Computing Challenges of the Future

Simone Campana (CERN)

Acknowledgements

Before I even start, I thank people for the fruitful discussions and material in preparation for this presentation. And apologies if I forgot someone

Tim Bell, Doug Benjamin, Latchezar Betev, Tommaso Boccali, Concezion Bozzi, Ian Bird, Predrag Buncic, Federico Carminati, Marco Cattaneo, Davide Costanzo, Xavier Espinal, Joao Fernandes, Maria Girone, Bob Jones, Gavin Mccance, Bernd Panzer-Steindel, David Rohr, Stefan Roiser, Jamie Shiers, Daniele Spiga, Graeme Stewart, Ikuo Ueda, Sofia Vallecorsa

Of course the selection of the material and they way I present it is my responsibility

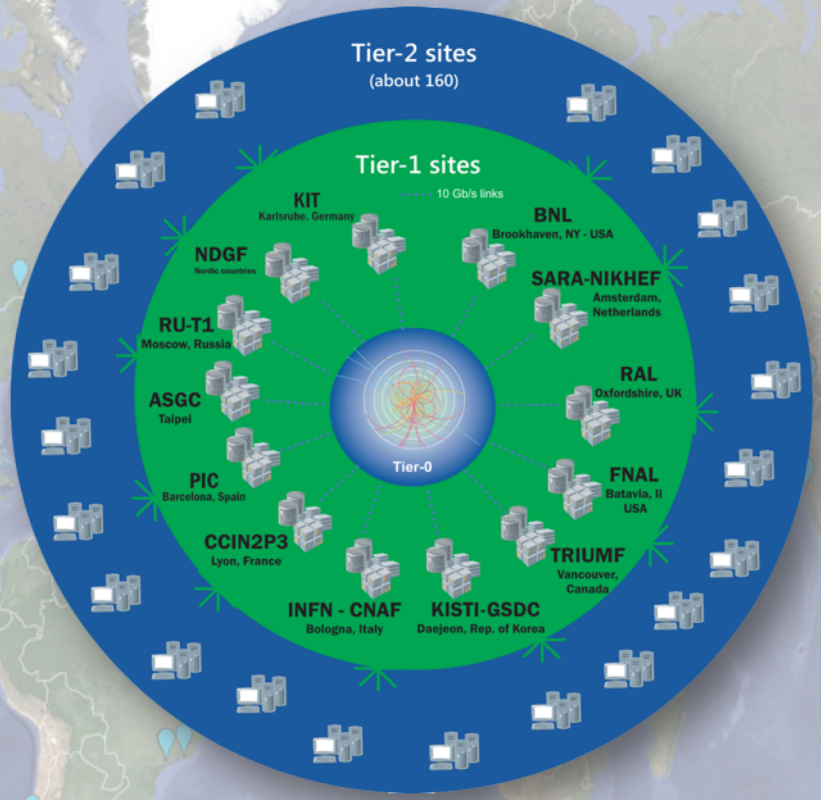
Introduction

The *HEP Software Foundation* delivered in 2018 a *Community White Paper* identifying the HEP computing challenges in the 2020s and defining a roadmaps for Software and Computing evolution

The HSF CWP is a bottom up approach. Scientific projects prioritize topics from the CWP based on their future needs and specific challenges. See for example the *WLCG Strategy Document* for HL-LHC

Those documents drive today a lot of the activity around future challenges in HEP
This presentation is an overview. Dedicated discussion on *Wednesday morning*

HEP computing embraced a large scale distributed model since early 2000s
Based on grid technologies, federating national and international grid initiatives



1 TB ~ 10-100 CHF

1 core ~ 100 CHF

HW lifetime: 3-5 years

167 sites, 42 countries

~1M CPU cores

~1 EB of storage

> 2 million jobs/day

10-100 Gb links

WLCG: an International collaboration to distribute and analyse LHC data

Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists.

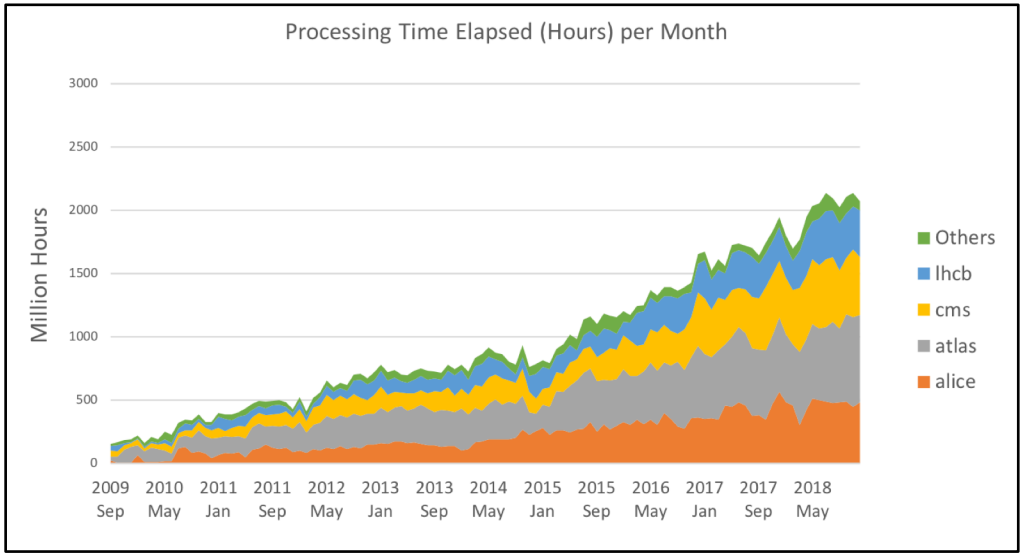
HEP Computing Today

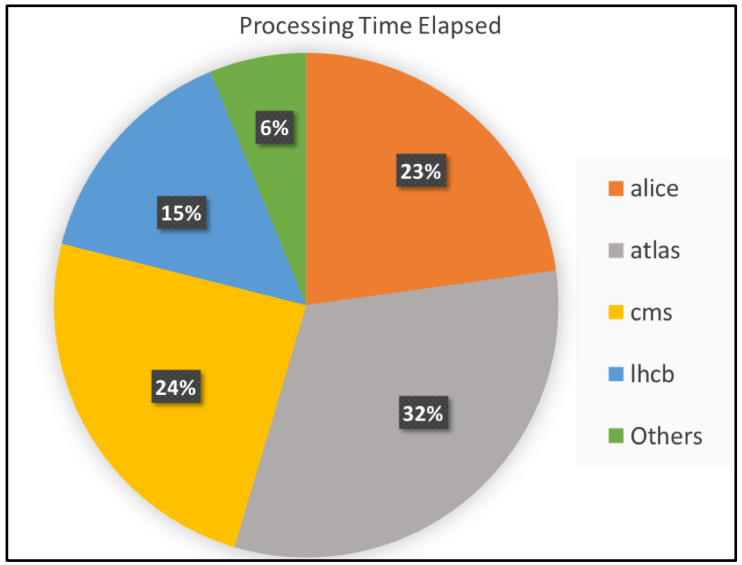
Grid resources are shared among many experiments and sciences

LHC experiments consume 95% of the accounted computing capacity

Non HEP sciences rely a lot on HPCs (not accounted here)

