IV JORNADAS CPAN GRANADA, DEL 26 AL 28 DE NOVIEMBRE DE 2012

Spanish Grid Computing contribution for the LHC data processing

Santiago González de la Hoz

on behalf of Spanish LCG community

Instituto de Física Corpuscular (IFIC)
Universitat de València - CSIC

Outline

- Introduction
  - Multi-petabyte scale computing for the LHC
    - Data transfers on the Grid
    - Processing on the Grid
  - WLCG: A global Collaboration
    - E-Infrastructures: Grids
- Spanish Grid Computing Contribution
  - Spanish Tier-1 and federated Tier-2s
  - Infrastructure & Resources
  - Tier-1 & Tier-2s performances
    - data, production, analysis, availability
- Future and perspectives
- Summary

Some history of scale

Date	Collaboration members	Data volume, archive technology
Late 1950's	2-3	100B, notebooks
1960's	10-15	kB, punchcards
1970's	~350	MB, tape
1980's	~100	GB, tape, disk
1990's	700-800	TB, tape, disk
2010's	~3000	PB, tape, disk

For comparison:
1990's: Total LEP data set ~few TB
Would fit on 1 tape today

Today: 1 year of LHC data ~25 PB

- **Where does all this data come from?**
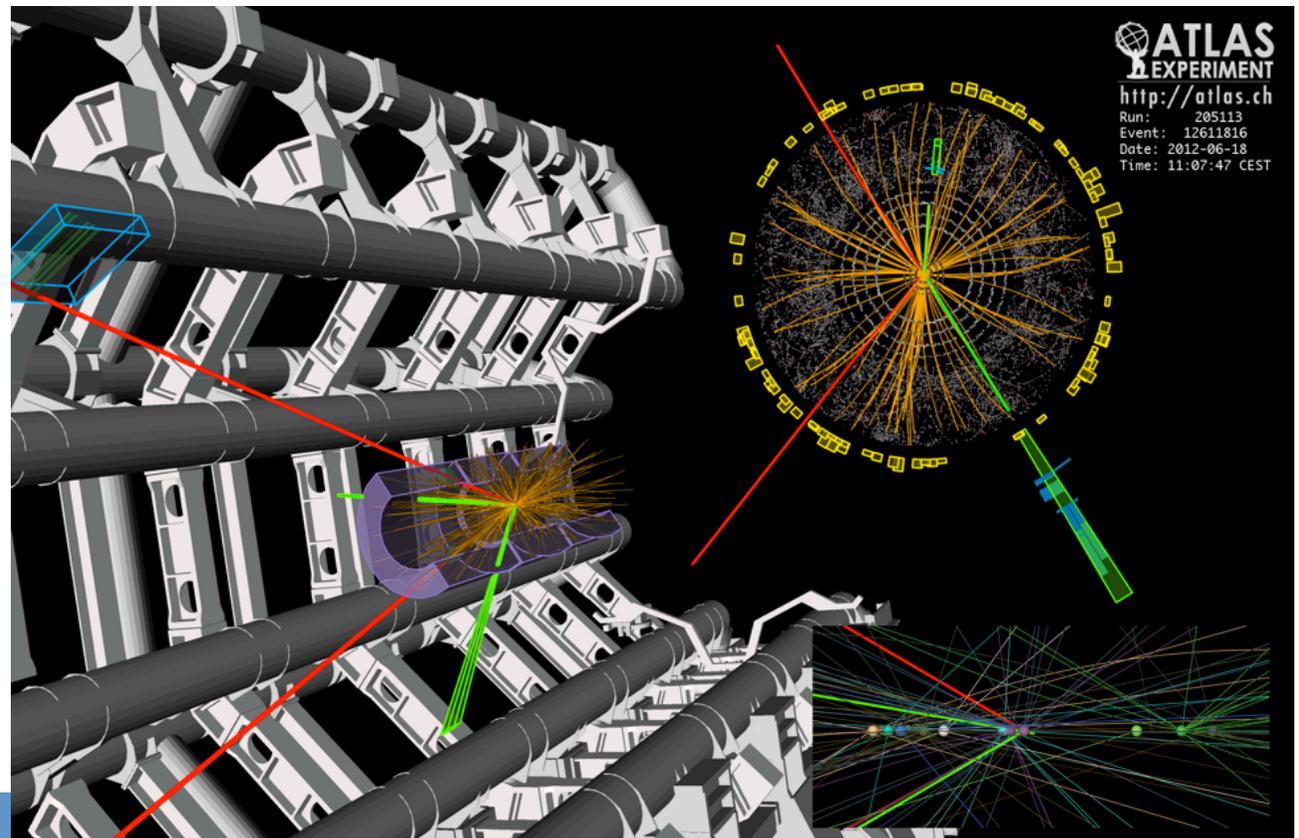
What is this data?

- **Raw data**:
  - Was a detector element hit?
  - How much energy?
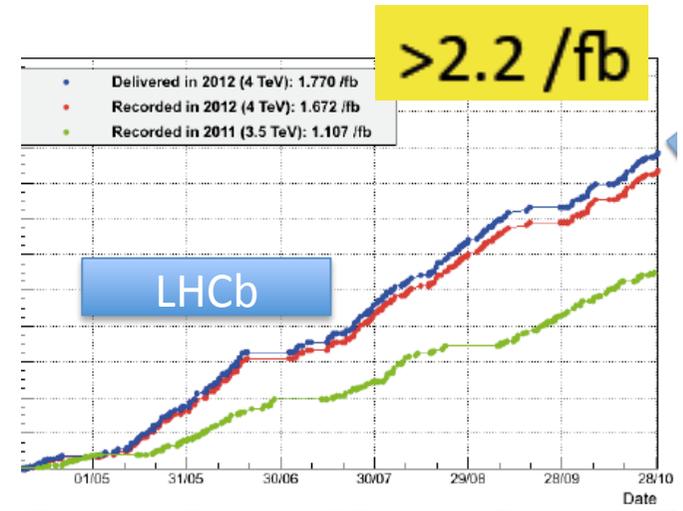 - What time?

- **Reconstructed data**:
  - Momentum of tracks (4-vectors)
  - Origin
  - Energy in clusters (jets)
  - Particle type
  - Calibration information
  - ...

- **150 Million sensors** deliver data ... 40 Million times per second!!!!!!

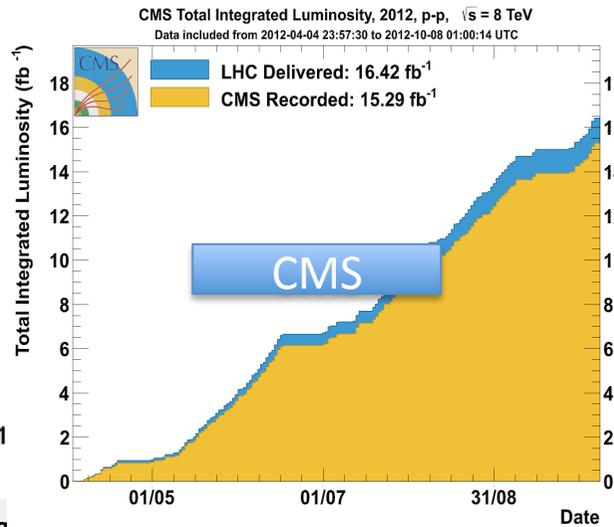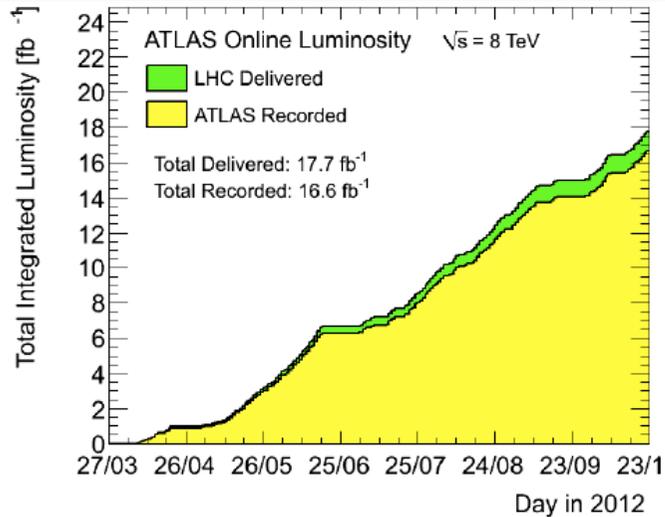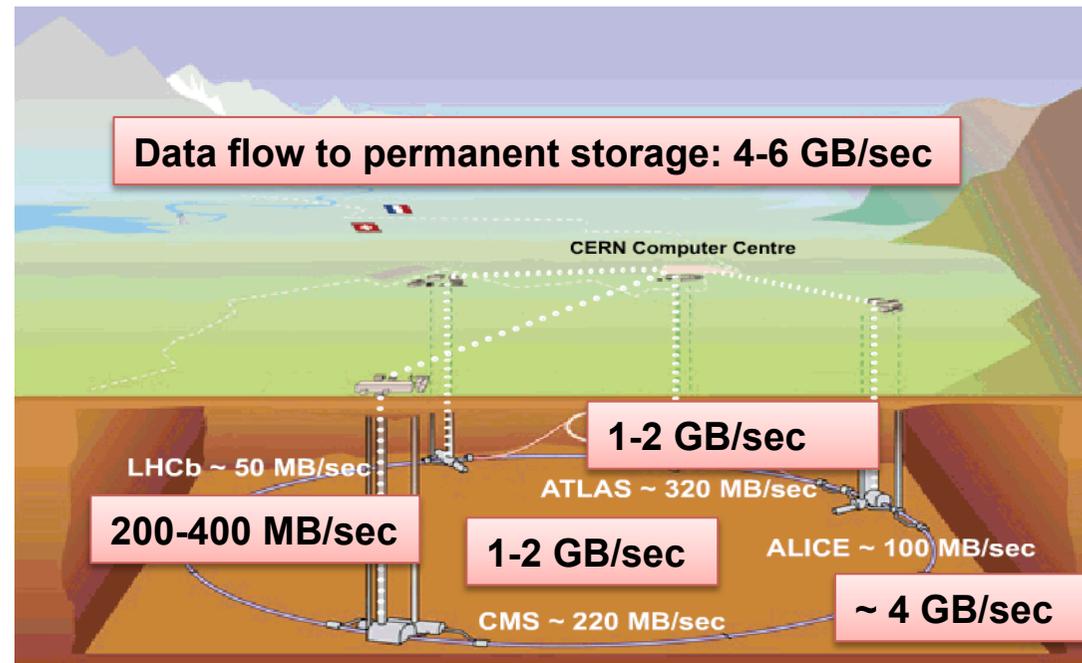- **Up to 6 GB/s** to be stored and analysed after filtering!!!!!!!

ATLAS EXPERIMENT
http://atlas.ch
Run: 205113
Event: 12611816
Date: 2012-06-18
Time: 11:07:47 CEST

2012 data taking

Peak luminosity seen by ATLAS: 7.73 10^{33} cm^{-2} s^{-1}

ATLAS Online Luminosity $\sqrt{s}$ = 8 TeV
LHC Delivered
ATLAS Recorded
Total Delivered: 17.7 fb^{-1}
Total Recorded: 16.6 fb^{-1}

CMS Total Integrated Luminosity, 2012, p-p, $\sqrt{s}$ = 8 TeV
Data included from 2012-04-04 23:57:30 to 2012-10-08 01:00:14 UTC
LHC Delivered: 16.42 fb^{-1}
CMS Recorded: 15.29 fb^{-1}

CMS

>2.2 /fb

Delivered in 2012 (4 TeV): 1.770 /fb
Recorded in 2012 (4 TeV): 1.672 /fb
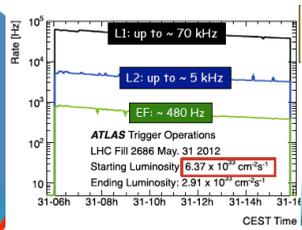Recorded in 2011 (3.5 TeV): 1.107 /fb

LHCb

ATLAS parameters for 2012	Actual p-p (Mar-Aug 2012)
Trigger rate (event/sec)	350 (prompt) + 150 (delayed)
Real Raw data (MB/event)	0.8
CMS parameters for 2012	Actual p-p (Mar-Aug 2012)
Trigger rate (event/sec)	400 (core) + 400 (parked)
Real Raw data (MB/event)	0.4
LHCb parameters for 2012	Actual p-p (Mar-Aug 2012)
Trigger rate (event/sec)	3000-5000
Real Raw data (MB/event)	0.02

Data flow to permanent storage: 4-6 GB/sec

CERN Computer Centre

1-2 GB/sec

LHCb ~ 50 MB/sec   ATLAS ~ 320 MB/sec

200-400 MB/sec   **1-2 GB/sec**   ALICE ~ 100 MB/sec

CMS ~ 220 MB/sec   **~ 4 GB/sec**

Data Handling and Computation for Physics Analysis

Online trigger and filtering

Selection & reconstruction

Offline Reconstruction

Processed Data (Active tapes)

Event summary data

100% | Raw data

10%

Batch physics analysis

1%

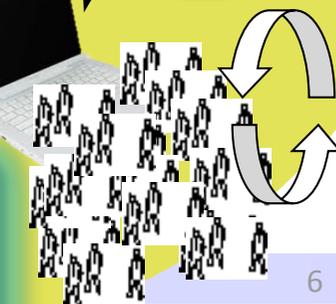**Event reprocessing**

Event simulation

Offline Simulation w/GEANT4
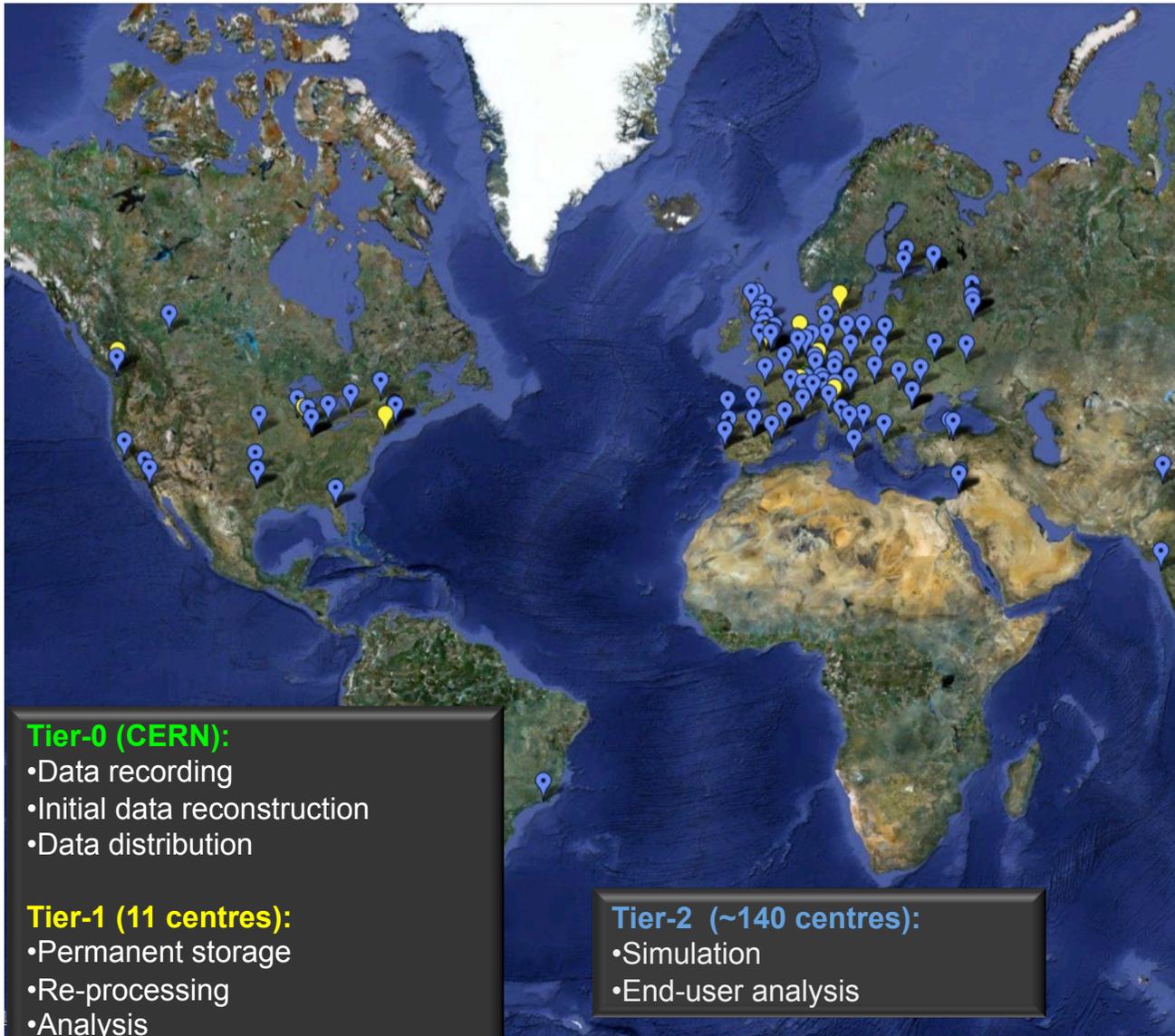
Offline Analysis w/ROOT

Analysis objects (extracted by physics topic)

Interactive analysis

WLCG: A global Collaboration

- WLCG has been leveraged on both sides of the Atlantic, to benefit the wider scientific community
  - Europe:
    - Enabling Grids for E-sciencE (EGEE) 2004-2010
    - European Grid Infrastructure (EGI) 2010--
  - USA:
    - Open Science Grid (OSG) 2006-2012 (+ extension?)

