Computación para el LHC

J.M. Hernández

"The Worldwide LHC Computing Grid (WLCG) has been a cornerstone of the success of the current LHC programme and represents a major achievement of the global particle physics community" (CERN's strategy 2026-2030)

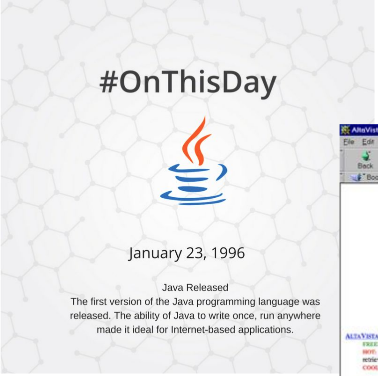
The situation

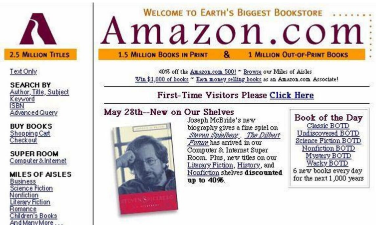
- LHC construction was approved on 16 Dec 1995
- With a target date for first beams of 2005
- For previous accelerators CERN had provided much of the computing resources
    - LEP total data volume (1989–2000) ~0.5 PB
- But there was no special budget for computing
- Would the computer centre operations budget suffice?
    - LHC scale $\mathcal{O}$(100 PB) annually
- If not, would the member states be willing to fund at CERN the shortfall?

The technological context (1996)

Work on computing for LHC started 30 years ago

- Altavista was the search engine of choice
  - Two years before Google first appeared
- The Nokia 8110 was the best cell phone available
  - With the novelty of SMS text messaging
  - 10 years before the first iPhone
- Amazon had been selling books online for a year, but it would be a further 10 years before Amazon announced the first cloud services
- First version of Java released
- A good home network connection was 64 kbits/sec
- A high-speed inter-site network was 622 Mbits/sec
- A big hard disk held 12 GB (5.25” format)
- The Pentium Pro (150 MHz) was the first PC chip with decent floating point





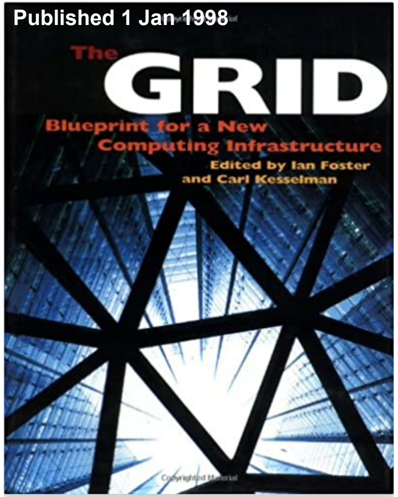
HEP computing in 1996

- With the advent of the Pentium Pro and the evolution to Linux the community was evolving rapidly towards a homogeneous PC/Linux model
- CERN and other sites had already 5 years experience of large clusters with integrated storage
- The tools had been developed to manage these clusters, distributed disk storage and, in the case of larger sites, mass storage (tape)
- And the scheduling of large numbers of jobs sharing access to large data volumes on a distributed facility

So, why not just go global,

absorbing capacity wherever it is available!

The Grid Concept (1997)



Meanwhile, in the wider world, the Grid concept was attracting interest

- Carl Kesselman (USC) and Ian Foster (Chicago) had been working on the idea for some time
- Their Globus Project had developed an open-source toolkit that included the basic functions needed to operate a computing Grid
- Early HEP pilots gave some confidence that distributed computing services for HEP would be feasible on the timescale of LHC
  - The Particle Physics Data Grid in the US
  - A distributed simulation facility in INFN
- But these were early days, for the LHC Grid several big issues had not yet been demonstrated

Searching for a Solution

- The MONARC Project (Models of Networked Analysis at Regional Centres for LHC) was started in 1997 with a wide representation from the community
- To assess requirements, technologies and costs for LHC computing within the constraints
  - The large data volumes: $\mathcal{O}$(100 PB) annually
  - The flexibility to incorporate resources of different capacity and functionality spread across the world
  - The wide area network bandwidth that could be expected in 2005
  - Using an architecture that looked feasible to be implemented in the time available

The MONARC project

Conclusions March 2000

- A distributed hierarchical model is feasible: CERN (Tier-0) ➔ Tier-1s ➔ Tier-2s
  - Driven by networking and capabilities
- Willingness in many national communities to participate in developing, funding and deploying this model
- Requires development of software for management and operation of the distributed model and its components

Models of Networked Analysis at Regional Centres for LHC Experiments

(MONARC)