Data from the EGI accounting portal

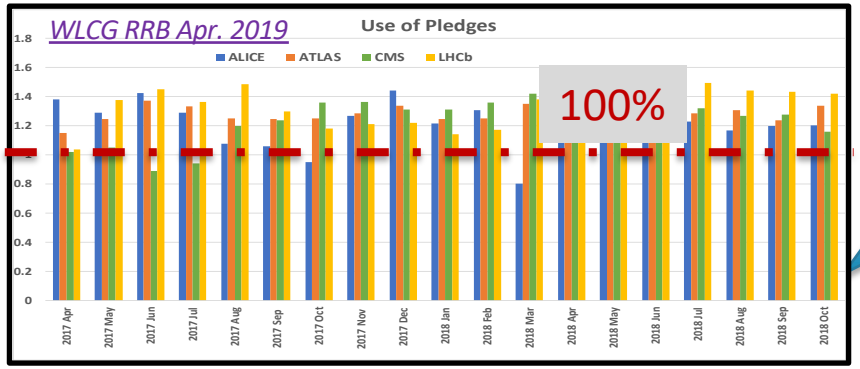




Today LHC is in a leading position steering the evolution of the scientific Grid infrastructure

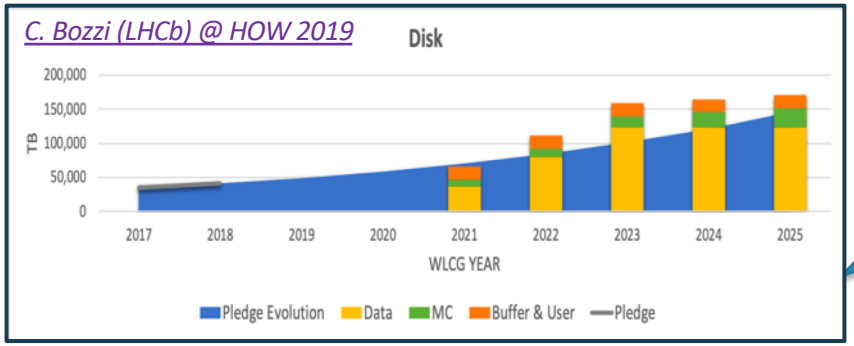
The LHC data volume challenge

2018



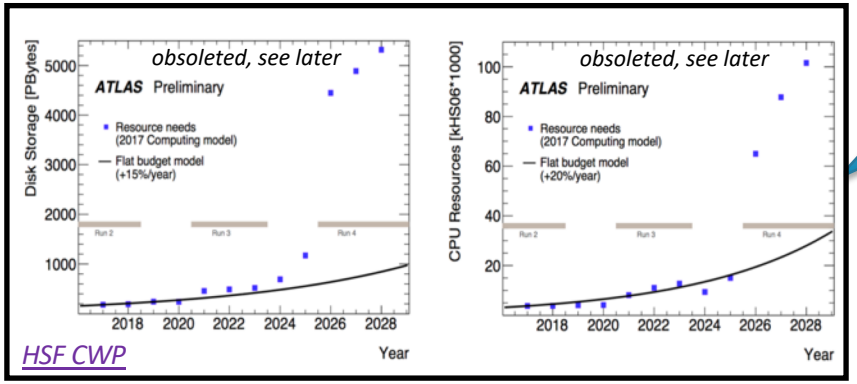
Today LHC generally gets the **requested** computing resources. Extra **opportunistic** compute capacity is available

Run-3



ALICE and LHCb will increase considerably the data rates in 2022 (LHC Run-3)

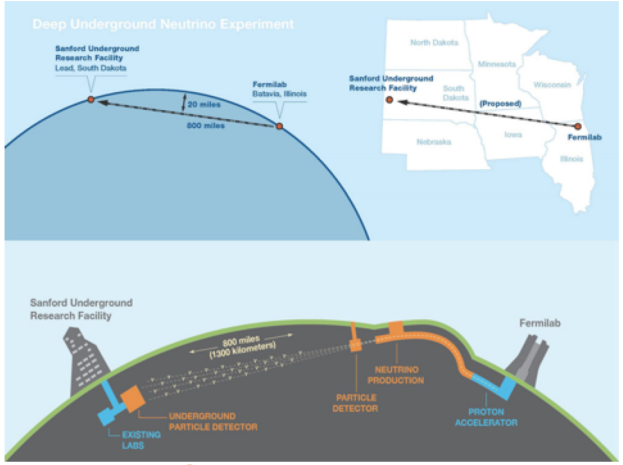
HL-LHC



ATLAS and CMS will increase the event rates by a factor 10 and the event complexity in Run-4 (HL-LHC)

No expected increase of funding for computing: “flat budget”

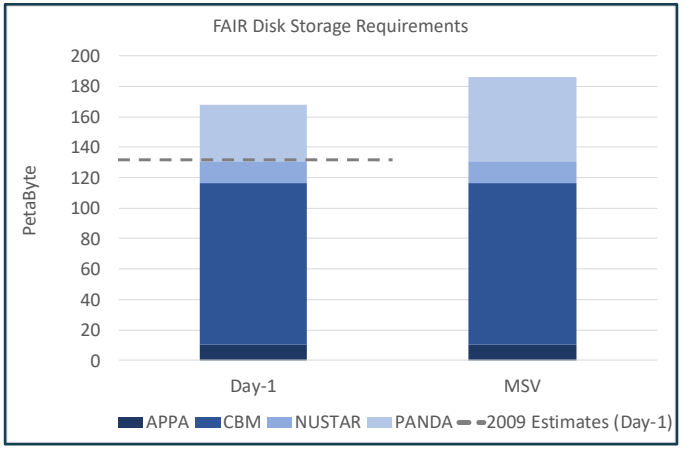
The HE(N)P data volume challenge



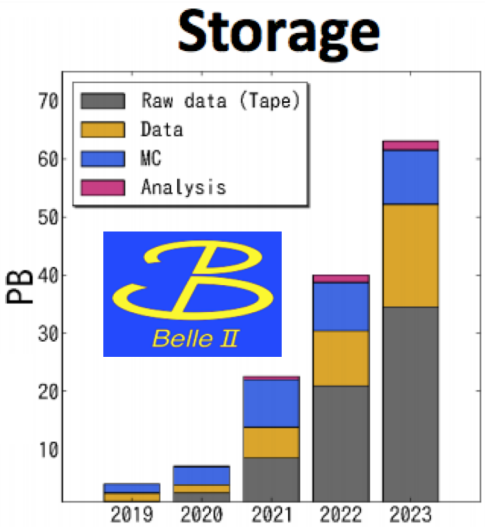
DUNE foresees to produce ~70PB/year in the mid 2020s

Several experiments will require relatively large amount of compute and storage resources. Several factors less than HL-LHC

Comparable data volume to LHC



J. Eschke @ ESCAPE kick-off



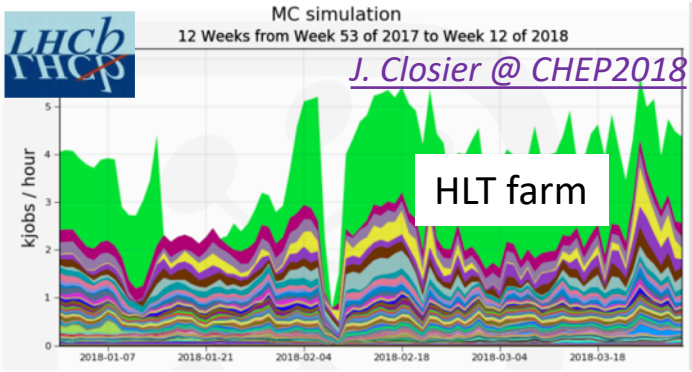
Y. Kato @ HOW 2019

Heterogeneous Facilities

Up to 25% of resources currently used by LHC experiments comes from non-Grid facilities

