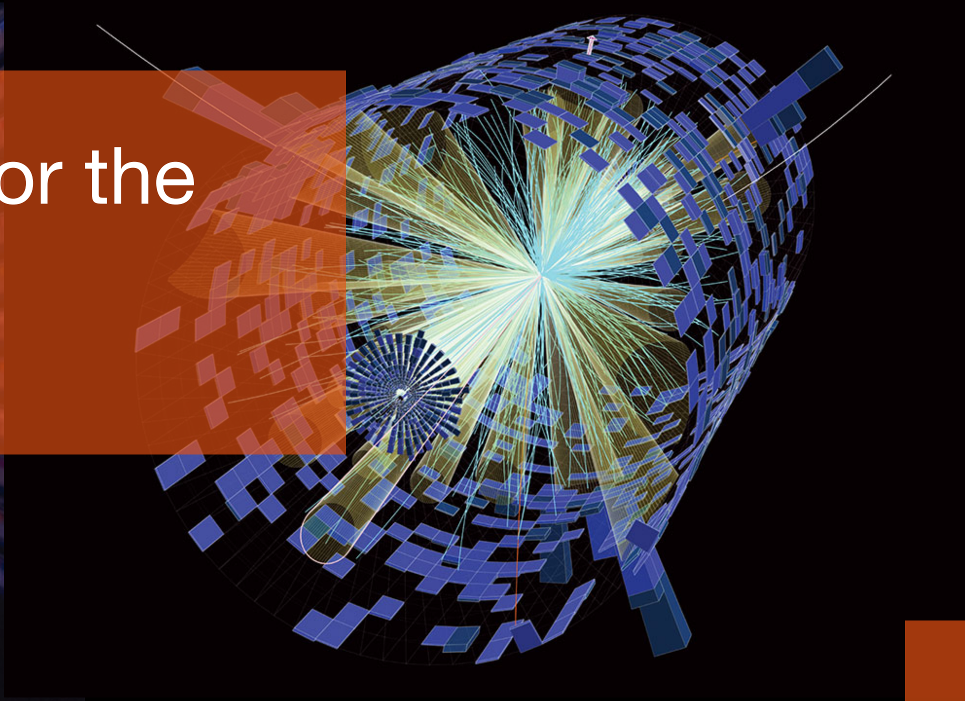
Institut national de physique nucléaire et de physique des particules

13/11/2024

Sabine Crépé-Renaudin, LPSC Grenoble CNRS/IN2P3

LPSC GRENOBLE | MODANE



Computing challenges for the future

IRN Terasscale

Introduction

Computing as a key element for HEP

- Software and computing infrastructure are crucial for HEP experiments
  - their quality have strong impact on the physics output
  - need high investments in HR and cost (same order as detectors but not the same time distribution)
- Software and computing needs also tightly linked to detector design and reconstruction and analysis technics

Some specificities to take into account

- Hardware needs to be changed every 5-7 years => needs continuous funding
- Hardware technologies evolving at high pace, software need to adapt accordingly and fast evolution of technics as well
  - see fast developments of AI and accelerated processor in the last years

Thus

- Difficult to go beyond 5 years to be reasonably predictive but one can also derive some general recommendation based on past and current experience
- Very broad subject, difficult to cover everything in a short talk
  - will start from today status and HL-LHC challenge
  - gives some key challenges and opportunities
  - illustrated with some recent developments
  - some derived recommendations

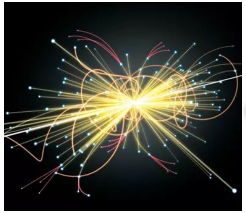
HEP computing
From where we start : key ingredients

Typical Computing in HEP

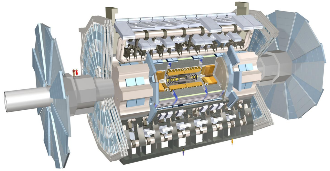
Typical for LHC run 1-2

Experimental Data

40.10^6 collisions/s



150.10^6 channels
1,6 MB/collision

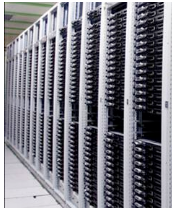


40 MHz
100 Tb/s

Niveau 1, hardware
~ 2.5 µs

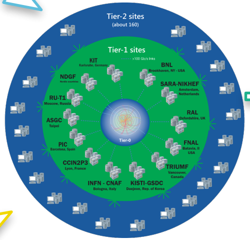
100 kHz
50 GB/s

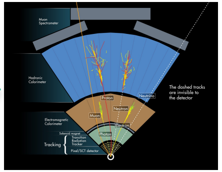
Niveau 2, software
(60k coeurs CPU)
~ 500 ms



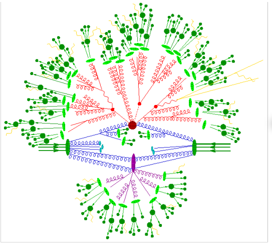
Triggers

1 kHz
1,6 GB/s



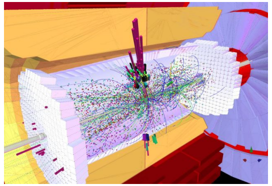


Simulated Data



Generation

Theory



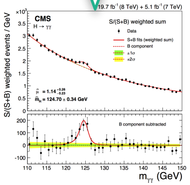
Simulation

Software

Digitisation
Trigger selection

Reconstruction
Analyse
Storage

Software
Middle/interware
Databases



Analyse

Main components

Hardware

- Trigger
- WLCG grid
    - grid sites : CPU, storage disks and tapes
    - network

Middleware, interware and databases

- Data management and distribution: FTS, xrootd, webdav, DDM, Rucio, DIRAC, …
- Computing task management : Panda, DIRAC, …
- Software distribution: CVMFS
- Databases: conditions database, detectors, softwares, datasets, sites etc
- Monitoring : sites, computing tasks, storages, transfers, networks
- Communication, tickets systems: GGUS, JIRA