- Many scientific applications

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

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

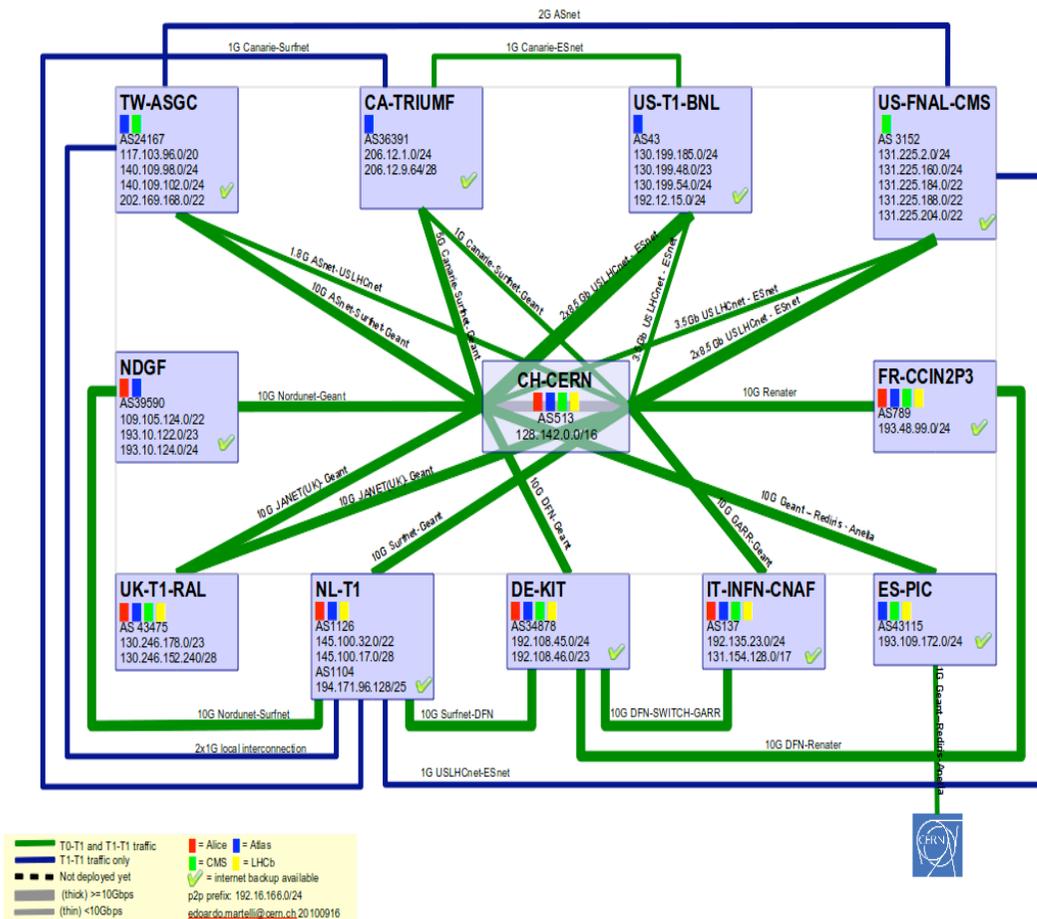**Tier-2 (~140 centres):**
- Simulation
- End-user analysis

● Tier 0   ● Tier 1   ● Tier 2
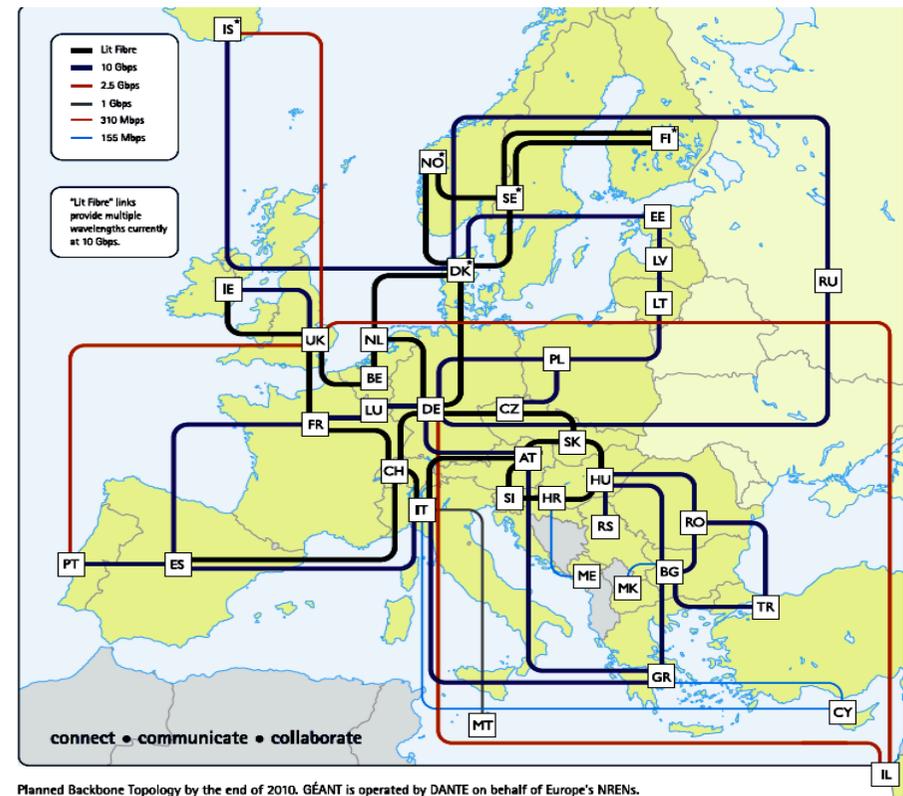
- ~150 sites
- >250k CPU cores
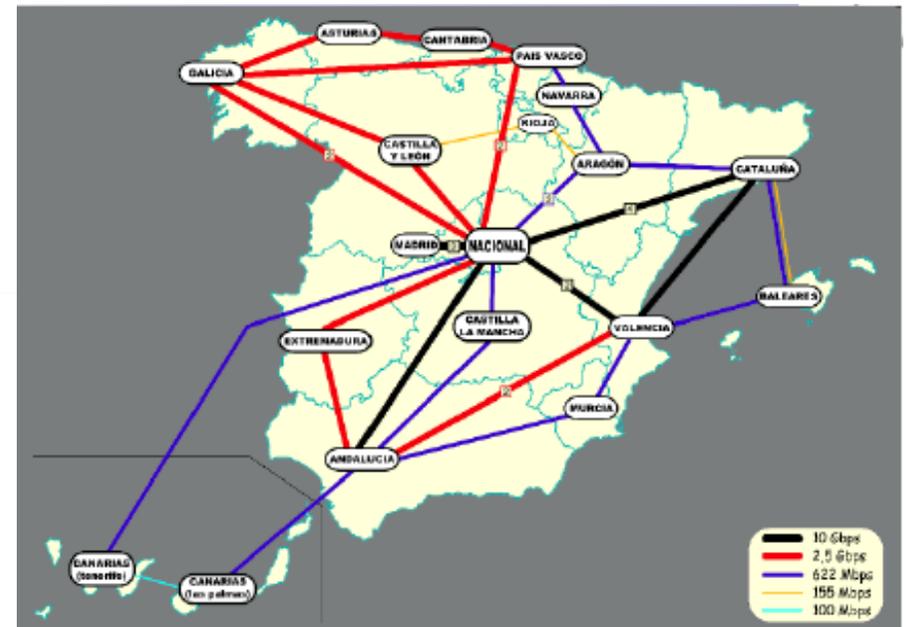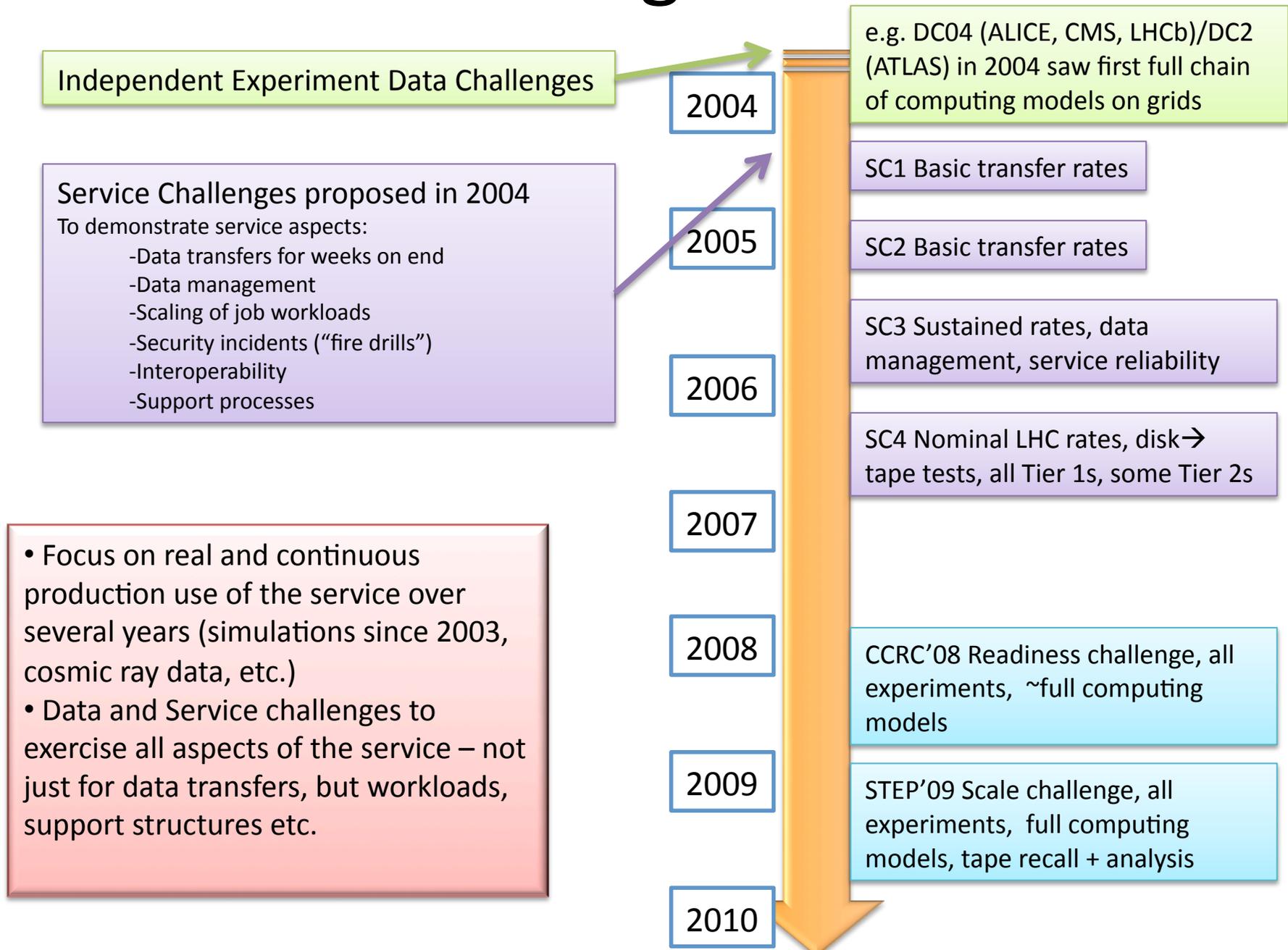- >150 PB disk

LHC networking

LHCOPN

- Relies on
  - OPN, GEANT, US-LHCNet
  - NRENs & other national & international providers

Planned Backbone Topology by the end of 2010. GÉANT is operated by DANTE on behalf of Europe's NRENs.
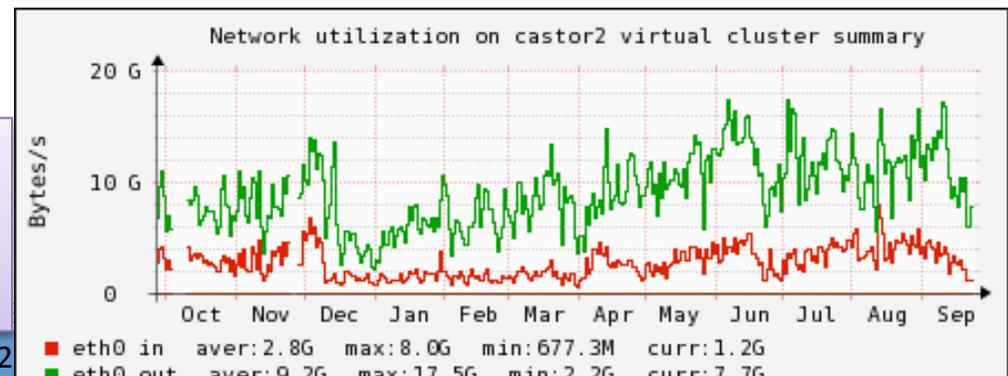
connect • communicate • collaborate

From testing to data

Independent Experiment Data Challenges

e.g. DC04 (ALICE, CMS, LHCb)/DC2 (ATLAS) in 2004 saw first full chain of computing models on grids

2004

Service Challenges proposed in 2004
To demonstrate service aspects:
- -Data transfers for weeks on end
- -Data management
- -Scaling of job workloads
- -Security incidents ("fire drills")
- -Interoperability
- -Support processes

SC1 Basic transfer rates

SC2 Basic transfer rates

2005

SC3 Sustained rates, data management, service reliability

2006

SC4 Nominal LHC rates, disk→ tape tests, all Tier 1s, some Tier 2s

2007

• Focus on real and continuous production use of the service over several years (simulations since 2003, cosmic ray data, etc.)
• Data and Service challenges to exercise all aspects of the service – not just for data transfers, but workloads, support structures etc.

2008

CCRC'08 Readiness challenge, all experiments, ~full computing models

2009

STEP'09 Scale challenge, all experiments, full computing models, tape recall + analysis

2010

Tier0 data written in 2012

CASTOR data written, 01/01/2010 to 31/8/2012 (in GB)

p-p data

Pb-Pb data

2010-2012 Data written:
Total ~19 PB in 2012
Close to 4 PB/month now

>30 PB in 2012?

15 PB in 2010

23 PB in 2011

USER
NTOF
NA61
NA48
LHCB
COMPASS
CMS
ATLAS
AMS
ALICE
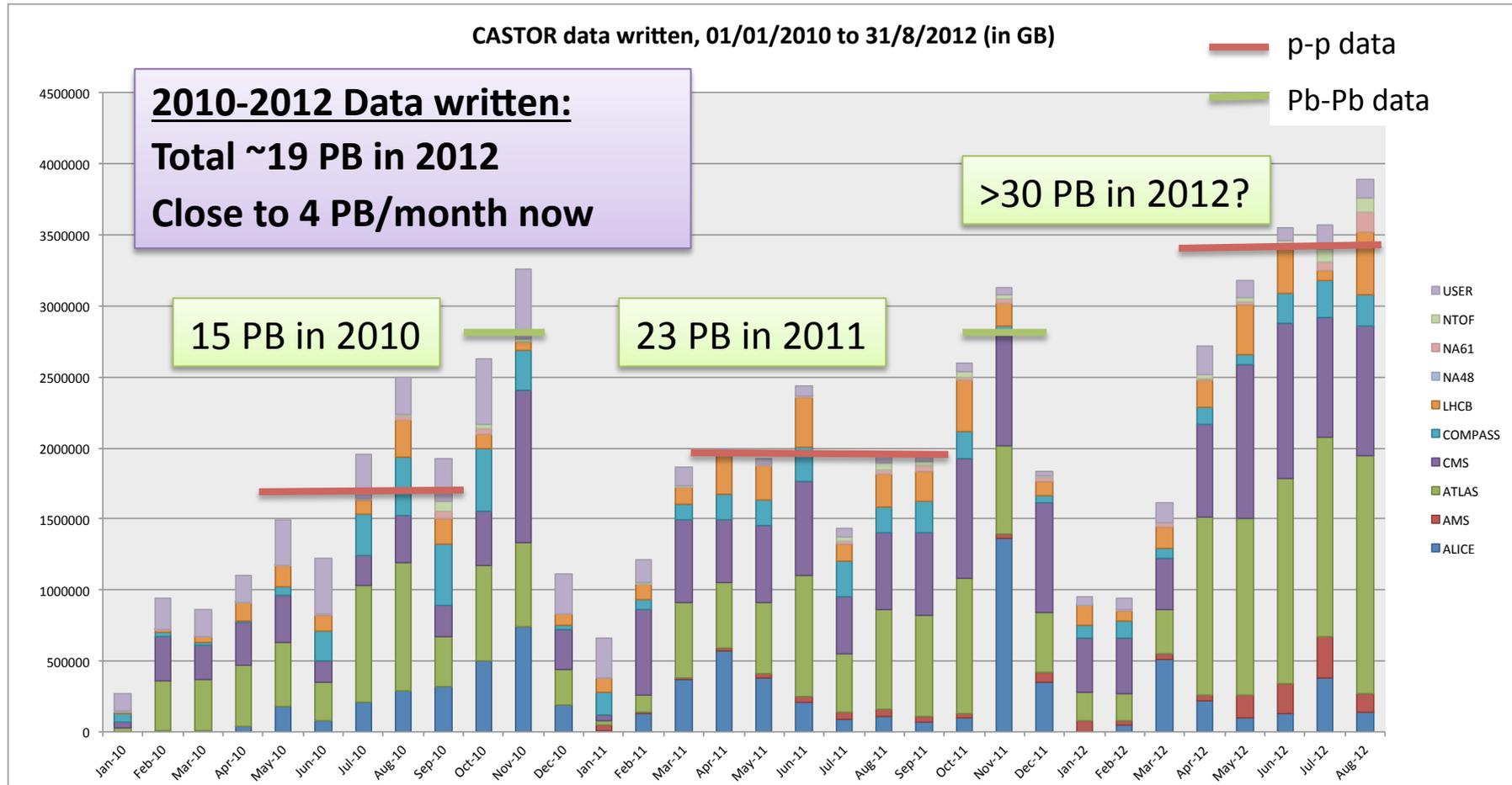
Data rates increased
3-4 GB/s input
~15 GB/s output

Network utilization on castor2 virtual cluster summary

eth0 in   aver:2.8G   max:8.0G   min:677.3M   curr:1.2G
eth0 out aver:9.2G   max:17.5G min:2.2G   curr:7.7G

