Triggering Discoveries in High Energy Physics III
*Vysoké Tatry, Slovakia*

# The NextGen Triggers Project
## Concept and initial activities

**Silvio Donato (University and INFN Pisa)**
**on behalf of the NextGen Project**

***[credits to Axel Naumann]***

NexTGen
Next Generation Triggers

Università di Pisa
INFN Pisa

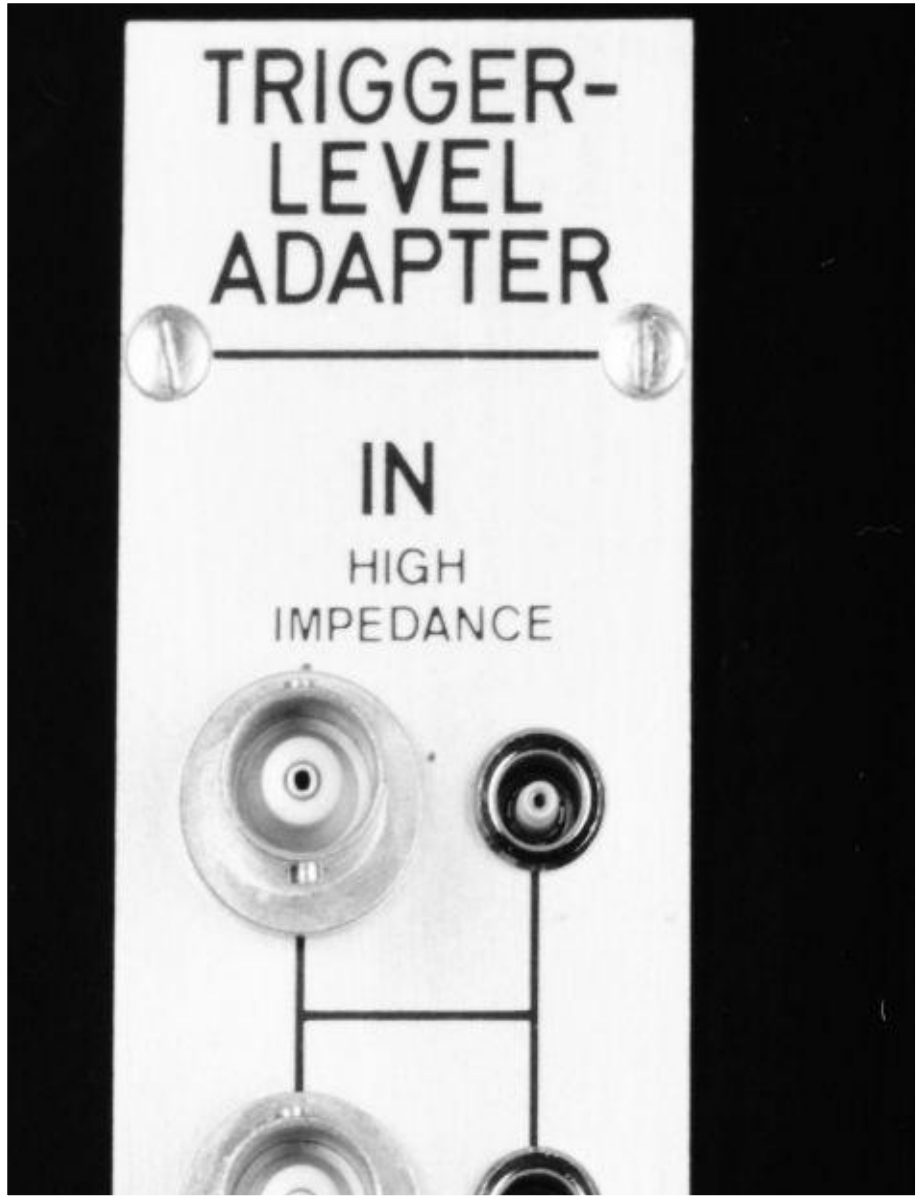
# IMPROVE PHYSICS REACH

## THROUGH TRIGGERS

**ε** Efficiency: more relevant data given data rate

**ops** More cycles / seconds means smarter selection

**.yml** Exploring trigger updates for new physics sensitivity, with improved models, generators, and simulations



# The Project
# NEXT GENERATION TRIGGERS

Five years: 2024-2028 enabled by external donation, combining

- ATLAS, CMS;
  limited participation of ALICE, LHCb
- CERN's Theory & IT departments
- CERN's Exp Physics Software group

Project goals (proposal)

- opportunity for wider R&D
- improve LHC experiments in 2028+
- invest in community

# HOW

## Common R&D + Training

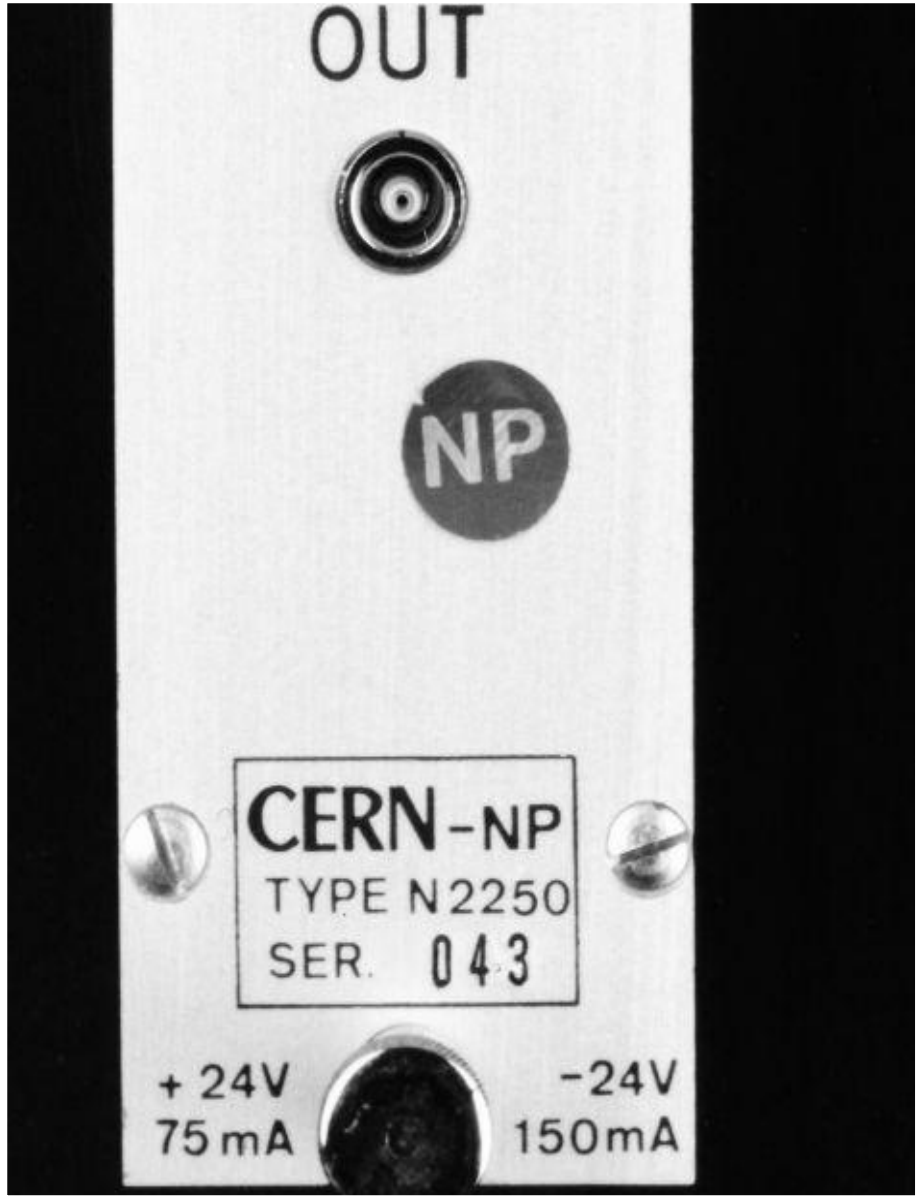
- Combining all parties: 2 (+2) experiments, IT, theory, experimental physics software

