Grid Computing for LHC

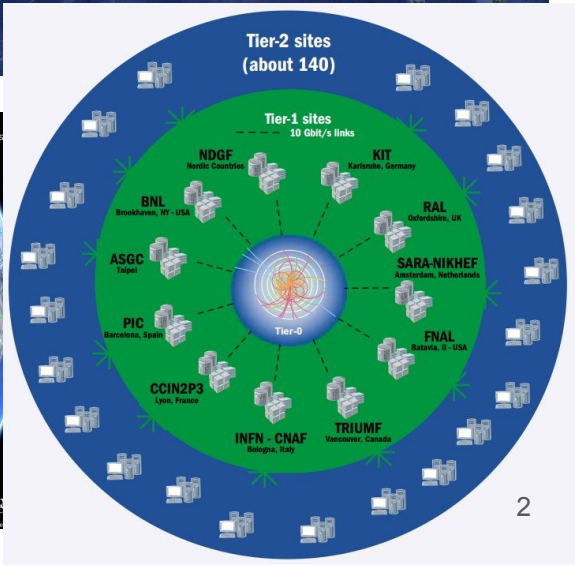
José M. Hernández
CIEMAT, Madrid

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
cfp CIEMAT física de partículas

Worldwide LHC Computing Grid

Distributed high-throughput computing infrastructure to store, process & analyze data produced by LHC experiments

- 167 sites, 42 countries, 63 MoU's
- ~1 million CPU cores
- ~750 PB disk storage
- ~1250 PB tape storage
- Optical private network (LHCOPN) and overlay over NRENs (LHCONE) with 10/100 Gbps links
- ~Tbps LAN bandwidth between compute and storage nodes at sites



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 July 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
- Countries pledge resources annually according to the experiment needs (C-RRB)

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. Fisher	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

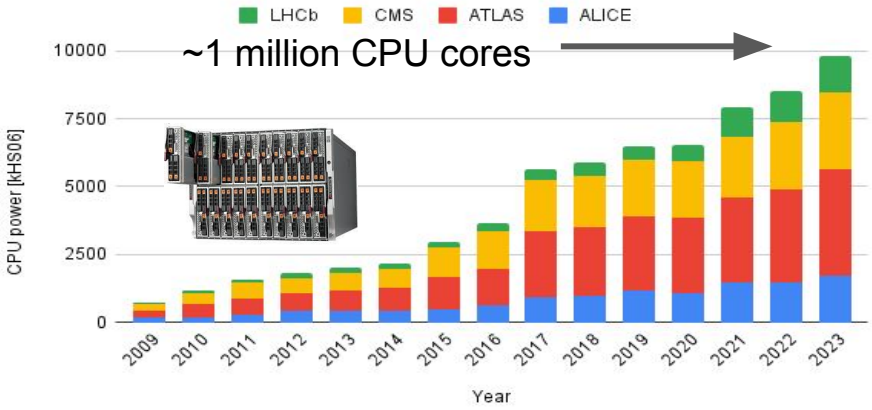
WLCG Computing resource evolution

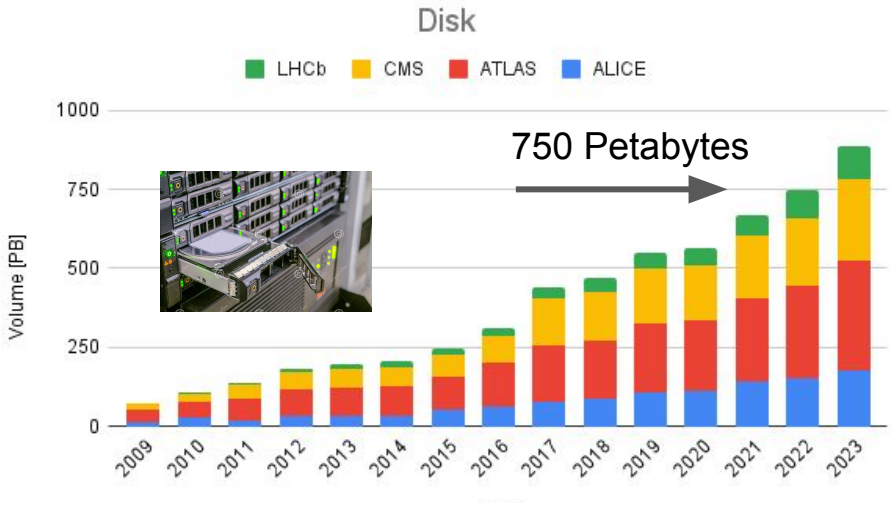
Currently:

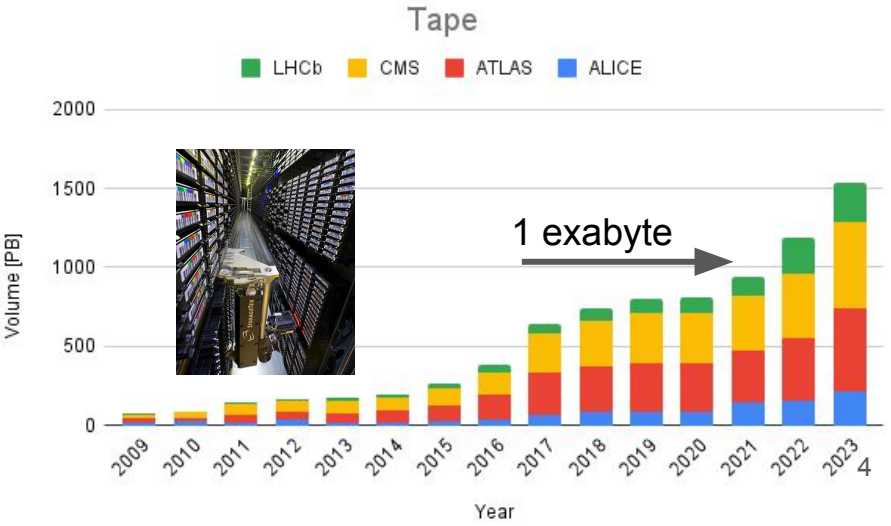
~1M CPU cores, ~2 exabyte storage

~20% annual growth with ~flat funding

(hardware getting cheaper)







Major features and capabilities of HEP computing infrastructure

- Networks
  - International and national, private and public
- Data management
  - Key to success, data transfers, storage systems, data management tools and data organization
- Compute
  - Provision of resources and workload scheduling, execution and monitoring
- Authentication and authorization
  - The mechanism of federation, single sign on, etc
- Operations support
  - Security, incidence response, problem tracking, daily operations, upgrade campaigns
- Diverse experiment-specific services and tools, applications

Distributed data-intensive high throughput computing (HTC)

Precursor of Big Data processing and Cloud computing

Spain in WLCG

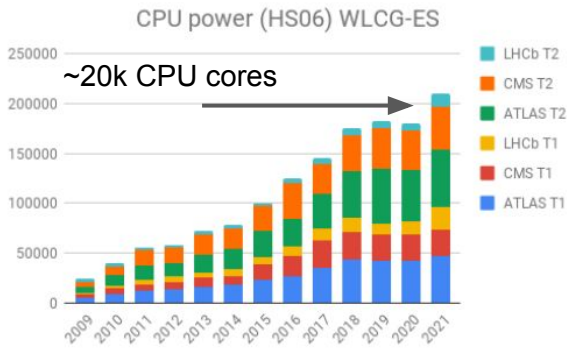
Spanish contribution:

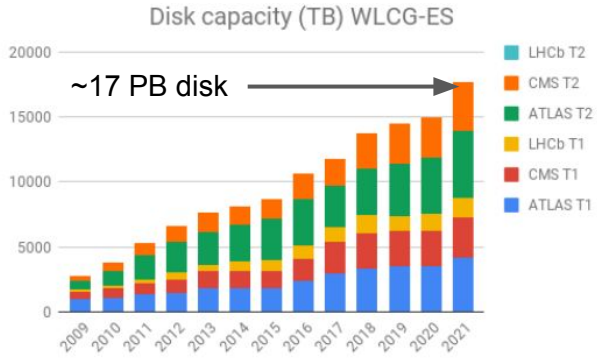
- ~5% resources T1 & T2 (MoU)
- 1 Tier-1 center (PIC, CIEMAT-IFAE)
- 6 Tier-2 centers
  - CMS: CIEMAT-Madrid (75%), IFCA Santander (25%)
  - ATLAS: IFIC-Valencia (60%) IFAE-Barcelona (25%), UAM-Madrid (15%)
  - LHCb: USC-Santiago, UB-Barcelona (decommissioned)

