



27th Conference on Computing in High Energy and Nuclear Physics (CHEP2024)

summary

Tania Korchuganova (Pitt)

SDCC-NPPS meeting, 14 Nov 2024

~~ATLAS weekly, 5 Nov 2024~~

Introduction

The **CHEP conference** series addresses the **computing, networking and software issues** for the world's leading data-intensive science experiments that currently analyse **hundreds of PB of data** using **worldwide computing resources**.

Hosted by the AGH University of Kraków, Institute of Nuclear Physics Polish Academy of Sciences and Jagiellonian University in **Krakow, Poland**

470 participants

351 talks & **143** posters



Scientific program: 9 tracks

Track 1 - Data and Metadata Organization, Management and Access

Track 2 - Online and real-time computing

Track 3 - Offline Computing

Track 4 - Distributed Computing

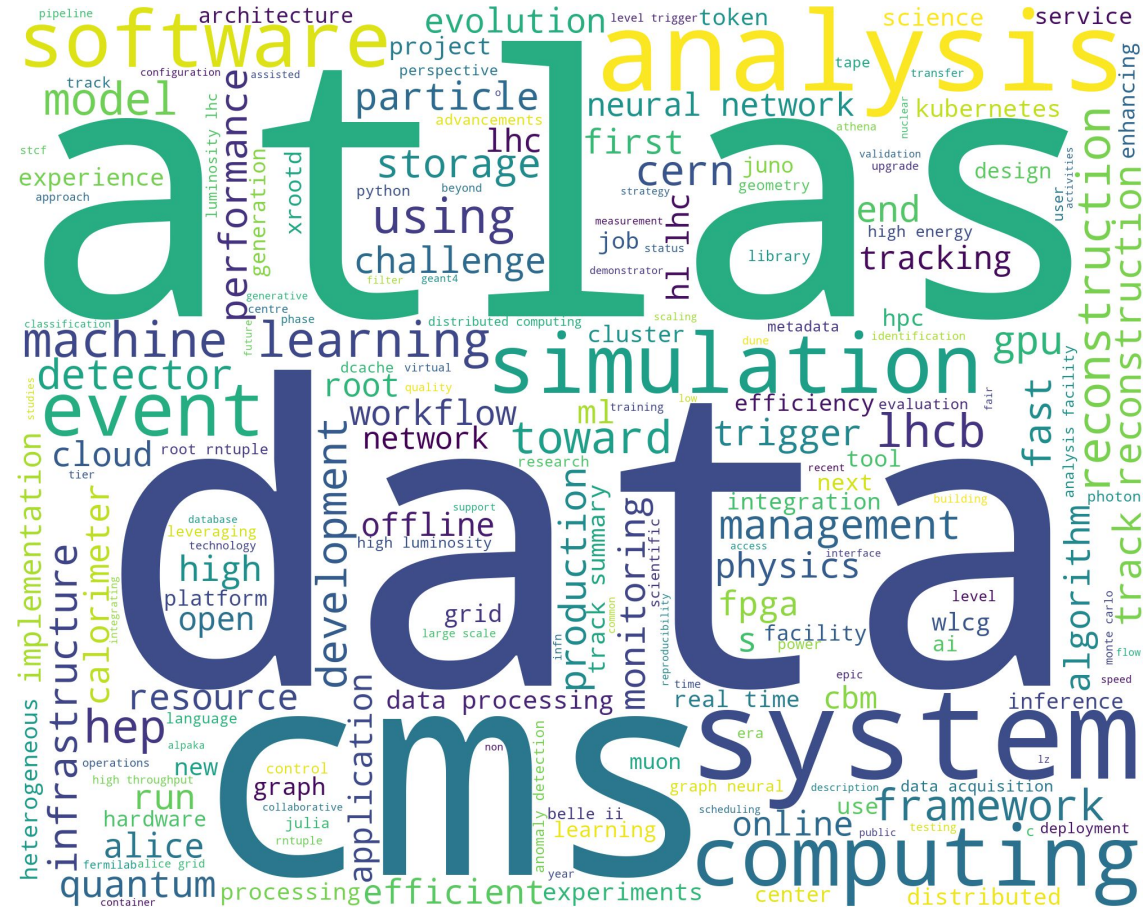
Track 5 - Simulation and analysis tools

Track 6 - Collaborative software and maintainability

Track 7 - Computing Infrastructure

Track 8 - Collaboration, Reinterpretation, Outreach and Education

Track 9 - Analysis facilities and interactive computing



Scientific program: 9 tracks

Track 1 - Data and Metadata Organization, Management and Access

Track 2 - Online and real-time computing

Track 3 - Offline Computing

Track 4 - Distributed Computing

Track 5 - Simulation and analysis tools

Track 6 - Collaborative software and maintainability

Track 7 - Computing Infrastructure

Track 8 - Collaboration, Reinterpretation, Outreach and Education

Track 9 - Analysis facilities and interactive computing

7 parallel sessions every afternoon



it is impossible to attend everything and one need to run between rooms for interesting talks



