Novembre 20, 2012

CC-IN2P3: Experiments and Servic

Workshop on Data Preservation,
Marseille

CC-IN2P3 within the French scientific framework

Main scientific organization in France

Nation… for Nu… …cs and …ysics

CC-IN2P3 federates the main computing resources
For :
- High energy physics
- Nuclear physics
- Astroparticle physics

Resources Mutualisation

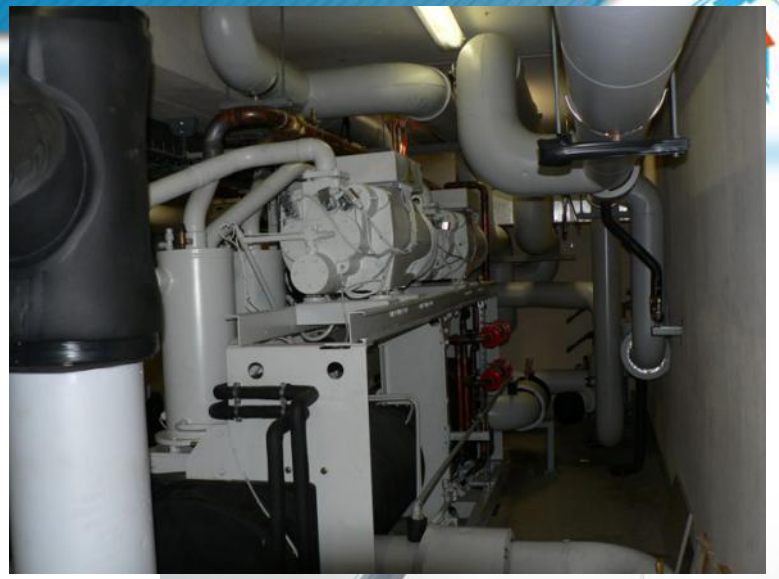
Dedicated Computing Center

+ some opening to other sciences

DSM cea Irfu — Atomic Energy Commission

Manpower: 76 people – 60 IT engineers
Budget: 10 - 12 M€ including salaries

Computing Centre: Evolution

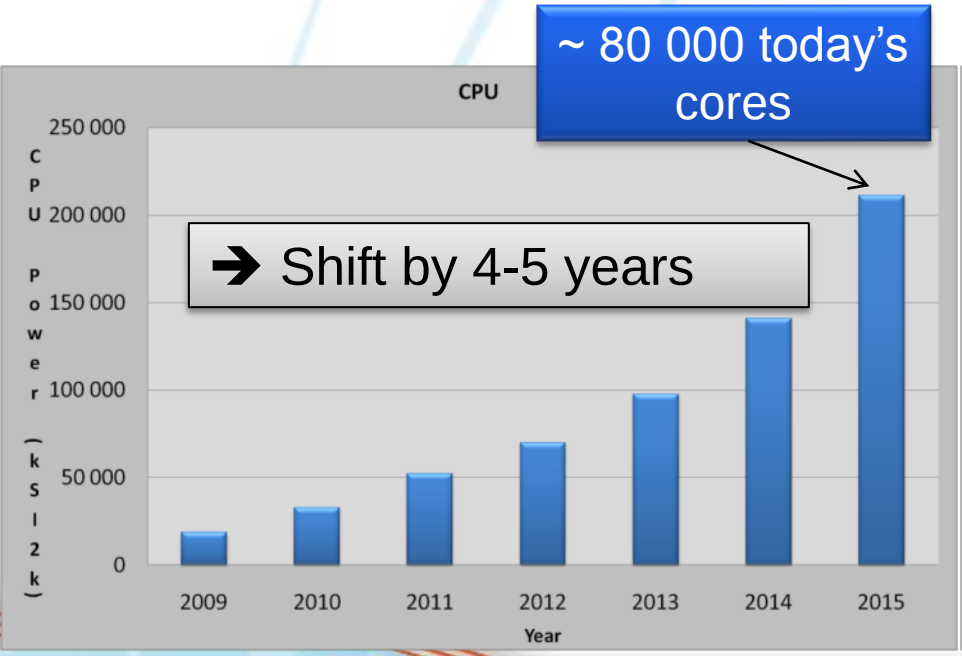


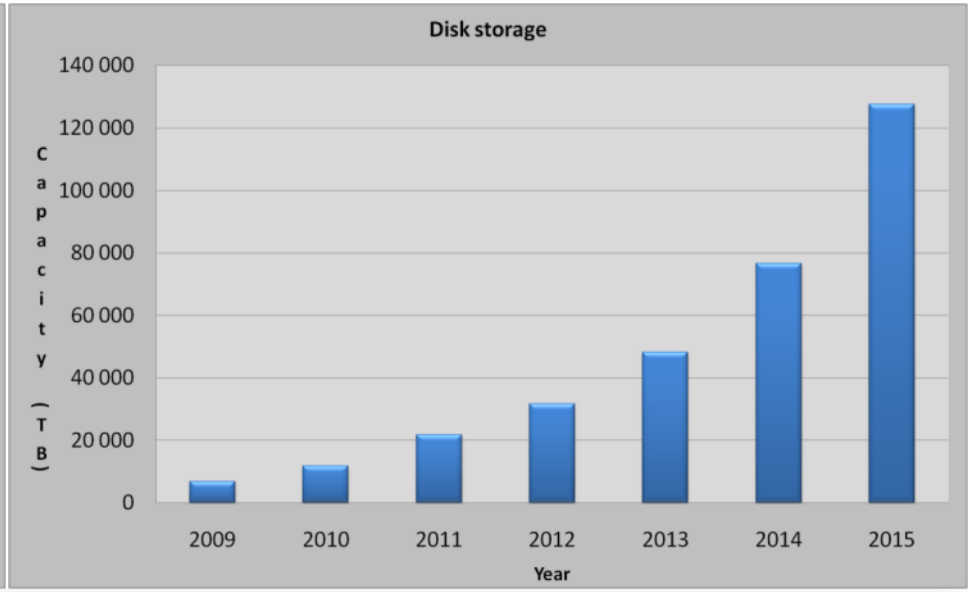
CC IN2P3 is installed since 1986 in Villeurbanne, Lyon
Big investments during 2004-2010 period
But the infrastructure has reaching a limit incompatible with our long term scientific commitments

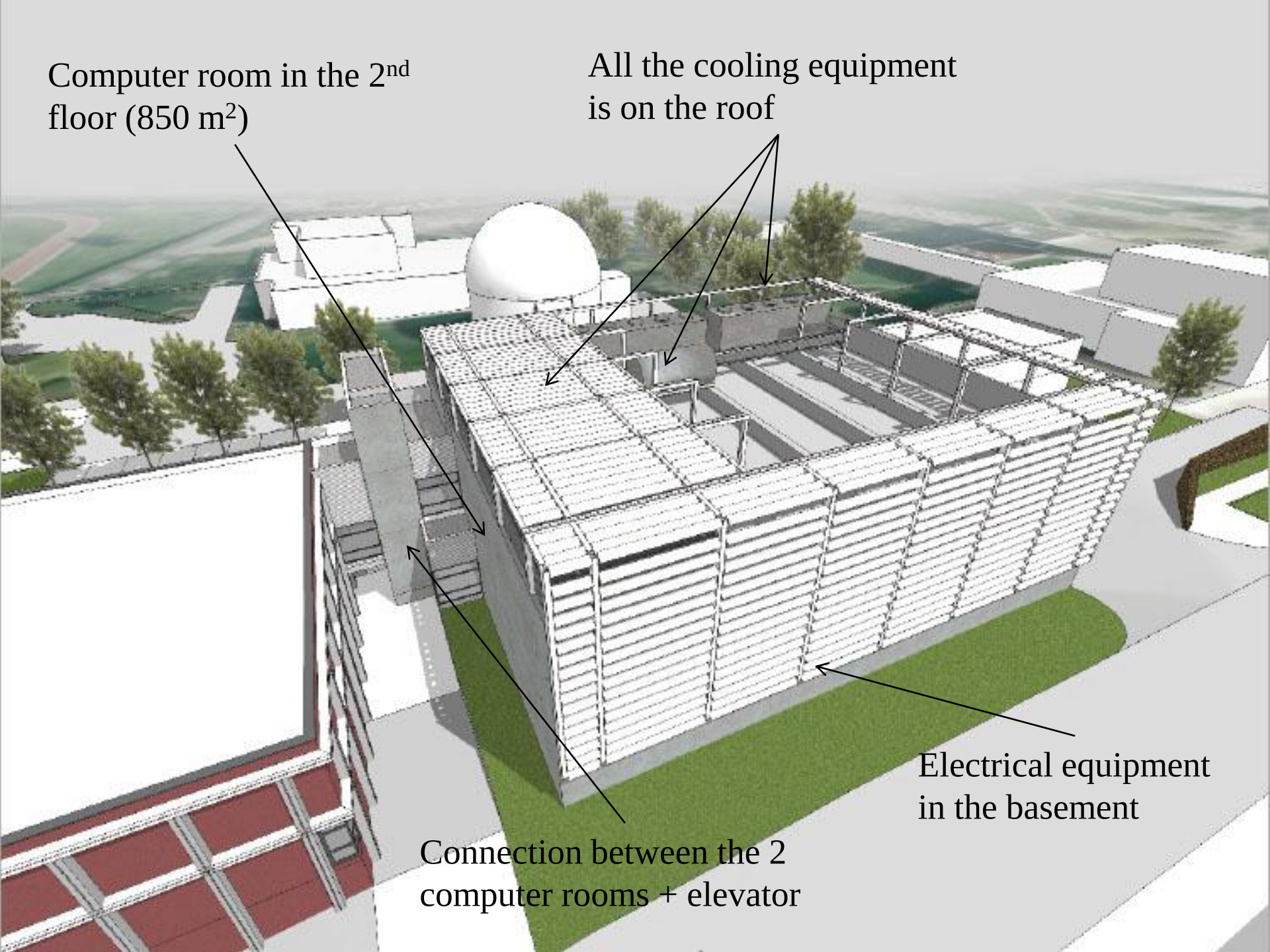
➔ Design goals for new datacenter

The design parameters used to define the size of the new datacenter has been the following:

- Serve ~40 experiments with "standards needs"
- Fulfill LHC computing commitments and provide first class analysis capability
- **Face the very significant growth of the Astroparticle community needs: LSST, EUCLID**
- Add some capacity for network, services, etc.