Data transfers

Global transfers > 10 GB/s (1 day)
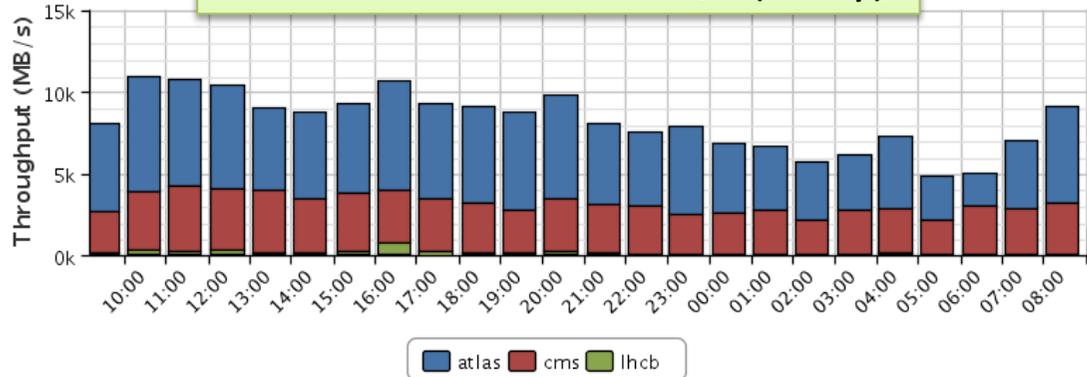
Global transfers (May-June)

2012-05-15 00:00 to 2012-06-12 00:00 UTC
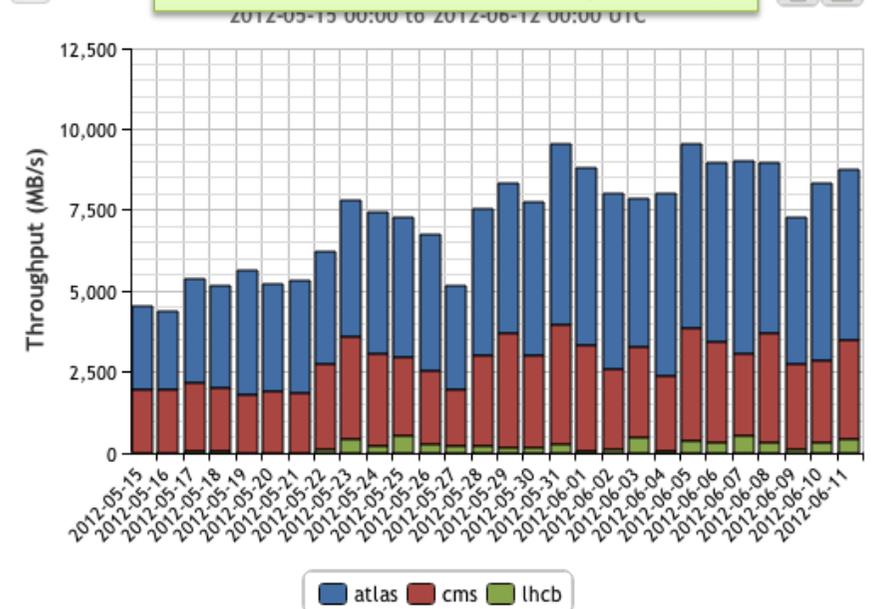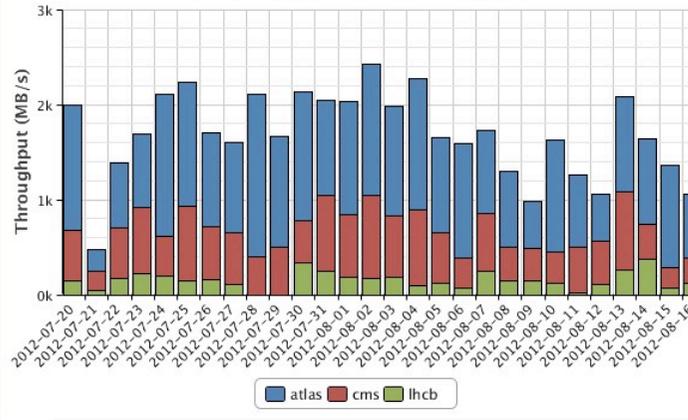
CERN → Tier 1s (with LHC on): 2 GB/s
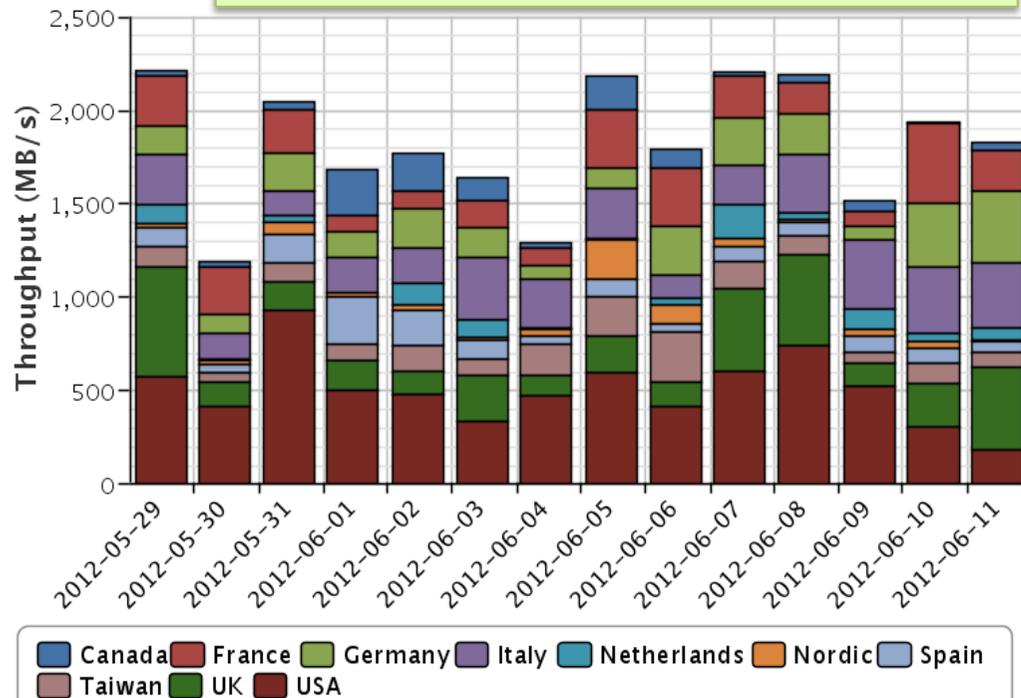
Planned rates:

Tier 0 – Tier 1: 650 MB/s + margin

Tier0/Tier1 to NDGF Tier1 dataset transfer time for 2012 data
Data are available for D3PD production and analysis in less than 4h, 60+% of datasets are available within 2 hours

4 hours

Throughput
2012-07-20 00:00 to 2012-08-17 00:00 UTC

Throughput
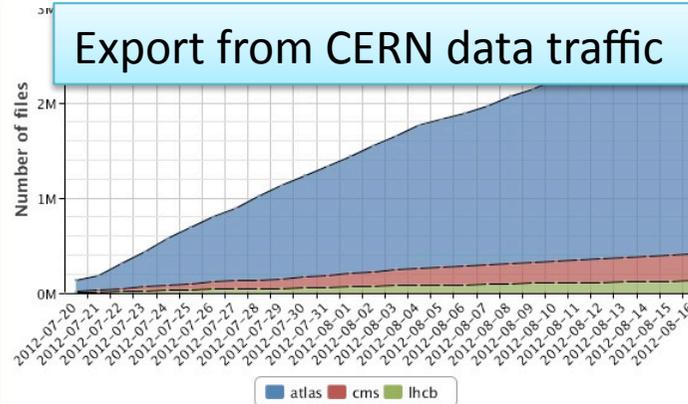2012-07-20 00:00 to 2012-08-17 00:00 UTC
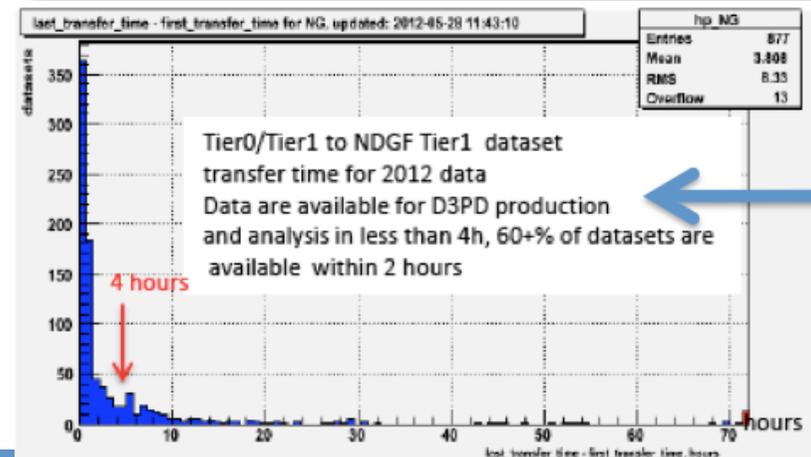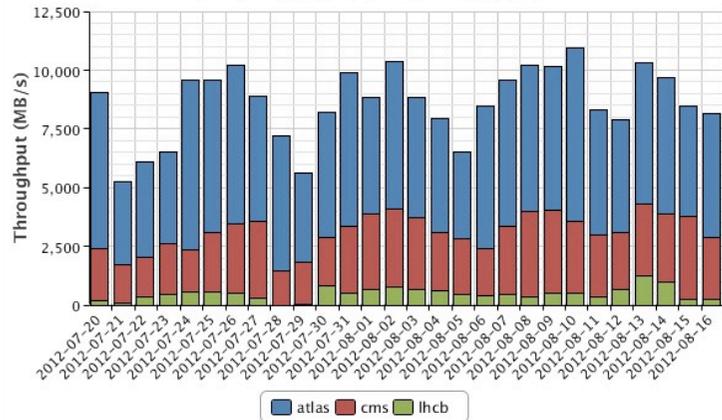
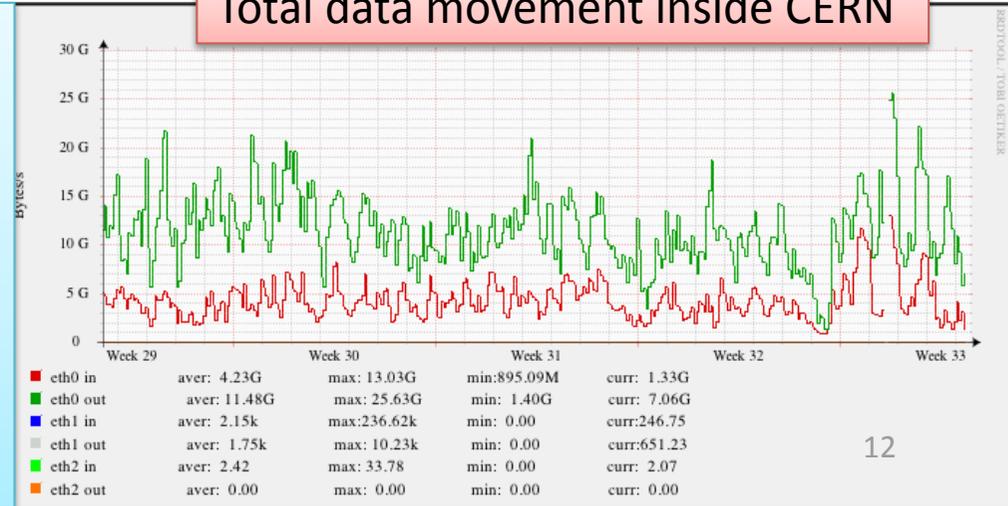**Data Transfers**

A key technology for LHC

Worldwide data traffic
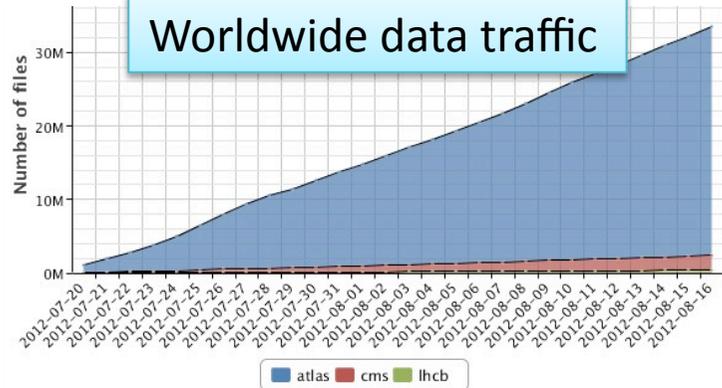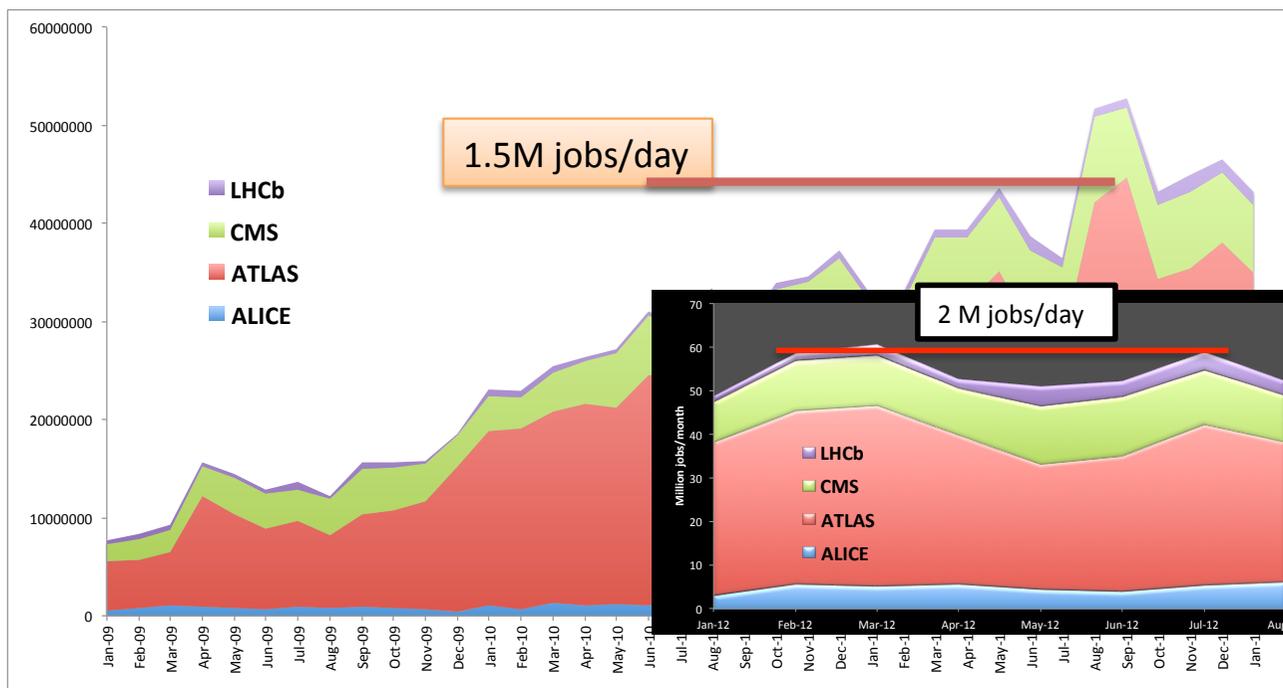
Export from CERN data traffic

Total data movement inside CERN

Last month:
4 PB, 2.6M files exported from CERN to Tier 1s
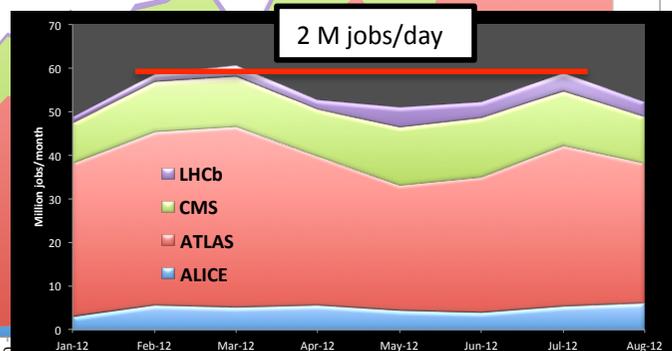20 PB, 33M files moved worldwide

Data movement inside Tier 0:
-Data in: 4.3 GB/s (peak 13 GB/s)
-Data out: 11.5 GB/s (peak 25 GB/s)
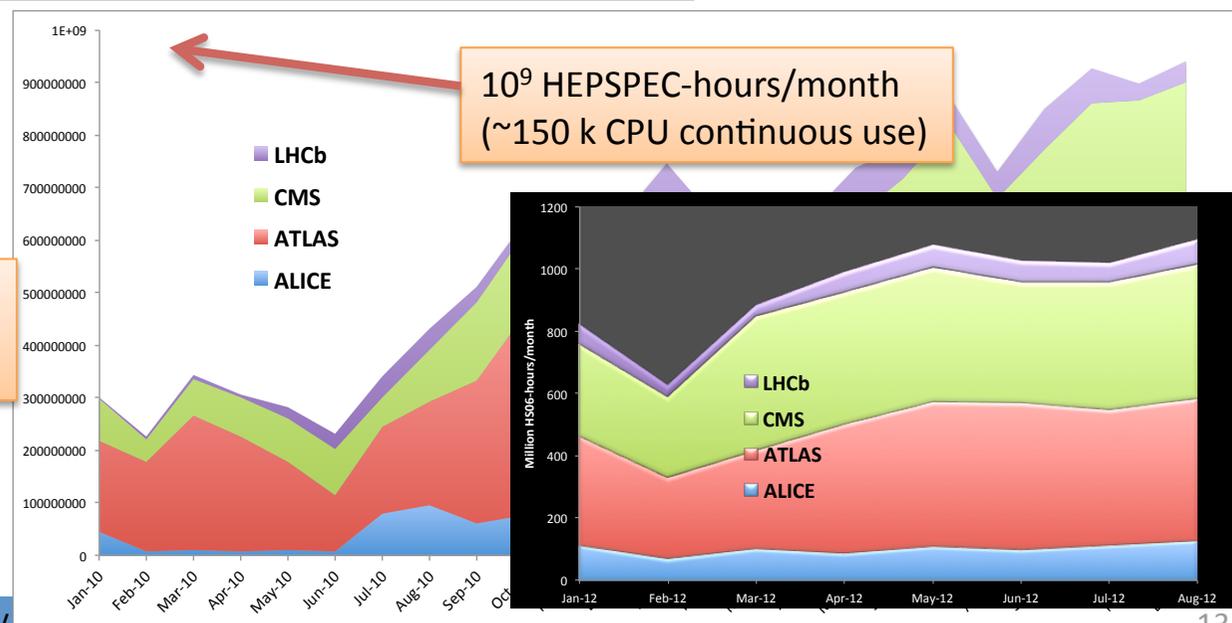-~1.5 PB/day data movement at CERN

eth0 in	aver: 4.23G	max: 13.03G	min:895.09M	curr: 1.33G
eth0 out	aver: 11.48G	max: 25.63G	min: 1.40G	curr: 7.06G
eth1 in	aver: 2.15k	max:236.62k	min: 0.00	curr:246.75
eth1 out	aver: 1.75k	max: 10.23k	min: 0.00	curr:651.23
eth2 in	aver: 2.42	max: 33.78	min: 0.00	curr: 2.07
eth2 out	aver: 0.00	max: 0.00	min: 0.00	curr: 0.00

12

Processing on the Grid

1.5M jobs/day

2 M jobs/day

Usage continues to grow...
- # jobs/day
- CPU usage

~ 150,000 years of CPU delivered each year

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

This is close to full capacity The resources are used!