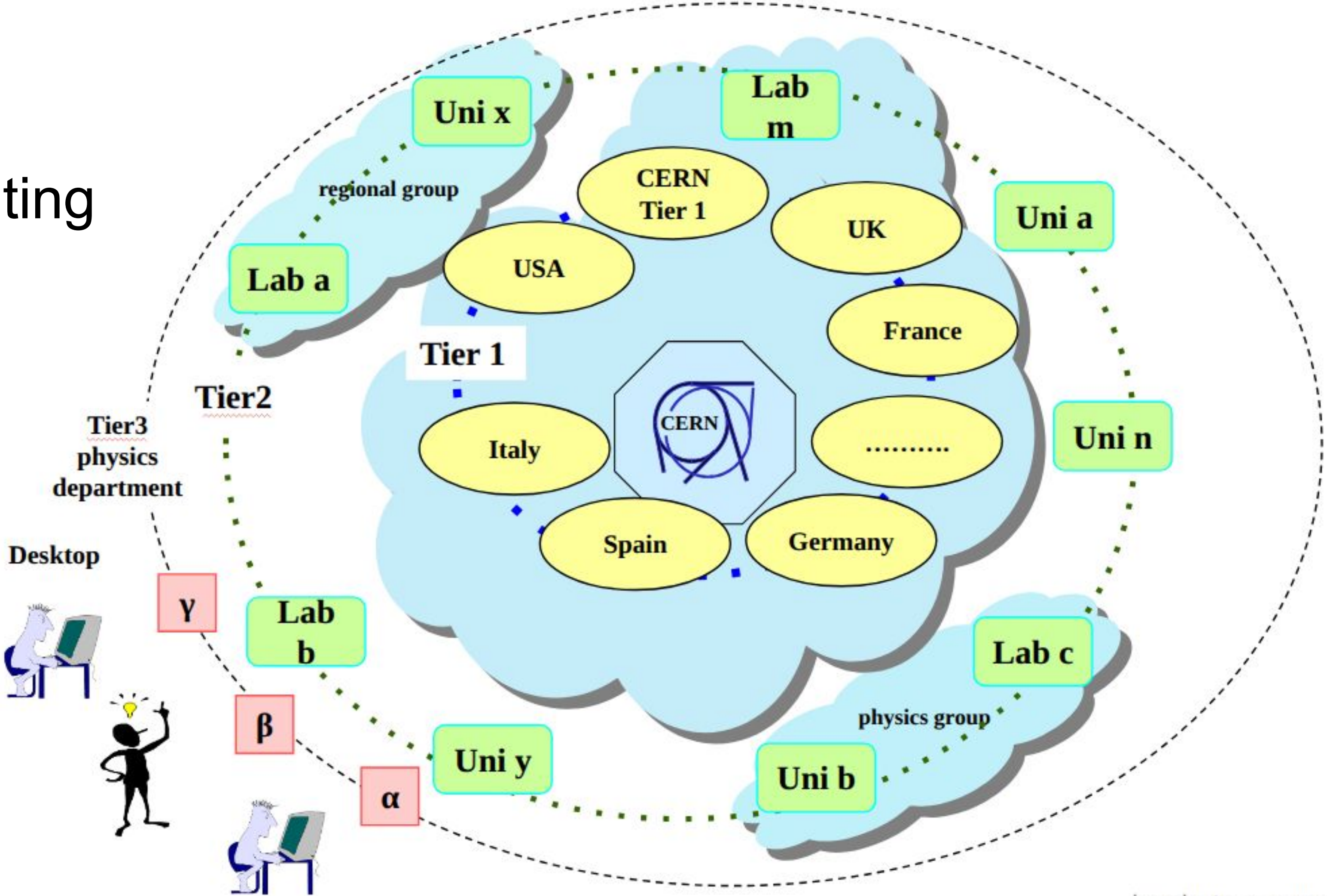
PHASE 2 REPORT

24th March 2000

MONARC Members

M. Aderholz (MPI), K. Amako (KEK), E. Auge (L.A.L/Orsay), G. Bagliesi (Pisa/INFN), L. Barone (Roma1/INFN), G. Battistoni (Milano/INFN), M. Bernardi (CINECA), M. Boschini (CILEA), A. Brunengo (Genova/INFN) J.J. Bunn (Caltech/CERN), J. Butler (FNAL), M. Campanella (Milano/INFN), P. Capiluppi (Bologna/INFN), F, Carminati (CERN), M. D'Amato (Bari/INFN), M. Dameri (Genova/INFN), A. di Mattia (Roma1/INFN), A. Dorokhov (CERN), G. Erbacci (CINECA), U. Gasparini (Padova/INFN), F. Gagliardi (CERN), I. Gaines (FNAL), P. Galvez (Caltech), A. Ghiselli (CNAF/INFN), J. Gordon (RAL), C. Grandi (Bologna/INFN), F. Harris (Oxford), K. Holtman (CERN), V. Karimäki (Helsinki), Y. Karita (KEK), J. Klem (Helsinki), I. Legrand (Caltech/CERN), M. Leltchouk (Columbia), D. Linglin (IN2P3/Lyon Computing Centre), P. Lubrano (Perugia/INFN), L. Luminari (Roma1/INFN), A. Maslennicov (CASPUR), A. Mattasoglio (CILEA), M. Michelotto (Padova/INFN), I. McArthur (Oxford), Y. Morita (KEK), A. Nazarenko (Tufts), H. Newman (Caltech), V. O'Dell (FNAL), S.W. O'Neale (Birmingham/CERN), B. Osculati (Genova/INFN), M. Pepe (Perugia/INFN), L. Perini (Milano/INFN), J. Pinfold (Alberta), R. Pordes (FNAL), F. Prelz (Milano/INFN), A. Putzer (Heidelberg), S. Resconi (Milano/INFN and CILEA), L. Robertson (CERN), S. Rolli (Tufts), T. Sasaki (KEK), H. Sato (KEK), L. Servoli (Perugia/INFN), R.D. Schaffer (Orsay), T. Schalk (BaBar), M. Sgaravatto (Padova/INFN), J. Shiers (CERN), L. Silvestris (Bari/INFN), G.P. Siroli (Bologna/INFN), K. Sliwa (Tufts), T. Smith (CERN), R. Somigliana (Tufts), C. Stanescu (Roma3), H. Stockinger (CERN), D. Ugolotti (Bologna/INFN), E. Valente (INFN), C. Vistoli (CNAF/INFN), I. Willers (CERN), R. Wilkinson (Caltech), D.O. Williams (CERN).

LHC Computing Model



les robertson – august 2001

Developing the middleware

- The idea of the Grid was stimulating many R&D projects on both sides of the Atlantic
  - Developing middleware
  - Operating testbeds
- Some, like the INFN grid and PPDG, had already practical experience in HEP
- But for LHC the development of solid software was needed to implement the MONARC model

The DataGrid Project

Supported by the EU for the development of grid software and deployment of real-world demonstrators was particularly important as it:

- Provided funding within HEP organisations and enabled collaboration with people and organisations outside the community
- Different ideas and highly motivated developers
- Focus for “homogeneity” of functionality and interfaces across several national projects.

