Track 4 Highlights

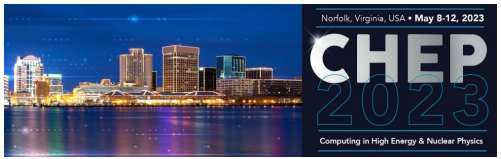
Distributed Computing

Rohini Joshi, Fernando Harald Barreiro Megino, Hideki Miyake, Katy Ellis

Themes of Track 4

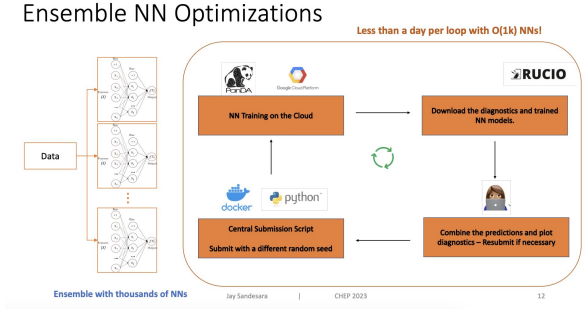
- Analysis workflows, monitoring and optimization (Mon 11:00)
- Computing strategies and evolution (Mon 14:00)
- Infrastructure and services (Tue 11:00)
- Monitoring, testing and analytics (Tue 14:00)
- Security and tokens (Tue 16:30)
- Distributed storage and computing resources (Thu 11:00)
- Workload management (Thu 14:00)

Link to all Track 4 contributions:
https://indico.jlab.org/event/459/sessions/2038/#20230508

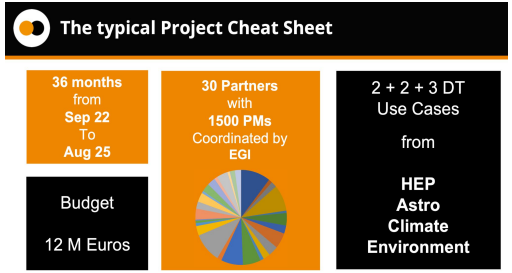


Analysis workflows, modeling and optimization

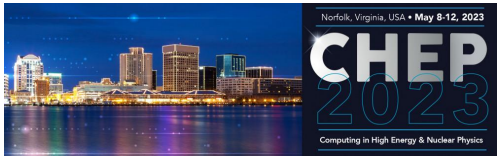
- ATLAS use cases for complex workflows, mapping distributed Machine Learning to heterogeneous resources like Grid, HPC and Cloud GPUs through PanDA
- Overview of the Digital Twin Infrastructure and how it will enable multiple use cases
- Analysis Grand Challenge benchmarking results using CMS open data



Jay Sandesara, ATLAS Data Analysis Using a Parallel Workflow on Distributed Cloud-Based Services with GPUs, slide 12



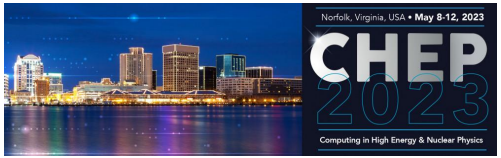
Patrick Fuhrmann, Digital Twin Engine Infrastructure in the interTwin Project, slide 11



Fri 12th May Plenary Session - Conference Highlights: Track 4

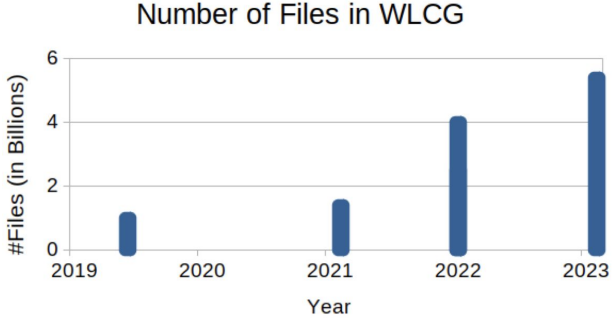
Computing Strategies and Evolution

- Big science computing models for old and new experiments: ATLAS, CMS, Einstein Telescope, JUNO, Ligo-Virgo-Kagra
- Neutrino, gravitational wave and high energy physics experiments getting ready for next decades of data taking
- Established data and workload frameworks preparing for upcoming challenges:
  - Larger datasets requiring more efficient storage strategies
  - Heterogeneous architectures and modernizing the SW
  - Making the best usage of varied computing resources: grid, HPC and cloud
  - Recurring topic this CHEP: sustainability awareness!