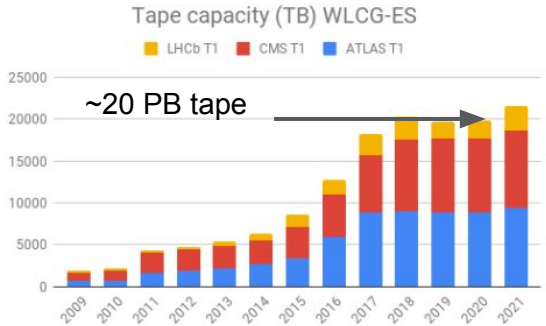
WLCG-ES: a success story

- Almost two decades contributing to LHC distributed computing infrastructure (Worldwide LHC Computing Grid, WLCG) and R&D at the highest level
  - ~5% of WLCG resources (20k CPU-cores, 15 PB disk, 20 PB tape), ~1500M CPU hours delivered since 2004
  - Providing 1 of the 13 Tier-1 sites worldwide (PIC)
  - Federated Tier-2 sites for ATLAS (IFIC, IFAE, UAM), CMS (CIEMAT, IFCA), LHCb (USC, UB)
  - Among the most reliable sites in WLCG
- A large effort from HEP community and institutions
  - ~24 M€ funding (direct costs) from HEP national program since 2001
  - Funding from institutions of the same order
    - Funding personnel, electricity, infrastructure
- Large community of experts in distributed high throughput computing
  - Contributions to LHC computing, development, integration, operations, management
  - Leverage expertise and infrastructure to support other projects in HEP/astro/cosmo (CTA, MAGIC, DUNE, DarkSide, PAU, Euclid, Virgo, etc)
  - **We have generated a big strategic asset for our community!**