Towards a collaboration, commitment and focus

- That was going to be expensive, so the blessing of the LHCC, CERN management, the national research funding agencies, was needed
- In 2001 Manuel Delfino, the CERN computing head, obtained the agreement of the LHCC to form a pan-HEP pilot project to:
    - Coordinate development activities, organise deployment of testbeds
    - Ensure the wide representation of stakeholders and provide a forum for working towards consensus
- Enable institutes and funding agencies to pledge resources and manpower
- And provide oversight by the LHCC referees

LCG-1 Service Opened September 2003

- Middleware package – components from –
  - European DataGrid (EDG)
  - US (Globus, Condor, PPDG, GriPhyN) → the Virtual Data Toolkit
- Agreement reached on principles for registration and security
- Certification and distribution process established and tested - June
- Rutherford Lab (UK) to provide the initial Grid Operations Centre
- FZK (Karlsruhe) to operate the Call Centre
- Pre-release middleware deployed to the initial 13 centres – July

- 1 September release :

Taiwan, Brookhaven, CERN, Bologna, Fermilab, Karlsruhe, Lyon, Budapest, Moscow, Prague, Barcelona, Rutherford UK, Tokyo

El contexto en España

2001: Nica fue pionero en organizar la comunidad española de LHC y conseguir fondos para LHC computing

- **Preparación de la infraestructura local de cálculo para el experimento CMS en el LHC del CERN**. FPA-2001-4074-E, 85.000 €, IP: Nicanor Colino. 2001.
- **Desarrollo de infraestructura datagrid para análisis de datos de LHC**. FPA2002-04208-C07, Oct 2003 - Sep 2005. 1.76 M€. Proyecto coordinado CIEMAT (N. Colino, coordinador), IFCA (J. Marco), USC (J. Saborido), UB (R Graciani), IFIC (J. Salt), UAM (J. del Peso)

2003: Creación del PIC (Puerto de Información Científica)

- Convenio CIEMAT-IFAE-UAB
- Manuel Delfino como director. Manuel había sido IT director 1999-2001
- Impulso de Manuel Aguilar, Marcos Cerrada, Nicanor Colino y otros

En el CIEMAT F.J. R. Calonge & co. estaban desplegando el Tier-2 de CMS

1 de septiembre de 2003

- El 1 de septiembre de 2003 entro yo en escena
- Pablo García se había involucrado en la computación para CMS
- Conseguimos que CMS acepte al PIC como T1 español
- Participamos exitosamente en el DC04 (CMS Data Challenge 2004)

CHEP'04
INTERLAKEN

CHEP04

27 de septiembre de 2004 a 1 de octubre de 2004
Interlaken, Switzerland

J. Hernandez (CIEMAT Madrid)

Autor en Distributed Computing Grid Experiences in CMS DC04

Autor en Role of Tier-0, Tier-1 and Tier-2 Regional Centres in CMS DC04

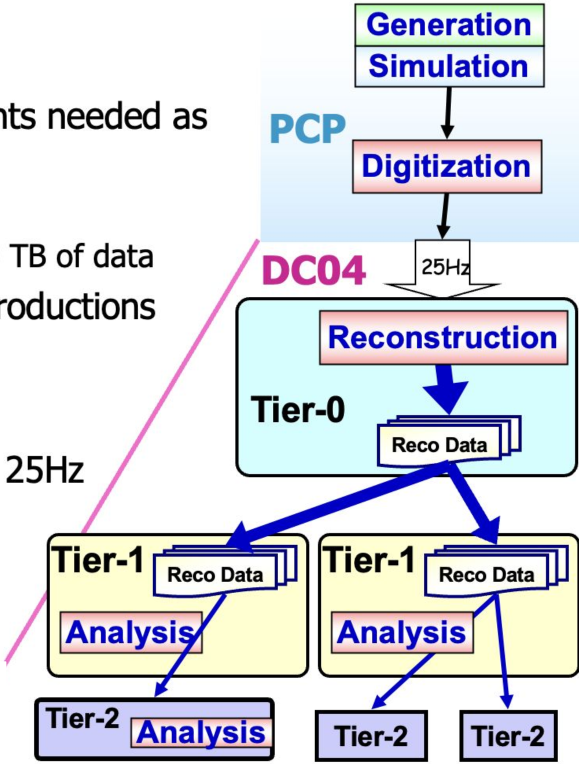
Autor en Software agents in data and workflow management

Autor en Tier-1 and Tier-2 Real-time Analysis experience in CMS Data Challenge 04

CMS Data Challenge 2004

Planned to reach a complexity scale equal to about 25% of that foreseen for LHC initial running

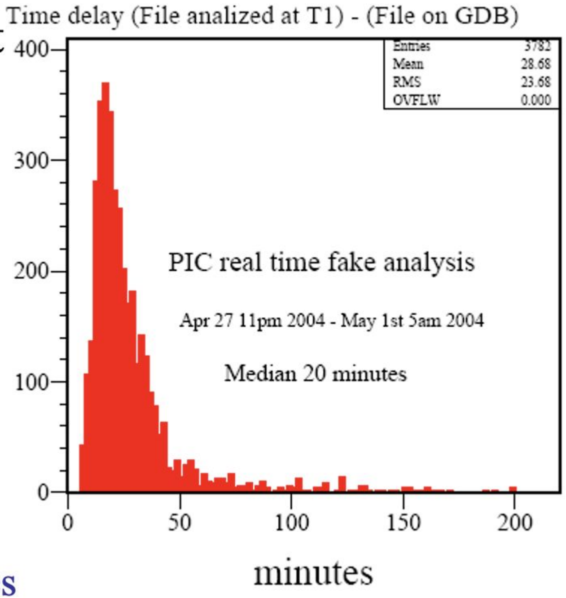
- Pre-Challenge Production in 2003/04 (PCP)
  - Simulation and digitization of ~70 Million events needed as input for the Data Challenge
    - started in July 2003, Digitization still running
    - 750K jobs, 3500 KSI2000 months, 700 Kfiles,80 TB of data
  - Classic and Grid (CMS/LCG-0, LCG-1, Grid3) productions



