

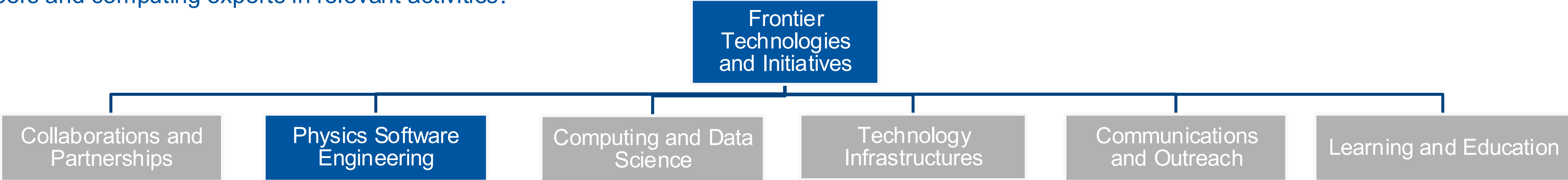
PHYSICS SOFTWARE ENGINEERING

STEFAN ROISER, ENRICO BOTHMANN, MAKS GRACZYK,
DANIELE MASSARO, KENNETH RIOJA, ZENNY WETTERSTEN

FTI GROUP MEETING, 11 DEC 2024

Physics Software Engineering: Functional overview

The Physics Software Engineering Section works on novel software engineering concepts and computing hardware performance aspects of commonly used algorithmic data processing applications in high energy physics (HEP) and provides support and opportunities for skills development and visibility of software engineers and computing experts in relevant activities.



KEY ROLES & RESPONSIBILITIES

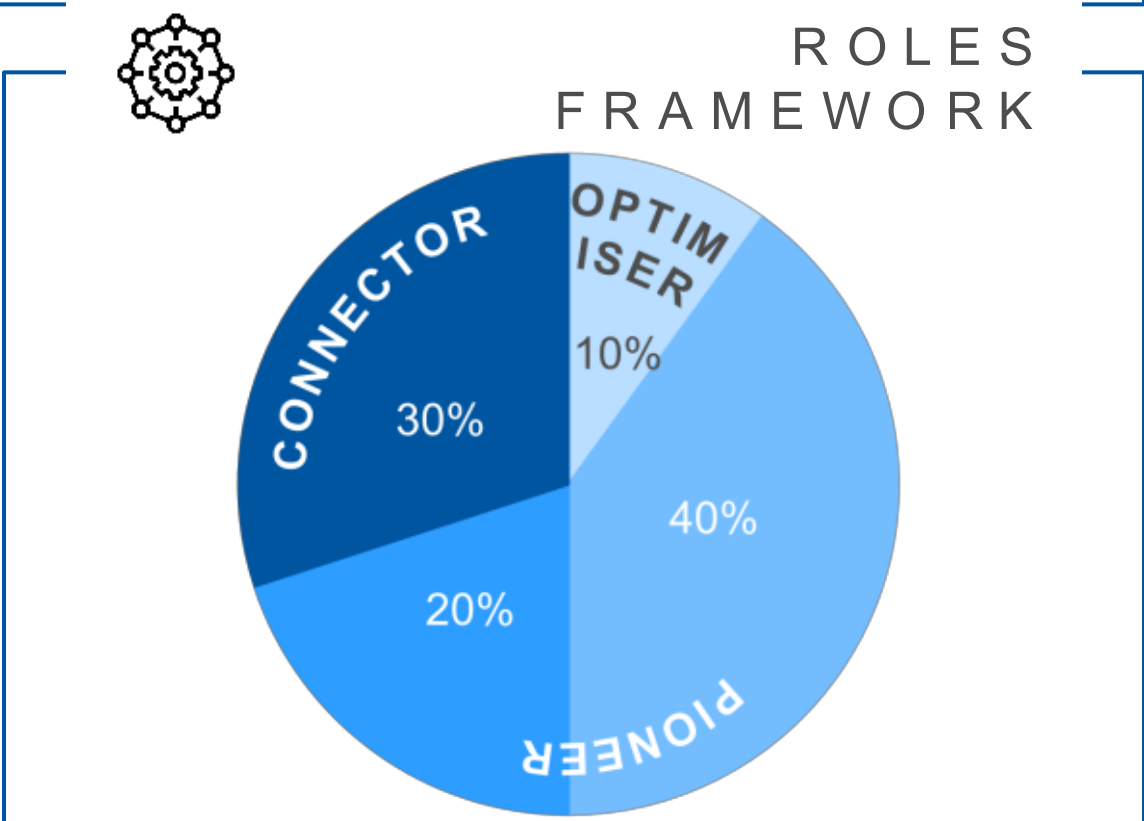
- Improve the performance of software applications commonly used in the HEP community by adopting new computing hardware architectures and engineering techniques
- Investigate and evaluate innovative computing architectures and software engineering concepts for their usability in the domain of physics data processing
- Measure and enhance the sustainability of HEP software at the level of the individual application on current and future heterogeneous computing architectures
- Boost the career paths of research software engineers and scientists via specialised training and recognition of their work in the field

MEASURES OF SUCCESS

- Adoption of software packages within the HEP community, in particular the LHC experiments and their usage on novel computing hardware platforms
- Improvements in data processing challenges of forthcoming data taking periods based on solutions provided by the Section
- Increased adoption of novel software engineering techniques, tools and platforms within the physics software engineering community
- Improved skills development, especially students and early career researchers, training resources and recognition mechanisms

KEY CAPABILITIES

Technology Assessment	Collaboration
Innovation	Training and Recognition
Impact on physics data processing	



