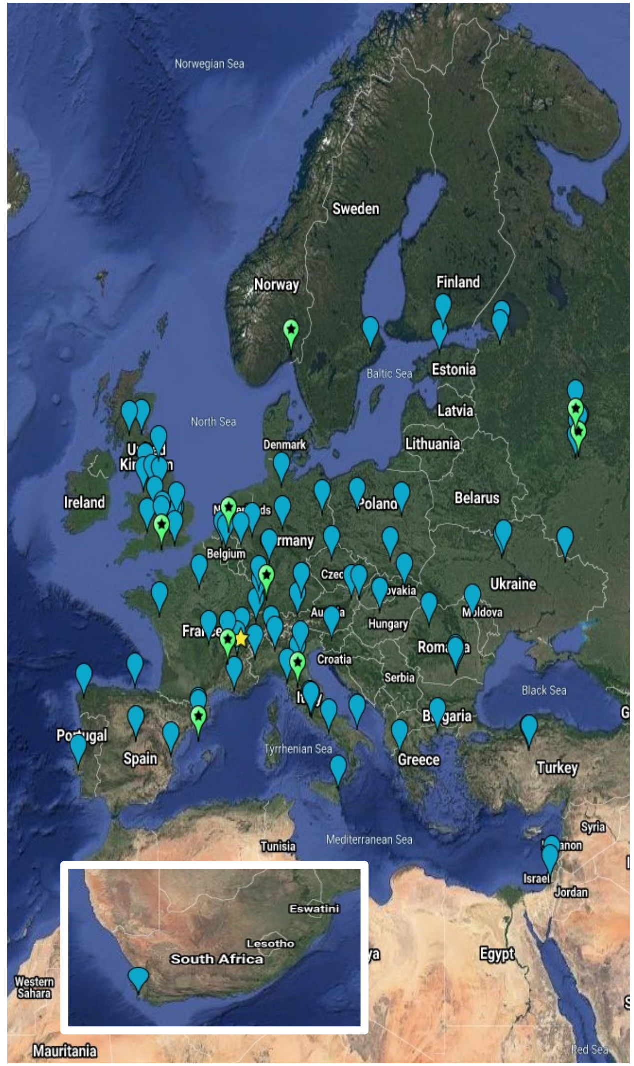
A holistic study of the WLCG energy needs for HL-LHC

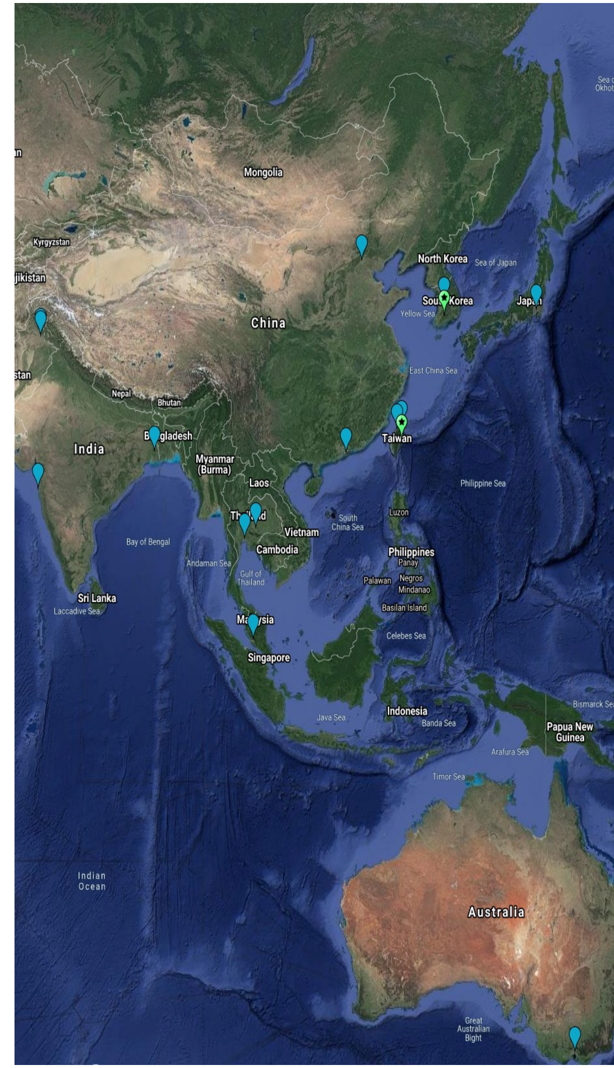
D. Britton (Glasgow), S. Campana (CERN), B. Panzer (CERN)

The WLCG Collaboration provides the distributed computing infrastructure for the needs of the CERN Large Hadron Collider experiments