## Experiment-specific R&D

- ATLAS+CMS define their R&D requirements
- Benefitting from common R&D, training

## Results are open

- Open access, open source, including training
- Embedded in experiments



# THE KEY OBJECTIVES

**More than technical work**

## Goals

- To get more physics information out of the HL-LHC data.
- To uncover as-yet-unseen phenomena by more efficiently selecting exotic and rare physics events thanks to better models and data processing techniques.

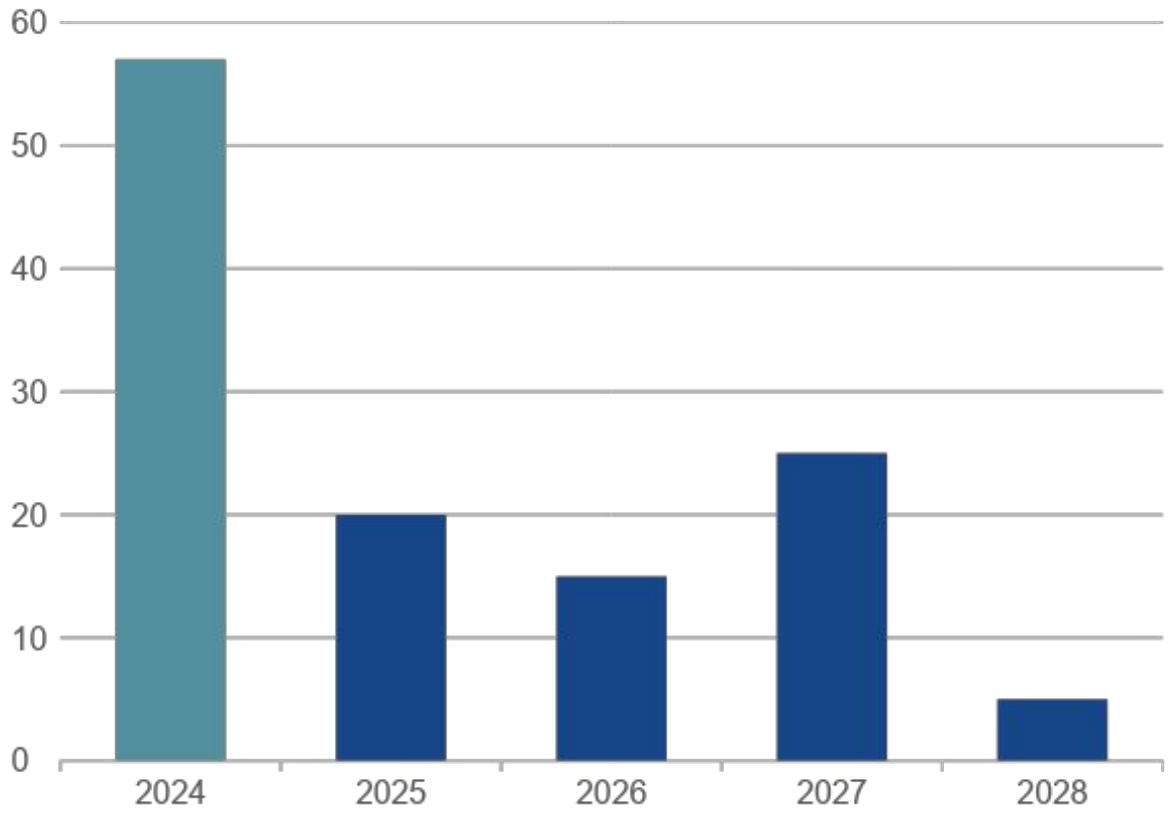
## Technologies

- ML and classical algos: invent, optimize, benchmark
- FPGAs, GPUs, high-performance computing, more efficient architectures

## Community

- Define common objectives across different experiments and Institutes
- Train future researchers on new computing techniques for real applications
- Promote Open Science principles and contribute to open source development of AI/ML technologies

# First action? Hire the Experts!




## Target effort

- ~280 FTEs planned in total across 5 years
- 57 new hires for 2024
- Mixed profiles: master, PhD, post-doc, staff

