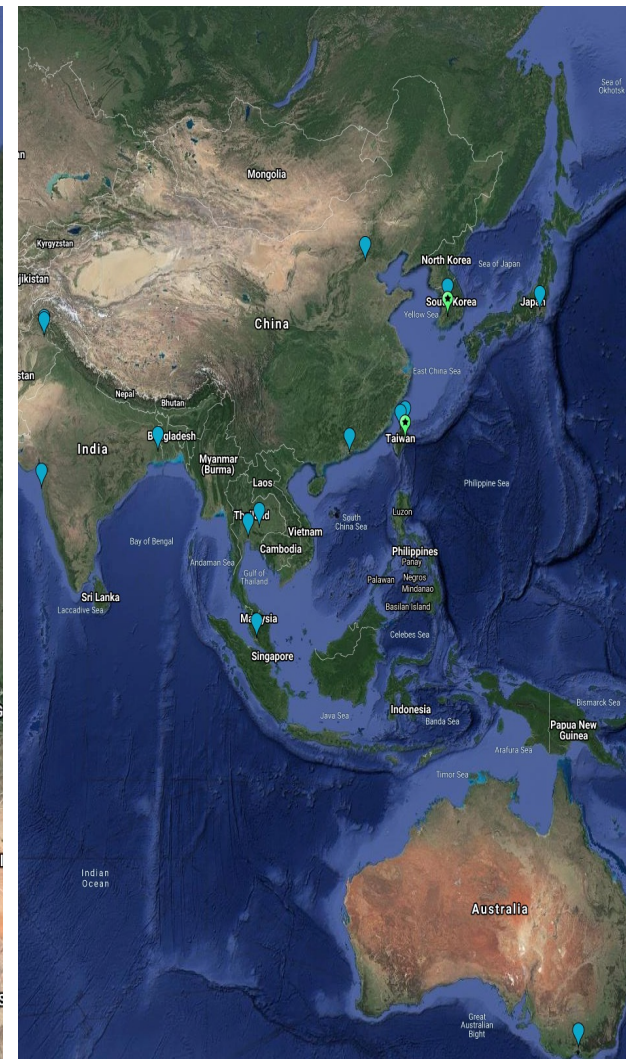
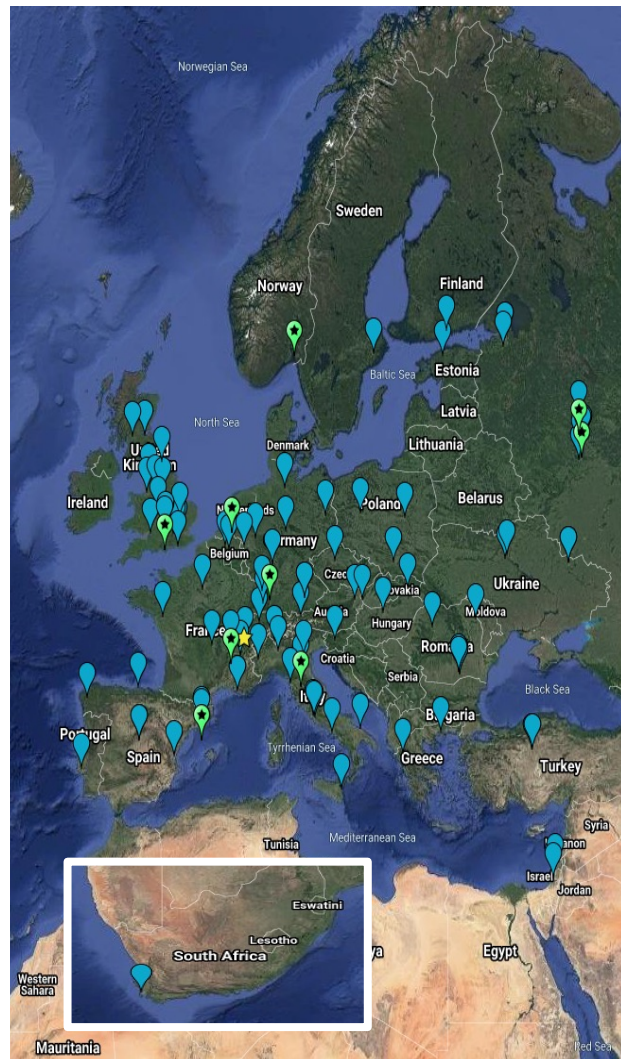
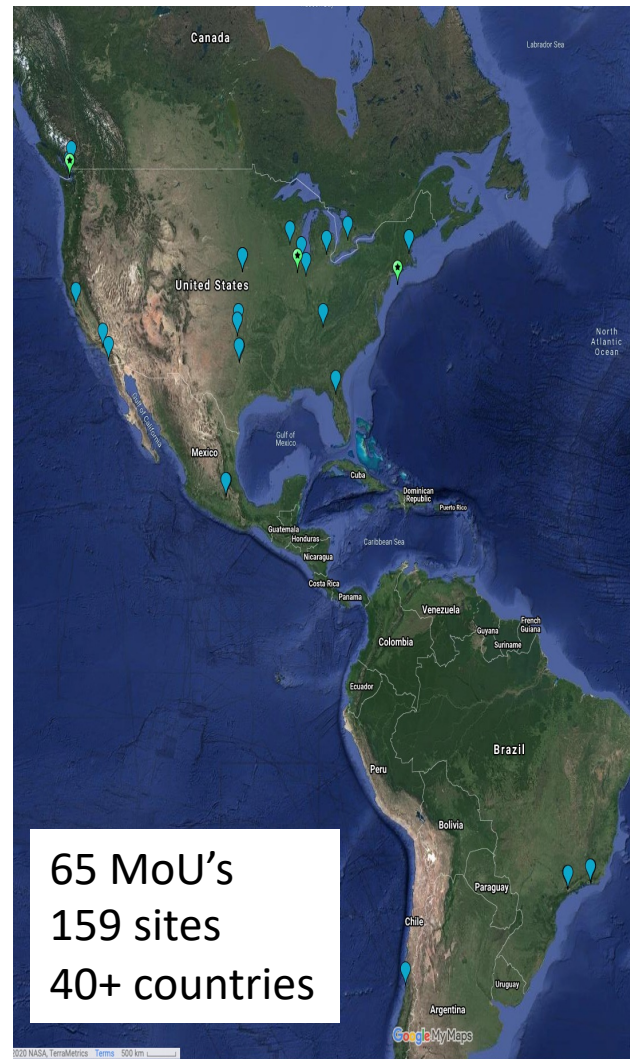


WLCG Status Report

S. Campana (CERN)

The WLCG Collaboration – May 2024



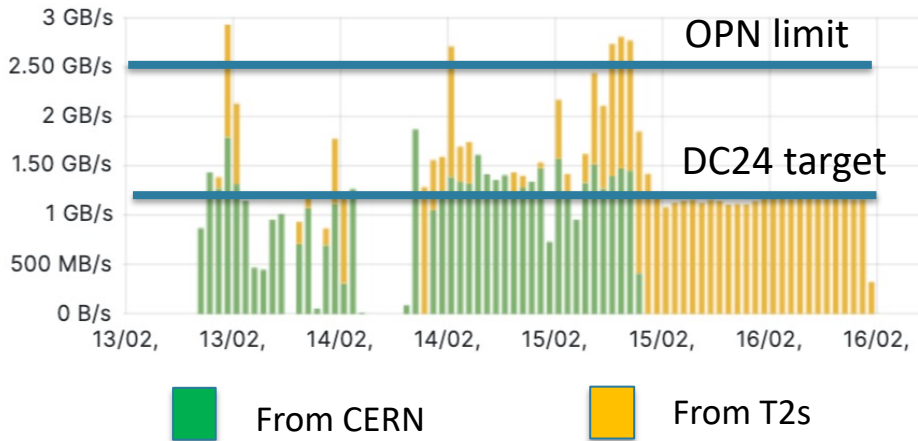
WLCG news

NCBJ – Swierk (PL)

Endorsed as WLCG Tier-1 for LHCb by the WLCG Overview Board



Throughput WLCG->NCBJ



Successfully passed scaling and stress test in Feb 2024

Used in production as a Tier-1 by LHCb since early 2024

WLCG news

IHEP – Beijing (CN)

- 100Gbps (best effort rather than reserved) network connection in place. Reserved network capacity being discussed
- Being tested now



Belgrade (Serbia)

- Signed the WLCG MoU in Dec 2023
- Plan to become a T1 for CMS



WLCG news

Russia

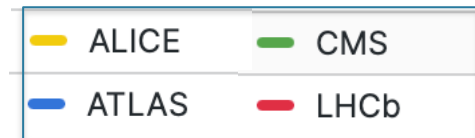
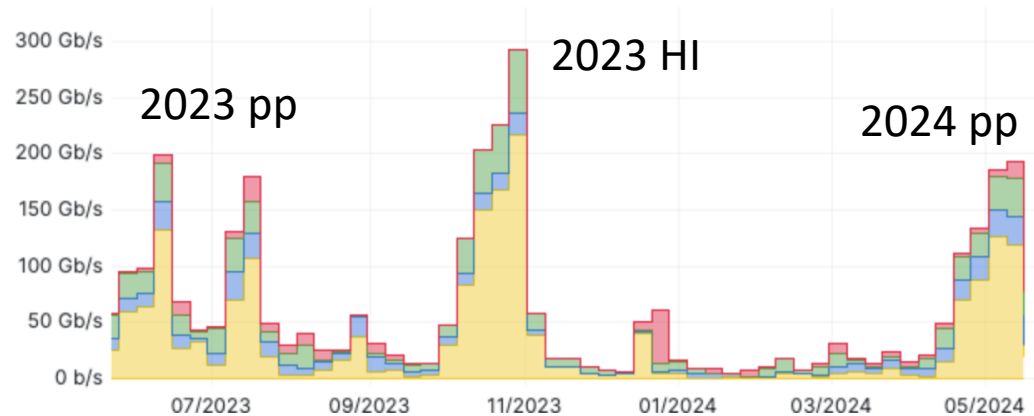
The CERN Council agreed to terminate the Cooperation Agreement with Russia.

