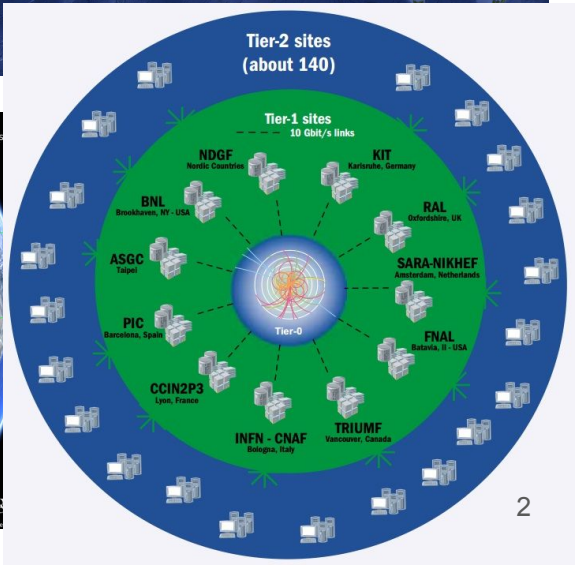
Grid Computing for LHC

- Deployment and operation of the distributed computing infrastructure for LHC
- R&D on computing services to manage data, processing, simulation and analysis

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
- ~1.4 EB disk storage
- ~2.9 EB tape storage
- Optical private network (LHCOPN) and overlay over NRENs (LHCONE) with 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

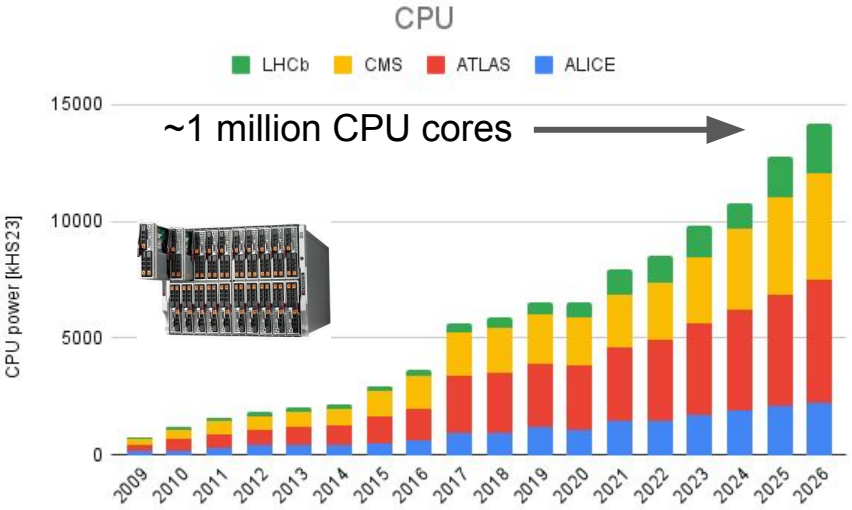
WLCG Computing resource evolution

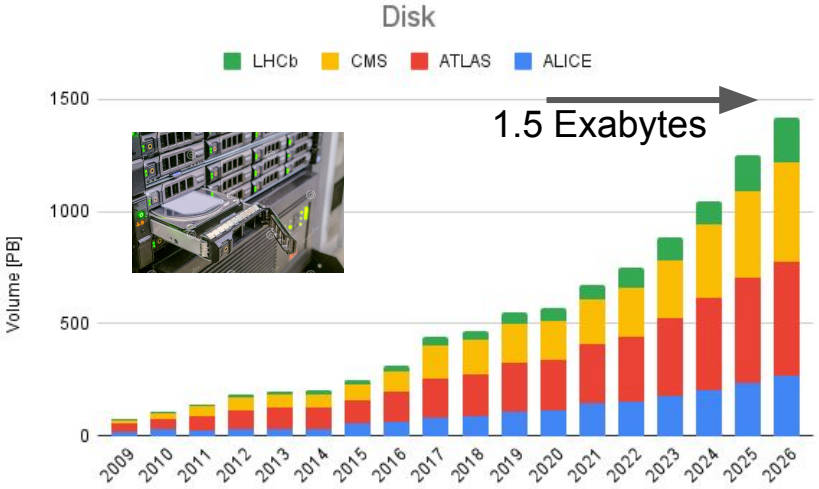
Currently:

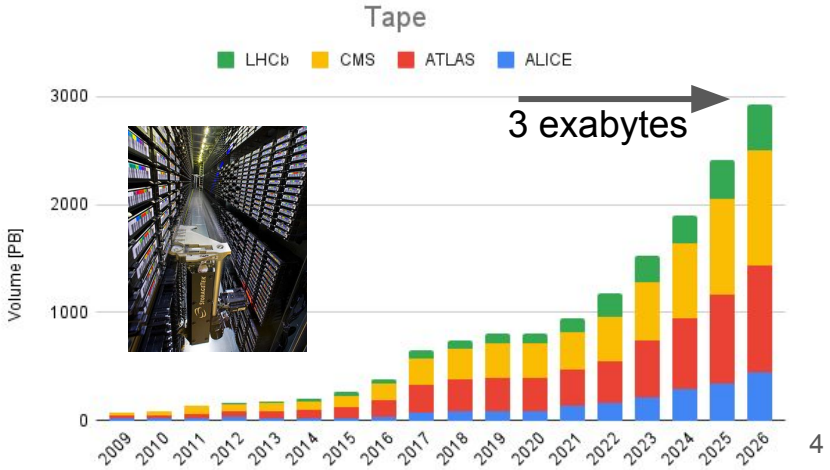
>1M CPU cores, >4 exabyte storage

~15% annual growth with ~flat funding

(hardware getting cheaper)







