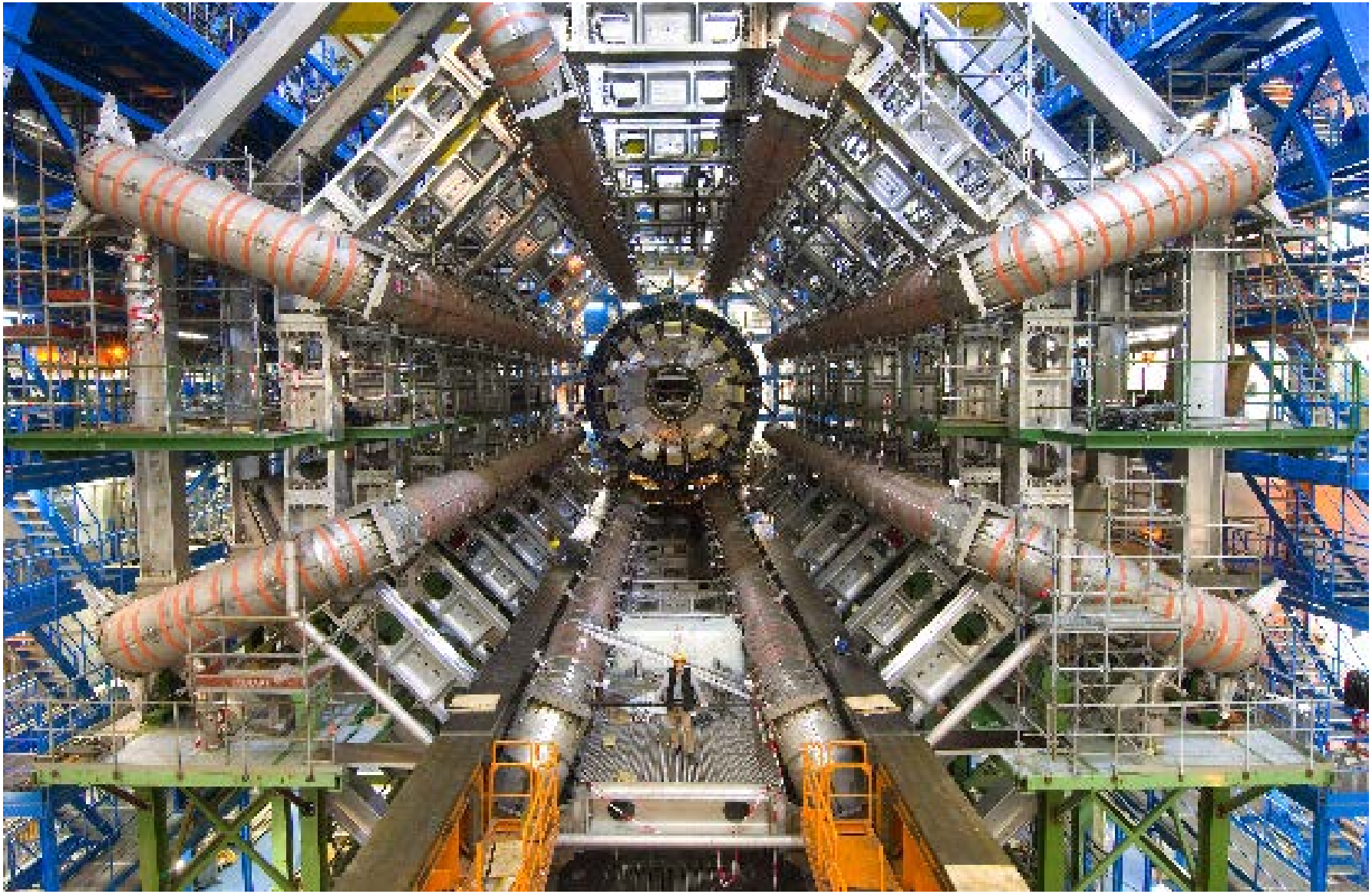
- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator
- Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade
- Restart early 2015 at 6.5 + 6.5 TeV

LHC – Large Hadron Collider

- Proton-proton collider
- 27 km circumference
- Started operation in 2010 with 3.5 + 3.5 TeV, continued in 2011, 4 + 4 TeV in 2012
  - World's most powerful particle accelerator
- Run 1 until early 2013
- Now in Long Shutdown 1 – machine upgrade
- Restart early 2015 at 6.5 + 6.5 TeV

Four Large Detectors

- ATLAS, CMS, ALICE, LHCb
- Some ATLAS facts:
  - 100 million channels
  - 25 m diameter, 46 m length, 7’000 tons
  - 3’000 scientists (including 1’000 grad students)
  - 40 MHz collision rate
  - Run 1: 300 Hz event rate after filtering
- All LHC experiments: 30 PB in 2012, 100 PB in total