Computer room in the 2nd floor (850 m2)
All the cooling equipment is on the roof
Electrical equipment in the basement
Connection between the 2 computer rooms + elevator

The new computer room is inaugurated in September 2011: with its innovative and modular design is an enormous advantage for IN2P3 scientific projects

LHC

HESS

Auger

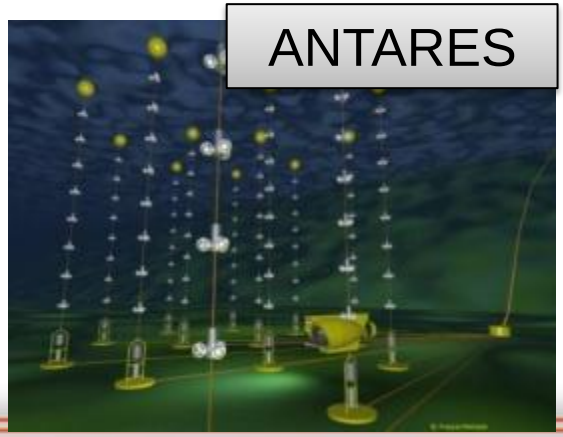
AMS

~70 groups are using CC-IN2P3 resources

Planck

Supernovae

Fermi



ANTARES

VIRGO

1.7 MHS06.hours this year

Services

Mass Storage System (HPSS v7)

4 tape libraries SL8500 (40 000 slots)

17 PB stored

Tape technology: T10K-A (phased out next year), T10K-B, T10K-C (ramping up)

Dynamic staging : key point ➔ scheduling of staging by TReqs(*) (traffic can go up to 60-80 TBs / day between clients and HPSS).

More and more hidden from the end user:
dCache, xrootd, iRODS, SRB used as a front end.

Credit : J.Y. Nief – CC-IN2P3

(*) TReqs : Derived from the ERADAT system from BNL

Storage services

- Front end to HPSS for direct I/O access (~ 10 PBs of disk space):
  - dCache: mostly used by LCG.
  - xrootd: used by HEP (including LCG with Alice) and non HEP experiments.
  - Increasing role for xrootd within LCG at CC-IN2P3
- Data management (data sharing + multi sites federations):
  - iRODS: powerful tool. Front end to HPSS and unlimited number of IT and storage services. 4 PBs of data so far.
  - Used for HEP, astro, biomedical apps and Arts & Humanities data management.

DAS servers: Sun X4540 (being phased out) and Dell R510 + MD1200 (increasing) (54 usable TB / server)

Parallel File System

Currently using GPFS – 1.2 PB

Widely used by our experiments – very important for Astroparticle high throughput applications (Supernovae – Planck – etc.)

We expect an even increasing role of // FS in the future
LSST or Euclid usage..

For the cloud: going to be used as a central repository for virtual machines

We are looking for a potential replacement to GPFS (SONAS, Panassas, Isilon, Fluid File System …)

Storage services prospects

- Tapes still around for a long time!
  - Tape density will increase much more than disk and SSD in the next few years.
- Major challenges for the next 10 years:
  - Place of the « cloud storage » ? ➔ definition of the needs.
  - Exabyte scale: which technologies ?
  - Huge parallel file systems.
  - Breakthrough in databases technologies (SciDB like…).
  - Long term data preservation

CPU service

The current building block for CPU service is DELL C6220
4 servers in a 2U box
2 x 6 core Intel X5650 @ 2.6 GHz – 96 GB
351 HS06 / server

Total computing power @ CC-IN2P3 : 161 kHS06
~11 000 physical core ➔ 22 000 Hyper-Threads

Following a detailed assessment of the batch systems available on the market :
We selected Oracle Grid Engine (GE) in replacement of BQS our home grown batch system

➔ Now fully deployed

Quality management

We invested a lot during the past 2 years in quality management

Quality manager appointed

Computer room manager appointed

19 people trained to ITIL standards

New OTRS incident tracking system

Big effort in documentation and procedure writing

Next to come : CMDB and Disaster Recovery Plan

New Architectures: Partnerships

Test new hardware / new architectures

Publish results in White Papers

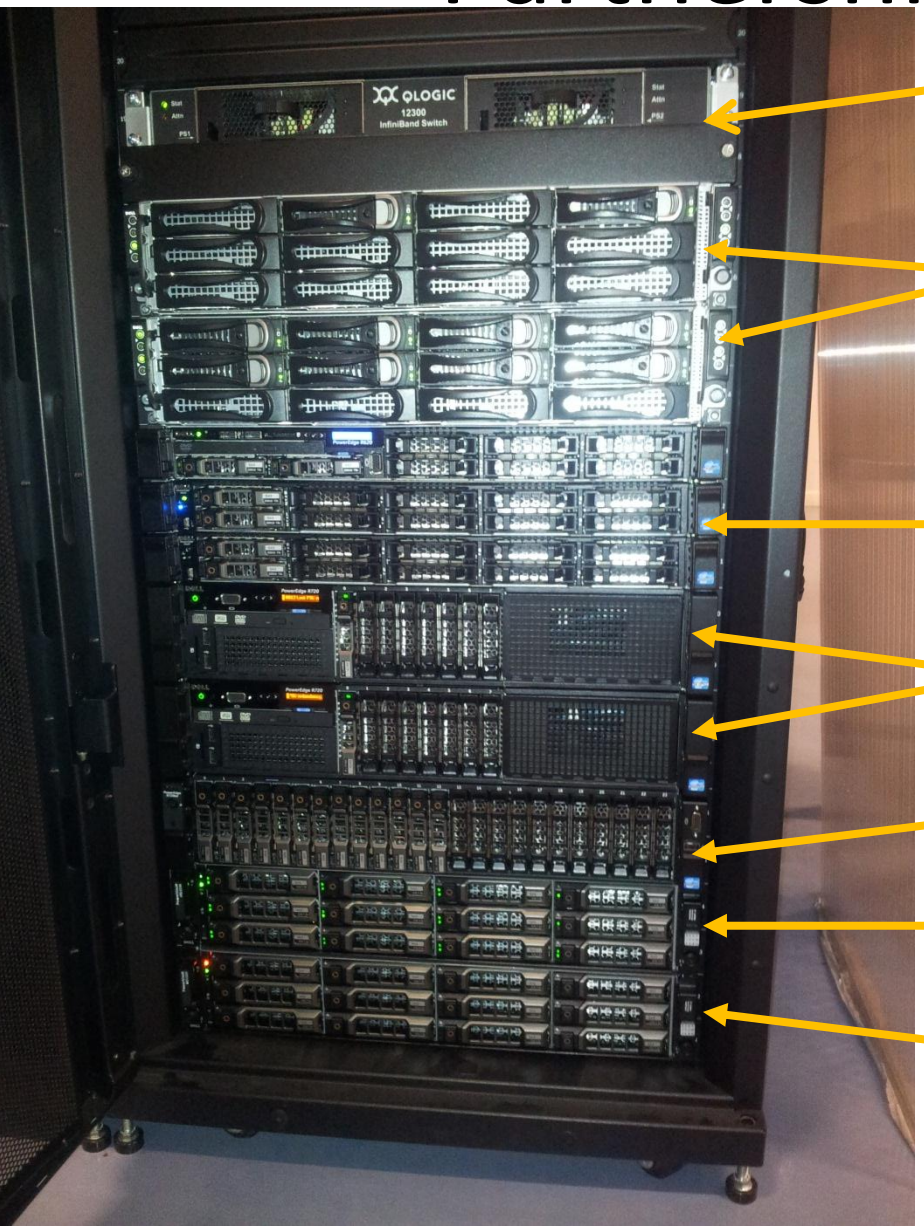
Partnership with DELL to test new Technologies:
GPU (Graphics Processing Unit):
LQCD – ALICE – CTA – …

Distributed file system : gLuster

Impact of LAN interconnect on performances : 1, 10 Gb/s - InfiniBand

Functionalities of specific hardware storage system : MD1200 / MD3200
Multipath access

Partnership with DELL



Network

C6220

Services : 3 x R620 + R410

R720 + 4 GPU Tesla M2090

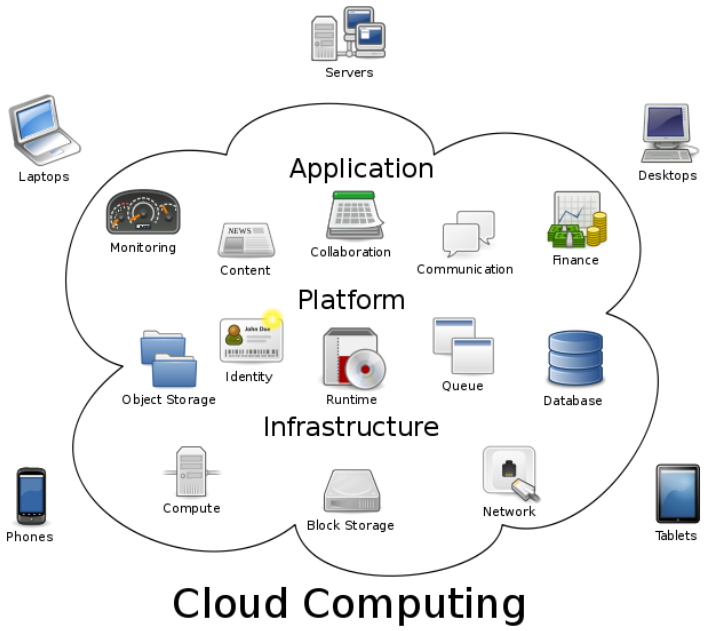
R720 XD – 12 to 24 HDD

MD3200

MD1200

Direct High Capacity Storage for sequential applications

Development of a Cloud Infrastructure



IaaS : To provide access to a large number of on demand virtualized resources.