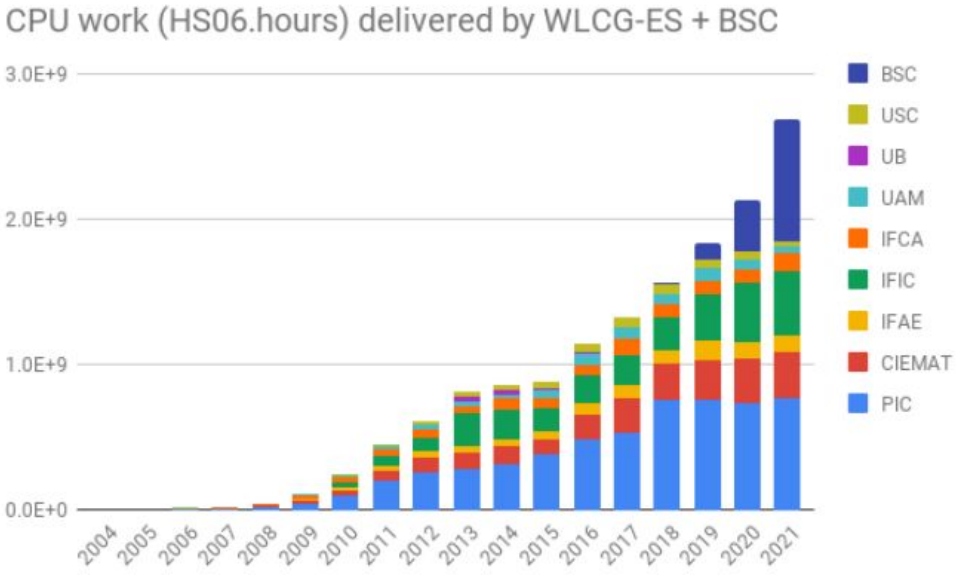
Resources provided by Spain to WLCG

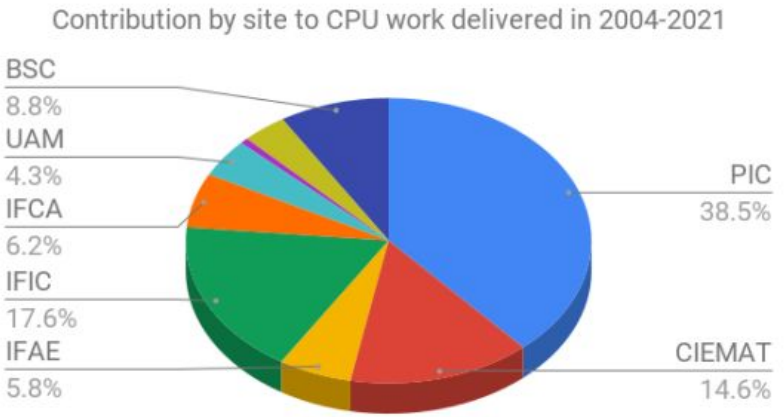






CPU delivered by Spain to WLCG



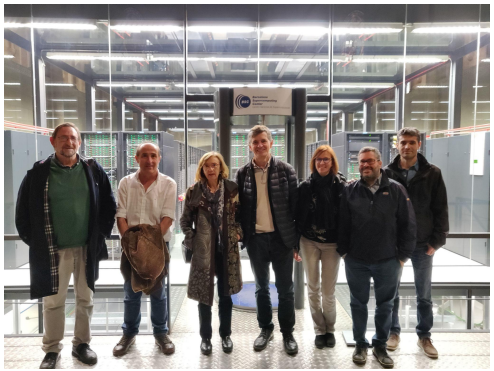


~1500 Million hours delivered during 2004-2021

(15x10^9 HS06.hours; average CPU core power ~10 HS06)

Barcelona supercomputing center & LHC computing

- The BSC is the largest supercomputing center in Spain
  - MareNostrum4 (150k CPU cores); MareNostrum5 10x larger (expected from 2023)
- BSC - WCLG-ES agreement
  - LHC computing designated as a BSC “strategic project”
  - Access to dedicated resources (up to 7% of MareNostrum4)
  - Providing CPU for LHC simulation (~50M hours/year, ~50% of WLCG-ES CPU)

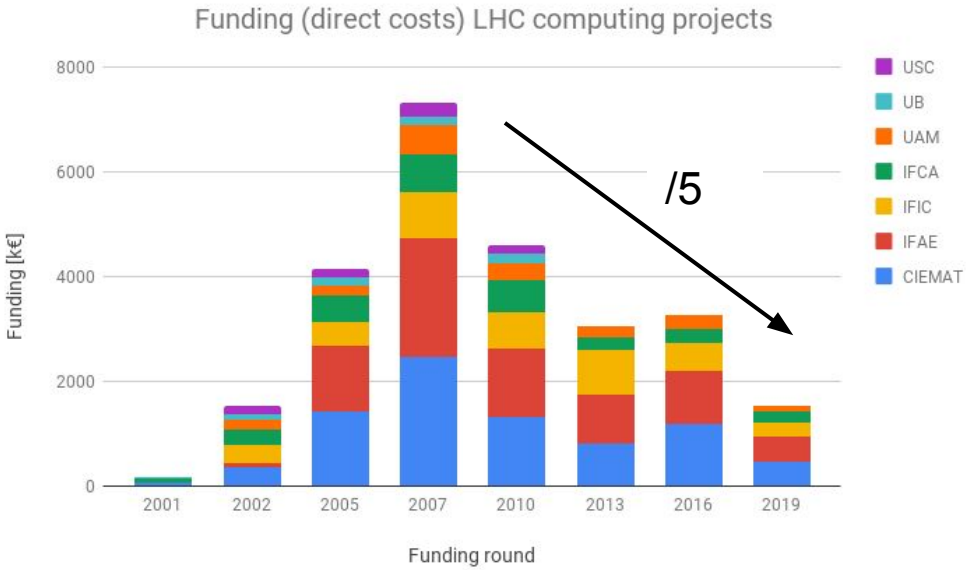


Exploiting supercomputers for LHC

- Data intensive computing with HPC facilities is a **challenge**
  - Limited/no network connectivity in compute nodes
  - Limited storage for caching input/output event data files
  - In practice only run cpu-bound workflows (MC simulation) with little I/O
- Our applications are not really suited for HPC
  - No large parallelization (no use of fast node interconnects)
  - No substantial use of accelerators (GPU)
- Substantial **integration** work to make HPC work for HTC
  - No one-fit-all solution: each facility is different
  - Little effort available in the LHC experiments; in charge of the local communities
  - Experiments do not accept capacity from HPCs as pledged resources unless they can be used transparently as any other WLCG site
- Not suitable resource **allocation** model
  - We would need a guaranteed share of resources rather than apply for allocations