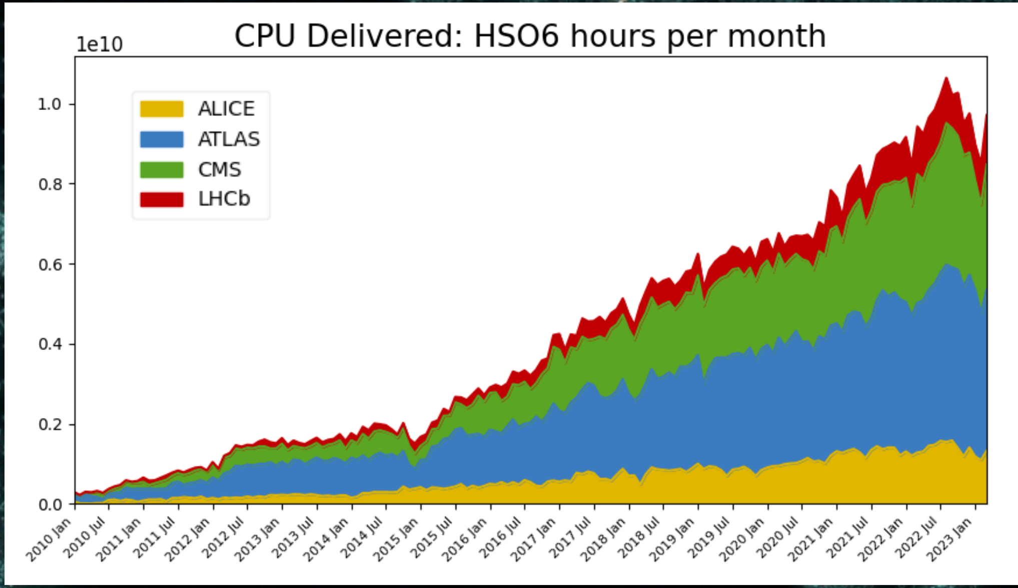
WLCG
Worldwide LHC Computing Grid

65 MoU's
164 sites
42 countries





Worldwide data processing and data transfers 24/7/365, increasing over time



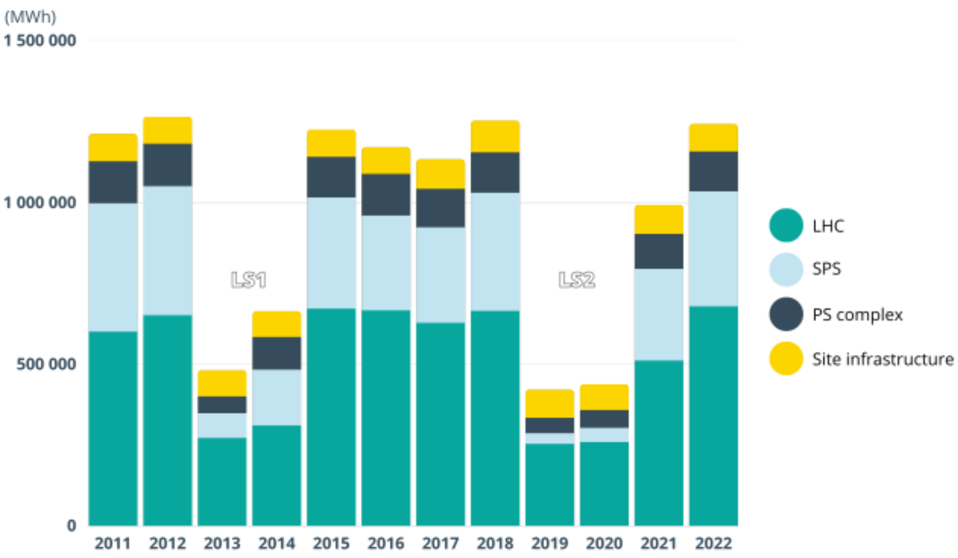
WLCG
Worldwide LHC Computing Grid

LHC computing energy needs

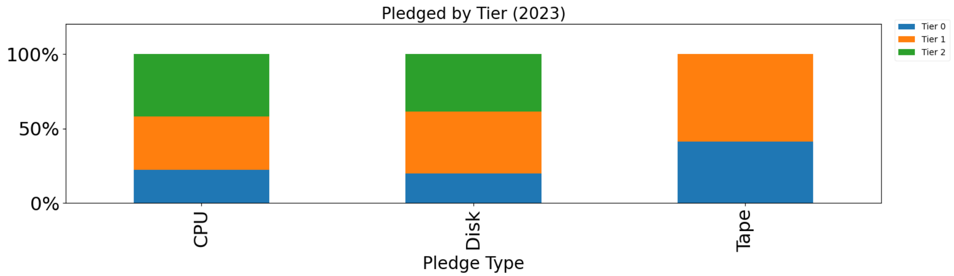
The CERN energy consumption is ~1.25TWh/year during LHC runs

The CERN IT infrastructure contributes <5% in average

CERN provides ~20% of the WLCG resources



https://hse.cern/content/energy-management



In terms of energy needs for the LHC program, the impact of computing is non negligible and motivates this study

Disclaimers

This is ”A **holistic** study of the WLCG **energy** needs for HL-LHC”

WLCG is a distributed infrastructure built on top of loosely coupled facilities in different countries and continents

- We do not attempt to make a detailed study on a country by country (or site by site) basis but rather a general one
- We discuss only energy needs and not impact on e.g. carbon footprint as this is also country specific

There are however interesting studies [1] about the impact on CO2 in WLCG federations.

[1] https://indico4.twgrid.org/event/25/contributions/1211/

Factors impacting the WLCG energy needs

The amount of consumed energy in WLCG to support the scientific program of the LHC experiments and its evolution depends on different factors:

- **The LHC luminosity:** more luminosity (and therefore, data) and more complex event data => increased need in more compute resources and more storage.

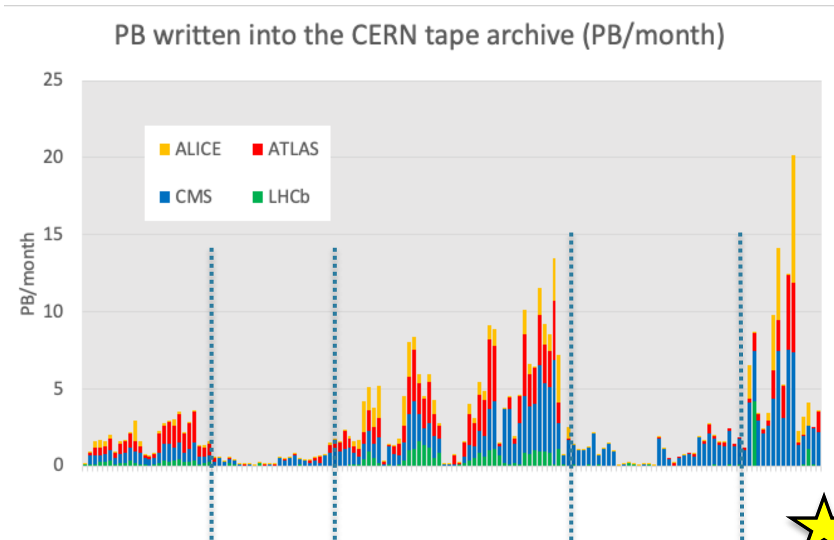
- **The evolution of the computing models and the software:** the progress made with different computing R&D activities and particularly software efficiency and performance will play a key role defining the resource needs

- **The hardware technology**, modern hardware improves energy efficiency and also reliability

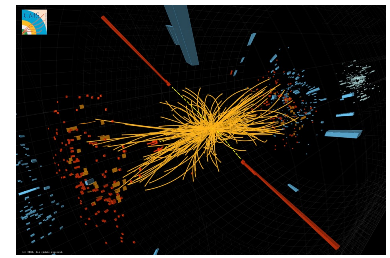
- **The facilities** hosting the WLCG hardware are progressively being modernized to improve the Power Usage Effectiveness (PUE) as part of their renovation.

WLCG
Worldwide LHC Computing Grid

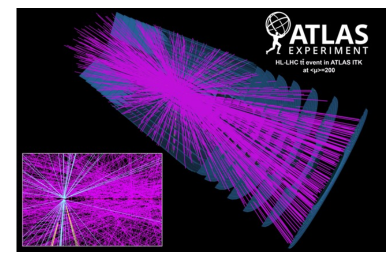
LHC Luminosity and Run conditions



2011 Higgs candidate



HL-LHC simulated event



2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038

LHC Run-1	LS1	LHC Run-2	LS2	LHC Run-3	LS3	HL-LHC Run-4	LS4	HL-LHC Run-5
L=29 fb-1		L=156 fb-1		L~350 fb-1		L~900 fb-1		L~1200 fb-1

(see next slide)

LHC luminosity and parameters

The conditions we used for this study are in the table below

2023, 2024 and 2025: < 100/fb luminosity, < 50 pileup, 6x10^6 seconds of pp time

(they are not the most up-to-date numbers, but good for this analysis)

