

WLCG-ES computing: Perspectives and Strategy

José Hernández, Andrés Pacheco, José Salt

slides for the
I Workshop de Computing y Software de la Red Española de LHC



Centro de Investigaciones
Energéticas, Medioambientales
y Tecnológicas



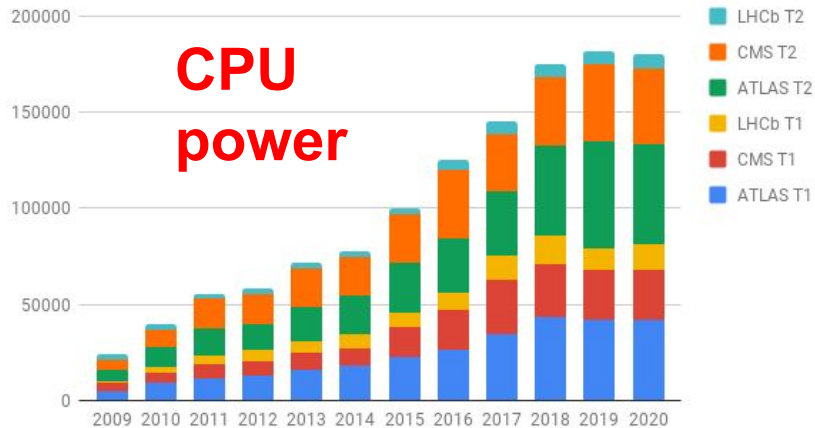
Outline

- WLCG-ES past and present
- Perspectives
 - Medium term (Run 3)
 - Long term (HL-LHC)
- Strategy

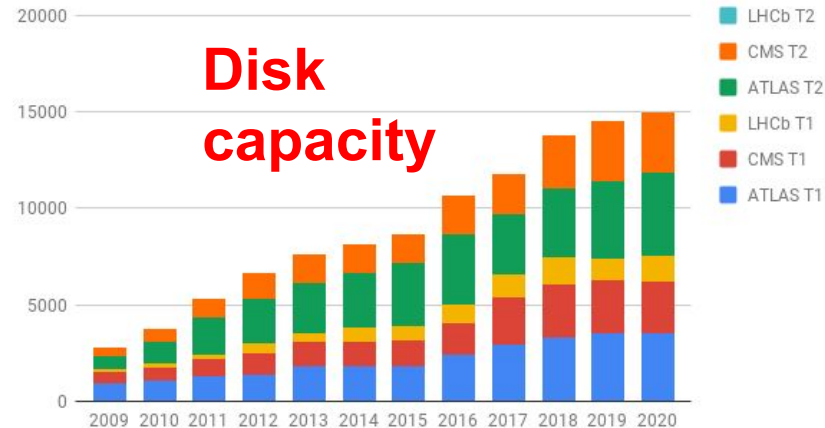
WLCG-ES: a success story

- Almost two decades contributing to LHC distributed computing infrastructure (Worldwide LHC Computing Grid, WLCG) at the highest level
 - ~5% of the WLCG resources, ~1200M CPU hours delivered
 - 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 plan 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
 - **We have generated a big strategic asset for our community!**
- Congratulations to everyone!

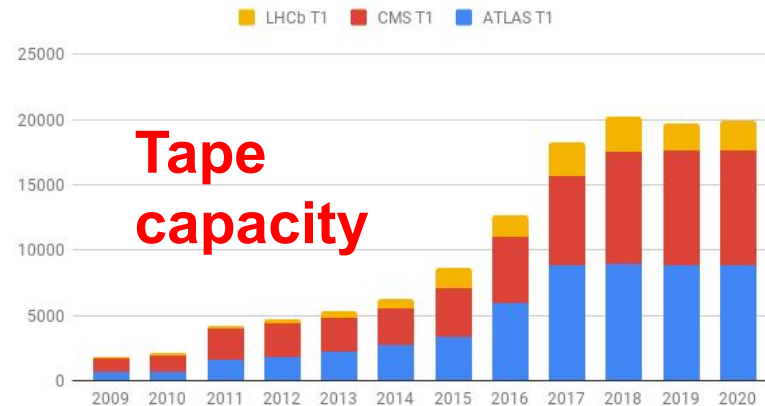
CPU power WLCG-ES (HS06)