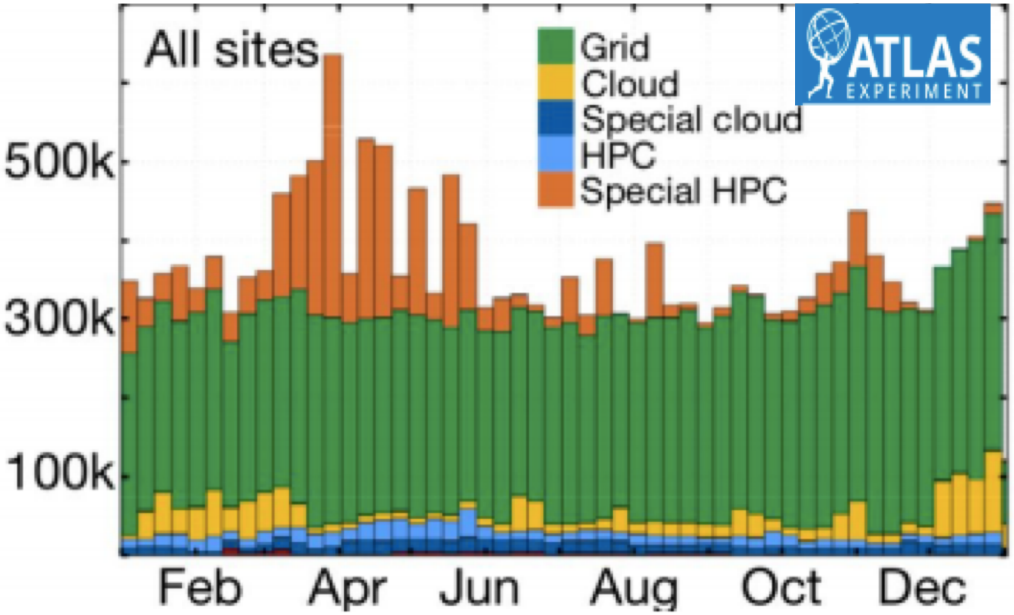
- Cloud Computing
- High Performance Computing
- HLT farms

Today contributing to the "opportunistic" pool of available capacity



J. Closier @ CHEP2018

Avg # running cores per resource type (2018)



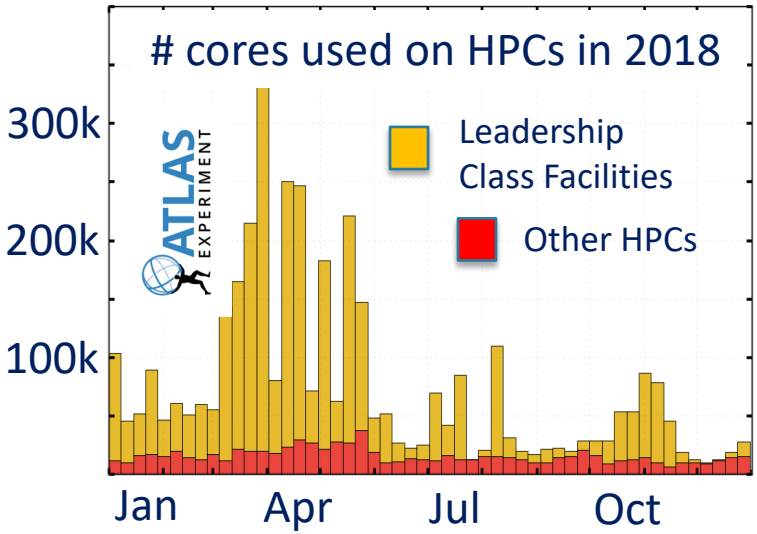
D. Costanzo @ HOW 2019

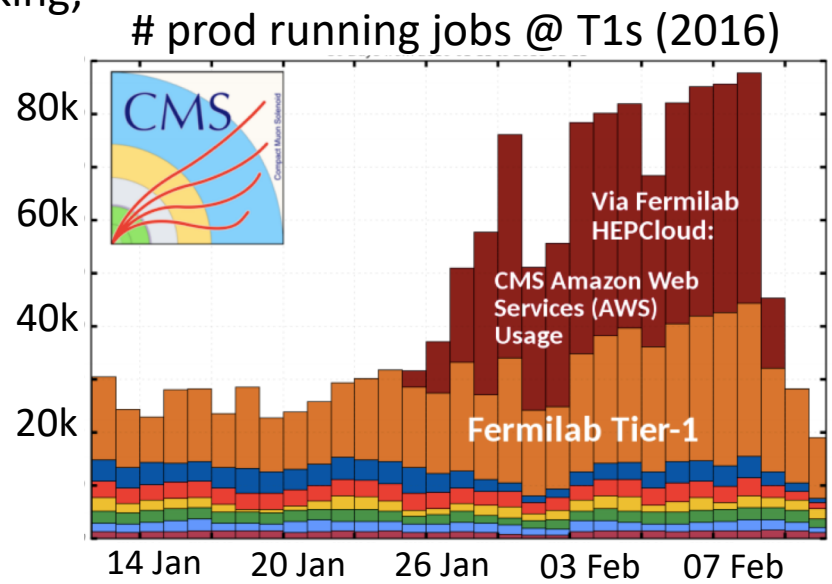
Efficient usage of those resources will be a key challenge in the future

Clouds and HPCs

High Performance Computing centers**:** very heterogeneous in hardware and policies. Large pools of resources and growing. **Strengthen the link between HEP and HPC** communities: find effective common and scalable solutions to the challenges of **data intensive processing**

Commercial Clouds: virtually infinite resources and flexibility, cost effectiveness still to be demonstrated**.** Challenges: interfaces, vendor locking, networking, procurement, economic model





Google, Amazon via HepCloud *Azure via DoDas*

Storage and Data Management

Storage: the main challenge in the 2020s …

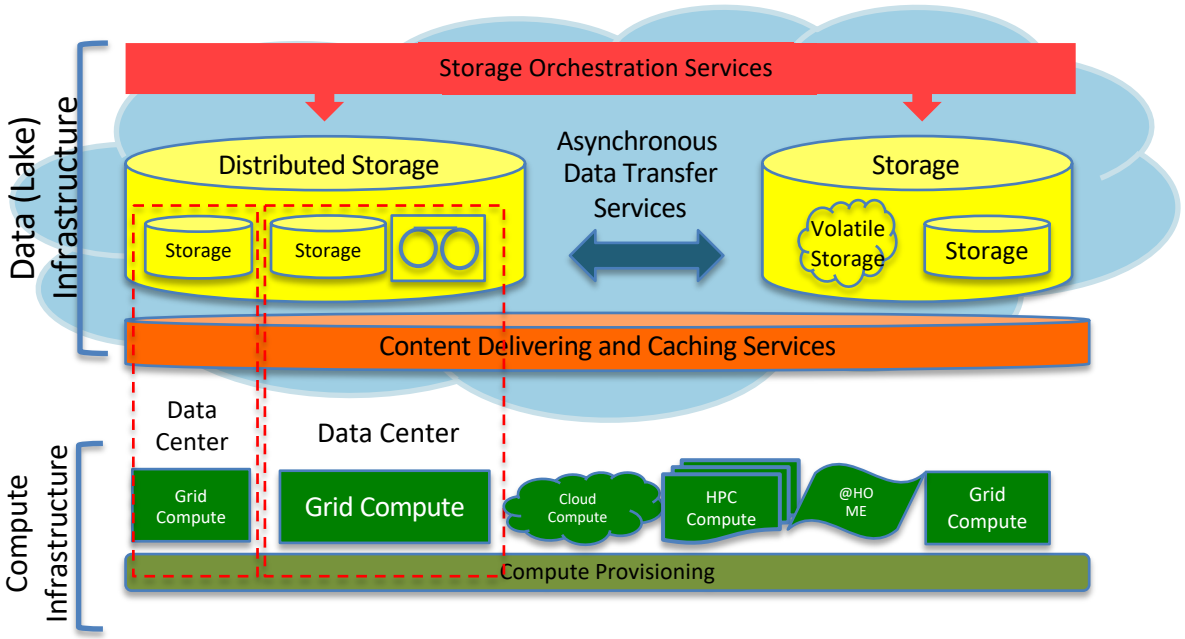
- No opportunistic storage
- Data is **THE** main asset of HEP, but storage needs are hard to fulfill
- Storage service are challenging to operate

