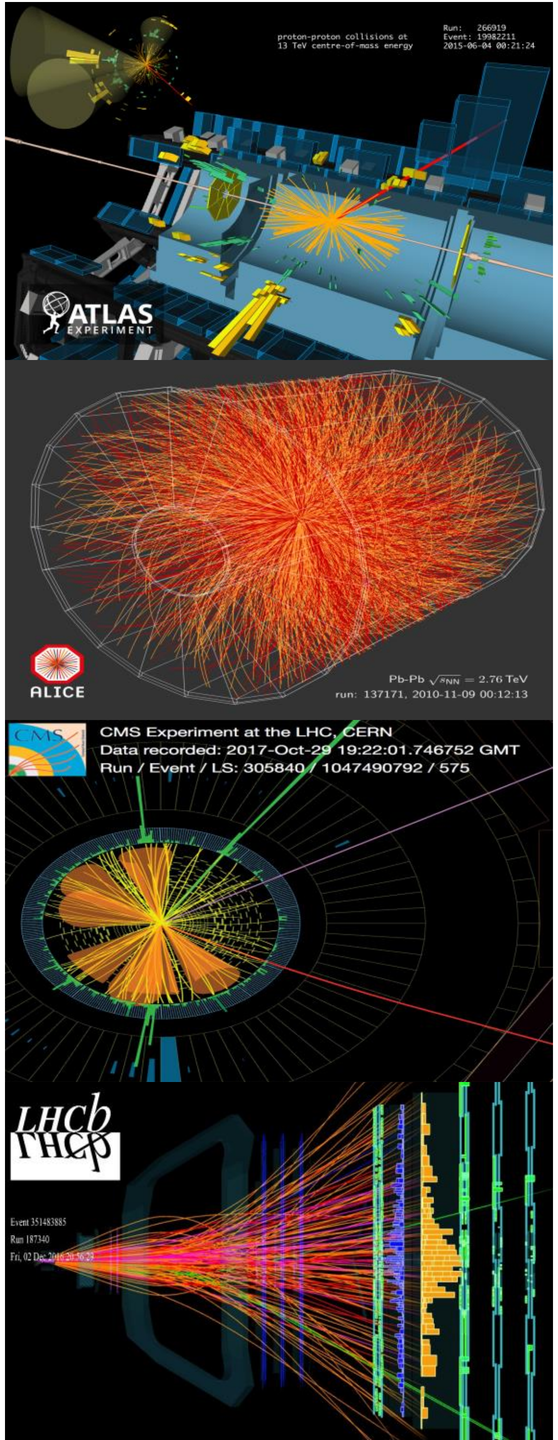
TRIUMF

The Worldwide LHC Computing Grid

Reda Tafirout

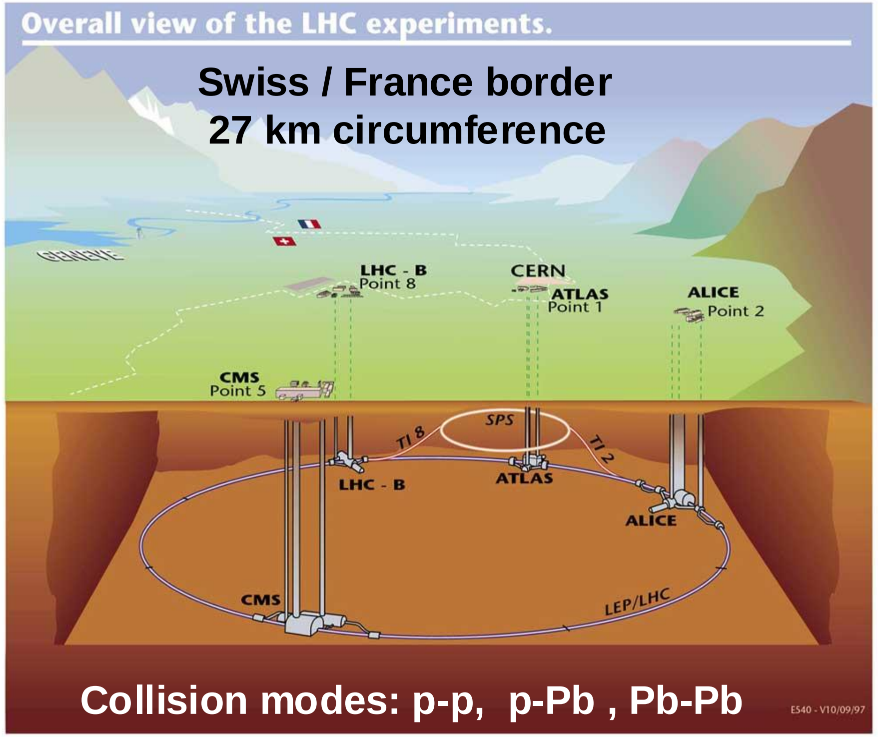
TRIUMF Science Week 2024



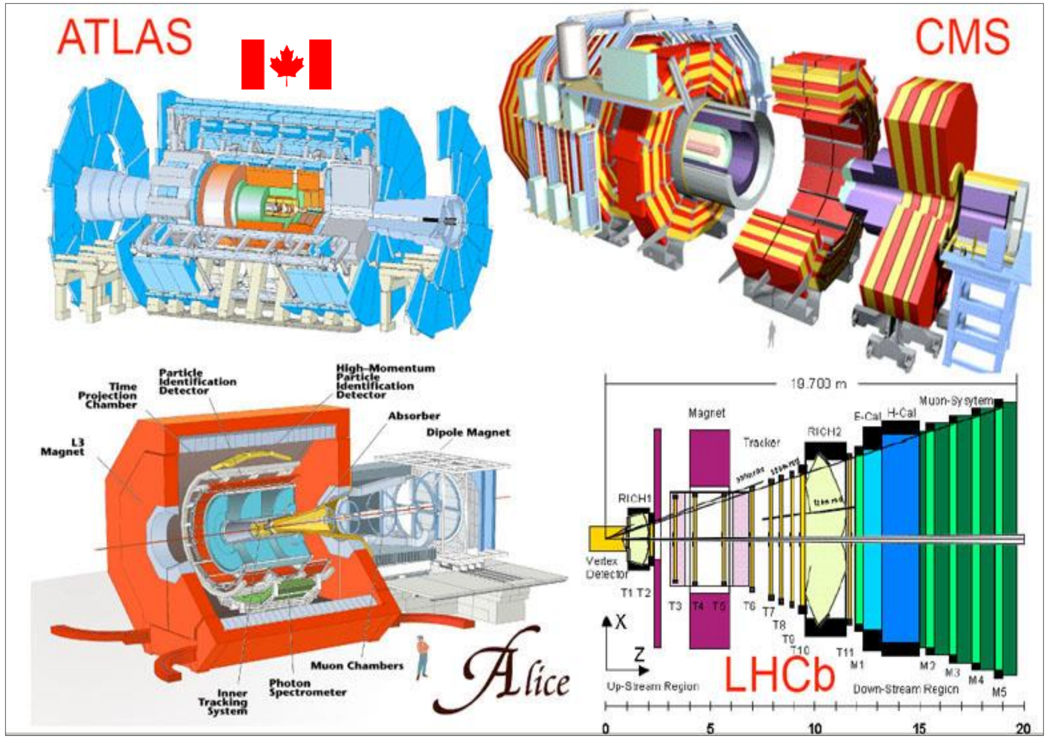
Discovery, accelerated

LHC Scientific Program Tools

Powerful Particle Accelerator



Sophisticated Detectors



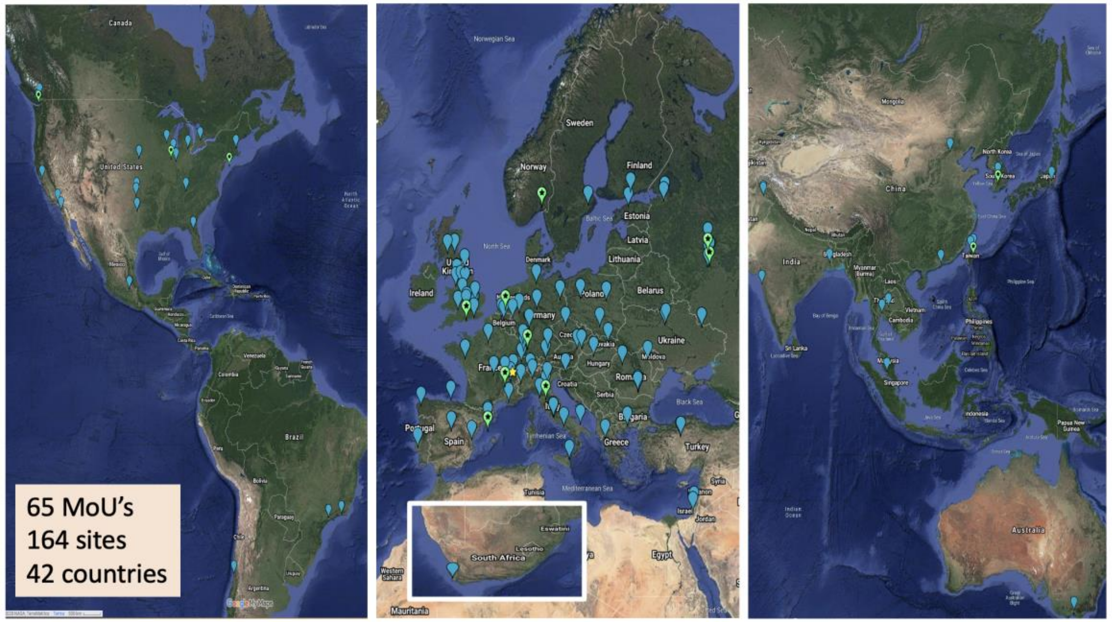
Large-scale Distributed Computing

WLCG
Worldwide LHC Computing Grid

WLCG: global collaboration & effort

- Coordinated resource sharing across multiple organizations
- 164 sites in 42 countries (tiered structure)
- MoU's between parties participating in the project
  - CERN (Tier-0), Tier-1 centres, Tier-2 centres
  - Baseline services & roles
  - Service levels & up-times
  - Management structure & Grid operations
- Pledged resources updated yearly based on experiments requirements.