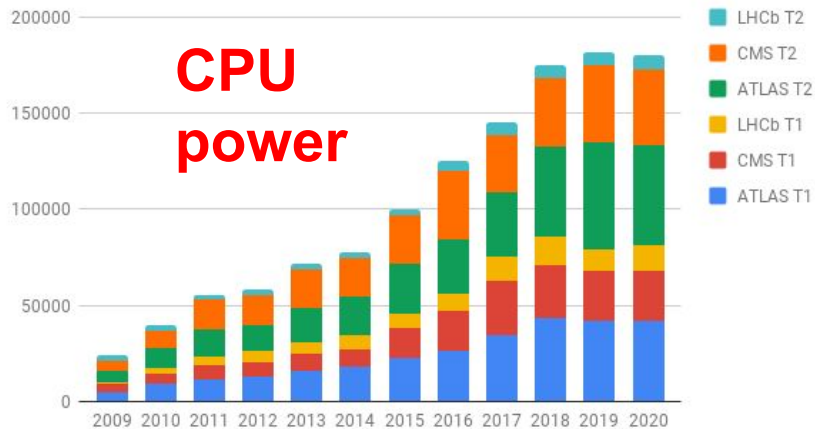
Disk capacity WLCG-ES (TB)



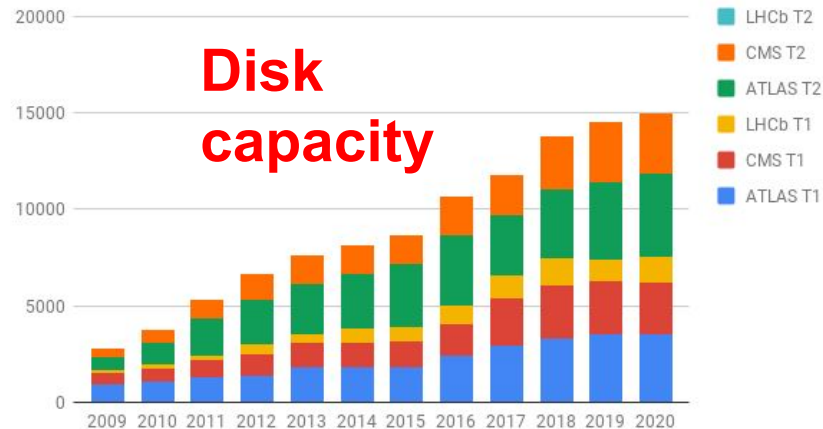
Tape capacity WLCG-ES (TB)



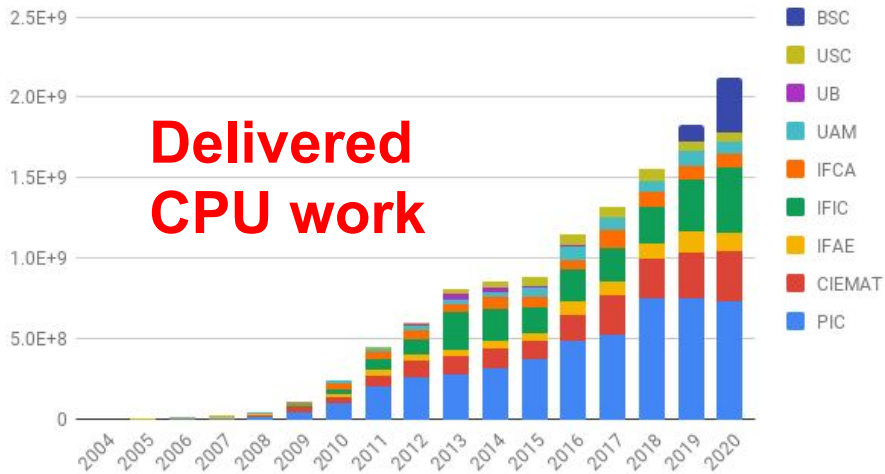
CPU power WLCG-ES (HS06)



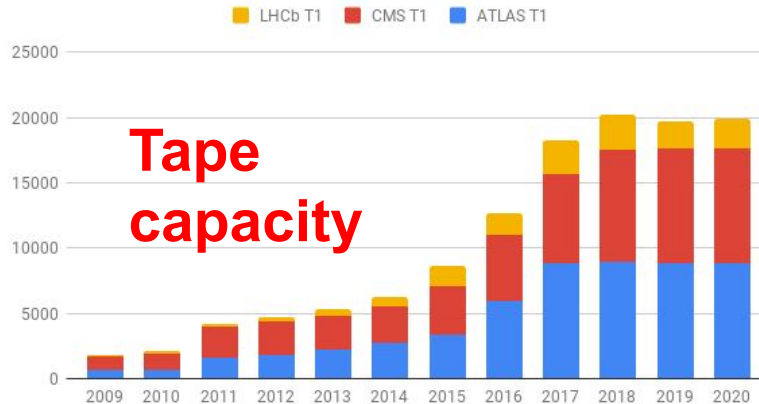
Disk capacity WLCG-ES (TB)



