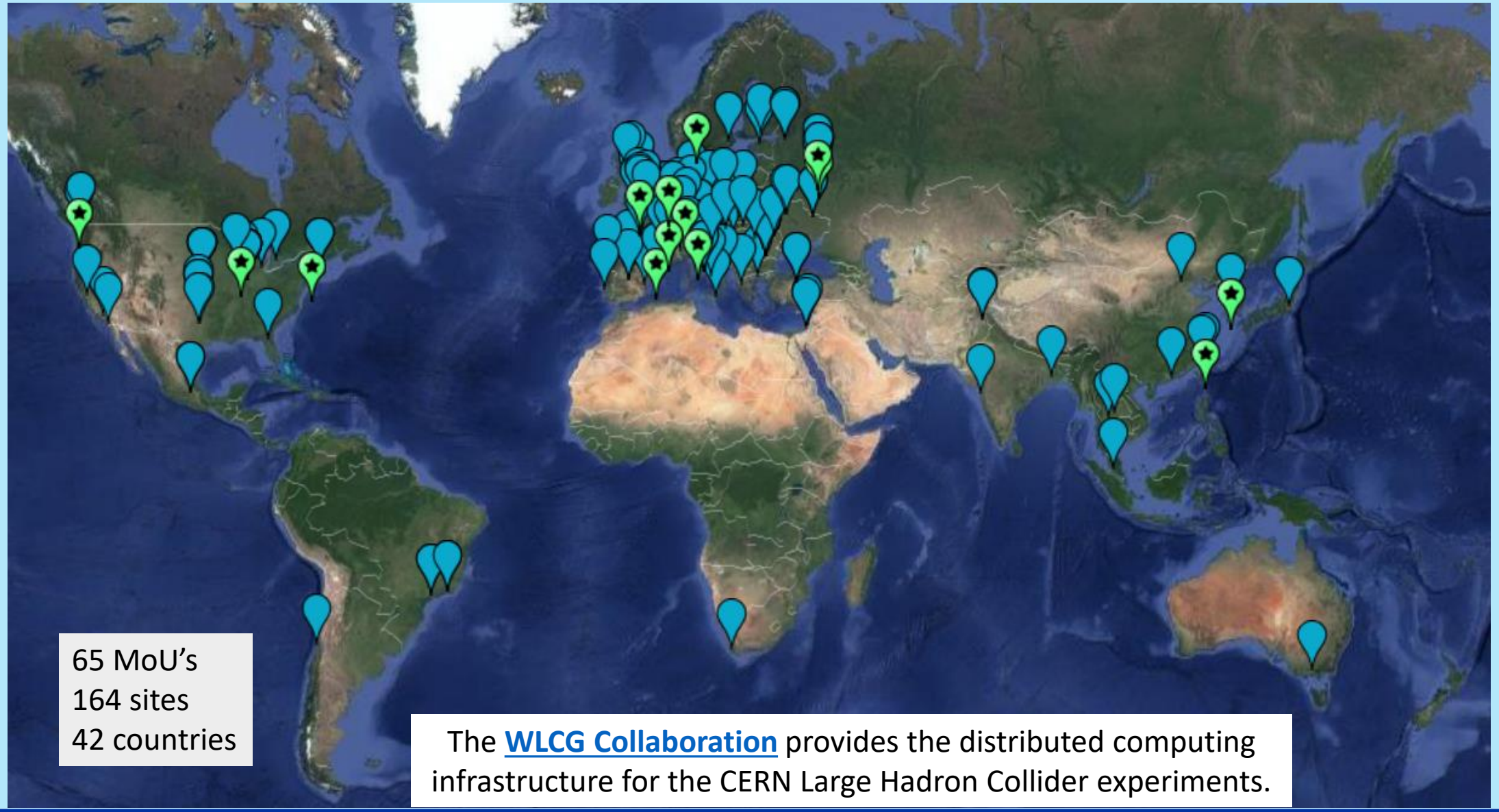
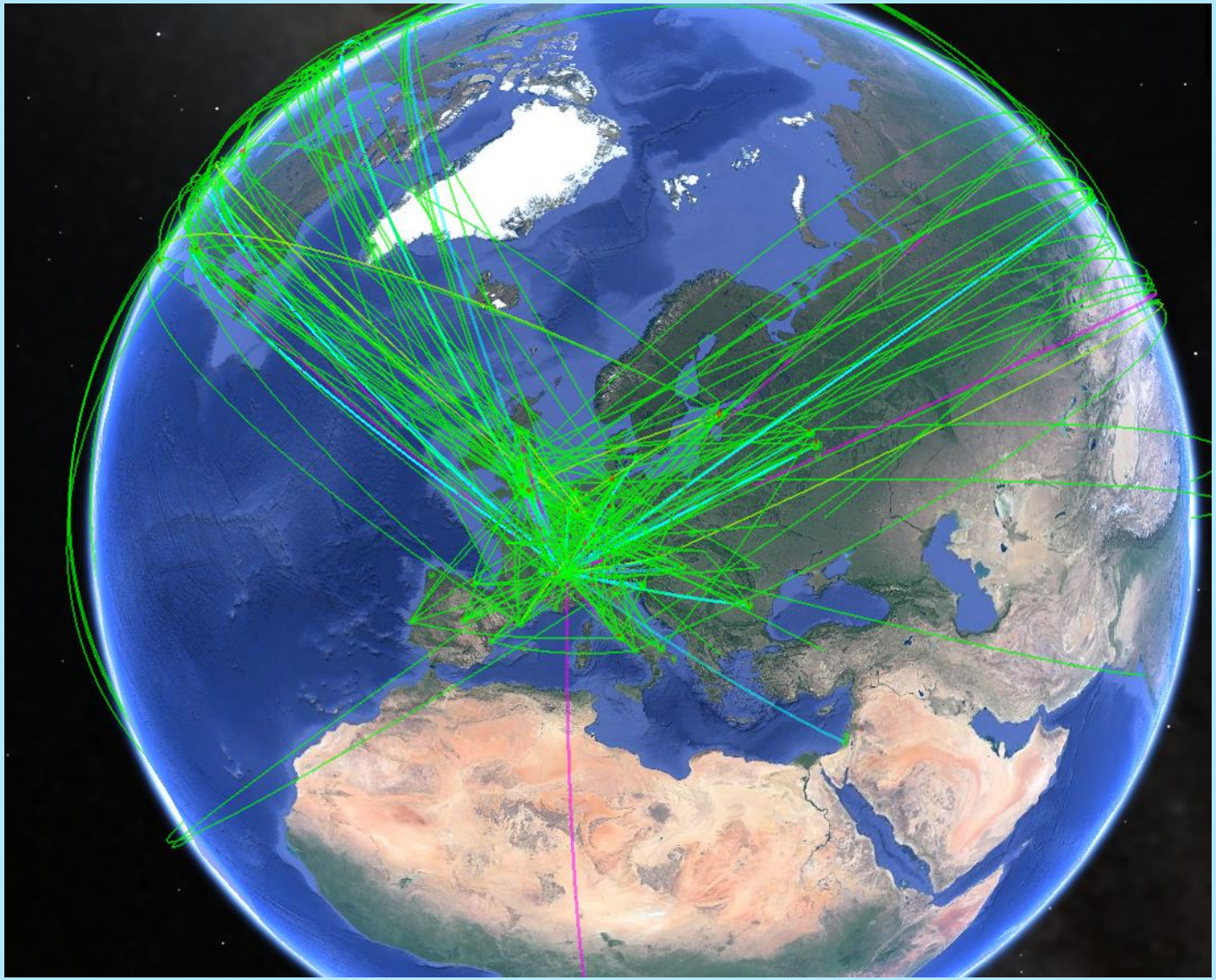
WLCG evolution towards the High-Luminosity LHC