GOVERNANCE STRUCTURES

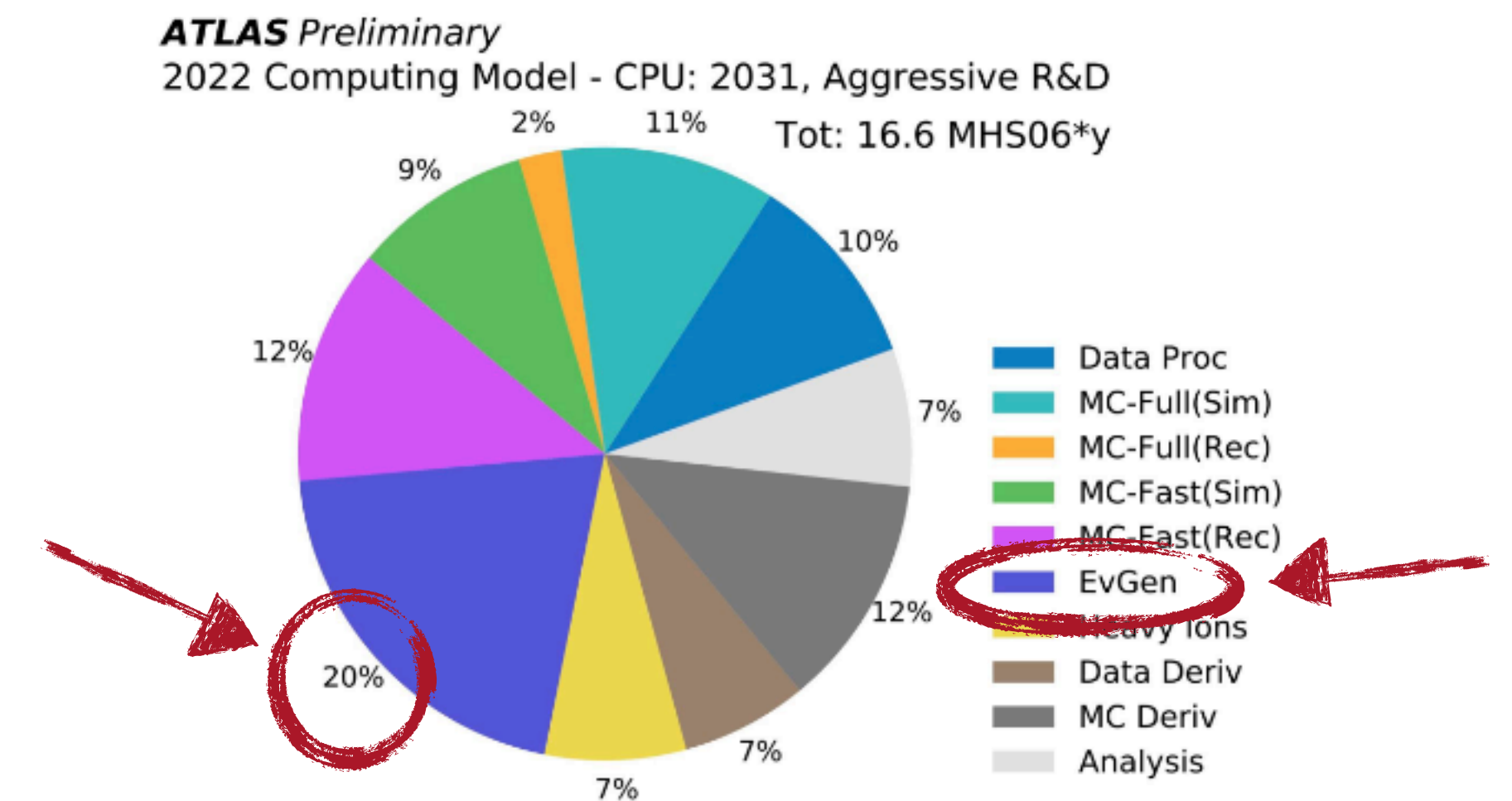
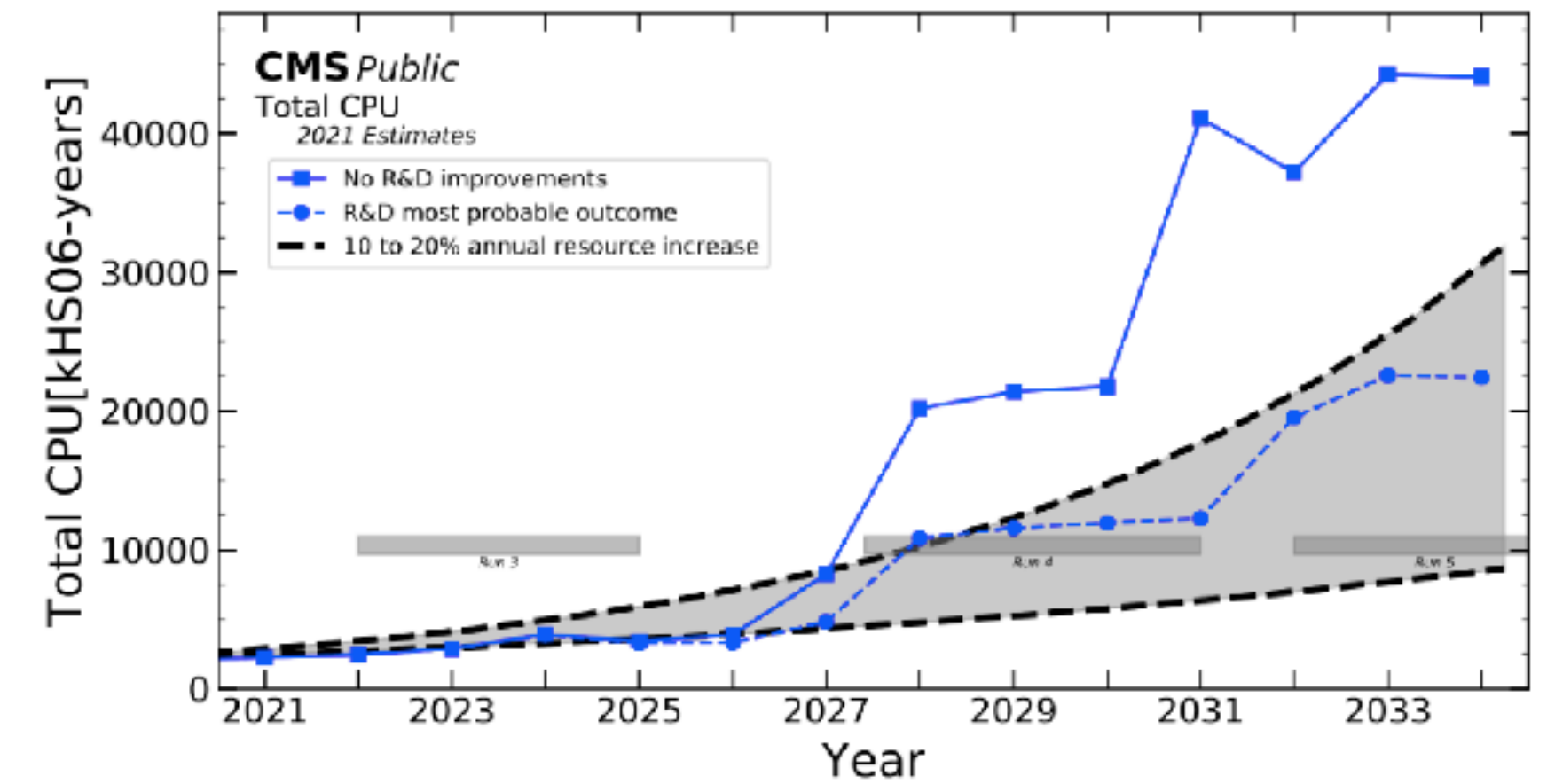
- Bi-weekly Section leaders meetings
- Weekly section meetings to assess status and plans
- Regular meetings with projects and collaborations (CERN EP + TH, NGT, LHC MC WG, MCNet, HSF, ICFA, JENA, EIROforum, EVERSE, SYCLOPS, ...)