# NextGen Workshops

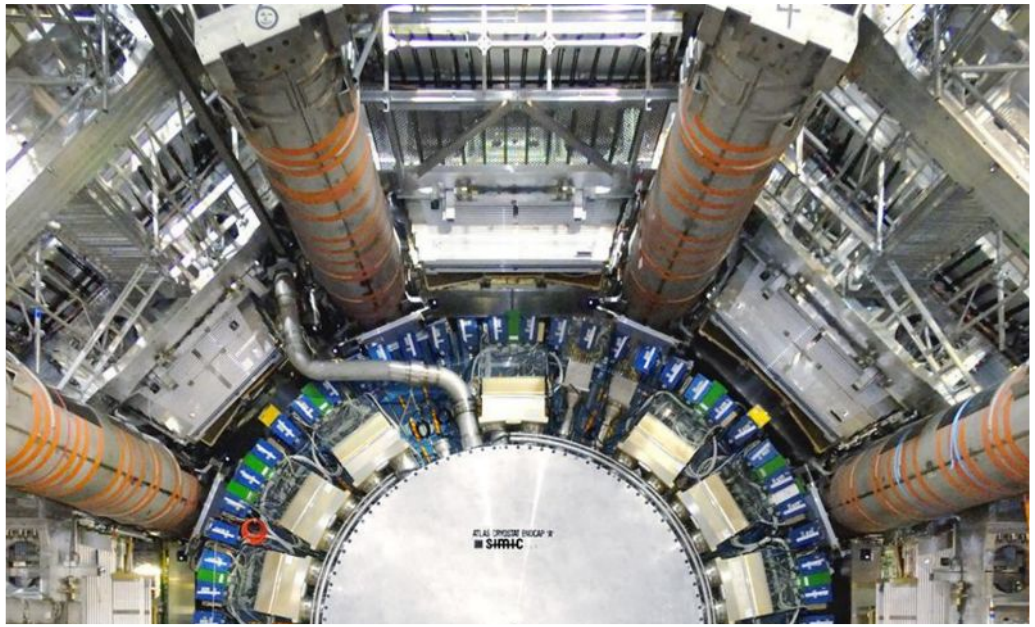
**1st Technical Workshop**

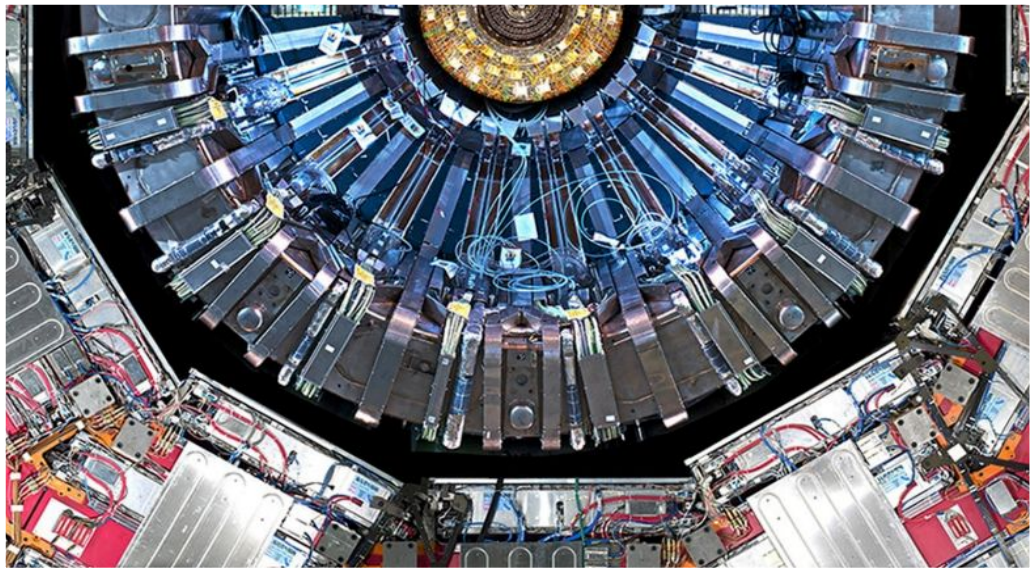
- Discussing technical progress across tasks
- November 25-27, 2024 @ CERN
- https://indico.cern.ch/event/1421629

# Topical Workshop Examples

**Bringing the tasks' communities together**

- NGT Algorithm Workshop – Lattice QCD at large scale on exascale computing facilities
- hls4ml HEP Community Forum