… but also a great opportunity

We have decades of experience in developing and operating large storage solutions and managing large data volumes

DOMA projects
(Data Organization, Management, Access)

- A set of R&D activities evaluating components and techniques to build a common HEP data cloud



HEP computing services ecosystem

There is an opportunity to leverage commonality across HEP and beyond. This is happening already. Compromise between experiment specific and common solutions

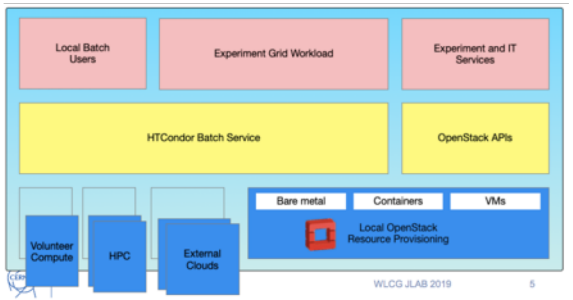
Dirac User Workshop

Rucio Community Workshop

Compromise in adopting products from open source community projects and in-house development

Openness and collaboration are the key.
Watch for sustainability!

Common access layers to HPC and Clouds



Different Sciences, Shared Infrastructure

Most of the facilities supporting HEP and other science projects are the same.
The Funding Agencies do not want to deploy several computing infrastructures

Several HEP and Astro-Particle projects indicated their wish to share many aspects of the computing global infrastructure and services, while maintaining flexibility to adopt what suites them more. E.g. APPEC *[ID-84]*, DUNE *[ID-126]*.

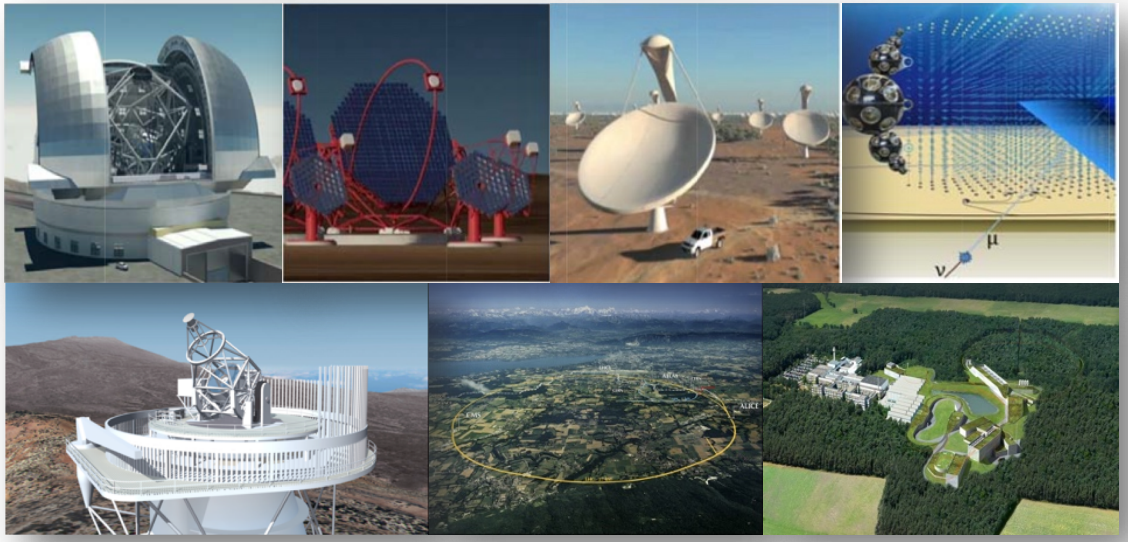
The idea to generalize the infrastructure related aspects of WLCG and open them to more scientific communities has been presented [ID-53] and generally well perceived. It is being prototyped with DUNE and Belle-2

ESCAPE

European Science Cluster of Astronomy & Particle physics
ESFRI research infrastructures

Horizon 2020 funded project

EUROPEAN OPEN SCIENCE CLOUD



Goals:

Prototype an infrastructure adapted to the Exabyte-scale needs of the large science projects.

Driven by the sciences

Address *FAIR* data management

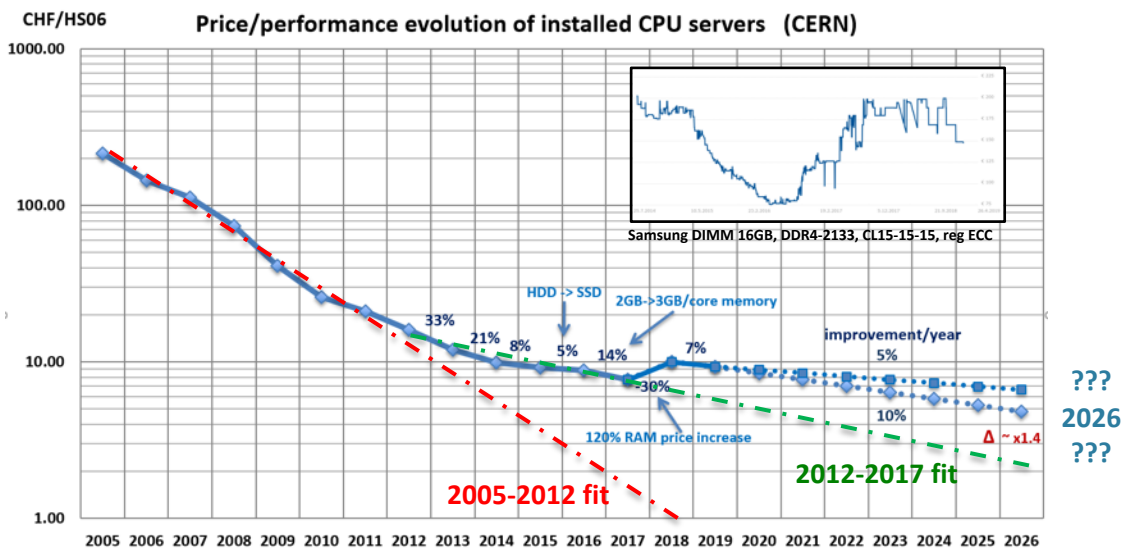
Science Projects

HL-LHC	SKA
FAIR	CTA
KM3Net	JIVE-ERIC
ELT	EST
EURO-VO (LSST)	EGO-VIRGO (CERN,ESO)