WLCG
Worldwide LHC Computing Grid



2024 pledge: ~2 EB tape, 1 EB disk, 550 kcores

Tier-1: TRIUMF / SFU

Tier-2: Alliance sites (Victoria, SFU, Waterloo)

TRIUMF

SFU SIMON FRASER UNIVERSITY

Digital Research Alliance of Canada

WLCG: many key components

- Middleware services
- Information system (sites configuration)
- Security & authentication: IGTF (CA's)
- VO management service
- File transfer service & network provisioning
- Workload management
- Data management
- Accounting & monitoring
- Constant evolution while ensuring continuity

IGTF International Grid Trust Federation AP | EU | TAG

dCache

CRIC Computing Resource Information Catalog

FTS

HTCondor High Throughput Computing

EOS

GGUS

slurm workload manager

ARC

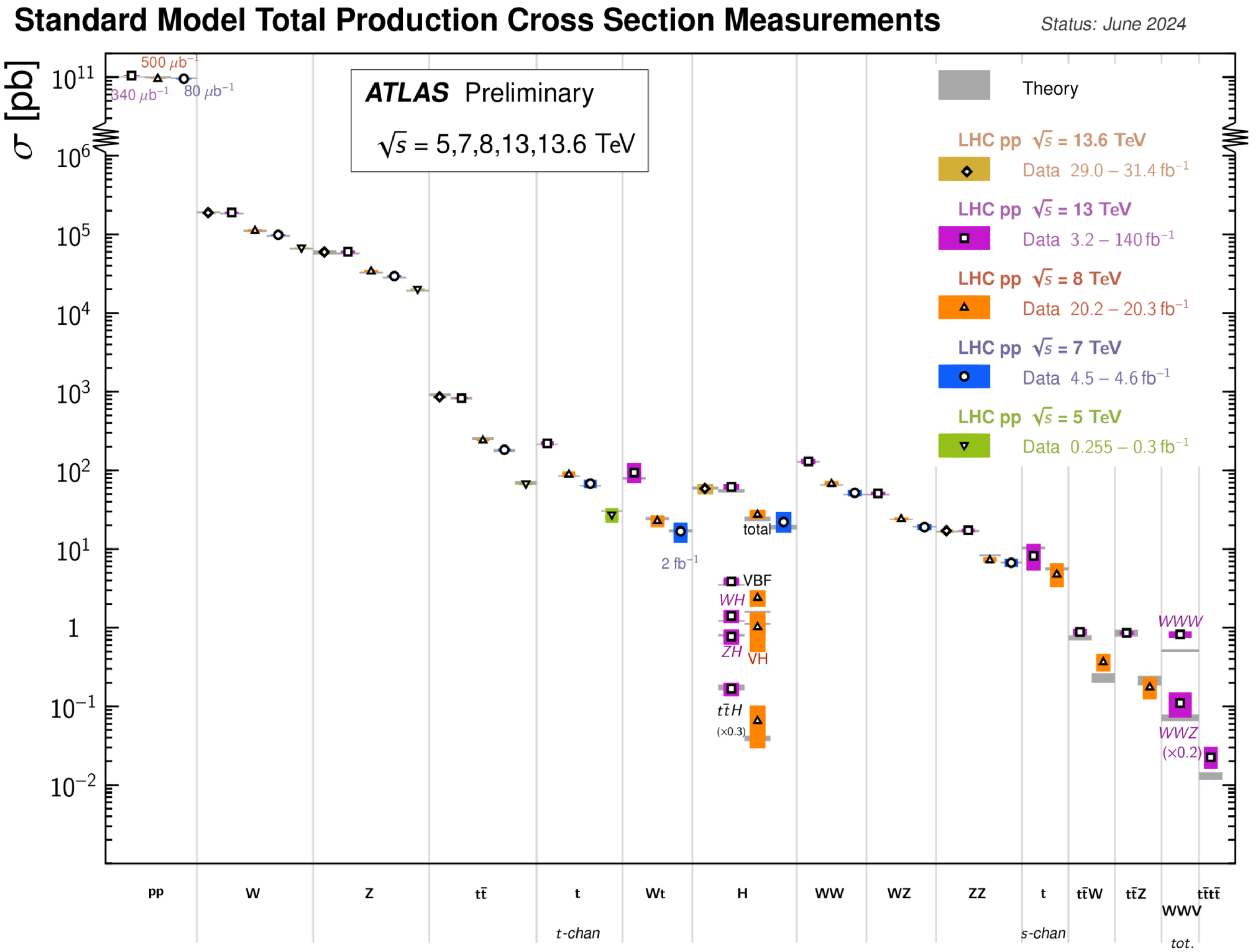
DIRAC

PanDA

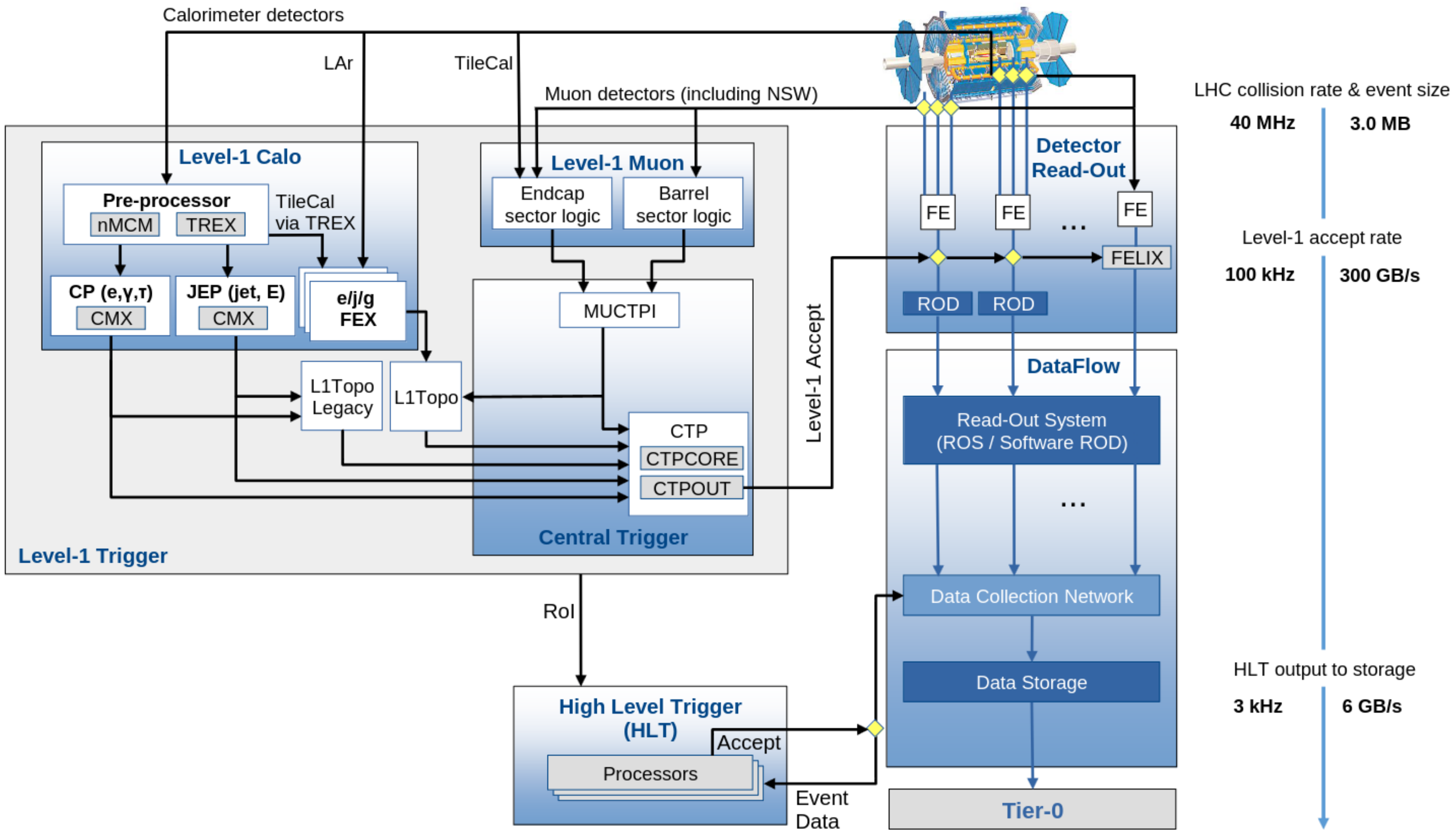
SCIENTIFIC DATA MANAGEMENT RUCIO

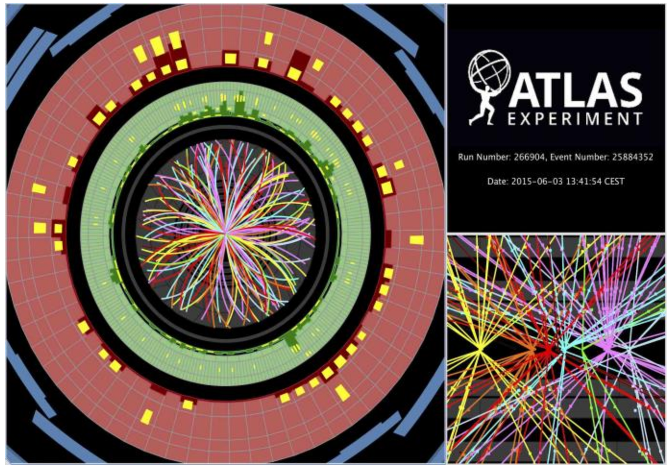
CernVM-FS

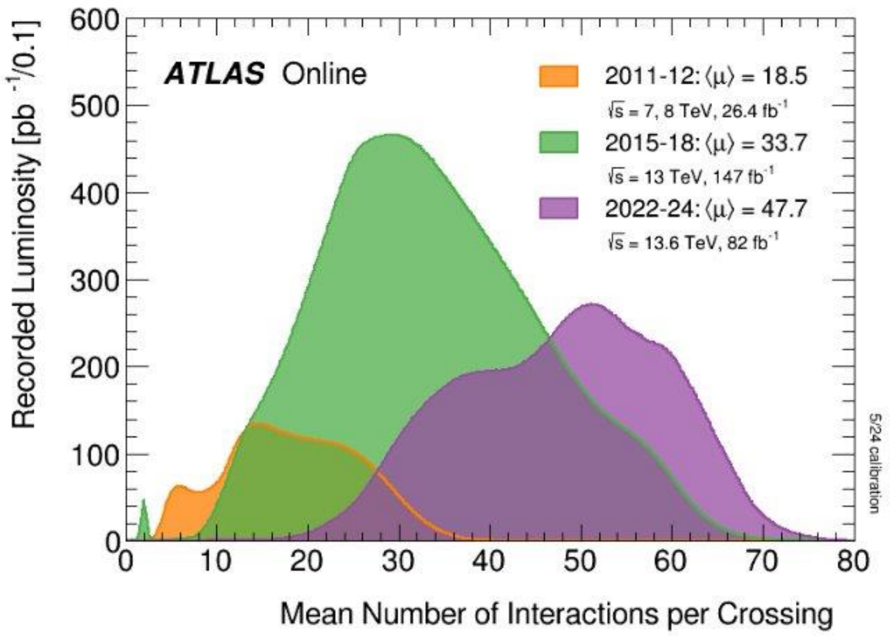
- **A lot of data needed**
- **SM measurements**
- **BSM searches**
- **Large-scale simulations needed**