Major features and capabilities of HEP computing infrastructure

- High-speed Networks
  - International and national, private and public
- Massive data management
  - Key to success, data transfers, storage systems, data management tools and data organization
- Large-scale task execution
  - 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

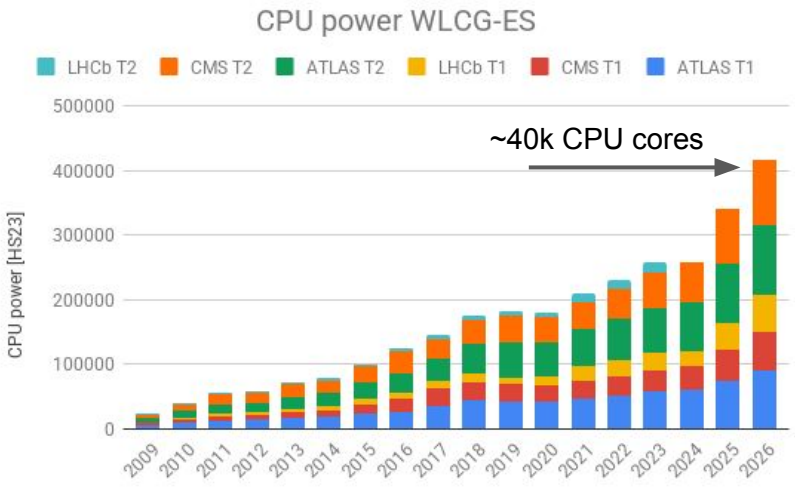
Spanish contribution:

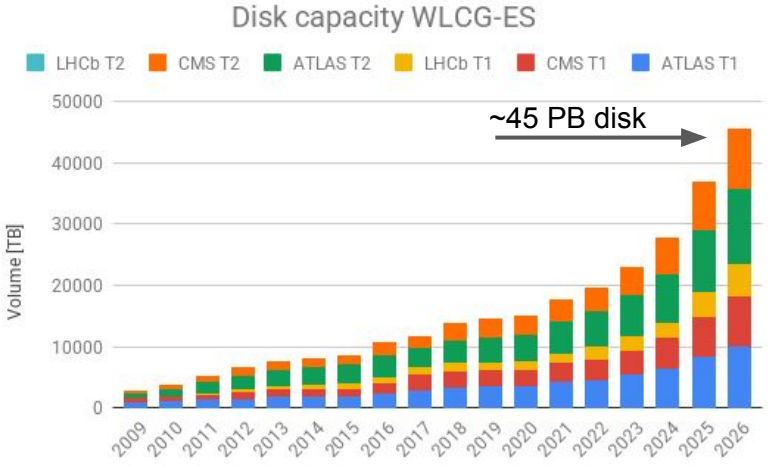
- ~5% resources T1 & T2 (MoU)
- 1 Tier-1 center (PIC, CIEMAT-IFAE)
- 5 Tier-2 centers
  - CMS: CIEMAT-Madrid (75%)
    IFCA Santander (25%)
  - ATLAS: IFIC-Valencia (60%)
    IFAE-Barcelona (25%)
    UAM-Madrid (15%)

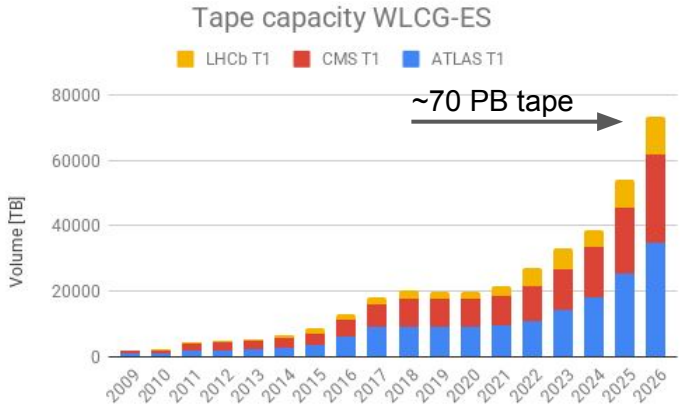
WLCG-ES: a success story

- 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 (40k CPU-cores, 45 PB disk, 70 PB tape), ~2900M CPU hours delivered since 2004
  - Providing 1 of the 13 Tier-1 sites worldwide (PIC) (CMS 7 T1s, LHCb 8 T1s, ATLAS 9 T1s)
  - 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
  - ~26 M€ funding (direct costs) from HEP national program since 2001
  - Co-funding from institutions of the same order
    - Funding personnel, electricity, infrastructure, resources
- 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