M. Litmaath, CERN

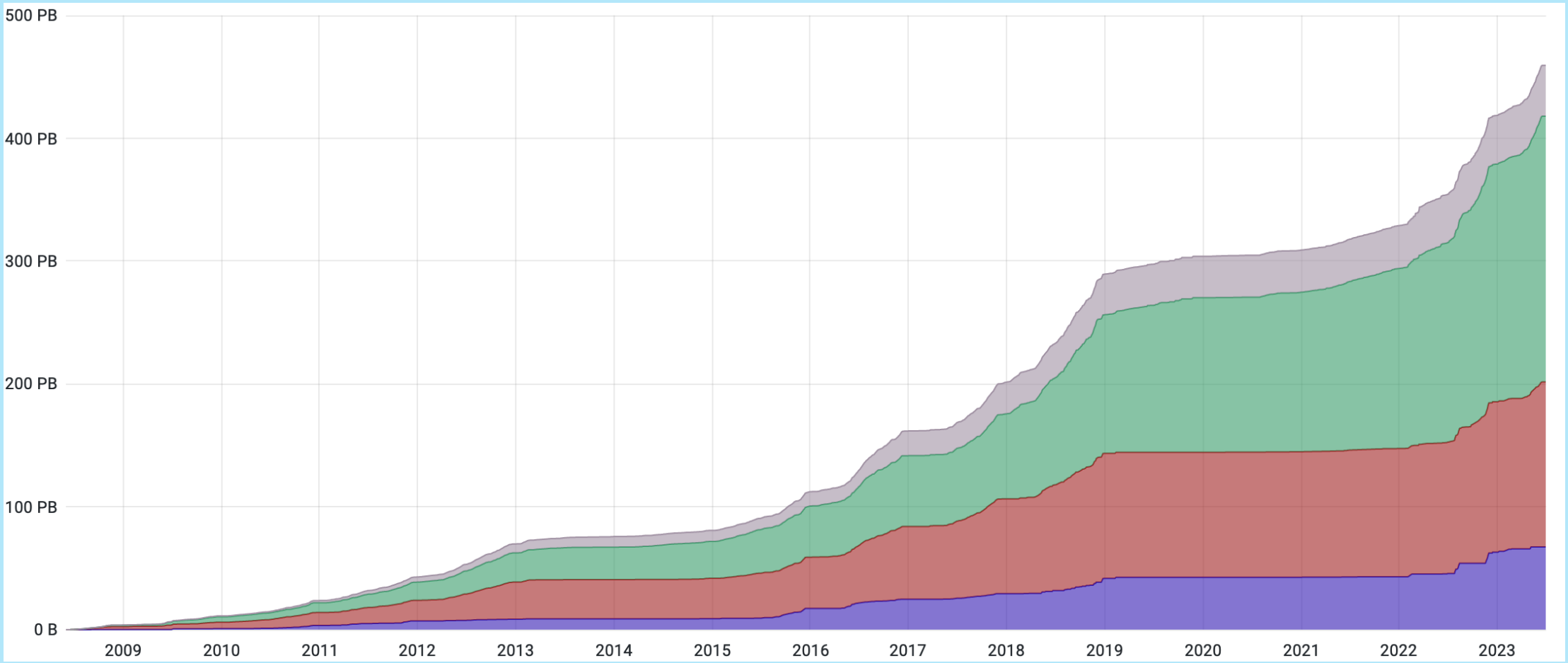
What is the Worldwide LHC Computing Grid?