# NEXT GENERATION TRIGGERS

## R&D Topics

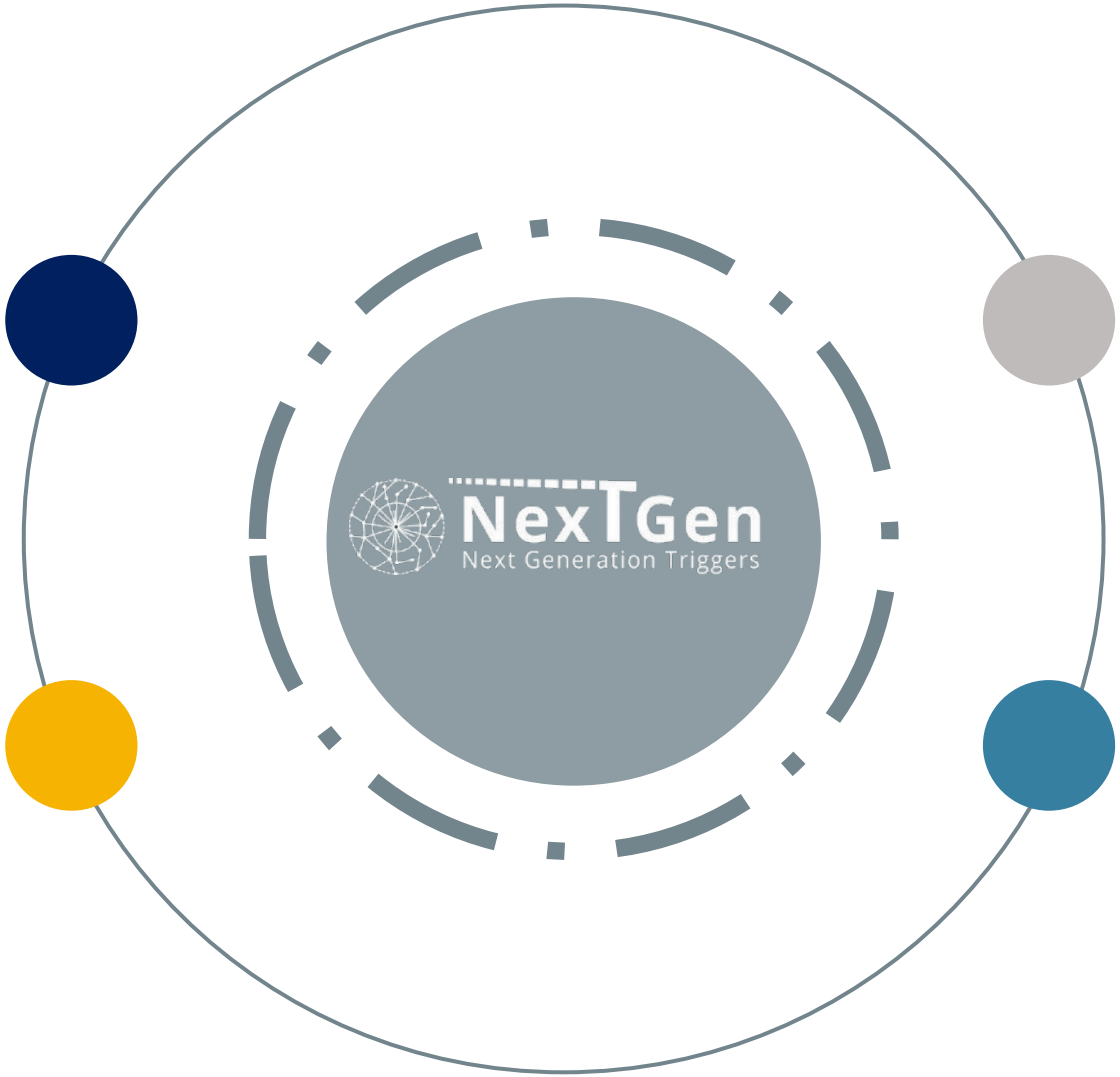
PROGRESS YOU MIGHT CARE ABOUT

# THE NEXT GENERATION TRIGGERS

PROJECT IS BROKEN DOWN INTO **FOUR WORK PACKAGES:**

**WP1**

Infrastructure,
Algorithms and Theory




**WP2:**

Enhancing the ATLAS Trigger

and Data Acquisition

**WP4:**

Education Programmes

and Outreach

**WP3:**

Rethinking the CMS Real

Time Data Processing

# WP1: Main goals

**Overarching rationale**

- Providing constituent R&D to NextGen themes
- Interleaving experiments, IT, theory, software engineering, ML, accelerators

**Infra** O(100) GPUs, including wide range of GPUs (vendors, specs) for benchmarking workflows

**ML** Optimization of ML models for inference hardware;  offering training + optimization service

**Lattice-quantum-field theory** with HPC optimization and novel algos (ML, quantum-inspired)

**Generators** Wider physics reach, applicability to experiments' triggers

**Heterogeneous compute** Novel approaches for accelerators, reducing vendor lock-in

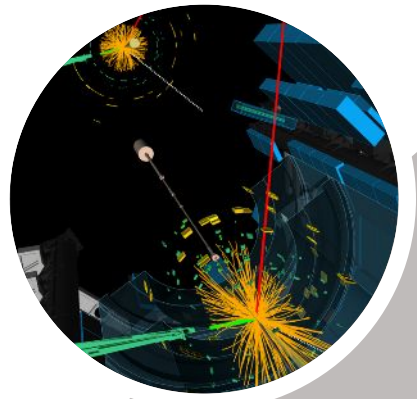
# EXAMPLE: OPTIMIZING MEMORY LAYOUT FOR INFERENCE AND ACCELERATOR USE

**Array-of-structs**

vec<jet{pT, eta, phi}>

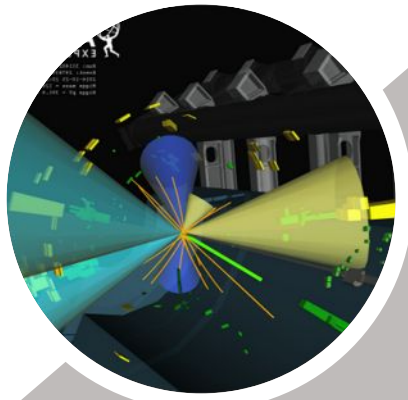
**Struct-of-arrays**

vec<pT>, vec<eta>, vec<phi>



All LHC experiments implement this transformation - differently!

Rely on **C++ reflection**?



Decouple experiments' frameworks from vendors' inference libraries, optimized for **inference from GPU**

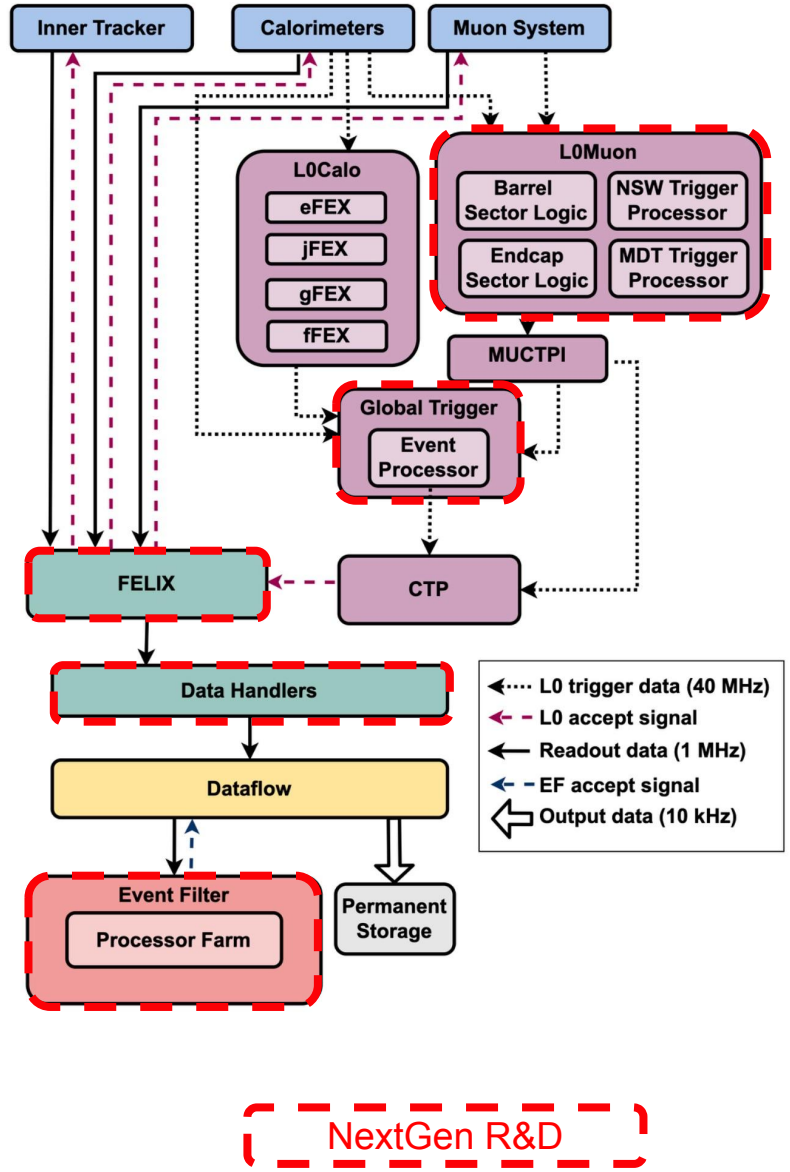
# WP2: Main goals

## Diagram: overall T-DAQ architecture for Phase-II

- 1 MHz L0 rate
- 4.6 TB/s to Event filter
- Event filter data reduction 1 MHz -> 10 kHz