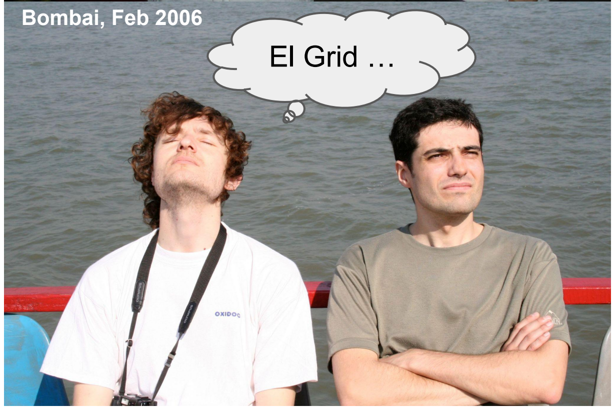
- Data Challenge (DC04)
  - Reconstruction of data for sustained period at 25Hz
  - Data distribution to Tier-1,Tier-2 sites
  - Data analysis at remote sites
  - Demonstrate the feasibility of the full chain

Results: Timeline analysis

- The time delay between the availability at Tier-0 of a file and the analysis at PIC is **20 minutes on average.** The minimum time is around **5 minutes**.



- The main contributions to that time:
  - the time of transfer of the file from the Tier-0 to the Tier-1 ~ 13 minutes on average.
  - the time for replication from the CASTOR SE to the disk SEs ~ less than 1 minute.
  - the time for job preparation ~ about 1.5 minutes
  - the time for job submission, including the local copy of the metadata files ~ 3 minutes
  - the overhead for submitting to the grid the simplest job is ~ 2 minutes



Bombai, Feb 2006

Fermilab, Abril 2005

WLCG Memorandum of Understanding

The European Organization for Nuclear Research (CERN)

and

The Spanish Ministry of Education and Science (MEC)

declare that they agree on this Memorandum of Understanding for collaboration in the deployment and exploitation of the Worldwide LHC Computing Grid.

Done in Geneva

on May 31st, 2007

Done in Madrid

on 6 Julio 2007

in two originals in the English and Spanish languages, each version being equally authentic, it being understood however that in case of contradiction, ambiguity or differences in interpretation, the English text shall prevail.

For CERN

Jos Engelen
Chief Scientific Officer

For MEC

Miguel Angel Quintanilla Fisac
(Real Decreto 567/2006, BOE 108, 06/05/2006)

- Signed by Spain in 2007
- Contribute with 5% of computing resources at the T1 and T2 levels for the ATLAS, CMS and LHCb experiments

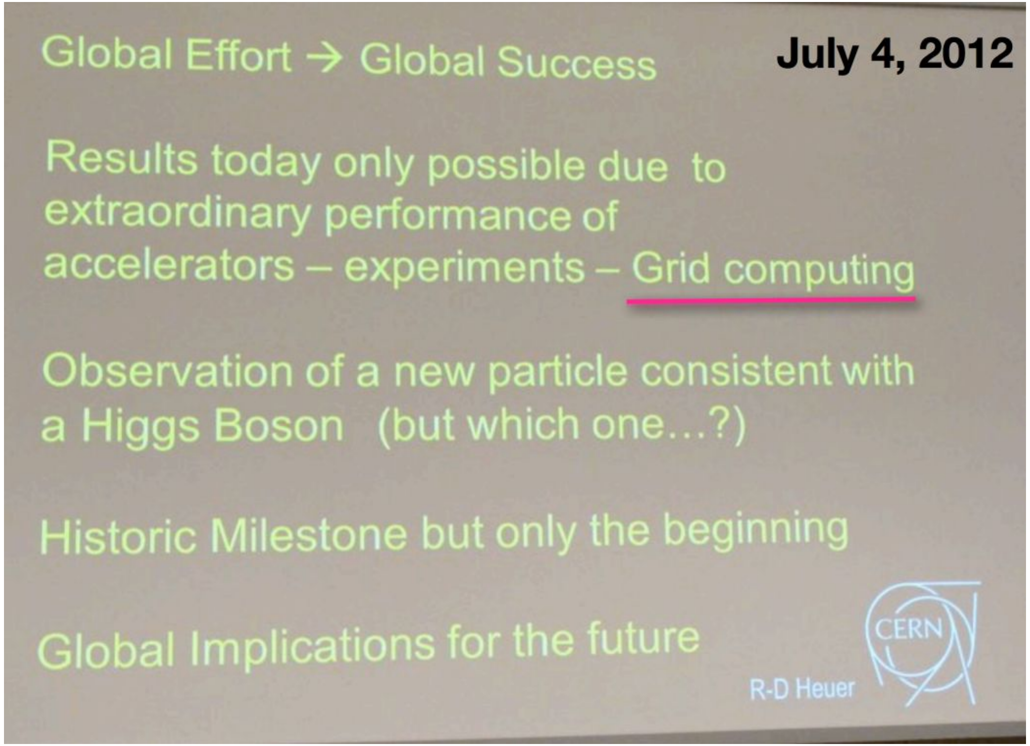
Tier1

Centre	*Experiments served with priority*				*Representative to WLCG Collaboration*	*Funding Agencies*
	ALICE	*ATLAS*	*CMS*	*LHCb*		
Canada, TRIUMF		X			M. Vetterli	CFI
France, CC-IN2P3	X	X	X	X	F. Malek (deputy: F. Hernandez)	CNRS/IN2P3 and CEA/DSM/DAPNIA
Germany, FZK-GridKA	X	X	X	X	K.-P. Mickel	BMBF/FZK
Italy, CNAF	X	X	X	X	M. Mazzucato (deputy: L. Dell'Agnello)	INFN
Netherlands LHC/Tier1	X	X		X	J. Templon	NIKHEF
Nordic Data Grid Facility (NDGF)	X	X	X		L. Fischer	NDGF
Spain, PIC		X	X	X	M. Delfino (deputy: G. Merino)	MEC
Taipei, ASGC		X	X		S. Lin	Academia Sinica
UK, RAL	X	X	X	X	N. Geddes	PPARC
USA, BNL		X			M. Ernst (alt.: R. Popescu)	DOE
USA, FNAL			X		V. White	DOE