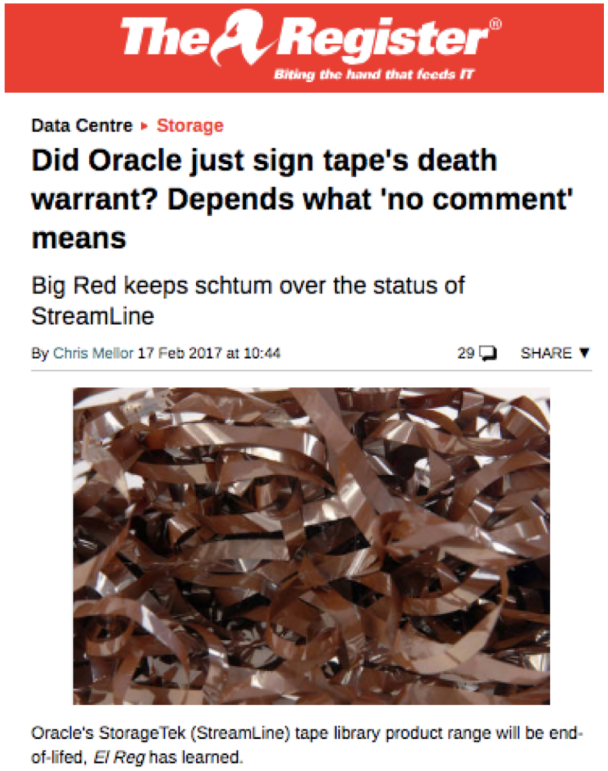
Data centres: CERN, INFN, DESY, GSI, Nikhef, SURFSara, RUG, CCIN2P3, PIC, LAPP, INAF

Challenges from hardware trends

- Cost of hardware decreasing exponentially vs time 😄 … but not as steeply as before 😳
- In general, trends driven by market (revenues) rather than technology



B. Panzer @ Scientific Computing Forum



The Register®
Biting the hand that feeds IT

Data Centre ▸ Storage

Did Oracle just sign tape's death warrant? Depends what 'no comment' means

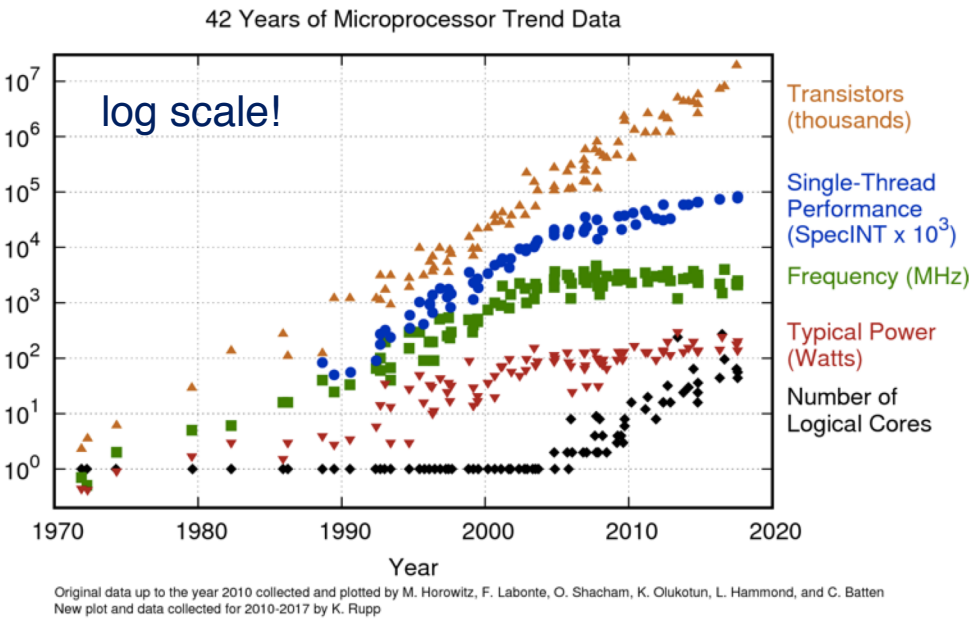
Big Red keeps schtum over the status of StreamLine

By Chris Mellor 17 Feb 2017 at 10:44 29 SHARE ▼

Oracle's StorageTek (StreamLine) tape library product range will be end-of-lifed, *El Reg* has learned.

General loss of long term predictability

Challenges from Hardware Trends



New Summit HPC @ ORNL:
9.96 Pflops in CPU (**non X86_64**)
215.7 Pflops in GPUs

https://www.karlrupp.net

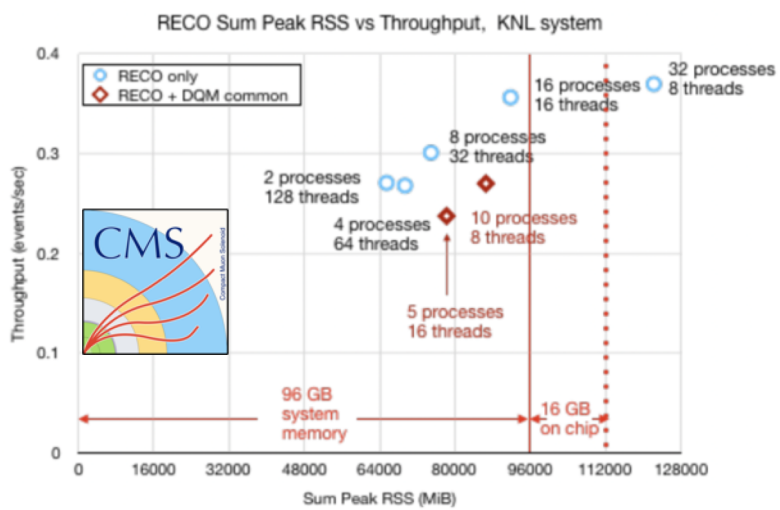
We need to take advantage of the huge GPU capacity at HPCs
(also to get good allocations)

In general, we need to continue to modernize our applications to take advantage of hardware evolution

HEP Software opportunities and challenges

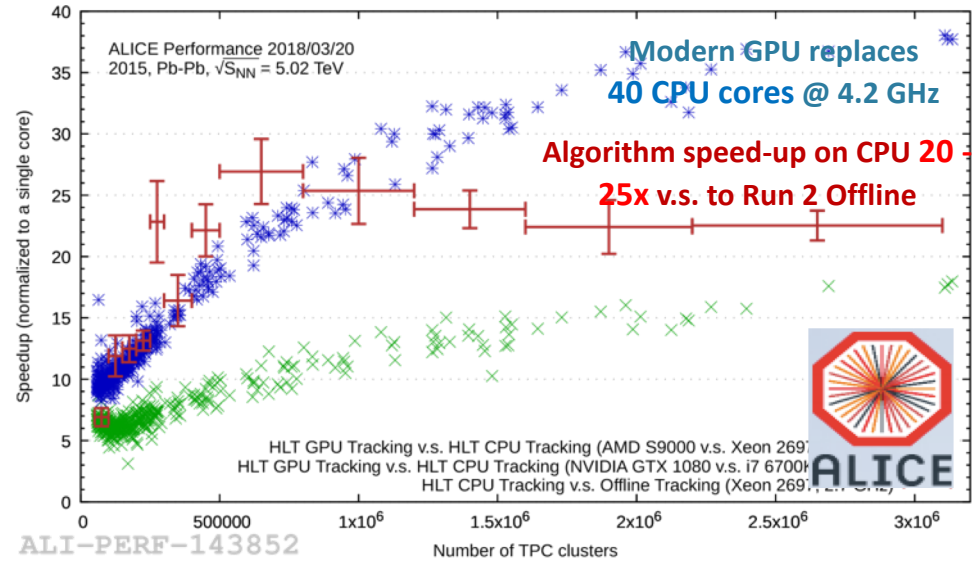
Software is probably the biggest opportunity to address the possible future HEP shortage of computing resources. See HSF input [ID-79]

Multithreading allows flexibility to optimize memory usage and throughput



CMS reconstruction MT+MP

Speed up from GPU usage + from algorithmic improvements + tuning on CPUs



D. Rohr @ HOW 2019

Increasing gap in skills between early career physicists and the profile needed for programming on new architectures, e.g. parallelization. Collaboration between physics and SW engineers communities on core components (see e.g. see [ID-114] from MCNET)

Lot of legacy in HEP software. To implement radical changes need a multi-year planning and institutional commitment as we have in other areas

Career prospects for those who specialize in software should be as bright (or no more gloomy) than anyone else in HEP. SW development is a critical part of our future success, let's reward it!