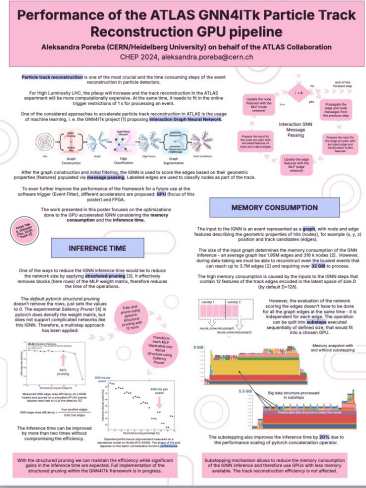
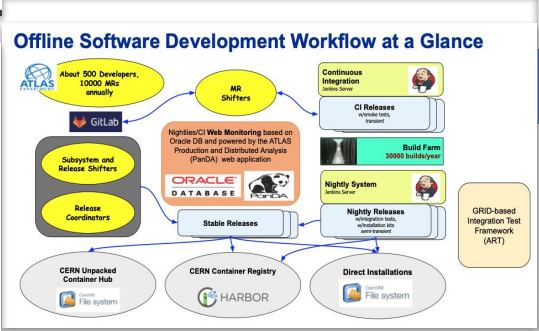
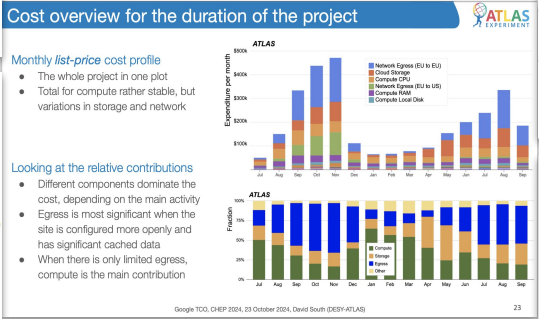
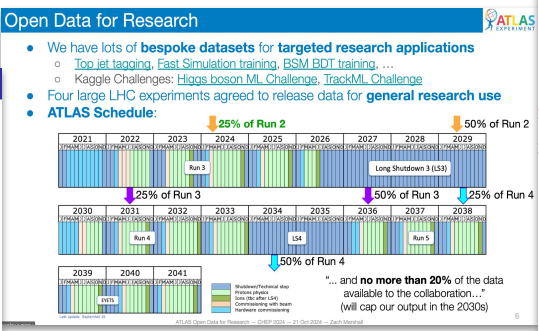
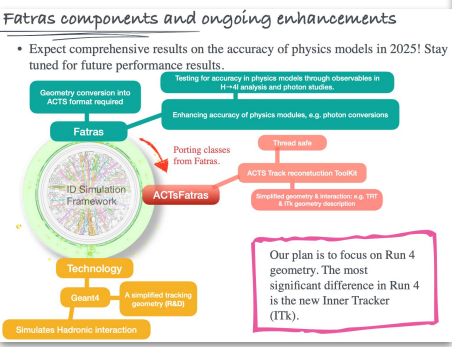
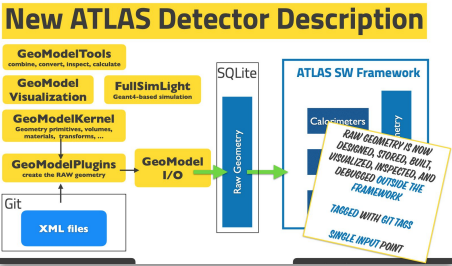
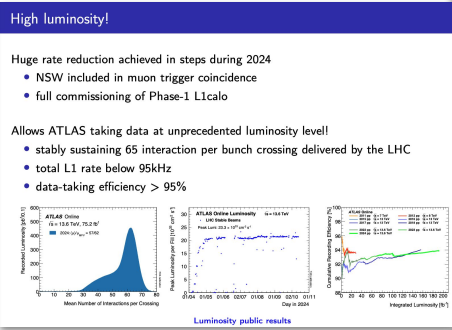
very subjective summary of the conference