Annex 2. WLCG Tier2 Centres and Federations of Centres that together constitute a Tier2 Centre

Institution	*Experiments served with priority*				*Representative to WLCG Collaboration*	*Funding Agencies*
	ALICE	*ATLAS*	*CMS*	*LHCb*		
Spain, ATLAS Federation - IFAE, Barcelona - IFIC, Valencia - UAM, Madrid		X			J. Salt (alt: A. Pacheco Pages, J. del Peso)	MEC
Spain, CMS Federation - CIEMAT, Madrid - IFCA, Santander			X		F. Matorras (alt.: N. Colino)	MEC
Spain, LHCb Federation - UB, Barcelona - USC, Santiago				X	R. Graziani Diaz (alt.: J.J. Saborido Silva)	MEC

Higgs Boson discovery



The Worldwide LHC Computing Grid: the present

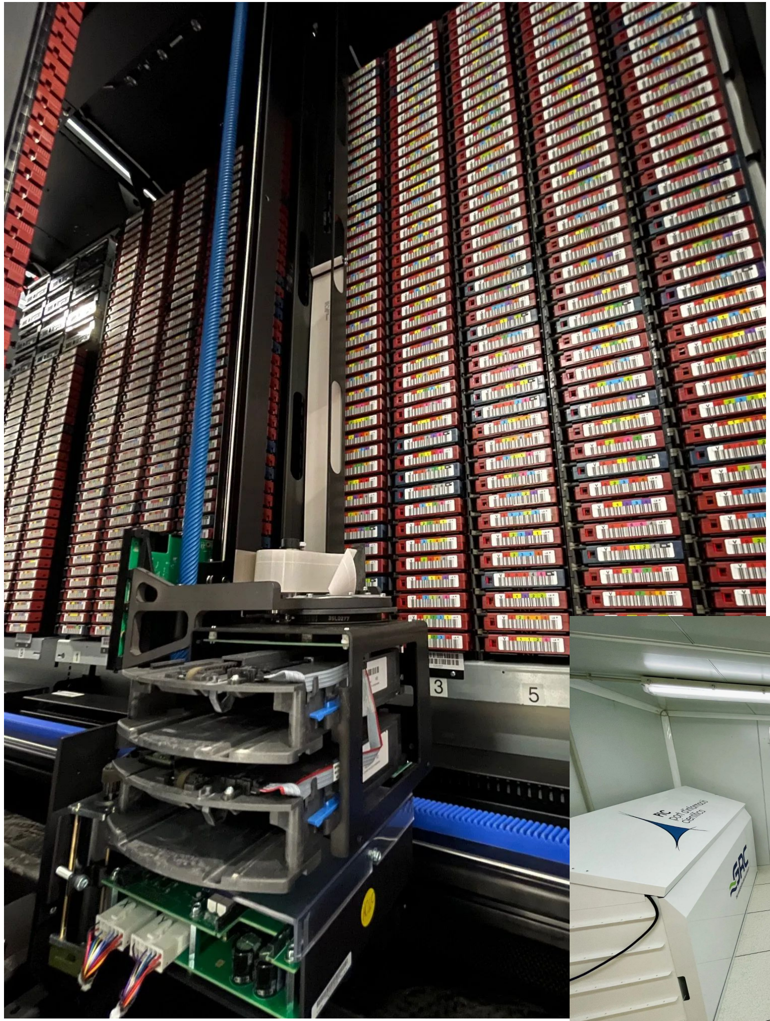
~170 data centres; ~40 countries; shared by the four LHC experiments

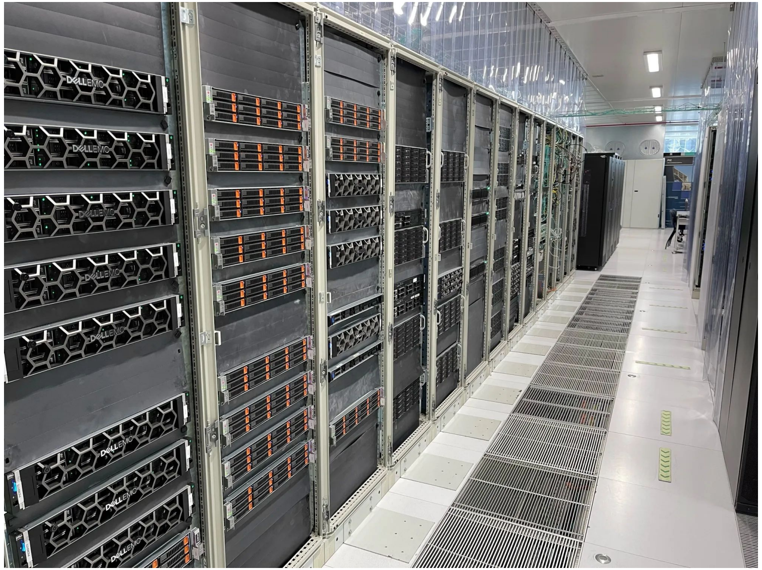
>50 GB/second; >50 million files/week

>2 million tasks/day; >1 million cores;