Softwares

- Trigger
- MC generation
- Simulation
- Reconstruction
- Analyse
- Monitoring

➜ 10s millions of code lines

Skilled people

- for all the components
- technicians, engineers and physicists

dCache

GEANT4 A SIMULATION TOOLKIT

ROOT Data Analysis Framework

XRootD

CTA (Tape Archive)

SHERPA

EOS (File System)

H7

CVMFS (Software Distribution)

acts

CRIC (Information System)

ROOT Data Analysis Framework

FTS (File Transfer System)

RUCIO

DIRAC (Workload Management)

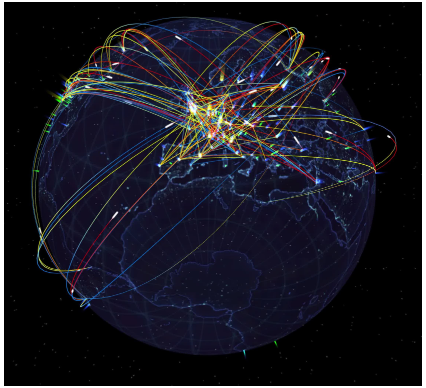
dask

IAM

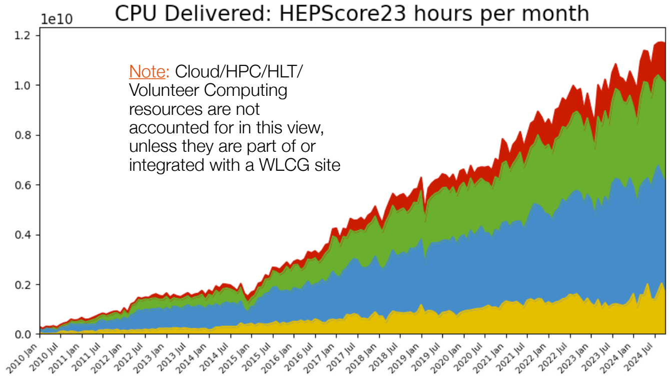
Scikit HEP

… …

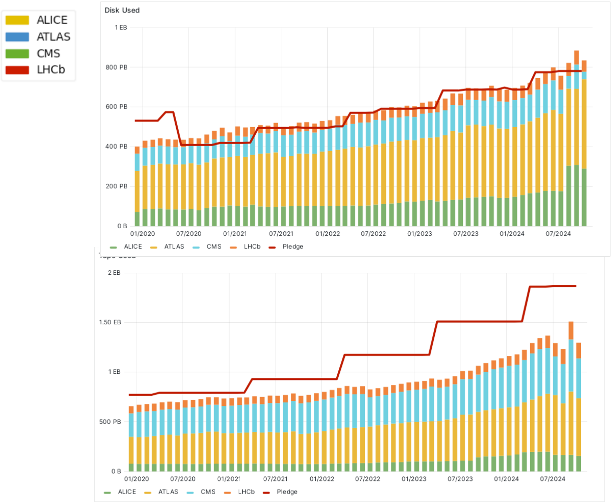
WLCG



Total: ~1,5 million CPU cores



Data Total: 1,9 ExaByte of data (disk + tape)



WLCG today

- 159 sites in 40 countries
  - Tier 0 at CERN, 14 Tier 1, Tier 2, Tier 3
- ➜ much less hierarchical than at the beginning,
- ➜ much more dynamic and automatised

Dedicated networks

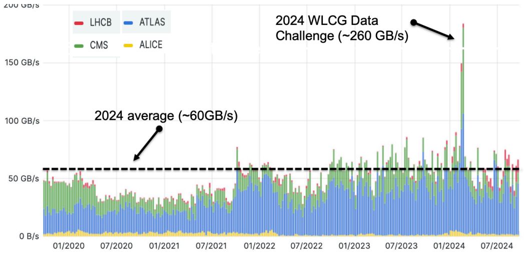
LHCOPN : 13 countries, 3 continents Tier 1 and Tier 0

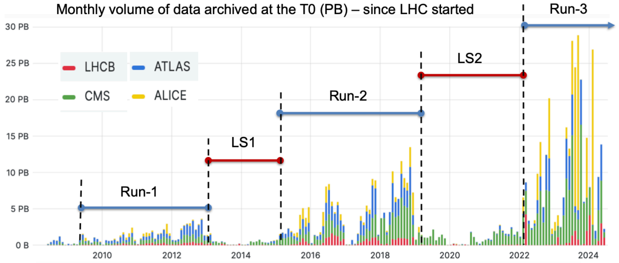
- 2.1 Tb/s to Tier 0

LHCONE :31 network providers 117 interconnected sites

- 10-400Gb/s

Average Transfert Rate 2024: 60 GB/s





Overall

A notable success

- Needs of the LHC experiments successfully met by the construction of a distributed Exascale federation with resources, services, software and dedicated teams worldwide

Within a flat budget

Collaborations strong enough to go through major crises:

- Covid pandemic: WLCG ran smoothly during the pandemic period
- Ukraine war: Russian ressources mostly compensated, Ukrainian sites back online
- and consequences:
  - cope with delay in delivery => Ex: 1 year delay to get network components, other hardware ~x2 wrt to pre-COVID times
  - energy cost increase

New developments that take advantage of technological progress

- automation of grid operations, optimal use of network and storage (no more hierarchy between sites)
- software adapted to use multicore/multithreaded computing
- new real time analysis trigger using GPUs (LHCb, ALICE)

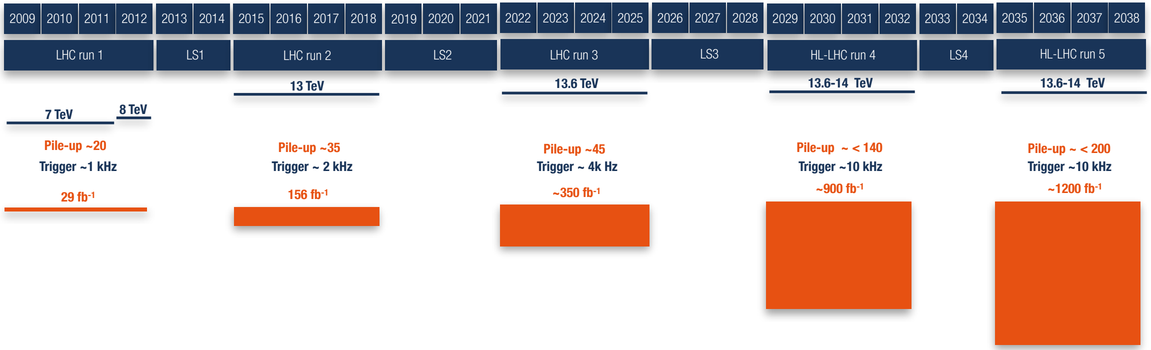
What are the HL-LHC needs ?