- All MoUs will be automatically terminated end Nov 2024, including the WLCG one
- Russian sites will not be part of WLCG after that
- Russian resources continue being used opportunistically and efficiently for the time being

JINR (Dubna) will be discussed by CERN Council in June 2024

Run-3 data taking

Experiments -> T0 network use (12 months)



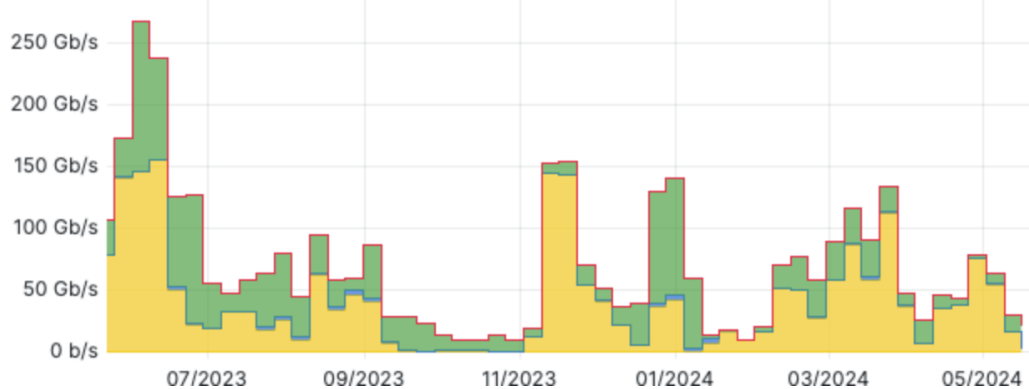
25 PB of data collected and efficiently transferred to the T0 in 2024 so far (+29 PB of ALICE pp CTF)

- ~97 PB in 2023, ~39 PB in 2022

Online farms used for offline processing when possible

- Input data imported from the T0

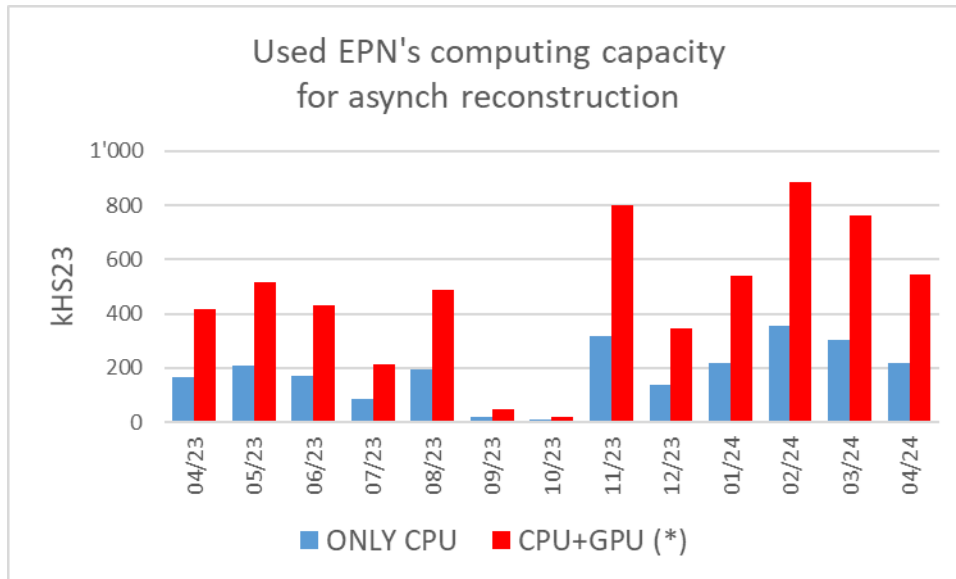
T0 -> Experiments network use (12 months)



Run-3 data taking

In ALICE, both CPUs and GPUs of the O2 system are used for asynchronous (offline) reconstruction, when not needed by online

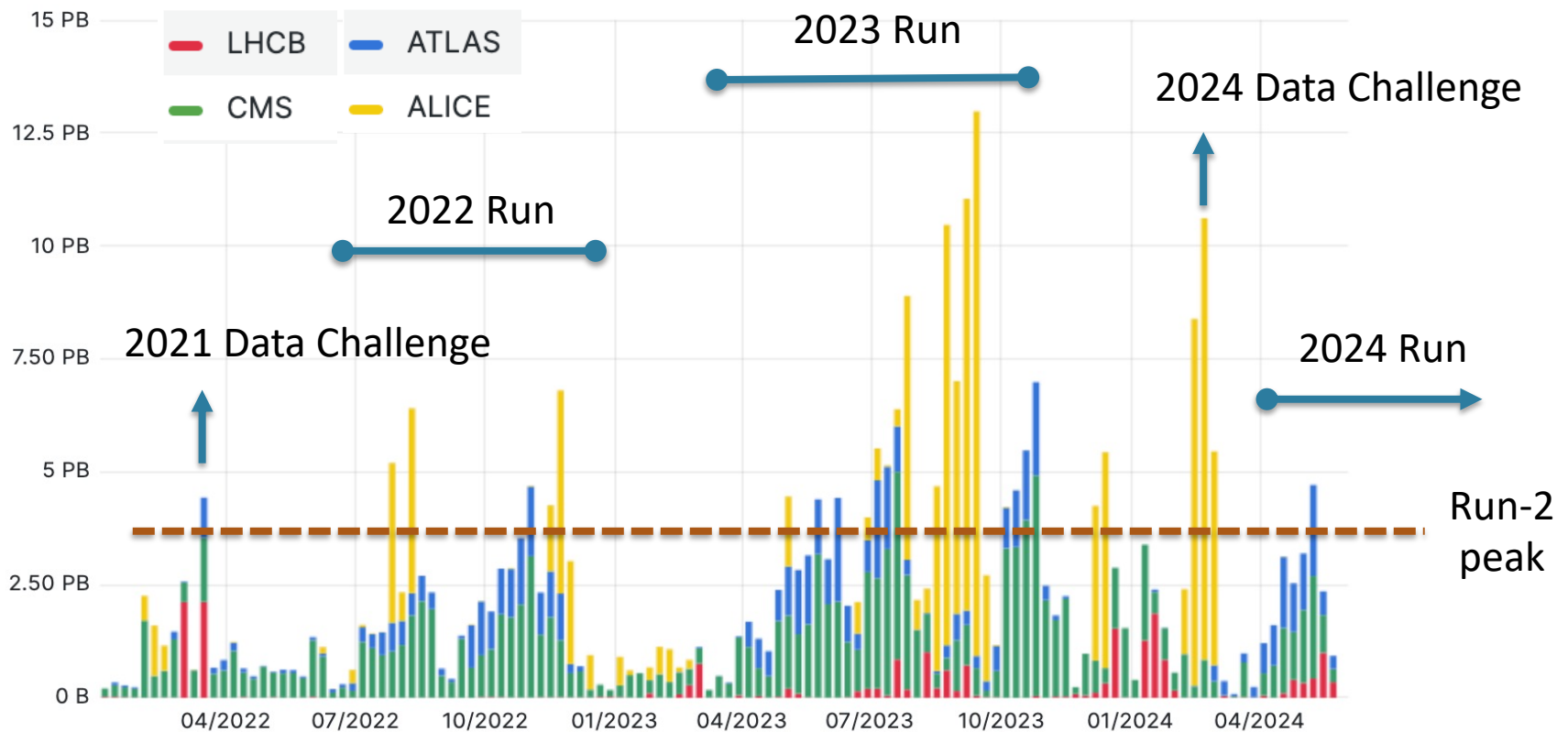
- 85% of EPNs compute power is in the GPUs
- 60% of the asynchronous reconstruction code has been ported to GPUs. Works on AMD and NVIDIA
- GPUs bring x2.5 speedup in asynchronous reco
- Testing use of GPUs at other sites planned. Requires non-trivial job tuning