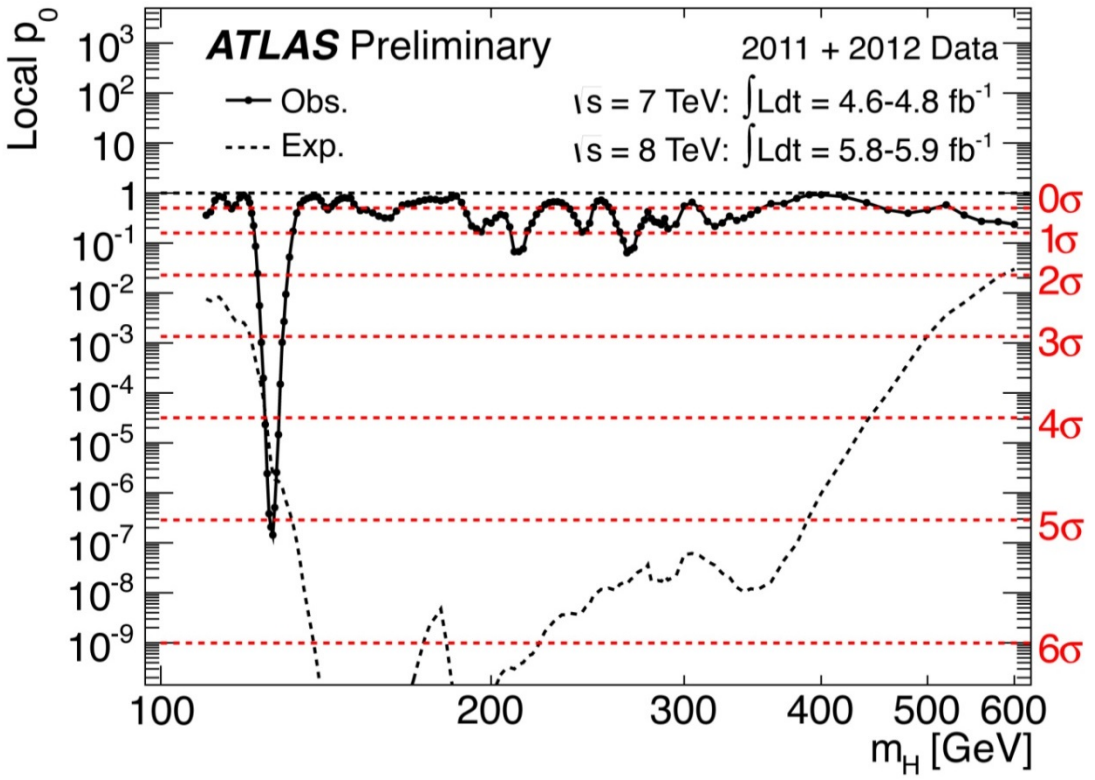


Results so far

- Many… the most spectacular one being

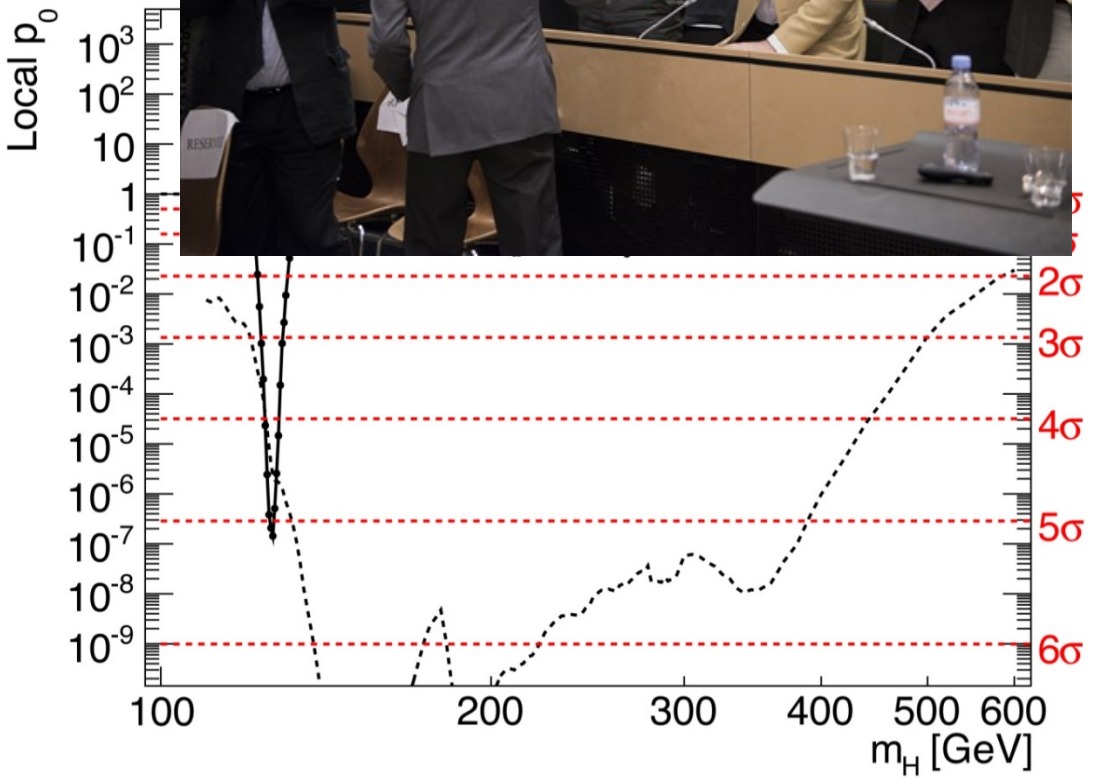
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”



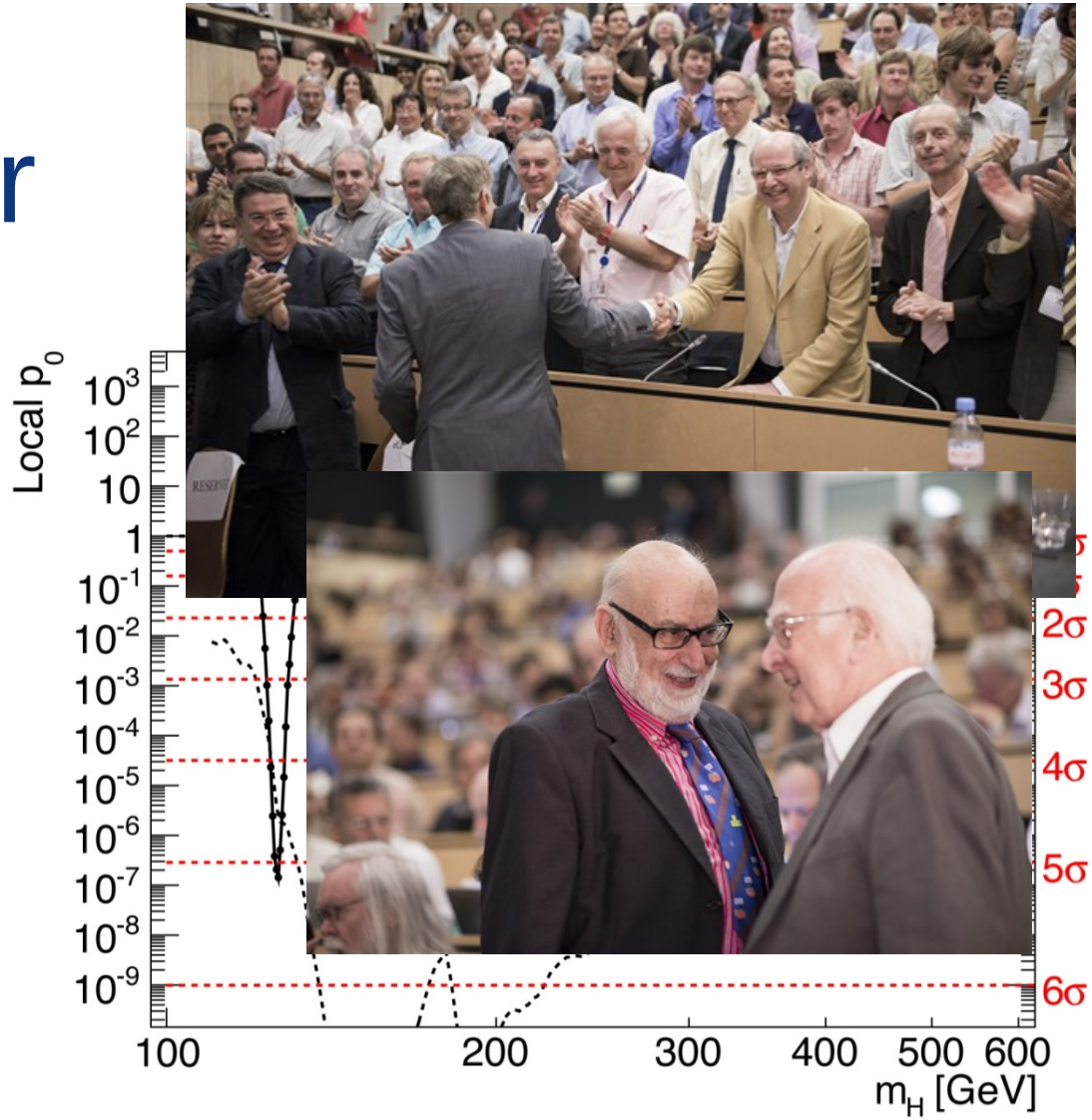
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”