## Main goal for the ATLAS NextGen work package

- ATLAS **R&D,** novel **(AI) approaches** and **innovation** for several sub-systems (dotted boxes)
- **Level-0** hardware trigger: **L0Muon** and **Global Trigger**
- **Event Filter** event processing: **Track reconstruction** in the **Inner Tracker** and **Muon detectors,** plus **ACTS** tracking software infrastructure
- **Novel trigger signatures and physics optimisation**



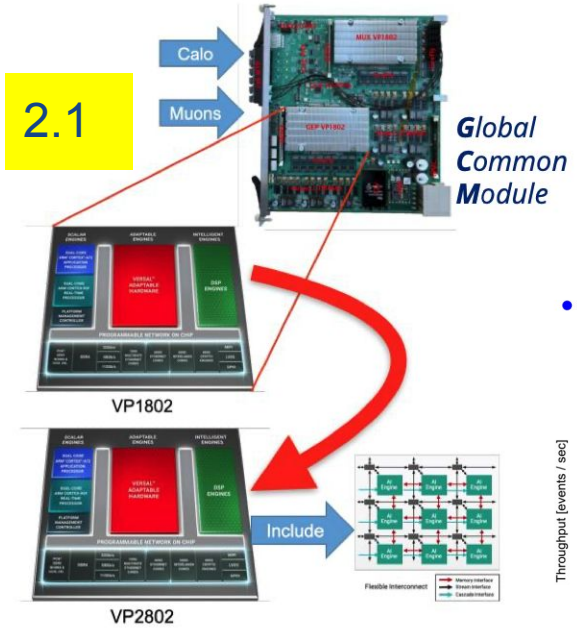
# WP2 – Highlights

- Work is embedded in ATLAS T/DAQ and offline groups
- Building on ongoing ATLAS work, first new ideas and initiatives
- Actively collaborating with WP1 and WP4
- Benefit from HW provided by WP1/IT, CERN openlab and ATLAS test-beds

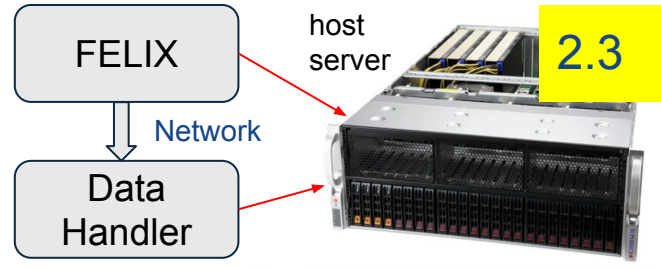
## Some examples of ongoing work and new initiatives

- Plan to procure L0Global board with novel FPGA (VP2802) with AI engine (2.1)
- Studies of ML based L0Muon trigger techniques started (2.2)
- First developments on readout, purchased prototype host server (2.3)
- Fast ITk tracking optimisation, GNN4ITk tuning and throughput tests (2.4)
- Novel ACTS Muon segment pattern and fitting (2.5)
- Enabling ACTS EF reconstruction, first results on full TRACCC chain (2.6)
- Developing trigger analysis kit, ATLAS survey to collect input (2.7)
- Study of L1 trackless b-tagging for ITk inner pixel replacement (2.4+7)

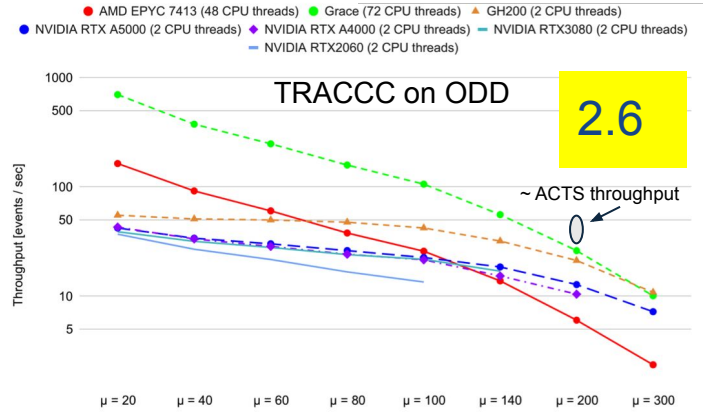
2.1



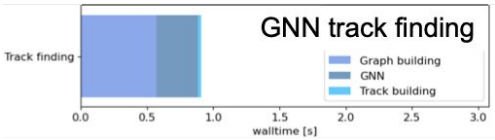

2.3



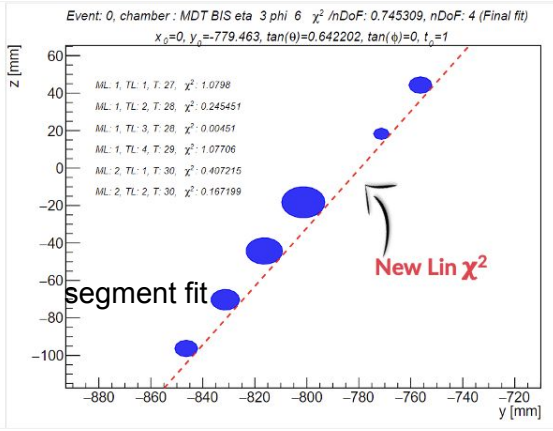

2.6




2.4



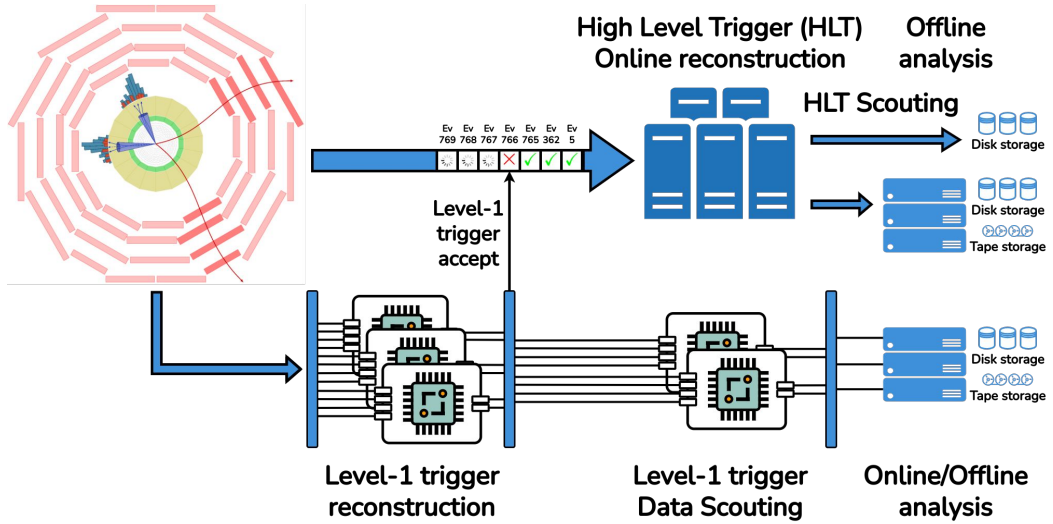

2.5



# WP3: Main goals

- At HL-LHC, up to 200 pile-up interactions: CMS is upgrading the L1T and HLT to enable the same physics program we are doing now (at 60 pile-up)
  - **L1-Trigger (L1T): ~750 kHz + Scouting, High Level Trigger: ~7.5 kHz + Scouting**