CPU time delivered WLCG-ES + BSC (HS06.hours)



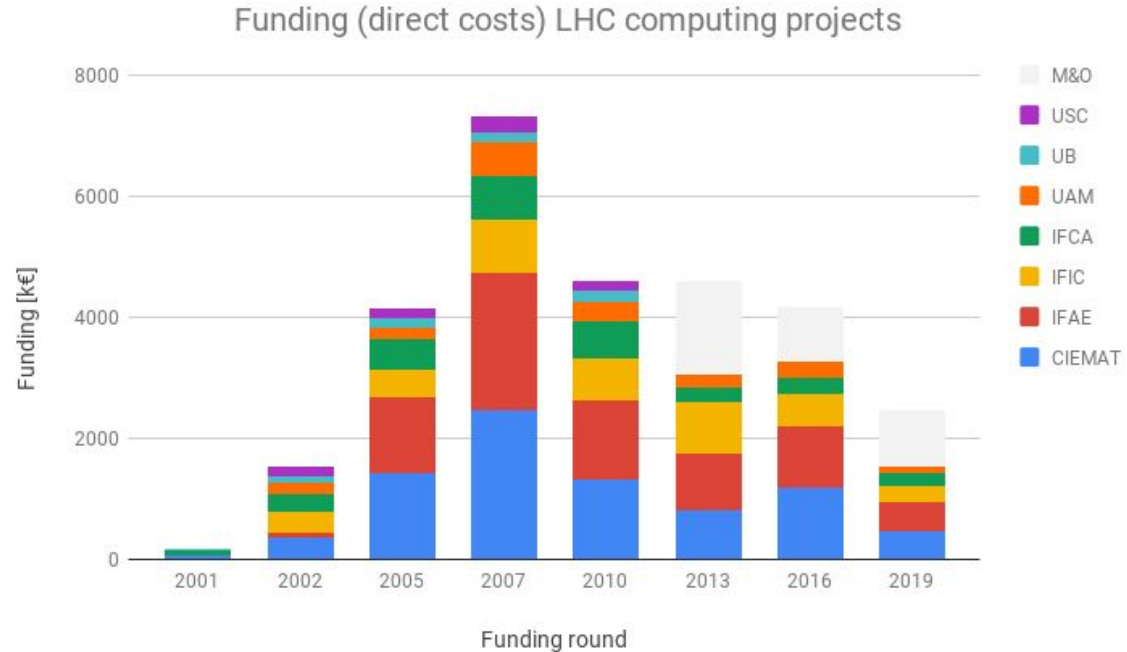
Tape capacity WLCG-ES (TB)