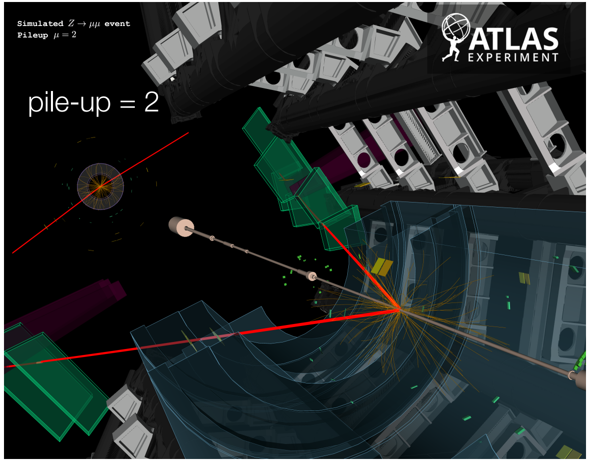
HL-LHC computing challenges (I)

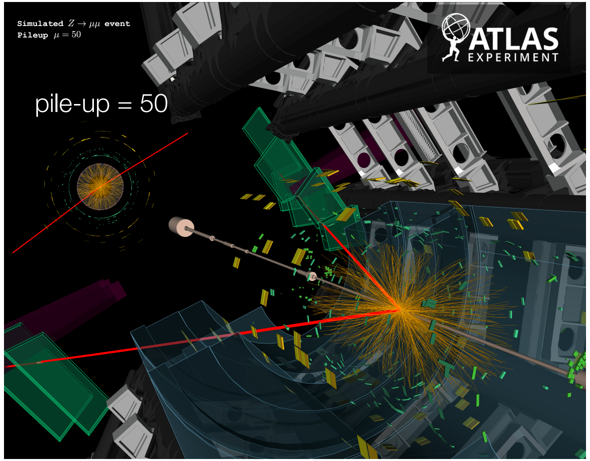
HL-LHC data

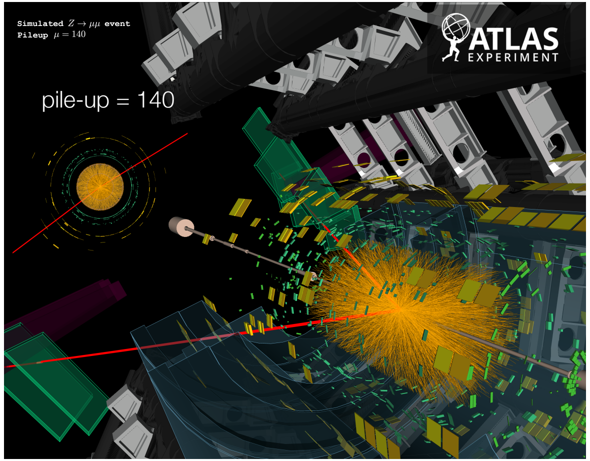
- integrated luminosity = 20 x run 2
- average pile-up ~200 = 6 x run 2

➜ more data and more complex data









HL-LHC computing challenges (II)

HL-LHC throughput

- Data throughput from detector back-ends:

  LHC: 100 GB/s => HL-LHC 1-15 TB/s
- Typical LHC "live-time": 5Ms/year

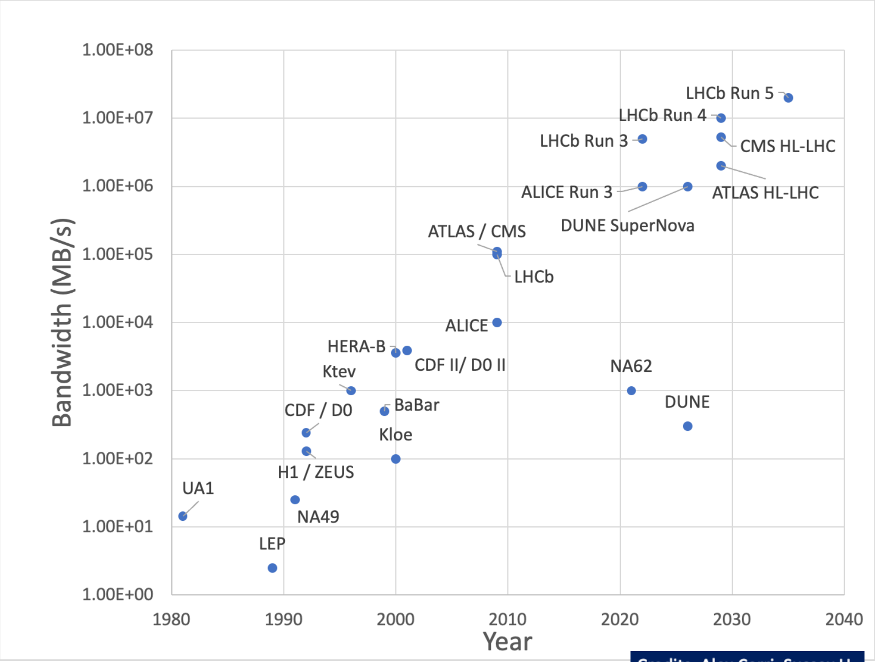
➜ Data volumes: 5-50 EB/year

➜ Triggers will have to significantly reduce data volumes

➜ Expected data throughput from CERN to Tiers1 = 4.8Tb/s !

Also an analysis challenge

	LHC run1-2	HL-LHC run
analysis dataset	10 TB	1000 TB
analysis resource	laptop	analysis facility



Credits: Alex Cerri, Sussex U.