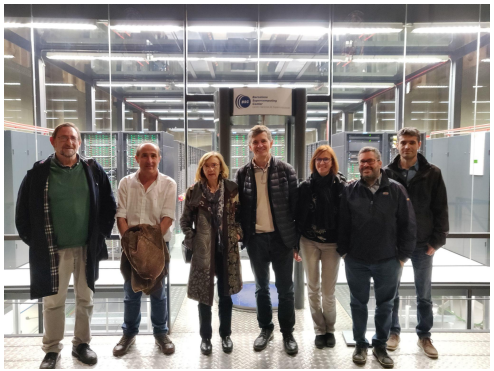






Barcelona supercomputing center & LHC computing

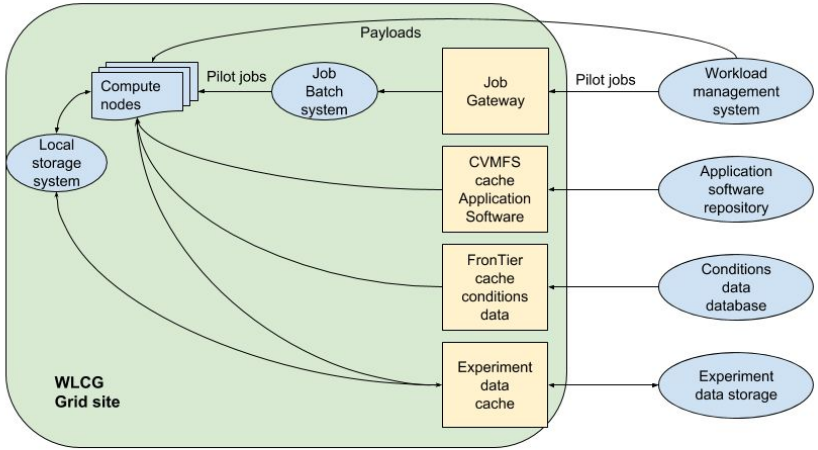
- The BSC is the largest supercomputing center in Spain
  - MareNostrum4 (150k CPU cores); MareNostrum5 (750k CPU cores, from 2024)
- BSC - WCLG-ES agreement in 2019
  - LHC computing designated as a BSC “strategic project”
  - Access to dedicated resources (up to 7% of MareNostrum4)
  - Providing CPU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