CERN data archiving

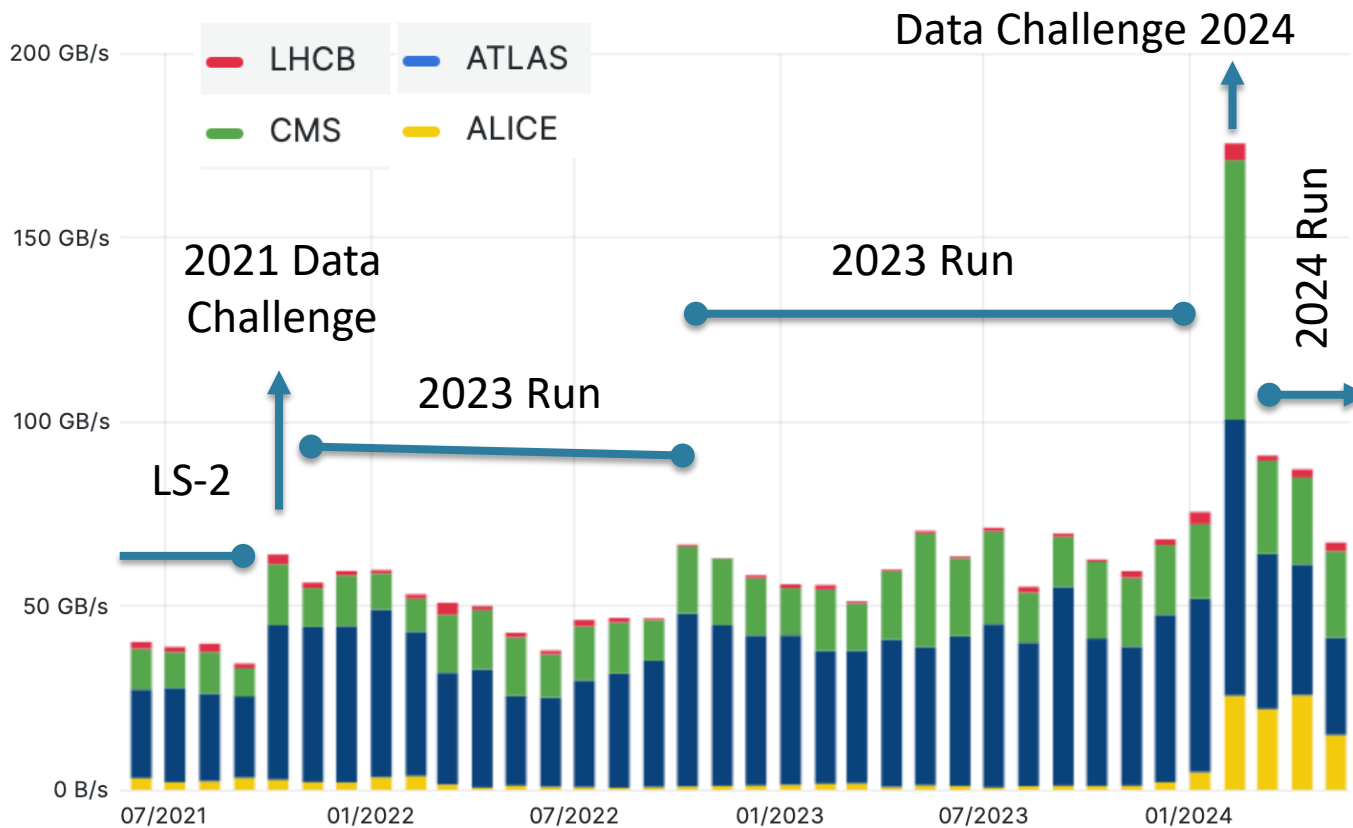
The 2024 Run started smoothly

Weekly volume of data archived at the T0 (PB) – 3 years



WLCG data transfers

Monthly data transfer throughput between WLCG sites (GB/s) – 3 years



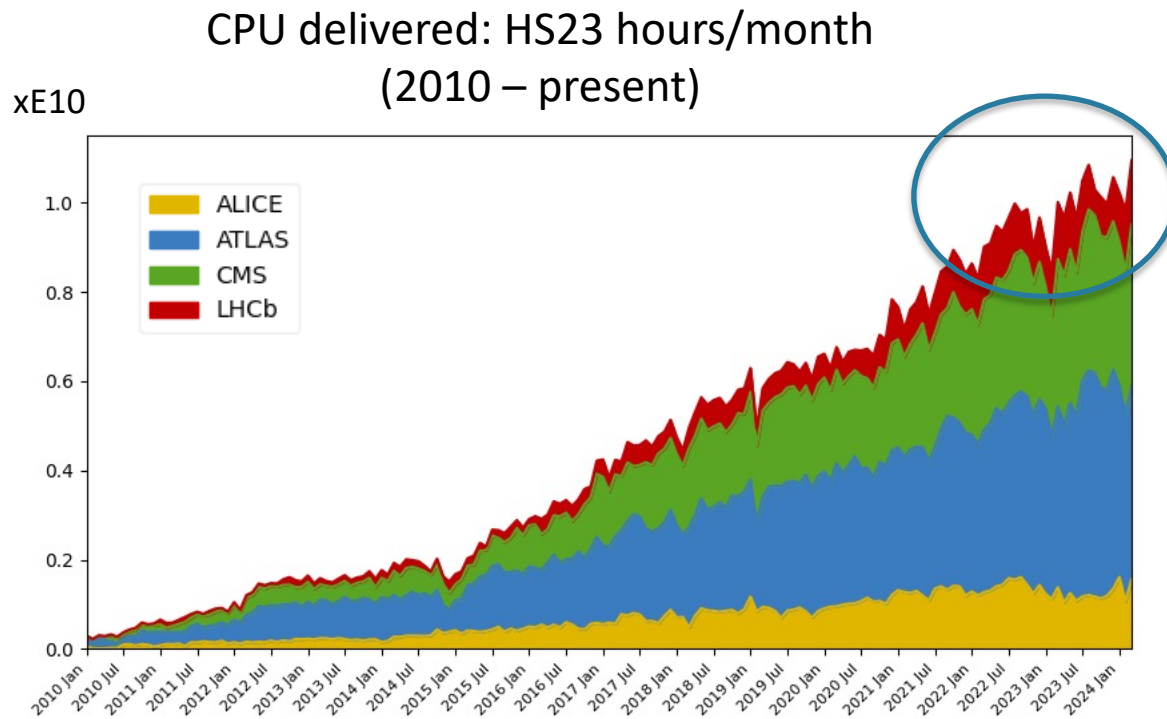
WLCG supports +40% more transfers since LS-2

Further scalability (x5) demonstrated in the Data Challenge

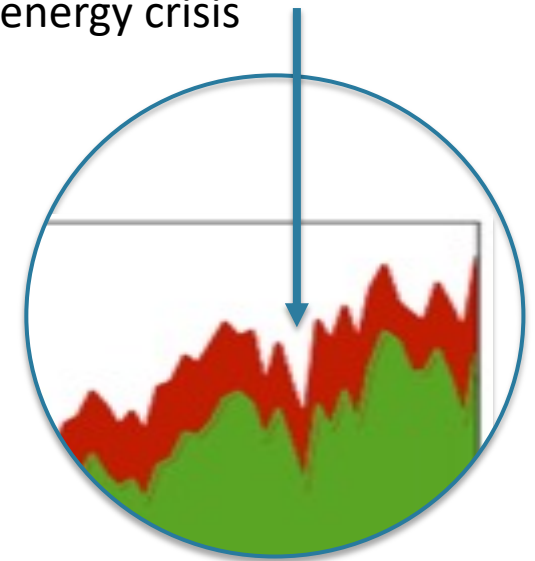
No strain to the services

WLCG data processing

Growing number of computing resources provided to the experiments (+20% since LS2)