PAIN POINTS ADDRESSED

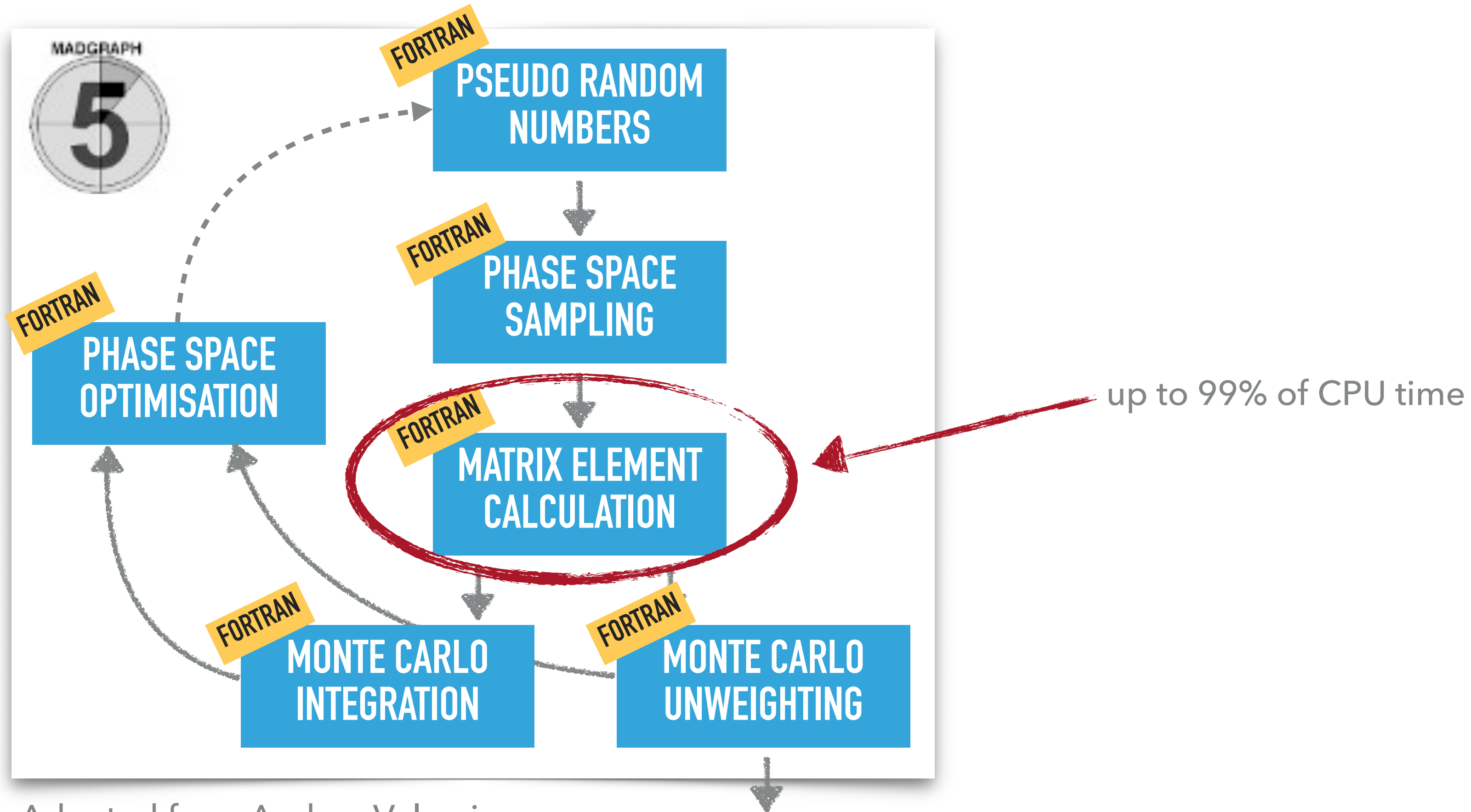
- Communicating the role CERN IT plays in fostering a culture of continuous improvement and cutting-edge solutions more clearly
- Improve integration and exchanges across the RCS sector on software engineering and code development for physics

HARDWARE ACCELERATION OF PHYSICS SOFTWARE APPLICATIONS

- ▶ Monte Carlo Event Generation will use considerable compute resources during HL-LHC, due to
 - ▶ increased number of simulated events
 - ▶ higher precision of calculations
- ▶ Strategy: overcome shortcomings by hardware acceleration
 - ▶ Via vector instructions on CPUs (SIMD)
 - ▶ Parallel execution on GPUs (SIMT)
- ▶ Status: A hardware accelerated release for leading-order calculations of Madgraph5_aMC@NLO is available for use by the experiments