Following HFS CWP, several funded initiatives: IRIS-HEP (US), IRIS (UK), IDT-UM (DE)
Preliminary discussions on the idea of a Scientific Computing Institute. See also [ID-5] for a Deep Learning focused proposal

Innovation, collaboration with other sciences and industry, access to and experience with modern technologies is a key to motivate our community.

Machine Learning

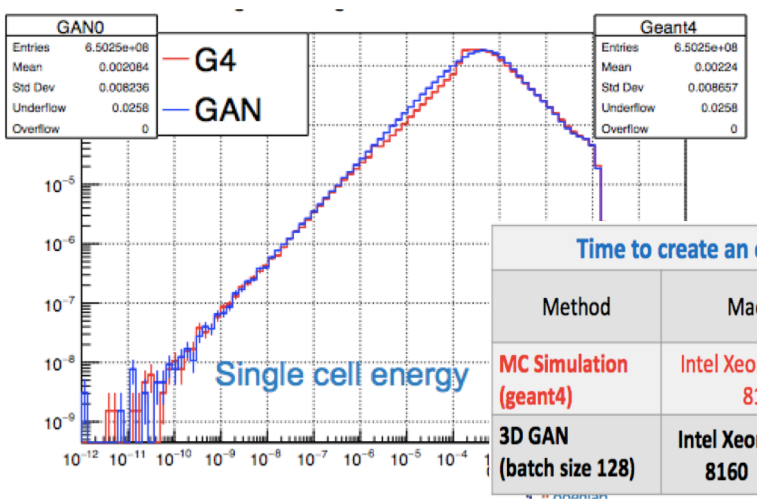
A concrete example of technology attracting interest and showing great potential

- Adopted in HEP since the 90s e.g. Multivariate Analysis
- Rapid recent development of tools and techniques in last years driven by industry
  - optimized and even specialized for modern hardware
- Different ML applications in HEP in different phases of maturity

ML Applications and R&D (from the HSF CWP)

1. Simulation
2. Real Time Analysis and Triggering
3. Object Reconstruction, ID, and Calibration
4. End-To-End Deep Learning
5. Sustainable Matrix Element Method
6. Matrix Element Machine Learning Method
7. Learning the Standard Model
8. Theory Applications
9. Uncertainty Assignment
10. Data Quality Monitoring
11. Operational Intelligence

HSF
HEP Software Foundation



Time to create an electron shower		
Method	Machine	Time/Shower (msec)
MC Simulation (geant4)	Intel Xeon Platinum 8180	17000
3D GAN (batch size 128)	Intel Xeon Platinum 8160 (TF 1.8)	1

S. Vallecorsa @ CERN IT Technical Forum

Example of Collaboration with Industry

CERN openlab [ID-162]

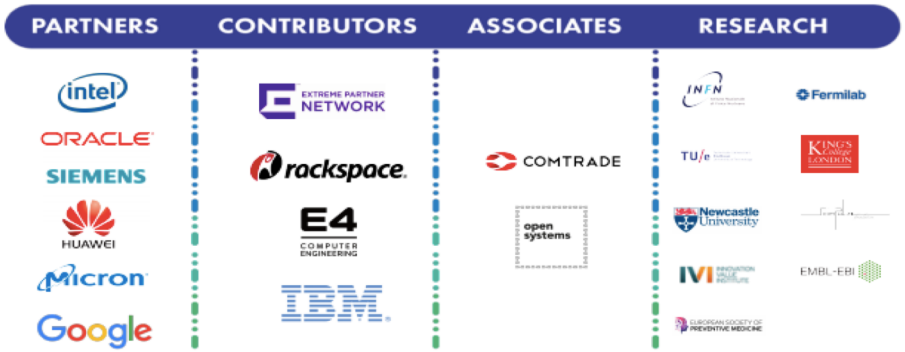
A science – industry partnership to drive R&D and innovation

Evaluate **state-of-the-art technologies** in a challenging environment and improve them

Test in a **research environment** today what will be used in many business sectors tomorrow

Training

Dissemination and **outreach**



CERN openlab

2

Besides aforementioned R&D activities, allows investigations of long term disruptive technologies e.g. Quantum Computing *[ID-59]* and *[ID-128]*

Aspects of data stewardship

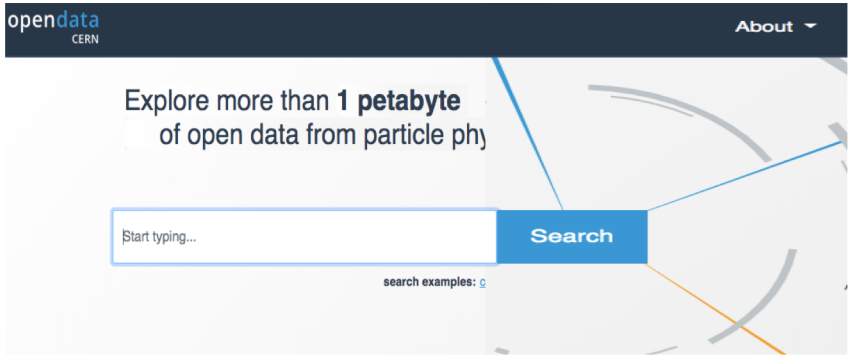
Data Preservation: preserving data and associated "information" for designed purposes

- E.g. ANALYSIS and educational outreach

Open Data – data released for re-use

- Data that is "open" is not necessarily accessible with zero latency

Significant resource/support requirements involved



(FAIR) **Data Management Plans**: bring the above together:

- Findable, Accessible, Inter-operable, Re-usable

DPMs must describe how data is preserved (including resource requirements)

- are increasingly an (EU) funding agency requirement
- need to be produced and updated at regular intervals

Do you have a Data Management Plan?