HL-LHC projections

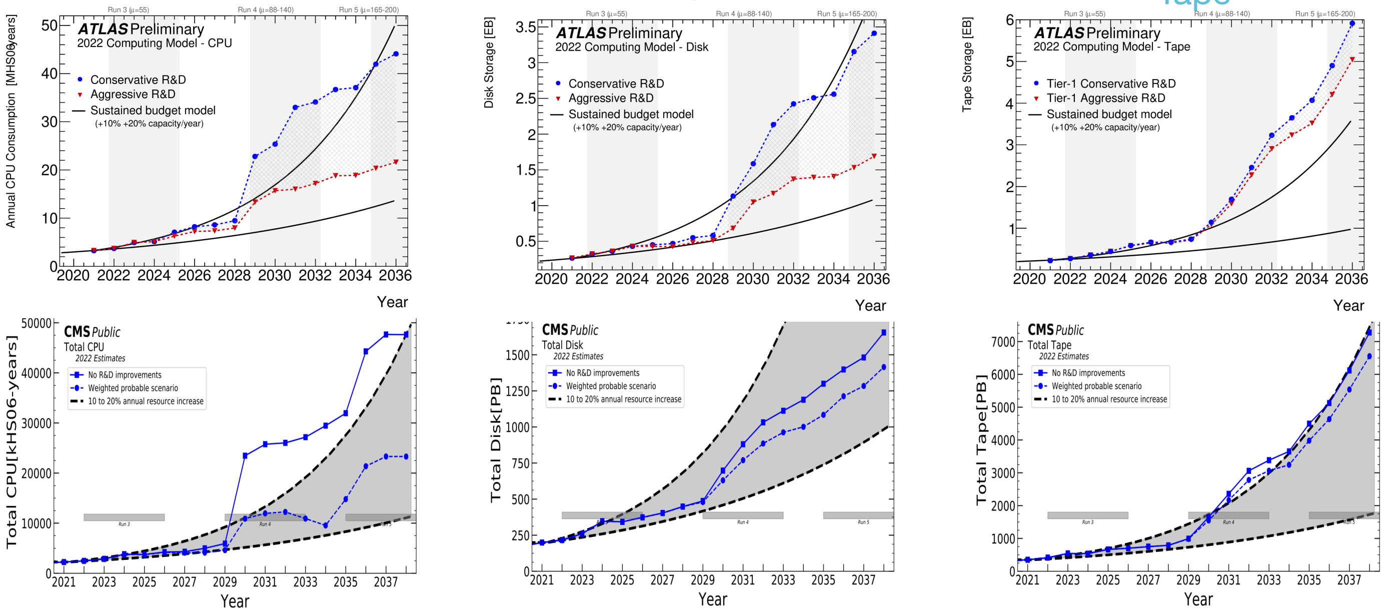
First projection in 2015 ➜ *Needs = 10x what the flat budget allows !!*

Latest evaluation in 2022 after a lot of work ➜ *Flat budget allows to cover ~ the needs but only in the most optimistic scenario*

CPU | Disk | Tape

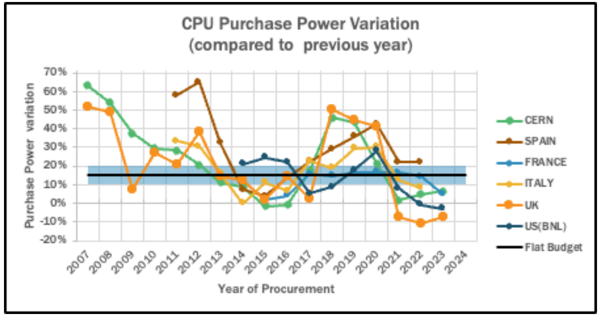
ATLAS

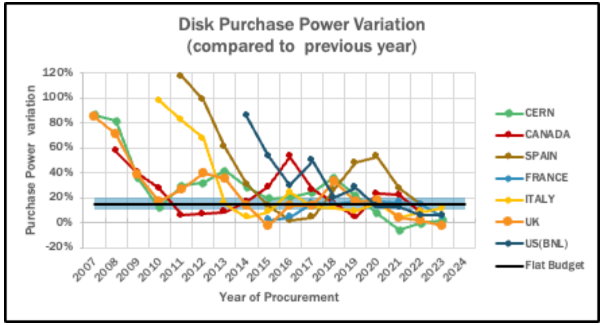
CMS



A moving context: challenges and opportunities

Hardware costs





Increase of WLCG ressources within flat budget

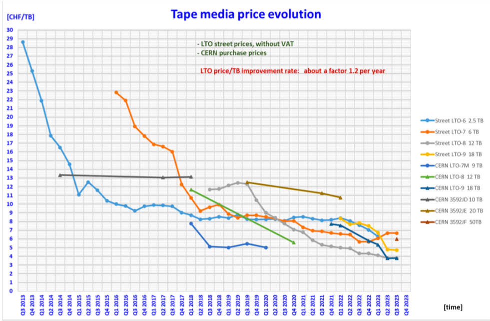
- WLCG model relies on decrease of CPU, disk and tape cost => need ~15% increase of ressource every year for same budget to fulfil experiments needs
- Works well since beginning of LHC (one exception in 2017 with the outstanding performance of LHC)
- Observed average ressource increase in last 5 years (several countries report the evaluation)
  - CPU: +14%/year, disk: +15%/year, tape media: +20%/year

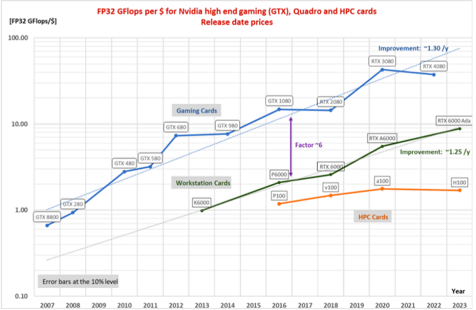
Start to follow up GPU

- Trend in GigaFLOPS/USD is favourable for video games but flattened for HPC cards price
  - Very volatile markets, long procurement times (52+ weeks of delivery time). High demand worldwide (AI/ML/ChatGPT)