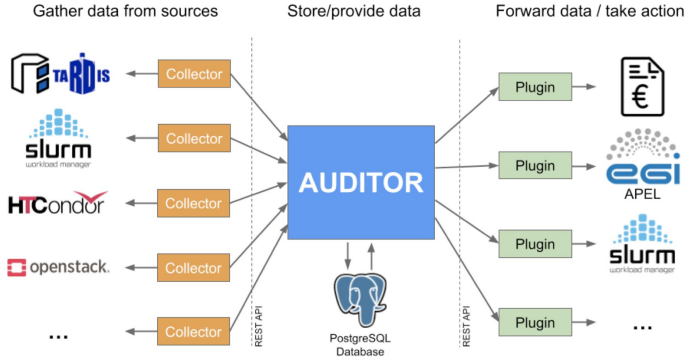
Infrastructure and Services

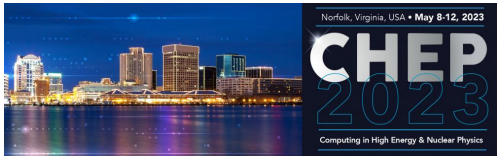
- Systems supporting multiple experiments across multiple sites in a clear and efficient way
  - Central orchestration of services 👍
- Extreme CVMFS!
- Accounting of opportunistic resources with AUDITOR



Left: Laura Promberger, CERNVM-FS at Extreme Scales

Right: Michael Boehler, AUDITOR: Accounting for opportunistic resources

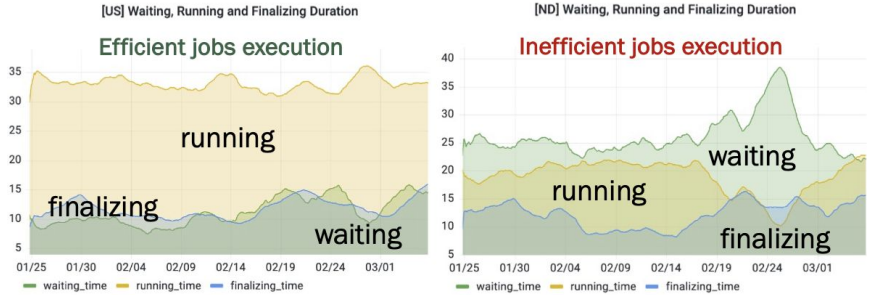




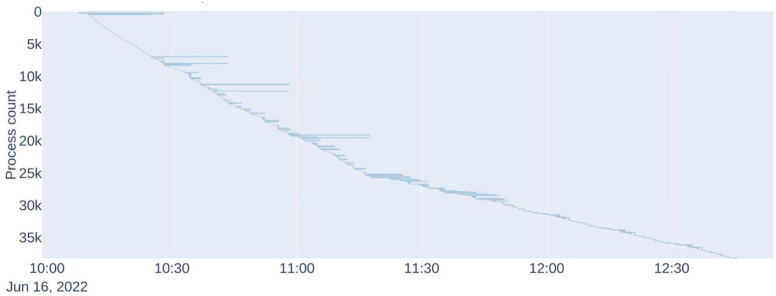
Fri 12th May Plenary Session - Conference Highlights: Track 4

Monitoring, Testing and Analytics

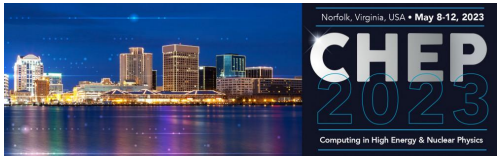
- Significant updates from ATLAS and CMS on their monitoring systems and overview of the new ALICE monitoring - SiteSonar
- Monitoring of multicore ALICE jobs spawning parallel processes
- Data popularity forecasting and job queue optimization for ATLAS
  - New metrics can help improve utilization