Conclusions (1/2)

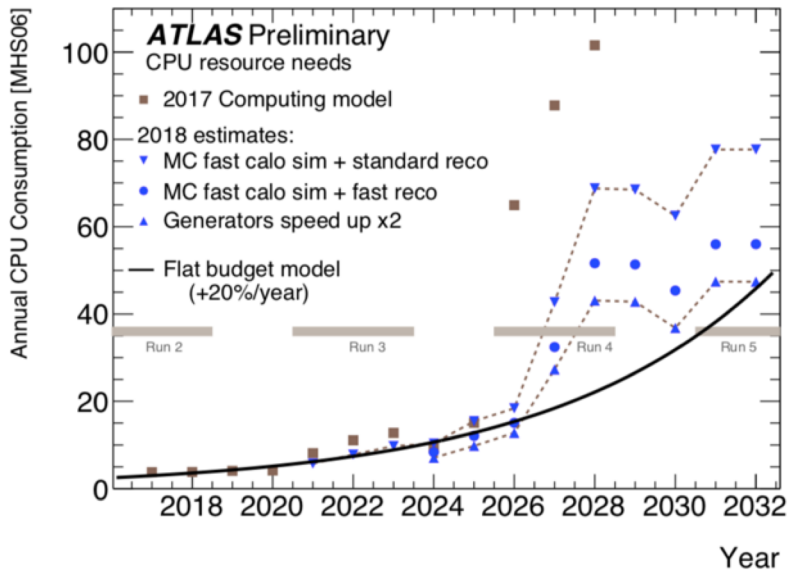
Computing resources for HEP in the 2020s will be more constrained w.r.t. the current decade: existing projects entering in a challenging phase, many new projects, competition from other sciences

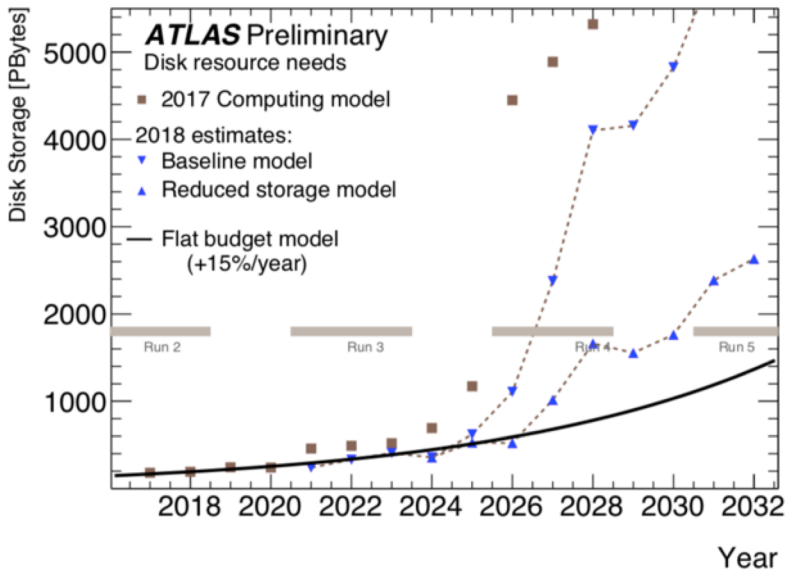
We have the possibility leverage this intellectual challenge and turn it into opportunity. Innovation is the key aspect as it offers:

- an occasion for young (and not so young) generations of physicists to master modern computing technologies and specialize
- solutions of our problem, maximizing the physics we get for our computing and software

Conclusions (2/2)

We faced the 2010s compute challenge with success. The 2020s present a new one.
A lot of work has been done for LHC Run-3 (e.g. Alice/LHCb): now under control
HL-LHC work started, there is a concrete plan and lots of opportunities

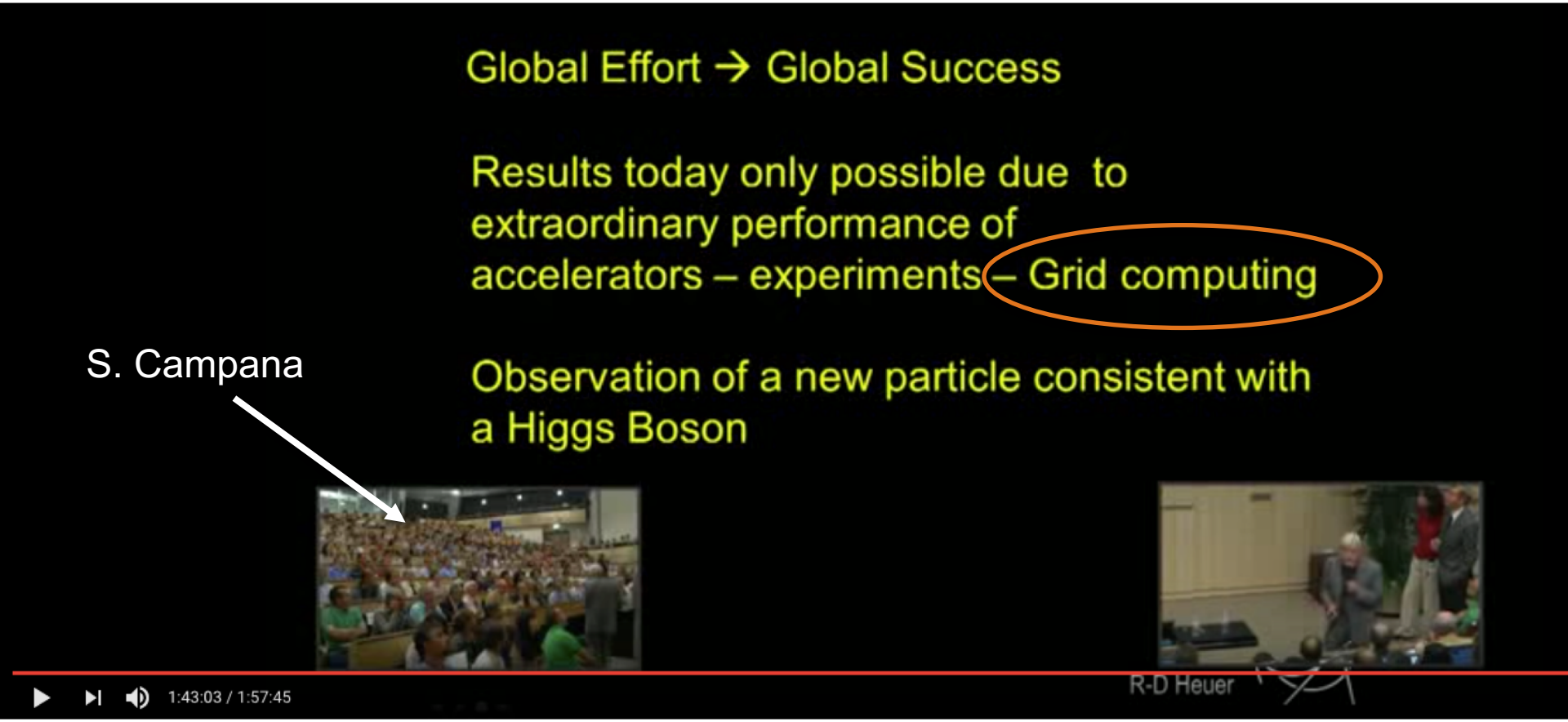




The plan needs to be implemented. Everyone on board!