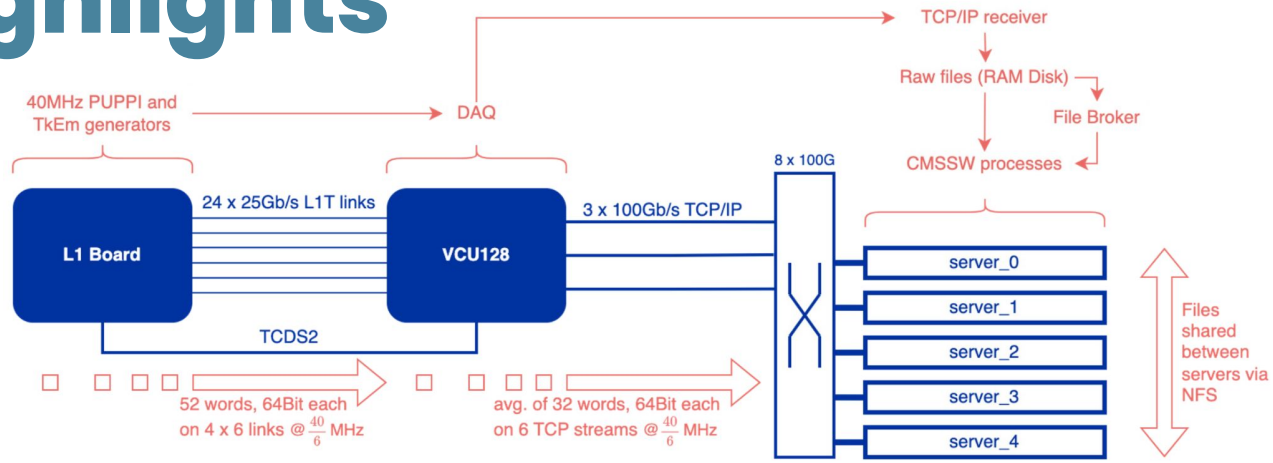
- What if New Physics is buried under the bulk of background events we are throwing away due to the trigger selections? **NGT goal is to extend the CMS discovery and precision measurement reach by:**
  - Redesign the data collection and scouting strategy to **reduce the need to reject events in the Level-1 and High-Level CMS triggers** aiming at complementing the current workflows
  - **Replace the trigger filtering task** with an event processing task **similar to what happens with offline** events stored on disk
  - Achieve all this by exploiting **advanced AI solutions and heterogeneous frameworks**



# WP3: Highlights

## R&D has started successfully

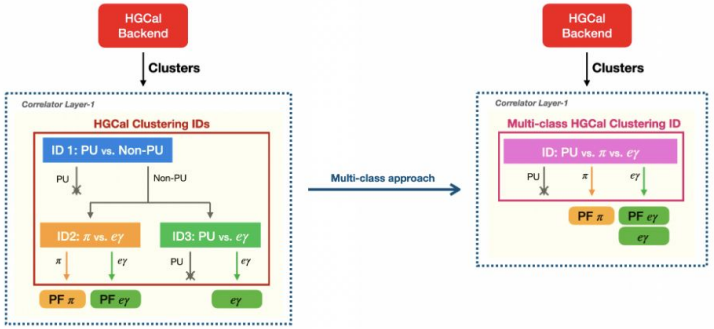
- Work embedded in CMS L1T/HLT offline groups
- Building on ongoing CMS work
  - L1T Run-3 Scouting Demonstrator,
    first L1T AI-models in Run-3 data taking,
    GPUs in production in HLT Farm since Run-3
- Actively collaborating with WP1
  - Specific needs for hls4ml/Conifer for new R&D
  - Heterogeneous frameworks and fast inference




Overview of the new demonstrator system for the HL-LHC L1 Scouting

## Some examples of ongoing work and new initiatives

- **L1T Scouting for HL-LHC:** benchmark analyses defined and tested on new hardware **(first demonstrator in place!)**
- **AI@L1T:** ultrafast jet-taggers models, **multiclass clusters ID for e/γ,π, pile-up, …**
- Operational practises for unsupervised anomaly detection trigger for L1T Run 3
- Systematic investigation of RECO performance in Phase-2 HLT, Phase-2 Offline and Run-3 Offline to identify bottlenecks
- Towards evolving CMSSW into a distributed application for Phase-2 HLT Farm: implementation of a client server prototype
- Evaluating impact of RAW data compression for HLT




New ML-based algo in L1T FPGAs for e/γ, π, pile-up clustering ID

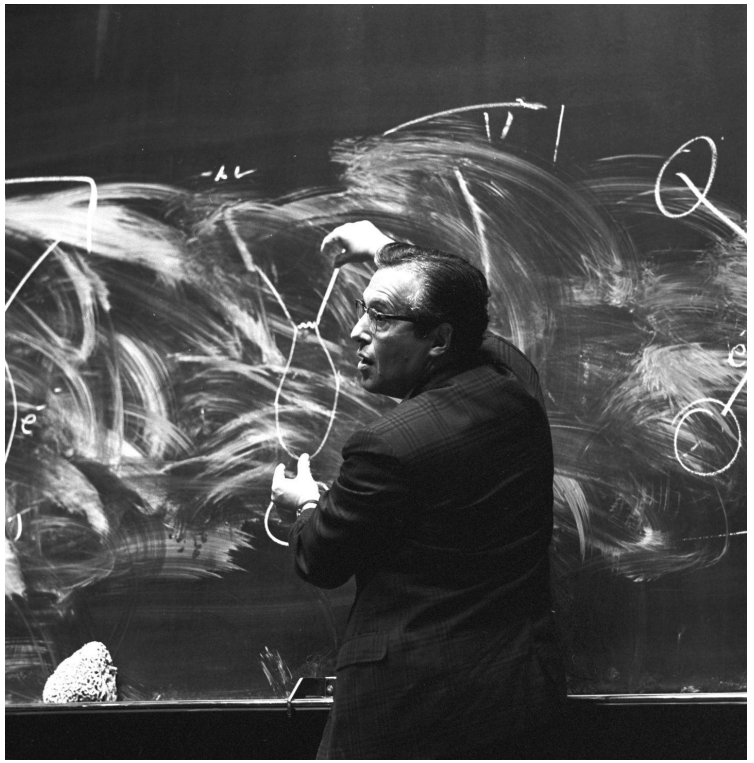
# WP4: Education & Outreach

**Vision: train, sustainably**

- Develop project members knowledge with an aggressive education programme and multiple training opportunities
- Reuse and enhance existing infrastructures to cover education and training needs of NextGen staff