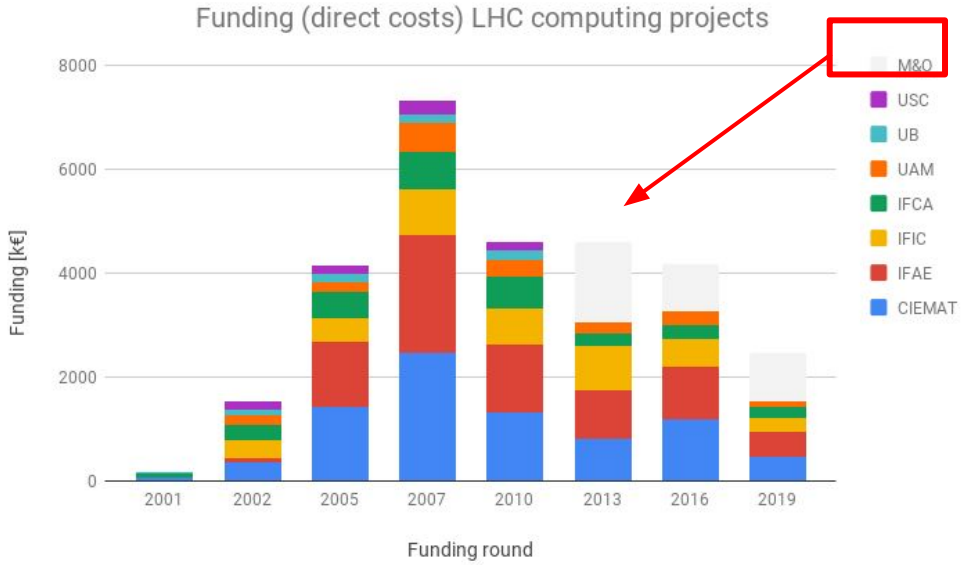
WLCG-ES sustainability challenge

- Decreasing funding from HEP national programme
  - From ~2.5M€/year in 2007 to ~0.5 M€/year in 2019
  - Contribution reduced from 5% to 4% to 3%
  - Aging equipment (~50% > 5 years)
- Big effort to complement funding and resources
  - From institutions
  - From national/regional scientific infrastructure calls
- Required funding ~1.5M€/year



WLCG-ES sustainability challenge

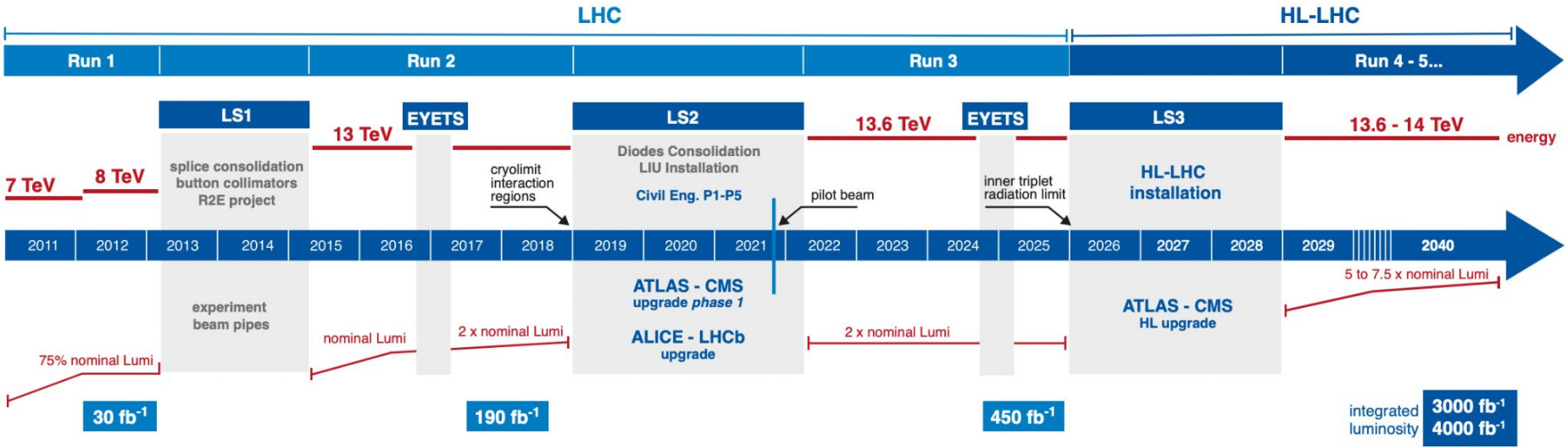
- Decreasing funding from HEP national programme
  - From ~2.5M€/year in 2007 to ~0.5 M€/year in 2019
  - Contribution reduced from 5% to 4% to 3%
  - Aging equipment (~50% > 5 years)
- Big effort to complement funding and resources
  - From institutions
  - From national/regional scientific infrastructure calls
- Required funding ~1.5M€/year