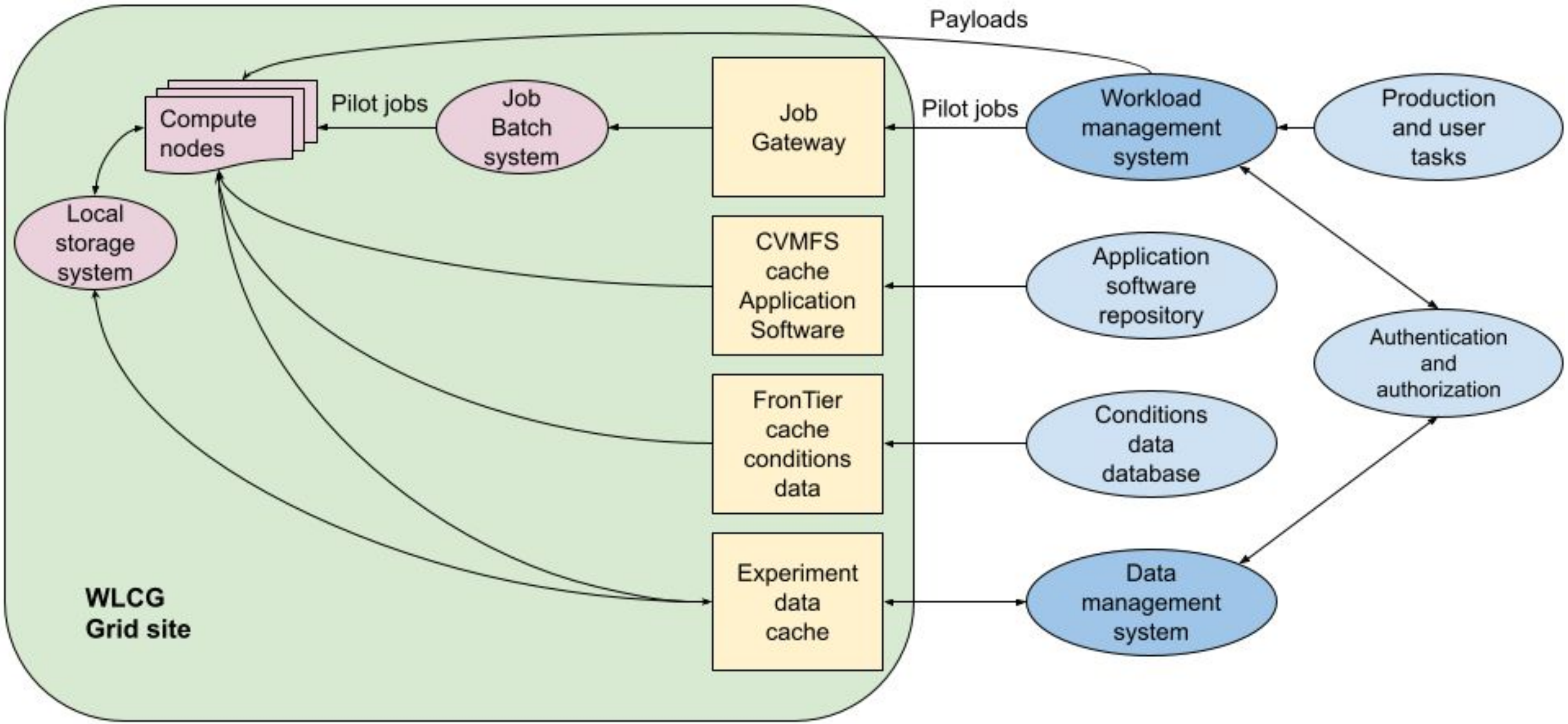
> 2 EB on disk and tape

WLCG
Worldwide LHC Computing Grid





Grid services

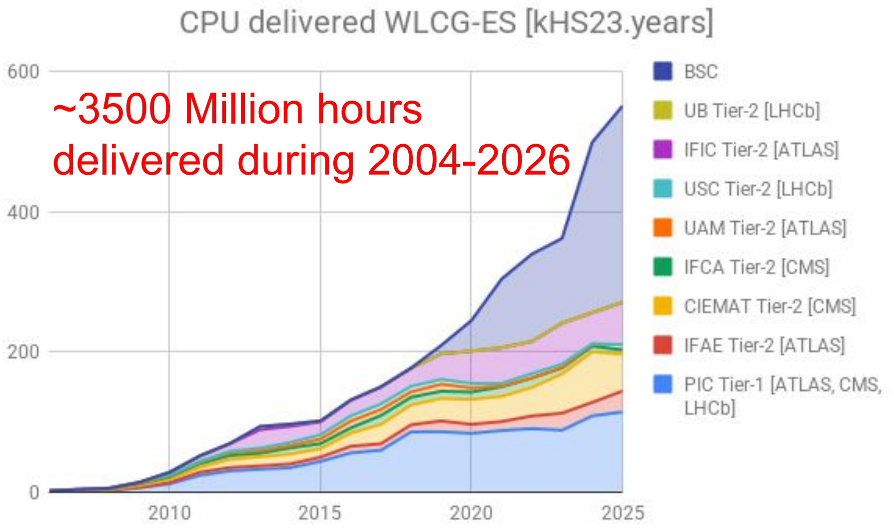


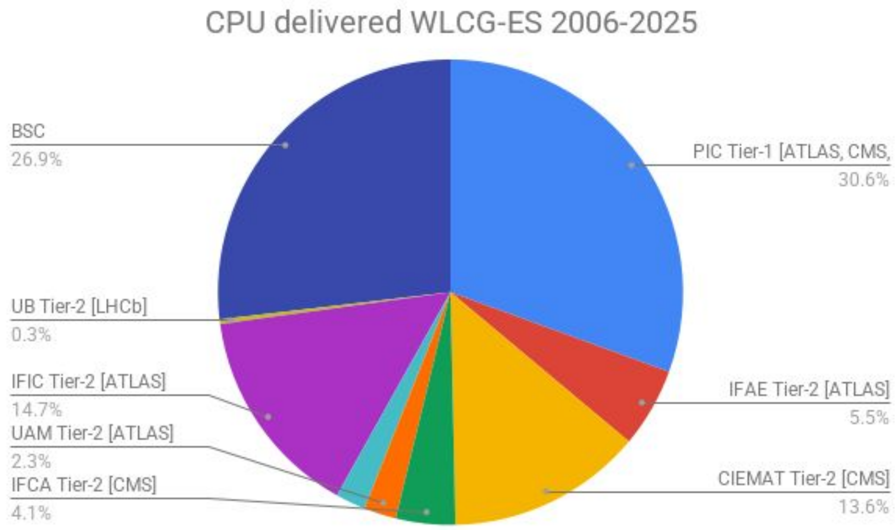
Spain in WLCG

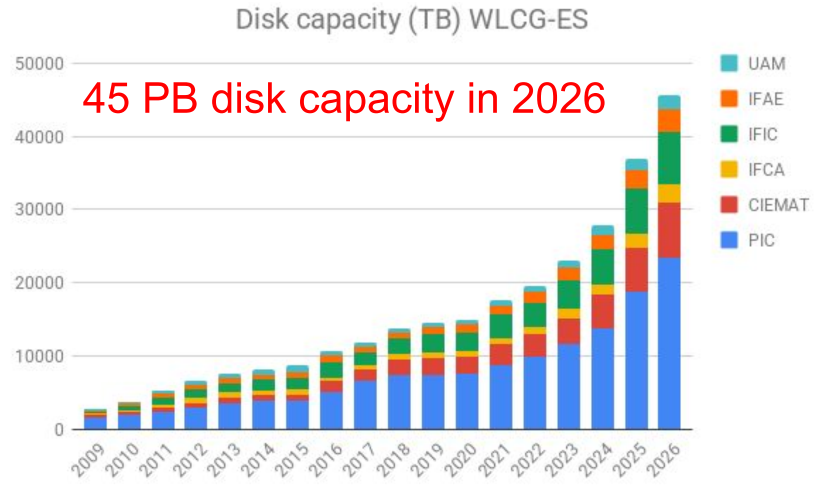
- ~5% resources T1 & T2
- 1 Tier-1 center (PIC, CIEMAT-IFAE)
- 5 Tier-2 centers
  - CMS: CIEMAT-Madrid (75%)
    IFCA Santander (25%)
  - ATLAS: IFIC-Valencia (60%)
    IFAE-Barcelona (25%)
    UAM-Madrid (15%)

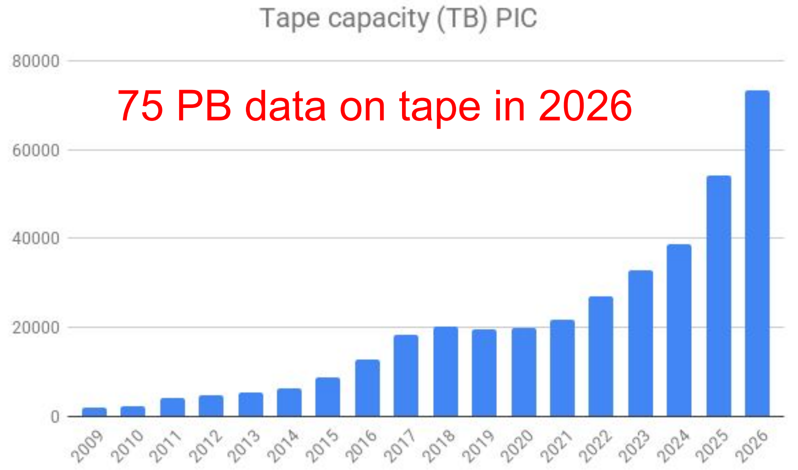
CIEMAT-PIC:
Gonzalo Merino, Pepe Flix, Antonio Pérez-Calero, Pau Tallada, Francesc Torradeflot, Jordi Casals, Jordi Delgado, Jorge Carretero