	Total Cost of Ownership and Evaluation of Google Cloud Resources for the ATLAS Experiment at the LHC David South						
	Large Hall					09:30 - 10:00	
10:00	Enhancing Data Management in Nuclear Physics and Cross-Domain Metadata Integration					Ivan Knezevic	
	Large Hall					10:00 - 10:30	
	Coffee break						
						10:30 - 11:00	
11:00	ROOT RNTuple and EOS: The Next Generation of Event Data I/O					Andreas Joachim Peters et al.	
	Large Hall					11:00 - 11:30	
	The WLCG Data Challenge					Katy Ellis	
	Large Hall					11:30 - 12:00	
12:00	Global Networking Challenges for the Coming Decade					Tony Cass	
	Large Hall					12:00 - 12:30	
	Lunch break						
						12:30 - 13:30	
	Benchmark Loren...	Monte Carlo Dr An...	Advanced m Brij Ki...	A RoCE-bas Grae...	Fast Jet Rec David...	A graph neu Hasa...	Recent Expe
	Advancem Bern...	Madgraph o Andr...	Whole-node Marta...	Front-End R Gabri...	Reconstructi Juan...	Charged Par Gagik...	DUNE Rucio Wern...
14:00	Zero-overhe Dr Vi...	Hardware ac Zenn...	Unprivileged Makl...	Near-Data C, Aash...	The Real-Tin M P...	TrackNET: D Mr P...	FTS as a par Rose...
	On-the-fly ds Florin...	Event gener Yutar...	Unified Expe Ewou...	Using the AT Joha...	Next generat Paul...	ML-Assisted Torri...	Distributed Lia L...
	GIL-free scal Jim P...	Monte Carlo Jame...	Designing O Dr Da...	Improving o Diana...	TGeoArbN - Ben...	Neural netw, Christ...	Data Movem Fabio...
15:00	Using and Visu, Julius...	Navigating Pha Salvato...	Addressing tok Kyle K...	The Clance pro Luis G...	The Belle II Raw Tristan...	Machine Learni Tam J...	
	Poster session: Presentation with coffee						
	Exhibition Hall						
16:00	zfit: general Jason...	WLCG Oper Pano...	Carbon, Pow Zach...	A Pilot Analy Enric...	Towards a G David...	Improvement Char...	Adoption of Marci...
	Model Buildi Lawr...	Implementat Haak...	Allocating C Alex...	Evolution an Thom...	tracc: GPU Beom...	GPU Acceler Nuno...	RNTuple: A Nick...
17:00	New RooFit Jona...	Enhancem Alex...	Heterogeneo Eman...	Operational Anton...	Application Yatho...	Low-latency Piero...	ML-based A Dr By...
	BATJl, the B Oliver...	Optimisation Natali...	Taking on RI Eman...	Building Scal Lincol...	Extending A Matte...	Performance Abhiri...	Advanceme, Maci...
	Parameter Esti, Ashik...	Modernizing AT Tatiana...	Heterogeneous Daniel...	Computing Acti Pablo...	Efficient metadi Mr Fab...	An online GPU Felix...	
18:00	Novel Fitting Ap Dr Yuri...	Optimization Marco...	Enhancing Net, Peiya...	Leveraging dist Federic...	So FAIR, so goe Lorenz...	SYCL-based o Bartos...	

ATLAS contributions

65 contributions (~13%):

- 2 plenary talks:
 - The First Release of ATLAS Open Data for Research
 - Total Cost of Ownership and Evaluation of Google Cloud Resources for the ATLAS Experiment at the LHC
- 44 session talks
- 19 posters
 - 2 winners of “Poster of the day”:
 - Performance of the ATLAS GNN4ITk Particle Track Reconstruction GPU pipeline
 - Evolution of Regional, Age and Gender Demographics in the ATLAS Collaboration

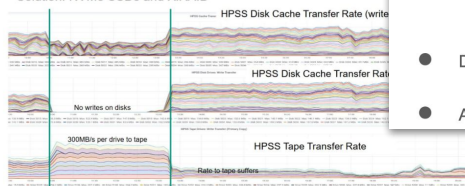


Track 1 - Data and Metadata Organization, Management and Access

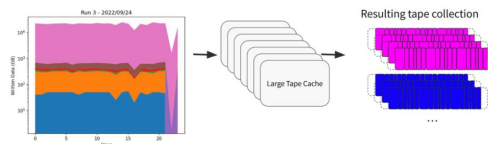
- WLCG Data Challenge 2024 (25% of HL-LHC)
 - Observed bottlenecks, planning the next steps, 50% challenge in 2026
- **Tape and Archive:**
 - CTA evolution, smart writing, which data formats to store on tape, ATLAS Data Carousel