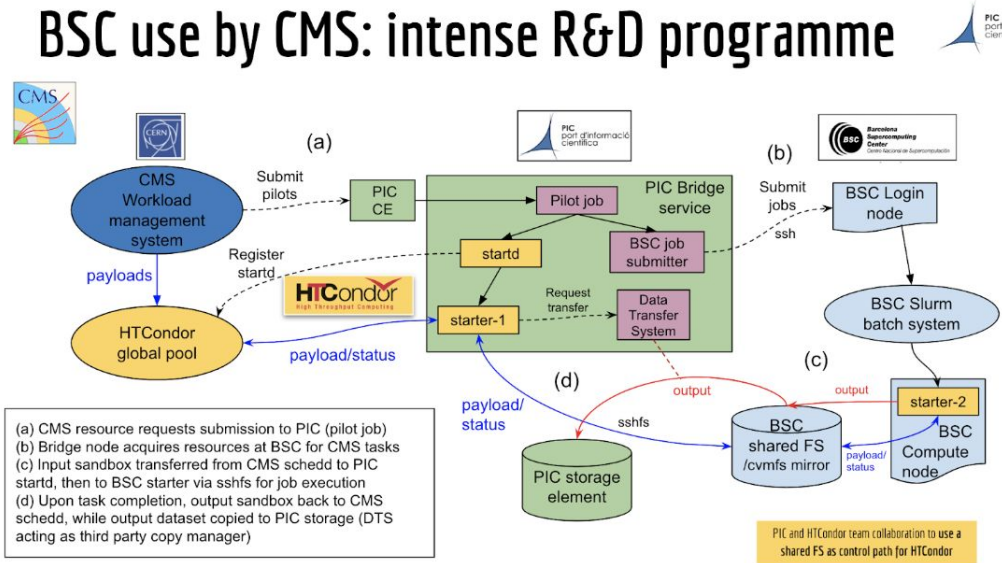


Exploiting supercomputers for LHC: HTC vs HPC

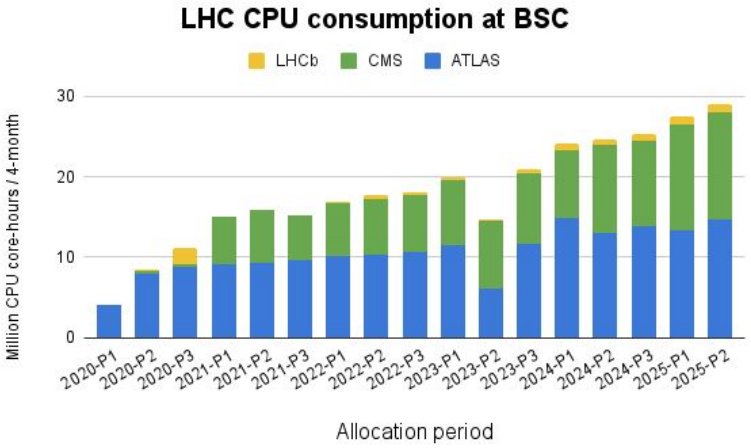
- Data intensive computing with HPC facilities is a **challenge**
  - Limited/no network connectivity in compute nodes
  - Limited storage for caching input/output event data files
  - In practice only run cpu-bound workflows (MC simulation) with little I/O
- Our applications are not really suited for HPC
  - No need for large parallelization (no use of fast node interconnects)
  - No substantial use of accelerators (GPU) yet
- 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
  - Experiments' planning requires a guaranteed share of resources rather than apply for allocations