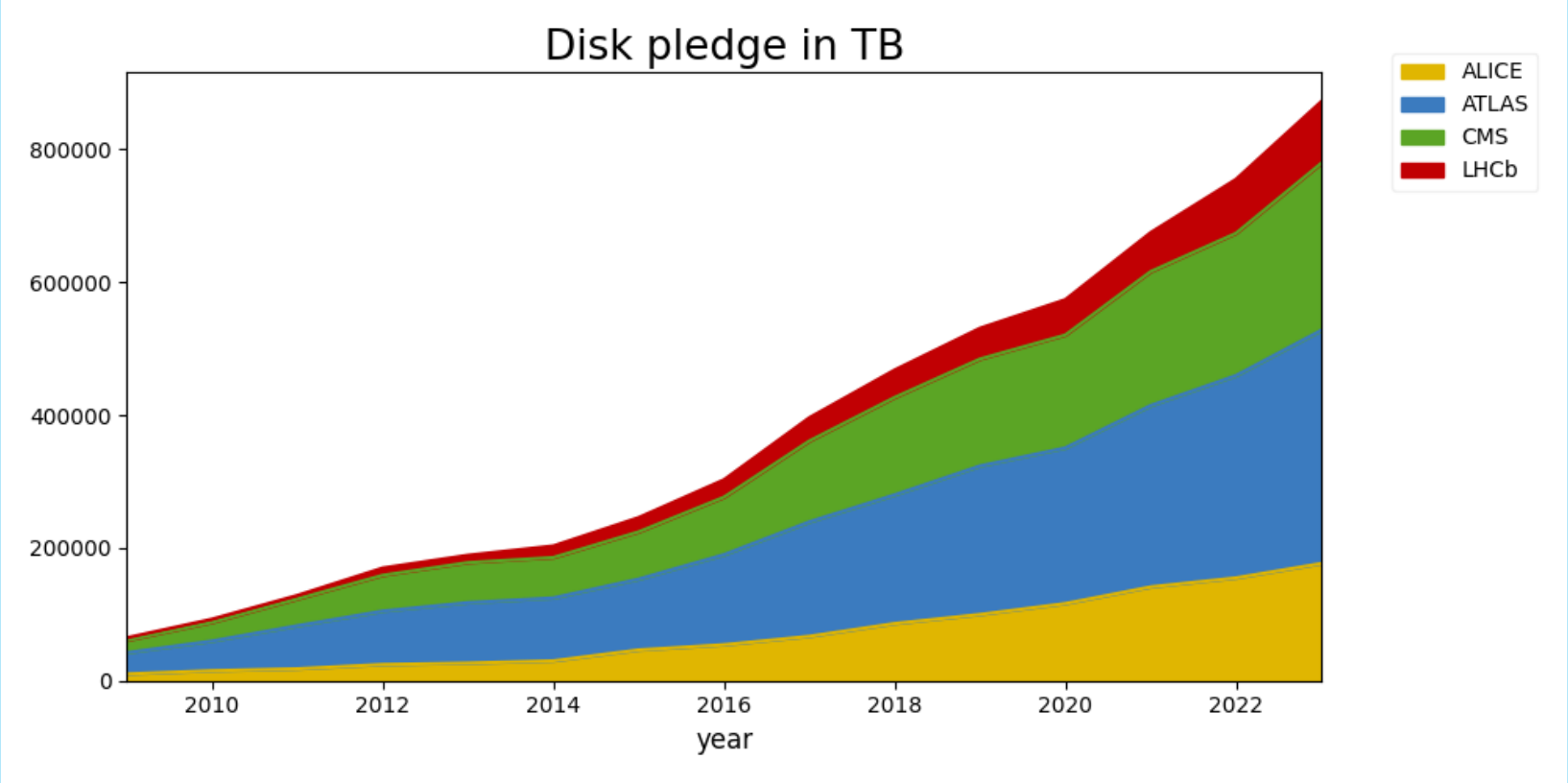
Worldwide data storage and processing 24/7