CPU delivered Jan 2012 –Sep 2012

Legend:
- CERN
- BNL
- CNAF
- KIT
- NL LHC/Tier-1
- RAL
- FNAL
- CC-IN2P3
- ASGC
- PIC
- NDGF
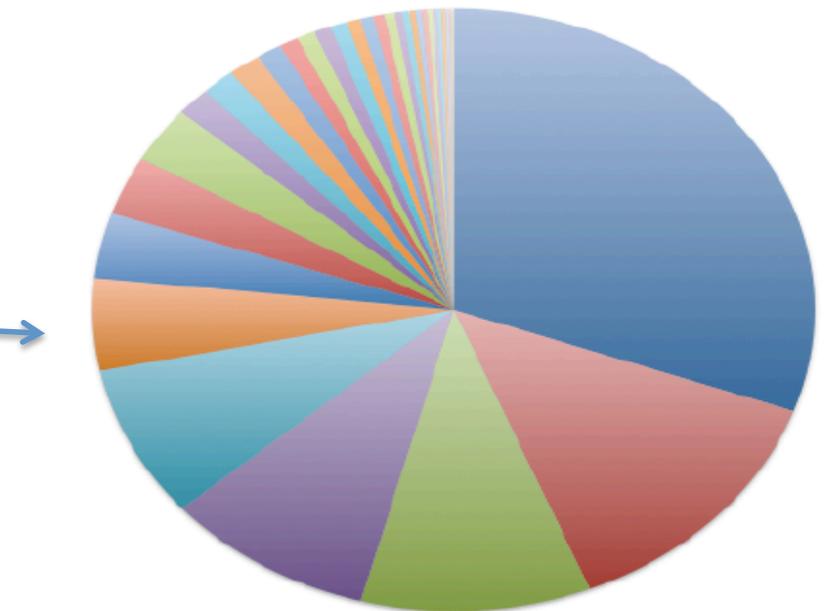- TRIUMF
- Tier 2
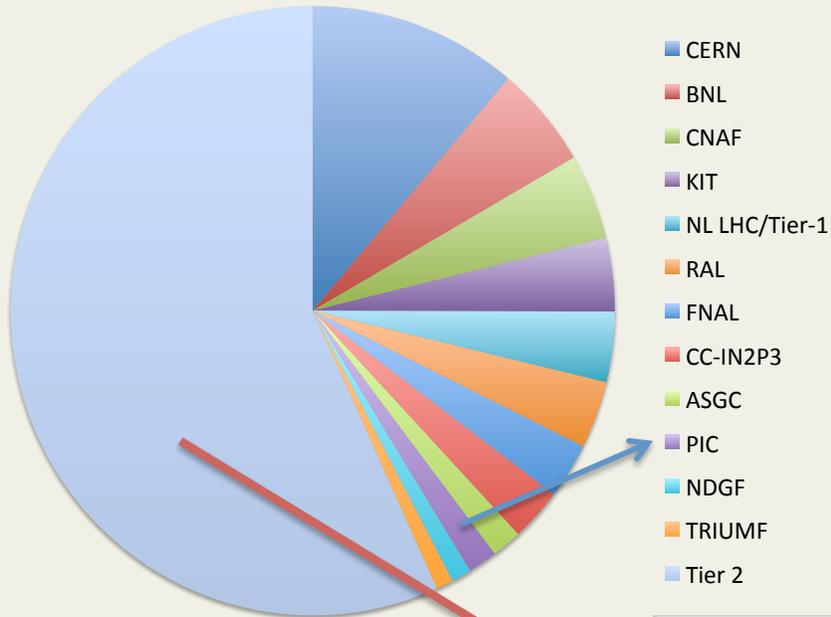
- The grid really works
- All sites, large and small can contribute
  - And their contributions are needed!

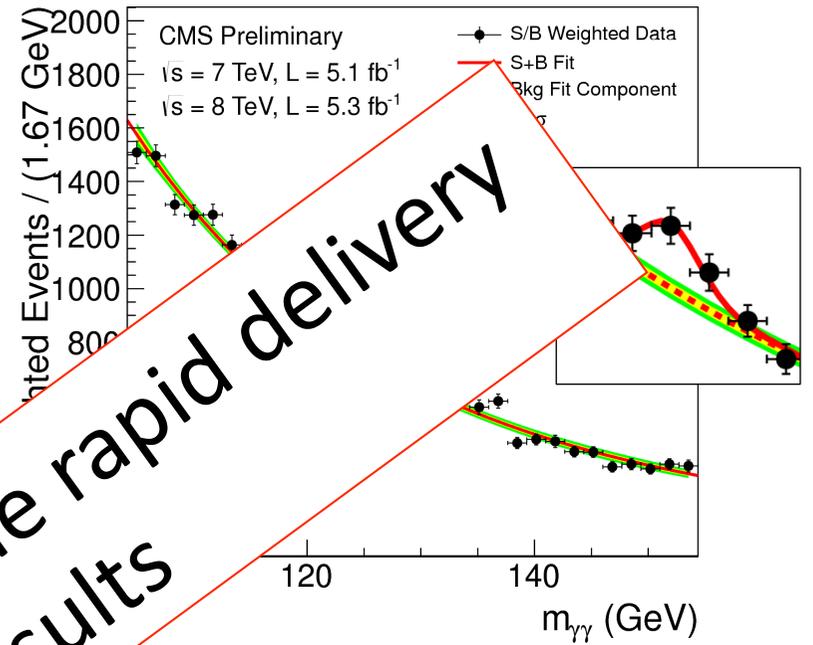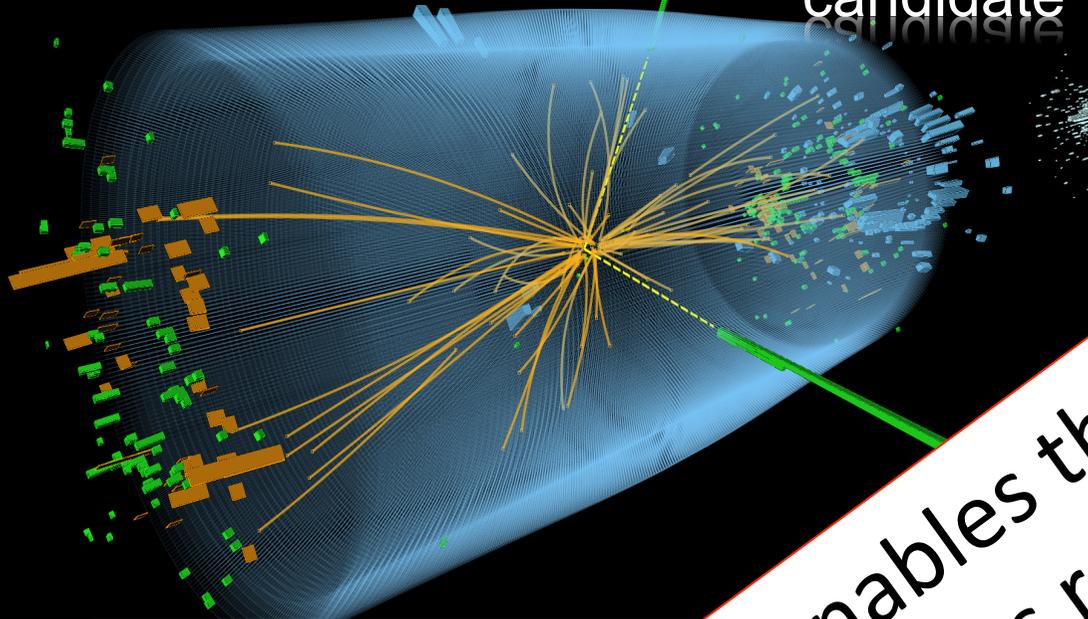Tier 2 cumulative CPU time delivered by Country (Jan 2012 - Sep 2012)

Spain (the 6th)

Legend:
- USA
- UK
- France
- Germany
- Italy
- Spain
- Russian Federation
- Czech Republic
- Canada
- Slovenia
- Romania
- Poland
- Belgium
- Switzerland
- Portugal
- Japan
- Sweden
- China
- Taipei
- Israel
- Estonia
- Australia
- Finland
- Republic of Korea
- Norway
- Hungary
- Turkey
- Brazil
- Austria
- India
- Pakistan
- Ukraine
- Greece

Grid computing enables the rapid delivery of physics results
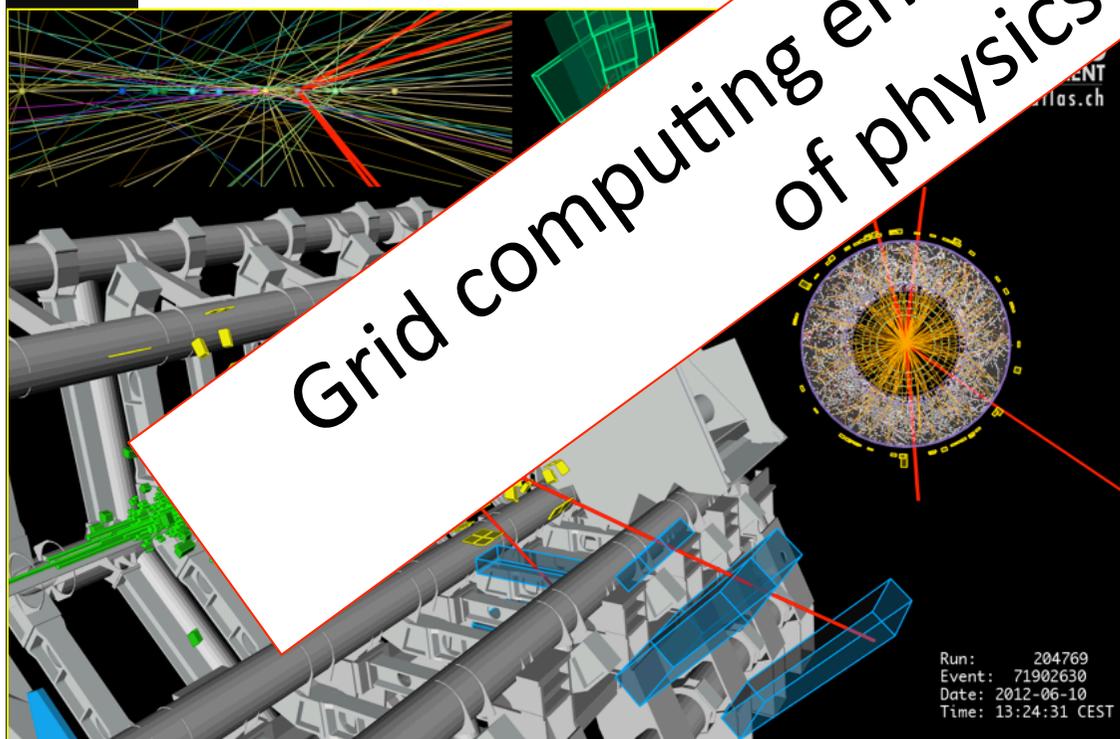
Impact of the LHC Computing Grid in Spain

ATLAS, CMS & LHCb Tier-1: PIC
ATLAS Federated Tier-2: IFIC, IFAE, UAM
CMS Federated Tier-2: IFCA, CIEMAT
LHCb Federated Tier-2: UB, USC

http://dashb-earth.cern.ch/

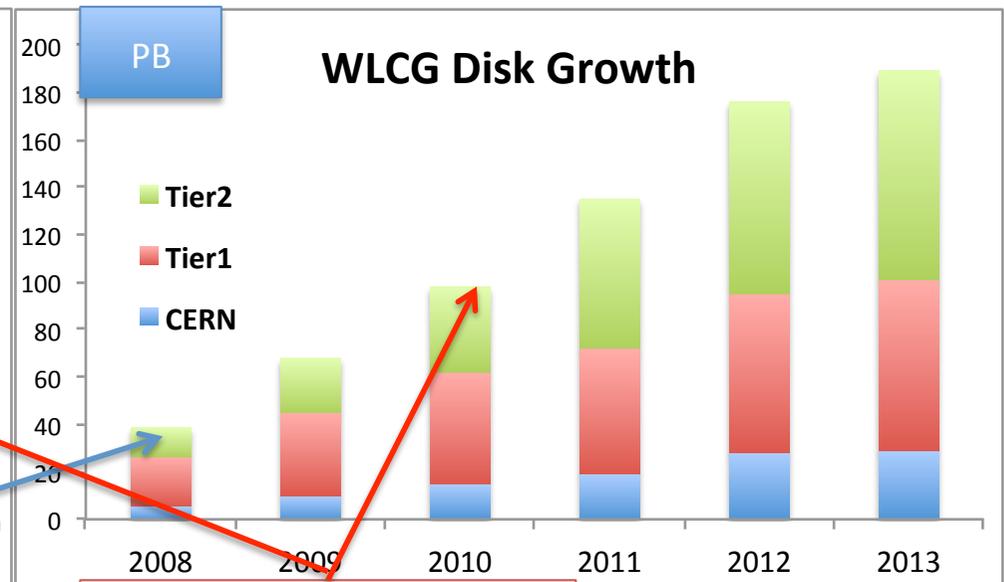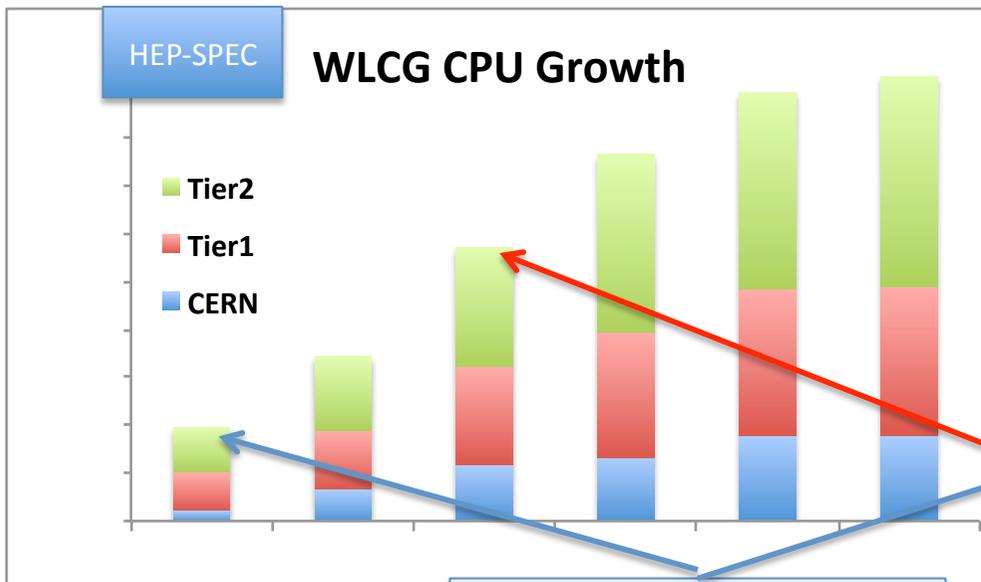# Evolution of capacity: CERN & WLCG

HEP-SPEC is the new HEP-wide benchmark for measuring CPU performance
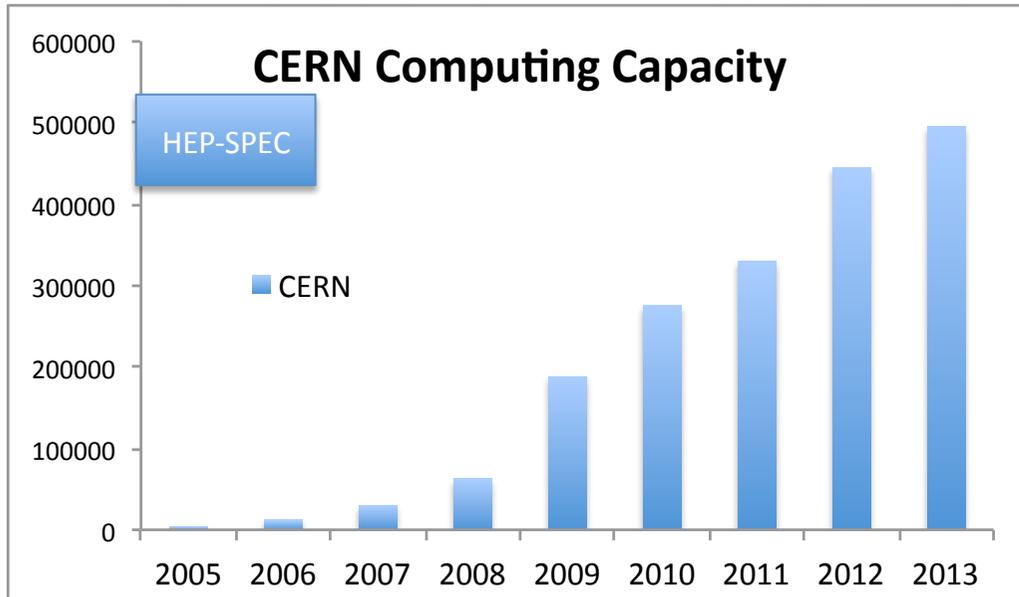
Modest growth until 2014

Anticipate x2 (?) requirement >2015

Anticipate x10-100 after 2018 ???