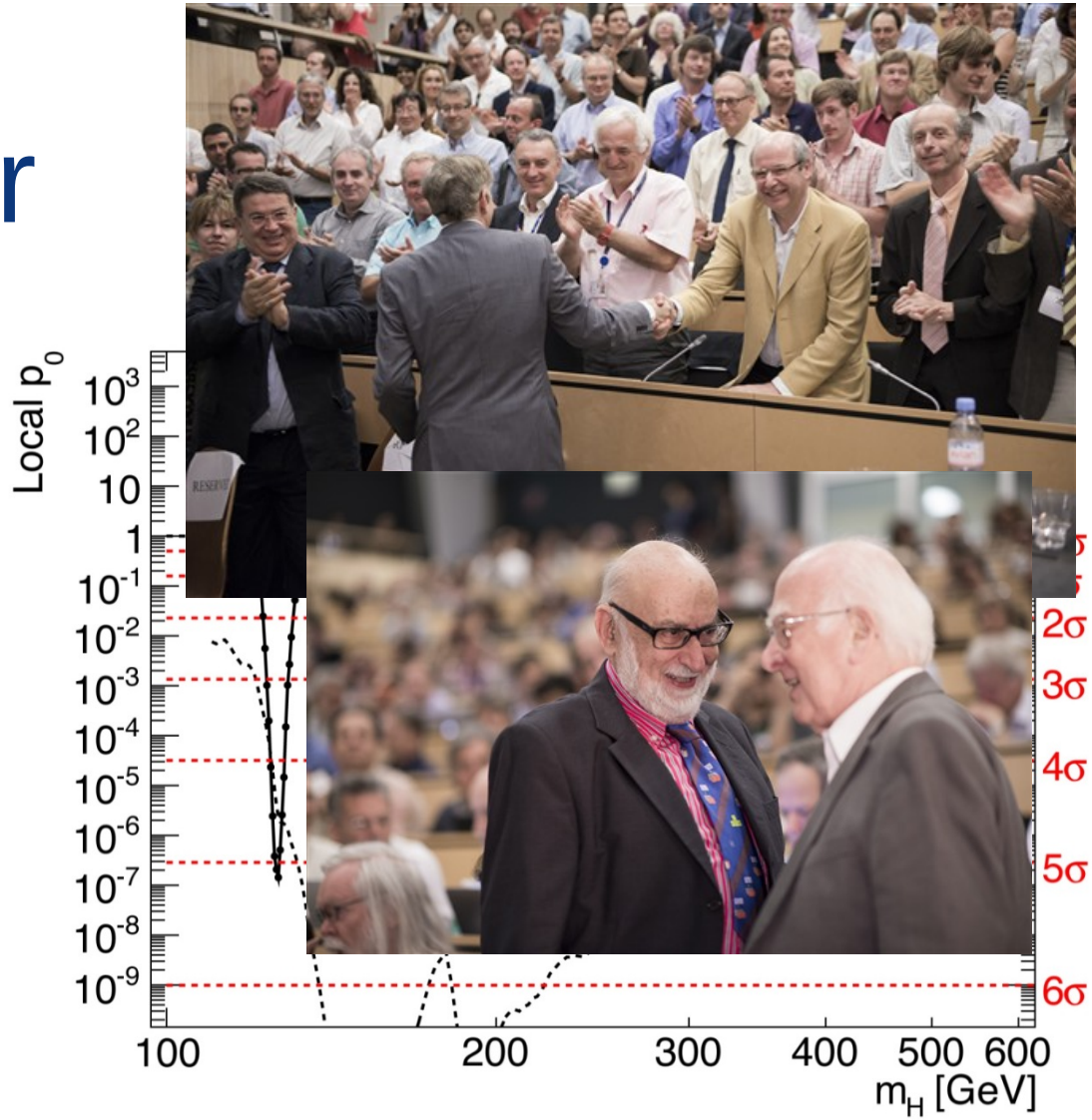
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”



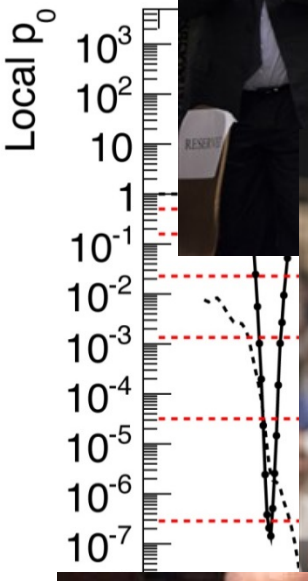
Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”
- March 2013: The particle is indeed a Higgs boson



Results so far

- Many… the most spectacular one being
- 04 July 2012: Discovery of a “Higgs-like particle”
- March 2013: The particle is indeed a Higgs boson
- 08 Oct 2013 / 10 Dec 2013: Nobel price to Peter Higgs and François Englert
  - CERN, ATLAS and CMS explicitly mentioned



Data Handling

- 30 PB per year demand 100’000 processors
- World-wide LHC Computing Grid (WLCG): 150 computer centres all around the world
  - CERN as Tier-0 largest and most important