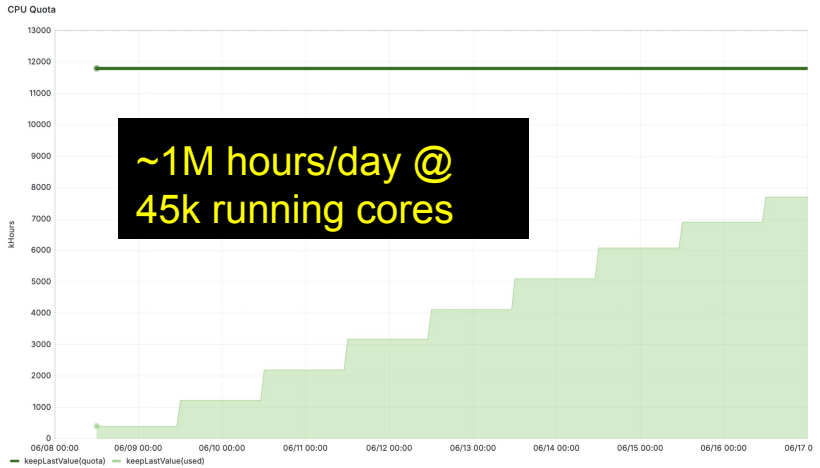
BSC integration by CMS

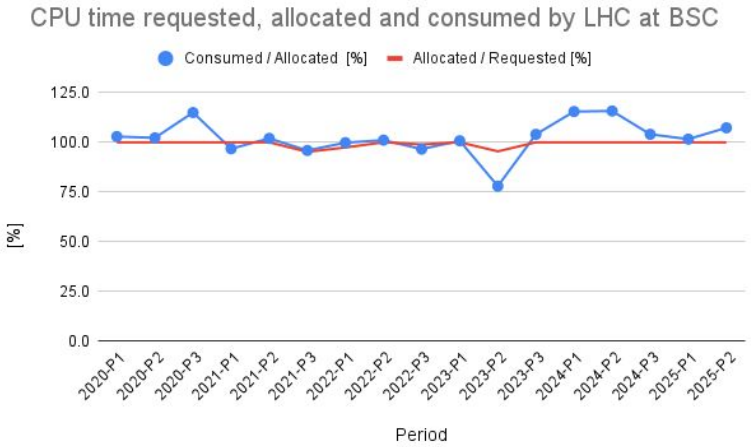


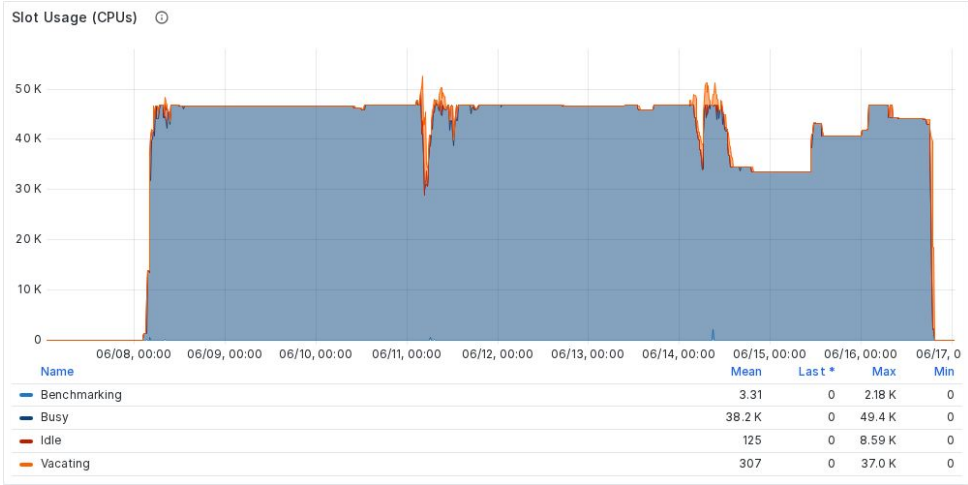


BSC exploitation

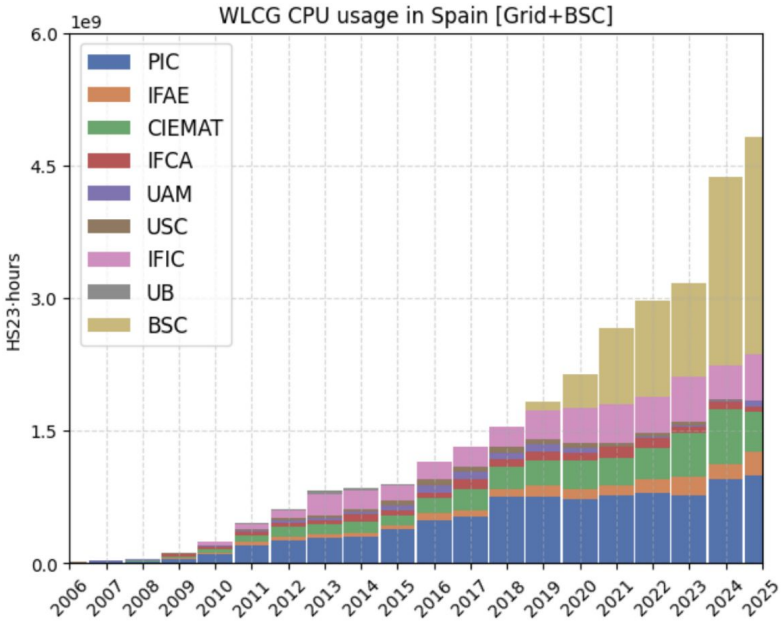


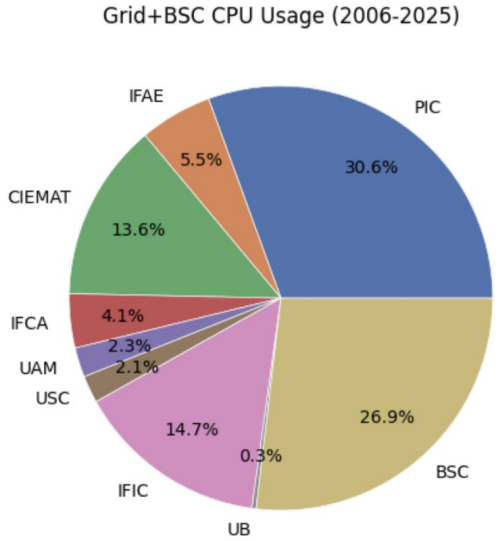






CPU delivered by Spain to WLCG





~3100 Million hours delivered during 2004-2025

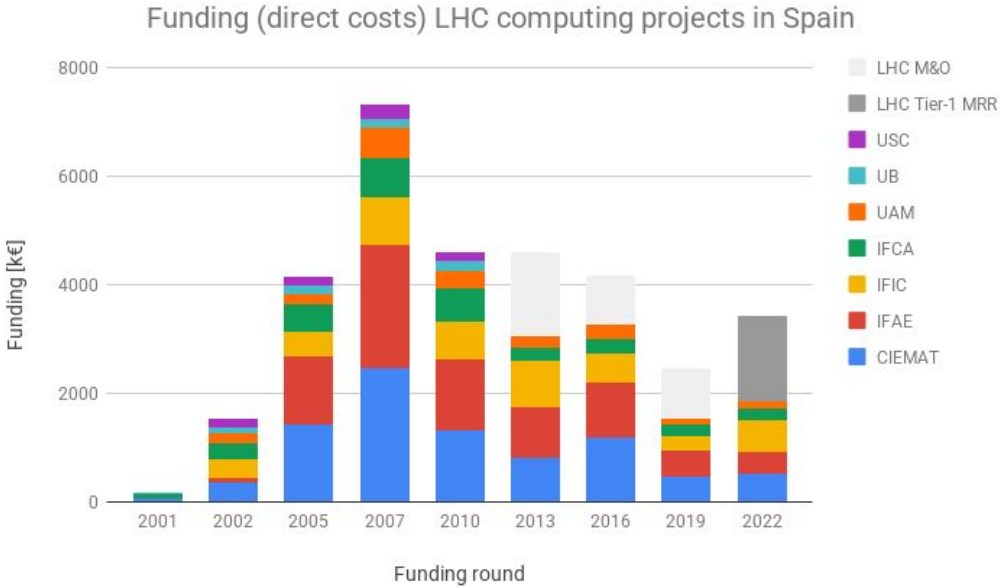
(31x10^9 HS23.hours; average CPU core power ~10 HS23)

Future of BSC strategic project

- The designation of LHC computing as a BSC strategic project has no expiration time, but not continuation guarantee either
- So far we got as much CPU as we requested
- No new CPU owned by WLCG-ES centers being purchased since BSC provides 50% of the pledged CPU
  - If we ever lost the BSC contribution, we would suddenly have a severe CPU under-provisioning
- Network limitations at BSC preclude a transparent usage by the LHC experiments, require significant integration effort by the Spanish community, and severely constrain the type of workflows
  - Currently limited to event generation and simulation
  - Workflows with input data, like data reconstruction, data analysis cannot be executed
  - BSC is the only HPC in EuroHPC with such strict network policies
- Meeting with BSC to be held in spring