ATLAS TDAQ Chain

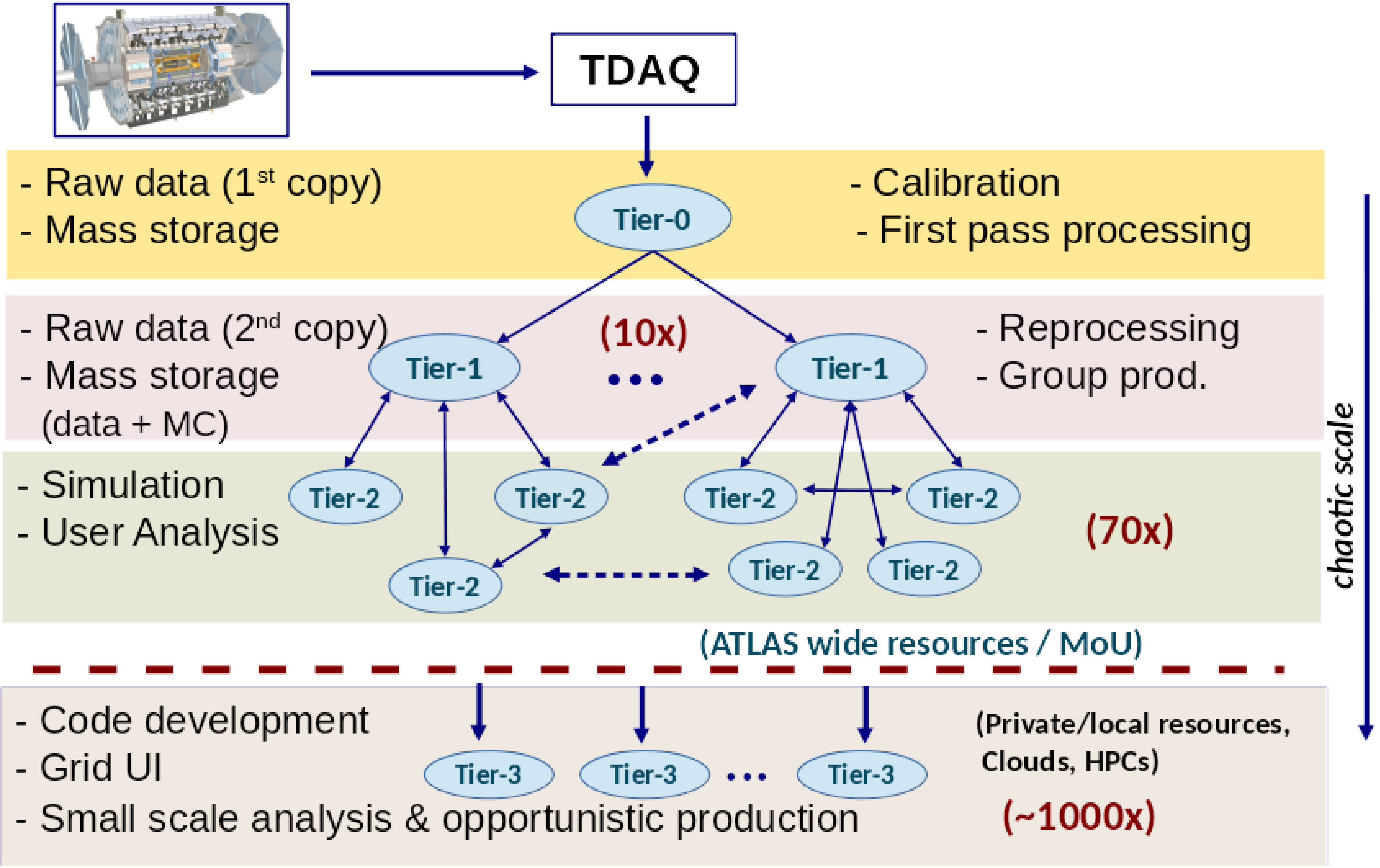




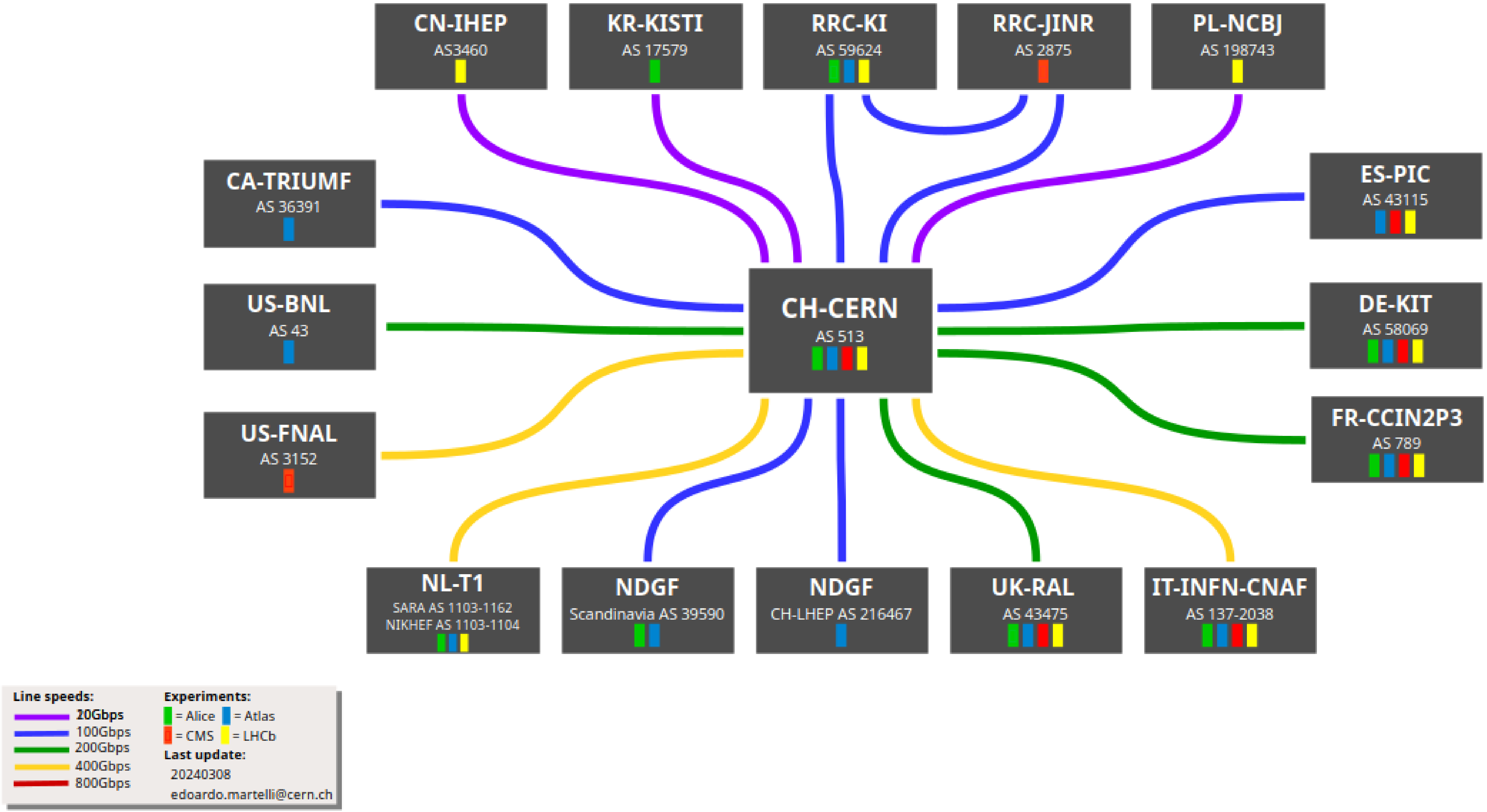


- From 40 MHz beam crossing to ~3 kHz (storage at T0)
- HLT farm scale: ~100,000 cores (dedicated)

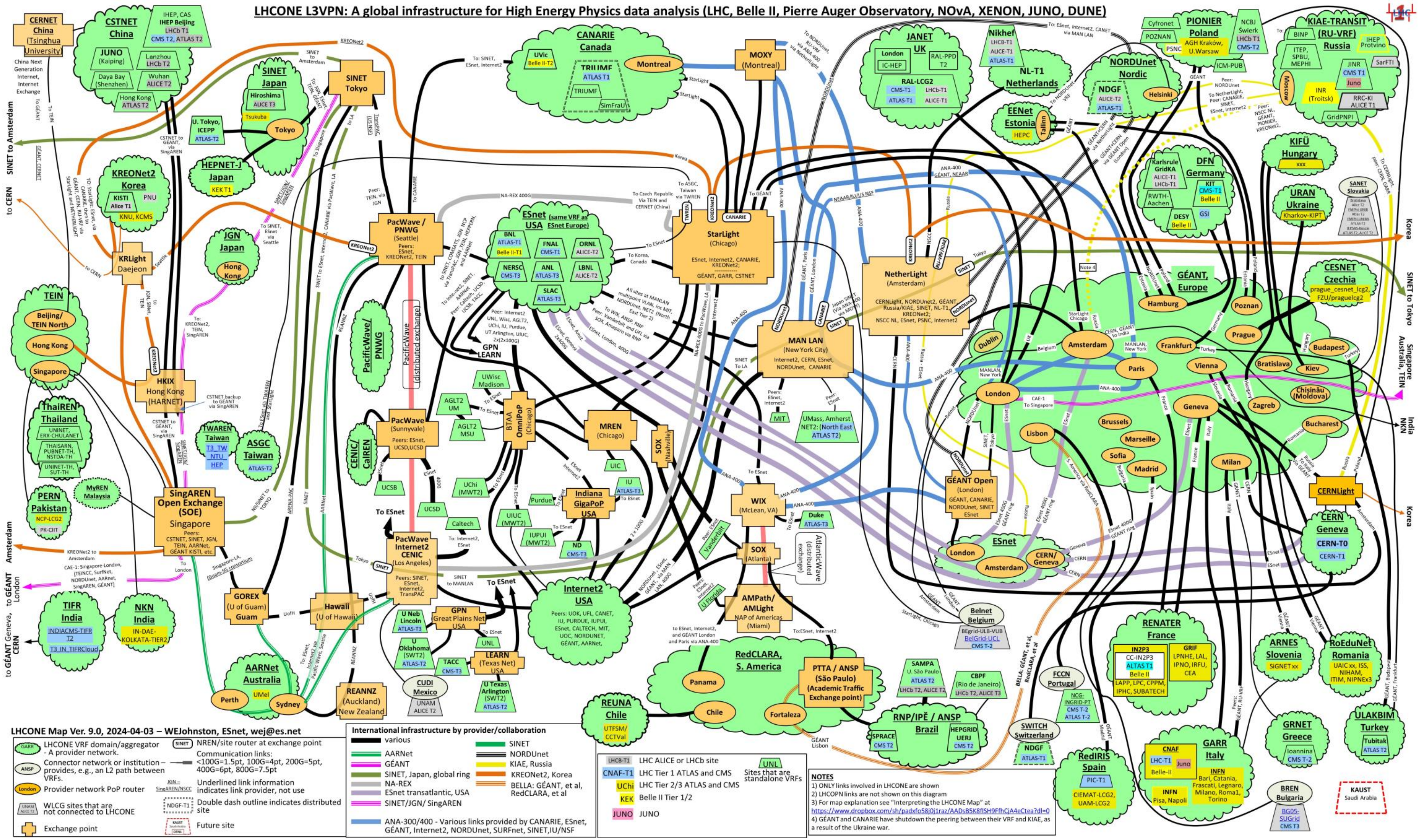
ATLAS Computing Model



LHC Optical Private Network (LHCOPN)