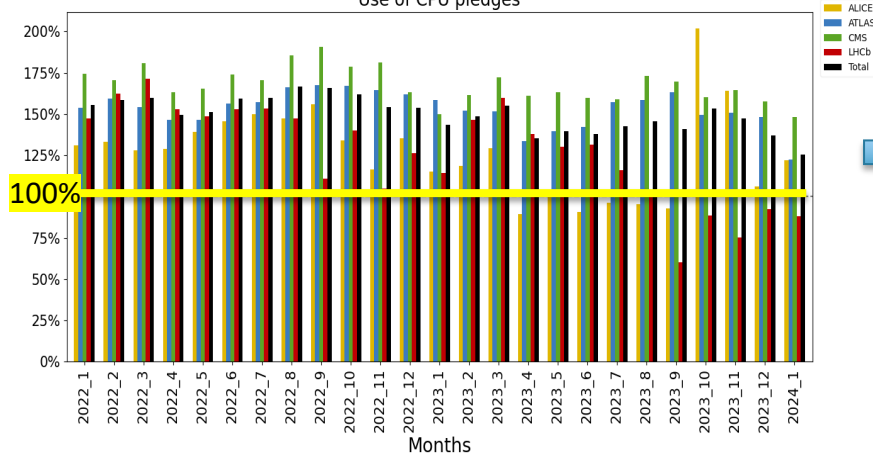
Drop in winter 2022/23:
energy crisis



Back to “normal” now

Data processing

Use of CPU pledges

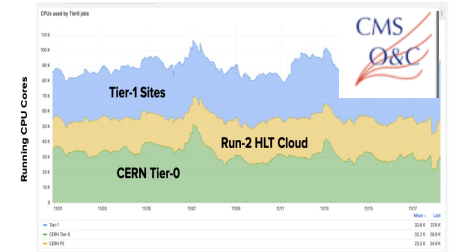
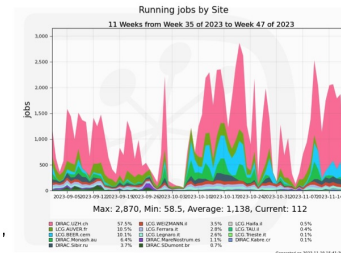
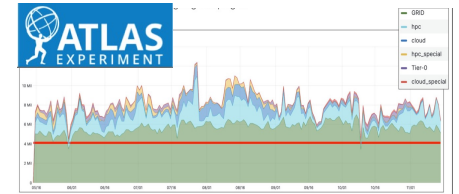
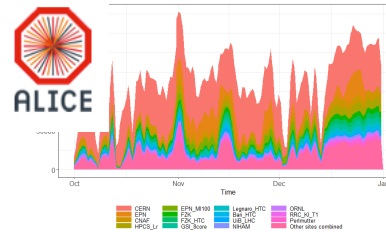


The experiments benefit from >40% extra capacity at WLCG sites. Plus HPCs, HLTs,..

The use of non-Grid resources is monitored by the experiments



WLCG is working on a consolidated view

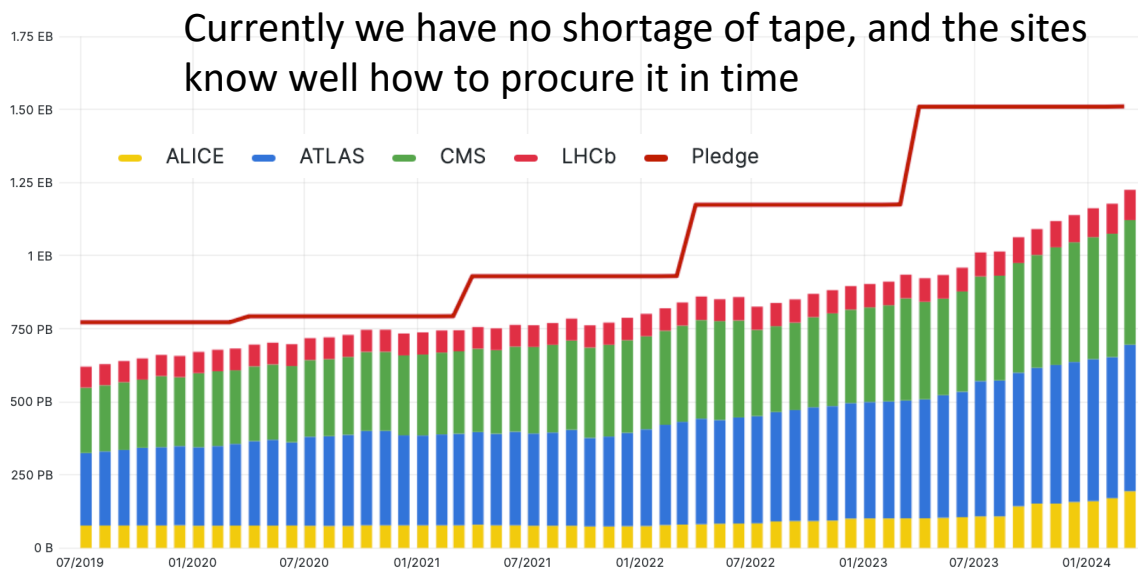


Run-3 overview

4 weeks of pp run shifted from 2025 into 2024: ~30PB of tape need to be anticipated from 2025 into 2024 at both T0 and T1s (+ some PB of disk)

Thanks to the Funding Agencies support we are confident we will not have a problem

Tape used (EB) wrt pledges



**2024 hardware:
installed and in use**

Delivery times now slightly shortened than in the past and the sites have a lot of experience with procurement

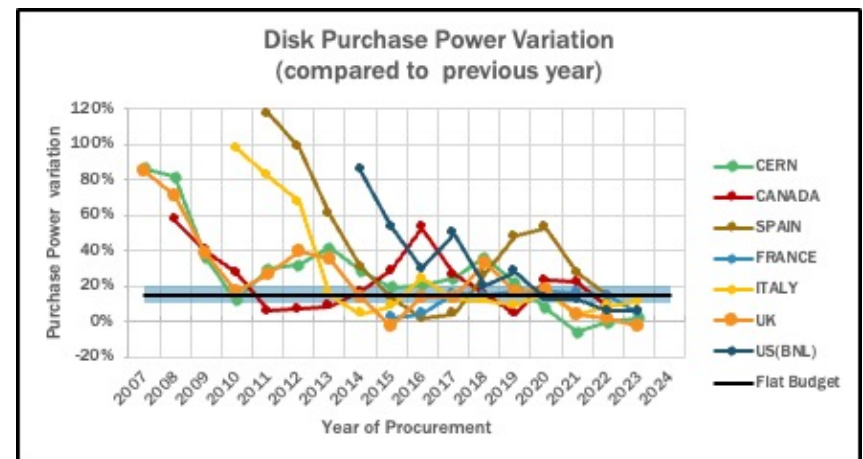
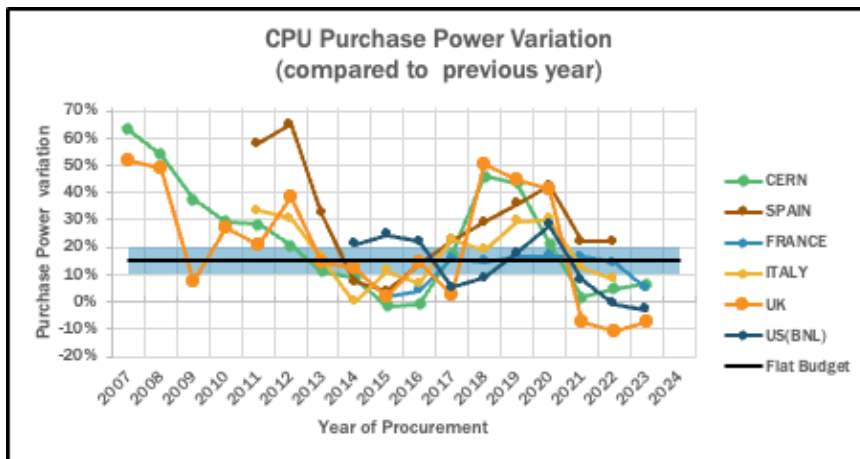
Hardware Cost Evolution

The WLCG "flat budget model" assumption: +15% CPU, disk and tape every year **with the same level of funding**

We now monitor the HW trends in many countries. Last 5 years average is compatible with the 15% assumption but look at the first derivative ...

CPU average variation (5 years): +14%

DISK average variation (5 years): +15%

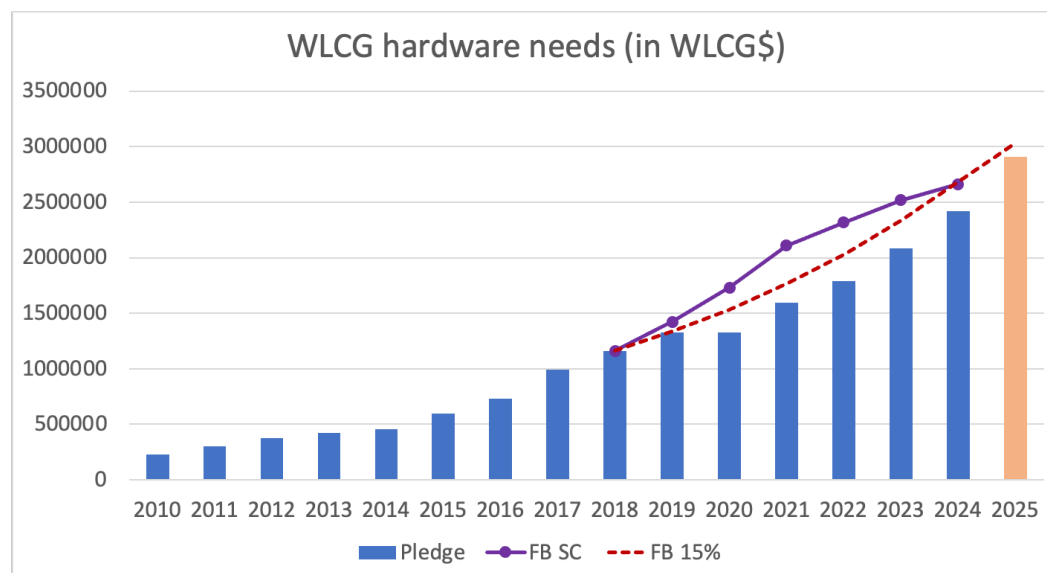


WLCG resource needs compared to Flat Budget models

We can convert the cost of hardware into a single WLCG currency, based on the CERN and UK data. In average: **1WLCG\$ = 1TB disk ~ 11 HS23 ~ 5TB tape**