Tape and Archive

- New **GridKa** Tape storage systems, switching from SP to **High Performance Storage System**
 - aggregating files from the same directory/creation time and computing+saving checksum
- **ATLAS** HL-LHC demonstrators with **Data Carousel** to improve tape bandwidth utilization
 - data-on-demand and **tape smart writing**, tests with ATLAS Tier-1s **FZK** and **RAL**
- A **Tape Remote Storage Element** in support of **LSST** backup challenge @SLAC (8 PB/year)
 - a specialized xroot server that orchestrates datasets backups, automatically creating a copy of a **Rucio** dataset as a single indexed file; works also for active restores



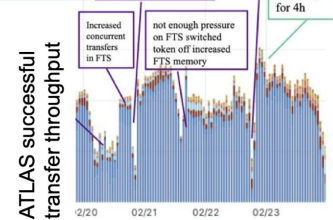
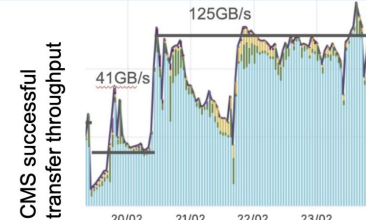
- **HTTP Archive Metadata** header added to **http data transfers** and received by tape endpoints to improve data **collocation** on tape and **tape scheduling**



DC24

- **WLCG Data Challenges:** established in 2021 to prepare for HL-LHC data rates. DC24: ATLAS, CMS, ALICE, LHCb, **BelleII** (poster) and **Dune**

Year	% of HL-LHC
2021	10
2024	25
2026?	50
2028?	100



- Observed bottlenecks (ATLAS, CMS): scalability of FTS, Rucio and token handling, risk to overload storage systems
 - Data Movement Manager prototype interface between Rucio and the software defined networking (SDN) service SENSE by ESNat. Dynamic provision of network path
- DC26: 2xDC24 rates (50% HL-LHC), tokens established and battle tested, tape. Intermediate mini challenges
 - FTS3 Token Support for a Proxy-less WLCG world
- ALICE and LHCb run a program mostly focused on CERN to Tier-1s traffic; Belle II and DUNE also participated

- **Data movement and management:**
 - Rucio & FTS is a standard now
- **Data formats & metadata:**
 - RNTuple is a hot topic, efficiency studies wrt TTree by ATLAS and CMS
- **DB and filesystems:**
 - AI in Oracle soon, CEPH is heavily used

Track 4 - Distributed Computing

- Moving from X.509 to tokens
 - The infrastructure is almost ready
 - Focusing on operation models
 - Improving IAM
- Monitoring & Operations
 - A lot of different technologies in use, dashboards-as-code is promising
 - CMS is trying to stick to MONIT stack & UEM
 - ATLAS HammerCloud with automatic recovery of blacklisted sites
- Optimizations
 - HEP score benchmark cross-check of sites core power
 - ALICE optimizing job placement & pushing sites to provide dedicated nodes for whole node scheduling
 - CMS improving CPU efficiency through Pilot overloading

WLCG IAM Token Issuers

- VOMS(-Admin) services were fully retired by July 2024 (see next page)
 - IAM provides backwards compatible VOMS endpoint
- Kubernetes replaces OpenShift deployment
 - Greater control
 - Fewer dependencies
 - Easier to replicate between data centers for yet better HA

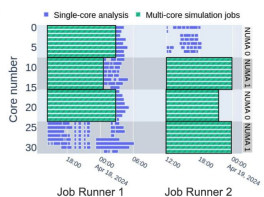


Data-flow: or 'death by a thousand acronyms'



Optimising job placement

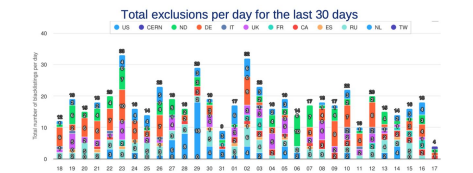
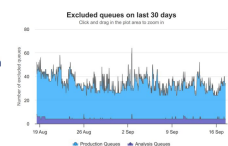
- Job-to-core mapping algorithm starts by jobs with larger allocations
 - Simulation jobs (8-core) are placed first and analysis (1,2,4-core) as backfill
 - Transparently balancing resource loads
- NUMA-aware pinning leads to improved **execution efficiency**¹
- Predictable execution time fostering better **scheduling decisions**



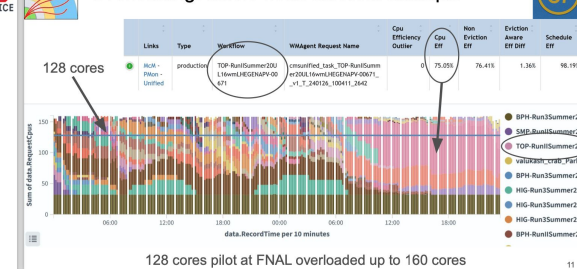
¹ CPU-level resource allocation for optimal execution of multi-process physics code