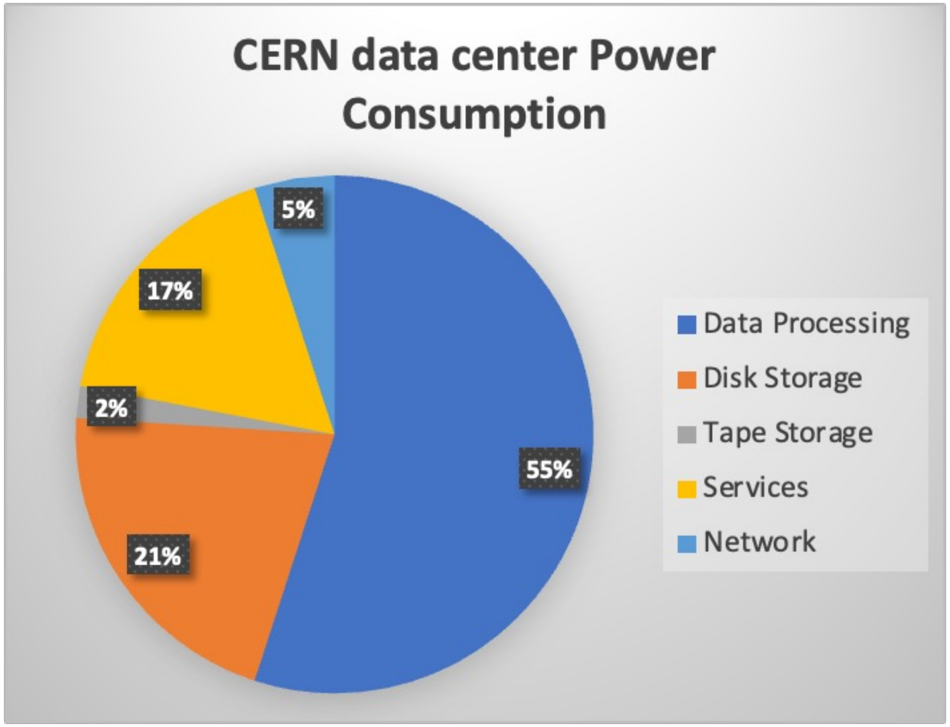
	Run-4 (2029-2032)	Run-5 (2035-2038)
ATLAS/CMS luminosity	<270/fb (<135/fb in 2029)	<340/fb (<170/fb in 2035)
ATLAS/CMS average pile -up	<140 (<70 in 2029)	<200 (<100 in 2035)
LHCb luminosity	15/fb	50/fb
Alice luminosity (pp)	100/pb	100/pb
Running time (pp)	6 M seconds	6 M seconds
Running time (ions)	1.2 M seconds	1.2 M seconds

WLCG data centers power consumption

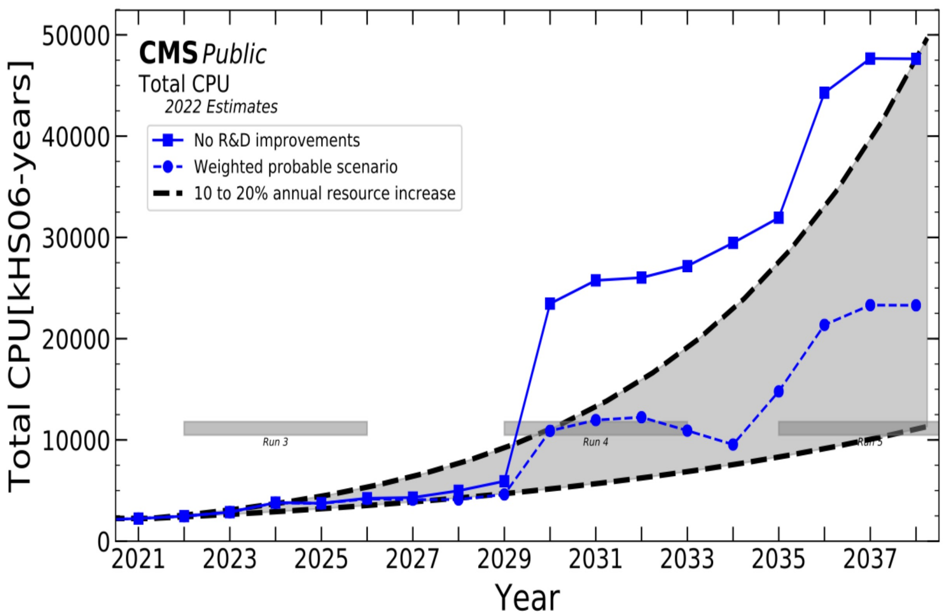
The pie chart shows the breakdown of the power consumption at the CERN data center

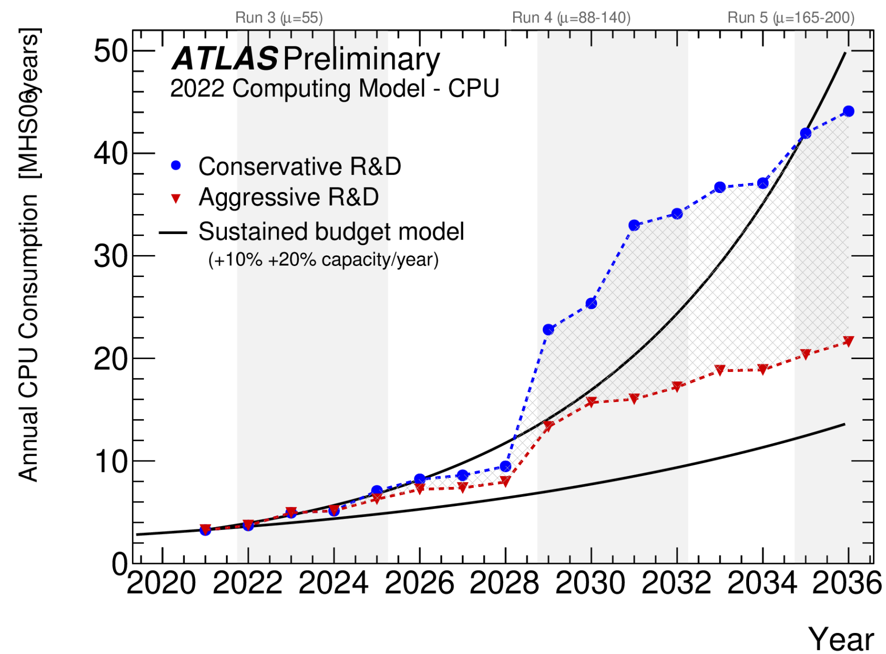
Most of the power is consumed for data processing (CPUs). Large part of the “services” are in fact CPUs

In this study we will focus on the energy needs for CPUs



ATLAS and CMS CPU needs for HL-LHC





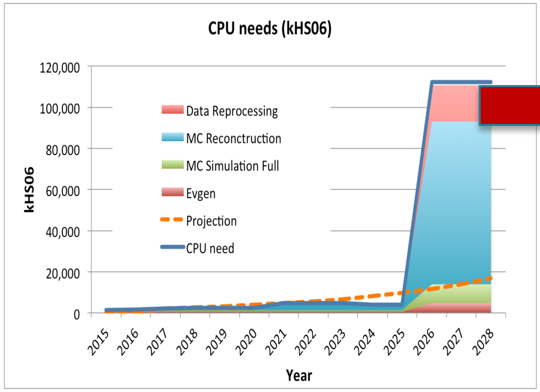
The projected CPU needs of ATLAS and CMS at HL-LHC. Estimates produced for the 2021 HL-LHC computing review and updated in 2022 to reflect the changes in the LHC schedule

ALICE and LHCb will require considerably less in Run-4, while no firm estimates are available for Run-5. We will consider ATLAS and CMS only in this study.