What we thought was needed at LHC start

What we actually used at LHC start!

PIC Tier-1

as of October 2012

PIC Current resources (1-Oct-2012)

CPU (HS06)	36314
Disk (TB)	4164
Tape (TB)	4717

Tier1s has reached 4 PB disk & Tape Storage !

Federation Pledges

Year: 2012

Pledge Type	ALICE	% of Req.	ATLAS	% of Req.	CMS	% of Req.	LHCb	% of Req.	SUM	% of Req.
CPU (HEP-SPEC06)			13,209	5%	7,395	5%	5,763	5%	26,367	5%
Disk (Tbytes)			1,377	5%	1,122	5%	485	5%	2,984	5%
Tape (Tbytes)			1,836	6%	2,601	6%	306	5%	4,743	6%

Showing 1 to 3 of 3 entries

Evolution of the resources for CPU, Disk and Tape for the Spanish Tier1

PIC Pledges:

	2009	2010	2011	2012	2013	2014
CPU (HS06)	10364	17238	23272	26367	30804	33558
Disk (TB)	1702	1968	2438	2984	3550	3692
Tape (TB)	1844	2136	4234	4743	5345	6370

PIC Tier-1

People

Engineers:

- Esther Acción
- Marc Caubet
- Fernando López
- Elena Planas
- Xavier Espinal (now at CERN IT Department)
- Paco Martínez (now at Telefónica I+D)
- Gerard Bernabeu (now at Fermilab CD Department)
- Arnau Bria (now at Centre de Regulació Genòmica)
- Bruno Rodriguez (recent hire)
- Carles Acosta (recent hire)
- Jordi Casals (recent hire)
- Vanessa Acín
- Technician: Ricard Cruz

Tier-1 Liaisons:

- Víctor Méndez (LHCb)
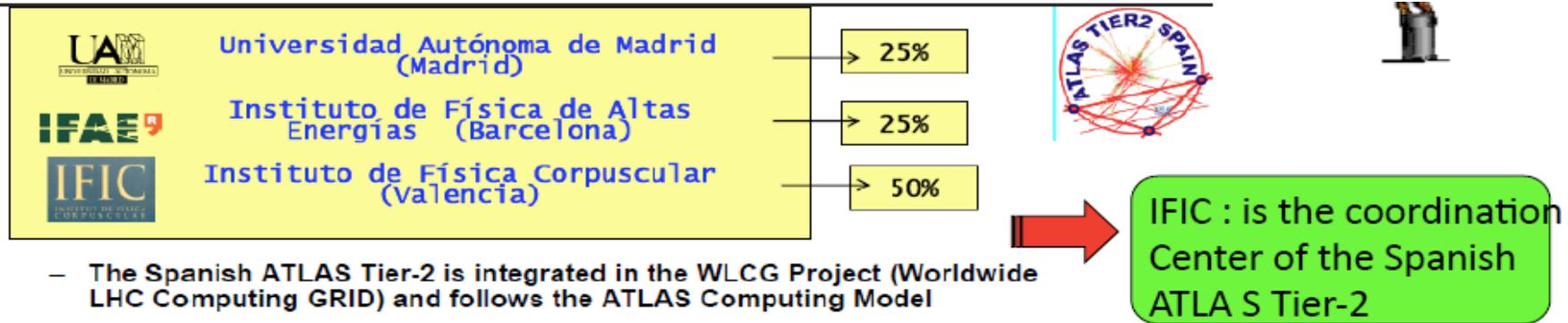- Alexei Sedov (ATLAS)
- Antonio Pérez (CMS)

Senior Staff:

- Gonzalo Merino (IP Tier-1 project)
- Manuel Delfino (co-IP Tier-1 project, PIC Director)
- Andreu Pacheco
- Mari Carmen Porto

PIC Computer Center today

ATLAS Federated Tier-2

Universidad Autónoma de Madrid (Madrid) → 25%

Instituto de Física de Altas Energías (Barcelona) → 25%

Instituto de Física Corpuscular (Valencia) → 50%

IFIC : is the coordination Center of the Spanish ATLA S Tier-2

- The Spanish ATLAS Tier-2 is integrated in the WLCG Project (Worldwide LHC Computing GRID) and follows the ATLAS Computing Model

as of October 2012

Deployed Equipment:

ES-ATLAS-T2	CPU(HEP-SPEC06)	DISK(TB)
IFIC	7066	1248
UAM	3780	555
IFAE	4950	579
Total	15796	2382

Tier2s has reached 2 PB disk Storage !

14 FTE

as of October 2012

¡The Spanish ATLAS Tier-2 has Got 2 PB Disk Storage!

Table 5.1: The evolution of the pledges in CPU and disk for the ATLAS Spanish Tier-2

ES-ATLAS-T2	2006	2007	2008	2009	2010	2011	2012	2013	2014
CPU(HEP-SPEC06)	92	243	1750	5390	10308	13900	13300	16050	22200
DISK(TB)	14	63	387	656	1107	1880	2350	2800	3000

22

Federation Pledges

Year: 2012

Pledge Type	ATLAS	% of Req.
CPU (HEP-SPEC06)	13,300	5%
Disk (Tbytes)	2,350	5%
Showing 1 to 2 of 2 entries		

ATLAS Federated Tier-2

People

- José Salt (IP project & coordinator, IFIC)
- Javier Sánchez (IFIC)
- Victor Lacort (IFIC)
- Álvaro Casani (IFIC)
- Farida Fassi (IFIC)
- Mohammed Kaci (IFIC)
- Miguel Villaplana (IFIC)
- Elena Oliver (IFIC)
- Santiago González (IFIC)
- Andreu Pacheco (IP project, IFAE)
- Carlos Borrego (IFAE)
- Ricard Cruz (IFAE)
- Jordi Nadal (IFAE)
- José del Peso (IP project, UAM)
- Sofia Sayzhenkova (UAM)
- Juan José Pardo (UAM)
- Francisco Gallardo (UAM)

IFIC Computer Center today

UAM Computer Center today

IFAE-PIC Computer Center today

CMS Federated Tier-2

Installed Capacities

Year: 2012   Month: 10

Infrastructure	Site Name	Physical CPU	Logical CPU	HEPSPEC06	Disk (GB)	Tape (GB)
EGEE	CIEMAT-LCG2	364	1,240	14,595	916,813	
EGEE	IFCA-LCG2	420	1,680	14,146	1,768,105	
Total		784	2,920	28,741	2,684,918	

Showing 1 to 2 of 2 entries

Tier2s has reached 1 PB disk Storage !

Federation Pledges

Year: 2012

Pledge Type	CMS	% of Req.
CPU (HEP-SPEC06)	15,750	5%
Disk (Tbytes)	1,300	5%

Showing 1 to 2 of 2 entries

IFCA Instituto de Física de Cantabria — 50 %

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas — 50 %

Evolution of the resources for CPU, and Disk for the Spanish Tier2

CMS Pledges:		2009	2010	2011	2012	2013	2014
	CPU (HS06)	5120	9040	16000	15750	20000	20000
	Disk (TB)	420	670	1000	1300	1500	1500

CMS Federated Tier-2

People

- Francisco Matorras (IP project, IFCA)
- Nicanor Colino (IP project, Ciemat)
- José María Hernández (Ciemat)
- Antonio Delgado (Ciemat)
- Javier Calonge (Ciemat)
- Juan José Rodríguez (Ciemat)
- Miguel Cárdenas (Ciemat)
- Álvaro García (IFCA)
- Pablo Orviz (IFCA)
- Ibán Cabrillo (IFCA)
- Rafael Marco (IFCA)
- Miguel Ángel Núñez (IFCA)
- Ana Rodríguez (IFCA)
- Jesús Marco (IFCA)
- Isidro González (Oviedo)
- Javier Fernández (Oviedo)

IFCA Computer Center today

Ciemat Computer Center today

LHCb Federated Tier-2

Installed Capacities

Infrastructure	Site Name	Physical CPU	Logical CPU	HEPSPEC06	Disk (GB)	Tape (GB)
Year: 2012		Month: 10				
EGEE	UB-LCG2	194	776	5,432	1,476	0
EGEE	USC-LCG2	170	942	7,065	984	0
Total		364	1,718	12,497	2,460	0

Showing 1 to 2 of 2 entries

Federation Pledges

Year: 2012

Pledge Type	LHCb	% of Req.
CPU (HEP-SPEC06)	2,800	7%
Disk (Tbytes)	1	

Showing 1 to 2 of 2 entries

UB Universitat de Barcelona — 40 %

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA — 60 %

Evolution of the resources for CPU, and Disk for the Spanish Tier2

LHCb Pledges:	2009	2010	2011	2012	2013	2014
CPU (HS06)	2960	3420	2340	2800	2800	2800
Disk (TB)	1	1	1	1	1	1

LHCb Federated Tier-2

People

- Ricardo Graciani (IP project, UB)
- Juan José Saborido (IP project, USC)
- Marcos Seco (USC)
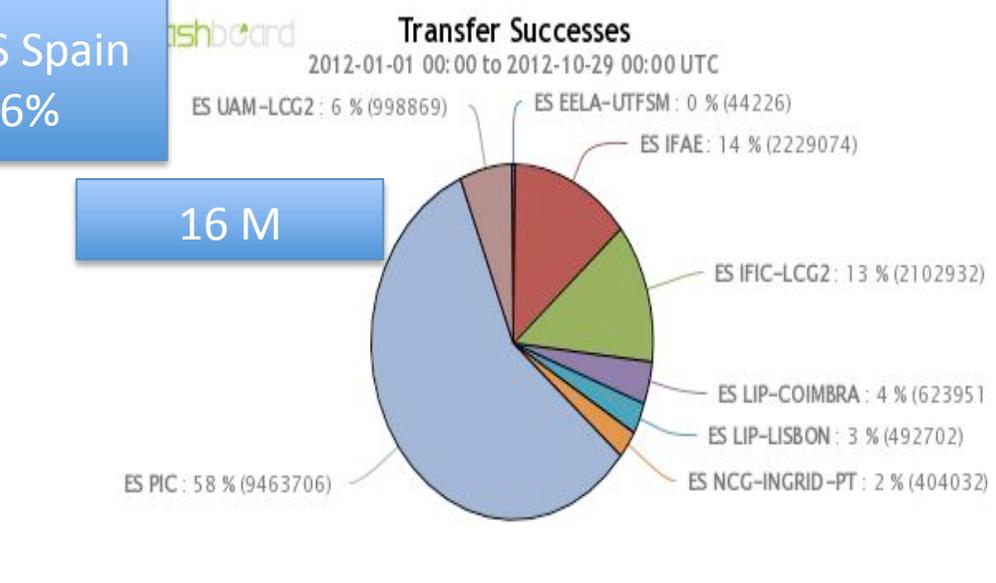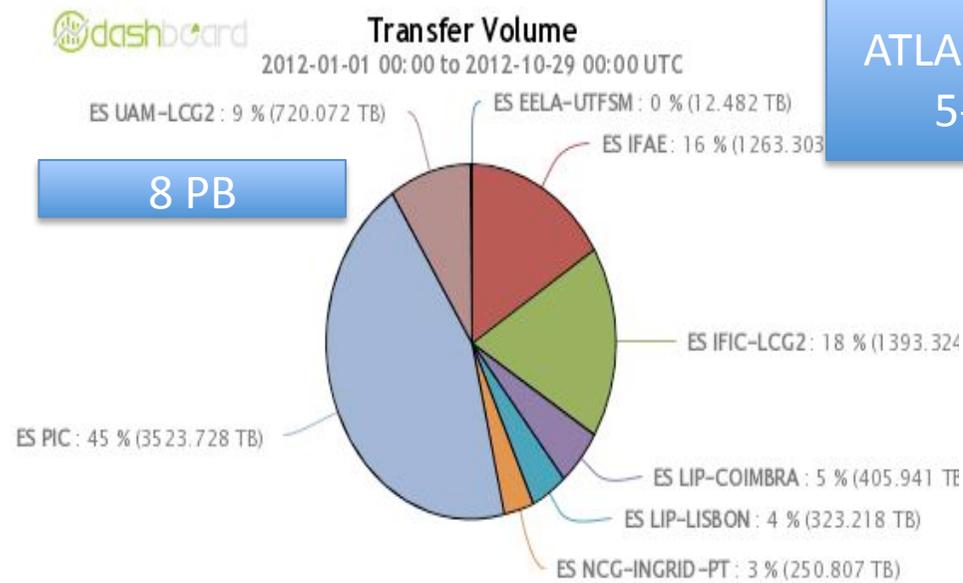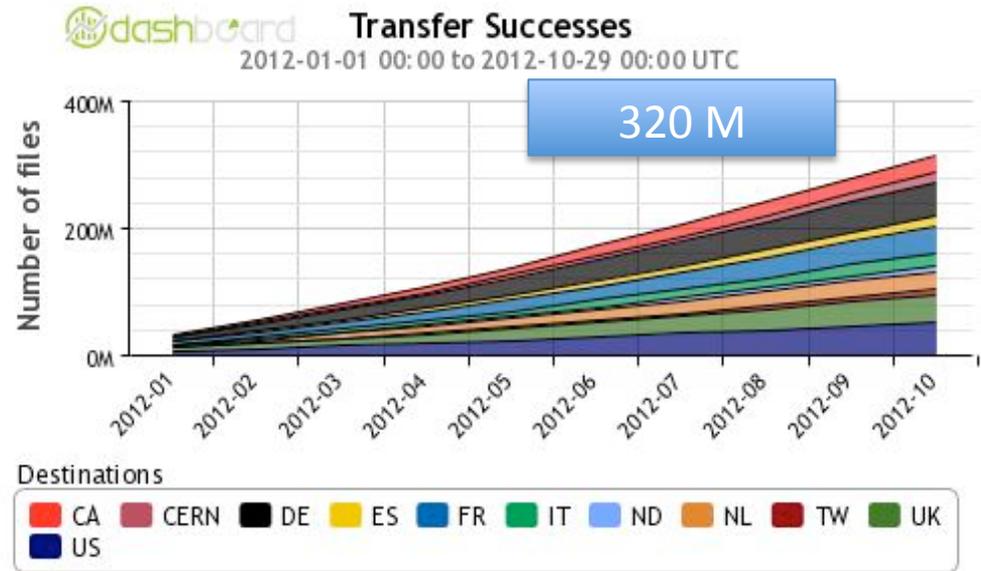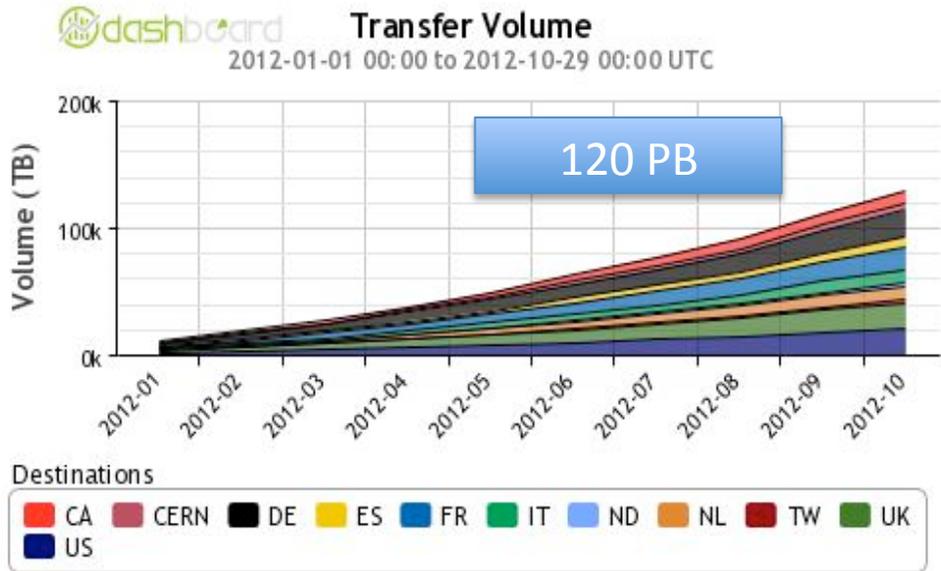- A. Casajus (UB)
- Victor Fernández (USC)
- S. Gómez (UB)

USC Computer Center today

UB Computer Center today

Performances: Data Distribution 2012

ATLAS

Transfer Volume
2012-01-01 00:00 to 2012-10-29 00:00 UTC

120 PB

Volume (TB)

Transfer Successes
2012-01-01 00:00 to 2012-10-29 00:00 UTC

320 M

Number of files

Destinations
CA CERN DE ES FR IT ND NL TW UK US

Destinations
CA CERN DE ES FR IT ND NL TW UK US

Transfer Volume
2012-01-01 00:00 to 2012-10-29 00:00 UTC

ATLAS Spain 5-6%

ES UAM-LCG2 : 9 % (720.072 TB)

ES EELA-UTFSM : 0 % (12.482 TB)

ES IFAE : 16 % (1263.303

8 PB

ES IFIC-LCG2 : 18 % (1393.324

ES PIC : 45 % (3523.728 TB)

ES LIP-COIMBRA : 5 % (405.941 TB

ES LIP-LISBON : 4 % (323.218 TB)

ES NCG-INGRID-PT : 3 % (250.807 TB)

Transfer Successes
2012-01-01 00:00 to 2012-10-29 00:00 UTC

ES UAM-LCG2 : 6 % (998869)

ES EELA-UTFSM : 0 % (44226)

ES IFAE : 14 % (2229074)