Auto-exclusion

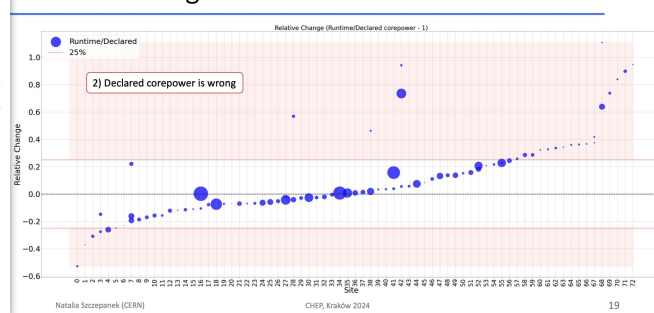
- Consider amount of recent successful/failed job
- Set of rules triggers the **automatic exclusion / recovery** of sites (queues) from the pool of resources available to users
 - ~ 10 - 40 queues excluded at given time
 - ~ 10 - 30 daily exclusions / recoveries



Overloading: whole node slot real example



Relative Change for Different ATLAS sites



Track 4 - Distributed Computing

Workflow management systems:

- ATLAS PanDA is under refactoring & modernization
- DIRAC is being rewritten from scratch, first release of DiracX is in 2025
- ALICE JAlEn is evolving, started to support ARM, and has integrated HPC

Outside HEP:

- SKA** (all purpose radio telescope) expects 400PB/year by 2030, using common tools like IAM, FTS
- HERD** (High Energy cosmic-Radiation Detection), start expected in 2027, using DIRAC & Rucio
- DUNE** (neutrinos) starts in 2029, expects 30PB/year, using CMS stack (HTCondor + GlideinWMS) + production system justIN and Rucio
- Vera Rubin Observatory** is in final steps of construction, start is in 2025, uses PanDA, Rucio, IAM

Stream 1: Code-base optimization & modernization

- Upgrades and Migration:** Updating to the latest versions of operating systems, Python, and package dependencies to enhance performance and security. Also, supporting Kubernetes-based deployment for all PanDA WMS components.
- Code Refactoring:** Focus on refactoring the code to improve readability and maintainability, and remove outdated or obsolete code to simplify the codebase.
- Reinforcement of Coding Standards:** Implement and enforce coding standards with the help of code validation tools to ensure quality and consistency throughout the development process.
- Kubernetes-Based CI/CD:** Develop a Kubernetes-based testbed to facilitate and optimize Continuous Integration (CI) and Continuous Deployment (CD) processes, supporting automated testing and deployment.

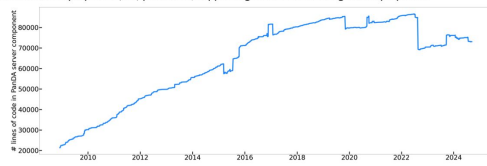
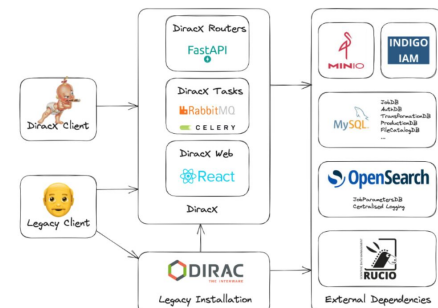


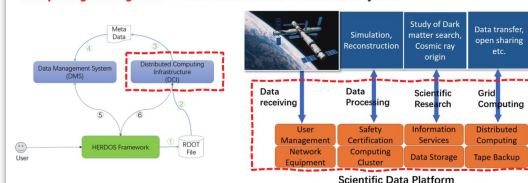
Figure 2 – # lines of code in the PanDA server component over time



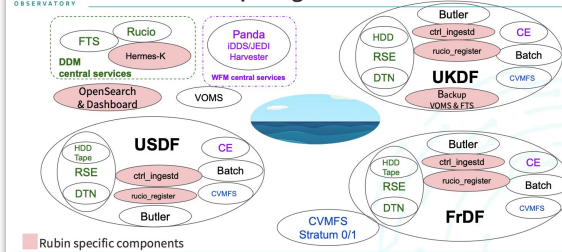
Distributed Computing Infrastructure

HERD DCI is a distributed computing system for,

- Data processing** -> Distributed computing system
- Data access** -> Distributed data management system
- Data distribution** -> Network and data transfer system
- Data privilege management** -> Authentication and authorization system