Concerns for the next years

- Much more fluctuations of the market => less predictable prices
- Increase of demand and crises tend to increase prices (in 2024 CPU price increased wrt 2023) and hardware time delivery (x2 these last years)





Hardware evolution

Computing architecture evolution

- Moore's Law: transistor density still doubles every two years
- Clock speed stalls since 2000 too much power used
- Dennard's scaling: power used by silicon device is independent on the number of transistor but proportional to the transistor area

New processors

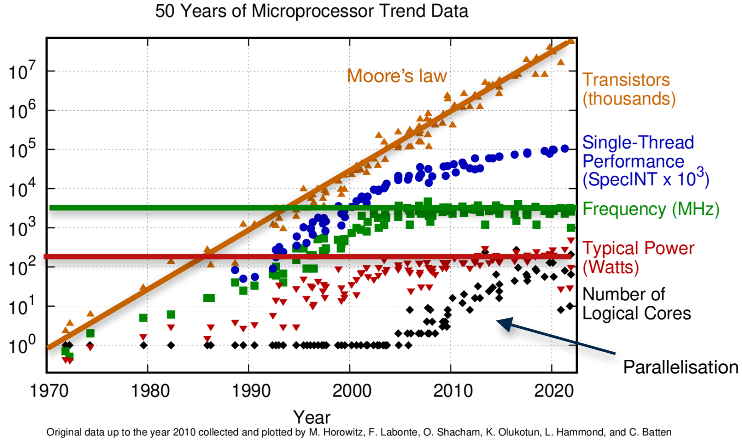
- GPU: multi-core servers with co-processors and complex memory configuration
    - in 2023 70% of Top500 machine power is from accelerators
    - power consumption controlled
    - multiple competing infrastructure and different programming langage
- non-x86 CPU architectures share increasing (AMD, ARM)
    - more energy efficient than x-86 CPU
    - HPC FUGAKU is based on ARM

Consequences

➜ evolution towards more parallelism

➜ evolution towards heterogeneous system

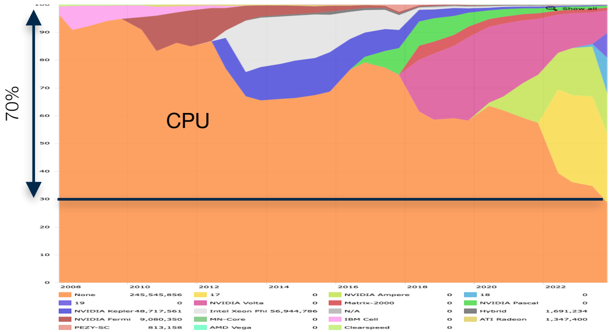
Next revolution: Quantum Computing



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2021 by K. Rupp

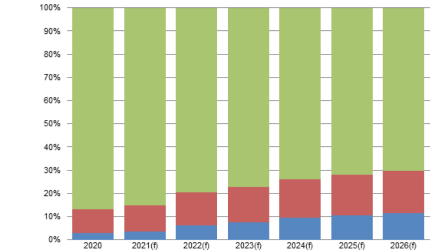
https://github.com/karlrupp/microprocessor-trend-data?tab=readme-ov-file

Accelerator performance share in Top 500



https://www.top500.org/statistics/overtime/

CPU share



	2020	2021(f)	2022(f)	2023(f)	2024(f)	2025(f)	2026(f)
Intel	86.9%	85.3%	79.5%	77.4%	74.1%	72.0%	70.3%
AMD	10.1%	11.2%	14.3%	15.2%	16.3%	17.4%	18.2%
Arm and others	3.0%	3.5%	6.2%	7.4%	9.6%	10.6%	11.5%

https://www.digitimes.com/news/a20211007GS400.html&chid=2

Heterogenous resources

HPC

- Huge investments of countries in HPC machines, entering exascale area
- Challenges
  - very heterogeneous in hardware and policies
  - mostly GPU now
  - not generally suited for data-intensive processing
    - also a network issue !
  - security policies

Commercial clouds

- Clouds flexible, large ressources available
- Challenges:
  - interfaces
  - networking
  - procurement, economic model and vendor locking
- Cost effectiveness ? potentially interesting for special tasks or peak needs

Different hardware

- CPU, GPU, FPGA, ASIC
- Vendors: Intel, AMD, ARM, Power, NVIDIA
- and different programming libraries !

➜ Need portable code

- Portability libraries with abstraction layer to hide the backend implementation and use their parallelism efficiently

➜ Open new possibilities for edge/online data processing and reduction (also with AI)

➜ What is the good balance between

- Investing in new possibly opportunistic resources like HPC not completely adapted to our needs
- Investing in our own infrastructure

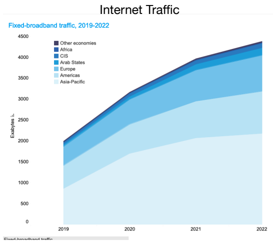
Growing computing needs

In our research fields

- Several experiments have or will have non negligible computing needs in the near future
- Also in astro-particle and cosmology: Vera Rubin/LSST, CTA, KM3NET, Euclid, ET and SKA and multi-messanger approach
- Nuclear physics: dynamic structuration ongoing, EuroLabs, FAIR, EIC, Lattice QCD…

In other research fields and in society

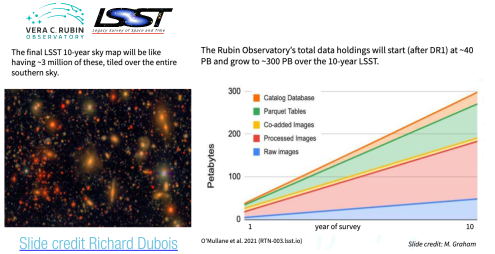
- Importance of data growing everywhere: improved technologies easier to use and computing ressources more easily available, developments of AI etc