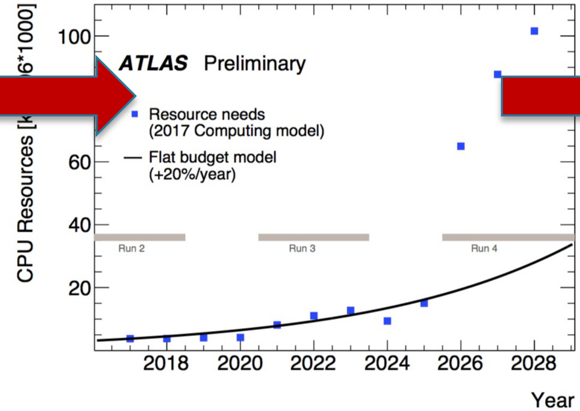
HL-LHC computing resource needs evolution

WLCG
Worldwide LHC Computing Grid

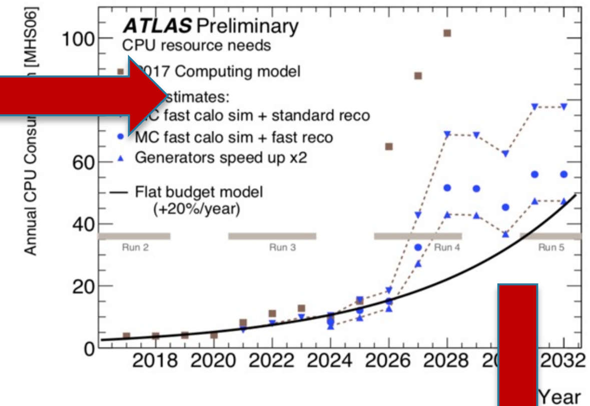
2015



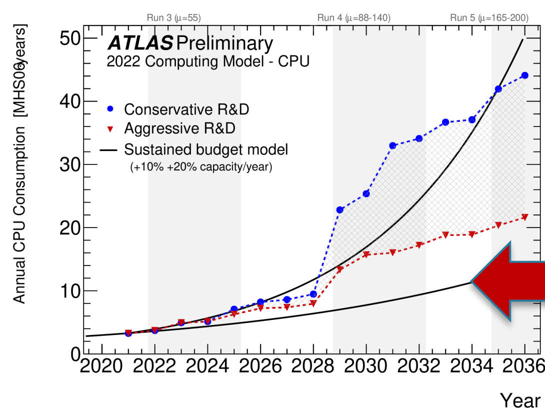
2017



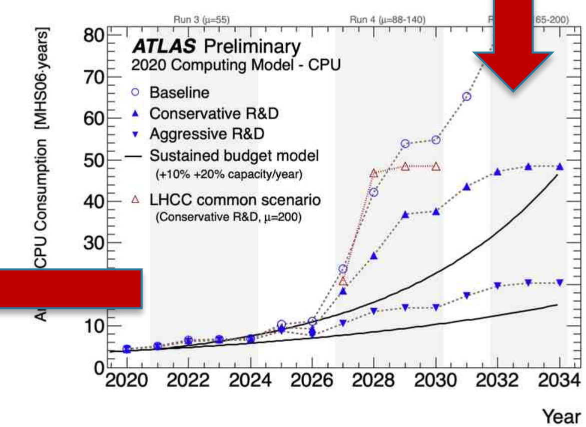
2018



2022



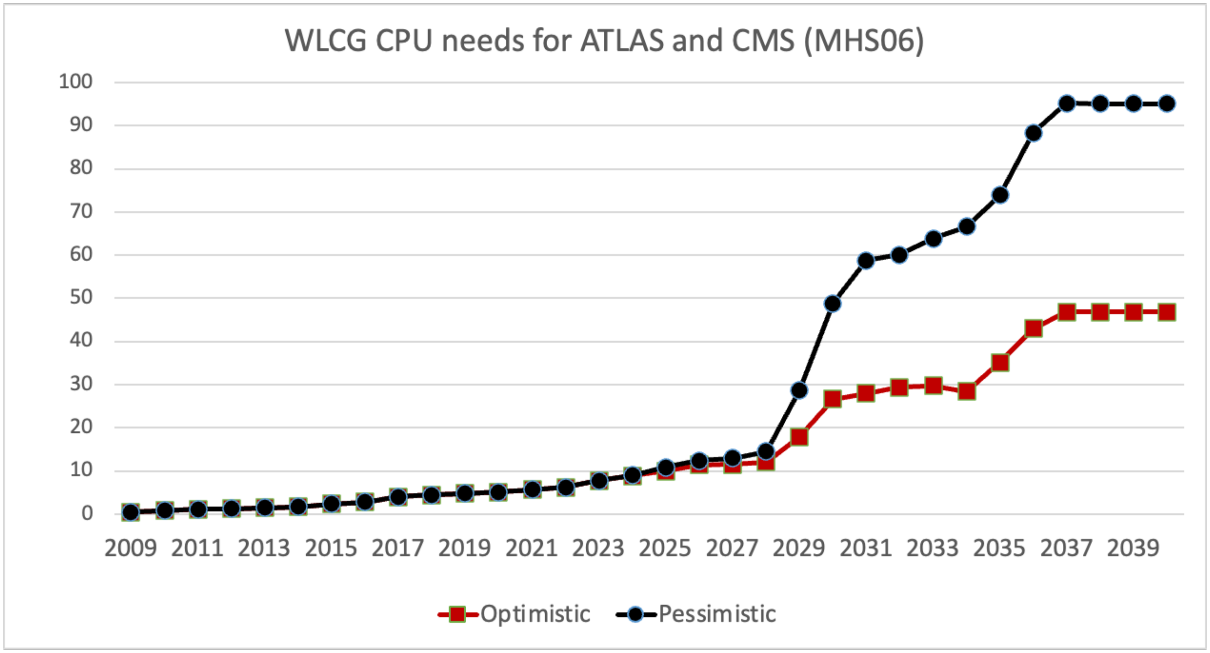
2020



2015 projections: resource needs = 10x more than budget allows

2022 projections: resource needs compatible with budget (optimistic scenario)

ATLAS and CMS CPU needs for HL-LHC



ATLAS (conservative R&D) + CMS (no R&D improvement) = **PESSIMISTIC** scenario

ATLAS (aggressive R&D) + CMS (weighted probable scenario) = **OPTIMISTIC** scenario