Funding of LHC computing projects from national program 2007-2022

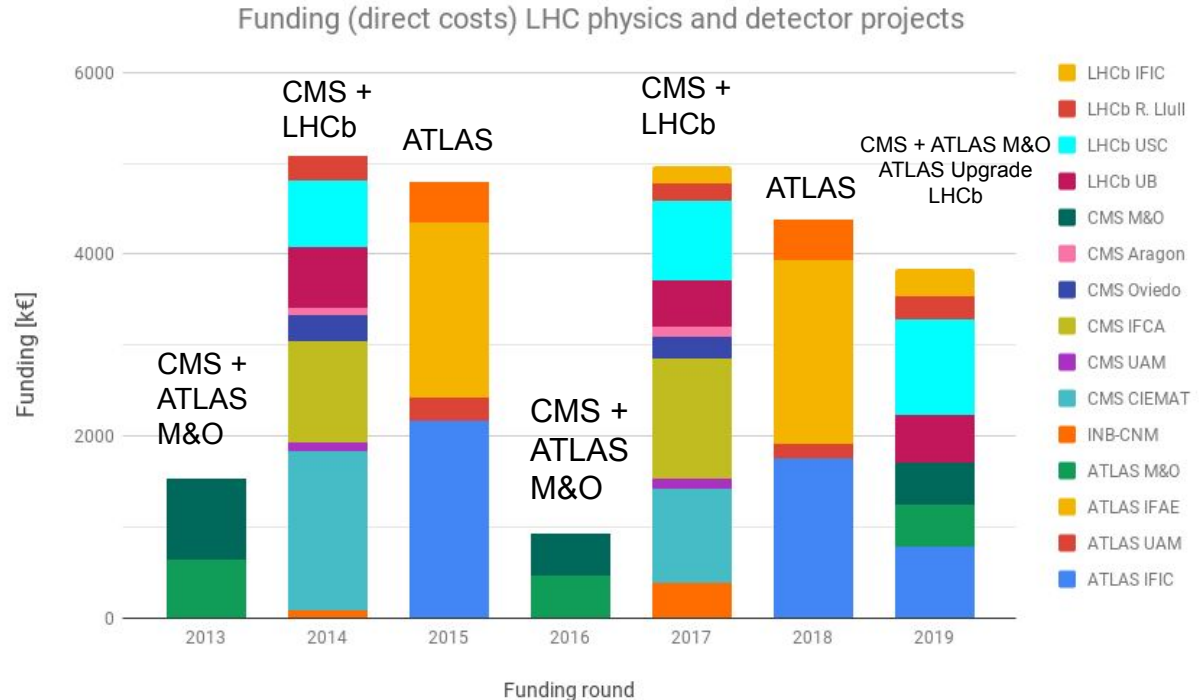
Decreasing funding

- From ~2.5M€/year in 2007 to ~0.5 M€/year in 2019
- Money for LHC M&O fees effectively reduces funding available for computing even further



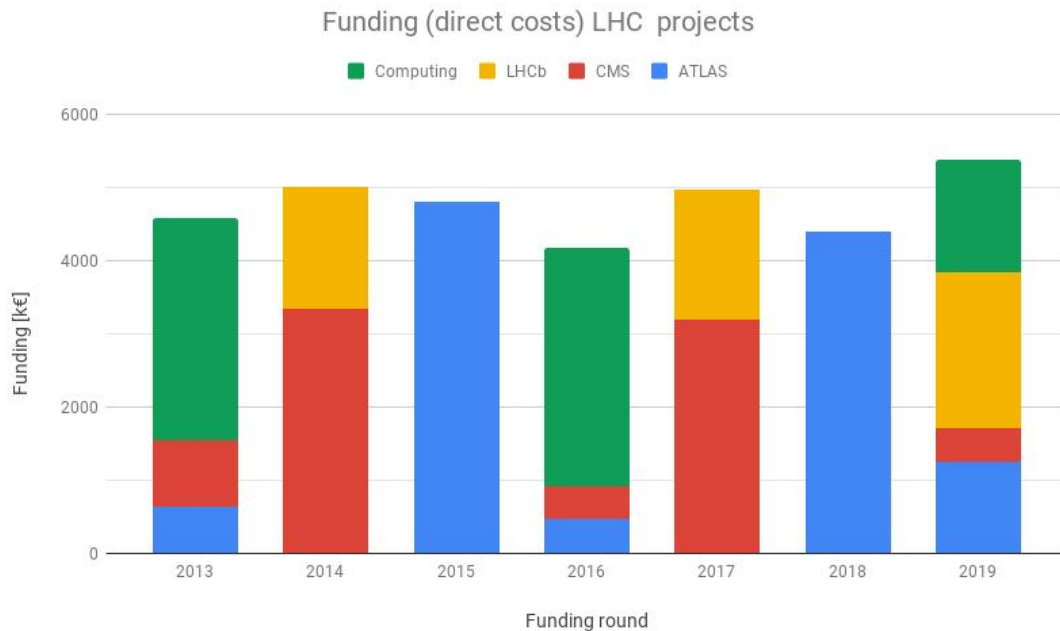
Funding LHC physics+detector projects 2013-2019 (no LHC computing)

- ~constant funding
 - ATLAS ~ 5M€/3years
 - CMS ~4M€/3years
 - LHCb ~2M€/3years
- Part of M&O for ATLAS and CMS in computing rounds
- LHCb funding round shifted to the same year as LHC computing



Funding LHC physics+detector+computing (2013-2019)

- Constant funding ~4.5M€/year
- Physics + detector: ~ 85% total LHC funding, ~ 3.75 M€ / year
- **Computing: ~ 15% total LHC funding, ~0.75 M€ / year**
- 2019 round very competitive
 - LHCb ~2M€
 - Computing ~1.5 M€
 - ATLAS upgrade ~ 0.8 M€
 - CMS&ATLAS M&O ~0.9 M€