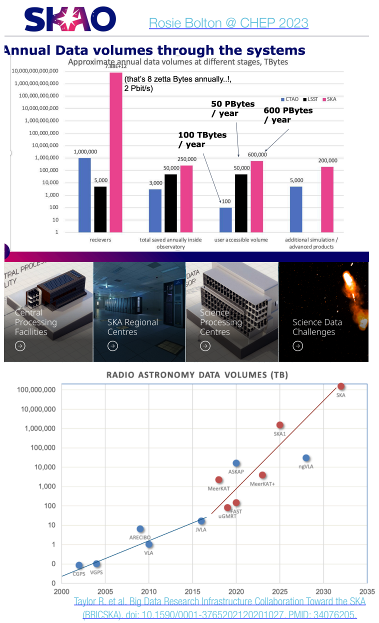


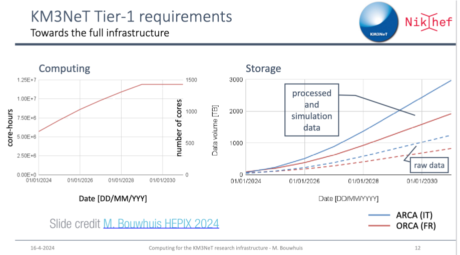
https://www.itu.int/itu-d/reports/statistics/2023/10/10/ff23-internet-traffic/

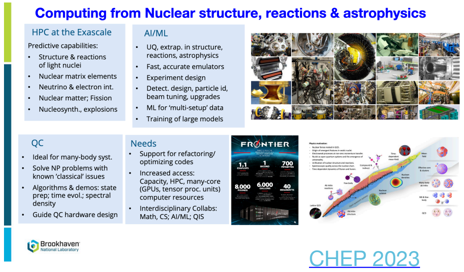
Taking into account Open Science movement

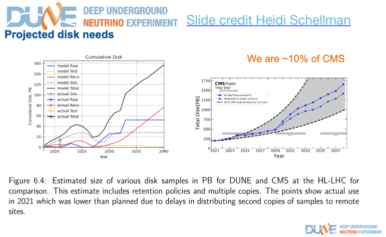
- Data preservation and sharing
- Open source software and FAIR (findable, accessible, interoperable, reproducible) data
- Important investments for instance in Europe with the construction of the EOSC (European Open Science Cloud











Artificial Intelligence

Use of AI since decades in HEP

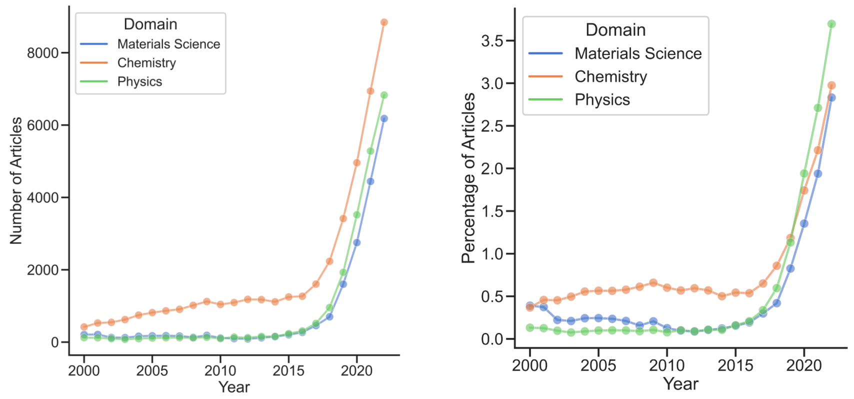
- signal/background separation, particle identification, multivariate analysis… ex BDT widely used since 90’

Modern AI and computing ressources allow new usage

- graph NN, generative models, unsupervised classification, low latency inference, LLM, foundation models…
- CPU, GPU, FPGA implementation
- AI assisted code generation

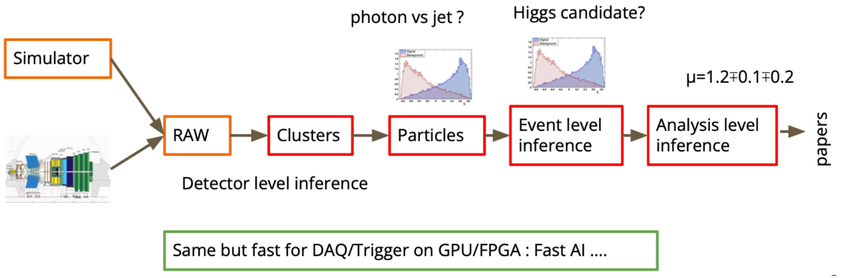
➜ AI usage and developments at all stages of computing in our fields

ML Publications in Science



Ben Blaiszik, “2021 AI/ML Publication Statistics and Charts”. Zenodo, Sep. 07, 2022. doi: 10.5281/zenodo.7057437.

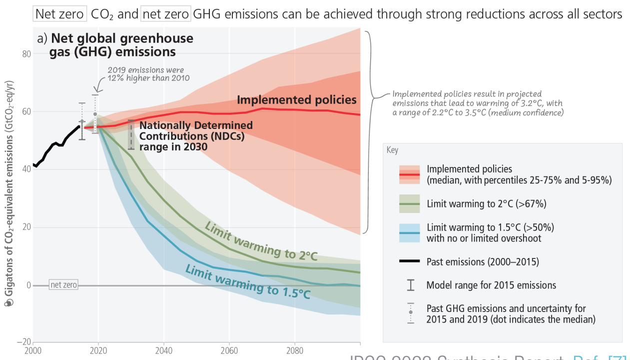
CHEP 2023 3



© David Rousseau

Sustainability

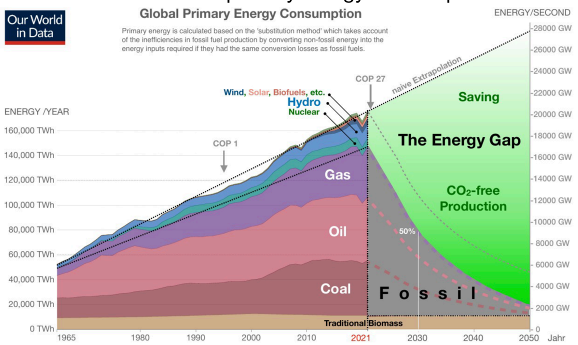
Evolution of Greenhouse Gaz Emissions



IPCC 2023 Synthesis Report, Ref. [7]

The energy gap

Evolution primary energy consumption



https://ourworldindata.org/grapher/electricity-prod-source-stacked

The contribution of data-centers to greenhouse-gas emission is sizeable and growing, no exception in our fields of research

Paris agreement

- Zero emission around 2100 => -50% by 2030 !
- To get the 50% in 2030 you have equivalently
  - to expand CO2 free energies by a factor 12
  - to increase energy efficiency by a factor 2
  - to save energy by a factor 2
- will have a mix of these

➜ What translation in our computing models and infrastructure ?