CERN Data Centre

Machine Inventory 2 Apr 2014 Wed 15:31:20

Service information

full name: **Machine Inventory**

short name: dcbynum

group: IT-CF-FPP

site: CERN

email: Data.Centre.By.Numbers@cern.ch

web site: http://hwcollect.cern.ch/

Service availability (more)

availability:

percentage: 100%

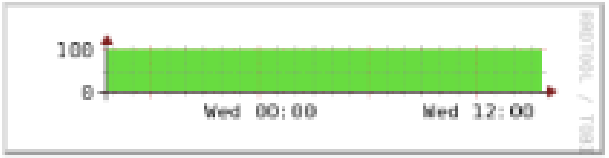
status: **available**

last update: 15:30:32, 2 Apr 2014 (48 seconds ago)

expires after: 1440 minutes

rss feed with status changes

availability in the last 24 hours (more):



Additional service information (more)

Number of 10GB NICs:	3,073
Number of 1GB NICs:	19,234
Number of cores:	97,696
Number of disks:	73,872
Number of memory modules:	67,591
Number of processors:	18,452
Number of servers:	10,718
Total disk space (TiB):	108,149
Total memory capacity (TiB):	342

Part of (subservice of):

none / not declared

Subservices

none / not declared

Clusters, subclusters and nodes

none / not declared

Depends on

none / not declared

Depended on by

none / not declared

CERN Data Centre

Machine Inventory 2 Apr 2014 Wed 15:31:20

Service information

full name: **Machine Inventory**
short name: dcbynum
group: IT-CF-FPP
site: CERN

email: Data.Centre.By.Numbers@cern.ch
web site: http://hwcollect.cern.ch/

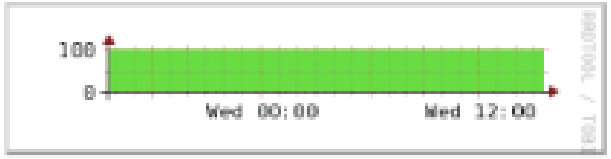
Service availability (more)

availability:
percentage: 100%
status: available

last update: 15:30:32, 2 Apr 2014 (48 seconds ago)
expires after: 1440 minutes

rss feed with status changes

availability in the last 24 hours (more):



Additional service information (more)

Number of 10GB NICs:	3,073
Number of 1GB NICs:	19,234
Number of cores:	97,696
Number of disks:	75,872
Number of memory modules:	67,591
Number of processors:	18,452
Number of servers:	10,718
Total disk space (TiB):	108,149
Total memory capacity (TiB):	342

Part of (subservice of):

none / not declared

Subservices

none / not declared

Clusters, subclusters and nodes

none / not declared

Depends on

none / not declared

Depended on by

none / not declared

CERN Data Centre

Machine Inventory 2 Apr 2014 Wed 15:31:20

Service information

full name: **Machine Inventory**

short name: dcbynum

group: IT-CF-FPP

site: CERN

email: Data.Centre.By.Numbers@cern.ch

web site: http://hwcollect.cern.ch/

Service availability (more)

availability:

percentage: 100%

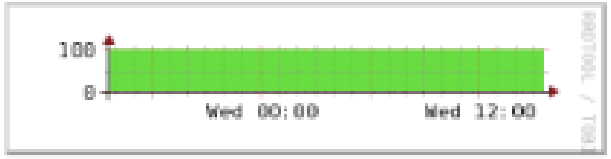
status: available

last update: 15:30:32, 2 Apr 2014 (48 seconds ago)

expires after: 1440 minutes

rss feed with status changes

availability in the last 24 hours (more):



Additional service information (more)

Number of 10GB NICs:	3,073
Number of 1GB NICs:	19,234
Number of cores:	97,696
Number of disks:	73,872
Number of memory modules:	67,391
Number of processors:	18,452
Number of servers:	10,718
Total disk space (TiB):	108,149
Total memory capacity (TiB):	342

Part of (subservice of):

none / not declared

Subservices

none / not declared

Clusters, subclusters and nodes

none / not declared

Depends on

none / not declared

Depended on by

none / not declared

Data Handling – Future (1)

- LHC Run 2 (starting 2015): higher energy
  - 8 TeV to 13 TeV
- More interesting collisions to retain after filtering
  - ATLAS: 300 Hz up to 1 kHz or more
- Moore's law helps, but not sufficient
- Large effort to improve software efficiency
  - Exploit multi-threading, new instruction sets, …
- Still need factor 2 in terms of cores, storage, …

Data Handling – Future (2)

- Challenges for CERN-IT
  - Where? CERN data centre full (3.5 MW)
  - How? No additional personnel
    - Traditional way of running centre does not scale

CERN Tier-0 Extension (1)

- Following open tendering process: Wigner research centre in Budapest/Hungary

CERN Tier-0 Extension (2)

Agile Infrastructure Project (1)

- Challenges:
  - Handle 15’000 servers
    - Part of them not (easily) physically accessible
  - Quickly react to changing requirements
    - Deploy new services and servers within hours rather than weeks or months
- Not possible with previous structure
  - Mostly vertical view – service managers responsible for (almost) entire stack
  - Strong coupling of services with hardware life-cycle
  - Configuration and monitoring: home-made developments of 10 years ago
    - Very successful at the time, but increasingly brittle
    - Lack of support for dynamic host creation/deletion