WLCG-ES sustainability challenge

- Decreasing funding from HEP national programme
  - From ~2.5M€/year in 2007 to ~0.5 M€/year in 2019
  - Spanish contribution to WLCG reduced from 5% to 4% to 3%
  - Aging equipment
- Big effort to complement funding and resources
  - From institutions
  - From national/regional scientific infrastructure calls
  - Agreement with BSC (50% CPU)
- The average required funding for WLCG-ES is ~1.5M€/year
  - Hardware resources and personnel for operations and R&D



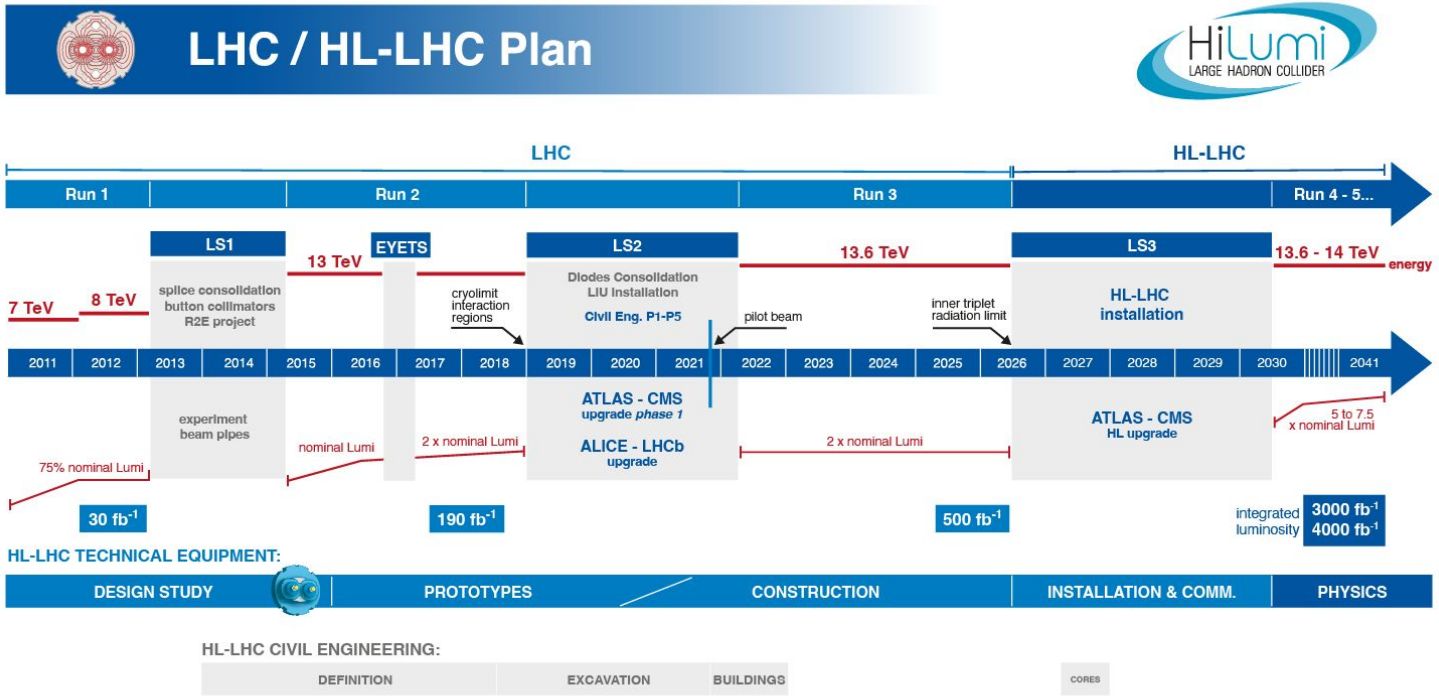
LHC M&O

- Maintenance and Operations costs of the LHC experiments
- Included in the funding of the computing projects, hence reducing the real money available for the projects

LHC Tier-1 MRR

- EU recovery funds for LHC upgrade (1.5 M€ for computing)
- Compensated the Tier-1 underfunding