PaaS : A toolkit for applications development, deployment and management on a virtual infrastructure

Applications

SaaS : The application itself can be executed on the virtual infrastructure through a web interface

Natural Hierarchy: IaaS – PaaS – SaaS

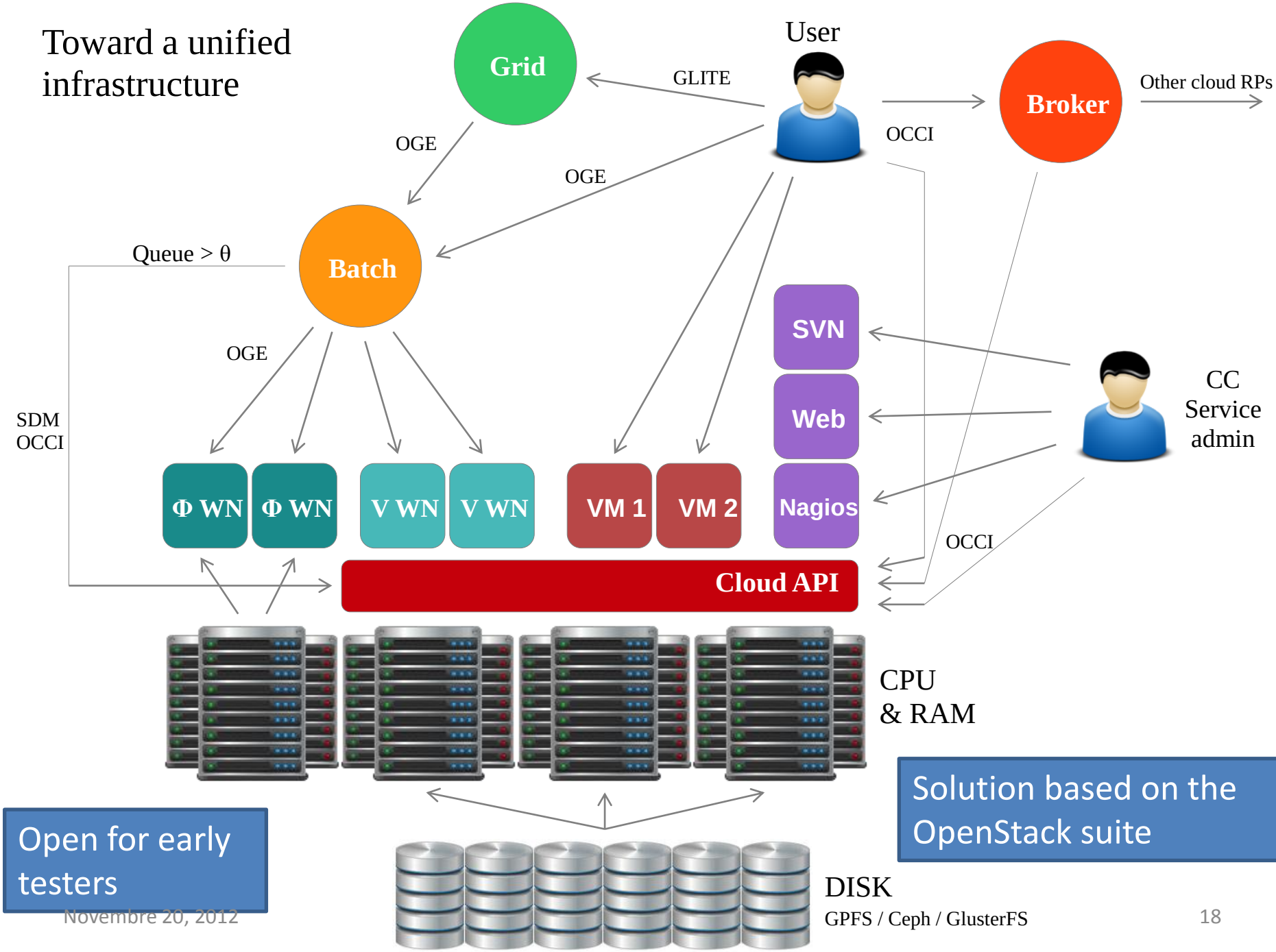
Grids will evolve in order to integrate virtualized resources

➔ French initiative to develop and academic Cloud at CC-IN2P3

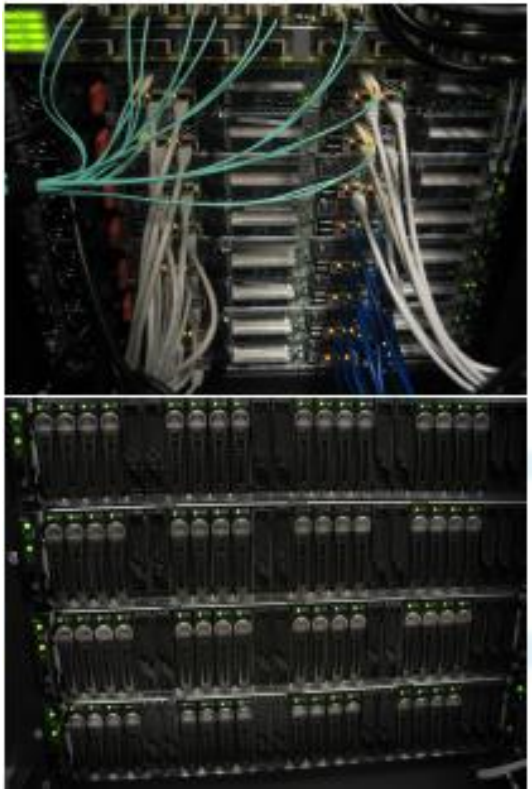
Initially focused on IaaS

Funded with the help of the French National Grid France Grilles

Toward a unified infrastructure



Today's Cloud setup @ CC



16 DELL Poweredge C6100
- 2 TB of local storage
- 96 GB RAM

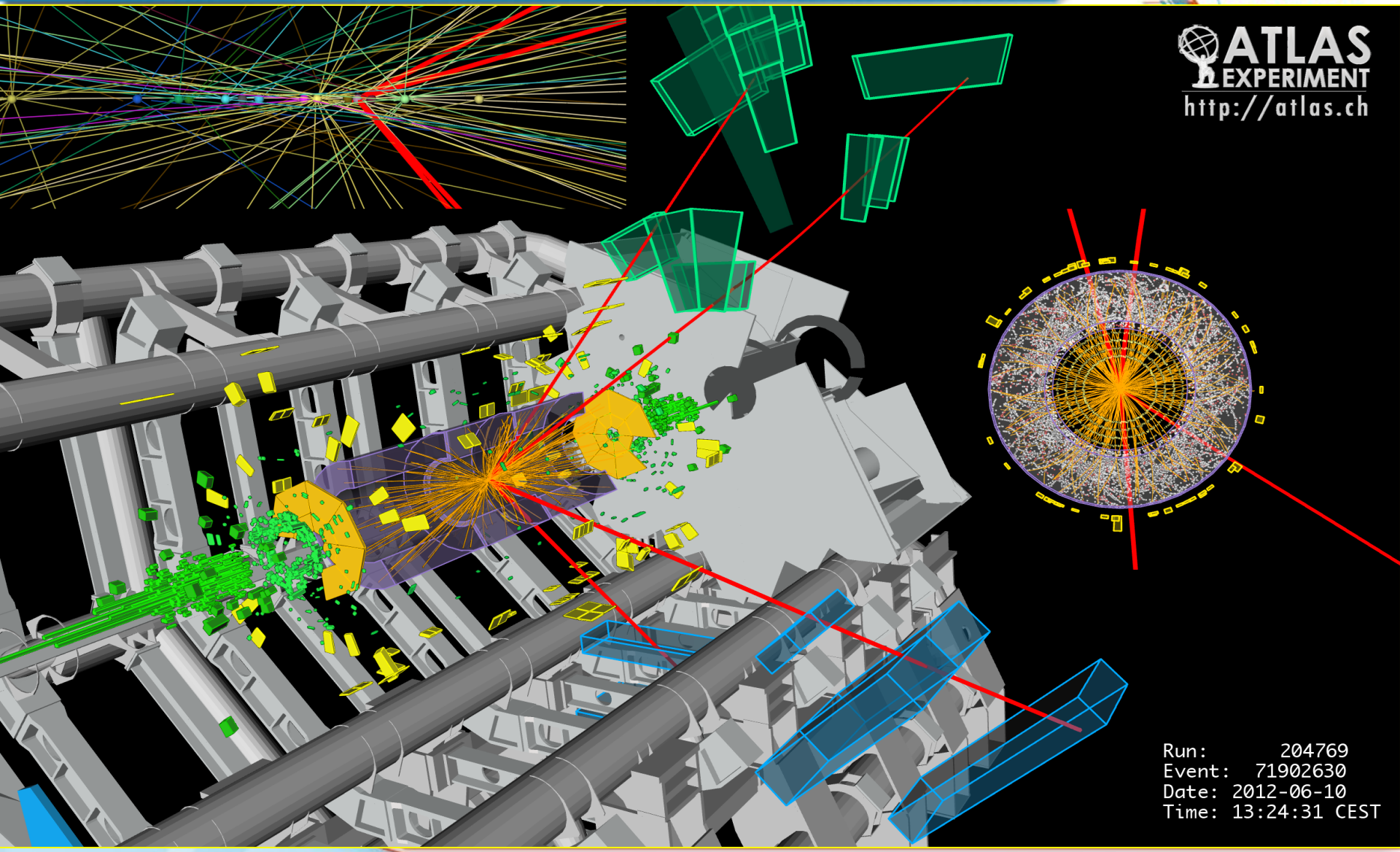
400 cores
10 Gb/s network
GPFS storage for image catalog