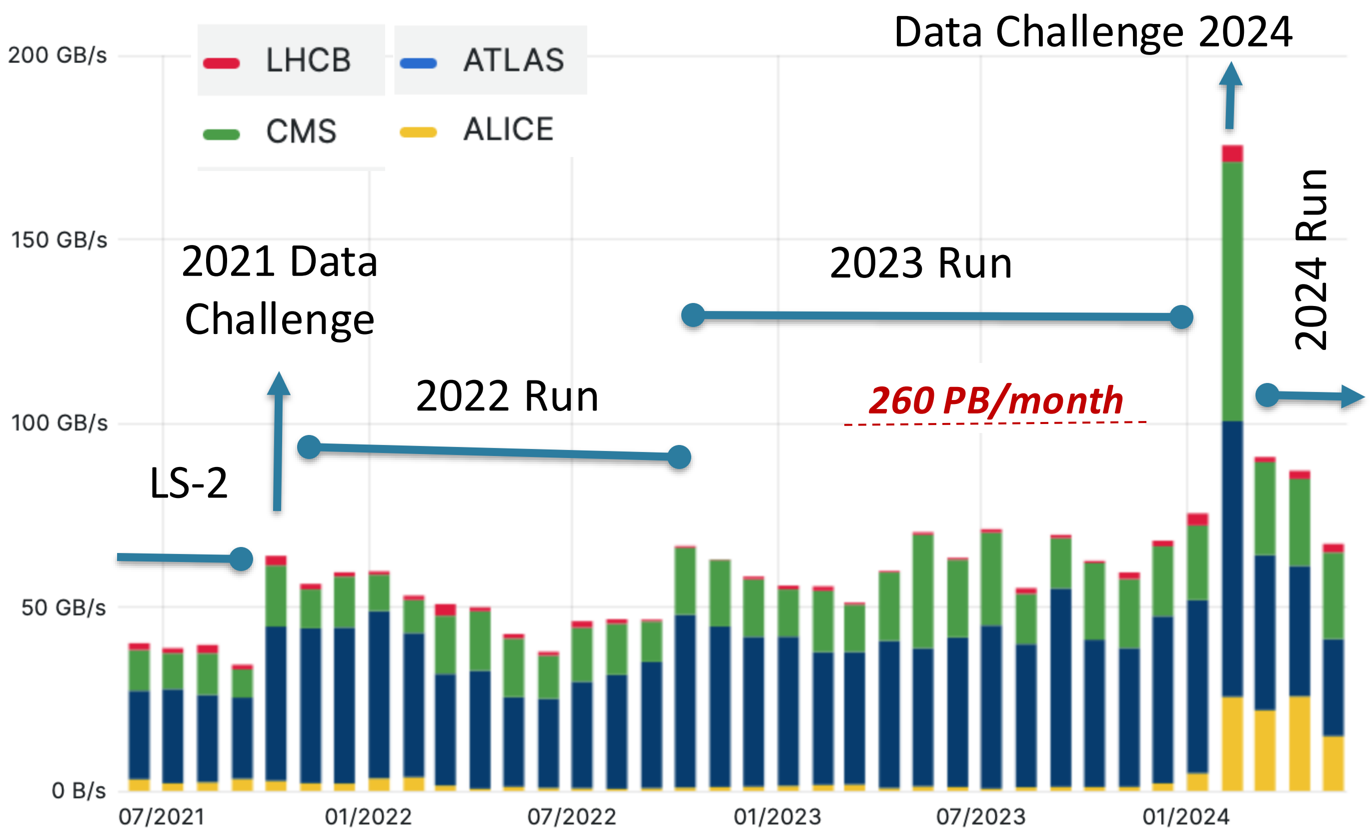


LHC Open Network Environment (LHCONE)



WLCG: global data transfers

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



WLCG supports +40% more transfers since LS-2

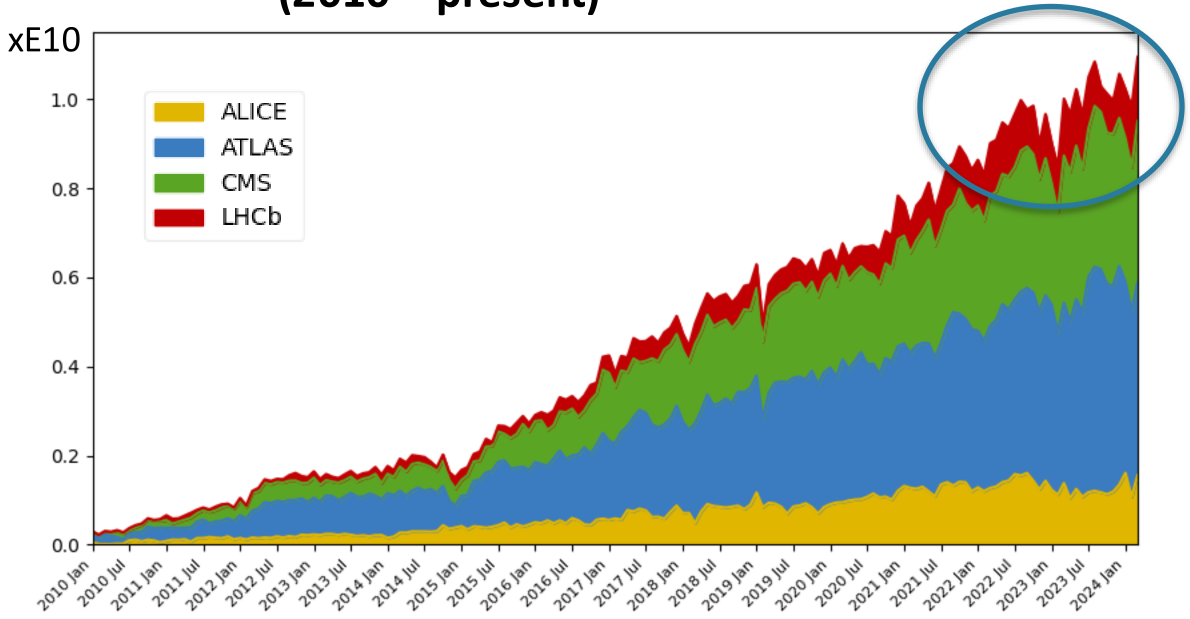
Further scalability (x5) demonstrated in the Data Challenge

No strain to the services

WLCG: cpu delivered globally

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

CPU delivered: HS23 hours/month (2010 – present)

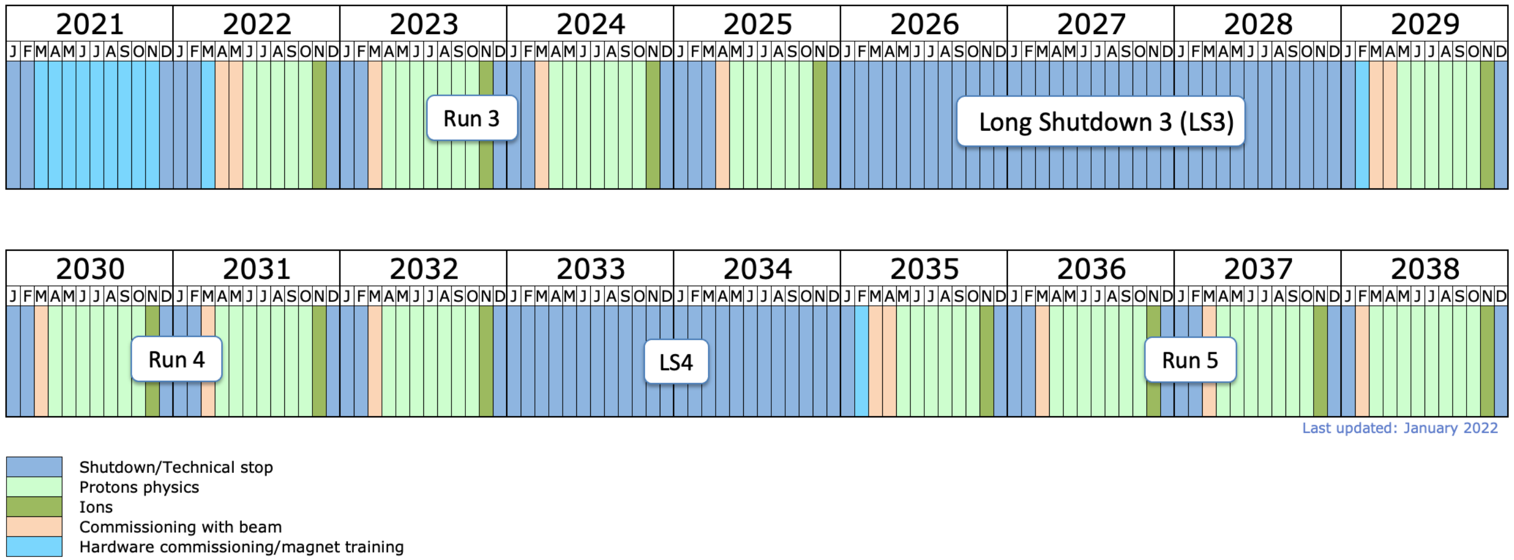


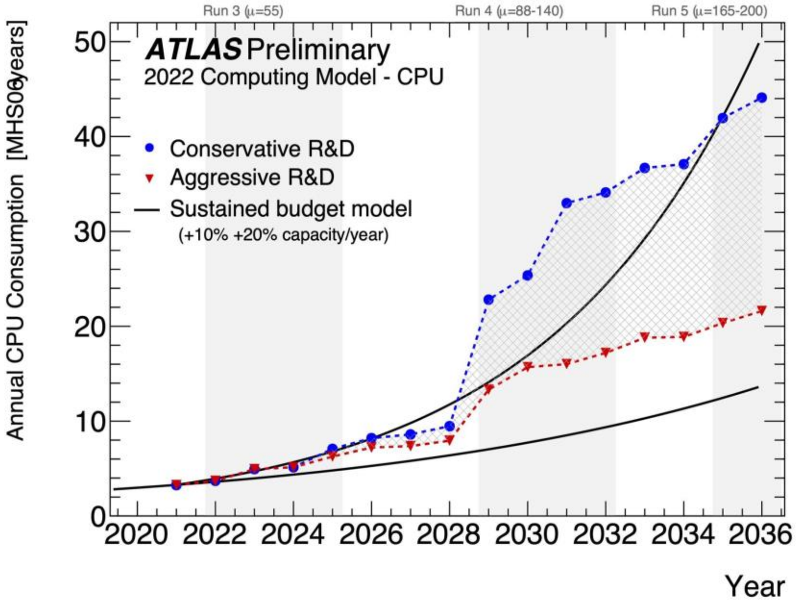
- Drop in winter 2022/23: energy crisis in Europe with high natural gas costs (supply issue from Russia)
- Back to "normal" now