Funding conclusions

- **Decreasing funding** for WCLG-ES computing over the last decade, while funding for LHC detector + physics projects kept constant
 - The inclusion of part of the CMS & ATLAS M&O costs in the computing projects has led to an effective reduction of the funding available for computing
 - Shifting LHCb funding round together with WCLG-ES in 2019 also contributed to further funding reduction for computing
- We were told to decrease our contribution to WLCG (from 5% to 3-4%)
 - Adjust to “Spanish physics contribution” to LHC experiments (~3% signing authors)
- Is the funding situation for WLCG-ES computing **temporary or permanent?**
- Reaching an **unsustainable situation**
 - Current funding only allows us to keep our contribution to 3-4% level in 2022
 - Difficult to follow the growth of LHCb requirements for Run 3 (>50%/year)
 - No funding for foreseen resource increases in 2023 (new projects only expected by mid 2023)

Changing paradigm for funding LHC computing?

- Strategic agreement with Barcelona Supercomputing Center
 - LHC computing is a BSC strategic project (one of the existing three)
 - It required a long negotiation process
 - Provide ~50% of the pledged CPU by Spain to LHC (CPU for simulation)
 - ~50 Mhours/year, equivalent to ~5000 cores
- Varying degrees of difficulty to integrate BSC in the LHC computing systems
 - No Internet connectivity in BSC nodes
 - ATLAS had an “easy” integration through the ARC-CE interface
 - CMS had to do considerable developments
 - LHCb is trying to use the same schema with the ARC-CE
- WLCG requires national resource pledges contributed through WLCG sites supporting all workflows
 - Integrate HPC resources as transparent extensions of WLCG sites
 - Expected commitment and support