We are now considering several solutions for the storage service

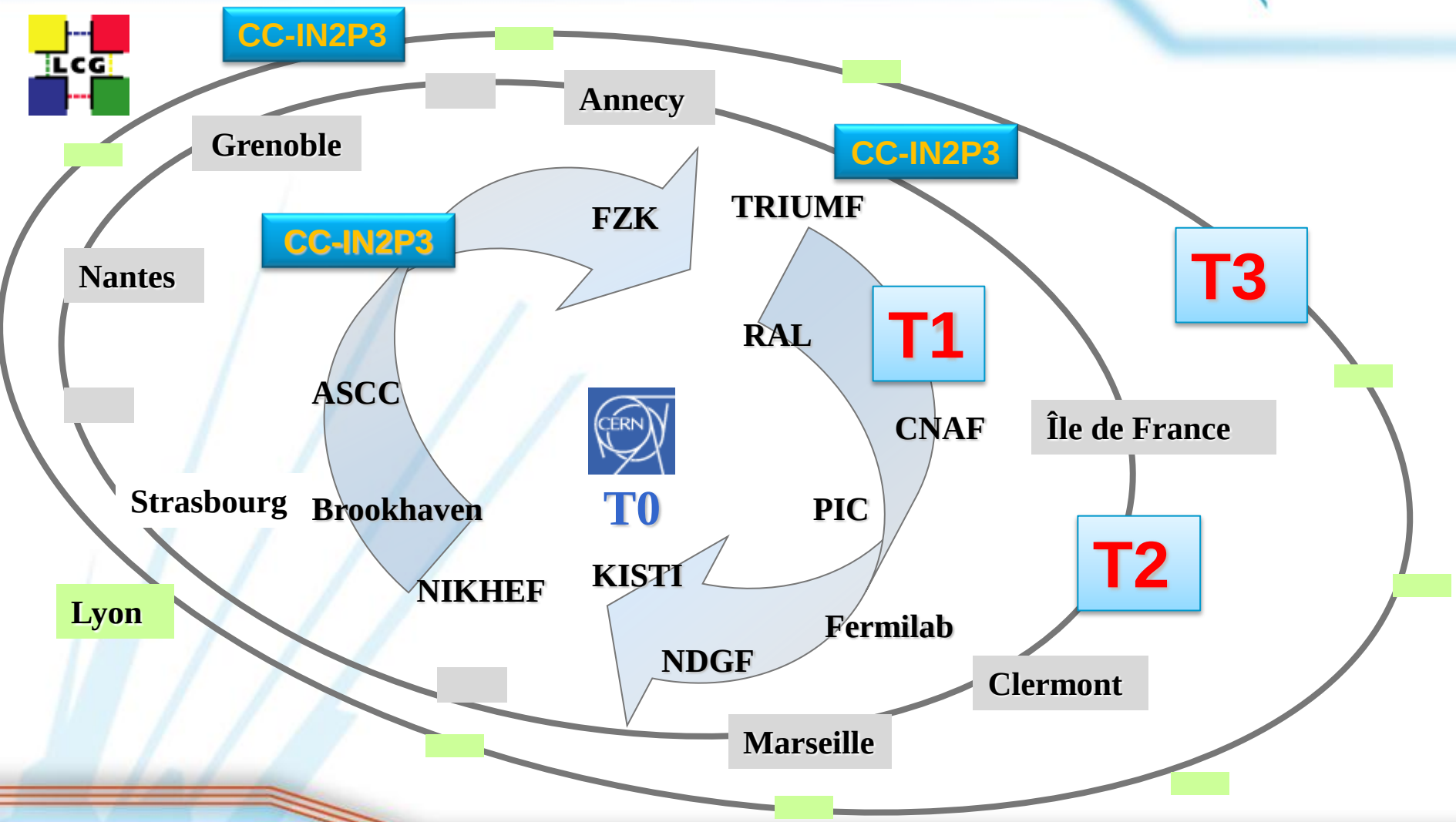
We would like to test IBM GPFS-SNC (Shared Nothing Cluster)
Use the local storage attached to the computing node to provide a global high capacity storage, potentially multi-PB

Experiments – Users

LHC @ CC-IN2P3


ATLAS
EXPERIMENT
http://atlas.ch
Run: 204769
Event: 71902630
Date: 2012-06-10
Time: 13:24:31 CEST

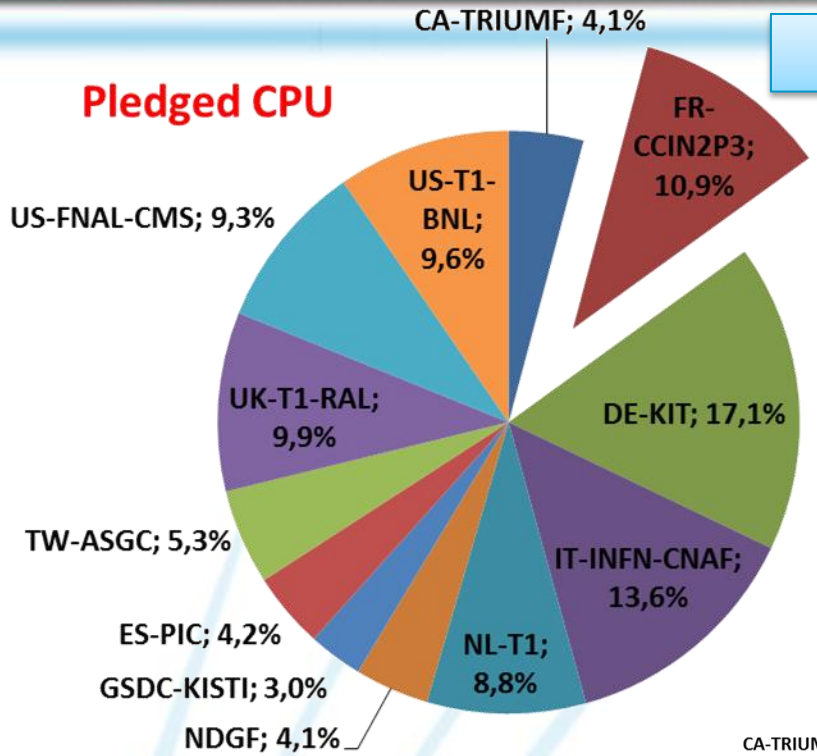
CC-IN2P3 within the global W-LCG organization