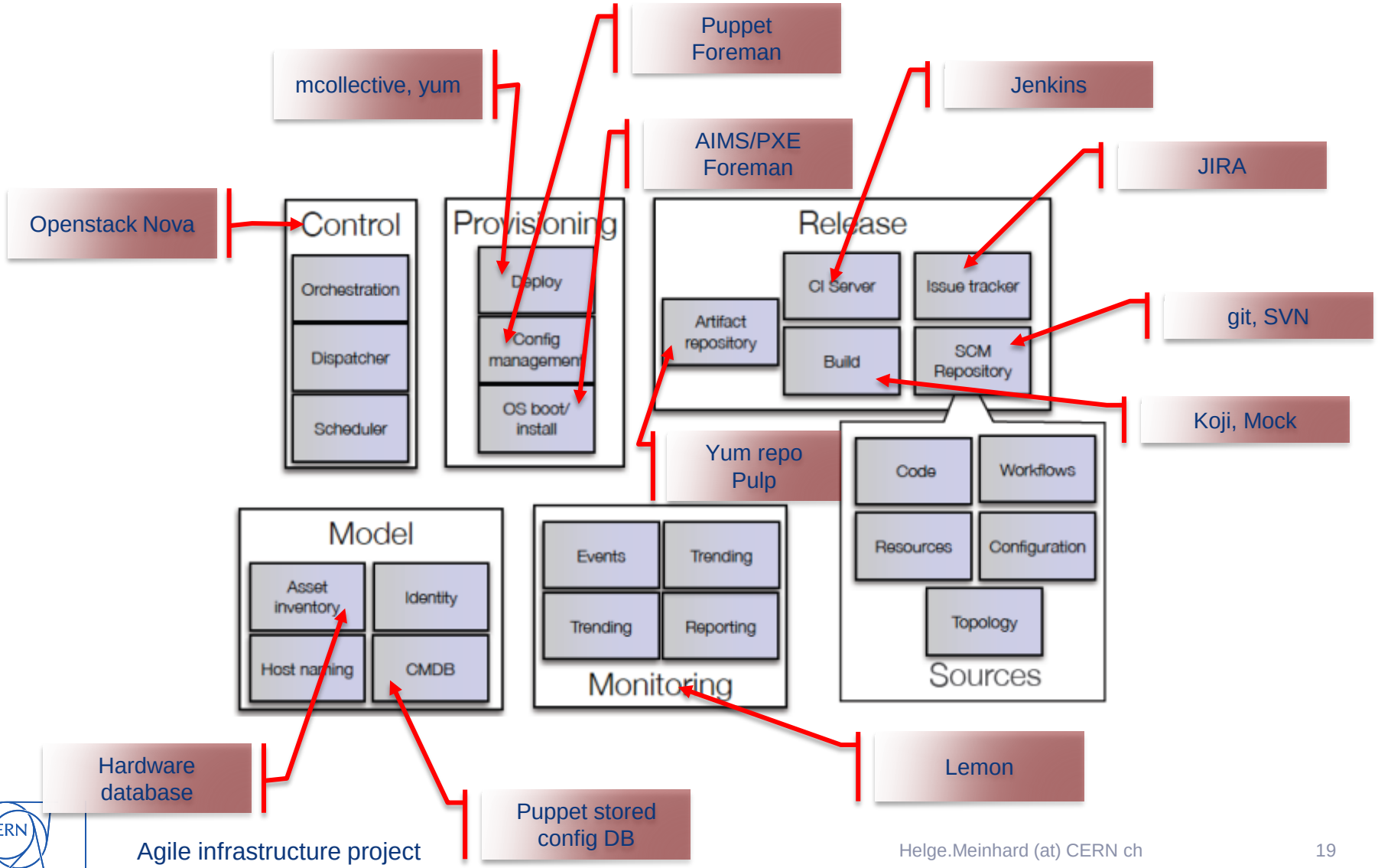
Agile Infrastructure Project (2)

- Launched a project in 2012 to move to a more ***horizontal*** approach
  - Services
  - Configuration
  - Software installation
  - Hardware
- Aim: improve
  - Operational efficiency
  - Resource efficiency
  - Responsiveness
- ***Virtualisation*** is key for ‘horizontalisation’
- Virtualisation + agility + provisioning = ***cloud***

Agile Infrastructure Project (3)

- Guiding principles
  - CERN is not special (any more) – join the community
  - ‘Tool-chain’ approach
    - Break problem space down into small pieces
    - Quickly identify suitable solution for each one – good enough, not necessarily best one
      - Be prepared to promptly reconsider if needed…
    - Minimal glue
  - ‘Devops’ approach – eat your own (dog food | medicine)
  - Preference for open-source solutions
    - Benefits all parties

Agile Infrastructure Project (4)



Agile Infrastructure Project (5)

- Key areas
  - Private cloud services
  - Configuration
  - Monitoring
  - Registration, burn-in, software installation
  - Scheduling and accounting
- ~ 15 people in the core team, mostly part-time
- Massive deployment started in 2013

Private Cloud Services (1)

- Earlier smaller-scale (production!) projects with Xen/KVM, Hyper-V; SCVMM, OpenNebula
- Chose *Openstack* for the project
  - Very large, active community with attractive mix of company support and user influence
  - Moving fast – new functionality becoming available very rapidly
- Followed Openstack releases
  - Essex, Folsom, Grizzly, Havana (migration completed)
  - Watching out for Icehouse
- Using Nova (multiple cells), Glance, Cinder, Keystone, Ceilometer, …

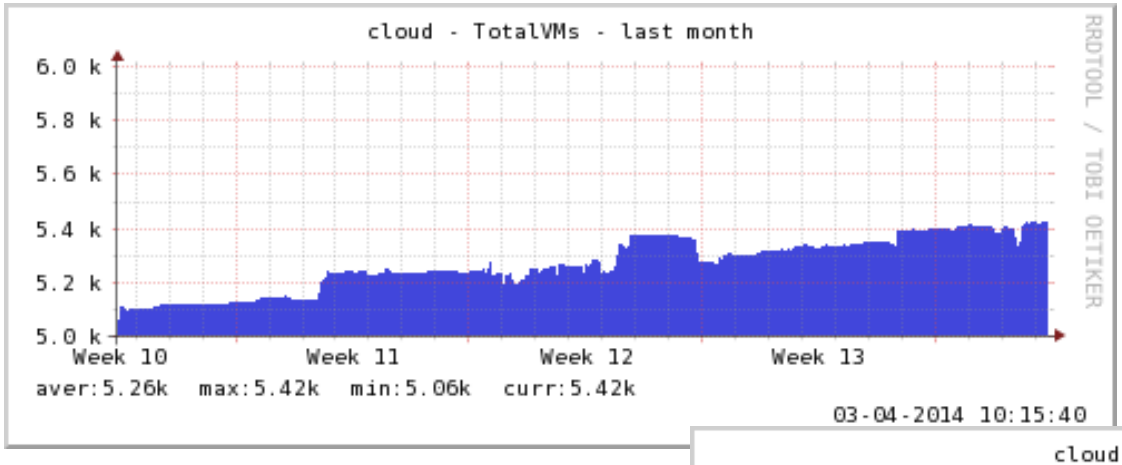
Private Cloud Services (2)