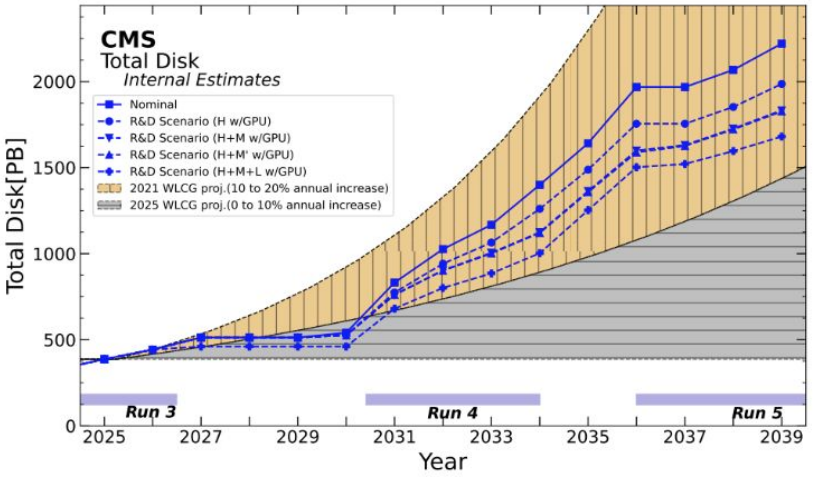
HL-LHC plan

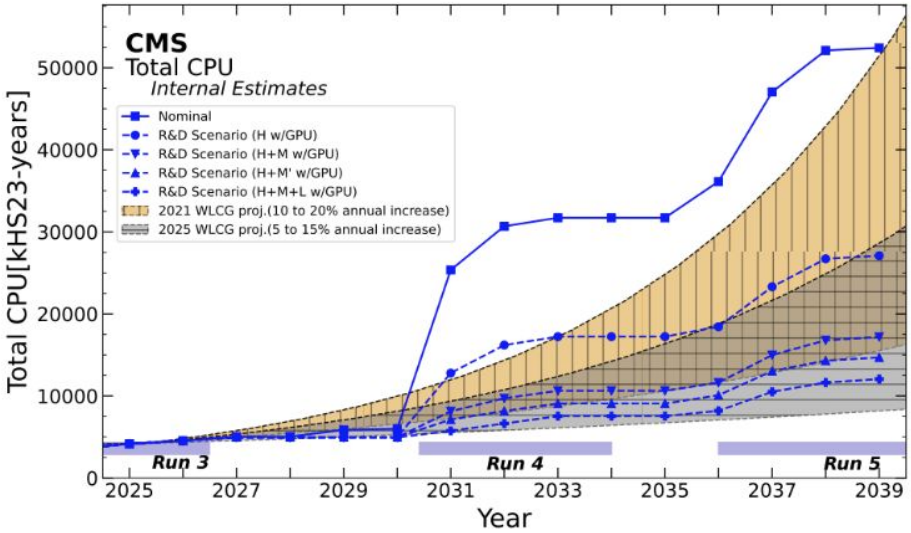


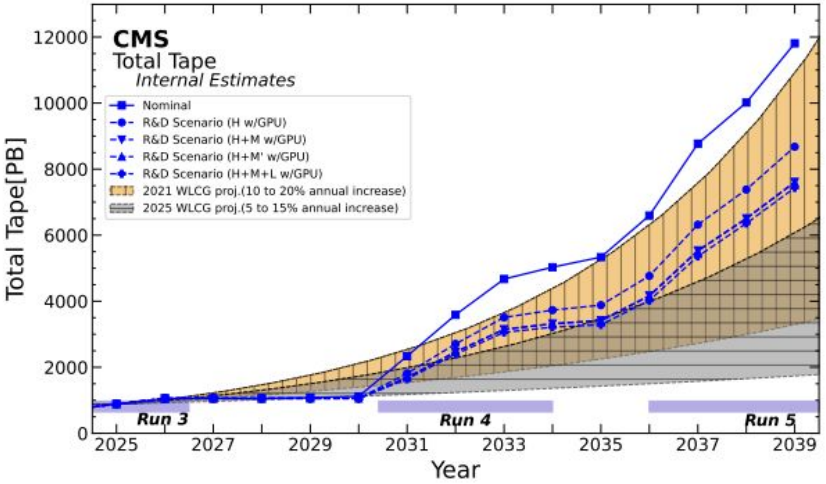
- Run 3 (2022-2026): **~2x more data**. **Evolutionary** changes in computing models
- Run 4 (HL/LHC, 2031+): **~10x more data**. **Revolutionary** changes required

CMS resource projections

Significant R&D needed to keep resource needs affordable with flat budget







HL-LHC computing challenge

- Intense R&D ongoing 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
  - Deployment of dedicated infrastructure for final stage analysis (Analysis Facilities)
- The goal is to face the resource challenge with constant funding
  - ~15% annual increase achieved during Run3 from hardware getting cheaper
  - But sustaining the current level of funding is needed!!!

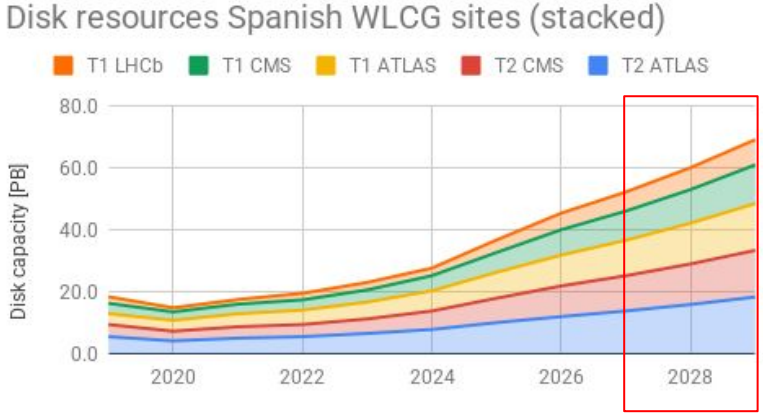
LHC computing projects for 2027-2029 in Spain