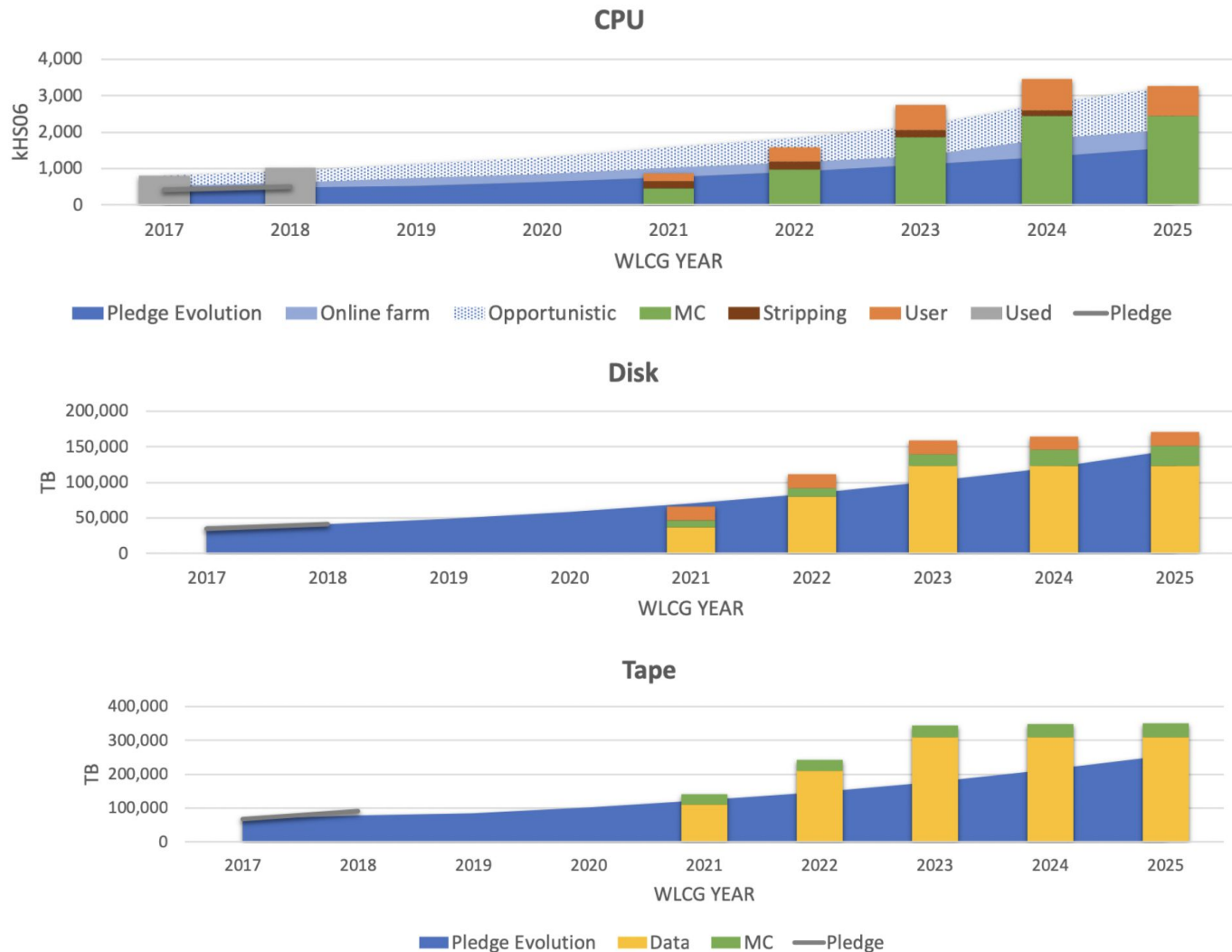
What about storage resources

- **Data is our distinctive asset!**
 - Storage is the most expensive resource as well
- Ongoing developments in WLCG to reduce costs of storage infrastructure
 - Data lakes, reduce data duplication, analysis via caches, central processing through buffers
- Need proper funding
- Spanish Supercomputing Network (RES) extended with Data Nodes
 - Provide storage capacity and services to scientific projects
 - PIC is one of the new data nodes of the RES
 - Pilot program with 3-year projects
 - Unclear funding
- What is the **funding model** for the computing infrastructure in Spain in a longer term?
 - Current storage pledges are 15 PB disk and 20 PB tape