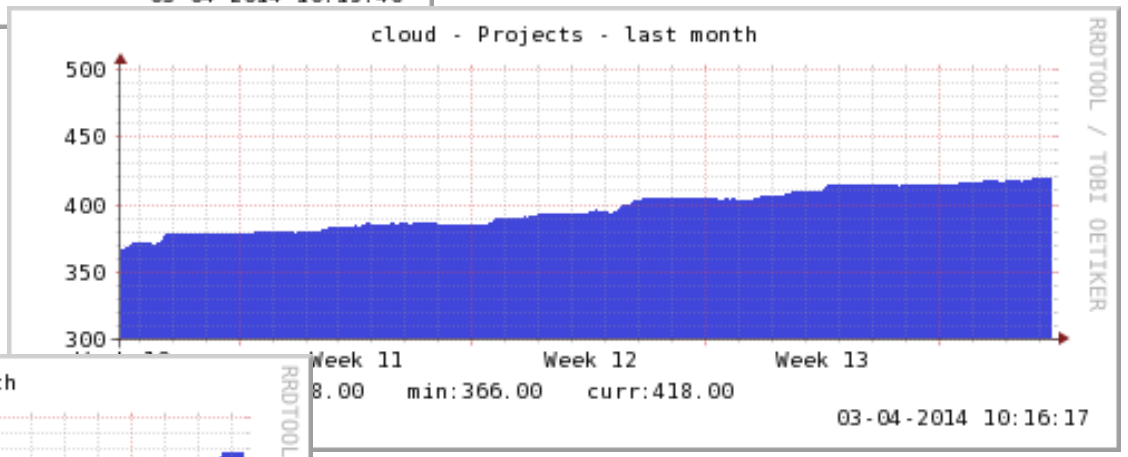
- Fully integrated with Active Directory, CERN's network data base, account and quota management, …
- Production service – documentation, support lines, notifications, operator and sysadmin support, …
- Focused on 'cattle' use-case first, now addressing 'pets'
- Linux (KVM) and Windows (Hyper-V) as hypervisors and guests
- Target: >= 90% of CERN's servers

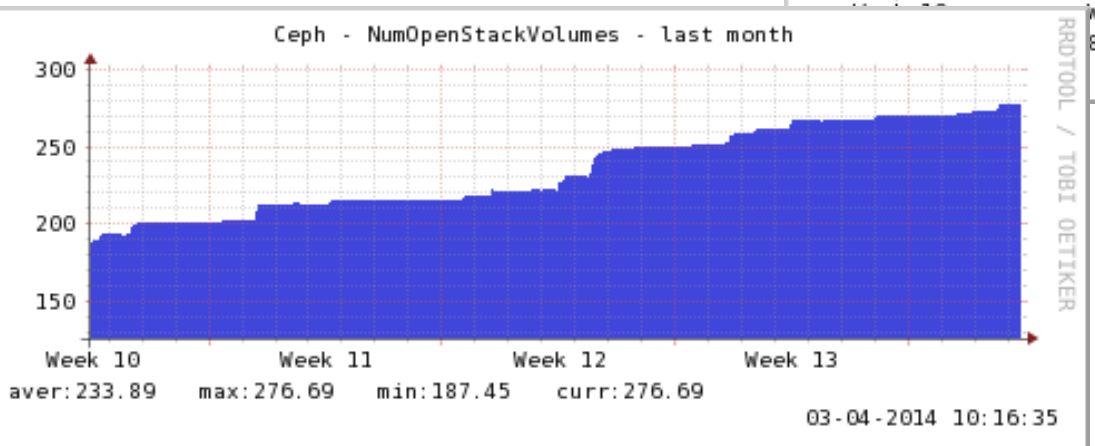
Private Cloud Services (3)

- Volume service (requirement for live migration) deployed (Cinder-based)
  - Linux: Large (3 PB) CEPH installation as backend
- As of 03-Apr-2014: 2’615 hypervisors, 5’515 VMs
  - Including major part of large-scale batch service (4’500 physical servers total)
  - Rapid growth (100 or more hypervisors per week)

Private Cloud Services (4)







Configuration (1)

- Dynamic cloud requires dynamic configuration system
- Previous system (Quattor) not dynamic and scalable enough, high maintenance
- Chose *Puppet* as the centre of configuration services
- In addition: PuppetDB, Foreman, mcollective, git
- Currently 17 Puppet servers (including 5 VMs), can be scaled out
- Serving 8’216 hosts (physical and virtual) as of 02-Apr-2014; 80…150 Git commits to configuration files per week

Configuration (2)

- We know how to scale out further – targeting 50k hosts
- Strong emphasis on QA – all services to have machines in QA (10% level) for configuration and software installation
- Currently being addressed
  - Security improvements, including handling of secrets
  - Workflow automation, continuous integration
- Some tools written ourselves (e.g. state management)

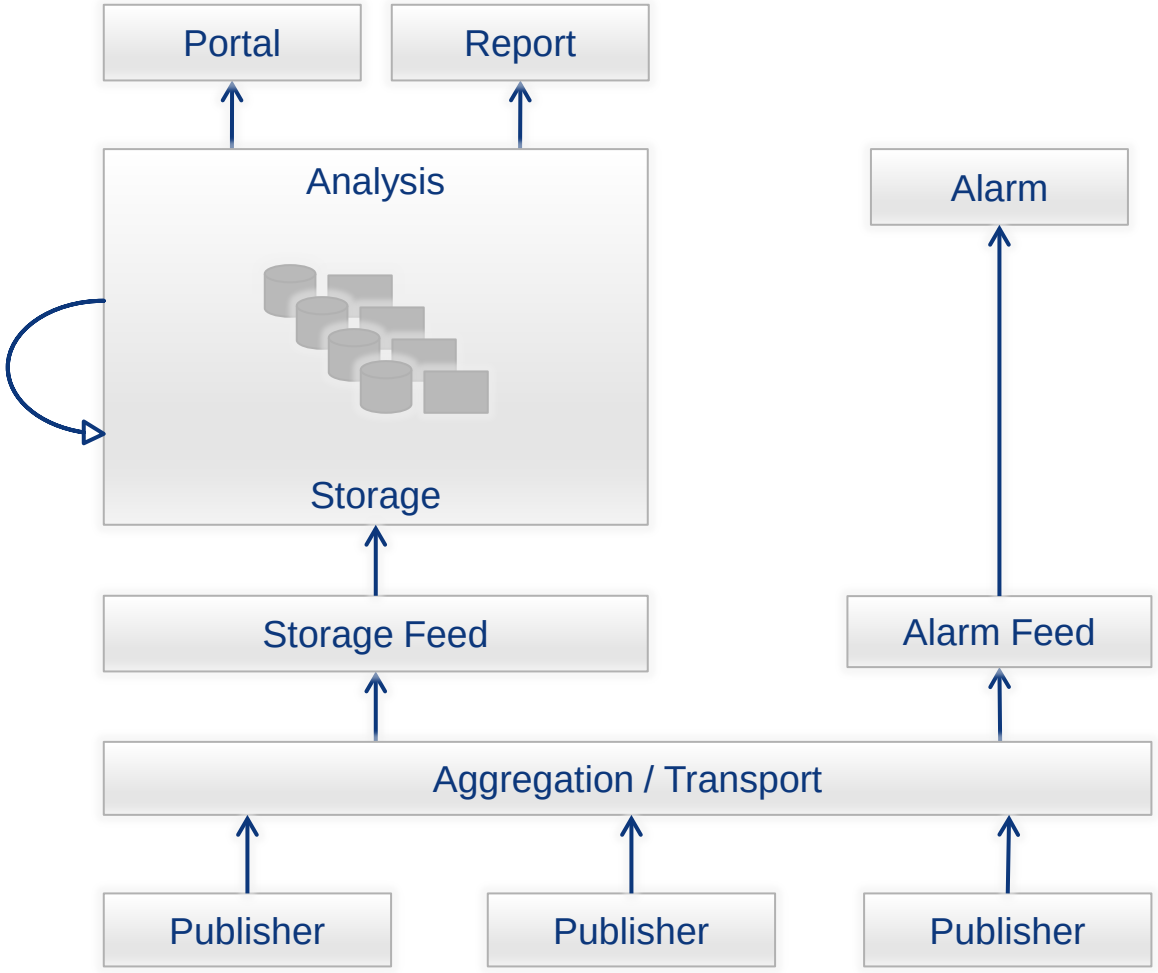
Configuration (3)