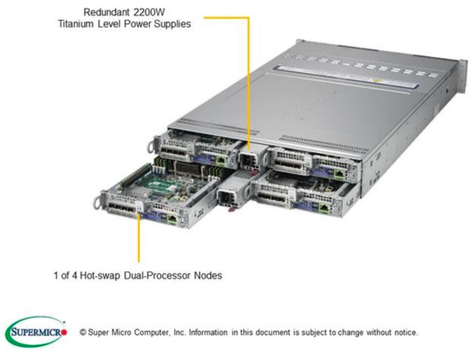
HS06 into Watts

Consider a dual AMD 7302 processor, 4 TB SSD, 256 GB memory, 10 Gbit NIC

Incorporate 4 separate servers into a common chassis with common redundant power supply (minimize the infrastructure over-head)

The mentioned configuration has, per server:

- a performance value of 1040 HS06,
- an idle power value of 120 W
- and a full load value of up to 420 W



Consider 80% CPU efficiency (as an average in WLCG) => **350 W/kHS06**

(this is the number we use for a processor of "today" in this study)

HS06 into Watts

The underlying semiconductor manufacturing technology for processors is continuously improving and the feature size is shrinking

- increase of performance or the reduction in energy usage or a compromise between the two

Advertised PPA Improvements of New Process Technologies Data announced during conference calls, events, press briefings and press releases							
	TSMC						
	N7 vs 16FF+	N7 vs N10	N7P vs N7	N7+ vs N7	N5 vs N7	N5P vs N5	N3 vs N5
Power	-60%	<-40%	-10%	-15%	-30%	-10%	-25-30%
Performance	+30%	?	+7%	+10%	+15%	+5%	+10-15%
Logic Area Reduction % (Density)	70%	>37%	-	~17%	0.55x -45% (1.8x)	-	0.58x -42% (1.7x)
Volume Manufacturing				Q2 2019	Q2 2020	2021	H2 2022

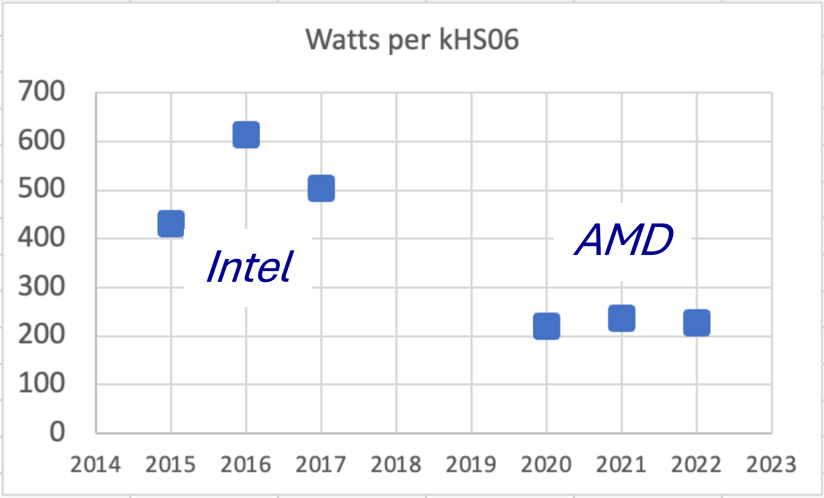
The energy consumption per unit of computation (W/HS06) decreased by ~50% over the last 5 years. We assume this trend will continue in the next years

The hardware replacement strategy also plays a role: we assume a 5 years lifecycle

CPU Power Consumption: Technology Trends

WLCG
Worldwide LHC Computing Grid

Year	Glasgow Hardware	Loaded Power Watts	HS06	Watts per kHS06	KWh per kHS06-year
2015	*2*Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60GHz*	160	371	431	3776
2016	*2*Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz*	210	342	613	5373
2017	*2*Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz*	210	416	505	4422
2020	*2*AMD EPYC 7452 32-Core Processor @ 2.30GHz*	390	1766	221	1934
2021	*2*AMD EPYC 7452 32-Core Processor @ 2.40GHz*	420	1766	238	2083
2022	*2*AMD EPYC 7513 32-Core Processor @ 2.60GHz*	480	2112	227	1991



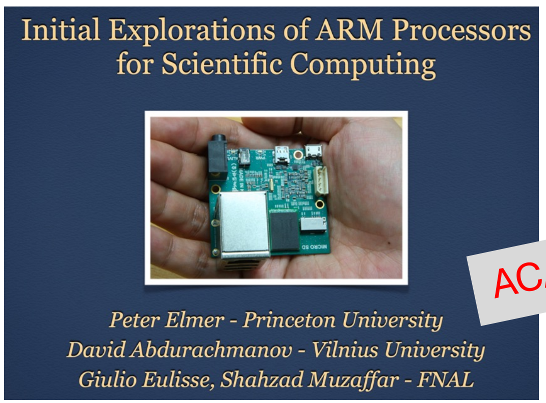
The 50% decrease in (W/HS06) of the last 5 years was not a gradual process.

A step-change happened with a new technology. **What about the future then?**

ARM as next Technology Step ?

ARM (Advanced RISC Machine) chips have low power consumption and heat generation and used extensively in portable, battery-powered devices, such as smartphones, laptops.

LHC experiments have kept an eye on this technology for over a decade. But …



ACAT 2013

Benchmarks - Simulation

Type	Cores	Power	Events/ min/core	Events/ min/Watt
Exynos4412 Prime @ 704GHz	4	4W?	1.14	1.14
Xeon E5520 @ 2.27GHz	2x4	120W?	3.50	0.23
Xeon E5-2630L @ 2.0GHz	2x6	190W?	3.33	0.21

☹ Low power but slow in 2013

☹ Porting HEP software on non-X86 arch. non trivial

☺ Recent huge increase in performance of mobile devices over last 10 years

☺ LHC experiments have software releases for ARM (used in some HPC) and have done some (but not all) physics validation

ARM_64 vs AMD(x86)

GridPP compared [1] performance and power consumption of two ~**similar spec-ed and similar priced** machines: AMD(x86_64) and ARM_64

x86_64: Single AMD EPYC 7003 series (Milan)
- CPU: AMD EPYC 7643 48C/96T @ 2.3GHz (TDP 300W)
- RAM: 256GB (16 x 16GB) DDR4 3200MHz
- HDD: 3.84TB Samsung PM9A3 M.2 (2280)