16 M

ES IFIC-LCG2 : 13 % (2102932)

ES LIP-COIMBRA : 4 % (623951)

ES LIP-LISBON : 3 % (492702)

ES PIC : 58 % (9463706)

ES NCG-INGRID-PT : 2 % (404032)

Performances: Data Distribution 2012

Performances: Data Distribution 2012

Performance: Jobs Distribution 2012

Performance: Jobs Distribution 2012

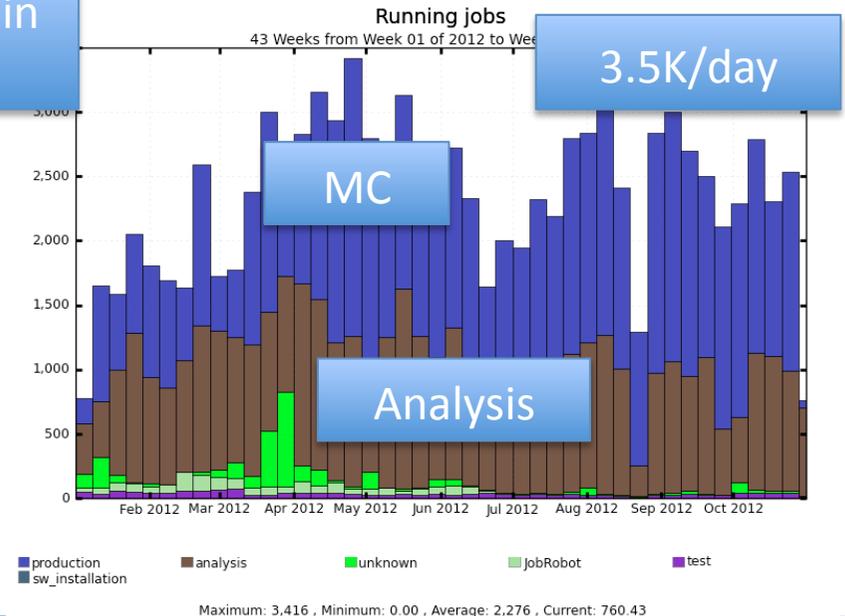# Performance: Jobs Distribution 2012

LHCB

Cumulative Jobs by JobType
43 Weeks from Week 52 of 2011 to Week 44 of 2012

10 M!!!

MCSimulation	3.9	DataReprocessing	0.6	sam	0.2	unknown	0.0
user	3.5	DataReconstruction	0.5	DataSwiming	0.1		
DataStripping	1.0	Merge	0.4	WGProduction	0.0		

Max: 10.2, Min: 0.08, Average: 4.88, Current: 10.2

Generated on 2012-10-30 18:08:02

Running jobs by JobType
43 Weeks from Week 52 of 2011 to Week 44 of 2012

25K/day

LHC stop

MC

Analysis

Max: 26.3, Min: 0.56, Average: 12.0, Current: 11.9

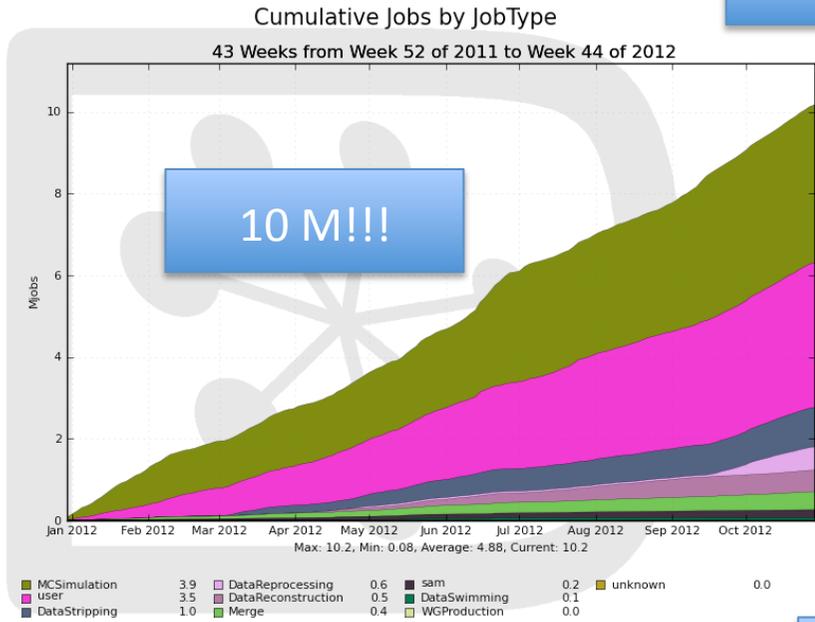|					
MCSimulation	38.0%	DataStripping	12.3%	WGProduction	0.0%
DataReconstruction	19.1%	Merge	0.8%	unknown	0.0%
DataReprocessing	15.4%	DataSwiming	0.3%		
user	13.9%	sam	0.1%		

Generated on 2012-10-30 18:05:34 UTC

Cumulative Jobs by Country
43 Weeks from Week 52 of 2011 to Week 44 of 2012

LHCB Spain 4-6%

0.6 M!!!

ES

UK	1.8	RU	0.9	IE	0.1	CY	0.0
FR	1.7	DE	0.9	SU	0.1	HU	0.0
CH	1.5	ES	0.6	RO	0.1	ANY	0.0
IT	1.4	PL	0.1	BG	0.0	GR	0.0
NL	1.0	BR	0.1	IL	0.0	MULTIPLE	0.0

Max: 10.2, Min: 0.08, Average: 4.88, Current: 10.2

Generated on 2012-10-30 18:11:54 UTC

Running jobs by Country
43 Weeks from Week 52 of 2011 to Week 44 of 2012

5% 1k/day

Max: 26.3, Min: 0.56, Average: 12.0, Current: 11.9

UK	17.2%	DE	9.3%	IE	0.6%	HU	0.0%
FR	17.0%	NL	7.5%	SU	0.6%	CY	0.0%
CH	15.0%	ES	5.0%	BR	0.5%	ANY	0.0%
IT	13.9%	PL	1.5%	BG	0.1%	GR	0.0%
RU	10.8%	RO	1.0%	IL	0.0%	MULTIPLE	0.0%

Generated on 2012-10-30 18:02:56 UTC

Performance: Availability & Reliability

WLCG availability = time_site_is_available / total_time

WLCG reliability = time_site_is_available / {total_time – time_site_is_scheduled_down}

Federation Site	Phy. CPU	Log. CPU	HS06	Relia bility	Availa bility	Unkn own	Reliability History Jun-12	Jul-12	Aug-12
ES-ATLAS-T2 (Spain, ATLAS Federation)									
IFIC-LCG2	286	1,144	8,471	93 %	93 %	0 %	98 %	100 %	95 %
UAM-LCG2	92	370	3,780	95 %	94 %	0 %	100 %	93 %	97 %
ifae	640	3,564	45,992	97 %	96 %	0 %	100 %	98 %	100 %
ES-CMS-T2 (Spain, CMS Federation)									
CIEMAT-LCG2	364	1,240	14,595	100 %	100 %	0 %	98 %	98 %	98 %
IFCA-LCG2	436	2,320	19,534	87 %	86 %	9 %	96 %	92 %	94 %
ES-LHCb-T2 (Spain, LHCb Federation)									
UB-LCG2	224	776	5,432	97 %	97 %	0 %	91 %	94 %	82 %
USC-LCG2	156	826	6,195	89 %	89 %	0 %	94 %	88 %	86 %

Detailed Monthly Site Reliability (OPS tests)

Site	Apr 12	May 12	Jun 12	Jul 12	Aug 12	Sep 12
CA-TRIUMF	99	99	100	100	100	98
CH-CERN	78	100	100	100	95	100
DE-KIT	100	99	100	100	100	99
ES-PIC	100	100	100	98	100	100
FR-CCIN2P3	100	100	100	100	100	100
GSDC-KISTI	79	86	85	94	91	94
IT-INFN-CNAF	100	98	100	100	100	100
NDGF	99	97	99	100	100	100
NL-T1	99	98	97	94	100	99
TW-ASGC	100	99	100	100	98	98
UK-T1-RAL	100	100	100	100	100	100
US-FNAL-CMS	100	100	100	100	99	100
US-T1-BNL	100	98	100	100	100	100
Target	97	97	97	97	97	97

Colors: Green > Target   Orange > 90%   Red < 90%

Last months **ES TIER-1 & -2s had > 90% availability!!!!!!**

Performance: Availability & Reliability 2012

ATLAS
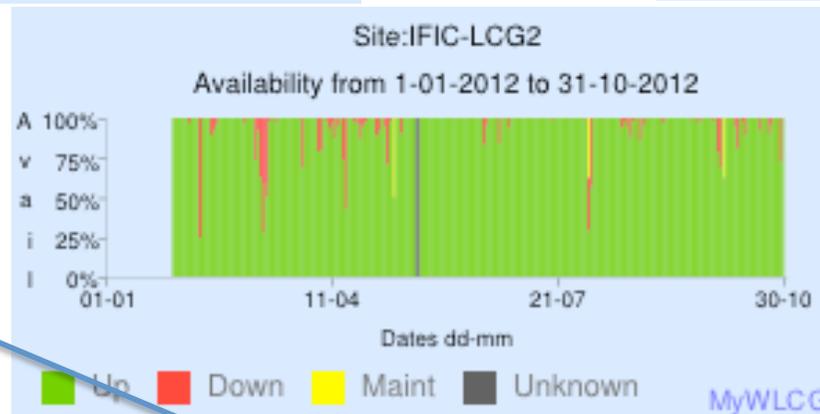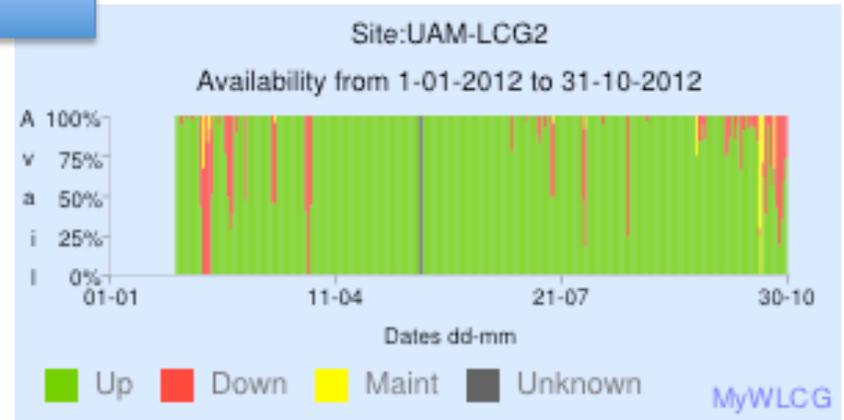
Hammercloud for ATLAS:
Sites are tested with 'typical' analysis & production jobs

- Good Availability & connectivity:
  a) To make direct transfers from Tier-2s to Tier-1s of different clouds. Even transfers with other Tie-2s globally
  b) Direct transfers from/to all ATLAS T1s and process data from many T1s
  c) **Candidate for more datasets replication**
  d) More Analysis jobs and "group production"

> 90%!!!
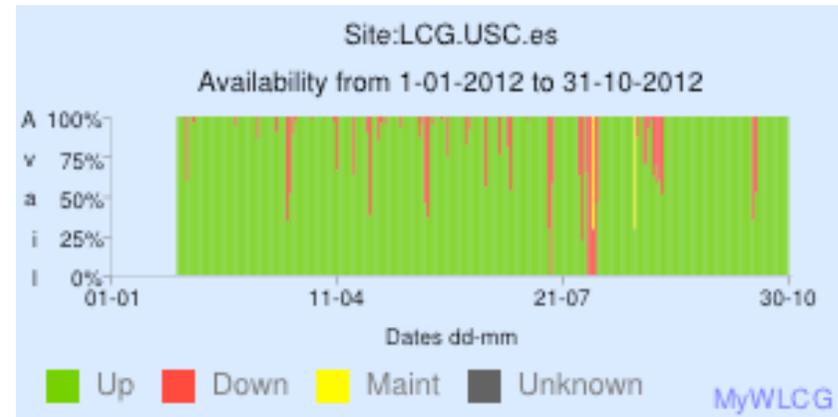
Performance: Availability & Reliability 2012

> **CMS**
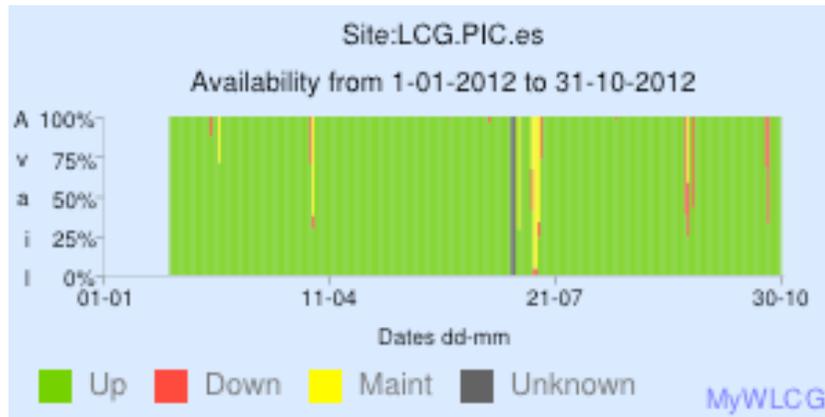
Readiness: Sites are tested with 'typical' experiment production & analysis jobs and data transfers
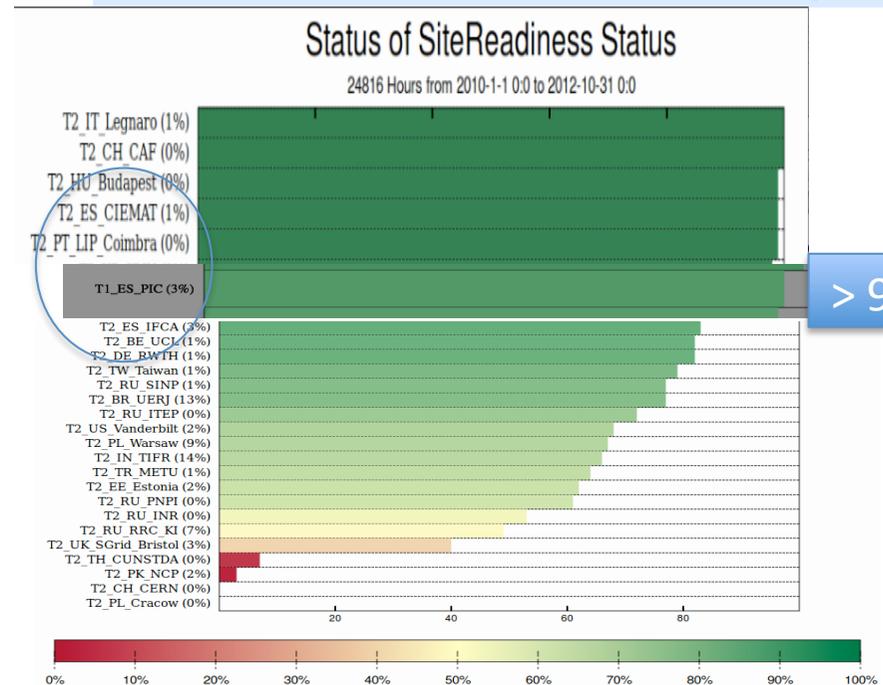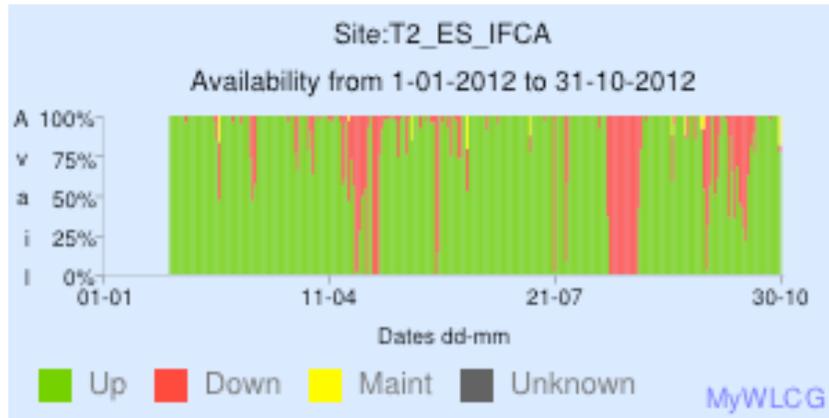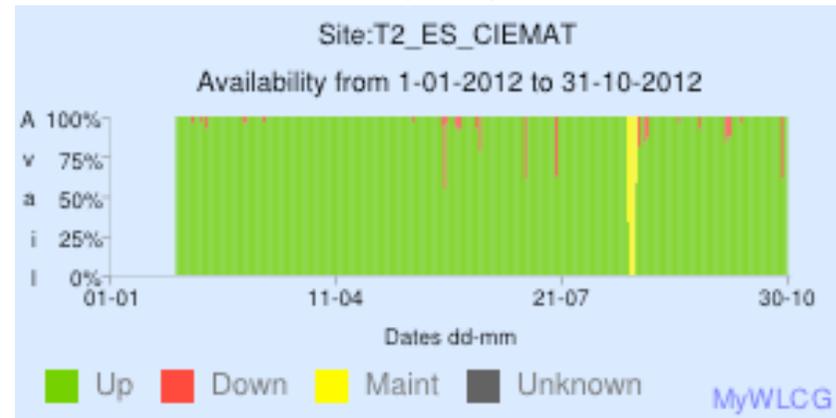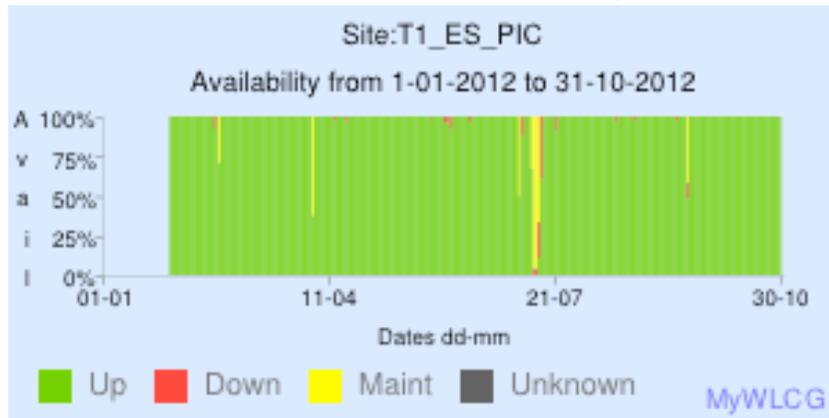
Performance: Availability & Reliability 2012

LHCb

Distributed analysis

ATLAS, LHCb & CMS has developed a specific tool for running distributed analysis: Ganga/Panda, CRAB

- users basically only have to define the dataset they want to analyze and the analysis algorithm.