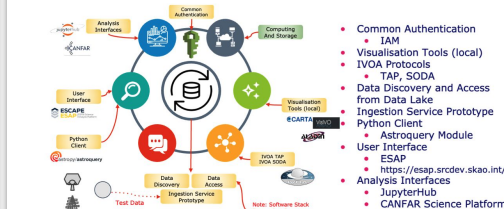


Distributed Computing Infrastructure

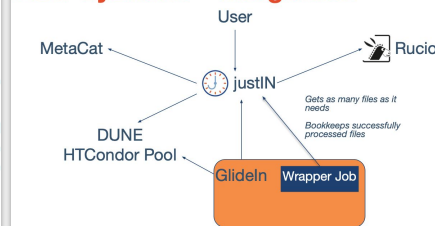


Software Stack (v0.1)

- SRCNet v0.1 defines a common set of tools / services for initial prototyping and to allow demonstration of:
 - Ingestion, Data Movement, Data life-cycle, Execution of "known" science use cases on SKA test data



New Systems – Integration



Other Hot Topics

- ML/AI and LLMs (~40 contributions) for:
 - code review, operations, user support, anomaly detections, teaching & training, navigating conference abstracts, fast simulation in LHCb, CMS FlashSim, reconstruction for FCC detectors, photons classification in ALICE, simulating CMS HG Calorimeter, tracking in GlueX, triggers, AthenaTriton, MC tuning in JUNO, particle flow in ATLAS detector
- Quantum computing
- RNTuple format
 - Improves I/O performance and data compression

LLMs for Enhanced Code Review

Alexey Ilyashchenko, Mohammad Al-Turany
GSI Helmholtz Center for Heavy Ion Research GmbH, Darmstadt, Germany

Peerbot [1]: GitHub Pull Request → Peerbot → LLM query instructions → Peerbot Backend → LLMs

Peerbot Features

- GitHub App for reviewing Pull Requests
- Local execution mode for diff or annotated comments
- Agent *ensemble approach* for comprehensive analysis
 - Initial review with any number of different models
 - Followed by a final model to refine the initial review
- Customizable model(s) via the llama setup
- Execution on low-end hardware and/or without GPU
- Customizable prompts

Peerbot Refining the Context

Adding relevant content to the LLM query can enhance the quality and relevance of the review. For example:

- Related issues and their summaries
- Coding standard/guidelines excerpts
- Historical code changes of the corresponding code
- Experiment & Framework-specific coding conventions

To select relevant information for the content and to avoid overloading context size of the model, relevant parts should be selected. Our current technique to achieve this is **Depth Retrieval: Augmented Generation (GRAMAG)**.

Peerbot Backend

allama (via python cli and HTTP requests) open source large language model server, written in Go, backed by **Brama.cpp** (C++)

- Efficient serving of large language models
- CPU/GPU/CPU+GPU hybrid inference to partially accelerate models larger than the total VRAM capacity
- Supports many model architectures: llama, gpt4all, qwen2...
- Support for multistate of model quantization techniques and precisions for faster inference and reduced memory use
- Usage Metrics

LLMs for Review Weaknesses

Allowing independent agents to generate separate reviews, followed by a final AI reviewer's synthesis, can improve the response quality. At the same time, the system should filter out superfluous reviews, presenting only relevant insights to the user, or none at all when there are none.

LLMs for Review Weaknesses

Allowing independent agents to generate separate reviews, followed by a final AI reviewer's synthesis, can improve the response quality. At the same time, the system should filter out superfluous reviews, presenting only relevant insights to the user, or none at all when there are none.

CodeRabbit [2]

Collaborative software development demands rigorous code review processes to ensure maintainability, reliability, and efficiency. We integrate Large Language Models (LLMs) into the code review process, utilizing both commercial and open models. We present a comprehensive code review workflow that incorporates open-weight LLMs, integrating various enhancements such as multi-agent capabilities and reflection. By harnessing the capabilities of LLMs, the review process can uncover faults and identify improvements that traditional automated analysis tools may overlook. This integration shows promise for improving code quality and reducing errors. We deploy **coderrabbit.ai** for commercial models & develop our own tool - **peerbot** for usage with local models.

CodeRabbit Features

- AI-based Pull Request summarizer and reviewer with chat capabilities via a web interface
- Combination of various OpenAI models
- Periodically open issues, new & closed projects
- Free to use for open source projects

LLM Review Strengths

- Unbiased, fast, scalable, multi-lingual, customizable, with a broad knowledge base, can be kept up-to-date with new data
- Good at identifying logical errors early, but other automated tools or even human review may overlook, avoiding issues later in the project lifetime
- Can potentially reduce the time and resources needed for code reviews