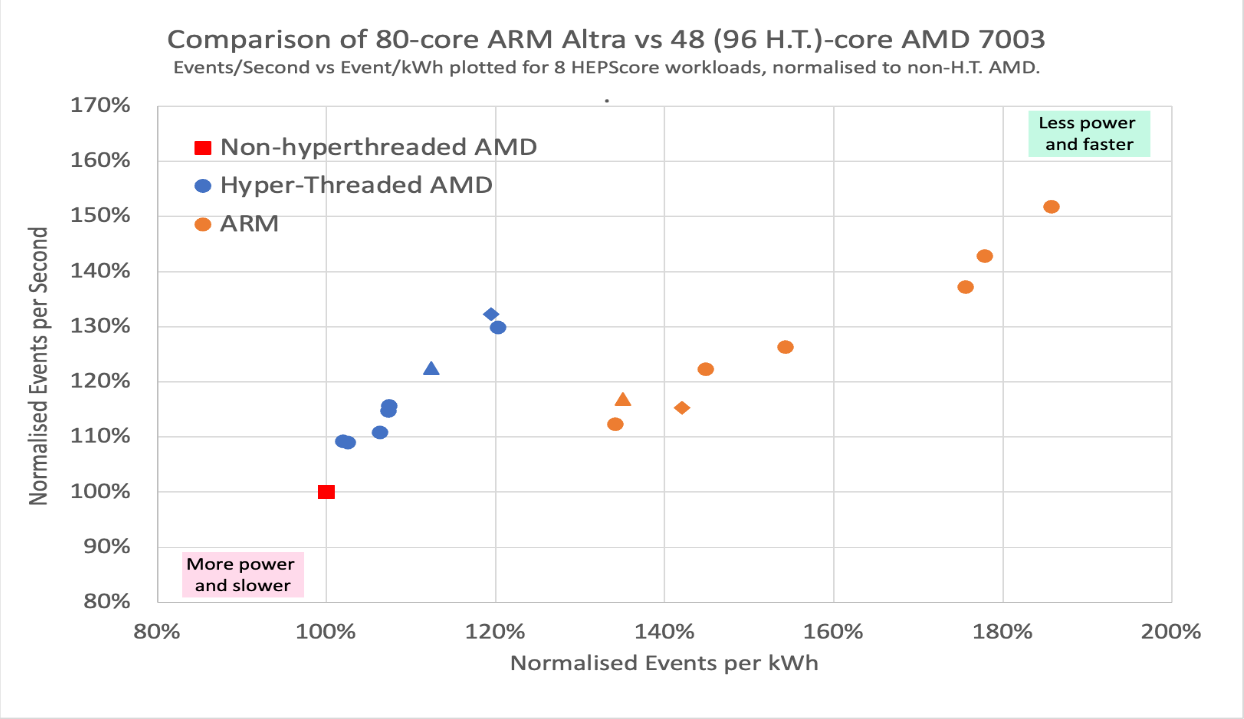
arm64: Single socket Ampere Altra Processor
- CPU: ARM Q80-30 80 core 210W TDP processor
- RAM: 256GB (16 x 16GB) DDR4 3200MHz
- HDD: 3.84TB Samsung PM9A3 M.2 (2280)

Performance: Ran the available new HEPSCORE benchmarks. This means the images were compiled for the two architectures by the experiments.

[1] https://indico4.twgrid.org/event/25/contributions/1211/

ARM_64 vs AMD(x86) results

- Hyper-threading AMD makes it faster and more energy efficient but gains are workload dependent
- ARM is much more energy efficient and generally faster than Hyper-Threaded AMD, again depending on workload



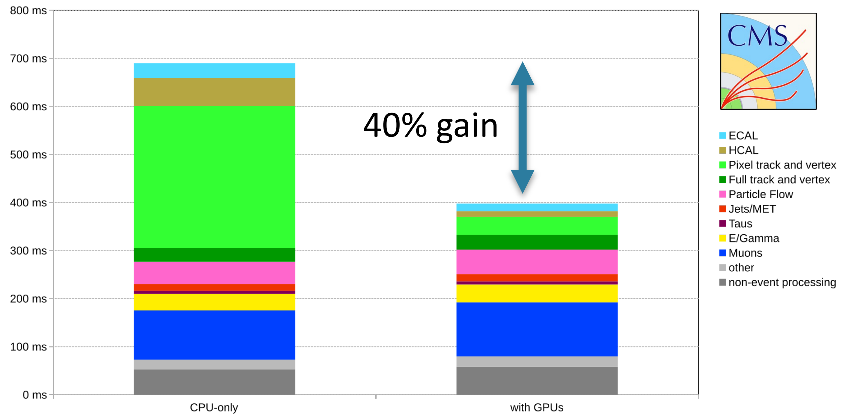
ARM looks like a potential step-changing technology. Motivates effort in porting and validating experiment software for a large number of workflows

GPUs?

CMS High Level Trigger (online): CPU + GPU (Nvidia T4, 2 per node) solution

In the HLT: 40% of event processing offloaded, +70% throughput and **+50% more events per kWh** wrt a CPU-only solution [3]

Time/Event



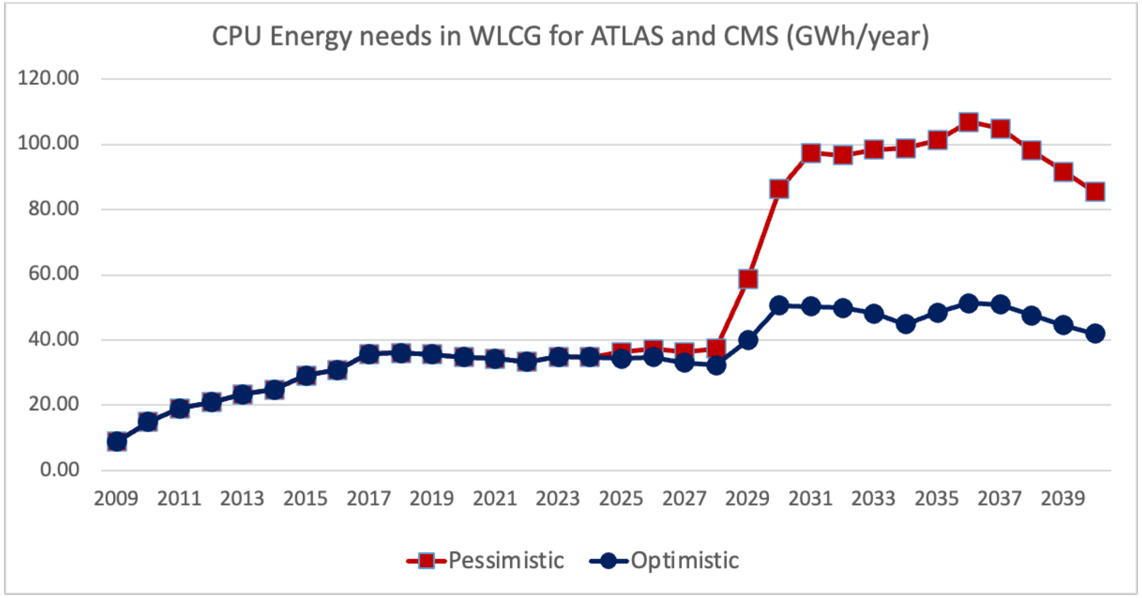
[3] https://indico.cern.ch/event/1106990/contributions/4991273/

CMS is working to profit from accelerator-ready code for offline processing already in Run 3: aim to 10% offload this year. ALICE is using CPUs+GPUs for asynchronous reconstruction in the O2 online farm

GPUs offer another opportunity to reduce the energy needs
Porting the LHC offline software to GPUs is a challenging task

ATLAS and CMS CPU power needs for HL-LHC

This is power now (GWh/year)