Run 3 and HL-LHC resource need projections

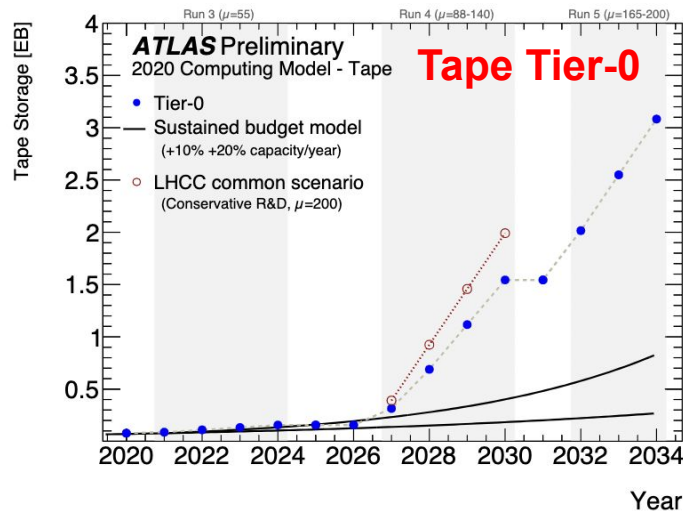
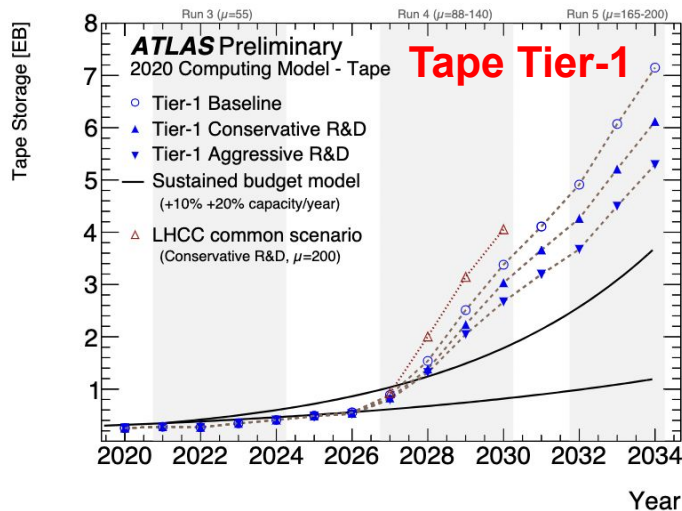
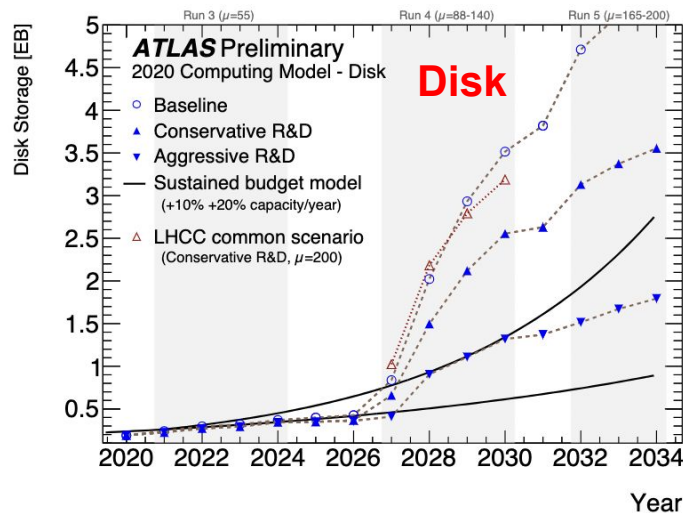
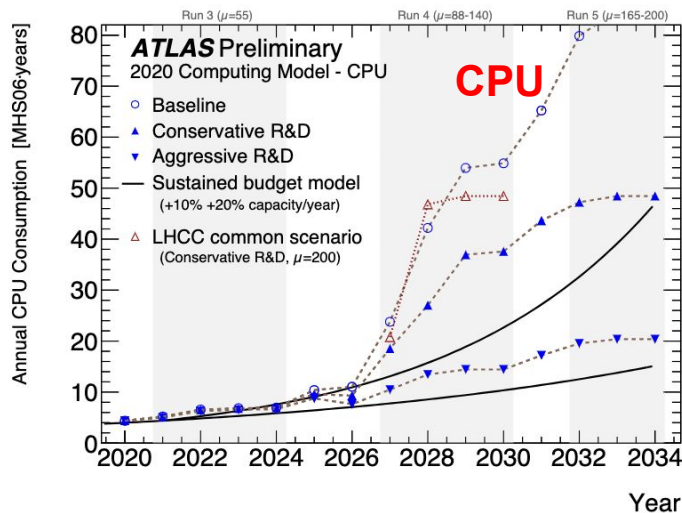
- Modest ($\sim 10\%$ /year) foreseen resource increases along Run 3
 - Exception is LHCb, since its high-lumi phase is Run 3
 - Large ($>50\%$ /year) increases in resource needs
- Large step function for HL-LHC
 - Smaller than thought/feared
- Quite some R&D underway
 - Small data formats ($\sim \text{KB}/\text{event}$) for analysis
 - Try to reduce cost of storage infrastructure (data lakes, caches, buffers)
 - CPU from outside WLCG resources (HPC, GPUs, Clouds)
 - Tabular data formats for easy massive filtering operations