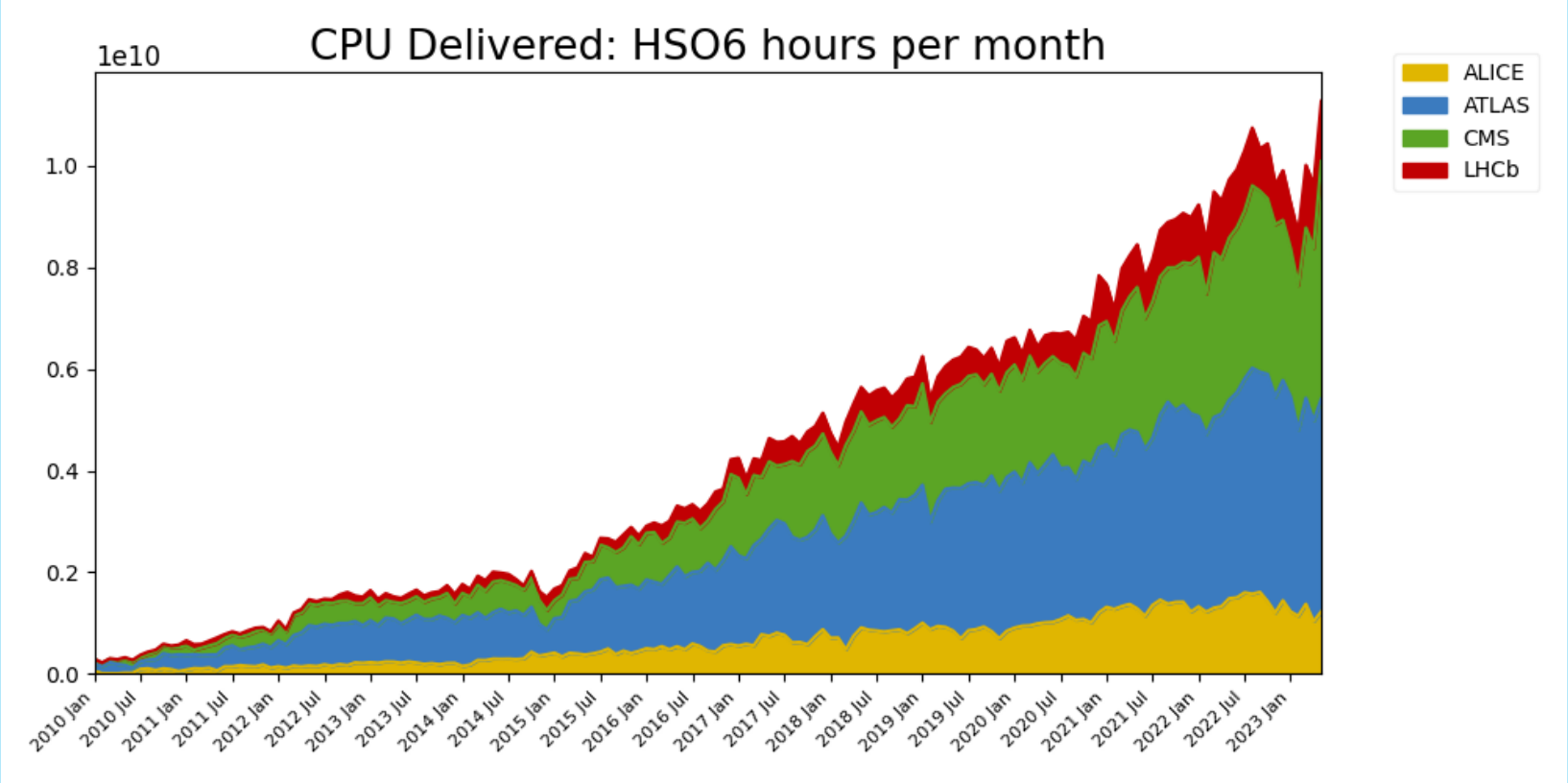
LHC experiment data on tape at CERN



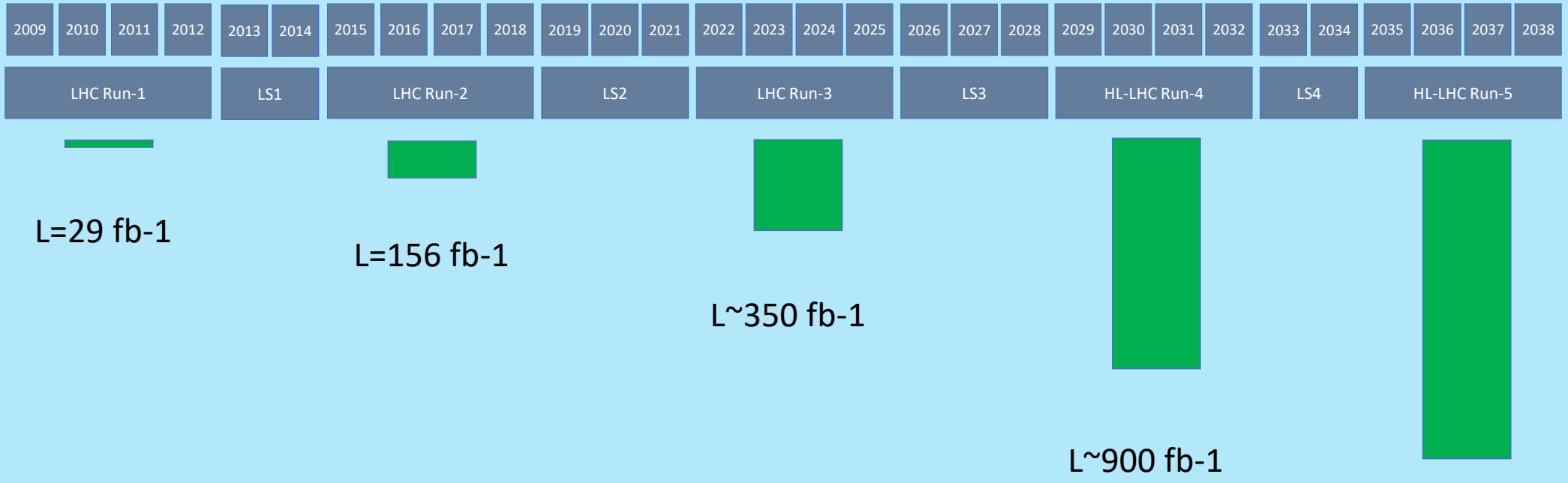
Disk pledge history



Corresponding increases in computing

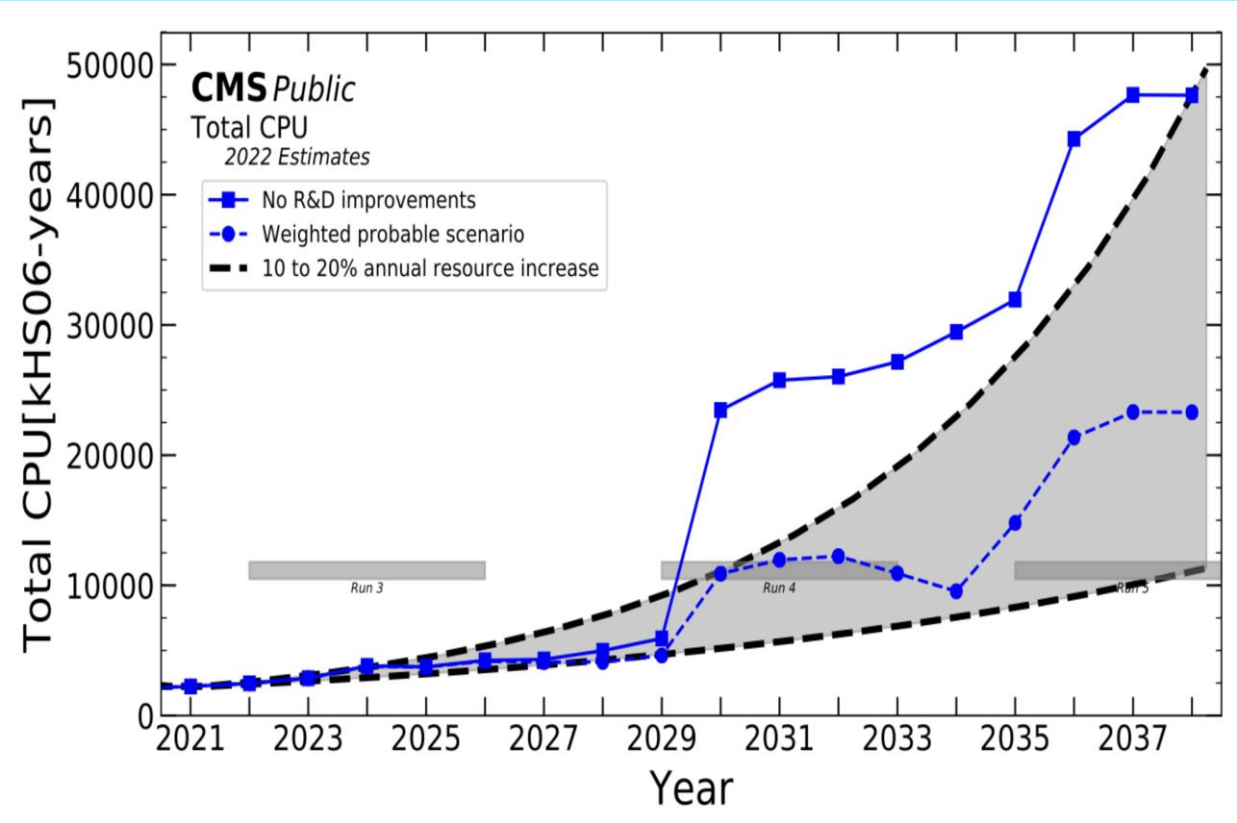


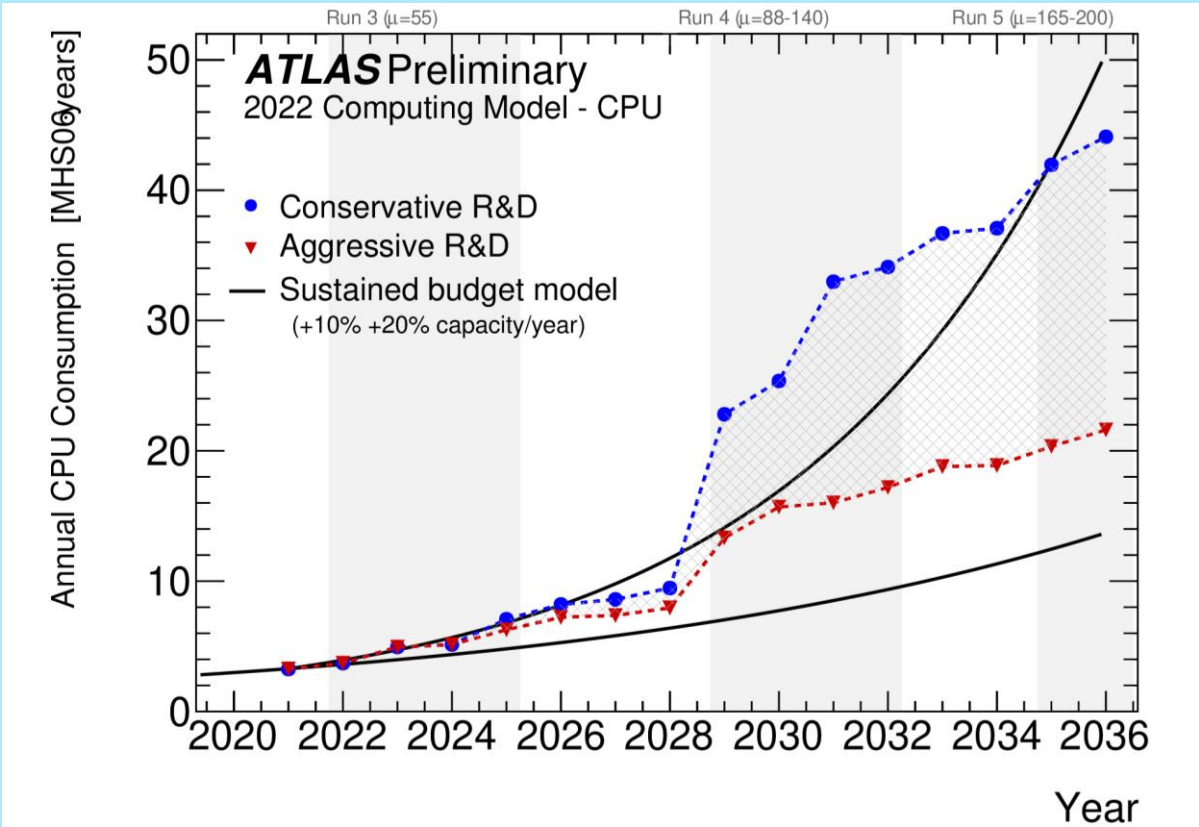
Ever increasing luminosity …



Focusing on ATLAS and CMS

… leading to ever increasing processing needs





… that require improvements in SW & services

Computing model & SW evolution

- Improved data formats for less space and faster processing
  - RNTuple
- Alternative architectures for cheaper and/or faster processing
  - ARM
  - GPU
  - Physics validation!
- SW development groups, foundations and projects foster R&D
  - SFT
  - HSF
  - IRIS-HEP
  - SWIFT-HEP
  - …

WLCG facility perspectives

- HW cost trends – do we still get +15% per year with flat budgets?
- Energy costs may rise significantly, depending on the country
  - Power *and* cooling
- Alternative architectures may be better in one or more ways
  - Work per second per monetary budget
  - Work per second per *energy budget*
- Energy may also be saved at peak periods