Funds from ministry for LHC upgrades in Spain

- Specific funding from ministry in 2021 (~6 M€) and 2022 (~6 M€)
  - Nuclear physics experiments at CERN (ISOLDE, n_TOF) managed to get ~50% of this funding!
- No funding for LHC computing upgrades!
  - The upgrade of the computing infrastructure towards HL-LHC is progressive, unlike for the detectors
- We estimate that about half of the CPU and disk servers of the Spanish WLCG infrastructure need to be renewed.
  - The estimated total cost to undertake this renovation would be ~1M €
- Is it still possible to get some money from the 2022 funding?

LHC / HL-LHC plan

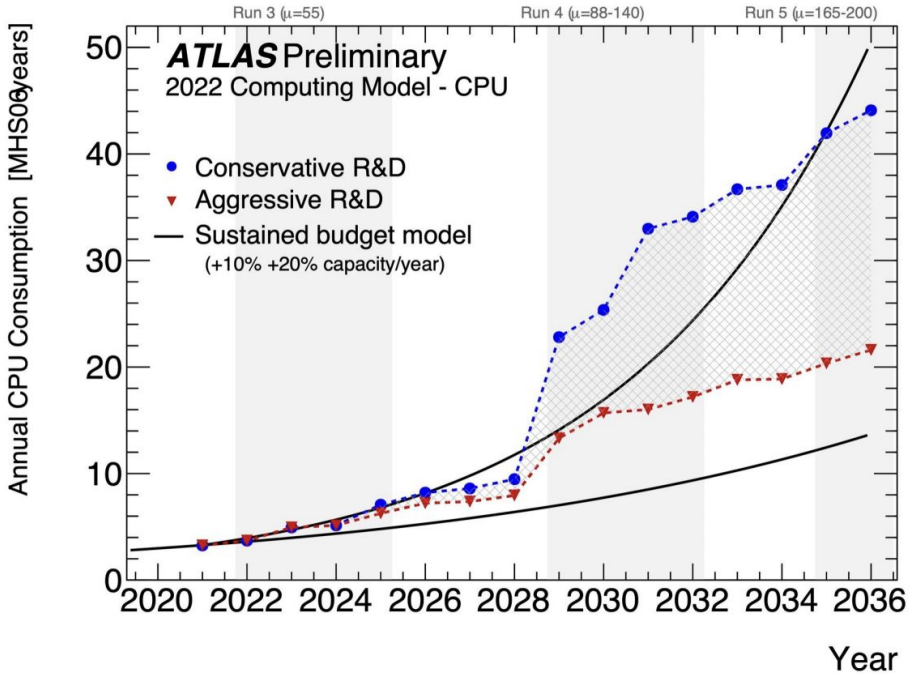


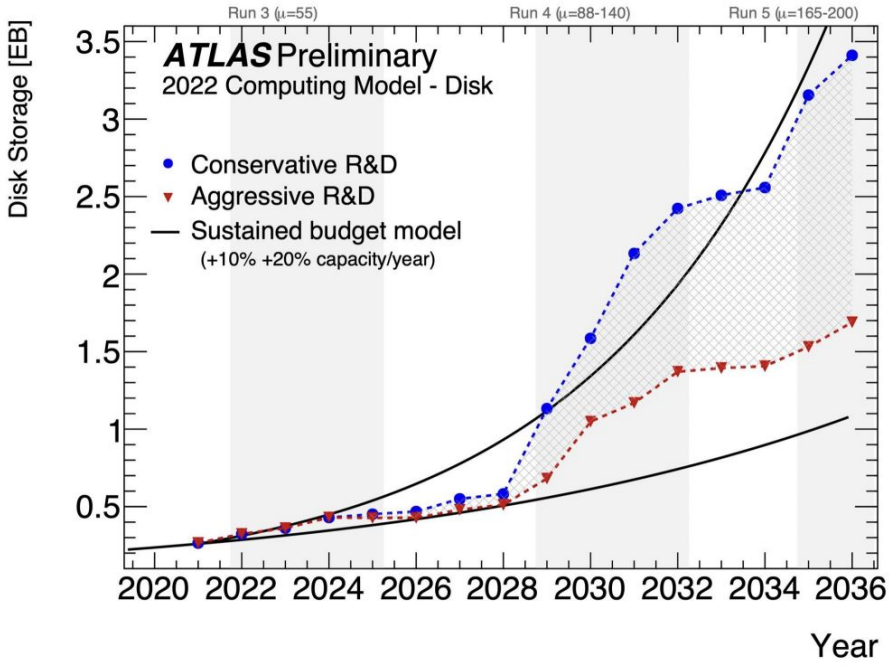
- Run 3 (2022-2025): **~2x more data**. **Evolutionary** changes in computing models
- Run 4 (HL/LHC, 2029+): **~20-30x more data**. **Revolutionary** changes required

HL-LHC computing challenge

- Intense R&D to drastically reduce computing resource needs for HL-LHC
  - Reduce data replication levels
    - Requires new infrastructure; data lakes and content delivery network with caching layer
  - Reduce data formats for analysis
    - ~1 kB/event. Requires good understanding of detectors and ~continuous reprocessing
  - Software improvements
    - Speed up and parallelize algorithms
  - Use new hardware resources
    - GPUs and supercomputing resources
- Attainable with constant funding
  - ~10-20% annual increase from hardware getting cheaper
  - But sustaining the current level of funding is needed!!!

Current computing resource needs projections for HL-LHC (ATLAS)





In view of the request for funding to the FPN programme this year, the WLCG-ES community has elaborated a document with the status and perspectives of the LHC computing in Spain

WLCG-ES status and perspectives report

May 2022
WLCG-ES community

Executive summary