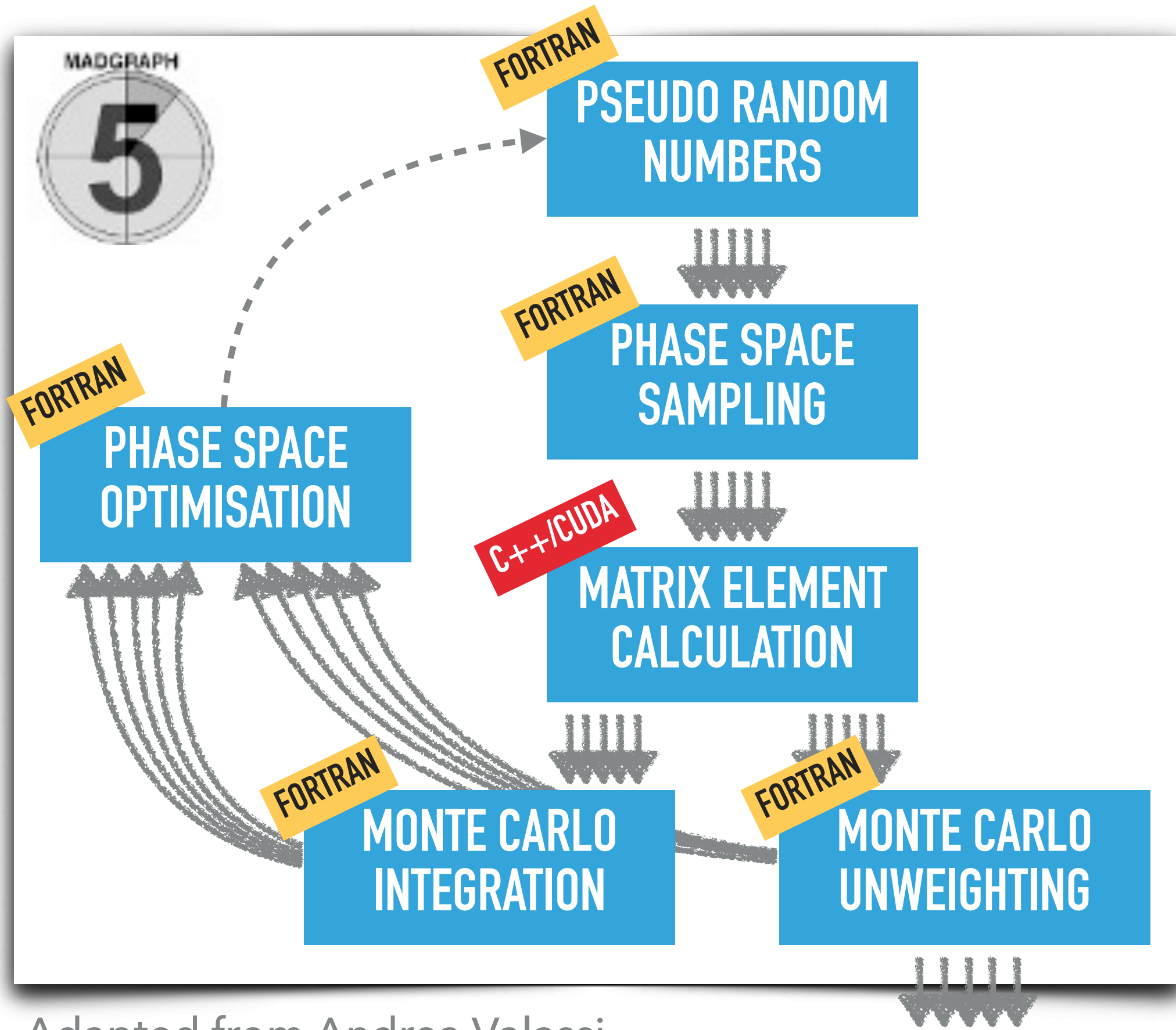


MADGRAPH5_AMC@NLO



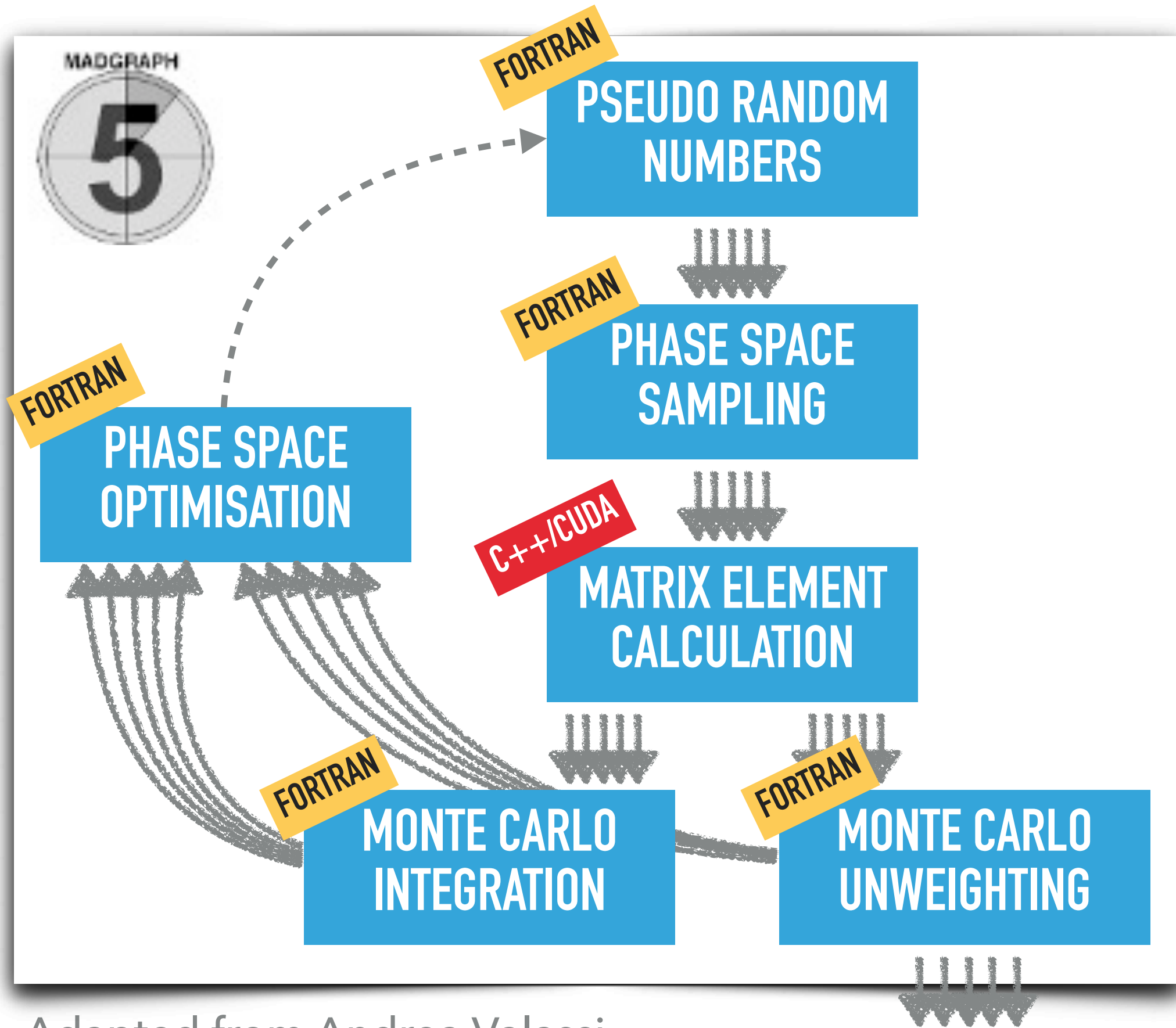
Adapted from Andrea Valassi

MADGRAPH5_AMC@NLO

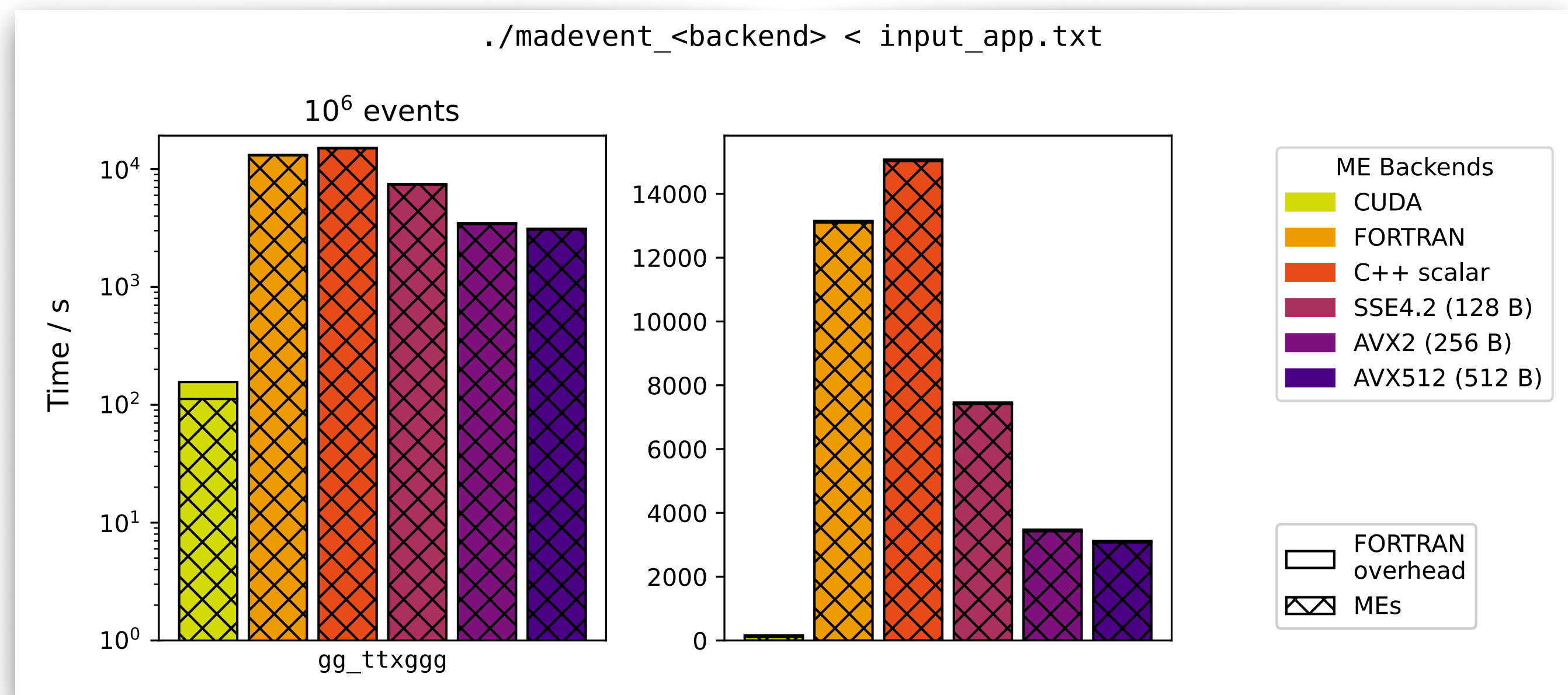


Adapted from Andrea Valassi

MADGRAPH5_AMC@NLO



Adapted from Andrea Valassi



Daniele Massaro

(RE-)USE OF WORK ON MADGRAPH HARDWARE ACCELERATION

CMS INTEGRATION

Quantifying the Computational Speedup with MG4GPU for CMS Workflow

Jin Choi
On behalf of the CMS Collaboration

Introduction
The most time-consuming aspect of Hard Scattering is the Matrix Element (ME) calculation. The Madgraph4GPU (MG4GPU) project is dedicated to porting the ME calculation to vectorized CPUs and GPUs.

Physics Processes
Conducted extensive testing of the most common and complex LO processes within the CMS framework, encompassing:
- Drell-Yan (DY) production, capable of generating up to four jets
- Top-Quark Pair (TT) production, capable of producing up to three jets

HPC Configurations
- Lxplus HTCondor Pool with AMD EPYC 16 CPUs + A100 / Intel Xeon 48 CPUs + H100
- Seoul National University HTCondor Pool with Intel Xeon 80 CPUs
- NERSC Permutler SLURM Batch with AMD EPYC 4x64 CPUs

Performance Results

Production Time	FORTRAN	CPP-AVX2	CUDA
DY+01234j	424h 36m	133h 38m	9h 32m
TT+0123j	253h 36m	155h 28m	3h 9m