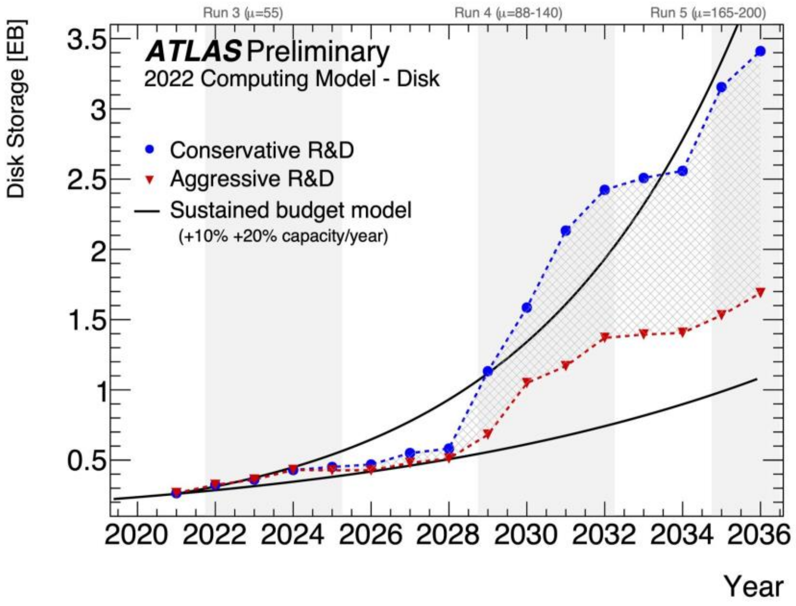
- HEPScore 2023 (HS23) is a common cpu benchmark unit (used in WLCG pledges & accounting)

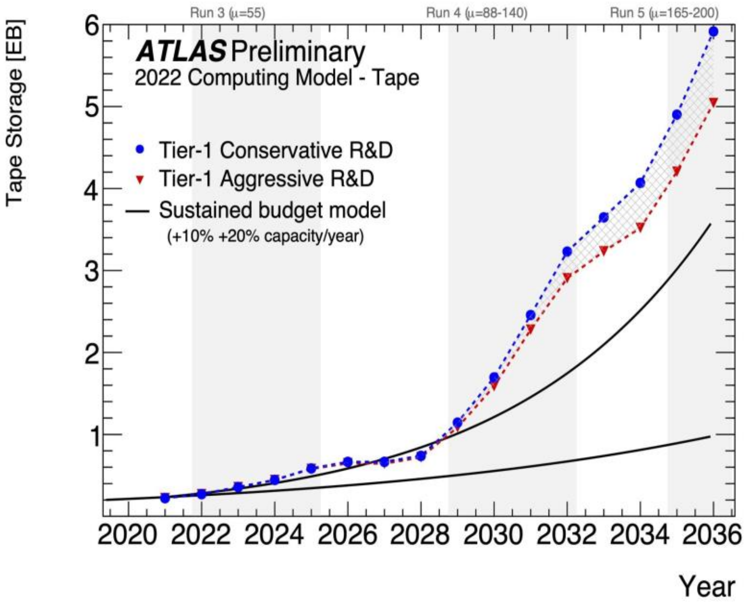
ATLAS Distributed Computing Needs & HL-LHC Challenges

- ~10x data rate increase in Run-4
- Flat budget model "problematic"
- Significant development effort is required:
  - Improve software performance
  - Leverage modern architectures
- Data challenges planned









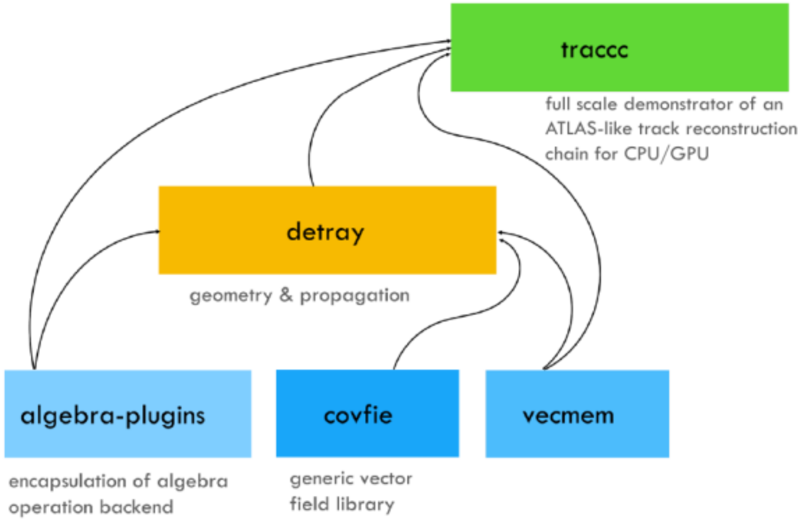
ATLAS Software & Computing Roadmap for the HL-LHC

- Roadmap has several components dealing with various topics:
  - Network infrastructure ready for Run 4
  - Detector Description, Simulation and Digitization projects
  - HL-LHC datasets replicas and versions management
  - Core Software and Heterogeneous Computing / Accelerators
  - etc.

- ATLAS Heterogeneous Computing & Accelerators Forum established recently

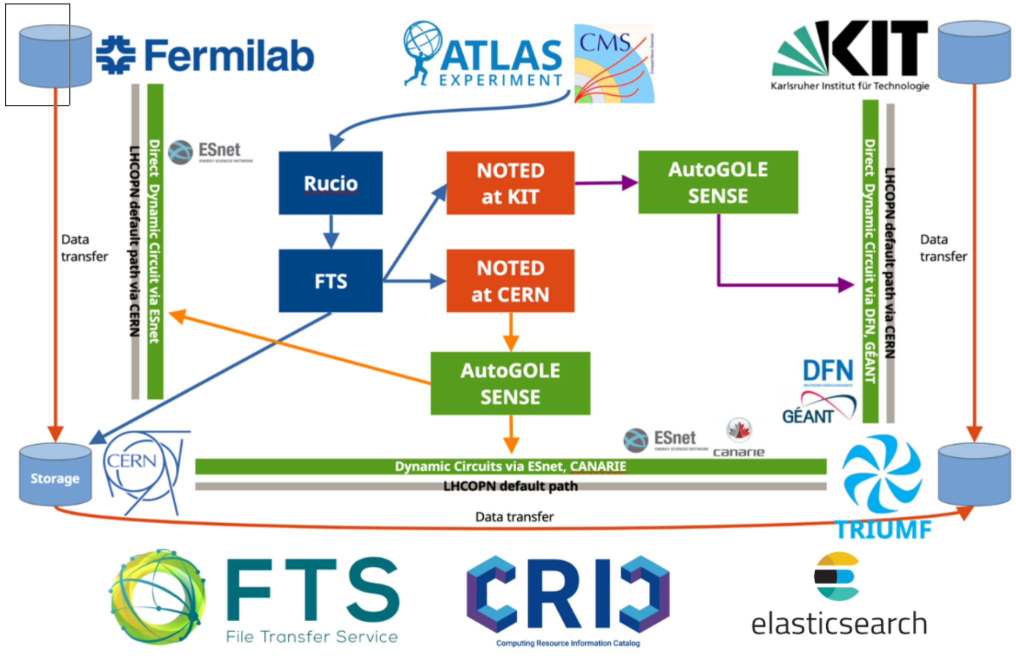
- To tackle the combinatorics in a high luminosity environment, investigate tracking on GPU. For this to succeed:
  - define a suitable Event Data Model,
  - develop a toolchain that supports e.g. CUDA kernels
  - provide GPU friendly implementations of the geometry and magnetic field.