Note on tape: new generation (LTO9) is expected to be **15% more expensive per TB** (see [here](#))

Full overview of hardware trends [here](#)



The WLCG resource needs of the experiments could be accommodated with a flat budget in 2024. But the trend does not look encouraging for the next years



WLCG/HSF [workshop](#): 13-17 May at DESY-Hamburg

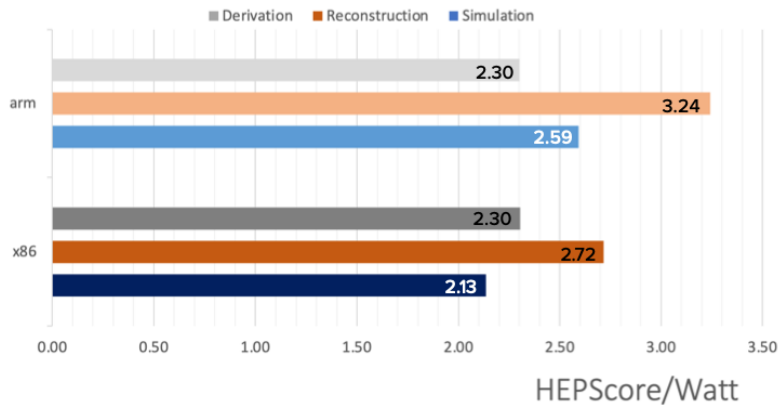
Main topics for WLCG:

- Finalisation of the 2024-2027 [WLCG strategy](#) and implementation plan
- Data Challenge 2024 post-mortem
- Analysis Facilities
- Operations and Facilities
- Collaboration Board to endorse the proposed changes to the WLCG MoU

A lot of focus on the preparation for HL-LHC



ARM CPUs



ARM CPUs process more events/Watt wrt X86

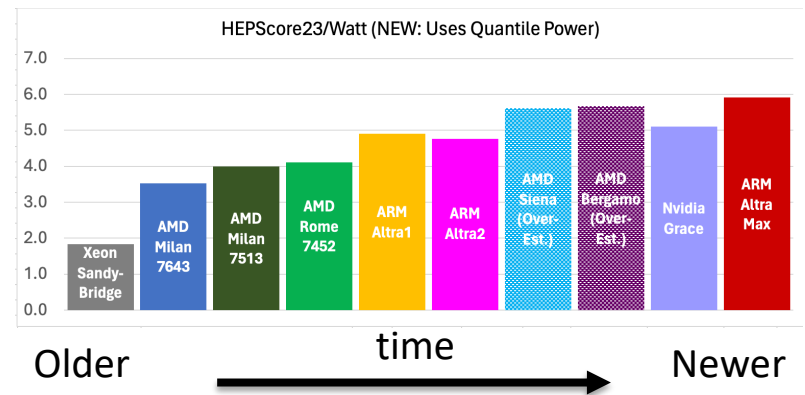
- This is the case for the HW in Glasgow
- Differences between workflows
- Modern AMDs perform similarly to ARM - again based on the HW in Glasgow

See [this presentation](#) for details

All experiments ported main workflows to ARM

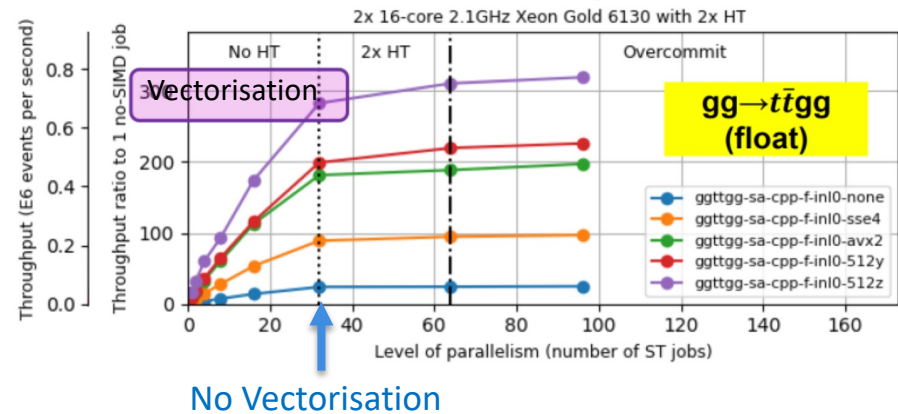
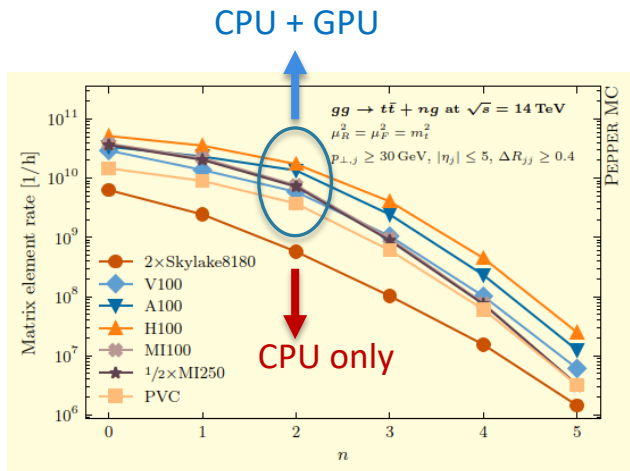
- ATLAS, ALICE: physics validation successful
- CMS, LHCb: in progress

WLCG is discussing about pledging ARM CPUs.
Resources available at various sites



Event Generators

A very good candidate for GPU acceleration with benefits for many experiments



Sherpa $gg \rightarrow t\bar{t} + n_g$

Matrix Element event throughput:
up to x10 gain when using GPUs

Available for production

Madgraph $gg \rightarrow t\bar{t} + n_g$ ($n=2$)

GPU-enabled Leading Order: being released to production.

By-product: enabling of CPU vectorisation: up to x8 gain in ME event throughput (x6 global). Note: all CPUs in WLCG provide vectorisation

GPU-related work brings immediate benefits also on CPUs

Networks and Data Challenges

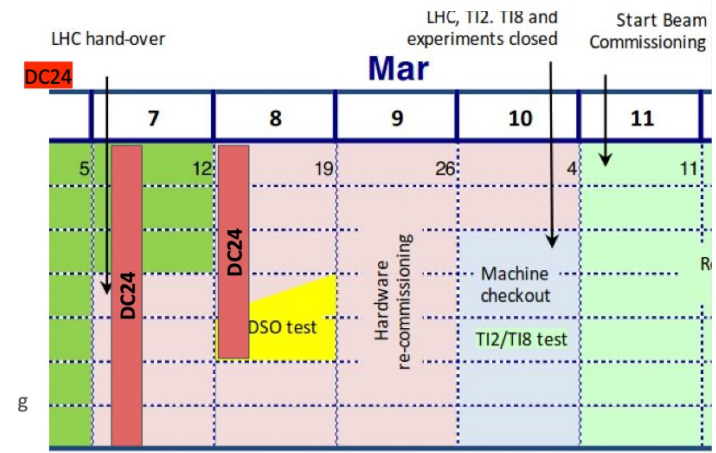
DC24 is one phase of the computing infrastructure [commissioning](#) for HL-LHC

DC21 lessons learned are documented [here](#).