- Different levels of parallelism employed, normalized 16 parallel madevent executions.
- CPP-AVX2 exhibits approximately 2-3x speedup; CUDA achieves 0(10)x.

Event Generation - Jet-Binned study (TT)
- Multiple events are generated with a single madevent execution
- As the complexity of events increases, the production time required to execute them also tends to grow.
- Refining ME requires a significant amount of time for FORTRAN/CPP-AVX2, resulting in substantial differences for inclusive TT and TT+3j, but not for CUDA.

Event Generation - Inclusive study
- Multiple events are generated with a single madevent execution, exhibiting a linear trend.
- CPP-AVX2 demonstrates approximately 1.5x speedup; CUDA achieves approximately 7x speedup when generating 100k events.

Conclusion
MG4GPU demonstrates substantial speedups in both gridpack production and event generation for the CMS workflow, making it a promising candidate for future large-scale simulation.

Reference
[1] A. Valassi et al, Developments in Performance and Portability for MadGraph_mC@NLO
[2] CMS Collaboration, Quantifying the computational speedup with madgraph4gpu for CMS workflow

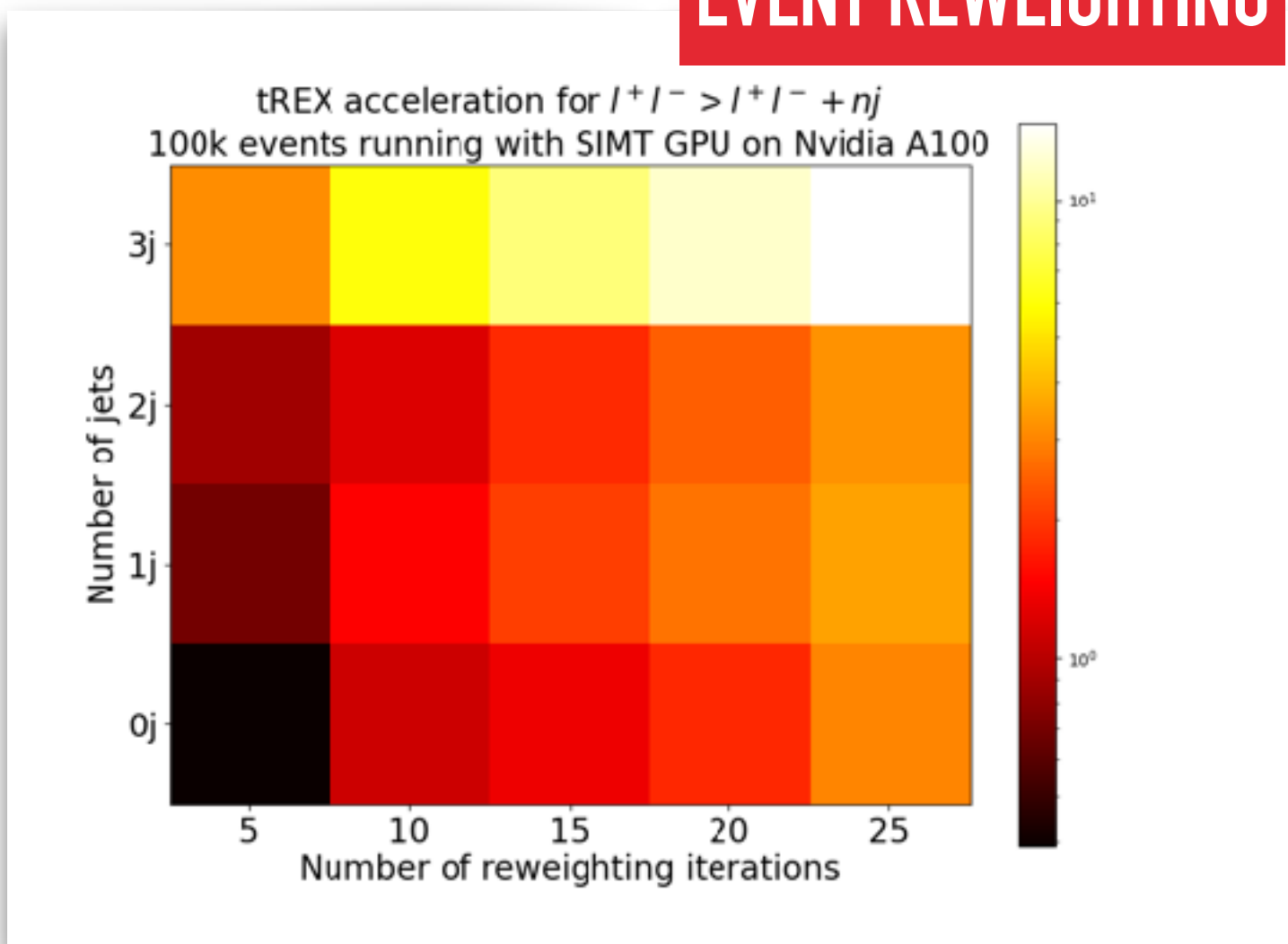
October 19 - 25, 2024
CONTACT: choij@cern.ch
cms-generator-conveners-mfga@cern.ch
cms-physics-conveners-GEN@cern.ch
CHEP 2024