LLM Review Weaknesses

- May produce unnecessary output, when no actionable changes are necessary or such are not detected by the model, which would waste developer's time
- Weaker models, especially with quantized weights, may more easily dive into hallucinations, potentially leading to inaccurate or irrelevant code suggestions
- Limited understanding of complex projects, which human reviewers inherently possess
- Potential for false positives, flagging issues that aren't actually problematic
- Inability to fully comprehend the broader architectural implications of code changes, especially in large, complex systems

Peerbot: Refining the Context

Adding relevant content to the LLM query can enhance the quality and relevance of the review. For example:

- Related issues and their summaries
- Coding standard/guidelines excerpts
- Historical code changes of the corresponding code
- Experiment & Framework-specific coding conventions

To select relevant information for the content and to avoid overloading context size of the model, relevant parts should be selected. Our current technique to achieve this is **Depth Retrieval: Augmented Generation (GRAMAG)**.

Peerbot Backend

allama (via python cli and HTTP requests) open source large language model server, written in Go, backed by **Brama.cpp** (C++)

- Efficient serving of large language models
- CPU/GPU/CPU+GPU hybrid inference to partially accelerate models larger than the total VRAM capacity
- Supports many model architectures: llama, gpt4all, qwen2...
- Support for multistate of model quantization techniques and precisions for faster inference and reduced memory use
- Usage Metrics

LLMs for Review Weaknesses

Allowing independent agents to generate separate reviews, followed by a final AI reviewer's synthesis, can improve the response quality. At the same time, the system should filter out superfluous reviews, presenting only relevant insights to the user, or none at all when there are none.

LLMs for Review Weaknesses

Allowing independent agents to generate separate reviews, followed by a final AI reviewer's synthesis, can improve the response quality. At the same time, the system should filter out superfluous reviews, presenting only relevant insights to the user, or none at all when there are none.

RNTuple in Practice

- For maximum optimization opportunities, RNTuple introduces a **new on-disk format** and a **new API**
- At the same time, RNTuple is smoothly integrated with the established ROOT/HENP ecosystem
 - RNTuple data stored in ROOT files
 - Consistent tooling
 - **BRWroot** support
 - **TFileMerger** & **hadd** support
 - **Disk-to-disk converter** **TTree → RNTuple**
 - RNTuple adopts TTree's I/O customization and schema evolution system
 - For RDataFrame code: no change required
 - Based on the specification: 3rd party readers available, e.g. for **Julia**
- For frameworks and power users, RNTuple provides a **modern API** for (multi-threaded) writing and reading
 - Follows C++ core guidelines
 - e.g. smart pointers, runtime errors signaled by exceptions
 - Reviewed by **Intel-CCE**, **Argonne**, **Brookhaven**, **Fermilab**, **PERLELAB**

A TTree with 100,000 entries is converted to the new RNTuple data format using the **hadd** tool.

DAOD per events size reduction

- Reduction seen for most of the domains with few exceptions for some branches
- Should be resolved before HL-LHC
- As example: the latest trigger optimization reduces the size of the trigger domain with factor 4 for both RNTuple and TTree

TTree: Container Sizes per Domain MC23 PHYS

Domain	Size (MB)	Percentage
Electrons	12.96	27.74%
Photons	1.93	4.25%
Truth	4.88	10.67%
Prior Detector	7.35	16.03%
Trigger	6.03	13.23%
Other	2.02	4.41%
Taus	2.08	4.54%
MET	2.08	4.54%
Inver Detector	2.08	4.54%
Total	39.84	100%

RNTuple: Container Sizes per Domain Job: MC23 PHYS

Domain	Size (MB)	Percentage
Electrons	6.52	14.04%
Photons	1.93	4.25%
Truth	4.88	10.67%
Prior Detector	7.35	16.03%
Trigger	6.03	13.23%
Other	2.02	4.41%
Taus	2.08	4.54%
MET	2.08	4.54%
Inver Detector	2.08	4.54%
Total	29.93	100%

Why quantum assisted generative AI?

- 'Classical' generative AI
- Quantum-assisted generative AI

Learnable



Twelve! months of Single Sign-On on Kubernetes

Authors: Antonio Nagel, Sebastian Lopskyer, Paul Van Dyrckx, Hannah Short
Presenter: Ismael Pardo Trobo

Why SSO


The Single Sign-On is one of the most critical services at CERN. It's essential for:

- Improve security**
 - One password for access to all organizations' confidential resources
 - Daily login with post-password session valid up to 12 months
- Improve security**
 - Centralized place to enforce policies or organization level risk (2FA)
 - Less risk of password leakage

Why modernize

- Stability issues caused by performance bottlenecks and operational overheads
- Puppet Module maintained by 1 person
- Scalability issues
- Inefficient resource allocation
- Service Managers were maintaining infrastructure as well as operational responsibilities
- Identity Server (Keycloak) and Cognito (InfiniteID) started to same place
- Impossible to scale and monitor separately