The Worldwide LHC Computing Grid (WLCG) infrastructure has been instrumental in the rich scientific harvest attained by the LHC. Massive offline computing has become an integral part of the HEP experiments, deserving effort and funding at the level of the detectors themselves. Spain has significantly contributed to the WLCG effort since its inception. We have provided about 5% of the computing resources with the highest levels of availability and reliability. Our participation in management, development and testing has also been very well recognized.

After years of limited funding, resulting in an overstretched capacity and an aging infrastructure, our share in WLCG resources had to be reduced in 2016 from 5% to 4% to match available funding. We estimate that about half of the CPU and disk servers of the Spanish WLCG infrastructure need to be renewed. The estimated total cost to undertake this renovation would be 1.5M €. As funding permits, we should try to return to the previous level of contribution to avoid our position and influence being endangered.

LHC computing faces new challenges in the upcoming Run 3 (2022-2025) and HL-LHC phase (2029+). The former will require evolutionary changes to deal with the expected increase of data of factor two, while the latter will probably need more revolutionary transformations in the computing models to cope with a ten-fold increase in data rate and complexity.

The forecasts of the ATLAS and CMS experiments of computing resource needs for Run 3 point to yearly increases of resources of 10-20%, consistent with what can be achieved with flat budgets, thanks to the price reduction of hardware from technology evolution. However, LHCb, which already in Run 3 needs to face an increase in data rates of about a factor ten, is anticipating a necessity of yearly growths of computing resources above 50%.

Based on the LHC experiments' requests and reasonable extrapolations, the cost of the acquisition of hardware and hiring of personnel in Spain for LHC computing until 2025 has been estimated. A total investment of 4.5 M€ is needed for the period 2023-2025 to maintain a 5% contribution from Spain to WLCG. Funding will be requested this year by submitting projects to the national R&D FPN programme.