CHEP 2025, Krakow
J Choi, S Bhattacharya (CMS)

QUANTUM ALGORITHMS
IBM, SOFIA, MICHELE

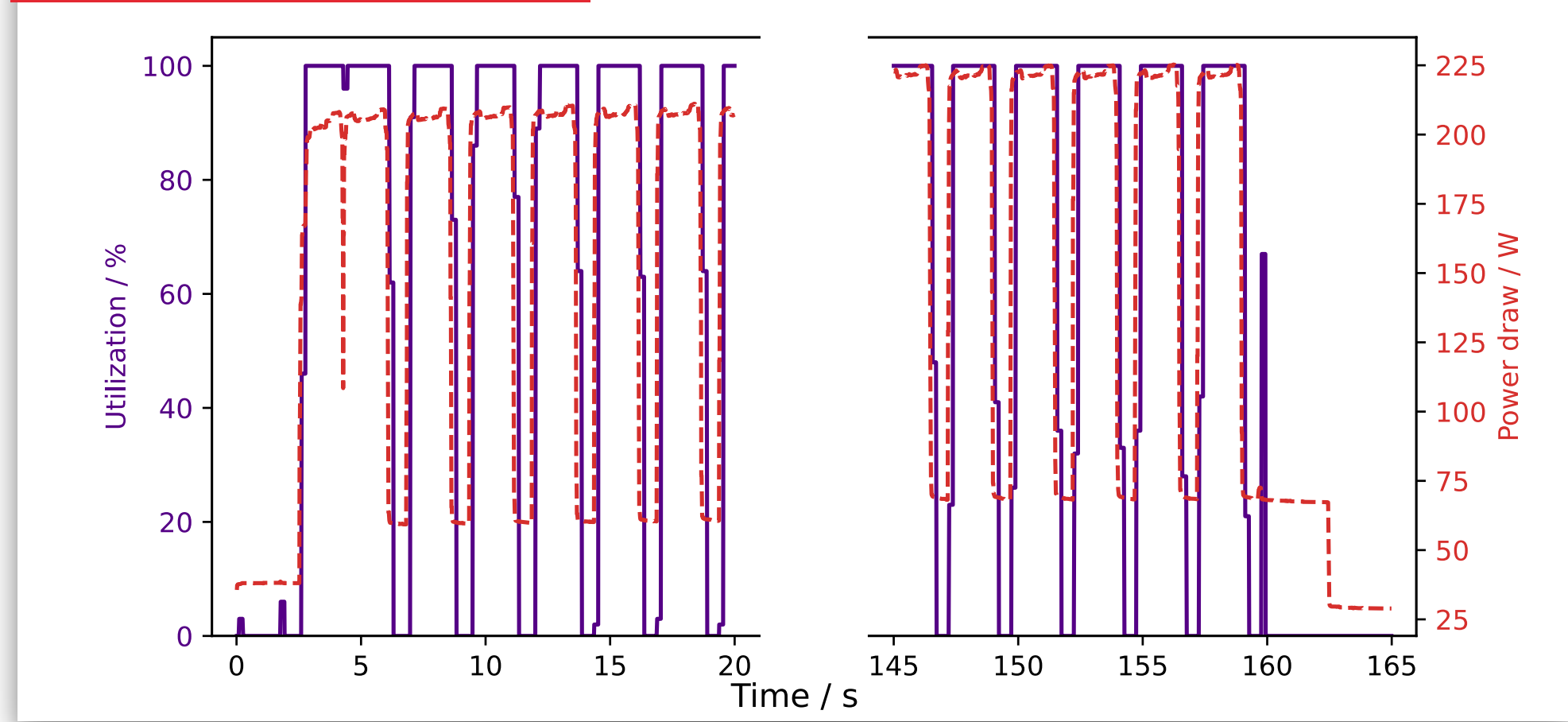
RE-USING MATRIX ELEMENTS
POWHEG COLLABORATION

EVENT REWEIGHTING



Zenny Wettersten

POWER CONSUMPTION



MADGRAPH ON FPGA

Héctor Gutiérrez, Luca Fiorini, Alberto Valero, Francisco Hervás
IFIC (Universitat de València - CSIC)
HIGH-LOW TED2021-130852B-I00
27th CHEP - Krakow
Poster Session - 21th October 2024

ABSTRACT
The escalating demand for data processing in particle physics research has spurred the exploration of novel technologies to enhance efficiency and speed of calculations. This study presents the development of a porting of MADGRAPH, a widely used tool in particle collision simulations, to FPGA using High-Level Synthesis (HLS). Experimental evaluation is ongoing, but preliminary assessments suggest a promising enhancement in calculation speed compared to traditional CPU implementations. This potential improvement could enable the execution of more complex simulations within shorter time frames.

FPGA IMPLEMENTATION
Developing applications for Alveo [2] involves two parts: programming the host, which runs on x86 processors, and programming the FPGA, which accelerates specific functions. Host development is similar to regular software development, using C++ and the OpenCL API to manage tasks on the FPGA, transfer data, and program the FPGA in real-time, optimizing its resources.