CIEMAT-Moncloa:
F. Javier R. Calonge, Raúl Fernández, Enrique Ugedo, Mario Uceda, Miguel Cárdenas, J.J. Rodríguez, JMHC


CPU delivered WLCG-ES [kHS23.years]
~3500 Million hours
delivered during 2004-2026
600
400
200
0
2010
2015
2020
2025
BSC
UB Tier-2 [LHCb]
IFIC Tier-2 [ATLAS]
USC Tier-2 [LHCb]
UAM Tier-2 [ATLAS]
IFCA Tier-2 [CMS]
CIEMAT Tier-2 [CMS]
IFAE Tier-2 [ATLAS]
PIC Tier-1 [ATLAS, CMS, LHCb]

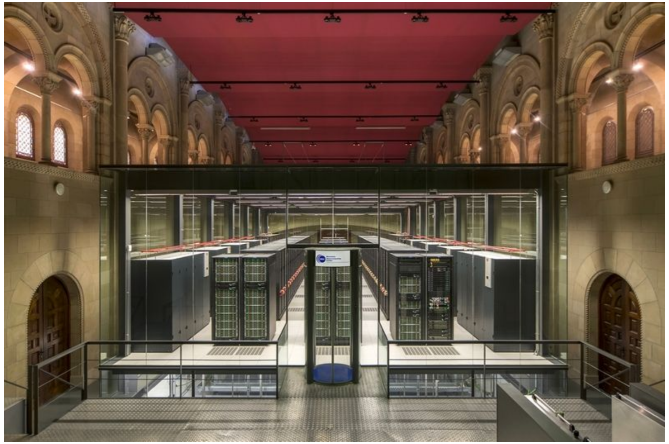

CPU delivered WLCG-ES 2006-2025
BSC
26.9%
UB Tier-2 [LHCb]
0.3%
IFIC Tier-2 [ATLAS]
14.7%
UAM Tier-2 [ATLAS]
2.3%
IFCA Tier-2 [CMS]
4.1%
PIC Tier-1 [ATLAS, CMS,
30.6%
IFAE Tier-2 [ATLAS]
5.5%
CIEMAT Tier-2 [CMS]
13.6%