- Most visible part for many service managers
  - Training sessions
  - Improvements to monitoring configuration services
- Migration out of old tools is a serious issue
  - Maintenance of old tools takes person-power
  - Co-existence of tool sets confusing
  - Agreed target date for complete shutdown: 31 October 2014

Monitoring (1)

- Way too many independent (i.e. partly overlapping, partly different) activities at CERN
- Need for common architecture supporting dynamically adding probes, data stores, data consumers
- Addressing both exception and performance monitoring
- Huge investment into probes to be preserved

Monitoring (2)



Monitoring (3)

- Technologies chosen:
  - Hadoop
  - ElasticSearch and Kibana
  - Flume
  - ActiveMQ
- Producers/probes
  - Probes from previous home-grown system
  - SCOM, Spectrum
  - Syslogs, application logs
  - …

Monitoring (4)

- Notification (alarm) system in production, linked with ticketing system
- Central dashboard
- O (10) of GB of monitoring data per day
- Being worked on: more dashboards, analytics

Extending to Public Clouds: The Helix Nebula project

- Aim: develop and exploit cloud computing infrastructure
  - For various European IT-intense research projects (CERN, ESA, EMBL, …)
  - Extend to enterprises, governments and society later
  - Infrastructure provided by various commercial and public European cloud providers
- (Slides courtesy of Bob Jones/CERN)

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

PIC port d'informació científica

To improve the speed and quality of research for finding surrogate biomarkers based on brain images

Suppliers

Atos CGI CloudSigma CSA cloud security alliance DANTE www.dante.net egi indra interoute SAP sixsq SWITCH T-Systems terradue 2.0 the SERVER LABS the IT architects

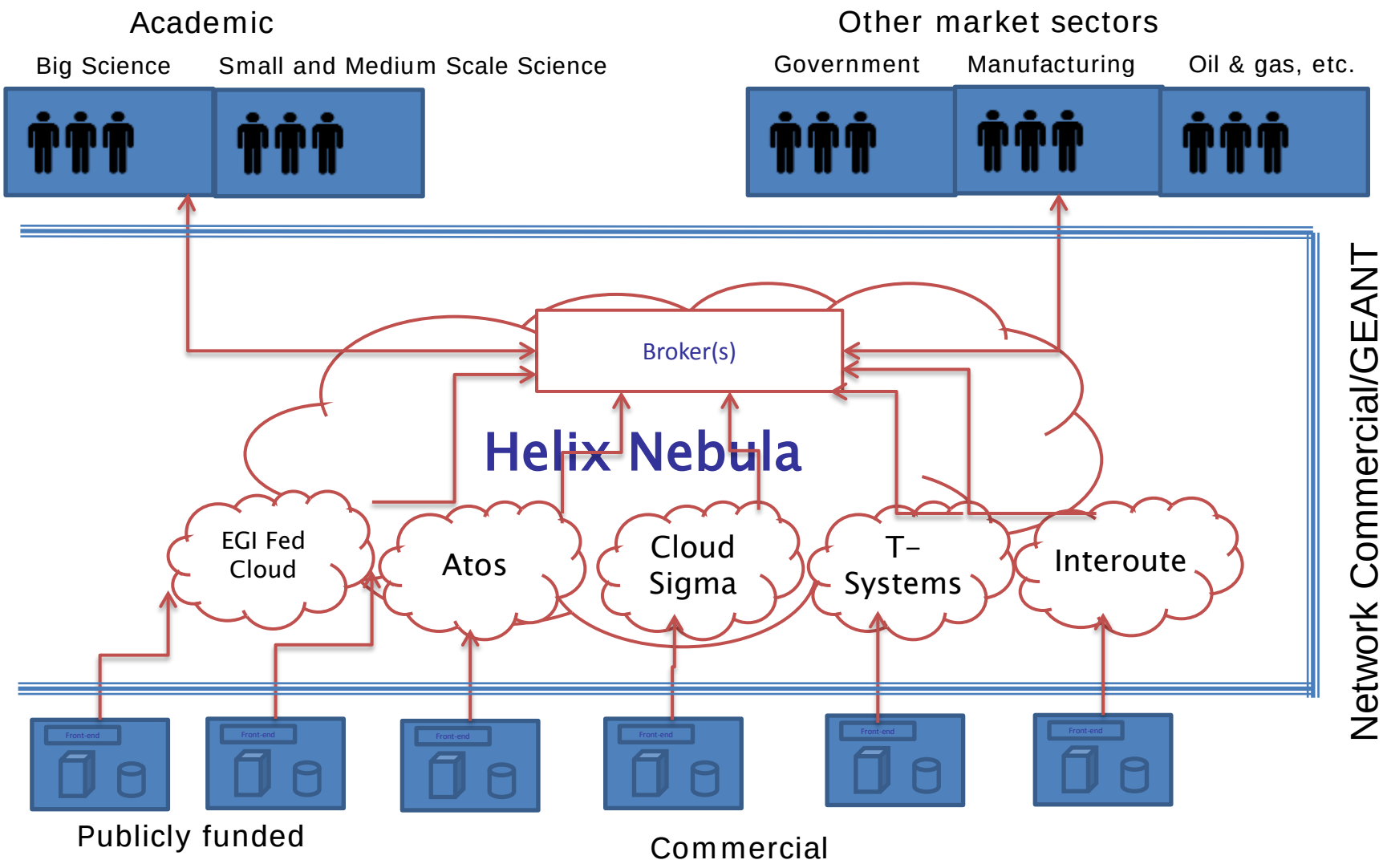
Adopters

AWST Capgemini CNRS ECMWF Ifremer OpenNebula.org orange Business Services THALES Telefónica it