**Exchange, impact beyond NextGen**

- Visiting scientists, public seminars
- Develop a coherent message for outreach and external communication, including the web site.
- Make NextGen knowledge accessible also outside the project
  - Conference participation and online presence
  - Support NextGen staff to contribute to Outreach initiatives, Schools, and Education activities





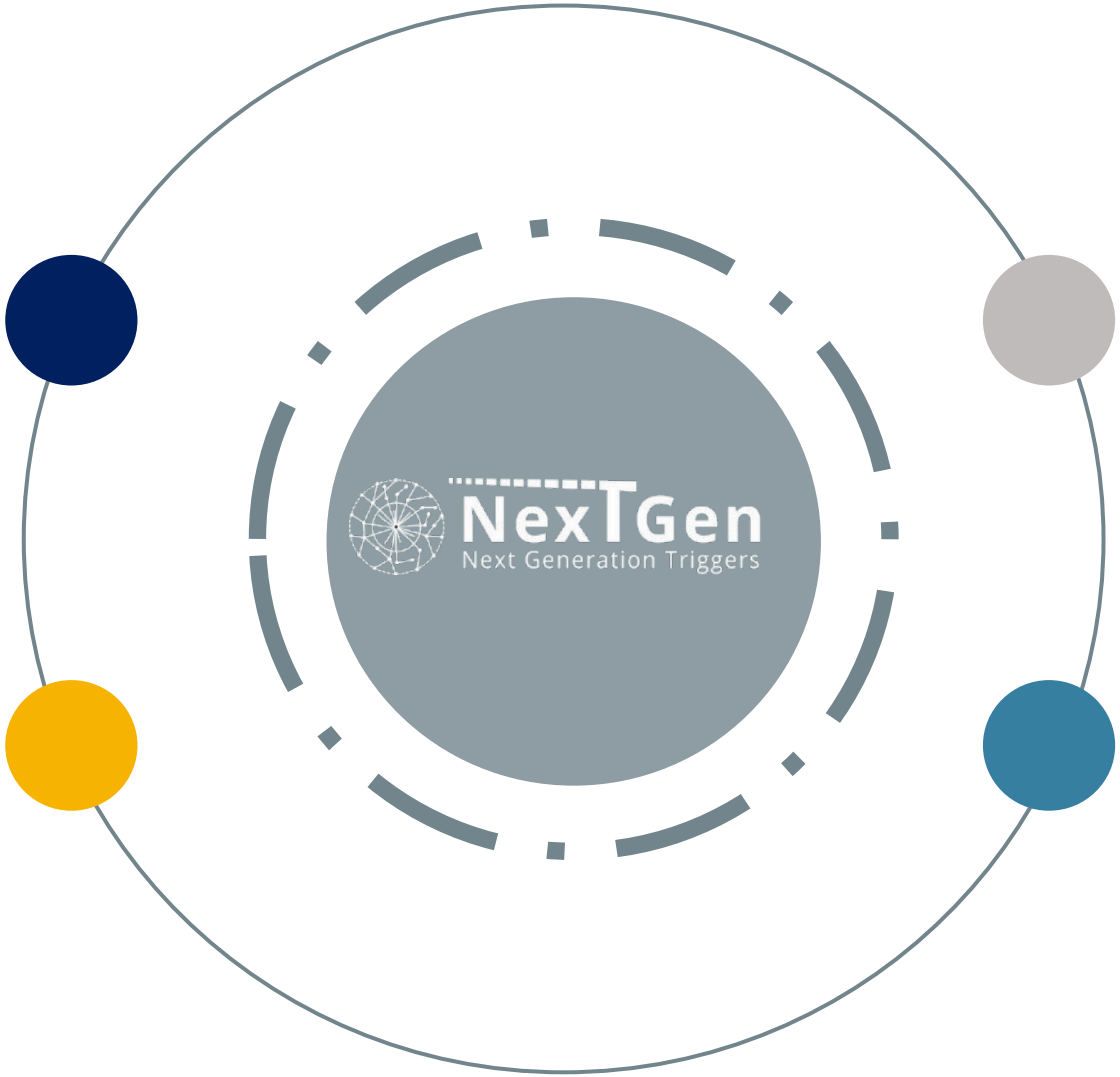
**Already in Full Swing!**


# INTEGRATION WITH CURRENT ACTIVITIES

NEXTGENTRIGGER IS IN SYNERGY WITH EXISTING ACTIVITIES:

**WP1**

collaboration with theoretical community and HEP Software Foundation




**WP2:**

fully integrated in the ATLAS collaboration, especially in TDAQ project

**WP4:**

target to foster the education in HEP, especially in trigger, to the whole community, not limited to NGT

**WP3:**

fully integrated in the CMS collaboration, especially in Trigger Studies Group and L1 trigger project