Disk capacity (TB) WLCG-ES
45 PB disk capacity in 2026
50000
40000
30000
20000
10000
0
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
UAM
IFAE
IFIC
IFCA
CIEMAT
PIC

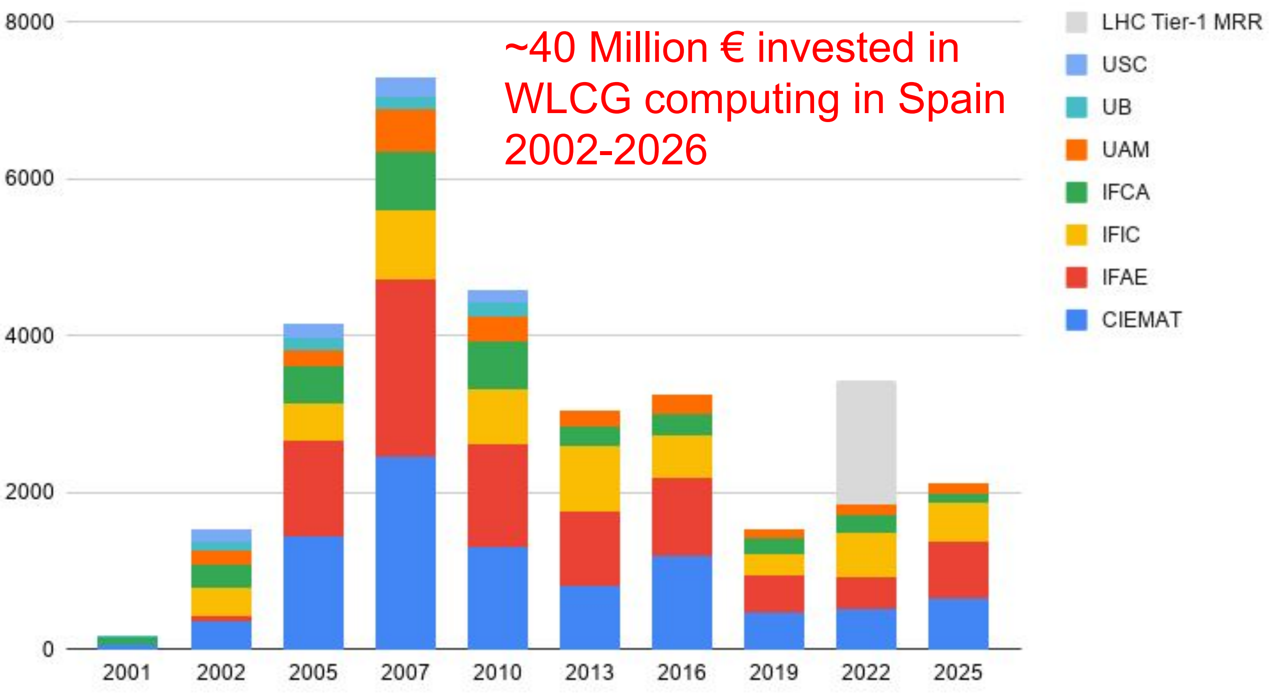

Tape capacity (TB) PIC
75 PB data on tape in 2026
80000
60000
40000
20000
0
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026

Barcelona supercomputing center & LHC computing

- The BSC is the largest supercomputing center in Spain
  - MareNostrum4 (150k CPU cores); MareNostrum5 (750k CPU cores, from 2024)
- BSC - WCLG-ES agreement in 2019
  - LHC computing designated as a BSC “strategic project”
  - Access to dedicated resources (up to 7% of MareNostrum4)
  - Providing CPU for LHC simulation (~50% of WLCG-ES CPU, currently ~100M hours/year)



Funding WLCG-ES



CIEMAT contributions to CMS computing







- Computing resources
  - CMS Tier-1 and Tier-2 sites (~5% of total), CMS Analysis Facility
- Development and integration of computing distributed services
  - Data management (PhEDEx), MC production management (MCRunJob, ProdAgent, WMAgent), User analysis management (CRAB[1-3]), HPC integration
- Monitoring and operation of computing services
  - Site monitoring, MC production, User analysis
- Management
  - Level 2 coordination (Integration, distributed analysis, computing operations, resource office, submission infrastructure)
  - Chair of CMS Computing Resource Board
- WLCG coordination and oversight
  - Grid Deployment Board, HEPix forum, WLCG management and overview boards, WLCG LHCC, WLCG CRSG

Applications of Grid Computing to other fields

The WLCG infrastructure has driven major breakthroughs in distributed high-throughput computing, data sharing middleware, and high-speed networking that now benefit a wide range of scientific, medical, and commercial fields

- Astrophysics, Cosmology, Gravitational Waves, Neutrinos, …

Cloud Computing Precursor: Grid computing paved the architectural and security pathways (such as virtual organizations and digital certificates) that shaped modern commercial cloud computing

LHC Computing evolution and the future

- WLCG computing model has evolved over the history
  - Less hierarchical architecture
  - More intensive use of data streaming over the WAN
  - Incorporation of heterogeneous resources: Clouds, HPCs, accelerator cards

- The HL-LHC poses a new challenge (2031+)
  - Large increase in data rates and complexity
  - Intense R&D required to keep resource needs under control
    - Storage consolidation
    - New analysis models and facilities
    - Strong movement towards GPU-based research computing infrastructure,driven by AI demands
    - Artificial intelligence technologies should play a major role