DC24 had 3 goals:

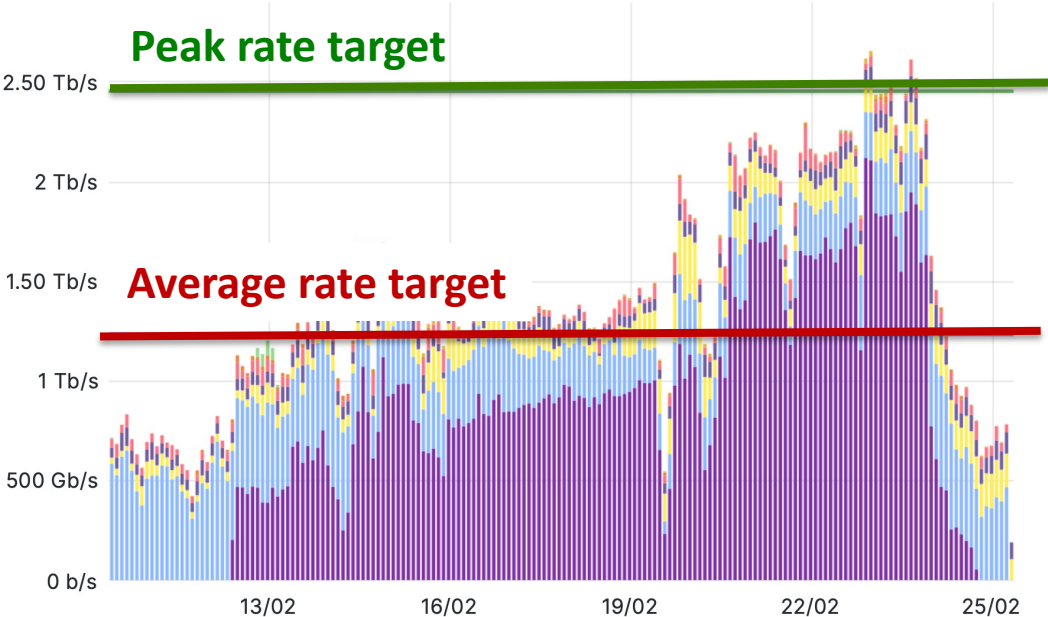
- Measure the end-to-end data transfer capabilities at WLCG sites (target is 25% of HL-LHC needs)
- Assess the progress integrating new technologies (e.g. tokens and monitoring)
- Assess the status of different R&D initiatives

DC24: from Feb 12 to Feb 23 in 2024



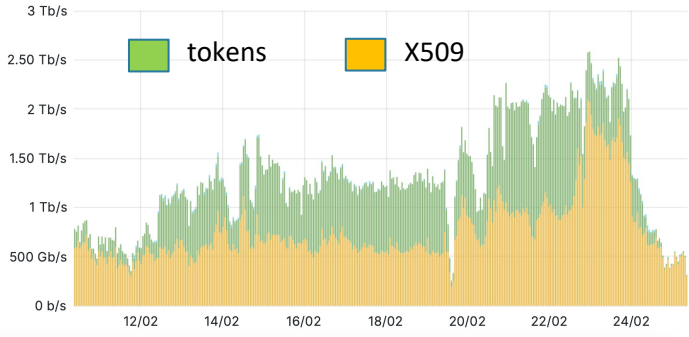
Data Challenge 2024 - Highlights

DC24 WLCG data transfers (Gbps) – 15 days: **all targets achieved**



- ATLAS
- CMS
- DC
- Belle-2
- ALICE (XRD)
- CMS (XRD)
- LHCb
- DUNE

New technologies (e.g. authentication **tokens**) introduced and validated



WLCG services successfully supports DUNE and Belle-2 computing models



DC24 Network R&D – one example

R&D proposed as follow up of DC21. Deployed in time for DC24

In production today for Run-3 data taking

Reduces cost while providing more bandwidth (x2) and expansion capabilities

Reduces latency (30%)

Being tested now for longer network paths

Transmission Testing

CNAF-CERN DCI

- spectrum sharing over GEANT and GARR dark fibres
- 4x100Gbps links between CERN and CNAF used for DC24 and now in production

Cost effective technique to get >1Tbps LHCOPN links already today

Geneva CERN DC

Bologna CNAF DC

TRX

Add/Drop

ROADM

8x AMP

3x AMP

GEANT Open Line System

GARR Open Line System

Milan PoP

1.6 Tbps

WLCG Worldwide LHC Computing Grid

LCG

Open Science Grid

WLCG/HSF Workshop 2024


8

DC24 Network R&D – more in general

Monitoring

DC24 perSONAR

WLCG/OSS/LHC have deployed and maintained a perSONAR infrastructure for our site network connectivity to each other.



For DC24 the plan was to utilize the testing networks and fix them prior to DC24. No major misconfigurations or network issues lead up to DC24.


- Some site issues were
- R&E networks were

The process of identifying network issues remains a challenge and work continues to better auto.

Consolidation

DC24 Scitags (Packet & Flow Marking)

In the January 2020 LHCONE/LHCOPN meeting at CERN, the experiments converged on three areas of interest for networking: visibility, optimization and orchestration. The Scitags Initiative has been working on making R&E network traffic "visible" anywhere in the network.

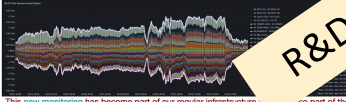


During DC24 we had over 80% of the CERN EOS Nebraska Tier-2 sending frannies for their associate USCMS. ESnet dash: <https://esnetbooster.stanford.edu/globe/IPv62eUj>. Currently have EOS and Xroot Scitags-capable and are on Storm and dCache. Target: have all production traffic labelled by DC28.

R&D

DC24 Site Network Monitoring

During DC21, we didn't have sufficient data about how much traffic each contributing site was experiencing and it made the analysis to identify bottlenecks difficult. For DC24 the WLCG Monitoring Task Force was charged with filling in this missing information. The larger sites participating in DC24 were given GGSIS tickets to enable site network monitoring, using SNMP to monitor total incoming/outgoing traffic.



This new monitoring has become part of our regular infrastructure. Also part of the purpose of the data challenges: to evolve our infrastructure to what is needed for HL-LHC.

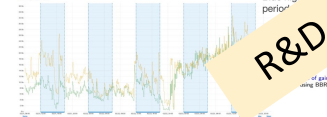
R&D

Protocols

DC24 TCP Protocol Explorations

For WLCG data transfers, optimizing the TCP protocols may lead to higher throughput for individual transfers and better sharing of available bandwidth.

BBR demo at DC24



R&D

DC24 Jumbo Frames Activities

Benefits of jumbo frames are evident on long distance transfers, less on the short distance. Operational issues are also evident, but they can be mitigated by sharing deployment experiences.

No special test was done on Jumbo frames. To be noted that servers have been happily using Jumbo for years and had no problem with it.


The preparation of DC26 will focus on a wider use of jumbo frames.

R&D

DC24 IPv6 Activities

One of the longstanding challenges for WLCG has been the transition to IPv6, which has been going on for more than a decade. DC24 gave us a chance to see where we were and what we needed to do to get there. Sites have been requested to provide perSONAR and storage dual-stack capabilities by November 2017.

Recently the request to dual stack services and worker nodes has been made, targeting June 2024.




While there is still significant work to be done, we are transitioning in the right direction. All T1 and 97% of T2 storage are transitioning to IPv6-only.

Consolidation

Traffic shaping

DC24 SENSE/Rucio (Network Orchestration)

The objective was to provide Rucio with capabilities to request network services via SENSE in order to: **a) improve accountability, b) increase predictability, and c) isolate and prioritize transfer requests.** This project used a dedicated Rucio as well as Xroot instances so it would not interfere with Production systems. Data was transferred across a mix of production and next generation network paths.



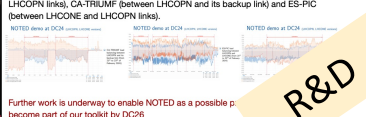
Between Fermilab, Catech, USCD Rucio-DMM/SENSE-FTS-Mirood multi-tenancy managed between multiple pairs of sites. The modify feature of DMM will be used to filter in response to Rucio requests. The following Quality of Service (QoS) on Server; Hard QoS at the network level. DMM Real-time API-drive transfers settings. Additional US-CMS Tier2 sites will be evaluated for deployment.

R&D

DC24 NOTED (Software Defined Networking)

NOTED (Network Optimized Transfer of Experimental Data) is an intelligent network controller to improve the throughput of large data transfers in FTS (File Transfer Services) by handling dynamic circuits.

During DC24 traffic was monitored and balanced for DE-KIT (between its LHCONE and LHCOPN links), CA-TRIUMF (between LHCOPN and its backup link) and ES-PIC (between LHCONE and LHCOPN links).



Further work is underway to enable NOTED as a possible production service to become part of our toolkit by DC28.

R&D

Plans already in place for DC26

All details [here](#)

Analysis Facilities – a ROOT based implementation

Based on RDataFrame: a ROOT high level API for data analysis

See [here](#) from Danilo Piparo at the WLCG workshop

RDataFrame and Analysis Facilities

RDF runs on distributed systems, already today: DistRDF

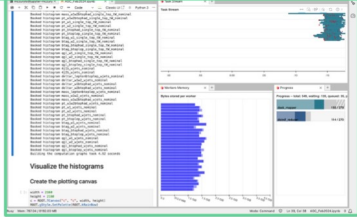
- ▶ Same analysis code of local execution
- ▶ **Shell and notebooks**
- ▶ Runs everywhere
 - Bare-metal clusters
 - HPC environments
 - **Analysis Facilities!**

See linked papers for more information

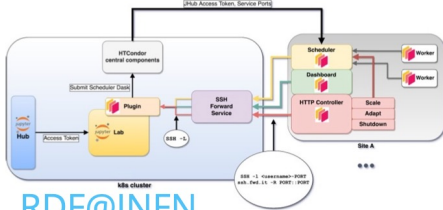
ROOT is ready to run at your favourite analysis facility, today

RDF+RNTuple@SWAN

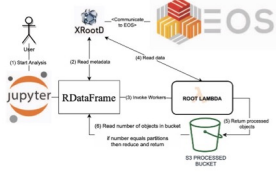
```
"root://eospublic.cern.ch/eos/root-eos/AGC/"
"root://eospublic.cern.ch/eos/root-eos/AGC/rntuple/"
```