Old Architecture



puppet

Powered on 17th of September 2021

New Architecture




puppet

Benefits

- Performance improved by 300%!
- Outsourced infrastructure management
- Authorized resources can't access their own domain
- Reduced infrastructure service incidentality 85%!
- Scalability provided
- Faster and easier operations
 - Keycloak is idempotent
- Implementation of Cognito principles
- Versioning and tracking of changes
- Easy rollback

New Old Architecture








Metric	Old Architecture	New Old Architecture
CPU	~10000	~3000
Memory	~10000	~3000
Network	~10000	~3000
Disk I/O	~10000	~3000


With Resource Monitoring and Alerting

With Resource Monitoring and Alerting

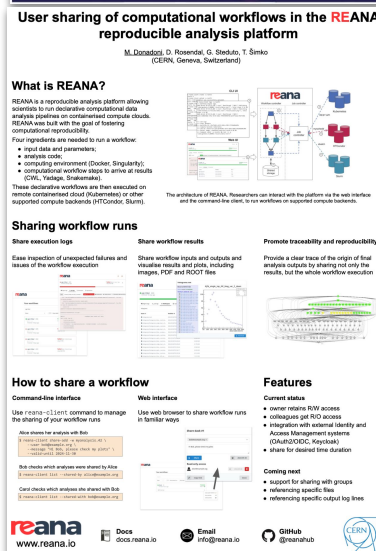
presented at CERN

CHEP 2024

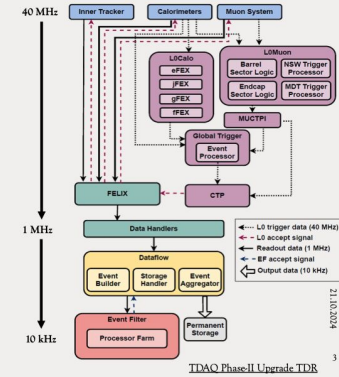









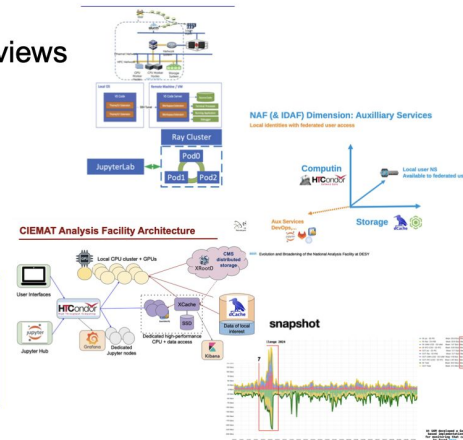
- Several national AFs provide notebook-based interfaces with possibility to scale out with Dask + Condor
- Reproducibility with virtual research environment in REANA
- Effort to federate AFs for US ATLAS with user session allocation across sites via network overlay + kubernetes
- Successful 200Gbps challenge as part of Analysis Grand Challenge targeting HL-LHC scale analysis



- Hardware based L0 trigger
- Software based Event Filter (EF) trigger
- **New trigger and increased readout capabilities:**
 - L0 trigger data reduction: 40 MHz \rightarrow 1 MHz (100 kHz in Run 3)
 - L0 latency 10 μ s (2.5 μ s in Run3)
 - Full Event Building at 1 MHz L0A rate
 - Event size of 4.6MB (1.5MB in Run 3)
 - EF data reduction: 1 MHz \rightarrow 10 kHz (3 kHz in Run 3)



- Chinese HEPS
- German DESY NAF
- Spanish CIEMAT
- Spanish Tier1 and Tier2



- Notebook-based interfaces are widespread
- Scale out with Dask + Condor

Summary

Very dense schedule

A lot of interesting talks and posters

For more please check [timetable](#) and [contribution list](#)



CHEP in numbers

5000 cups of coffee,
2500 cups of tea
1500 liters of water

176 kg of potatoes
470 kg of meat
140 kg of deserts





THAI
SYNCHROTRON
NATIONAL LAB



Chula
Chulalongkorn University

28th Conference on Computing in High-Energy and Nuclear Physics (CHEP 2026)

25-29 May 2026
Bangkok, Thailand

Phat Srimanobhas

Department of Physics, Faculty of Science,
Chulalongkorn University



Questions?

Notes

True cache in Oracle 23ai - looks interesting

Indico IAM is also using AngularJS, they will move to ReactJS

Alice wants to use only whole node scheduling, they pushing sites to move to only this type of resource allocation

Dashboards-as-code with [grafonnet](#) (still alpha, to be released soon) for Grafana, UEM at CERN IT uses it