➜ See for instance HECAP+ initiative

The way to HL-LHC and beyond

Development plan

Infrastructure

- new developments while keeping the infrastructure in production
  - ➜ scale-up tests only possible on top of the production grid
  - ➜ Data Challenge to test and push the infrastructure closer to HL-LHC request
- continuous developments since run 2 => validated new tools integrated straight away
  - ➜ some work for HL-LHC already in use in run 3

➜ see also WLCG strategy 2024-2027

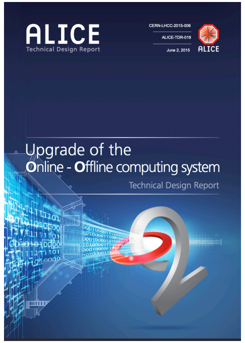
Software as a key element

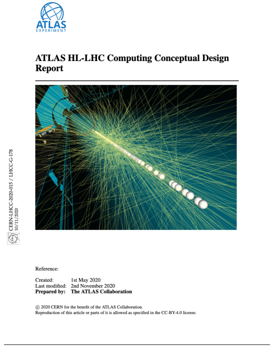
- is one of our best lever arm

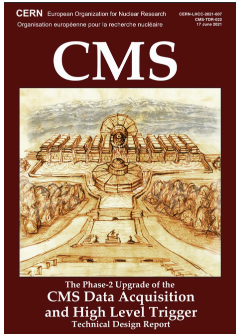
➜ need to tackle all elements: trigger, generation, simulation, reconstruction, analysis

➜ make the best use of new technology and technics

Towards HL-LHC computing models







A collaborative effort beyond HL-LHC

Enlarge collaborations

- **Common developments** allow to share expertise, to share effort and find common solution when possible
  - ➜ ease deployments and allow shared ressources and infrastructure
- Eased by international collaborations and programs
  - HEP Software Foundation (HSF), IRIS-HEP for software developments
  - WLCG/DOMA for WLCG infrastructure expands beyond LHC with DUNE, Belle-2, JUNO and VIRGO as WLCG observers
  - European programs for the development of the European Open Science Cloud (EOSC) => ESCAPE project for HEP, astronomy and nuclear physics, EVERSE project for software, GreenDigit etc
  - JENA computing workshop to discuss synergies across the 3 communities
    - 5 working groups created to draft a report by the end of the year to be discussed at next year JENA symposium with Funding Agency
      - HPCs, SW and heterogeneous architectures, Data Management and FAIR data, ML, Training – Dissemination – Education

HSF iris hep

WLCG Worldwide LHC Computing Grid DOMA

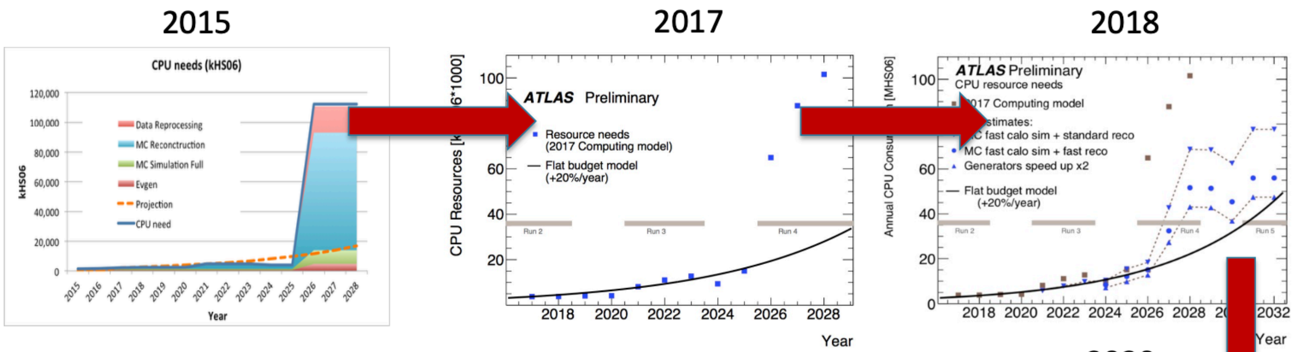
eosc ESCAPE EVERSE

JENAA Joint ECFA-NuPECC-APPEC Activities

ECFA European Committee for Future Accelerators NuPECC APPEC

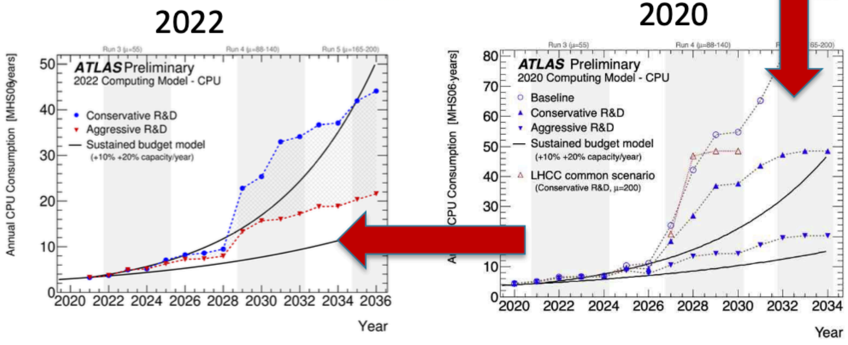
Results already visible !