  - CPU frequency reduction
  - Active capacity reduction
- Future: climate-neutral data centers…
- Now: **new T1 centers** coming for LHCb !
  - NCBJ, Poland
  - IHEP, China

And a new data center for CERN !

To be ready for first use cases in November…

Analysis Facilities

- HL-LHC computing models foresee Analysis Facilities as a new type of resources
- Some WLCG facilities may want to offer such resources instead of traditional set-ups, or even in addition
  - Proper recognition needed in pledges and accounting
- Input data must not too quickly become the bottleneck
  - Fast local storage and network are needed (cf. HPC *burst buffers*)
  - Possibly in the form of caches, i.e. local data losses can be tolerated
- Profiting again from alternative architectures
  - ARM
  - GPU
- Users may access such services through notebooks
  - Jupyter, …
  - Auto-scaling from the laptop to the grid…

Benchmarking & Accounting

- After 14 years of service, the HEP-SPEC06 legacy CPU benchmark is being phased out
- As of April 2023, its successor should be used for *new* HW: HEPScore23
- HEPScore is a flexible, modern benchmarking suite
  - Based on containers and a configurable variety of reference workloads
- **HS23** is a specific benchmark constructed from 7 modern HEP workloads
  - Released for x86_64 **and** ARM
  - Adopted by WLCG **and** EGI
  - Big implications for the whole **CPU pledge & accounting** chain
    - ARC and HTCondor CE, OSG GRACC, APEL, EGI Accounting Portal, WAU, CRIC
    - Overall coordination through the Accounting TF

See April 2023 GDB update

- Still to come: GPU support !
  - Vital for future pledges and accounting

High Performance Computing centers

- HPC centers are used by the experiments since many years
  - Mostly opportunistically so far
  - Also part of NDGF-T1 pledges
- Some countries may intend to pledge HPC rather than HTC resources
- Various barriers have been reduced over the years, some still exist
  - Operator access controls
  - Network access limitations
  - Lack of CVMFS
  - …
- Some can be worked around
  - ARC CE data handling
  - ARC Control Tower
  - …
- EU and US projects help make HPC look more like our HTC sites