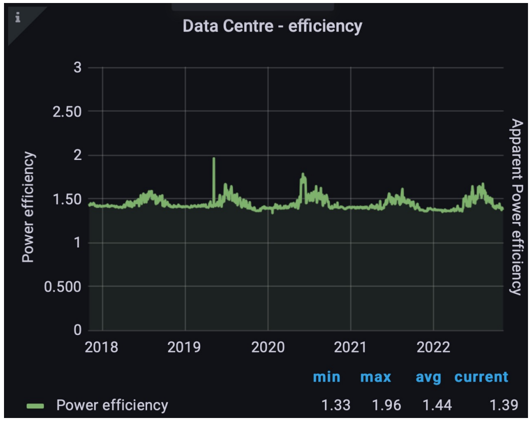
The peak of energy need happens in 2036 (start of Run-5): 400% higher than 2022 in the pessimistic scenario and 50% higher in the optimistic scenario

These estimates assume 50% improvement in hardware efficiency every 5 years, 5 years hardware lifecycle and a Power Usage Effectiveness of 1.45 for the facilities

Power Usage Effectiveness

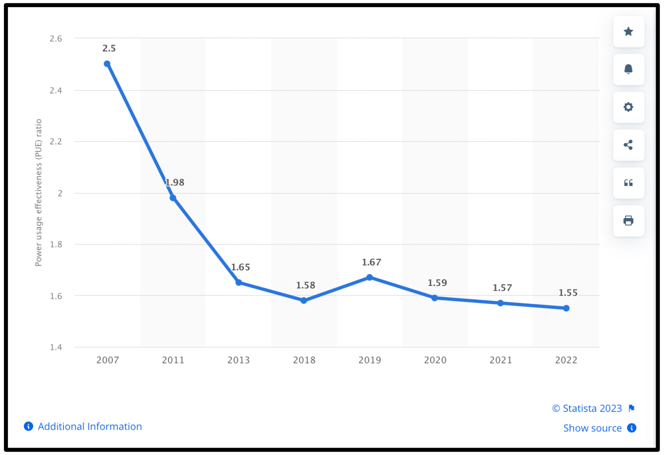
The overall power usage has to include the PUE factor:

- extra power is needed to provide the cooling of equipment



The average PUE for the CERN Meyrin (plot at the top) centre is about 1.45 over the year.

Worldwide, the average data centre PUE was 1.55 in 2022 [4] (plot at the bottom)



It is hard to quantify this value for all WLCG facilities.

- we used 1.45 for this study, constant over the years

[4] https://www.statista.com/statistics/1229367/data-center-average-annual-pue-worldwide/

Power Usage Effectiveness

The modernisation of the facilities generally reduces the PUE (the lower the better)

- modern data centres have a PUE below 1.4 [5]

[5] https://www.sunbirddcim.com/blog/whats-best-pue-ratio-data-centers

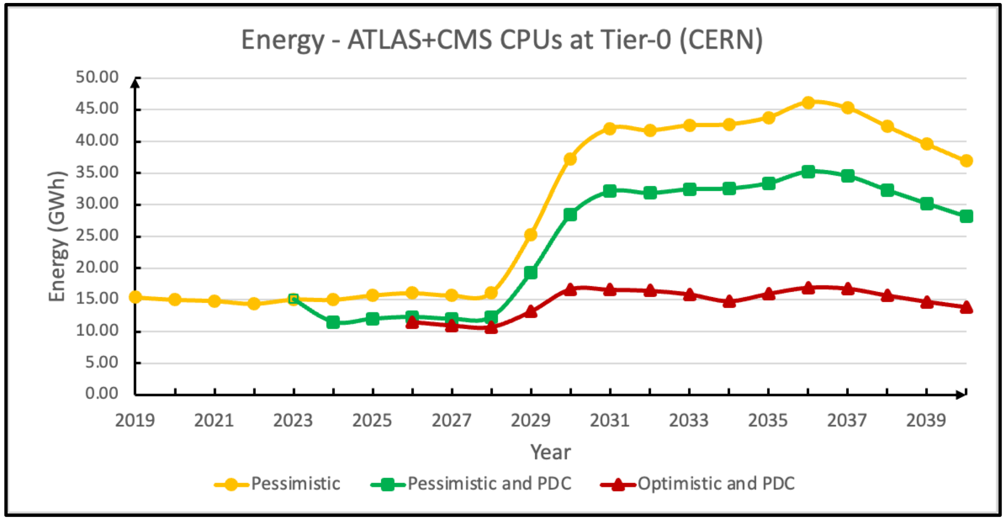
E.g. CERN is building a new data centre at its Prevessin site (ready from October 2023)

- Up to 12 MW, 4MW commissioned in 2023
- The Prevessin Data Center (PDC) PUE will be ~1.1
- Includes the possibility of heat recovery - up to 3MW out of the first 4 MW

Upgrade of WLCG facilities in various countries is a continuous process. Many plans in place

Power Usage Effectiveness

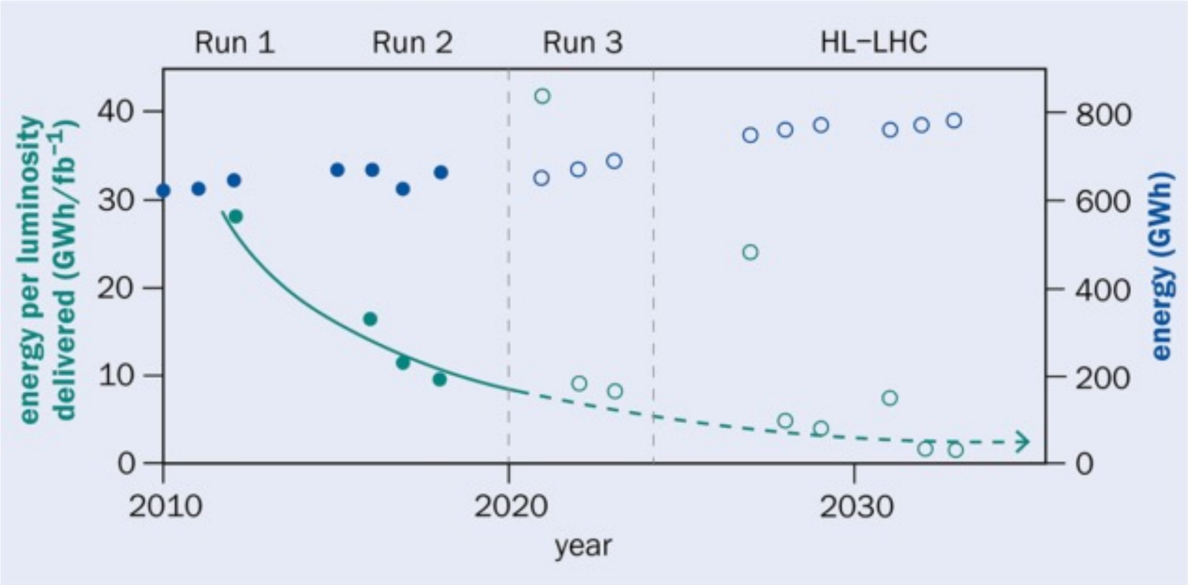
The effect of the Prevessin Data Center on the energy needs for ATLAS and CMS CPUs at **CERN** is on the right plot