HL-LHC computing resource needs evolution



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

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



Simone.Campana@cern.ch - CHEP 2023

[Simone Campana CHEP 2023]

Ongoing and new developments: few examples

Infrastructure

- data optimisation (lifetime model)
- data carousel (tape driven workflow with smart tape writting)
- datalake concept with possibility of simple caching at computing source and different quality of storage
- New AAI (token)
- Developments of analysis facility

Heterogenous ressources

- use of HPC resources
  - reached 40% of our used computing power at some period decreasing now
- Commercial Cloud usage
- Trigger and real time analysis impressive developments
  - ALICE continuous readout (timeframe 2.5-20ms) without trigger and O2 facility with FPGA, GPU and CPU
  - LHCb full software trigger: FPGAs-based clustering for Silicon Pixel detecto and HLT1 GPU based reconstruction

Sustainability

- Test of ARM/AMD processors
- adapt processor clock wrt to electricity origin
- Improve datacenter PUE

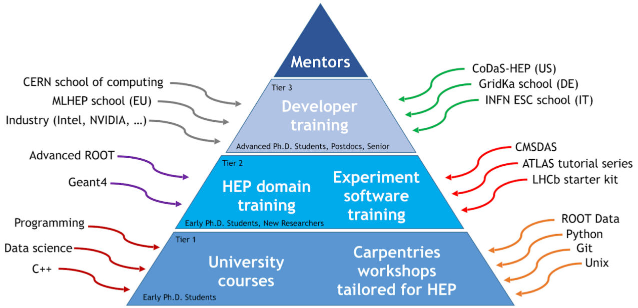
Software

- For all steps : optimise software efficiency
  - profiling
  - Move to GPUs and vectorise code when possible and effective
  - Optimise phase space sampling and integration algorithms, including AI use for generators
- development of fast simulation
- AI
  - everywhere: for simulation, reconstruction Cf ACTS and track reconstruction, for analysis
- Analysis
  - data format columnar analysis
  - use of Python ecosystem and industrial standards

Need Skilled and motivated people

Recognition

- Computing and software activities are fundamental in our research activities
  - they have huge impact on our physics results
  - their cost is similar than the one of the experiment
  - they are part of our experiments and are no more a « service »
- We need skilled and motivated people
  - Their work should be shared and publicised
  - Not only in dedicated meeting and conference !
  - Their expertise should be recognise also when physicists
  - we need to hire people with these skills



M. Ballroom CHEP 2023

Training

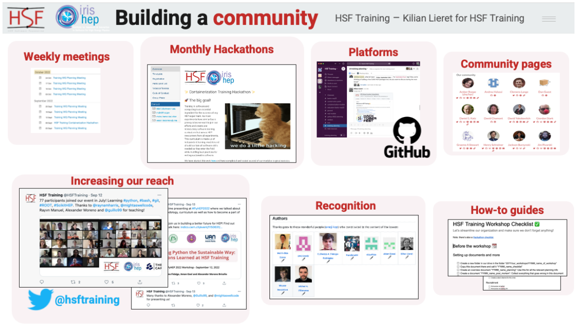
- Students often lack of software and computing skills
- Computing and software are evolving quickly => continuous learning is needed

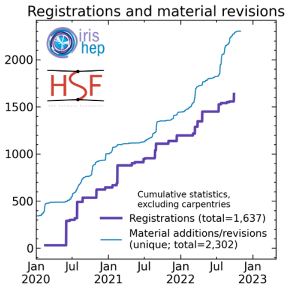
➜ tutorials, training are important and there are lot's of opportunities:

➜ tutorials, school and training locally, by their institution, in collaborations, organised by the HEP Software Foundation (HSF)…

➜ HSF as a forum to build the community and share knowledge

- Software Training in HEP





Conclusion

Conclusion

➜ *Software and Computing are key elements of our experiments*

Starting with a success

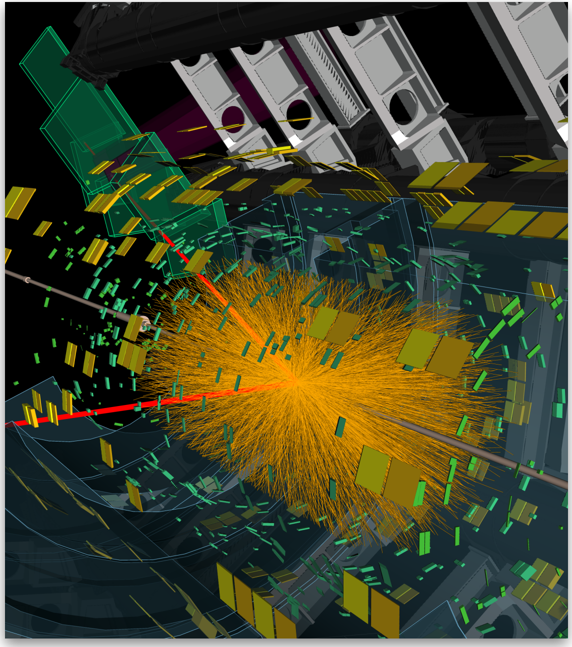
- Computing for LHC run1 and 2 with a complete and complex set of software, middleware and WLCG infrastructure at the Exa-scale successfully allows to store, process and analyse LHC data, leading to wonderful scientific results

Take up the future software and computing challenges

- Infrastructure and software will have to cope for high luminosity pile-up and throughput
- In a context that is rapidly evolving: hardware, technologies, cost ➜ constraints and opportunities
- Taking into account Open Science, data preservation and effort to reduce our carbon footprint

Huge amount of work and R&D already done and much more ahead

- New developments integrated in production as soon as validated
- Software, hardware and computing models need to be adapted for heterogenous ressources, more parallelisation => flexibility needed to adapt to ressources not build for and by us
- Make best use of new technologies and technics: AI, progress done outside our field
- Be prepared to the next (r)evolution like Quantum Computing



Will need

➜ **Coordinated efforts in the software and computing field and with detector design and reconstruction and analysis technics**

➜ **Continuous support of our infrastructure, large important code (ex Geant) and operation**

➜ **Solid continuous R&D program**

➜ **Skilled and motivated people (physicists + computing engineers) and training**

Thank you !