CC-IN2P3 within W-LCG

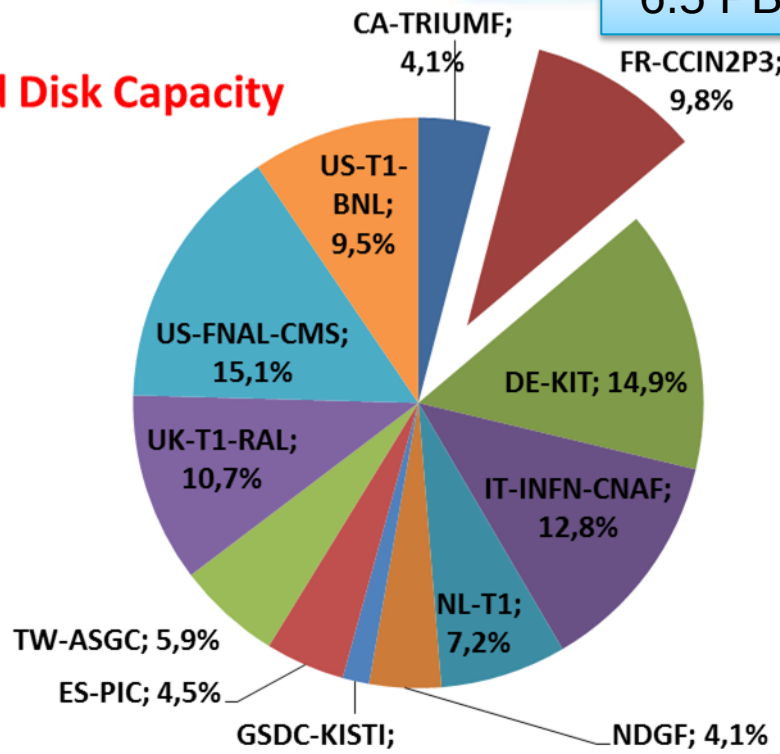
68.1 kHS06

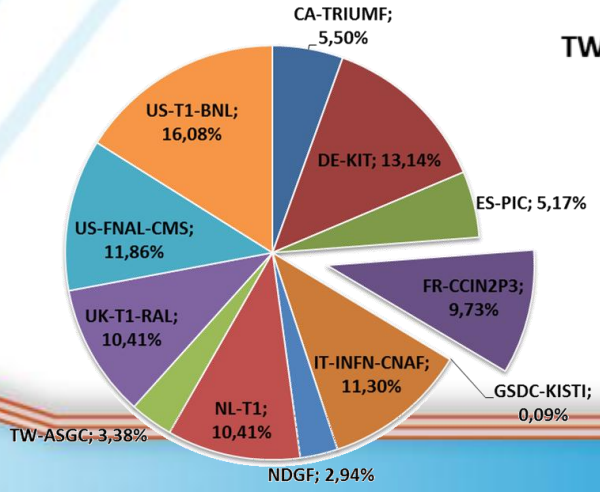
Pledged CPU



6.5 PB

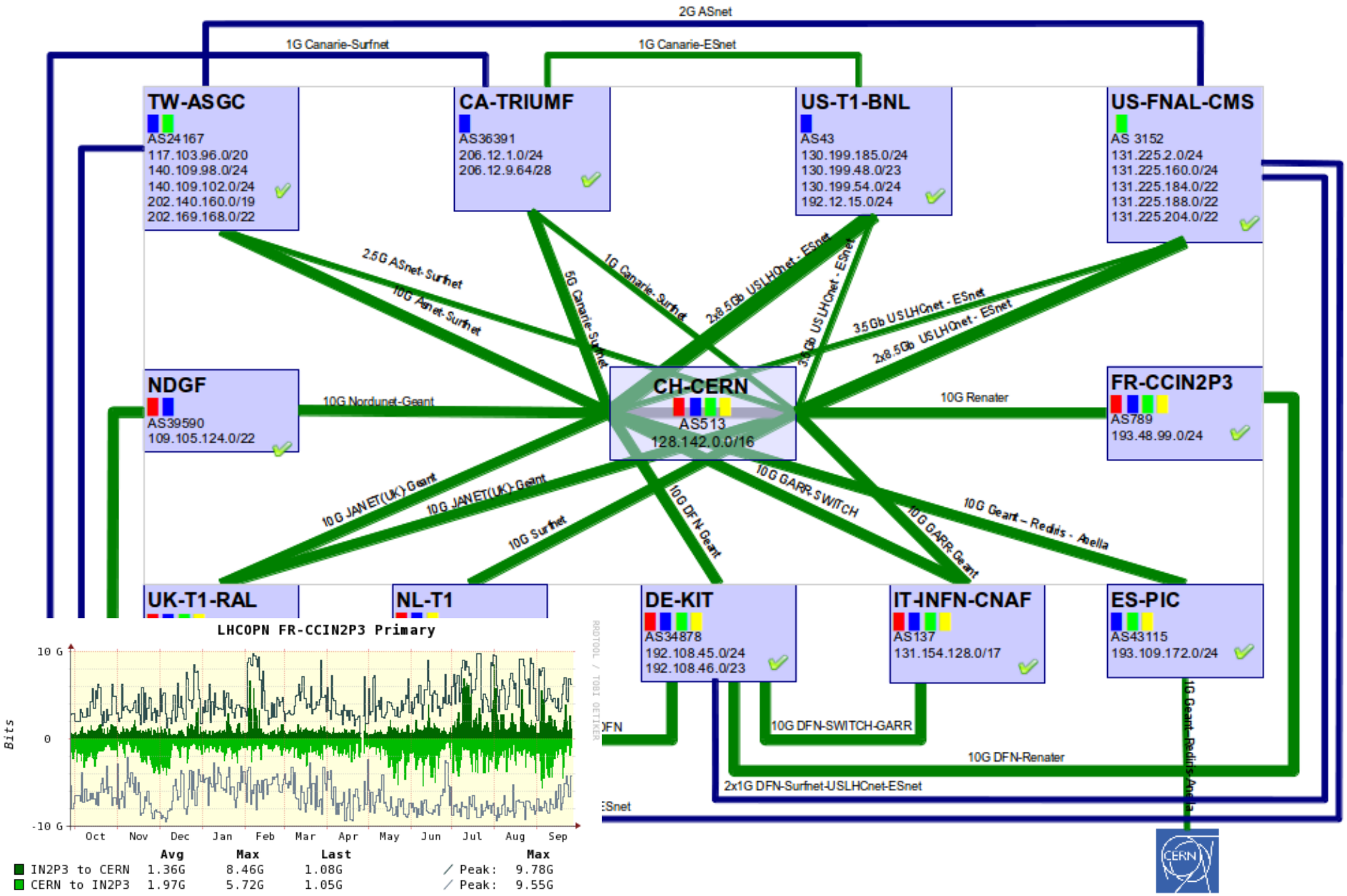
Pledged Disk Capacity





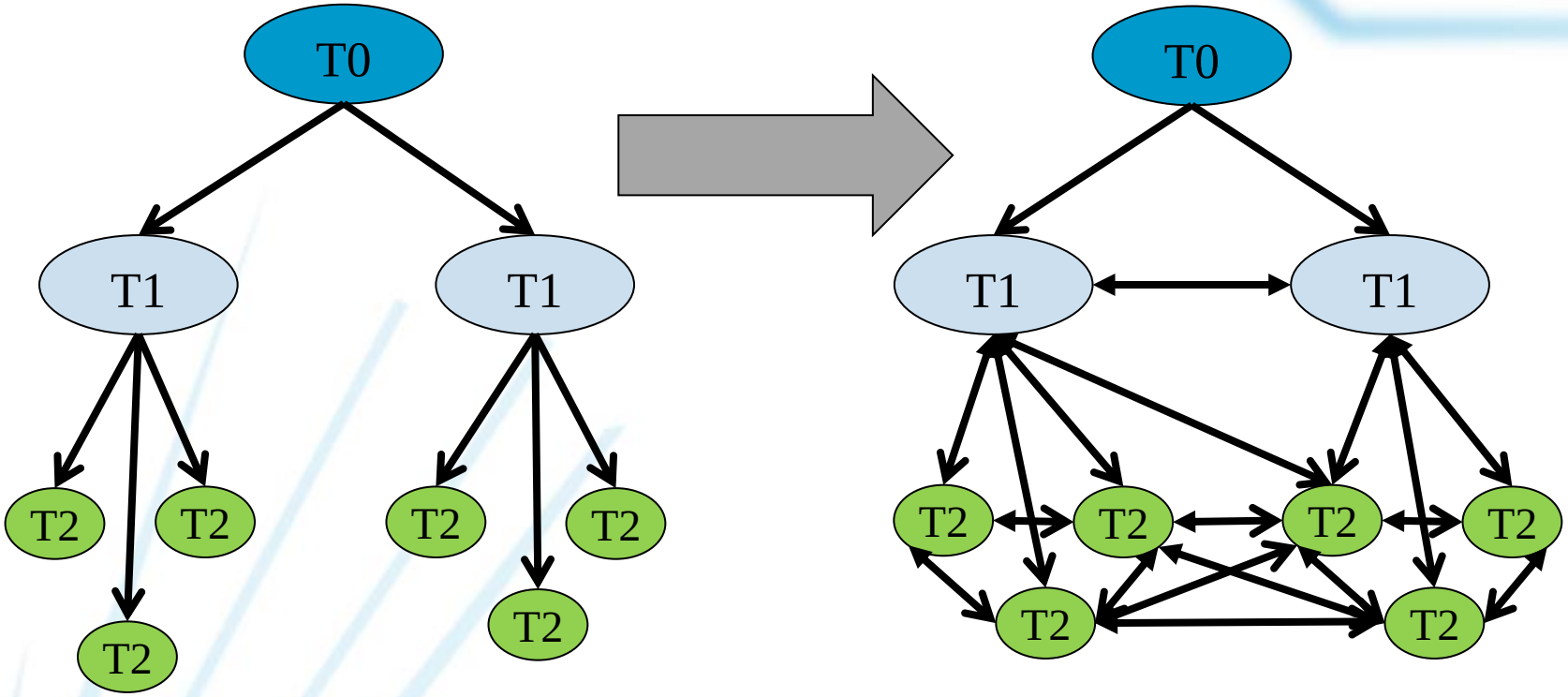
CPU delivered Jan-2011 – Sept-2012

LHCOPN


2G ASnet
1G Canarie-Surfnet
1G Canarie-ESnet
TW-ASGC
AS24167
117.103.96.0/20
140.109.98.0/24
140.109.102.0/24
202.140.160.0/19
202.169.168.0/22
CA-TRIUMF
AS36391
206.12.1.0/24
206.12.9.64/28
US-T1-BNL
AS43
130.199.185.0/24
130.199.48.0/23
130.199.54.0/24
192.12.15.0/24
US-FNAL-CMS
AS 3152
131.225.2.0/24
131.225.160.0/24
131.225.184.0/22
131.225.188.0/22
131.225.204.0/22
2.5G ASnet-Surfnet
10G Asnet-Surfnet
5G Canarie-Surfnet
1G Canarie-Surfnet
2x8.5Gb USLHCnet - ESnet
3.5Gb USLHCnet - ESnet
3.5Gb USLHCnet - ESnet
2x8.5Gb USLHCnet - ESnet
NDGF
AS39590
109.105.124.0/22
10G Nordunet-Geant
CH-CERN
AS513
128.142.0.0/16
10G Renater
FR-CCIN2P3
AS789
193.48.99.0/24
10G JANET(UK)-Geant
10G JANET(UK)-Geant
10G Surfnet
10G DFN-Geant
10G GARR-SWITCH
10G GARR-Geant
10G Geant – Redirs - Abella
UK-T1-RAL
NL-T1
DE-KIT
AS34878
192.108.45.0/24
192.108.46.0/23
IT-INFN-CNAF
AS137
131.154.128.0/17
ES-PIC
AS43115
193.109.172.0/24
10G DFN-SWITCH-GARR
10G DFN-Renater
2x1G DFN-Surfnet-USLHCnet-ESnet
10G Geant–Redirs-Abella
CERN

LHCOPN FR-CCIN2P3 Primary
Bits
10 G
0
-10 G
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Avg
Max
Last
Max
IN2P3 to CERN 1.36G 8.46G 1.08G Peak: 9.78G
CERN to IN2P3 1.97G 5.72G 1.05G Peak: 9.55G
Last update: Thu Sep 27 2012 15:57:39
CERN 2012