- jobs are sent to the sites where the dataset is available: follow the data

- retrieves output

- Spanish people involved in these activities and taking over of this workflow

- Job finished successfully, output files are copied in the **Tier3s**

  - institution-level that participate presumably most frequently in support of the particular interests of local physicists.
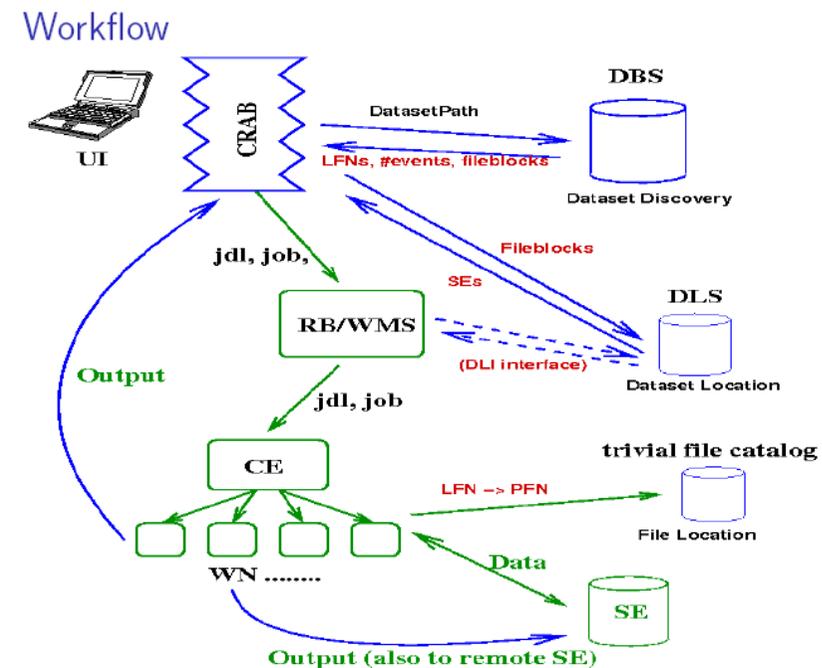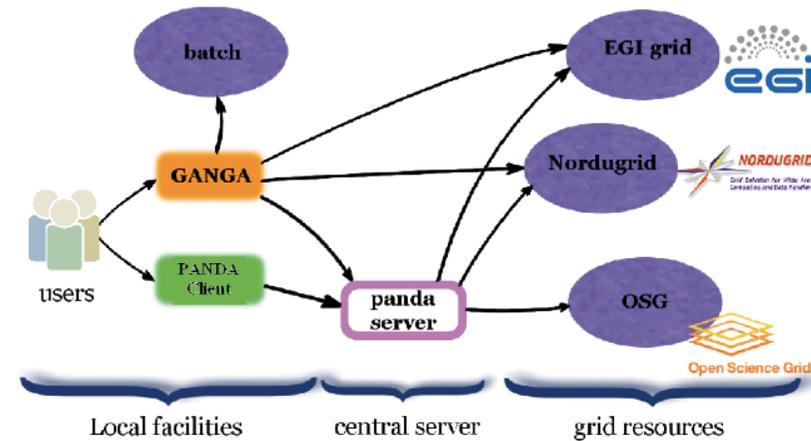
Future and Perspectives

- LHC:
  - The LHC running time increases from 21 to 30 weeks
  - The LHC efficiency is changed from 30% to about 40%, due to longer period of stable beams.
  - The extension of LHC running into 2013 leads to a delay in the resumption of LHC physic running until year 2015. So it well be no physics data taking in the 2014 (LS1).

- The increase in the volume of the data in 2012:
  - The lifetime of the current data ESD rolling buffer is reduced from 6 to 4 weeks. Volume on disk is reduced proportionally.
  - The assumed volume of data ESDs on disk after reprocessing in years 2013 and 2014 is reduced from 20% to 5% of the total data volume.
  - The increase in the needed group space in 2014 is reduced from 4 PB to 2 PB because of no data being taken in 2014, while group activity on 7 TeV and 8 TeV data should consolidate by then.
  - The simulation of 13 TeV collisions in 2014 in preparation for 2015 running will now be produced with full statistics in one campaign rather than two. Second campaign in 2015 to correct the simulation parameters.
  - All Grid computing resources, including Tier-0, will be available for throughout 2014 for simulation.
  - Calibration and Alignment jobs produce 0.3 PB/yr
  - Analysis and User activities is assumed to be constant over the year plus the simulated data.

- We need to make everything faster, more scalable, lower memory footprint, parallelized, threaded, vectorized – or we will be in trouble post LS1

Summary

- First years of LHC data – WLCG has helped deliver physics rapidly

- Just the start of a >20 year exploration of new physics

- Entering a phase of consolidation and evolution

  - LS1: opportunity for disruptive changes and scale testing of new technologies

- Challenges for computing – scale & complexity – will continue to increase dramatically

Summary

- Spanish Tier-1 & Tier-2s participate in more experiments critical activities (ATLAS; CMS; LHCb).
  - receiving data immediately (2 hours) to have a higher contribution to analysis activity.
  - availability is reported every month requiring a good connectivity for data transfer and production-analysis.
  - Thousand processing, reprocessing, calibration, etc. jobs run at Tier-1 every day.
  - thousands central production and analysis jobs run at Tier-2s every day
- Tier2 sites = Centres of physicist. We have the feedback directly from the users (Distributed Analysis Team)
- The Spanish computing resources for the LHC experiments are used and are very important for the physics analysis

THANKS TO THE SPANISH LCG COMMUNITY!!!!!!!!

Acknowledgements

- Many thanks to all colleagues, Manuel Delfino (PIC), Gonzalo Merino (PIC), Jose Hernández (CIEMAT), Nicanor Colino (CIEMAT), Francisco Matorras (IFCA), José Salt (IFIC), José del Peso (UAM), Andreu Pacheco (IFAE), Ricardo Graciani (UB), Juan José Saborido (USC), Víctor Fernández (USC) for slides and information used in this talk
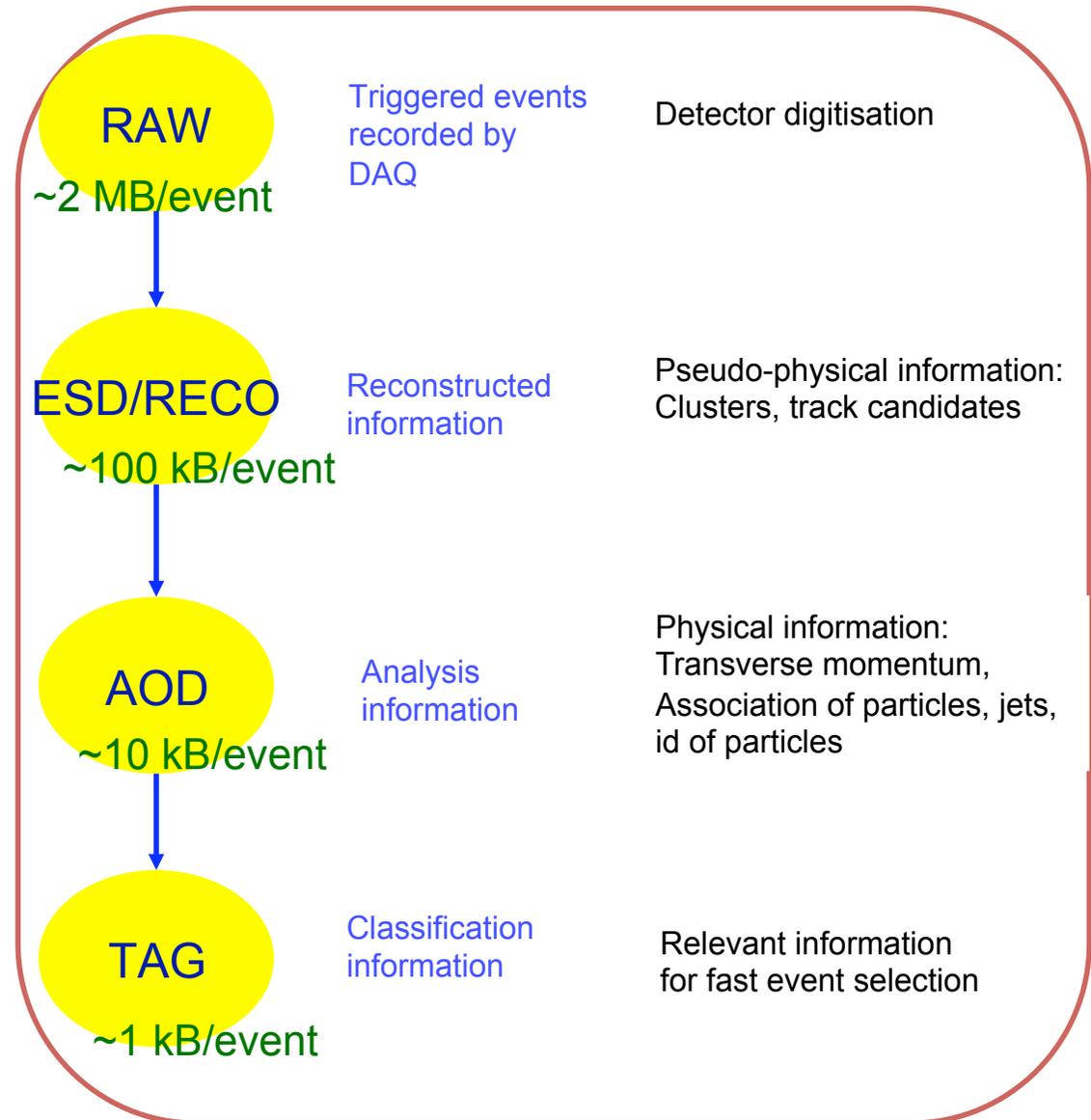
"Si lo que he contado les parece claro y transparente, han debido intrepretar erroneamente lo que he dicho"
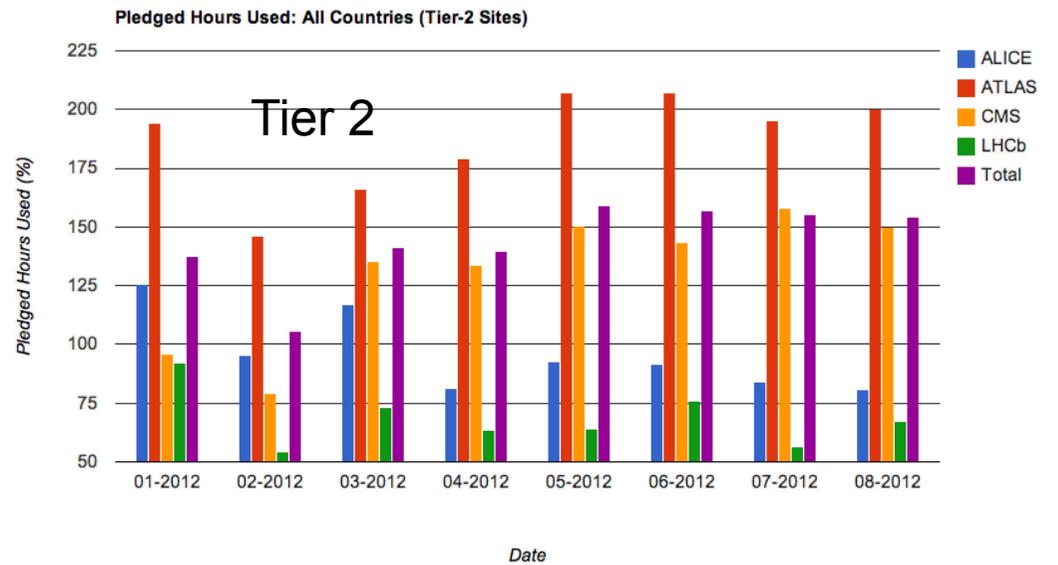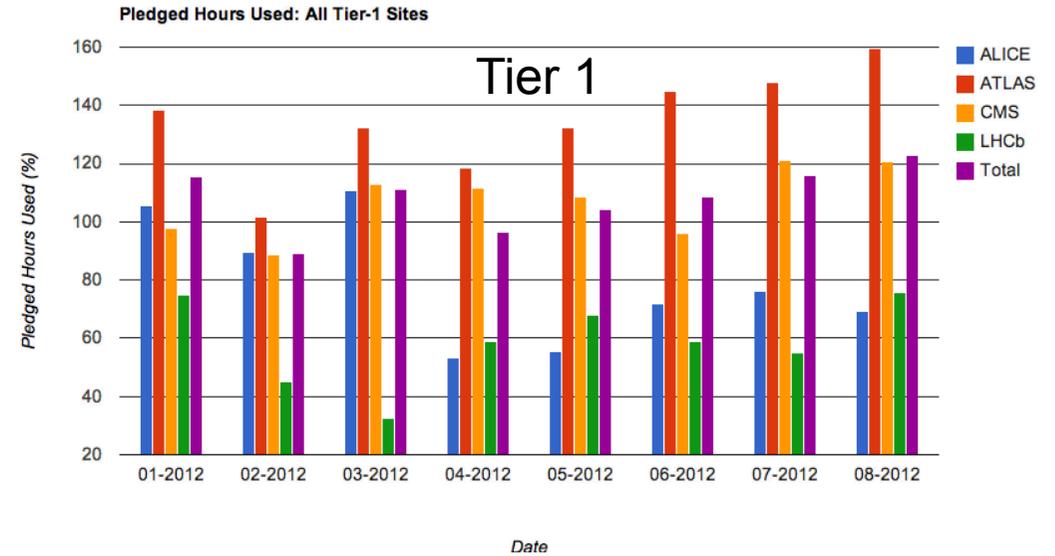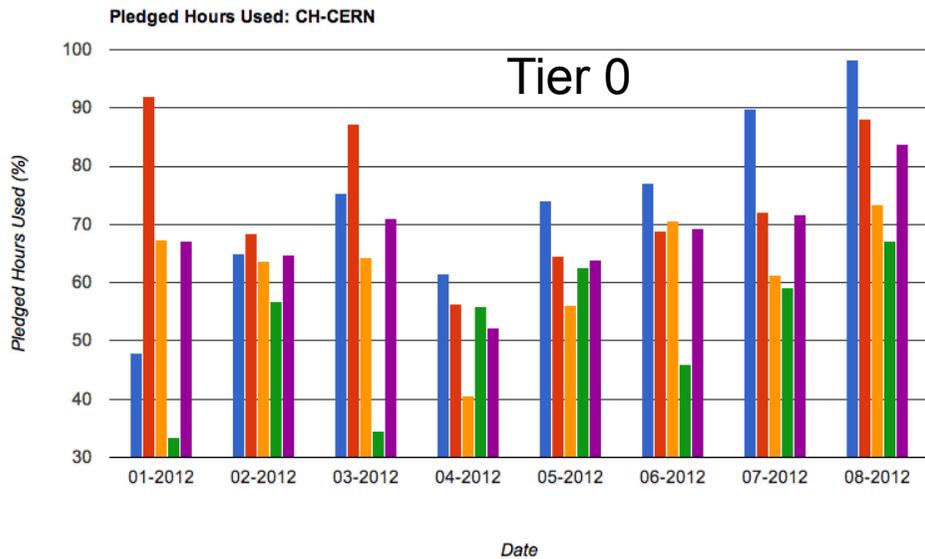ALAN GREENSPAN

Backup

Data and Algorithm

- HEP data are organized as *Events* (particle collisions)
- Simulation, Reconstruction and Analysis programs process "one event at a time"
  - Events are fairly independent → Trivial parallel processing
- Event processing programs are composed of a number of Algorithms selecting and transforming "raw" event data into "processed" (reconstructed) event data and statistics

RAW
~2 MB/event — Triggered events recorded by DAQ — Detector digitisation

ESD/RECO
~100 kB/event — Reconstructed information — Pseudo-physical information: Clusters, track candidates

AOD
~10 kB/event — Analysis information — Physical information: Transverse momentum, Association of particles, jets, id of particles

TAG
~1 kB/event — Classification information — Relevant information for fast event selection

Comparison: use/pledge

Infrastructure (see Javier's talk)

HEP-SPEC is the new HEP-wide benchmark for measuring CPU performance

#	CPU	Cores	Mem/core	HEP-SPEC2006/box	HEP06 TOT
32	Xeron E5472 3.00GHz	8	2 GB	73.62	2355.8
32	Xeon L5527 2.27 GHz	8	3 GB	147.20	4710.4
					7066.2
				Pledge 2011	6650

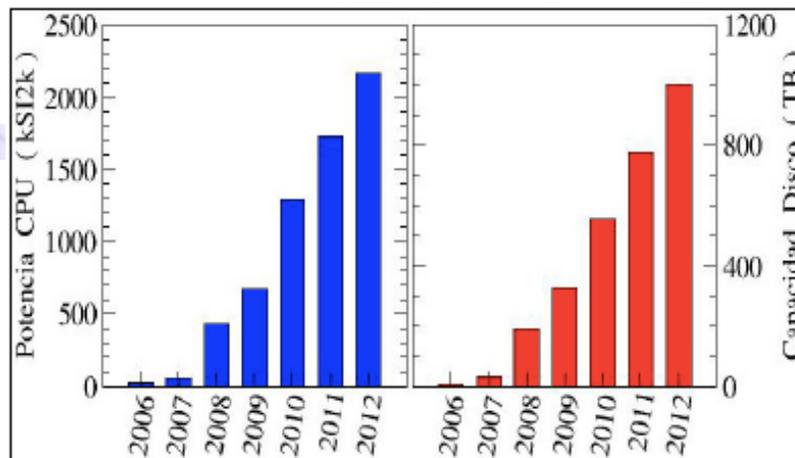SUN X4500: 5x(500GB) + 1x(1TB)   112 TB
SUN X4540: 14x(1TB)   476 TB
SuperMicro: 11x(2TB)   633 TB