RESULTS
Comparison of processing times: CPU vs GPU vs FPGA

Process: $e^+e^- \rightarrow \mu^+\mu^-$	CPU (3.9th Gen Xeon E5-2680v4 (7-13700))	GPU (NVIDIA GeForce RTX 3090)	FPGA (ALVEO U250)
200k-400k	~100s	~100s	~100s
800k-1000k	~200s	~200s	~200s
1.6M-2.0M	~400s	~400s	~400s

REFERENCES
[1] MadGraph (v4.1). MadGraph: A program for event generation in high-energy physics. Retrieved September 19, 2024, from <https://madgraph.hep.org/html/>
[2] Xilinx (v4.1). Vitis development flow. Xilinx. <https://www.xilinx.com/products/integration-solutions/vitis-dev-flow.html>

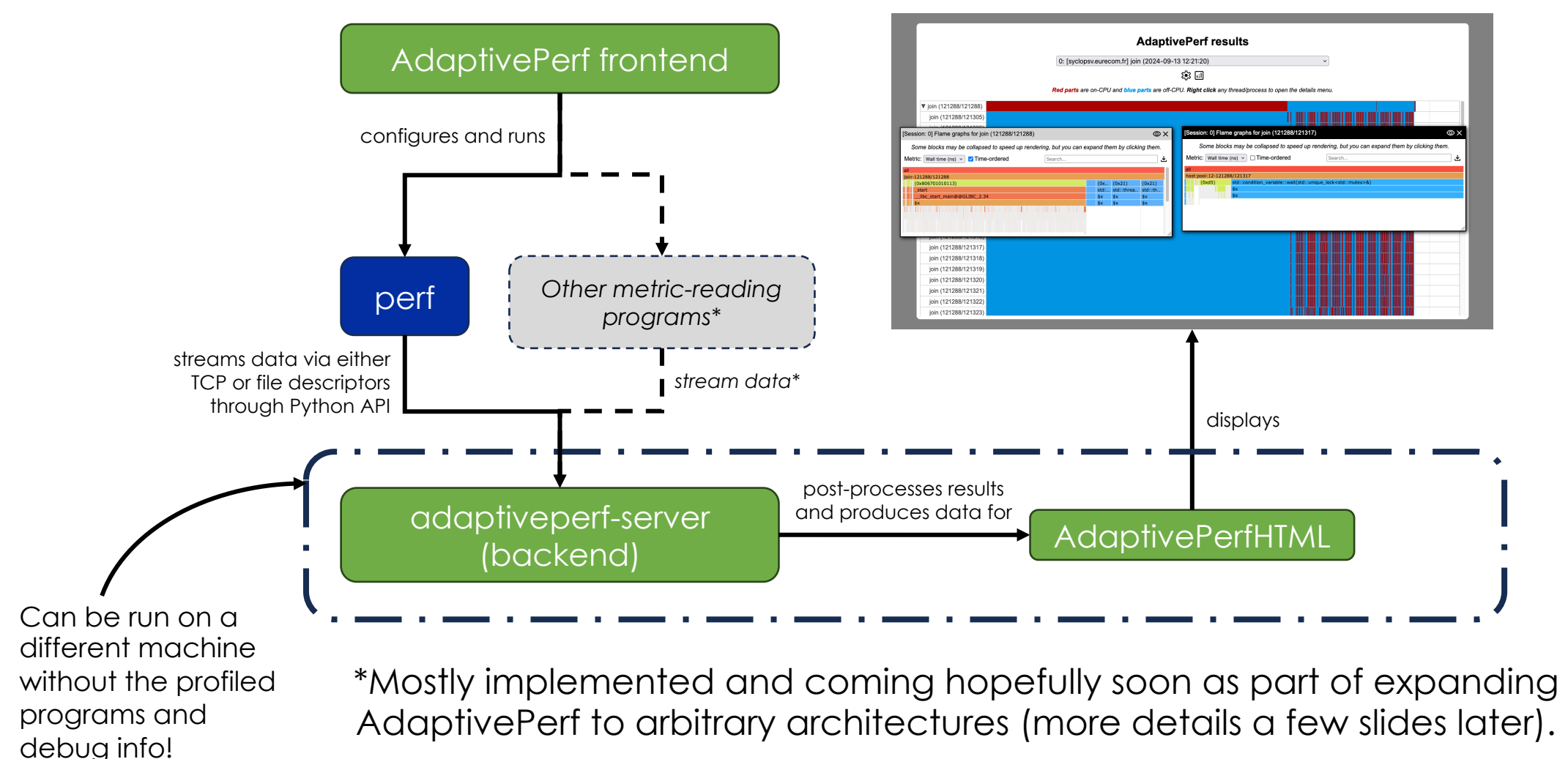
CHEP 2025, Krakow
Héctor Gutiérrez et al. (Valencia)



- ▶ Project on hardware acceleration with open standards, based on SYCL and RISC-V
- ▶ Involvement in work package on performance profiling and integration

DISCUSSING USAGE BY CERN EXPERIMENTS

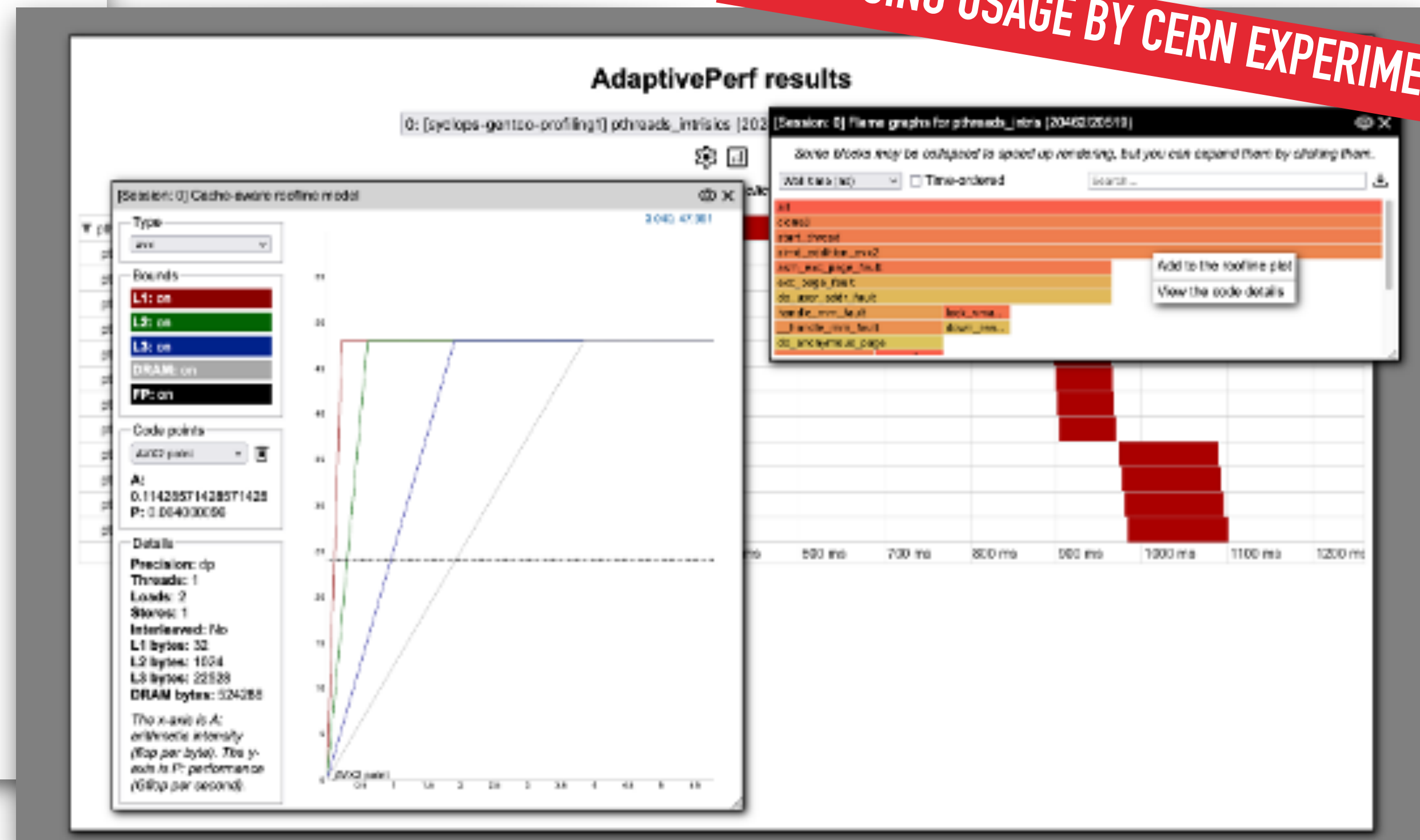
How does AdaptivePerf work under the hood?