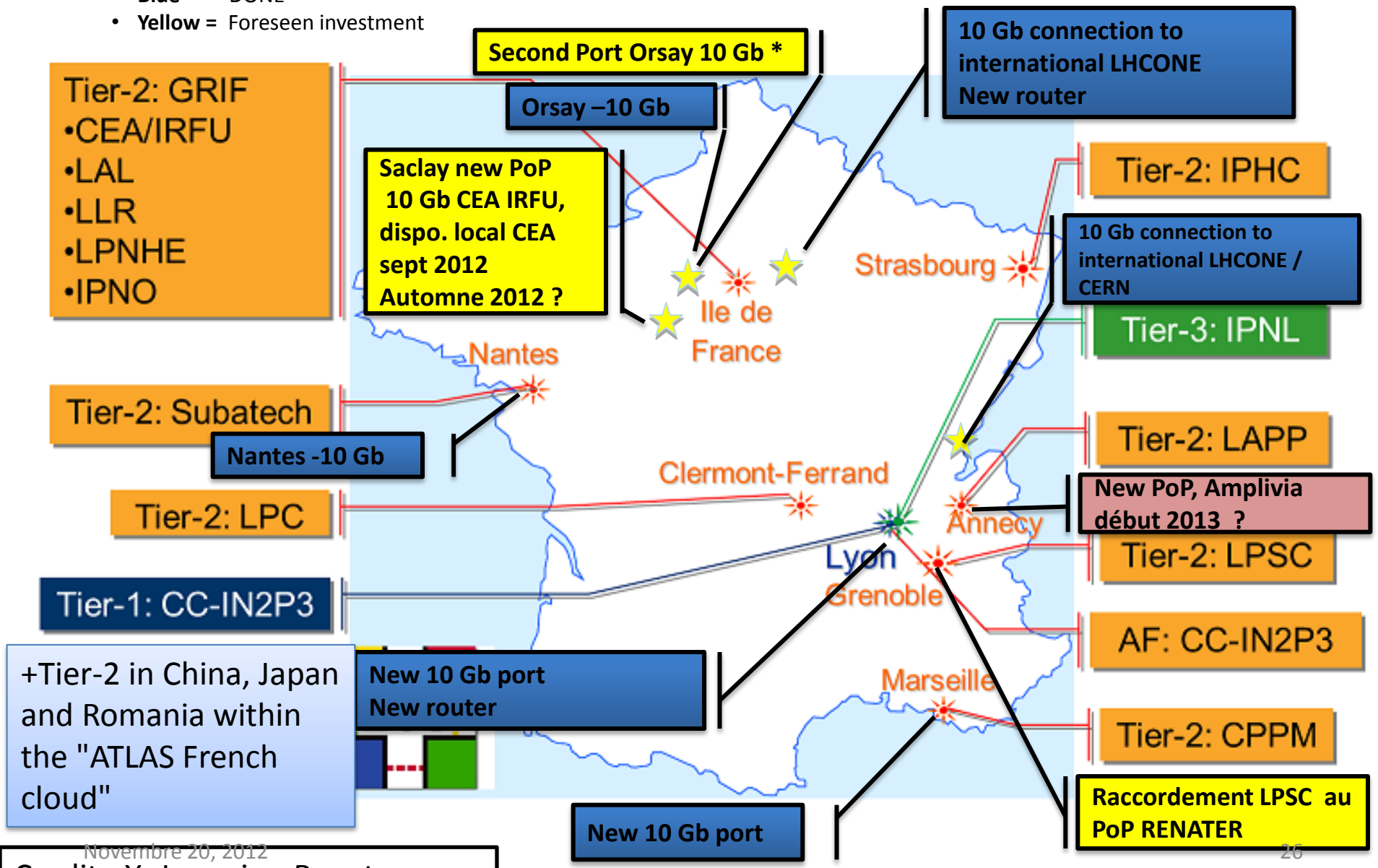
Evolutions of the Computing Models



➔ LHC Open Network Environment Project

French LHCONE improvement September 2012

- **Blue** = DONE
- **Yellow =** Foreseen investment

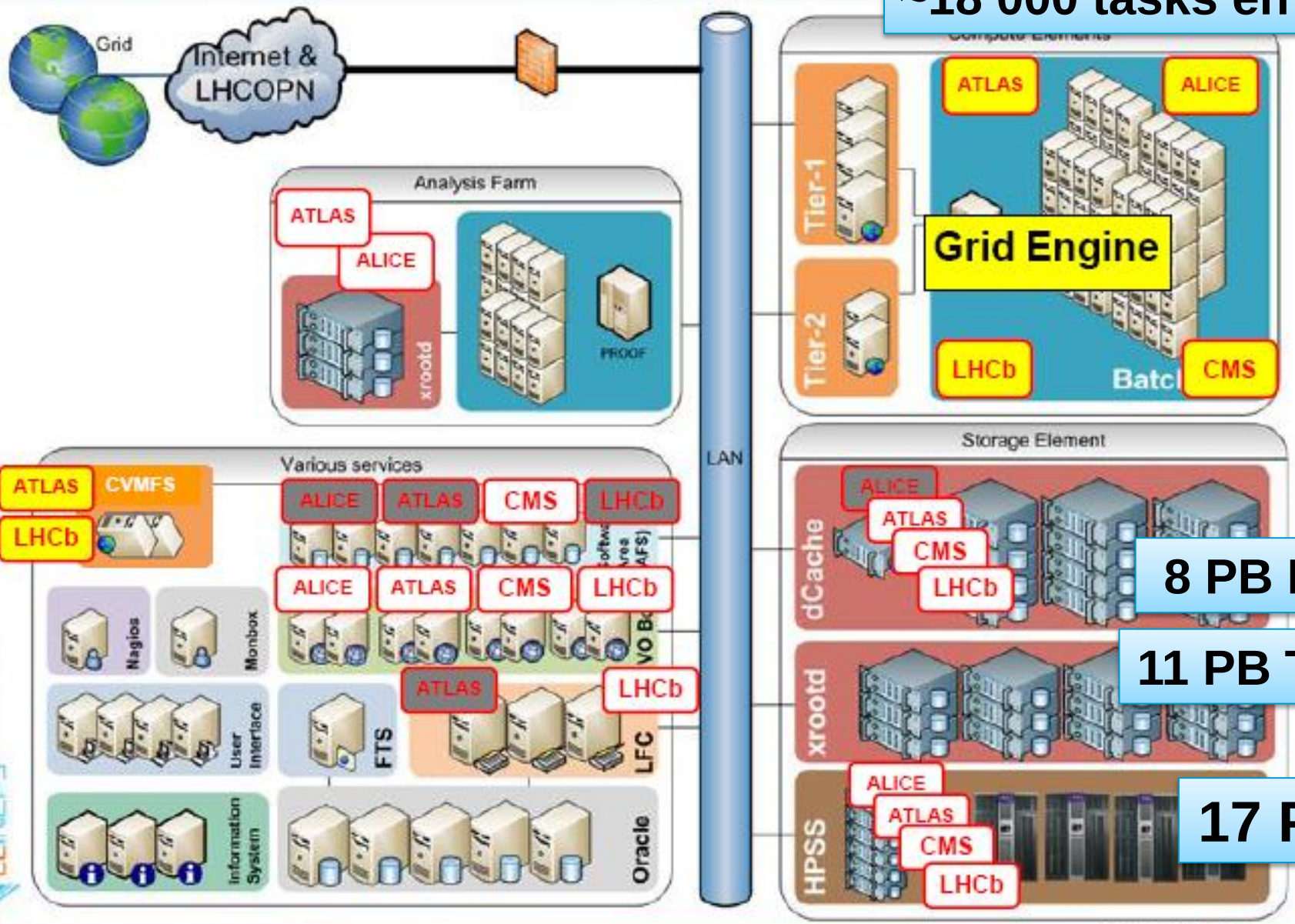


+Tier-2 in China, Japan and Romania within the "ATLAS French cloud"