TOT: 1221 TB

Pledge 2012   1175 TB

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

Atos · Capgemini · CloudSigma · CSA cloud security alliance · EGI · interoute · logica · OpenNebula.org · orange Business Services · SAP · the SERVER LABS · sixsq · Telefónica · terradue 20 · THALES · Trust it · T··Systems

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

CMS performance plots
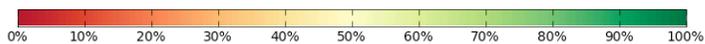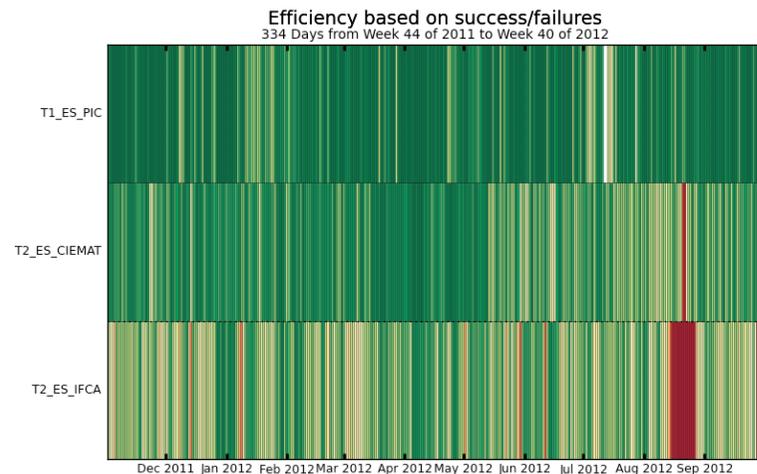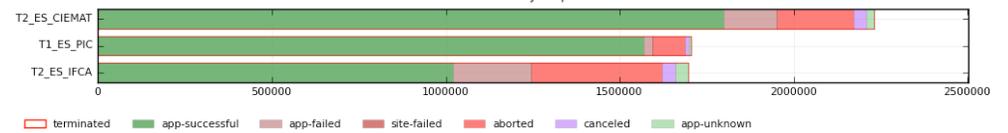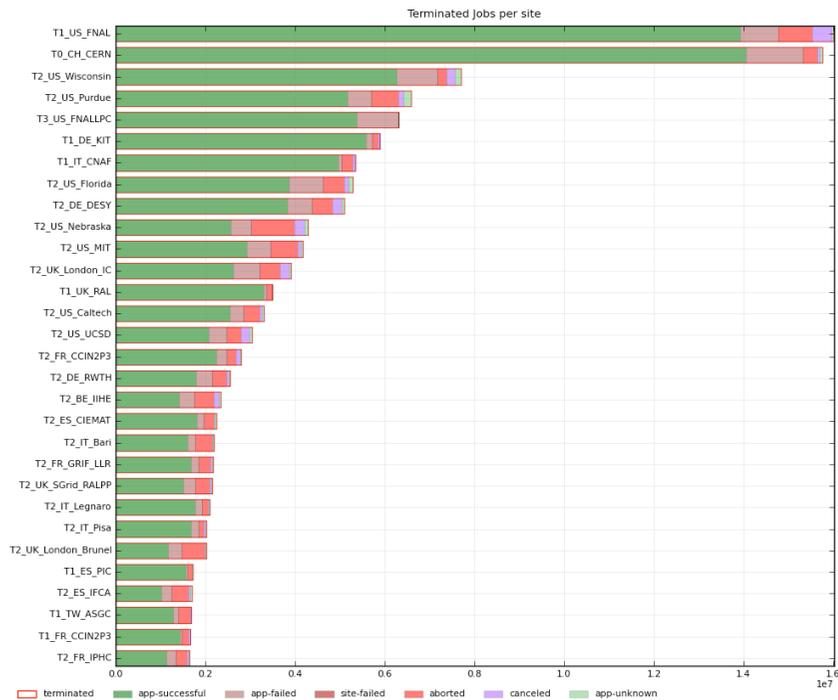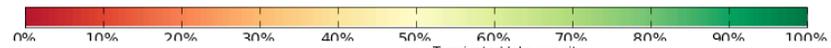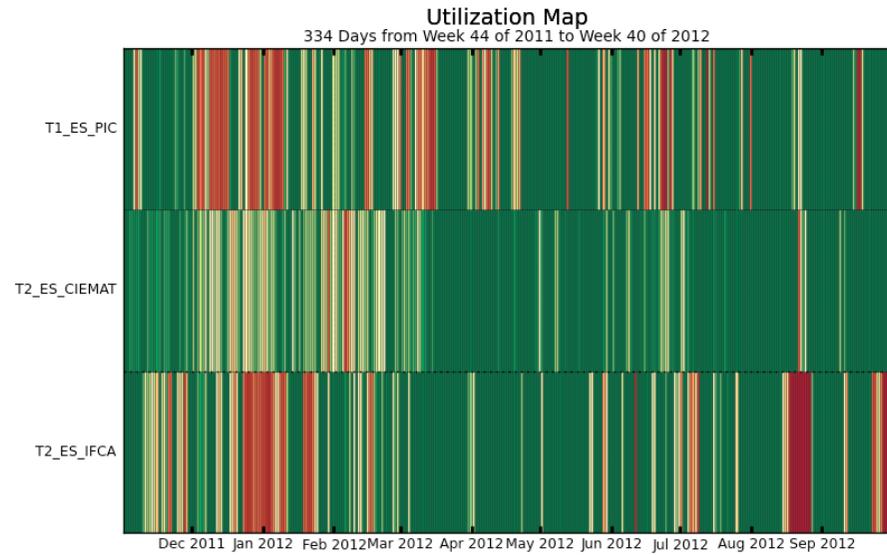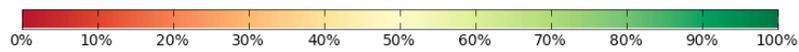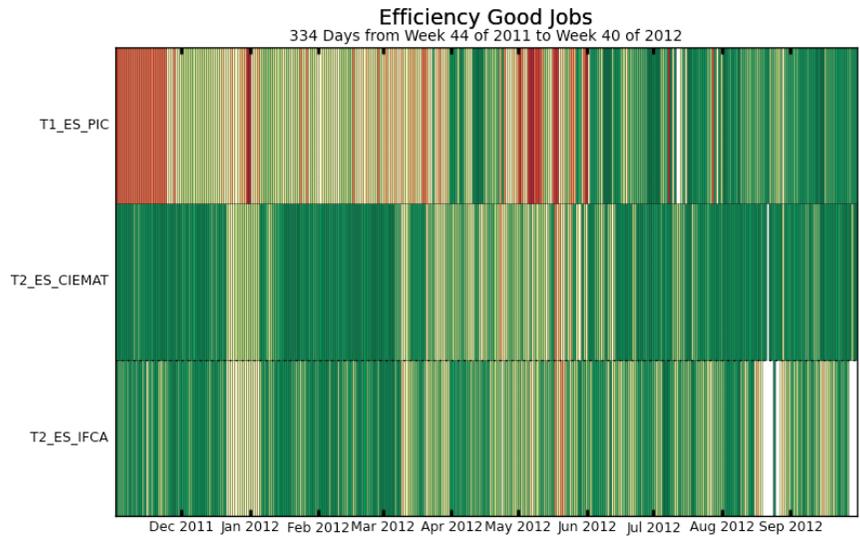
Future and Perspectives

- We have a grid because:
  - We need to *collaborate* and *share* resources
  - Thus we will always have a "grid"
  - Our network of trust is of enormous value for us and for (e-)science in general
- We also need distributed data management
  - That supports very high data rates and throughputs
  - We will continually work on these tools
- How much can come from open source community (or ...)?
  - We should determine what we can actually take from the open source (or commercial) world
  - Focus our (shrinking) efforts on the areas that we must develop ourselves
- But continue to question when/why we need to do that

The promise of cloud technology……

- Use of technology
  - Virtualisation
  - New "standard" interfaces (well, maybe one day)

- Services
  - Academic clouds
    - Grid → cloud? (or grids & clouds co-exist)
  - Commercial clouds
    - Outsourcing of services
    - Use for data processing, storage, analysis

- New types of services, new ways of providing services

PIC Tier-1

PIC port d'informació científica

as of October 2012

Installed Capacities

| Year: | 2012 | Month: | 10 | | | |

Infrastructure	Site Name	Physical CPU	Logical CPU	HEPSPEC06	Disk (GB)	Tape (GB)
EGEE	pic	640	3,564	45,992	5,758,424	7,357,035
Total		640	3,564	45,992	5,758,424	7,357,035

Showing 1 to 1 of 1 entries

PIC Current resources (1-Oct-2012)	
CPU (HS06)	36314
Disk (TB)	4164
Tape (TB)	4717

Federation Pledges

Year: 2012

Pledge Type	ALICE	% of Req.	ATLAS	% of Req.	CMS	% of Req.	LHCb	% of Req.	SUM	% of Req.
CPU (HEP-SPEC06)			13,209	5%	7,395	5%	5,763	5%	26,367	5%
Disk (Tbytes)			1,377	5%	1,122	5%	485	5%	2,984	5%
Tape (Tbytes)			1,836	6%	2,601	6%	306	5%	4,743	6%

Showing 1 to 3 of 3 entries

Evolution of the resources for CPU, Disk and Tape for the Spanish Tier1

PIC Pledges:	2009	2010	2011	2012	2013	2014
CPU (HS06)	10364	17238	23272	26367	30804	33558
Disk (TB)	1702	1968	2438	2984	3550	3692
Tape (TB)	1844	2136	4234	4743	5345	6370

Services Overview

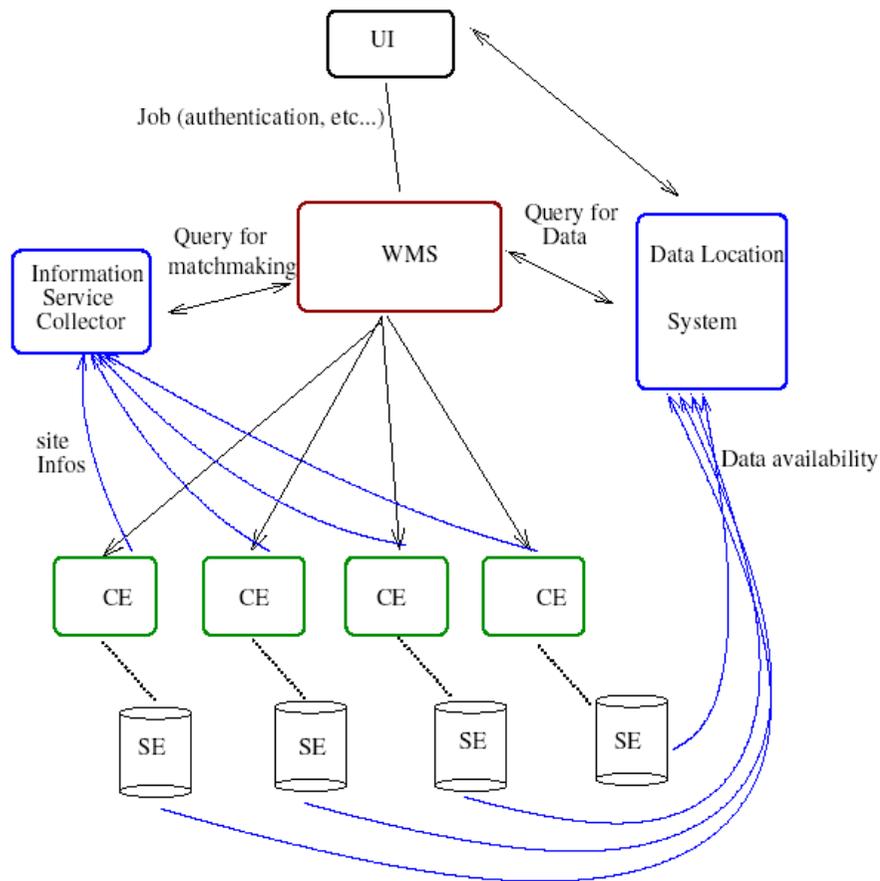Will not discuss here too much about services

Mostly based on WLCG services

• Additional tools developed by CMS

Grid Workload Management Systems
- Grid Workload Management
- Grid Job Logging and Bookkeeping
- Grid Monitoring

CMS Workflow Management
- Job Bookkeeping and Monitoring
- Data Production
- Analysis Tools
- Monte Carlo Production
- Parameter Set Database
- Other Misc Support

CMS Data Management
- Dataset Bookkeeping
- Data Location
- Data Transfer and Placement
- Local File Catalogs
- Data Access and Storage

F. Matorras, IFCA, Computing @ LHC: the CMS case
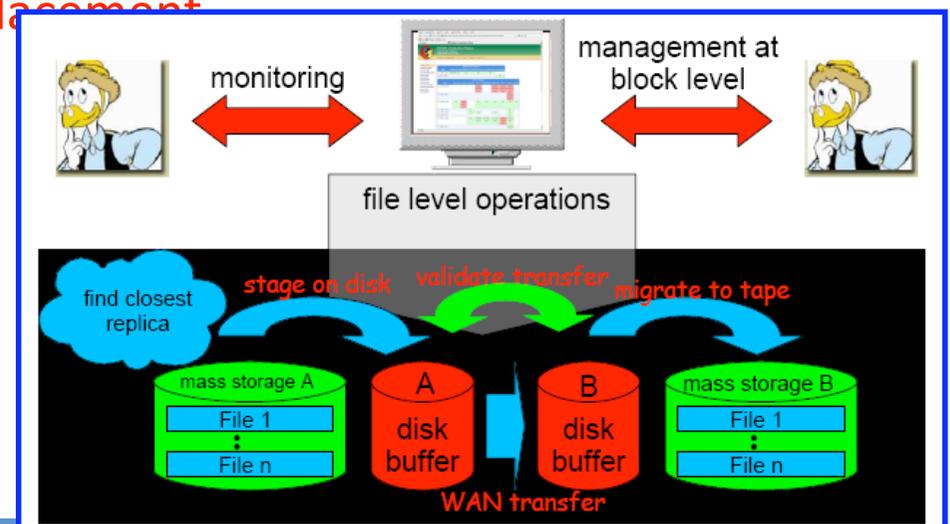
WMS & DMS Services Overview

- Data Bookkeeping System
  - "What data exist?"
- Data Location Service
  - "Where are data located?"
  - looks SE where a dataset is available
- Local File catalogue
  - provides the local file address
- Data Access and Storage
  - SRM and posix-IO
- Data Transfer and placement system

DMS

WMS

- Rely on Grid Workload Management
  - Reliability, performance, accounting
- Grid and CMS-specific job monitoring and bookkeeping
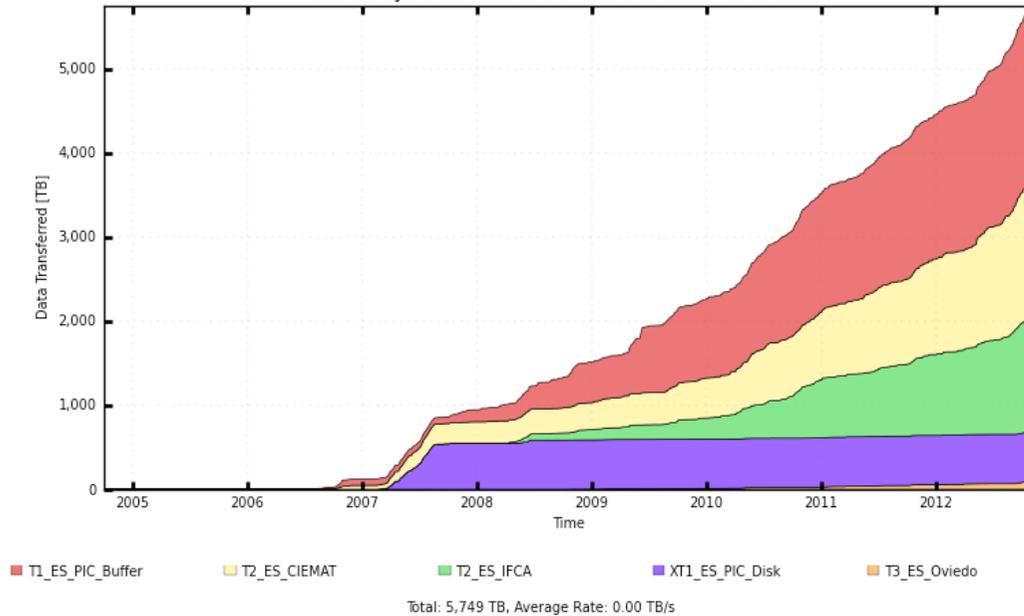
Data Transfer and Placement System

- **PhEDEx** (Physics Experiment Data Export)
- Large scale and wide area dataset replica management system
- Manages data flow following the transfer topology
- Reliable point-to-point transfers based on Grid transfer tools
- Nodes subscribe for data allocated from other nodes
    - data obtained from the "closest" available site
- Enables distribution management at dataset level rather than at file level
- Implements experiment policy on data placement
    - data placements drives analysis
- Allows prioritization and scheduling

CMS performance

CMS PhEDEx - Cumulative Transfer Volume
2958 Days from Week 39 of 2004 to Week 45 of 2012

Total: 5,749 TB, Average Rate: 0.00 TB/s

Terminated jobs
52 Weeks from Week 44 of 2011 to Week 44 of 2012

Maximum: 211,095 , Minimum: 0.00 , Average: 114,006 , Current: 36,106