- Current research projects funded by FPN programme end August 30th, 2026
  - Extensions granted for 6-12 months to smoothly overlap with new projects
  - 1 coordinated project CIEMAT-IFAE for Tier-1 (ATLAS, CMS, LHCb) and Tier-2 (75% CMS, 25% 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 2025 funding call
  - Funding resources required for 2027-2029
  - Global coordination between the projects (modeling of requests, goals, etc)

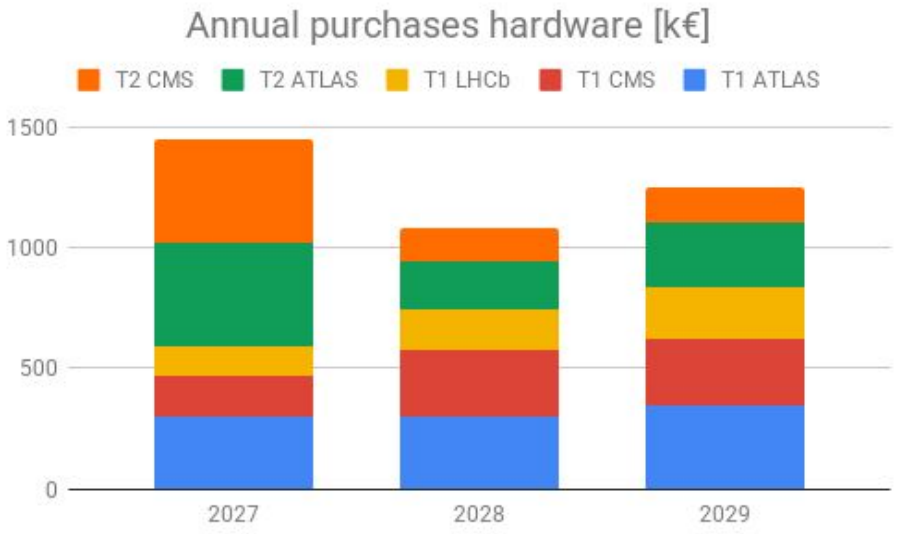
Modeling of resource cost

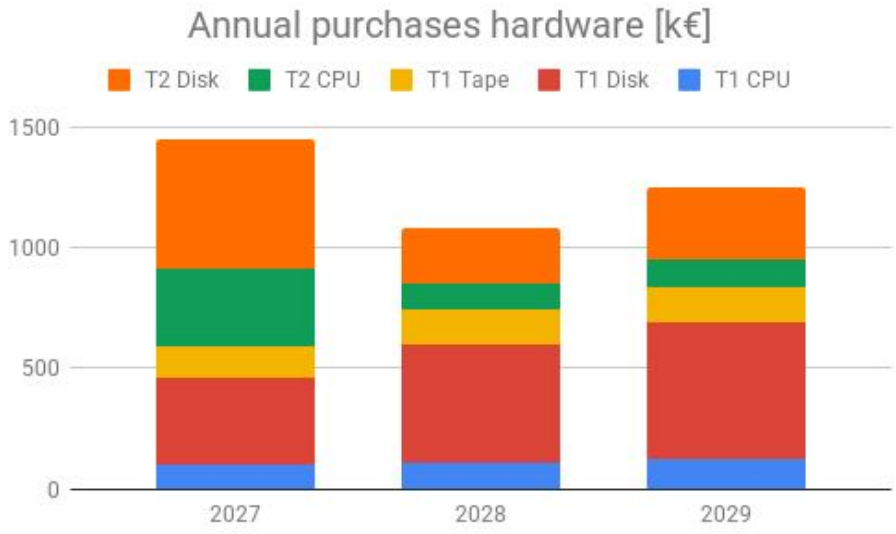
- According to the resource projections of LHC experiments, a moderate rise during the long shutdown (2027-2029) will be followed by a steep increase at the start of HL-LHC (2030)
- This round of projects request funding for the period 2027-2029
- If we do not increase resources during 2027-2029, we will not be able to fund and provide the large increase anticipated for 2030
- We propose an increase of 15% during 2027-2029, consistent with flat-funding, that would allows to increase by 50% the resources available in 2026
    - Push purchases to the end of the project to minimize aging of equipment and take advantage of hardware price drop
- Modeling of resource cost
    - Use official resource requests from the experiments (available until 2026), and project a 15% annual increase
    - Contribution from Spain of 5% at the Tier-1 and Tier-2 levels for ATLAS, CMS, LHCb
    - Assume a contribution of 50% of the required CPU from BSC
    - 5-year for disk and 8 year for CPU renewal policy for old equipment
    - Yearly hardware price decreases of 10%

Computing resources up to 2029 in Spain



Estimated expenses in computing resources for 2027-2029 in Spain





- ~ 1.25 M€/year investment in equipment required for 2027-2029
- In addition, 7 FTE (1 per site) for R&D, integration, operation: ~350 k€/year
- Overall: **~4.8 M€ all projects, 3 years**

New WLCG-ES computing projects 2027-2029: funding