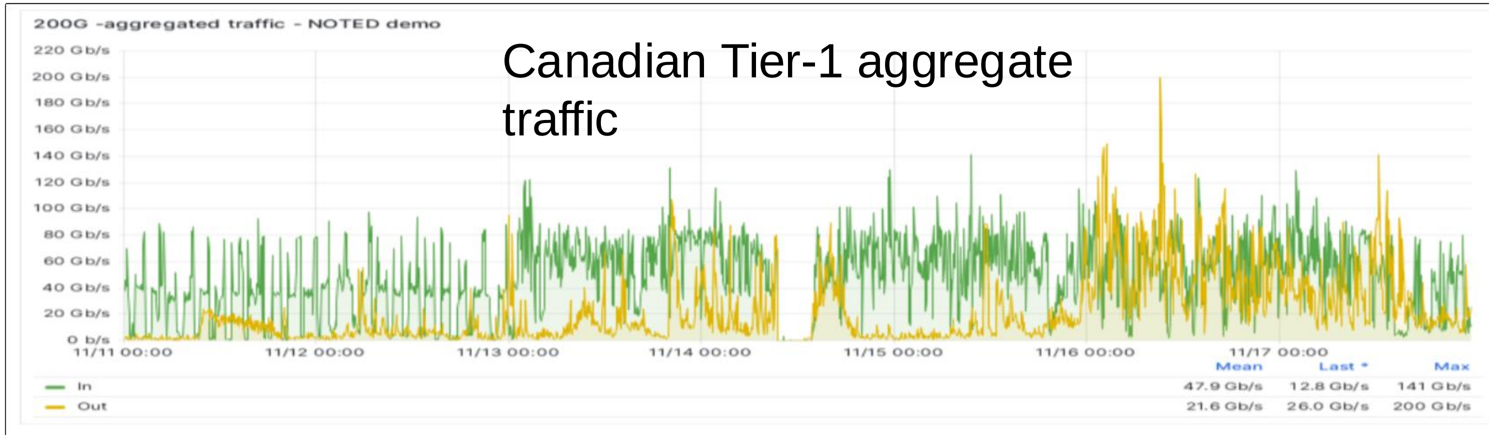
ACTS - A Common Tracking Software

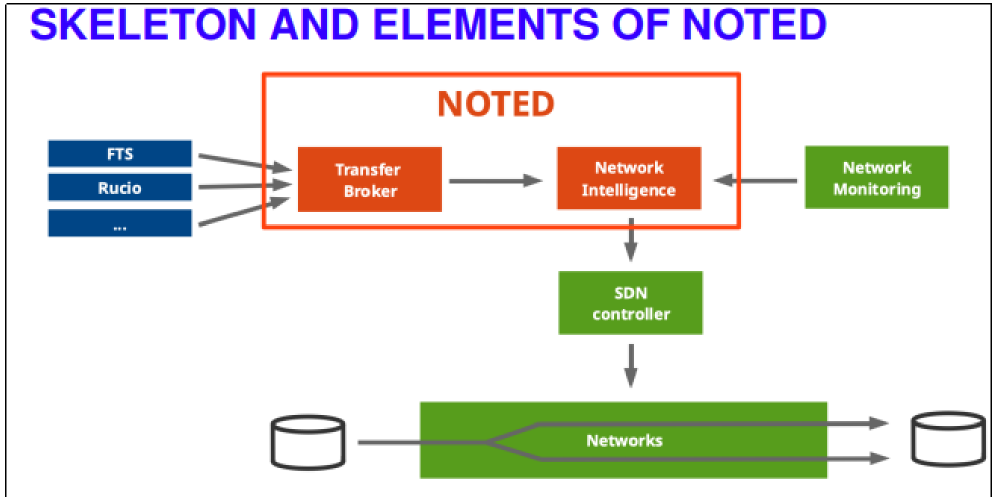


Dynamic network provisioning

- Demonstration of software defined networks and dynamic circuit provisioning based on demand (Supercomputing 2023 conference)
- Collaboration between TRIUMF, CERN, FNAL, KIT and network providers
- Traffic from both ATLAS & CMS
- Also work on network packets marking (scientific tags) in collaboration with HEPNet Canada



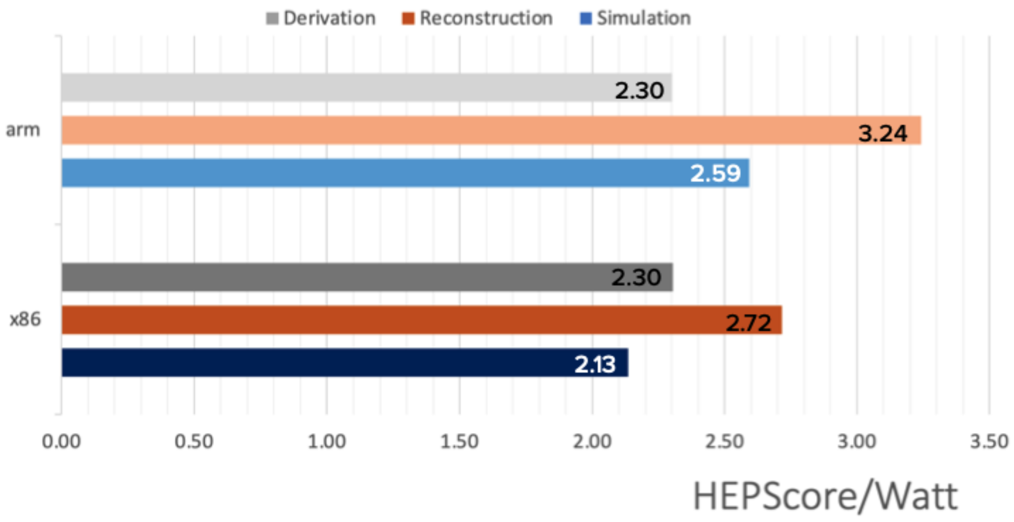


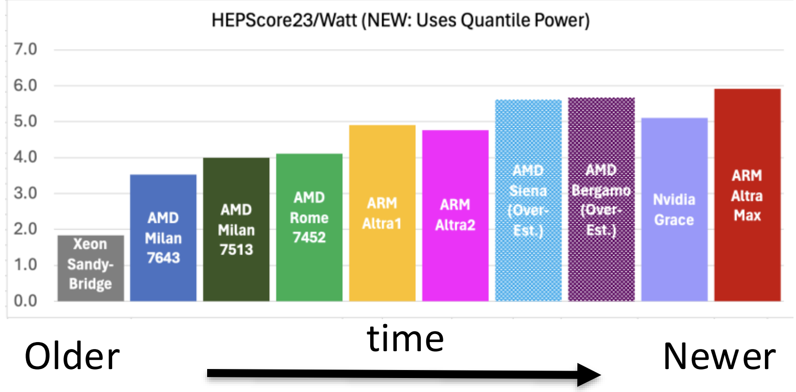


ARM CPUs & GPUs

ARM CPUs process more events/Watt wrt x86

- Studies done with hardware Glasgow
- Differences between workflows
- Modern AMDs perform similarly to ARM

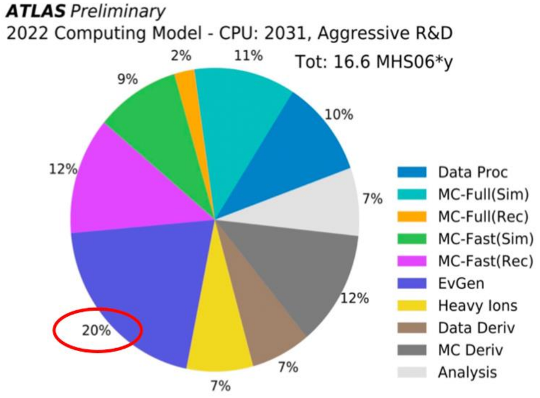




Experiment workflows ported to ARM

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

GPUs (in HPCs): massive parallelization, Machine Learning toolkits



		madevent		
CUDA grid size		8192		
$gg \to t\bar{t}ggg$	MEs precision	$t_{TOT} = t_{Mad} + t_{MEs}$ [sec]	N_{events}/t_{TOT} [events/sec]	N_{events}/t_{MEs} [MEs/sec]
Fortran	double	1228.2 = 5.0 + 1223.2	7.34E1 (=1.0)	7.37E1 (=1.0)
CUDA	double	19.6 = 7.4 + 12.1	4.61E3 (x63)	7.44E3 (x100)
CUDA	float	11.7 = 6.2 + 5.4	7.73E3 (x105)	1.66E4 (x224)
CUDA	mixed	16.5 = 7.0 + 9.6	5.45E3 (x74)	9.43E3 (x128)

NVidia V100, Cuda 11.7, gcc 11.2

- **WLCG discussing about pledging ARM CPUs. Resources already available at various sites**

Conclusion & outlook

- WLCG: global infrastructure developed and operated over the last two decades
- Notable achievement: needs of LHC experiments successfully met
- Recent WLCG strategy document developed to tackle key areas:
  - Technical evolution
  - Financial sustainability
  - Heterogeneous grid infrastructure
  - Interaction with other communities with similar challenges
- The HL-LHC era will be a challenging computing environment
  - Need to ensure sustained innovation and development while ensuring continuing operations