Credit : X. Jeannin - Renater

CC-IN2P3 resources for LCG

~18 000 tasks en //



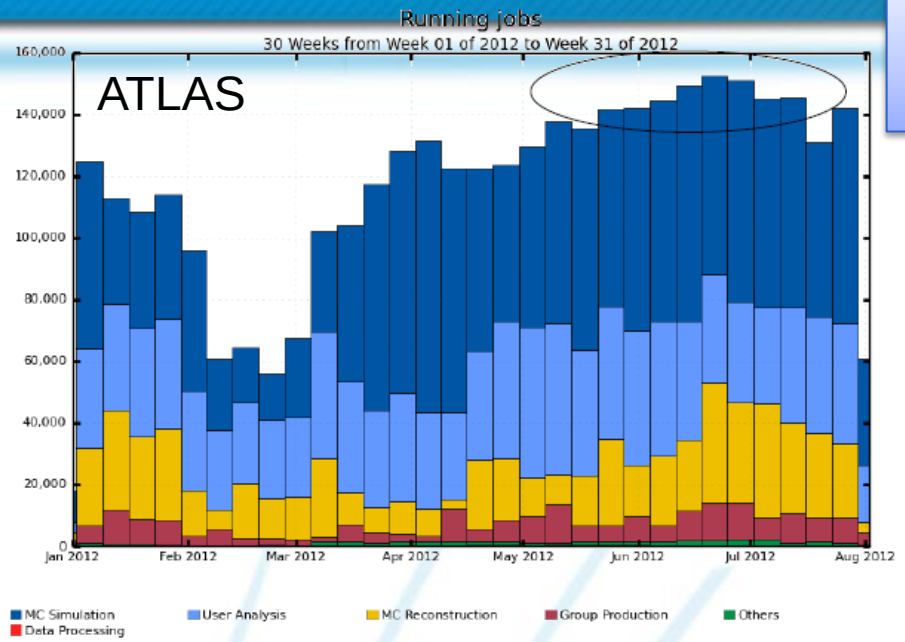
8 PB LHC

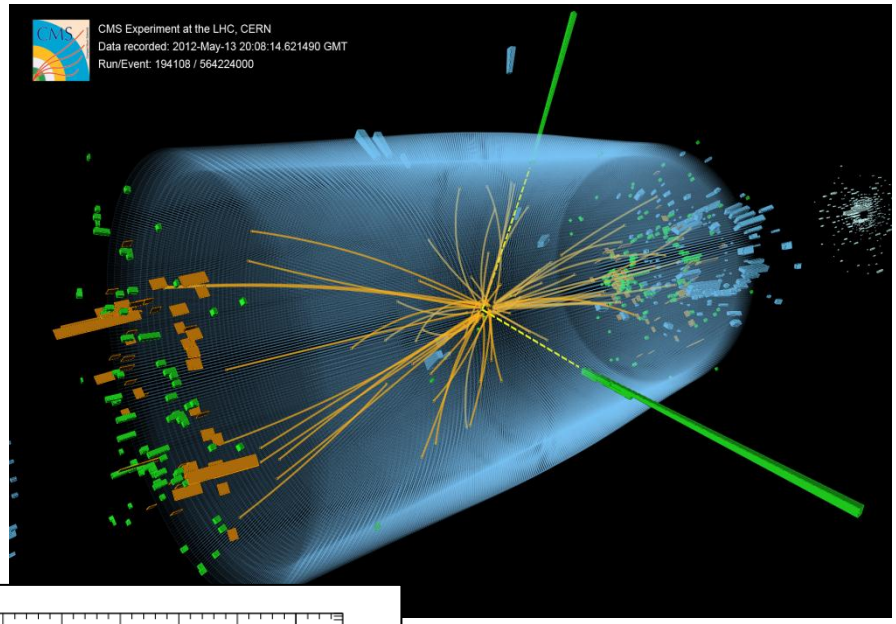
11 PB Total

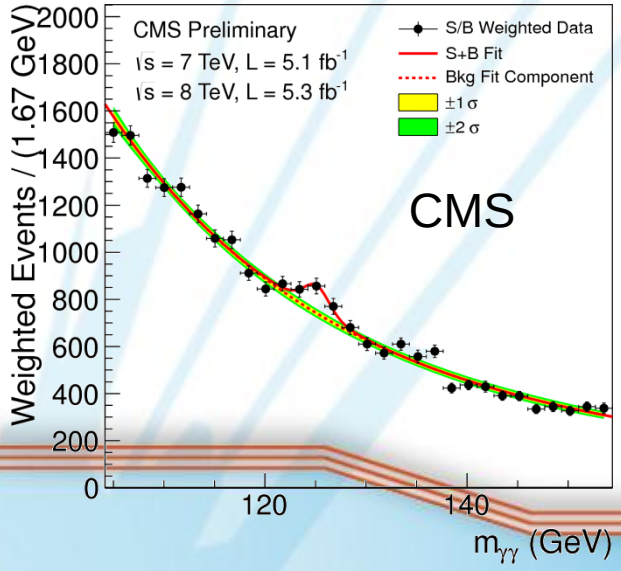
17 PB

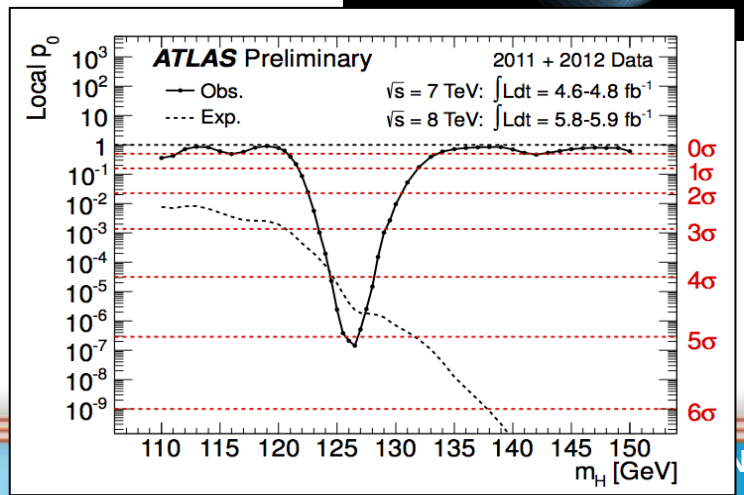
The Higgs field effect… on computing

Intense computing activity between May and July



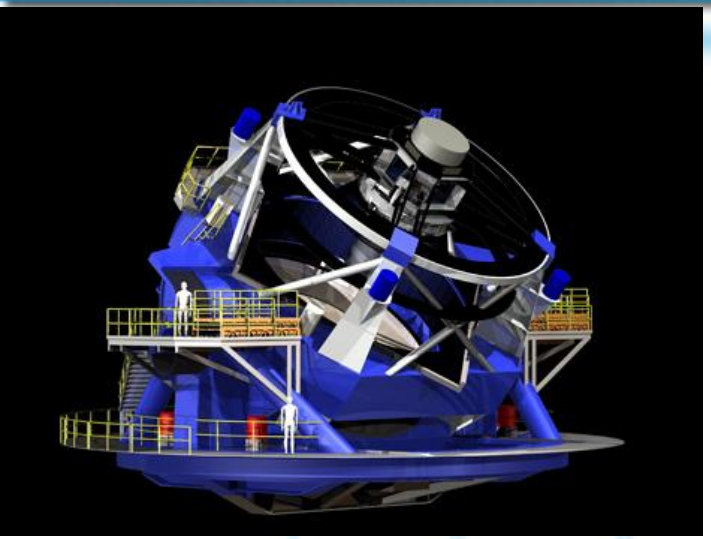






Large recognition of the W-LCG role in the Higgs discovery

Future Astropaticle experiments

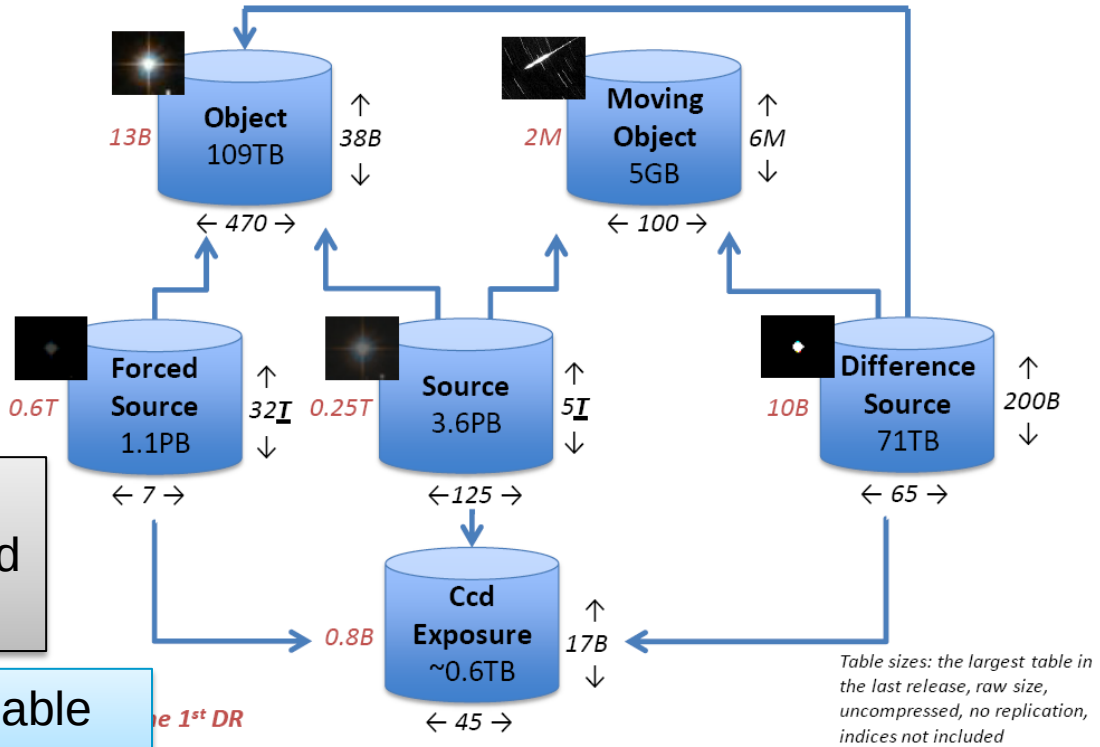


LSST : *Large Synoptic Survey Telescope*
3.2 Gpixels – 1 image every 15s
➔ 15 à 30 TB of data every night

The whole visible sky will be scanned 1000 times during the lifetime of the project (10 years)