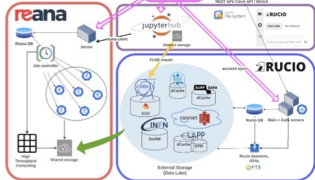
RDF@INFN



RDF@AWS



RDF@VRE



WLCG/HSF Workshop 2024 - D. Piparo, CERN EP-SFT – 14.5.2024

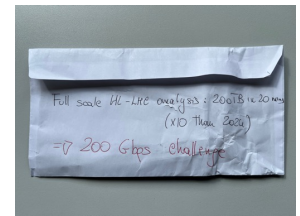
IRIS-HEP: analysis 200Gbps (Grand) Challenge

Launched by IRIS-HEP to commission analysis capabilities for HL-LHC

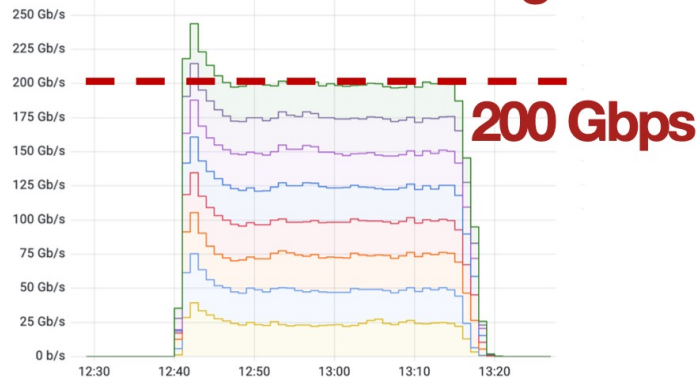
- Commission services at increasing scale + introduce innovative aspects

⇒ Show readiness at 25% of HL-LHC scale (same as for data challenge)

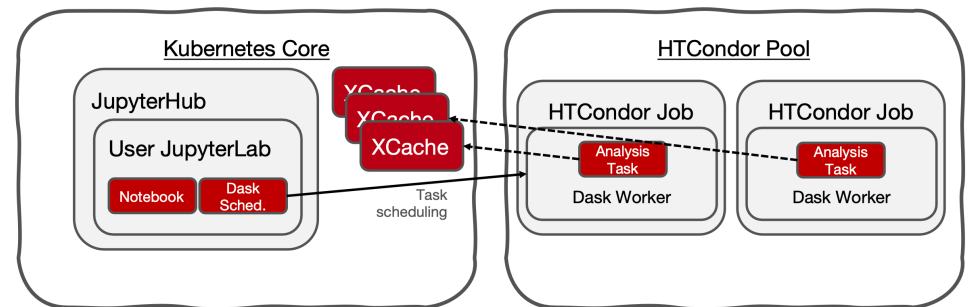
Analysis models evolving => metrics of success hard to quantify (25% of?)



Network monitoring



Based on the IRIS-HEP toolkit



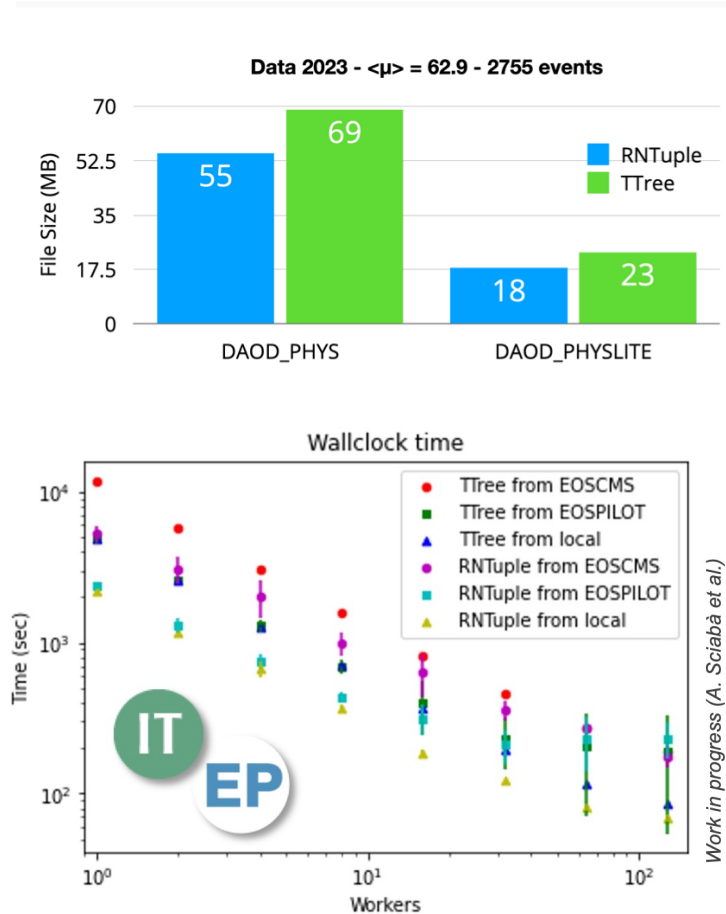
RNTuple

RNTuple is the successor of TTree, the ROOT columnar storage technology

Examples of recent commissioning progress:

- ATLAS now capable to read/write all data formats in RNTuple, 20% saving in size for DAOD_PHYS. Substantial progress also for the other experiments
- Expected RNTuple speed-up improvements measured in a real environment at CERN using a community standard analysis benchmark

Take home message: **RNTuple progress well on schedule** thanks to a very good collaboration between experiments, CERN EP-SFT and IT





Outcome of the workshop: a more complete identification of computing needs for the next decade is needed. Focus on synergies across the three communities.

Five thematic working groups launched: HPCs, SW and heterogeneous architectures, Data Management and FAIR data, ML, Training – Dissemination – Education

Working groups have been mandated to draft a bottom-up report by 30 Nov 2024.

The outcome will be discussed in ENA and summarised in a paper for the Funding Agencies. This will be input for the JENA Symposium in April 2025.

For WLCG this is an important step for the next European Strategy for Particle Physics

Conclusions

- The WLCG collaboration is broadening with new commitments
- Successful start of 2024 data taking. WLCG services running smoothly
- A flat computing budget is getting tight for supporting the experiment future needs, particularly for HL-LHC
- Very good technical progress towards HL-LHC: DC24, RNTuple, event generators (this is only a partial list). DC26 will be 50% of HL-LHC traffic
- Analysis Facilities is a rapidly evolving area with a lot of solutions available or being prototyped
- WLCG Strategy + JENA + .. we started preparing for the next update of the European Strategy for Particle Physics

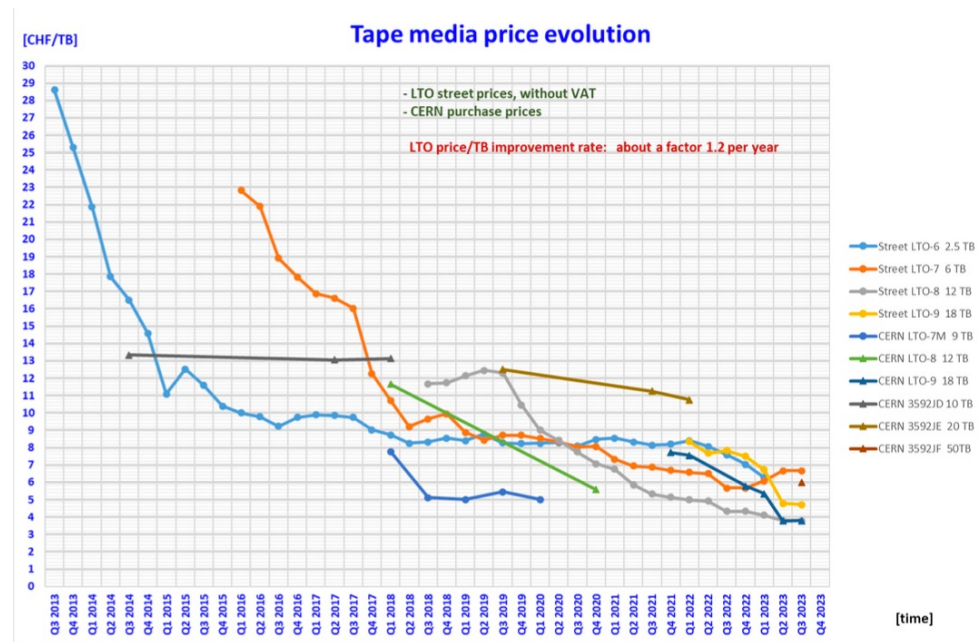
Backup

Hardware Cost Evolution

The cost of tape media decreased favourably for many years. Not true anymore for recent media (LTO9 expected to be 15% more expensive than LTO8 per TB)