CERN, update on the Higgs Boson searches at the LHC, 4th July 2012

https://www.youtube.com/watch?v=0Tnd8bxDYOs



Backup Material

Compute Cloud Resources

Commercial Clouds still need to demonstrate cost effectiveness when compared with on premise capacity. However:

- Offer elastic capabilities to absorb peak of activities for compute
- Price is decreasing …

Challenges: heterogeneous interfaces, networking, procurement and economic model, operational effort, **DATA MANAGEMENT**

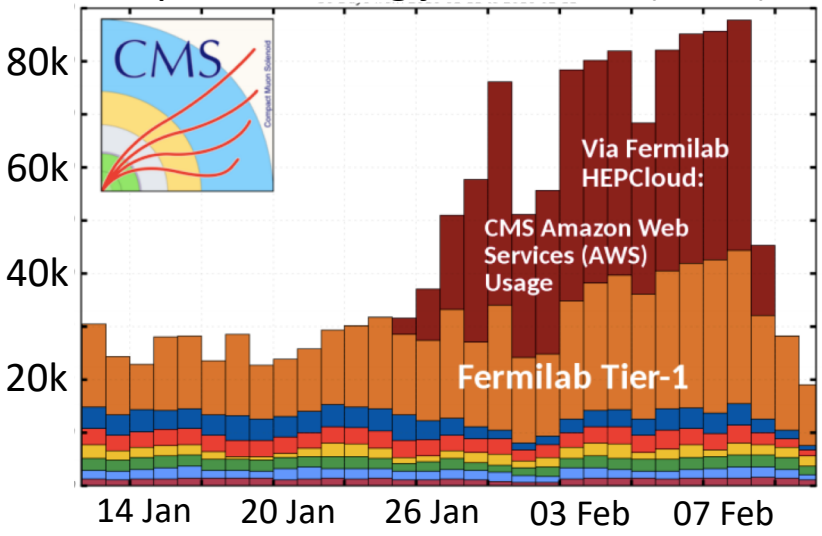
HELIX NEBULA THESCIENCECLOUD

Deployed in a hybrid cloud mode:

- Procurers' data centers
- commercial cloud service providers
- GEANT network and EduGAIN Federated Identity Management

B. Jones, J. Fernandes

prod running jobs @ T1s (2016)



Google, Amazon via HepCloud *Azure via DoDas*

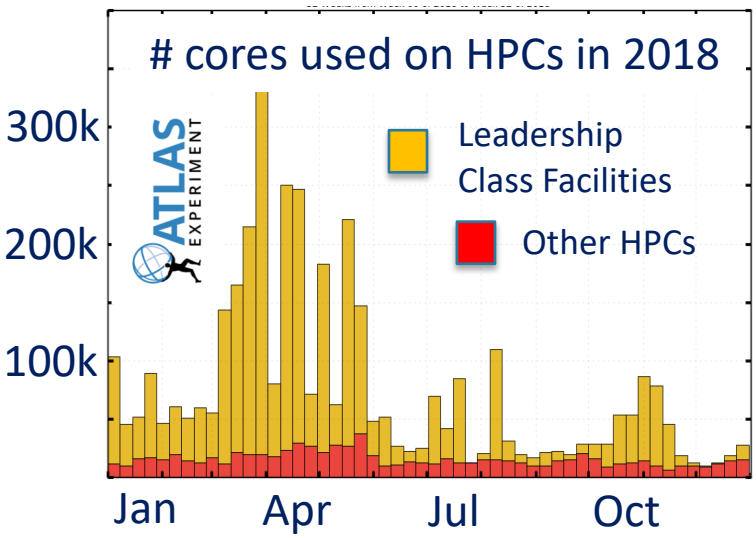
High Performance Computers

HPCs are in HEP computing, to stay and **grow**. HPCs are very heterogeneous. Major funding agencies are mandating a **very high profile for HPCs**

See e.g. *J. Siegrist (US DOE), Status of the DOE High Energy Physics Program (May 2018)*

Leveraging HPCs requires **dedicated investment** of effort:

- **stable allocations**, not just backfill, to make the investments pay; resource acquisition model is important
- **edge services** and tools for software distribution, provisioning, data handling. Negotiating policies is important



HEP and HPC centers need to engage:

find effective common and scalable solutions to the challenges of data intensive processing

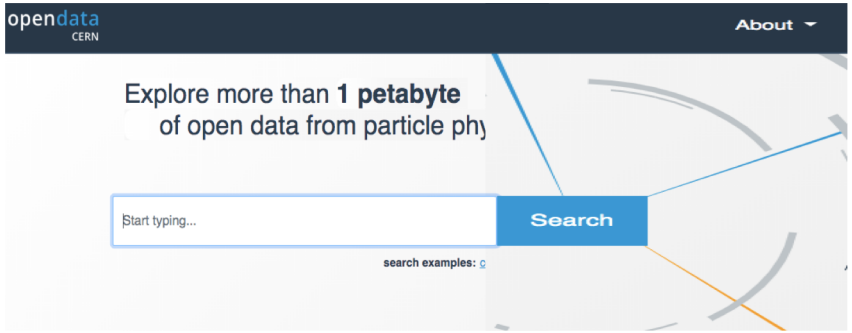
Mutual benefit if the link between HEP and HPC communities is strengthened. In Europe this would be through PRACE/EUROHPC

Aspects of data stewardship

Data Preservation – the (well established) active task of **preserving data and associated "information"** – even across significant changes in technology – for **designed purposes** such as ANALYSIS and educational outreach

Open Data – data that has been released for re-use / sharing typically with s/w + doc to make this possible. Data that is "open" is not necessarily accessible with zero latency

Significant resource/support requirements involved

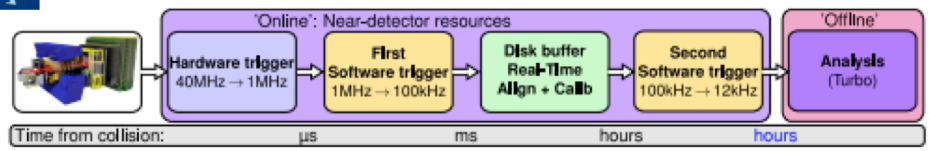


(FAIR) **Data Management Plans**: a relative newcomer but a way of tying the above together:

- Findable, Accessible, Inter-operable, Re-usable
- DMPs must describe (inter-alia) how data is preserved and can (should) include resource requirements
- (FAIR) DMPs are increasingly an (EU) funding agency requirement
- DMPs need to be produced and updated at regular intervals!

Do you have a Data Management Plan?

C. Fitzpatrick @ ACAT 2019



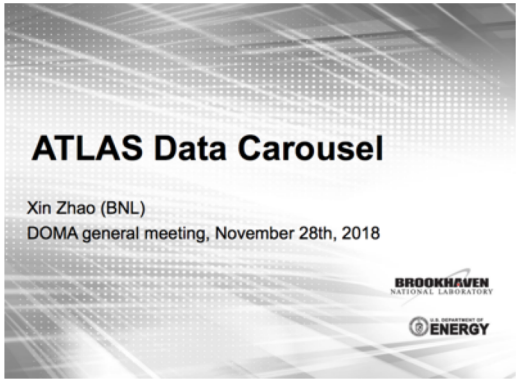
- Trigger *rates* aren't important, output *bandwidth* is
- Offline reprocessing previously needed to recover best quality
- After alignment: online == offline, why reprocess? Do analysis on trigger objects at HLT2, write only the relevant objects offline
- Significant reduction in event size → higher rates for the same bandwidth
- Keep all reconstructed objects, drop the raw event: 70kB

T. Boccali, M Klute

Data Tier	Size (kB)
RAW	1000
GEN	< 50
SIM	1000
DIGI	3000
RECO(SIM)	3000
AOD(SIM)	400 (8x reduction)
MINIAOD(SIM)	50 (8x reduction)
NANOAOD(SIM)	1 (50x reduction)

Analysis data formats

X. Zhao @ WLCG DOMA



Store on cheaper high latency media (tape today). Process in organized campaigns

Comments and Recommendations: 2017

7

- The experiments have worked to mitigate the impact of the increased LHC live time, but it is unreasonable to expect them to be able to keep pace with the rapid increase in LHC performance within a flat budget.
- C-RSG considers the requests of the experiments adequate for the physics program approved by LHCC.

D. Lucchesi @ 2016 LHC RRB