WLCG
Worldwide LHC Computing Grid

TRIUMF

Thank you

Merci

www.triumf.ca

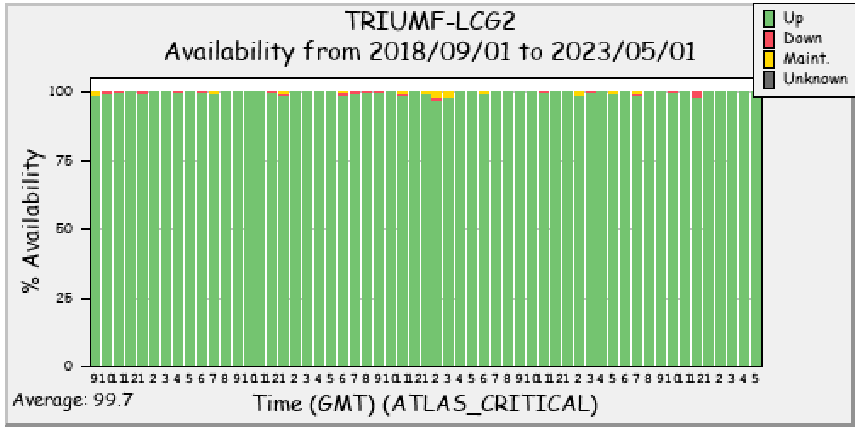
Follow us **@TRIUMFLab**

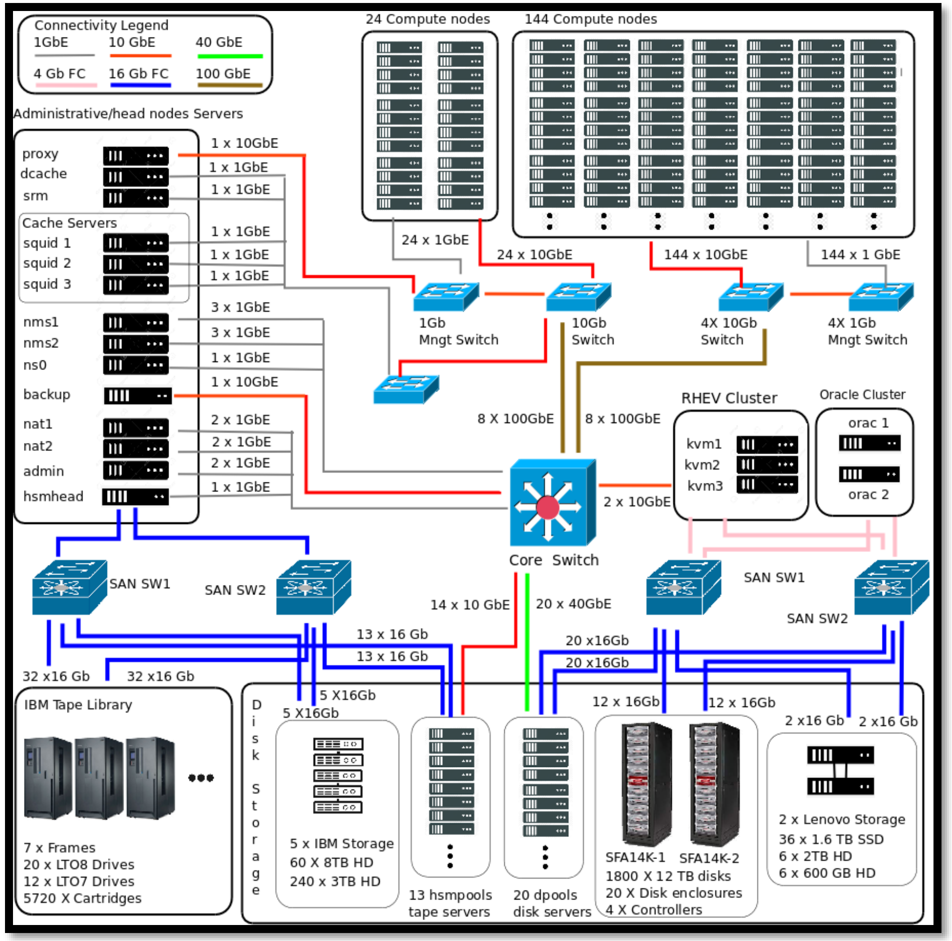
Discovery, accelerated

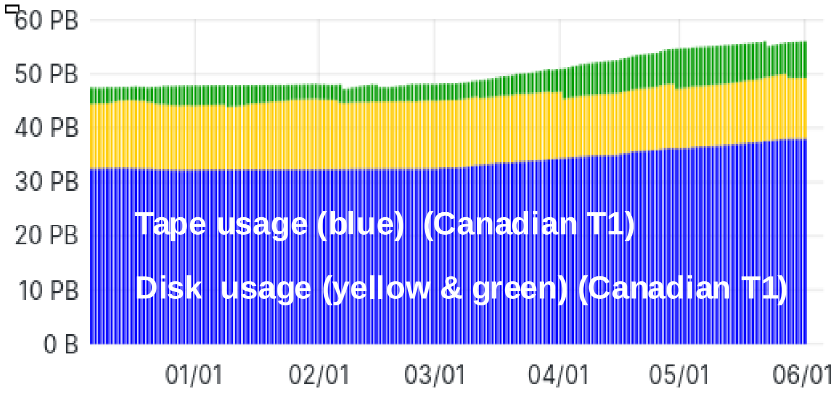
Additional Material

Canadian ATLAS Tier-1 Centre

- Dedicated facility operated 24/7 (per WLCG MoU)
- Key player and contributor within ATLAS Distributed Computing Operations:
  - Availability, Reliability, Scalability & Performance
  - Critical user support for the entire ATLAS collaboration
- Data storage, data processing, simulations and user analysis in a highly secure environment
- Initially located at TRIUMF since 2007
- Transitioned to SFU in 2018, co-located with Cedar/Alliance; continues to be under the perview of TRIUMF
- Operated as federation (new + old site) since 2017
- Current capacity: 9,300 cores ; 17 PB disk ; 46 PB tape

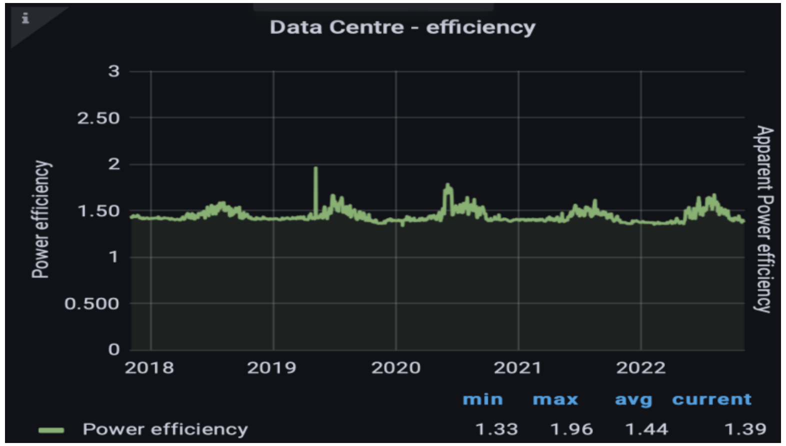






Energy efficiency

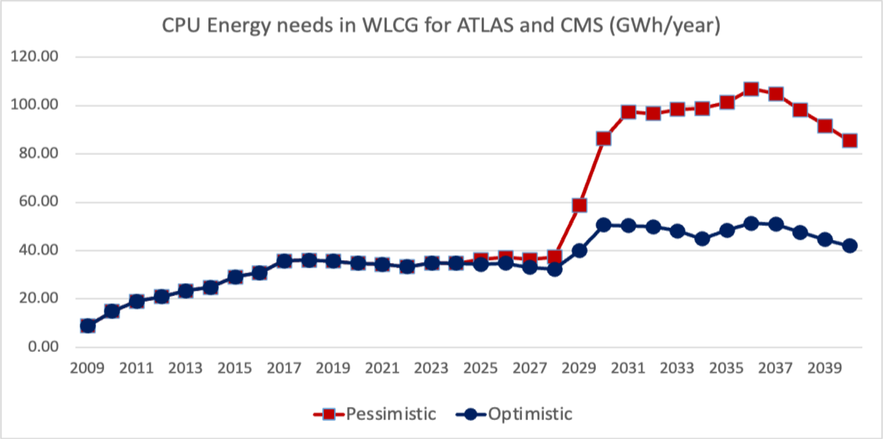
The electricity costs have been an unexpected development in the last couple of years. Environmental impact needs proper addressing!



What to do:

- Improve software performance
- Leverage modern architectures
- Invest in the facilities

There is no magic wand, however.



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.

RNTuple

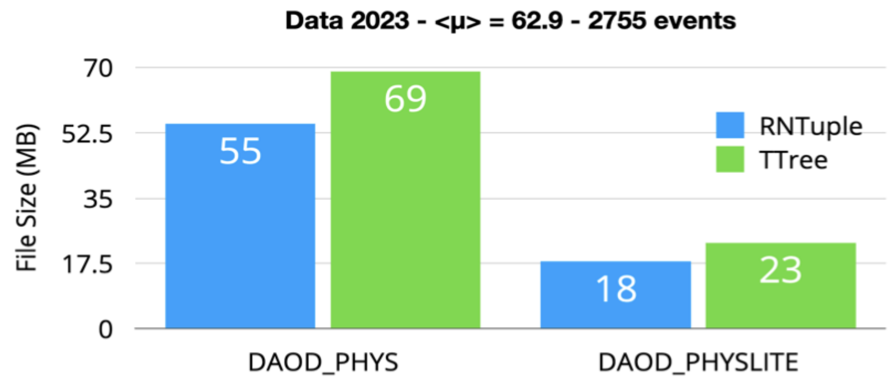
RNTuple is the successor of TTree, the ROOT columnar storage technology

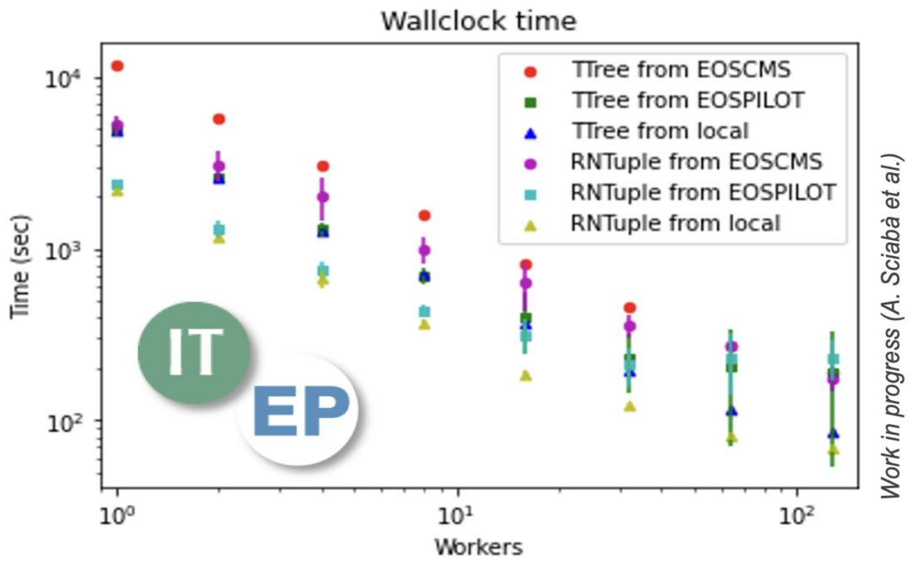
Examples of recent commissioning progress:

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

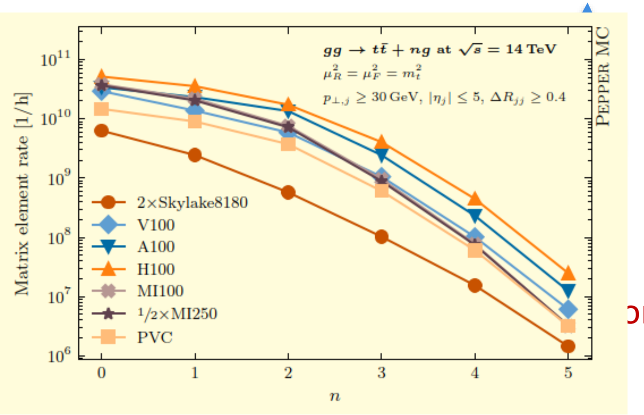




Event Generators

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

CPU + GPU

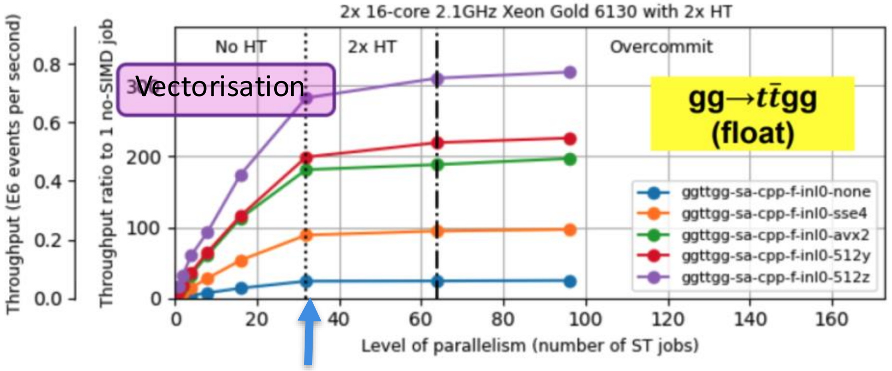


only

Sherpa gg->tt+ng

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

Available for production



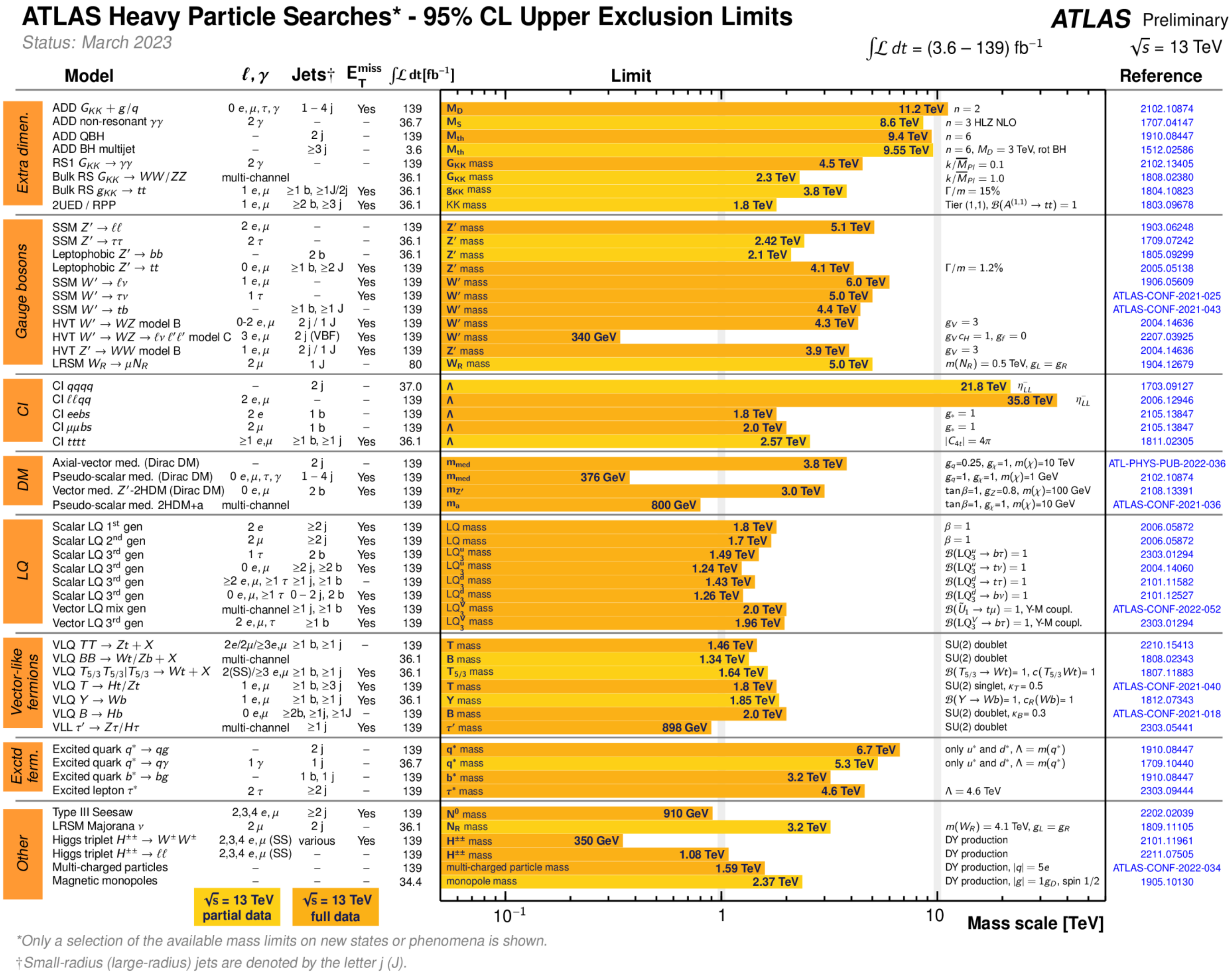
No Vectorisation

Madgraph gg->tt+ng (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

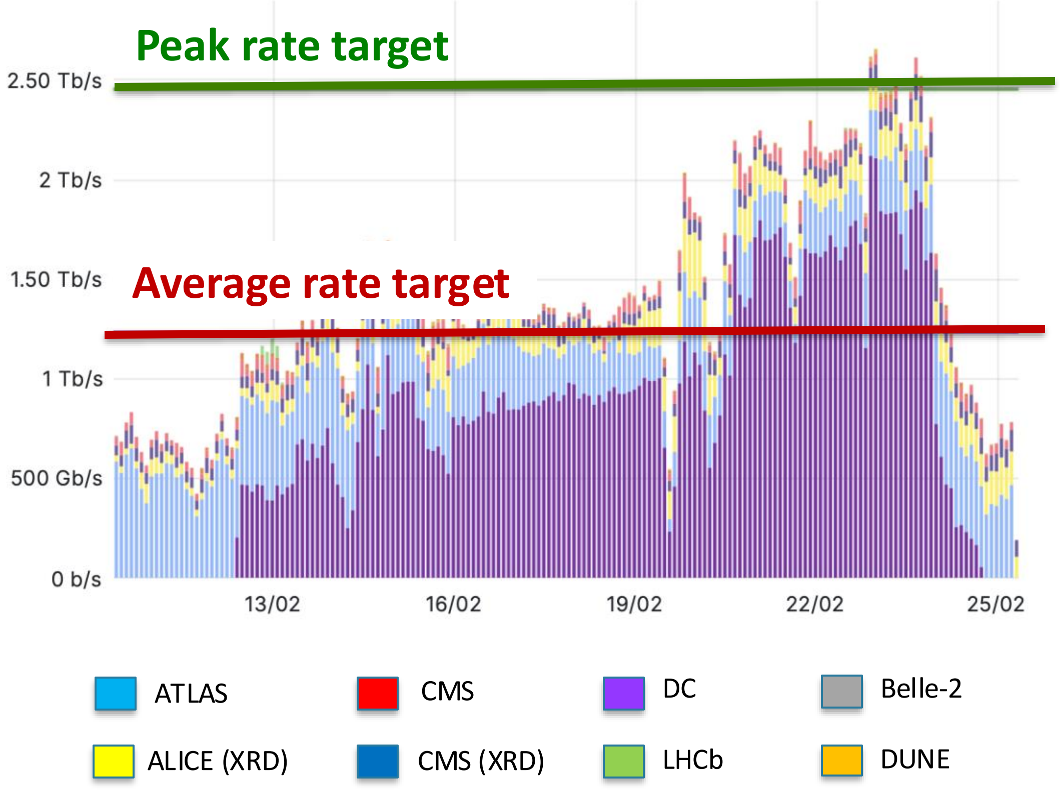
GPU-related work brings immediate benefits also on CPUs

ATLAS BSM searches

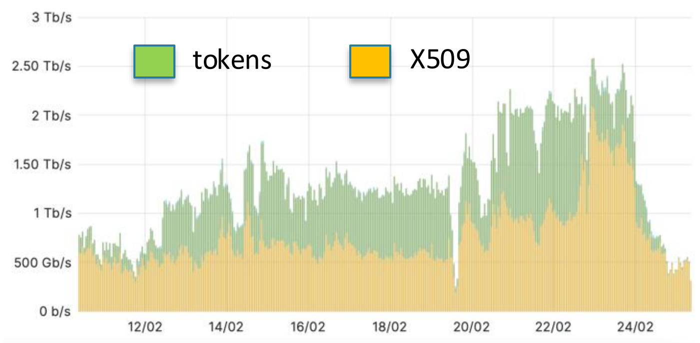


Data Challenge 2024 - Highlights

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



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



WLCG services successfully supports DUNE and Belle-2 computing models

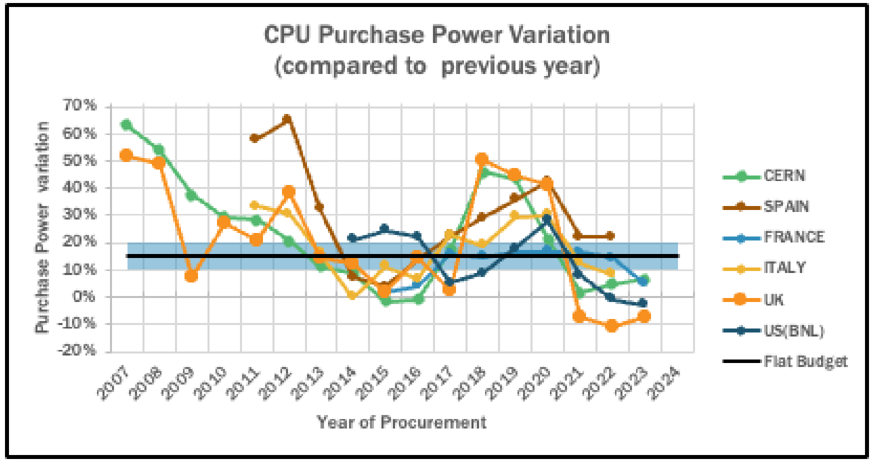
Hardware Cost Evolution

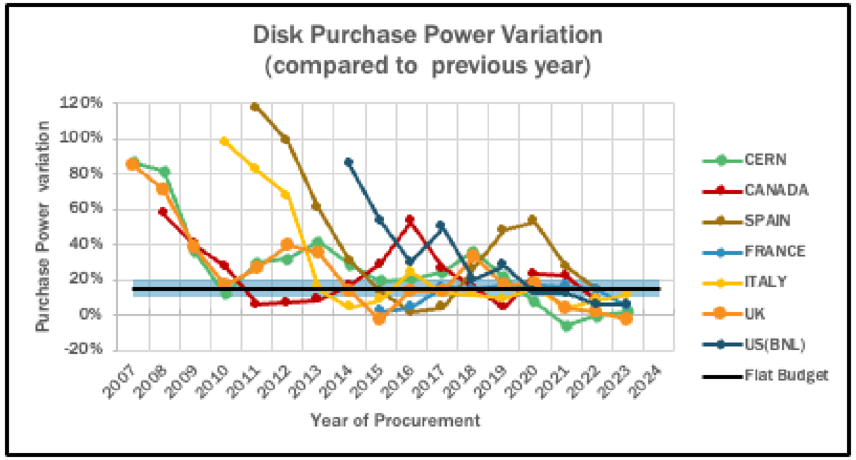
The WLCG "flat budget model" assumption: +15% CPU, disk and tape every year **with the same level of funding**

We now monitor the HW trends in many countires. 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%



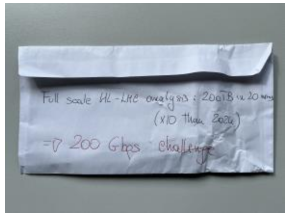


IRIS-HEP: analysis 200Gbps (Grand) Challenge

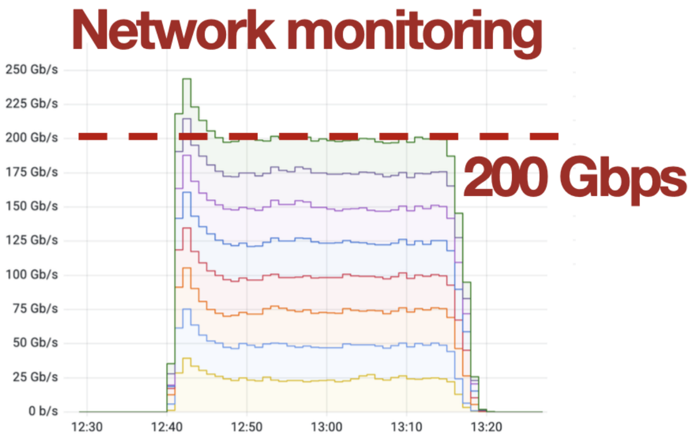
Launched by IRIS-HEP to commission analysis capabilities for HL-LHC

- 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?)



Based on the IRIS-HEP toolkit