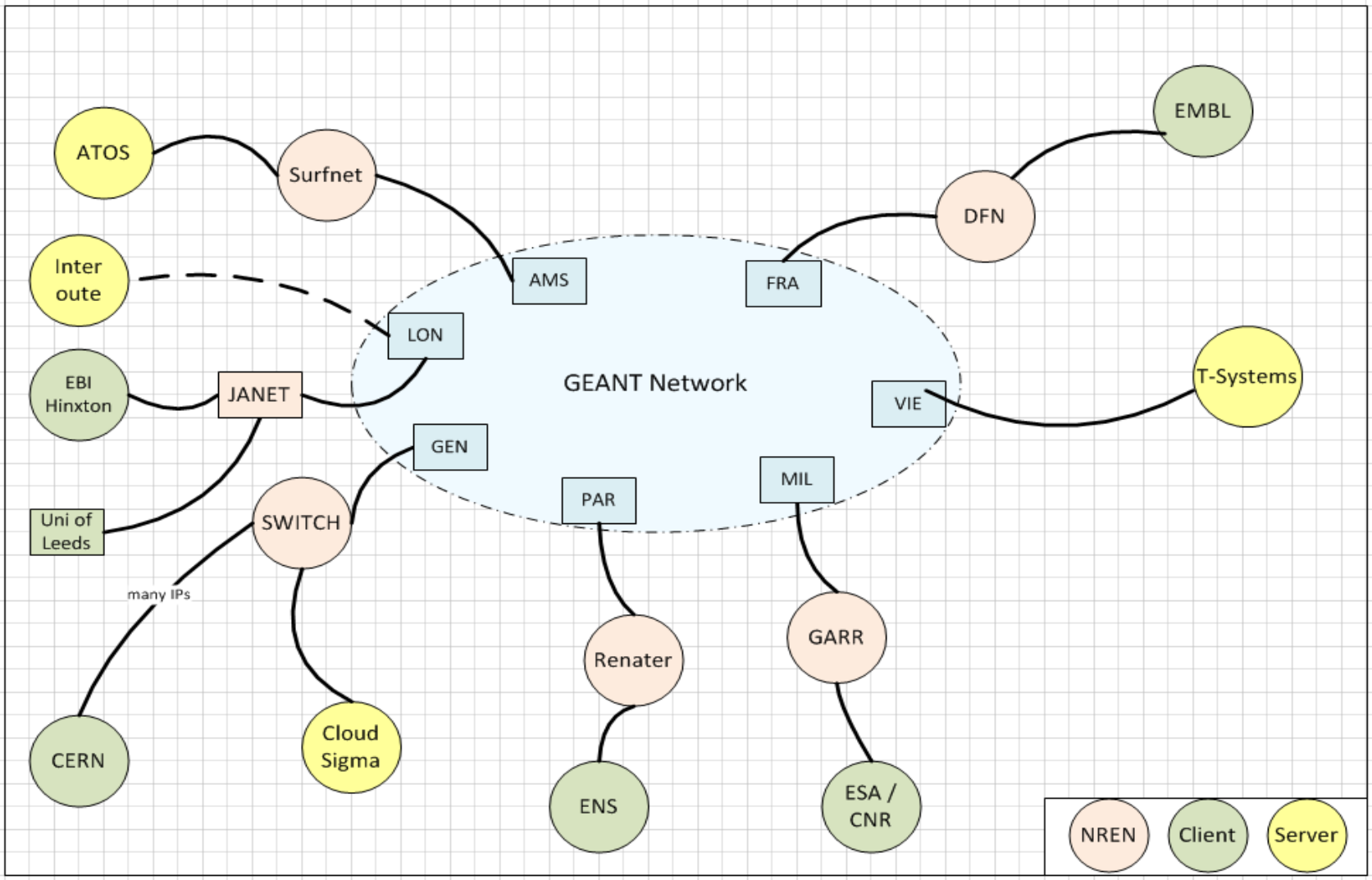
http://www.helix-nebula.eu
contact@helix-nebula.eu

@HelixNebulaSC
HelixNebula.TheScienceCloud

Hybrid Public-Private Cloud Model



Topology


ATOS
Surfnet
EMBL
DFN
Inter oute
AMS
FRA
LON
GEANT Network
EBI Hinxton
JANET
VIE
T-Systems
GEN
MIL
PAR
Uni of Leeds
SWITCH
many IPs
Renater
GARR
CERN
Cloud Sigma
ENS
ESA / CNR
NREN
Client
Server

Building the hybrid cloud

Testing the public-commercial cloud interoperability

- Deploy the ESA/CNES/DLR SuperSites Exploitation Platform on EGI Fed Cloud and then the CERN CMS/ATLAS flagship use cases across commercial suppliers and EGI Federated Cloud via a Blue Box broker

- Use the same evaluation criteria adopted for deployment on commercial cloud service suppliers

EGI Federated Cloud

Task Force

- Launched in Sep 2011
- 70 members from 40 institutions and 13 countries

Pre-production test-bed:

- 14 resource centres actively providing resources (900 cores, 16 TB storage)
- 30 active users from structural biology, linguistics, ecology, space science, software engineering

http://go.egi.eu/cloud

Helix Nebula Marketplace (HNX)

- **Builds upon the work of the Helix Nebula Initiative and EC support action**
- **Supported by European cloud providers**
- **Integrates with existing e-Infrastructures to form a hybrid cloud Market Place and reach out to Europe's research communities**
- **Trusted cloud services through compliance with EU regulations and legislation**
- **Simplifies procurement process across multiple services providers**

sixsq Atos CloudSigma

the SERVER LABS the IT architects · T-Systems · interoute · CGI · egi

hnx.helix-nebula.eu/

Conclusions (1)

- IT world rapidly changing
- CERN can’t follow all changes…
  but every now and then we’re catching up
- AI project is very challenging…
  but also motivating and exciting
- Have gone a long way already…
  but still a lot of work to do
- Huge amounts of to-do lists with technical items…
  but cultural change at least as demanding

Conclusions (2)

- The IT aspects are very interesting and challenging…
  but the final objective is physics discoveries at the LHC at its design energies as of 2015

- We're convinced to be on a good way!

- Stay tuned for more physics results from LHC…

www.cern.ch