Cloud computing

- Experiments have been using commercial cloud resources through special projects since many years
  - Sponsored by big providers and EU projects
- So far, resources on our own premises have been less expensive
  - That is expected to change in the coming years, though
- For computing, cloud procurements can make a lot of sense
  - Potential accounting issue: no HS23 ratings for the underlying HW…
- For storage, we must be able to abandon our data!
  - All **custodial** storage must remain with our T0 and T1 sites
  - Data replicas in the cloud help speed up processing and analyses
- Integrating such resources has security implications
  - Their CAs are not in IGTF
  - The Resource Trust Evolution TF is looking into sustainable recipes

Containers

- At most sites, pilot jobs of the LHC experiments run user tasks in Apptainer (was: *Singularity*) containers distributed via CVMFS
  - Provide legacy payloads with their expected environments
  - Isolate payloads to limit interference and potential fallout

- Containers are also used increasingly for service deployments
  - Can come largely pre-configured
  - Pods with micro-services can be restarted or scaled up as needed
    - If the service design allows for that
  - May help simplify the lives of site admins!

- Victoria T2 – all services already on Kubernetes except storage
  - R&D project: EOS deployment via K8s on CephFS

Storage

- Erasure coding has matured and may allow sites to recover lots of disk space while still providing good resilience to HW failures
- As an alternative to expensive tape storage, T1 sites may consider looking into custodial *disk* storage
  - KISTI have pioneered such a service for ALICE, in production since 2021
- Industry has been looking into new custodial storage technologies (e.g. optical) since many years, but progress has been slow
  - Also of interest to cloud providers
- Object stores and access protocols are becoming more popular
  - Ceph
  - S3
  - Experiment data model adjustments desired or required to take full advantage
    - RNTuple

Networks

- Dual-stack deployment of storage has reached ~94% at the T2 sites
- IPv4 will need to remain supported at least a few more years for *legacy* workflows that cannot handle IPv6
- IPv6 networks allow new features to be taken advantage of
  - Dynamic provisioning through *Software Defined Networks*
    - Allowing better use of available capacity instead of relying on overprovisioning
  - *Packet marking* and *flow labeling*
    - Allowing activities to be monitored separately, both between *and* within VOs
- New L3 VPNs like LHCONE are expected to be set up for partner projects
  - Our MW and network monitoring tools will also be shared where feasible
- Such networks also need to be interfaced with HPC centers and cloud providers
  - To avoid overloading general-purpose networks connecting sites to the internet

Monitoring

- Though many systems are used by experiments and sites, complete overviews at the WLCG level are tricky and incomplete

- For example, data traffic monitoring improvements (in particular for the *Xrootd* protocol) concern a number of stakeholders
  - Experiments
  - SE and FTS developers
  - CERN MONIT and IAM teams (see next page)
  - WLCG, OSG and EGI Operations

- MONIT is foreseen to be used even more for consolidation
  - With CRIC used as WLCG **information system**

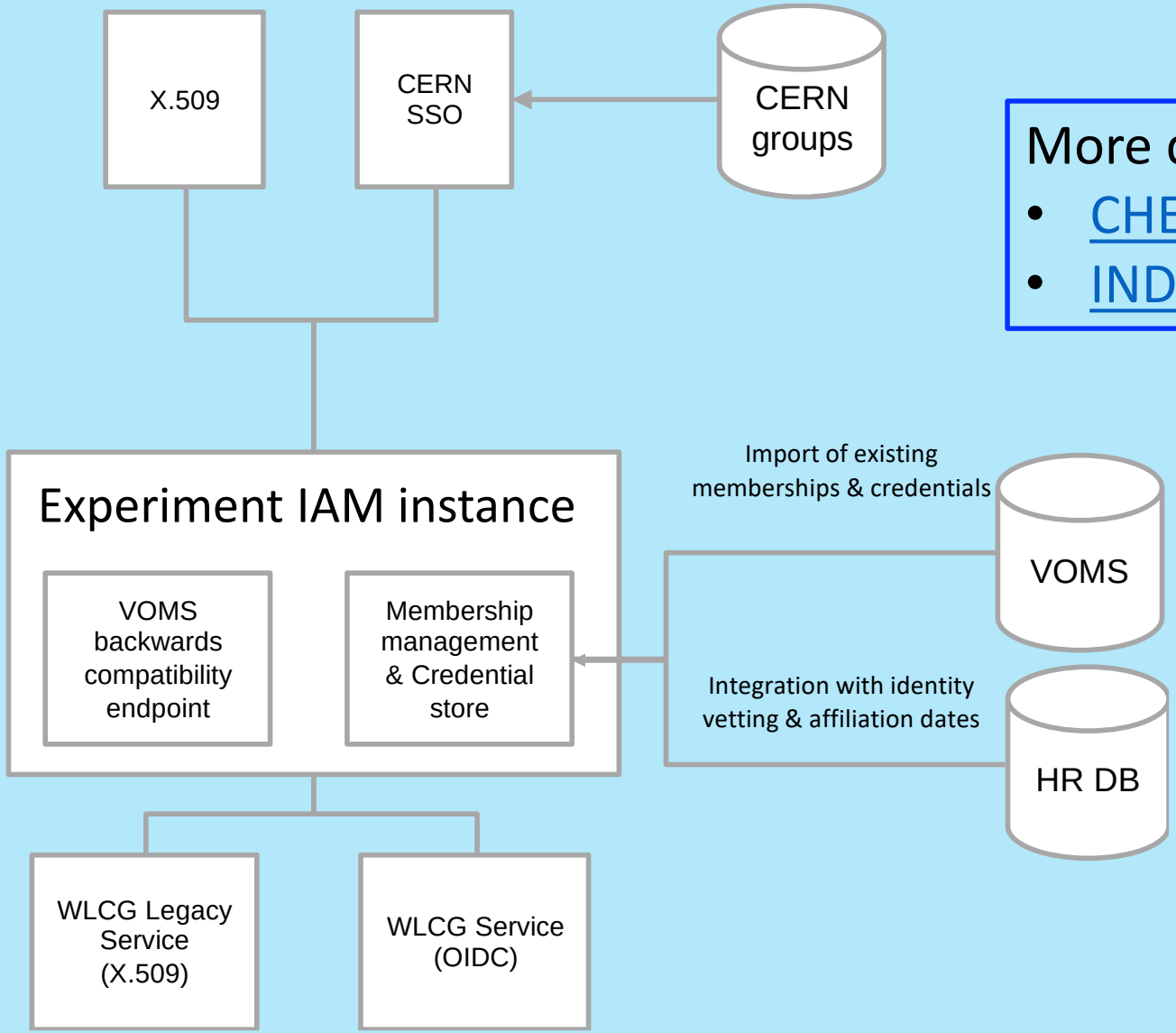
- New systems keep emerging and gaining popularity in many places
  - Prometheus
  - …