05/12/2024

AdaptivePerf talk for NGT task 1.7 meeting

5



- ▶ Build a virtual institute to promote best practices in research software engineering
- ▶ Lead WP "capacity building & recognition"

Thank you!!

The TeSS Training Portal
<https://tess.elixir-europe.org>

How can TeSS help you?

Search the portal for courses, events, videos, presentations, learning pathways, handbooks...
All types of resources at all levels for leveraging computational resources in the life sciences.

Search TeSS...

> 100 content providers

255 upcoming events

14,664 past events

Browse the catalogue

Events

Discover the latest training events and news from ELIXIR nodes and 3rd-party providers.

Start browsing →

Materials

Browse the catalogue of training materials offered by ELIXIR nodes and 3rd-party providers.

Start browsing →

Workflows

Create training workflows to visualise learning steps and link to resources specific to your training needs.

Start browsing →

Trainers

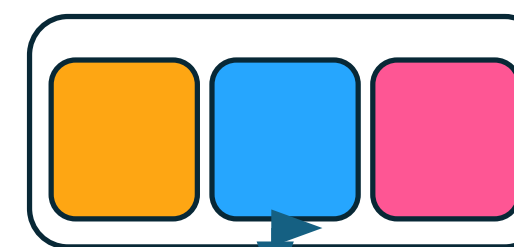
Browse a directory of trainers based on their skills and expertise.

Start browsing →

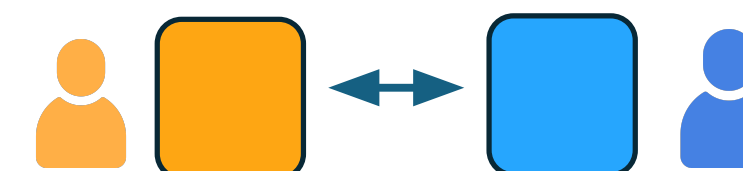
3,473 materials



multi-tenant TeSS

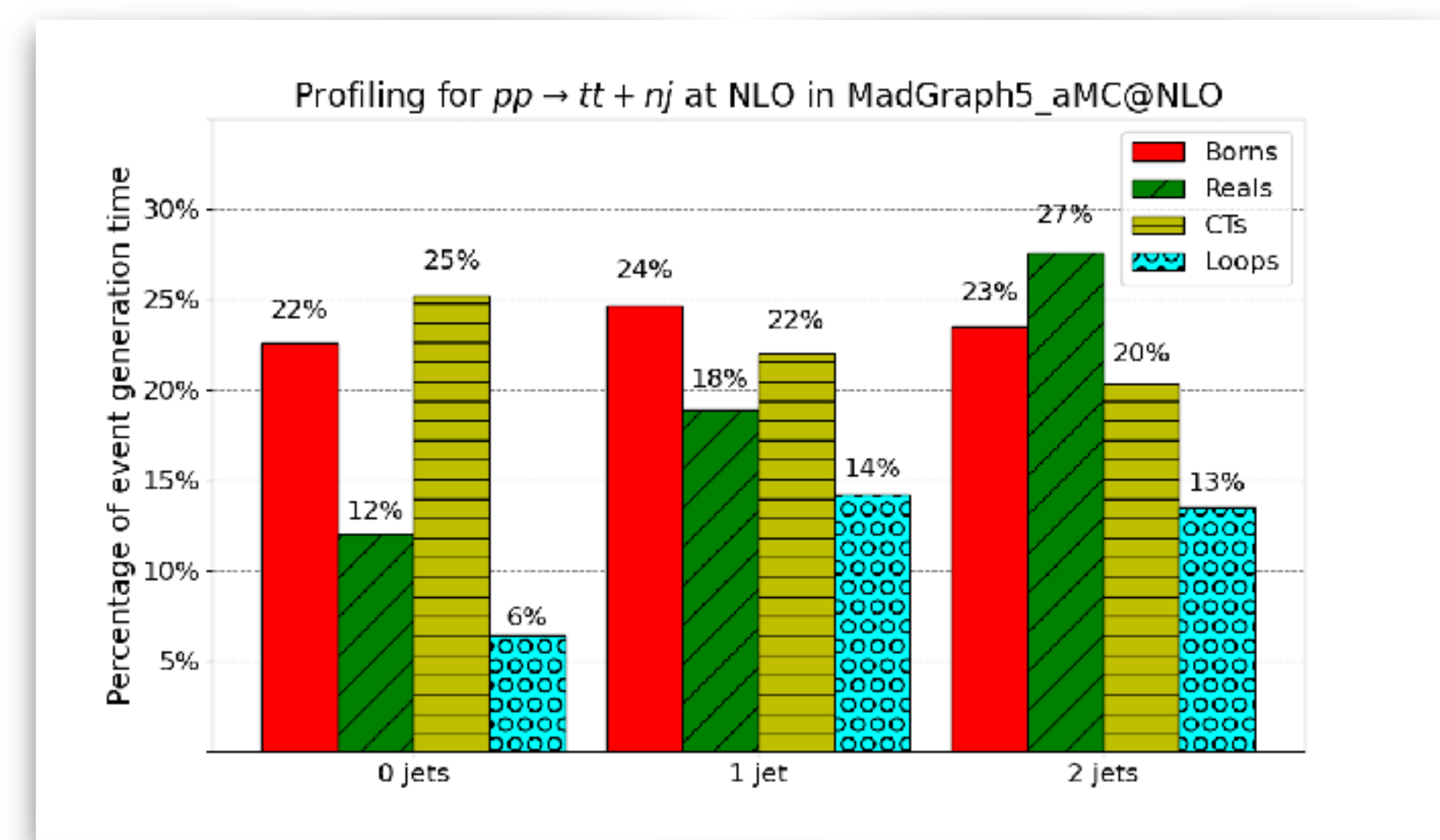


TeSS-eXchange



WHAT'S NEXT?

- ▶ MC event generation
- ▶ Deployment of leading-order for experiments
- ▶ Develop a prototype for next-to-leading-order calculations
- ▶ Expand the realm of projects with Sherpa/Pepper
- ▶ Development of new AdaptivePerf features and its usability by CERN experiments
- ▶ Create instance of a training catalogue, usable from within HEP and connect it to other sciences



Zenny Wettersten