The introduction of the PDC reduces by 30% the CPU energy needs

The successful completion of the R&D program reduces the needs by another 50%

Key Performance Indicator for LHC



https://hse.cern/environment-report-2019-2020/energy

The GWh/fb^{-1} metric has now been adopted by CERN as a key performance indicator (KPI) for the LHC. See this article for example

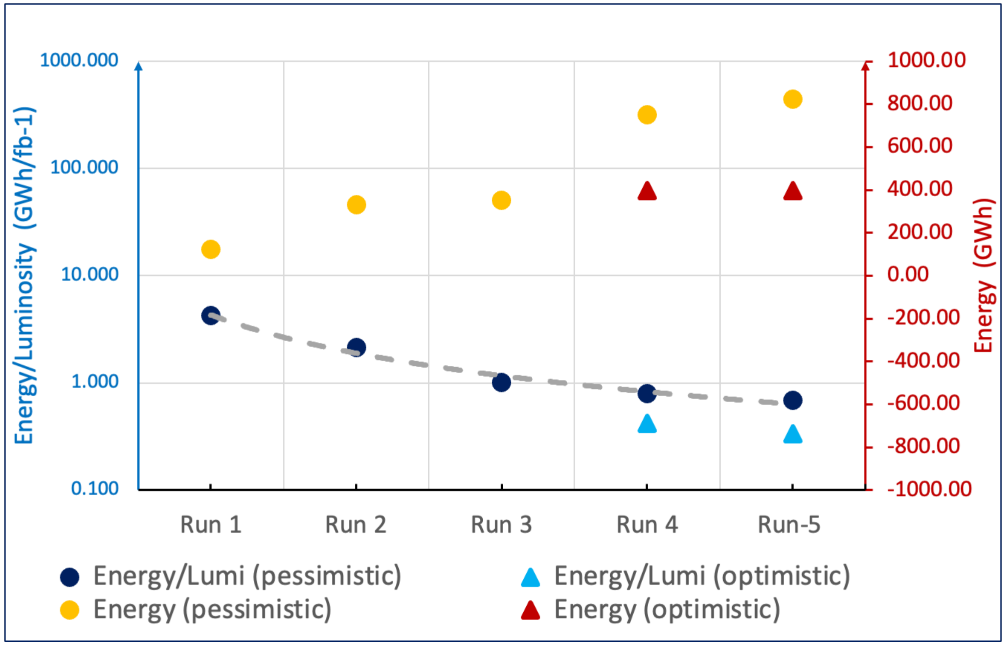
It represents the amount energy needed to **produce** a given amount of data

Looking at GWh/fb^{-1} for WLCG …

In WLCG GWh/fb^{-1} represent the energy needed to **analyse** the data

The scale on the right **(RED)** shows the energy and the scale on the left **(BLUE)** shows GWh/fb^{-1} (log!)

Energy needs in Run-4 and Run-5: +100% compared to Run-2 in the **pessimistic** scenario, only +10% in the **optimistic** scenario



GWh/fb^{-1} decreases a factor 10 between Run-1 and Run-5 (exponential trend fits well)
In Run-5, GWh/fb^{-1} in the **optimistic** scenario is half compared to the **pessimistic** scenario

Conclusions

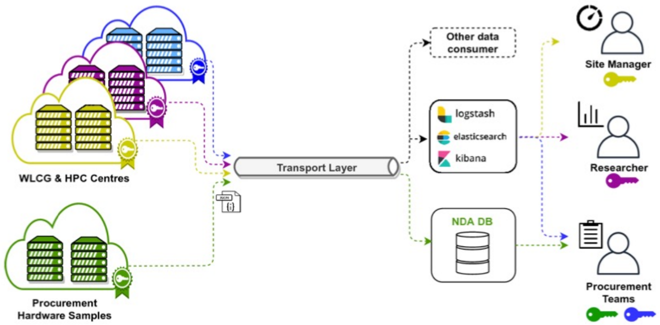
- This study shows the trends and does not pretend to make predictions
- The energy needs in HEP computing can be kept under control leveraging four pillars
  - The modernization of the facilities, going in the direction of more energy efficiency. Major capital investment
  - The improvements in the software and computing models. A gradual process bringing early benefits
  - The improvement in the hardware technologies and the optimization of the hardware lifecycle strategy. We need to invest in software portability
  - Turning off the air conditioning at CHEP
- I focused on the first three. Each pillar is important, but the improvements in software and computing models are an area where everyone in the WLCG community can contribute and where the largest gains should be expected
- In all scenarios, GWh/fb^{-1} decreases over time: more physics per kW

Backup

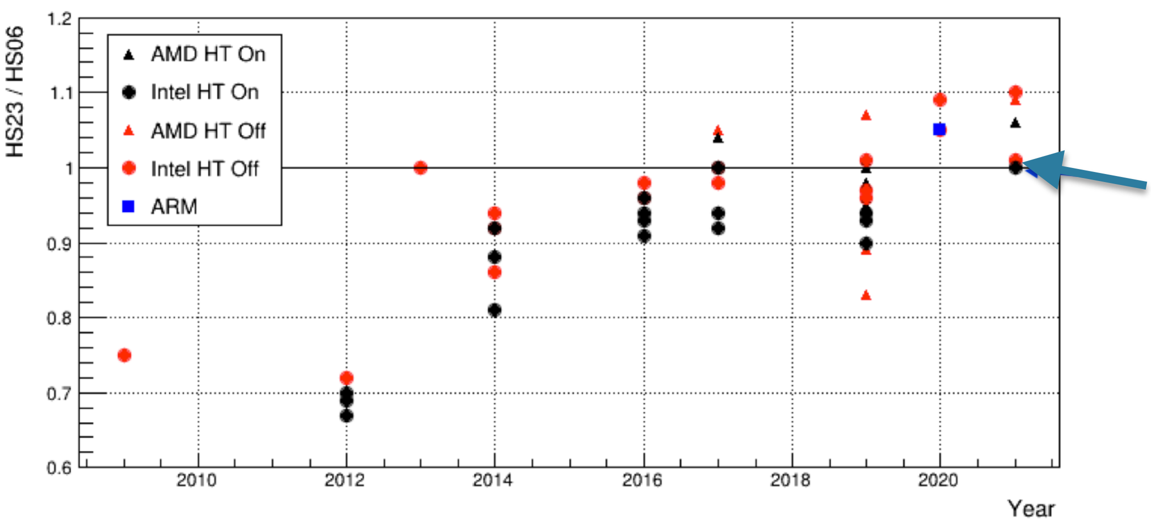
HEPSCORE23

The HEPSCORE23 (HS23) benchmark replaced HS06 from April 1st as planned

- Based on (seven) real workflows of the HEP experiments (not only LHC)
- HS23 runs on X86 and ARM (GPUs in future)



Benchmark results are centrally collected



We fixed HS23==HS06 on a modern reference machine (Gold6326@2.90GHz) for a simpler transition

See this contribution for details