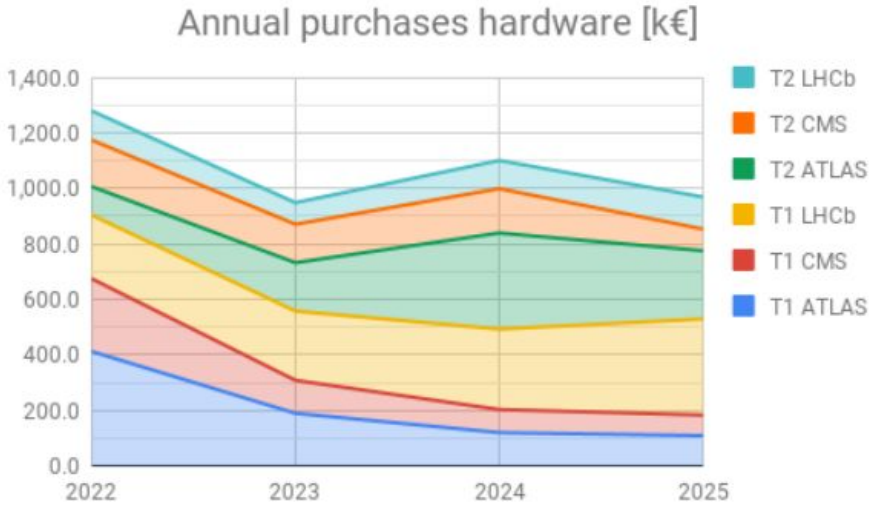
The HL-LHC phase will pose difficult computing challenges due to the large increase in data rate (x10) and complexity (200 simultaneous collisions per event). A vigorous R&D programme has been launched to optimize the experimental software, reduce data formats and exploit new technologies. The Spanish WLCG community is ideally placed to substantially contribute to this endeavor.

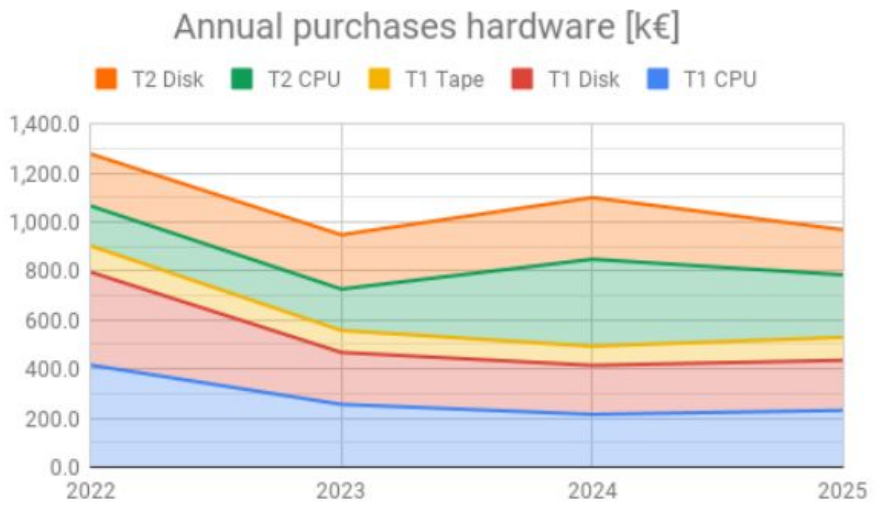
Computing resources for Run 3 in Spain

- Current research projects funded by FPN programme end May 30th, 2023
  - Intend to request extensions until end of 2023
  - 1 coordinated project CIEMAT-IFAE for Tier-1 (ATLAS, CMS, LHCb) and Tier-2 (CMS, ATLAS)
  - 1 coordinated project IFIC-UAM (75% Tier-2 ATLAS)
  - 1 single project IFCA (25% Tier-2 CMS)
  - No projects for Tier-2 LHCb
- Submit new projects in 2022 funding call
  - Funding resources for rest of LHC Run 3 (2023-2025)
  - Coordination between the projects (modeling of requests, goals, etc)
- Modeling of resource cost
  - Use resource needs of the experiments (CMS, ATLAS, 10-20 increase/year, LHCB ~50%/year
  - Recover a contribution of 5% at the Tier-1 and Tier-2 levels for ATLAS, CMS, LHCb
  - Assume a contribution of 50% of the CPU from BSC
  - 5-year renewal policy for old equipment
  - Yearly hardware price decrease of 15%

Computing resources for Run 3 in Spain

Estimated expenses in computing resources for Run 3 in Spain





- ~ 1M€/year yearly investment in equipment required for Run 3
- In addition, 7 FTE (1 per site) for development, integration, operation: ~350 k€/year
- Overall: ~4 M€ all projects, 3 years