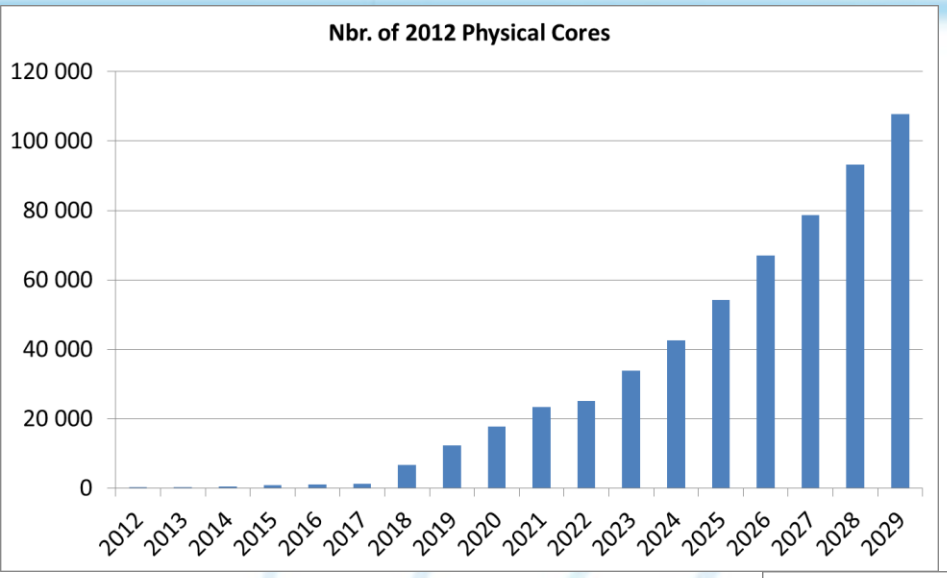
CC-IN2P3 will process 50% of LSST data (the other 50% will be processed at NCSA)

The whole LSST dataset will be available at CC-IN2P3



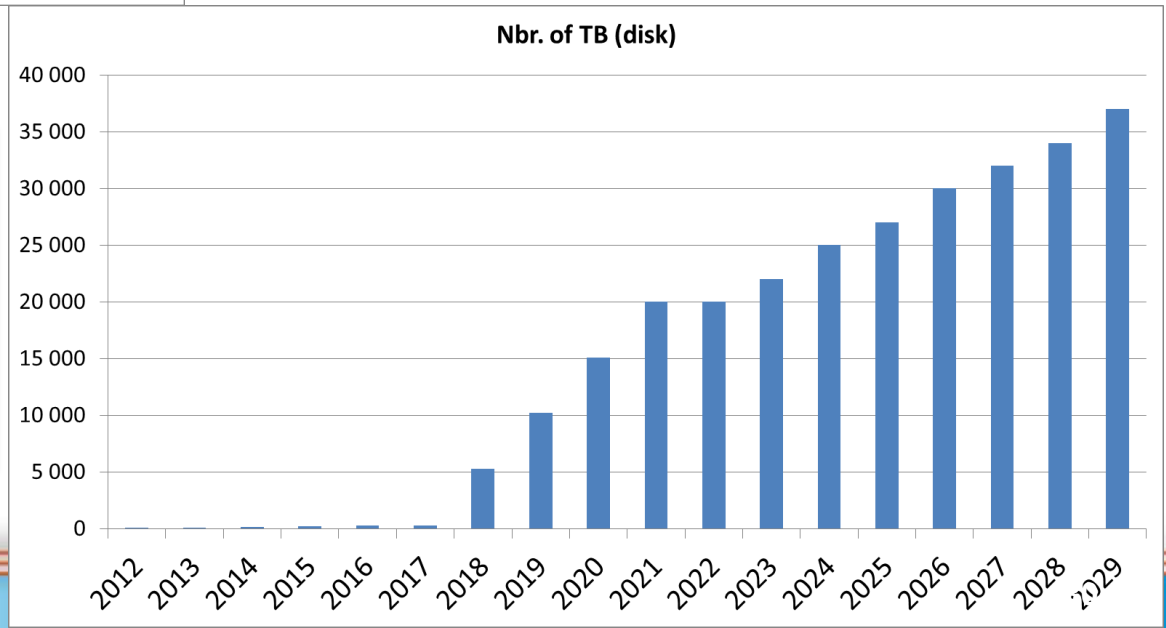
Crédit : LSST Collaboration

CC-IN2P3 contribution to LSST data processing

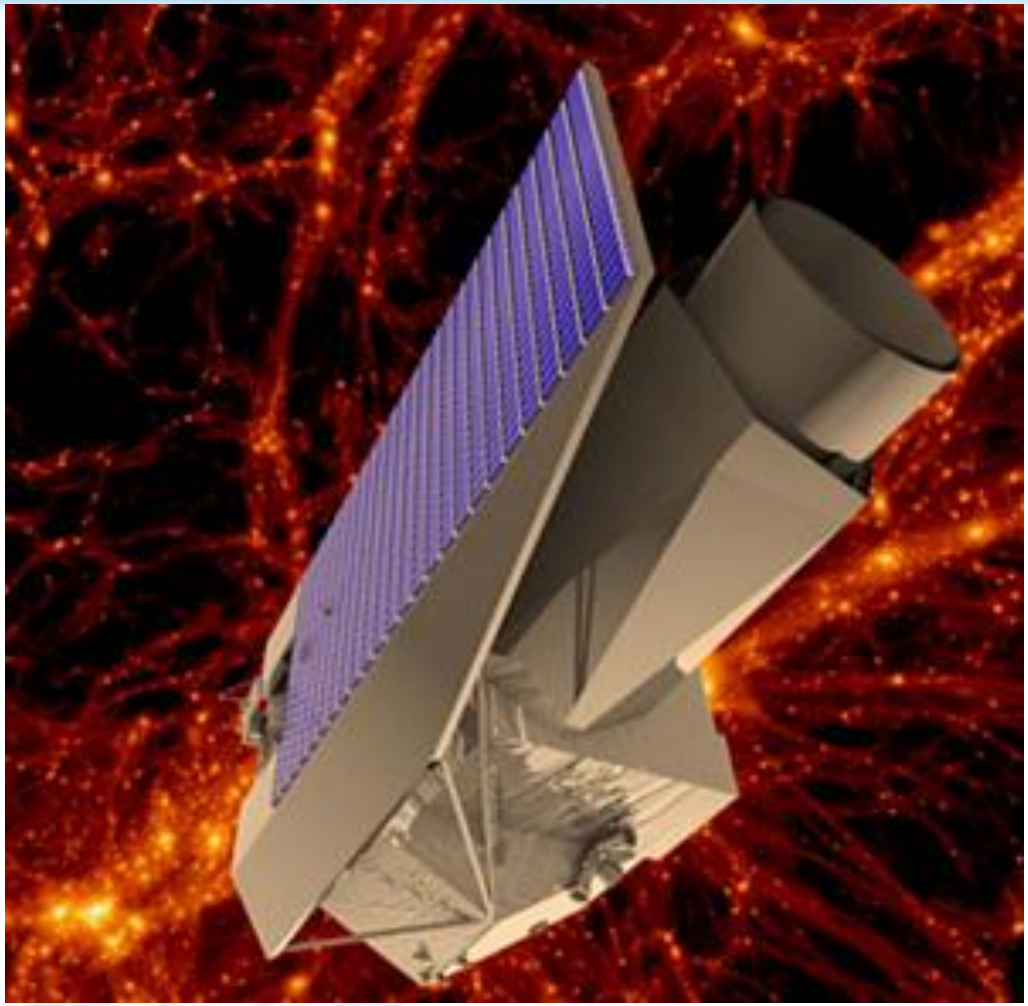


Hardware investment : ~1.5 M€ / year

CC-In2P3 will also provide computing resources for the physics analyses within the Dark Energy Science Collaboration



EUCLID



ESA space telescope dedicated to dark energy and dark matter studies

Idea to combine LSST and EUCLID data

If there is an agreement between the 2 projects there will be a huge load on the computing infrastructure to convert LSST data to EUCLID format

Having the LSST data at CC-IN2P3 will be a big advantage

Multidisciplinary

We are already opening to non HEP communities.

- Human Sciences
- Biomedical

Setup a Regional Grid TIDRA to accommodate “local” requests.

Would like to setup an **Academic Cloud**

- It will extend the Regional Grid TIDRA to other disciplines
  - Biomedical
  - Image processing

We are investigating how to open it to industrial users **<u>at full cost</u>**

This a complete change for CC-IN2P3 and we will have to handle it very carefully

- Develop innovative approach for multidisciplinary applications
- While keeping computing for HEP as a top priority

The new computer room can host such an academic Cloud

Example I: IRT

CC-IN2P3 is a partner of the IRT (Institute for Research and Technology) : LyonBioTech / BioAster

- Will provide a research and development infrastructure on infectious diseases
- Strongly oriented toward industrial applications
- Budget > 800 M€ in total
- Should become economically self-sufficient within 10 years
- Aim at creating 10 000 jobs (direct and indirect)

CC-IN2P3 and a regional startup (SysFera) to provide IT infrastructure

- 2.2 M€
- Budget for manpower

Very complex structure to be created from scratch with many partners

Several kind of applications but strong orientation toward database systems

Very important as it can promote CC-IN2P3 as the leading data processing center for biomedicine in France

Example II: eTRIKS

- European Translation Information & Knowledge Management Service Infrastructure based on the tranSMART platform (http://www.transmartproject.org)
  - Porting faster research results to pharmaceutical applications
- Biomedical knowledge management system within the Innovative Medicines Initiative (IMI)
  - Establish correlation between multiple biomedical data
    - Genome – proteome – imagery – medical data – etc…

~20 M€ project : 10 M€ from Europe – 10 M€ from the "pharmas"

➔ 1.7 M€ for CC-IN2P3 (530 k€ in hardware – the rest in manpower)

First phase has been accepted by IMI – eTRIKS is alone for the second phase

CC-IN2P3 will be in charge of the cloud platform, of the database infrastructure and of part of the user support.

What about Data Preservation?

CC-IN2P3 is the Computing Center for IN2P3

- Is a computing facility hosting data and Computing for many experiments but does not host a physics laboratory and is not dedicated to one experiment.
- Have some small scale Data Preservation within human and social sciences (ADONIS).
- Will have to work together with Experiments and/or DP teams to develop, test and deploy DP schemes.
- CC has a large expertise on data storage and retrieval that is an asset for DP