Alexei Klimentov, Operational Analytics Studies for ATLAS…



Marta Bertran Ferrer, Multicore workflow characterization methodology…



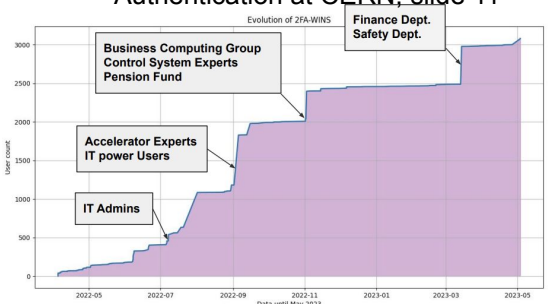
Fri 12th May Plenary Session - Conference Highlights: Track 4

Security and Tokens

- Collaborative, international operational security: sites invited to protect our infrastructure together
- CERN's multi factor authentication migration for further protection from password hacks
- Plans and status for the much anticipated WLCG move to tokens, followed by the implementations in some of the WLCG experiment frameworks
    - X509 to token migration should not be considered a drop-in replacement!

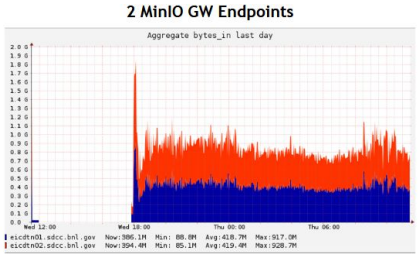
Migration Timeline - ~2500 users migrated

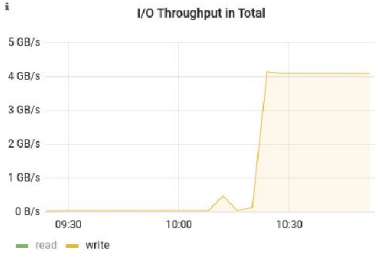
Adeel Ahmad, Enforcing Two-Factor Authentication at CERN, slide 11



Distributed Storage and Computing Resources

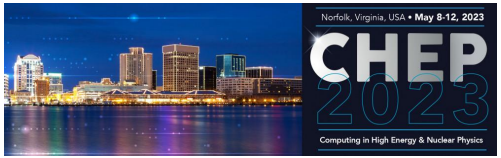
- Evolution of distributed computing technologies enables the combination of geographically distant sites as a big virtual laboratory showing comparable performance with single site
  - The difference is hidden from the user





Justas Balcas, Job CPU Performance comparison based on MINIAOD reading options: local versus remote

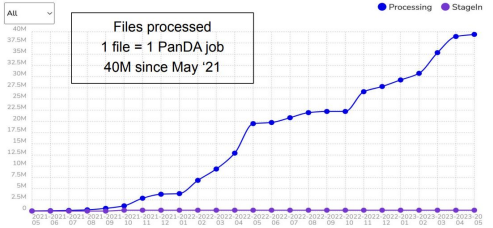
- As we heard in the plenaries from Larry Gladney, “...harnessing diverse intellectual resources brings overall benefit”, similar sentiment heard with respect to computing resources as we “Embrace heterogeneity”
  - Integrating GPU and non-X86 CPU with mainline grid computing



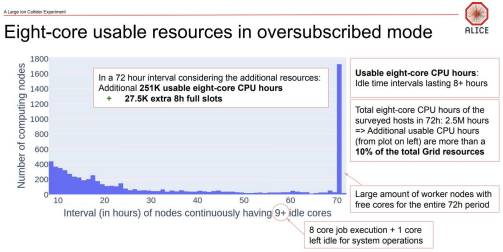
Fri 12th May Plenary Session - Conference Highlights: Track 4

Workload Management

- Reports of re-use of workload management systems (WMS) by new experiments
  - Vera Rubin Observatory - good progress with PanDA endorsed for DRP processing
  - CTAO - improvements with the DIRAC WMS, with Common Workflow Language support introduced
- Improvements to existing WMS
  - LHAASO Cluster Extension to include remote resources
  - Additional automation and intelligent dynamic scheduling of opportunistic resources with JAliEn for ALICE



Edward Karavakis, Integrating the PanDA WMS with the VRO, slide 8



Marta Bertran Ferrer, Dynamic scheduling using CPU oversubscription on the ALICE Grid, slide 6

Fri 12th May Plenary Session - Conference Highlights: Track 4