- For LHCb its HL-LHC is Run 3
- Large (>50%) annual increases 2022 to 2024
- By 2024, LHCb resources close to CMS
- Online reconstruction throwing away $\frac{2}{3}$ of raw data!

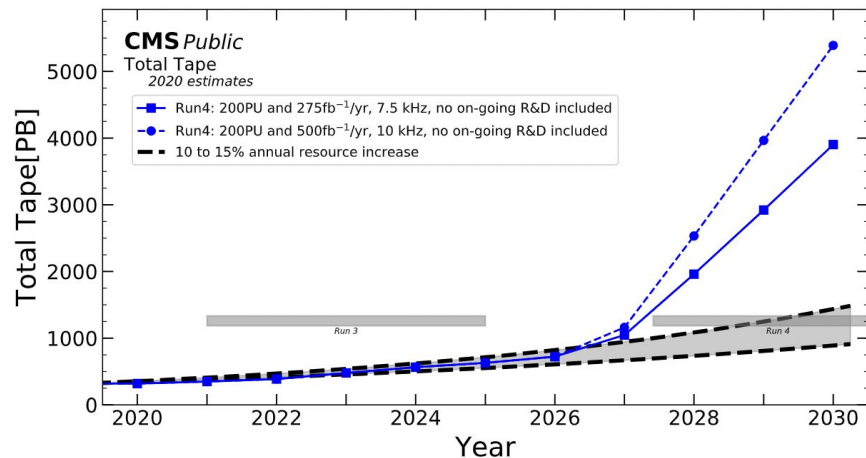
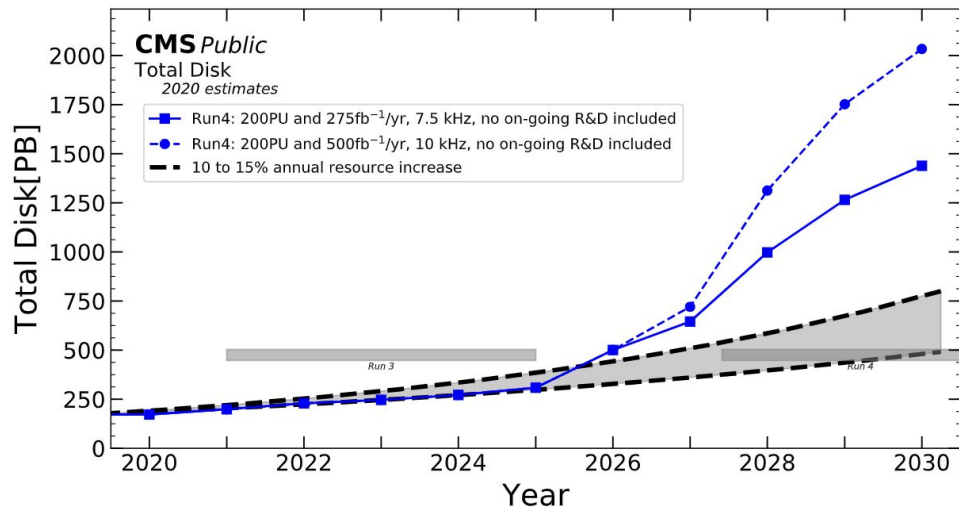
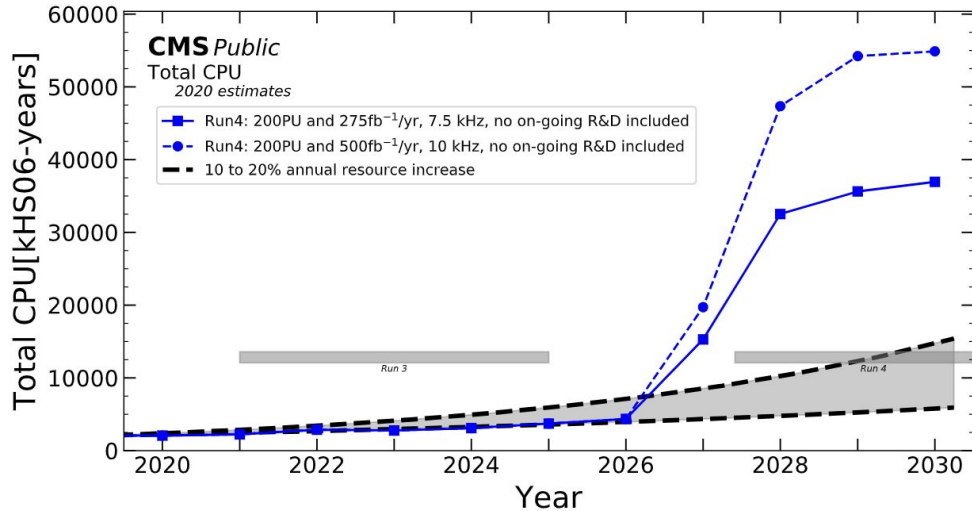


ATLAS computing resource needs projections

- ~10% annual increase till 2026
- Large increase for HL-LHC (2027+)



CMS computing resource needs projections



Our vision?

- Should we try to put forward our **vision/proposal** for LHC computing in Spain?
 - Prepare a document for next HEP national program manager?
- Medium term (Run 3, 2022-2024)
 - BSC contribution to CPU important, due to decreasing funding
 - Need still CPU for data processing (mainly at T1) and (interactive) analysis close to the physicists (mainly at T2s)
 - Need proper funding for data infrastructure
 - Build Spanish data lake
 - Contribution from RES data?
- Long term (HL-LHC, 2025+)
 - Which fraction of the infrastructure will we be able/willing to fund?
 - **Analysis Facilities are a very important aspect**
 - Specialized for highly parallel local and interactive analysis
 - Key to be competitive in data analysis

Conclusions

- We have developed a **great infrastructure and community within WLCG-ES!**
- Important message:
Computing must be considered as a detector, so it should be part of the funding strategy for the HL-LHC upgrades
 - Ongoing actions and ministry level for funding detector upgrades
 - LHC computing upgrade forgotten or dismissed?
- **Funding model**
 - Is there a way to get structural/consistent funding for computing infrastructure beyond the national R&D HEP funding program?
 - Role of RES?
 - Other national e-infrastructure funding programs? E.g. from EU post-covid19 reconstruction funds?
- Tighter organization/coordination at the technical level to evolve the new computing infrastructure towards HL-LHC
 - Spanish Data lake
 - Analysis Facilities
- **We need to be proactive!**