Evaluating the Total Cost of Ownership for tape is tricky

- buffers, tape drives, network, specialised personnel



Hardware Cost Evolution

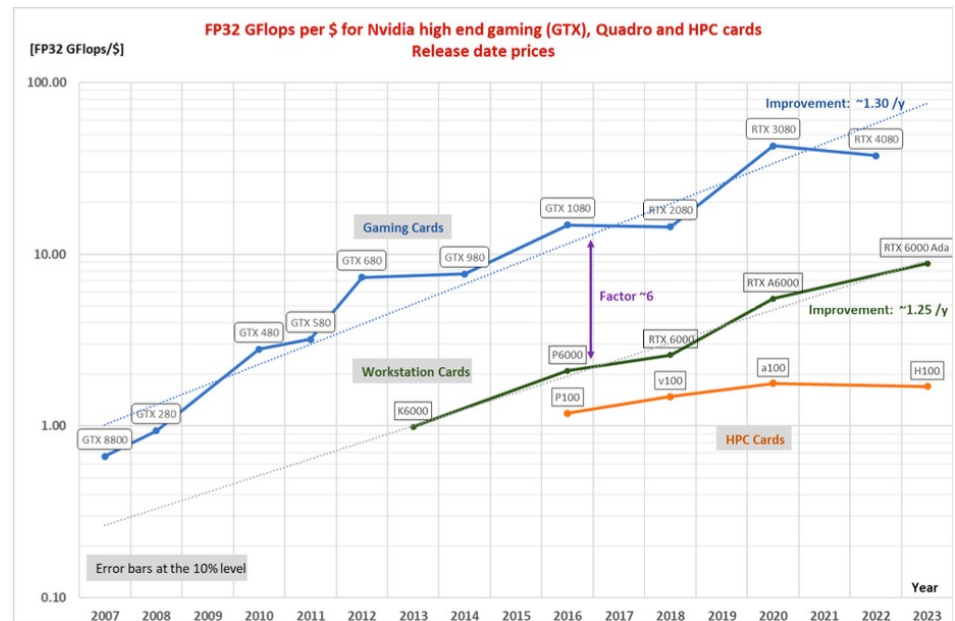
We started tracking also the GPU cost trends

- Not yet a pledged resource in WLCG but might become soon

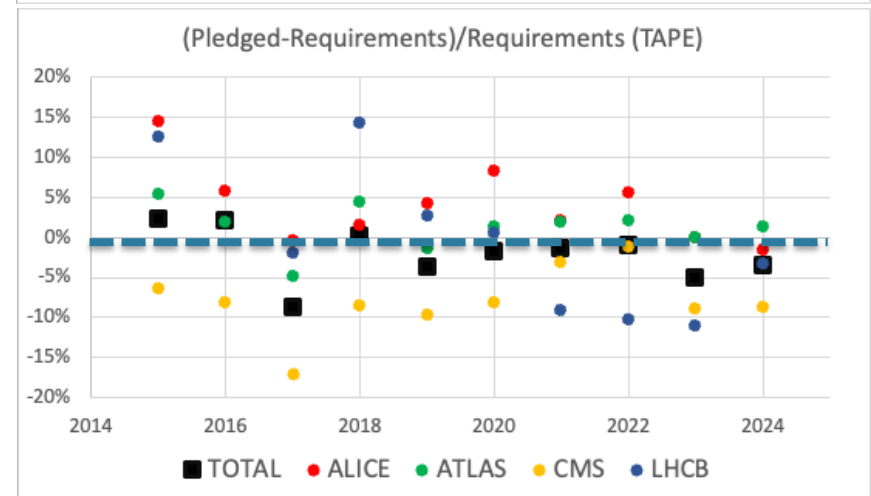
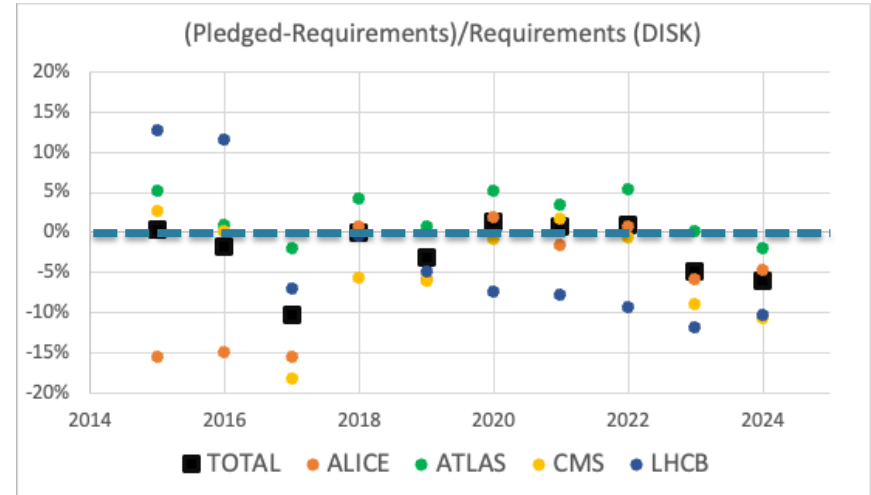
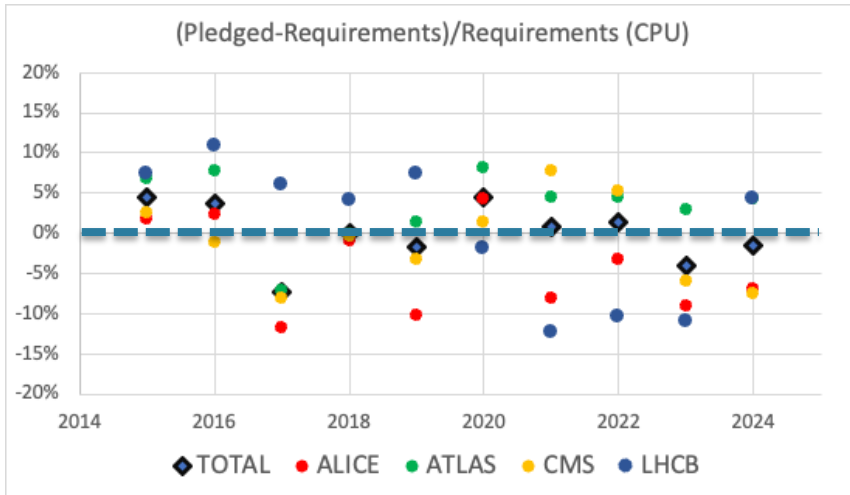
The trend in GigaFLOPS/USD is very favourable if you play videogames

The GigaFLOP/USD trend flattened for “HPC cards”

Very volatile markets, high prices and long procurement times (25+ weeks of delivery time). High demand worldwide (AI/ML/ChatGPT)



WLCG pledges: trend



Possible loss of Russian resources was partially compensated by other Funding Agencies (mostly tape). In the plots the contribution of Russian resources in 2023 and 2024 is **NOT** included.

2024 pledges

Tier	Resource Type	ALICE Balance	ATLAS Balance	CMS Balance	LHCb Balance	Total Required	Total Pledged	Global Balance
T0	CPU (kHS06)	0%	0%	0%	0%	2690	2690	0%
T1	CPU (kHS06)	-14%	0%	10%	-5%	3648	3617	-1%
T2	CPU (kHS06)	-1%	12%	-7%	23%	4421	4593	4%
T0	Disk (PB)	0%	0%	0%	0%	201	201	0%
T1	Disk (PB)	-13%	0%	-5%	-13%	418	394	-6%
T2	Disk (PB)	5%	-3%	-10%	-20%	427	407	-5%
T0	Tape (PB)	0%	0%	0%	0%	825	825	0%
T1	Tape (PB)	-4%	2%	-7%	-6%	1072	1041	-3%



Generally, the pledges match well the needs, within 5%

Run-3 overview and 2025 LHC conditions

The 2025 LHC conditions to be used at this RRB were agreed with the LPC in Dec 2023



2023 went differently than expected. The 2025 requests took this into account.

	2022 (what happened)	2023 (what happened)	2024 (planning for computing)	2025 (planning for computing)
ATLAS/CMS luminosity	~40/fb	~30/fb (exp. <90/fb)	<110/fb	<120/fb
ATLAS/CMS pile-up	~46	~62	<62 (peak 65)	<62 (peak 65)
LHCb luminosity	~0.1/fb	~0.15/fb (exp <10/fb)	15/fb	15/fb
ALICE luminosity (pp)	~20/pb	~10/pb (exp 50/pb)	100/pb	100/pb
Running time pp	~4 M sec	~2.5 M sec (exp. 4.5M sec)	5.2 M sec	6.3 M sec
Running time ions	No HI run	~1.6 M sec (exp. 2.4M sec)	1.7 M sec	1.4 M sec

Recent decision to anticipate 4 weeks of pp run from 2025 into 2024. Too late to fold this into the requests and to deal with it properly in procurement.

WLCG resource needs compared to Flat Budget models

We used the findings from the study on the previous slide to understand

- The difference between the “15% Flat Budget model” (more theoretical) and a “Measured Flat Budget model” (based on the real purchase power)
- How the WLCG resource needs compare to the two models