Authentication & Authorization

- Since a few years, WLCG has started transitioning from X509 user certificates and VOMS proxies towards using *WLCG tokens* instead
  - Arguably our biggest change ever!
  - Will also allow us to phase out the remaining Globus dependencies
- A timeline with tentative milestones was published in August 2022
- Coordinated by the Authorization WG and the Bulk Data Transfers WG
- Computing – in production for OSG & WLCG, WIP for *EGI Check-in* tokens
- Data – discussions ongoing between stakeholders
  - Experiments, CTA, dCache, DIRAC, Echo, EOS, FTS, IAM, Rucio, StoRM, XRootD, …
  - Rates, lifetimes, scopes, …
  - Plans for **Data Challenge 2024** (March) are shaping up!
- End game (~2026): users no longer need X509 certificates!
  - And do not need to know about tokens either!

New use case: message bus authN/authZ for *monitoring*

Experiment token issuer architecture



More details:

- CHEP 2023 contribution
- INDIGO IAM web site

Operations Coordination

- Coordinating the **evolution** of the WLCG infrastructure is part of the mandate of the Operations Coordination team, in close **collaboration** with stakeholders
  - Experiments
  - Infrastructure provider projects
  - Sites
  - Middleware projects

- Various activities serve in different capacities
  - Weekly operations meetings
  - Monthly Operations Coordination meetings
  - Task forces
  - Working groups
  - Grid Deployment Board and Management Board meetings
  - Workshops

Partners & Projects

- WLCG users and admins will rely on *federated identities* to have access to vital auxiliary services through EGI Check-in

- For example: GGUS
  - To be replaced next year by a **new WLCG Helpdesk !**

- Current and new infrastructure and SW **partnerships** and **projects** will continue playing important roles
  - EGI, NeIC, OSG, ASGC, IGTF, HEPiX, GÉANT, ESnet, ATCF, perfSONAR, …
  - EOSC-Future, PATh, EuroHPC JU, PRACE, …
  - *And all the **middleware** and **experiment software** projects!*

- Infrastructure and middleware will continue being shared with ***other communities***, in particular through MoUs with various collaborations
  - Belle II, DUNE, JUNO, Virgo, SKAO, …

Not so fast…

- Security incidents may affect part of our infrastructure at times
  - Security experts from our infrastructure provider projects and sites **collaborate** to minimize the fallout
    - For example through the SAFER trust group
  - WLCG sites can **collaborate** on incident response and prevention through the Security Operations Center WG
  - Sites can **contribute** through passive DNS SOC services

- We may be affected by upstream licensing or support changes
  - Case in point: the new RHEL source code policy
  - With consequences for AlmaLinux and Rocky Linux

Conclusions & Outlook

- Since many years, the Worldwide LHC Computing Grid has successfully provided the distributed computing infrastructure for the CERN LHC experiments.
- During that time, the WLCG has seen evolution in technologies as well as growth, to deal with ever increasing data rates.
- Those trends need to be made to continue, to allow the WLCG to take on High-Luminosity LHC data volumes as of 2029.
- Improvements across a wide range of services and software have been described, some of which already bring benefits as of today.
- Thanks to many **partners and projects**, not only do the LHC experiments profit, but many **other communities** as well !