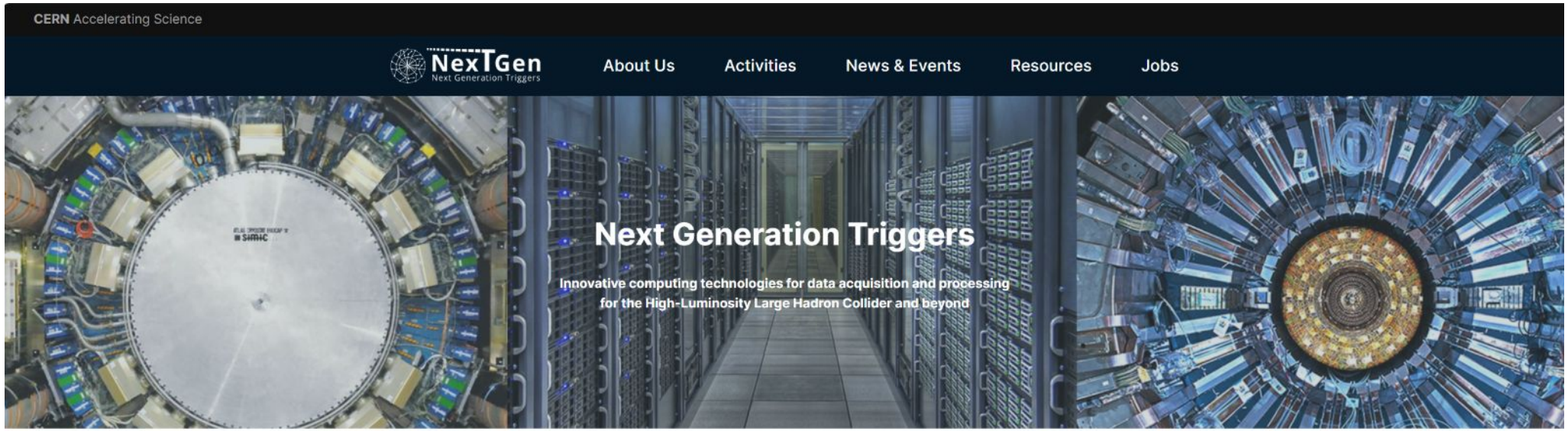
# RESOURCES

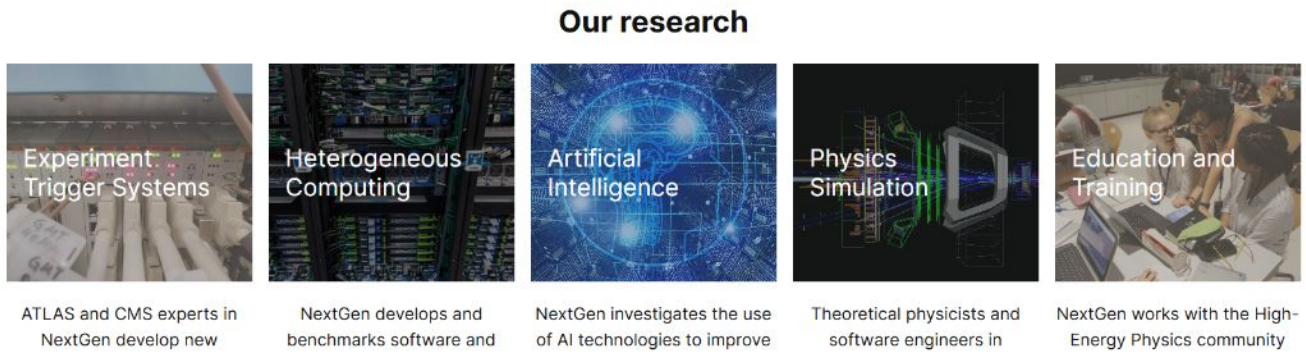
https://nextgentriggers.web.cern.ch

Project “Proposal” and other resources

1st Technical Workshop Nov 25-27
https://indico.cern.ch/event/1421629

# THE WEB SITE




https://nextgentriggers.web.cern.ch/

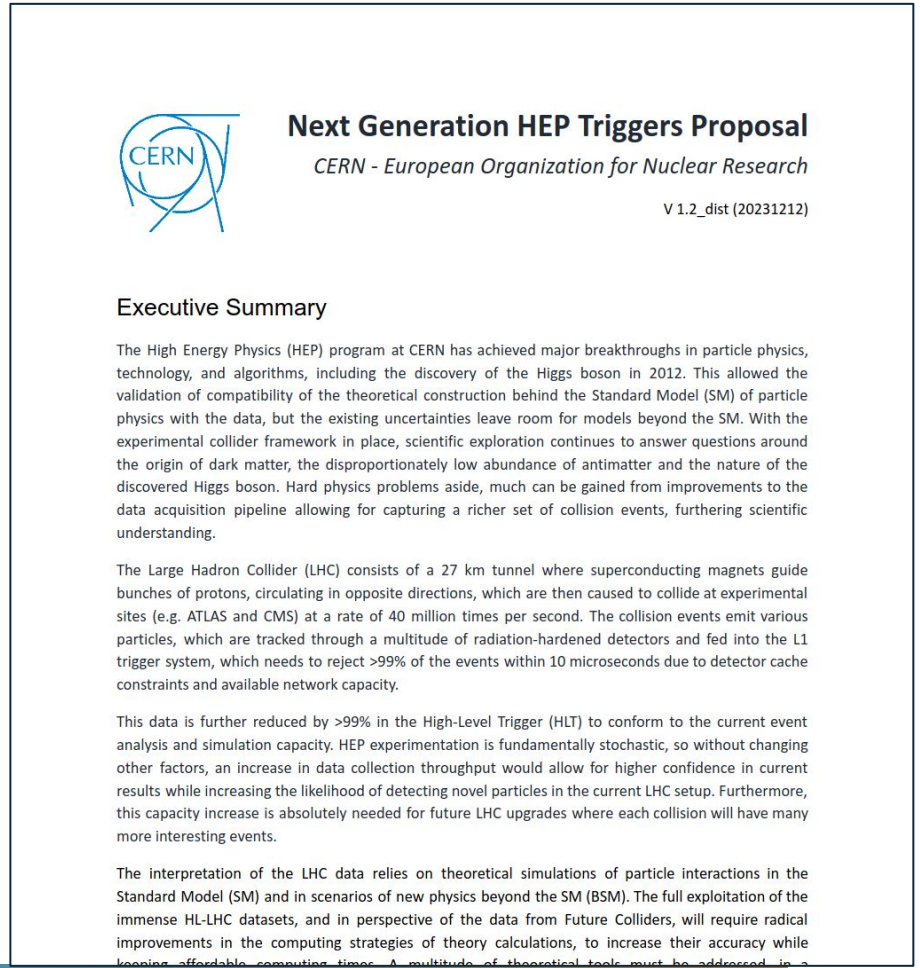
# THE PROJECT PROPOSAL

The original Project “Proposal”

approved 2023.

Detailed description of all working packages divided in **tasks** and their goal

Definition of the **milestones** in our path towards the next generation trigger!



CERN

**Next Generation HEP Triggers Proposal**

*CERN - European Organization for Nuclear Research*

V 1.2_dist (20231212)

Executive Summary

The High Energy Physics (HEP) program at CERN has achieved major breakthroughs in particle physics, technology, and algorithms, including the discovery of the Higgs boson in 2012. This allowed the validation of compatibility of the theoretical construction behind the Standard Model (SM) of particle physics with the data, but the existing uncertainties leave room for models beyond the SM. With the experimental collider framework in place, scientific exploration continues to answer questions around the origin of dark matter, the disproportionately low abundance of antimatter and the nature of the discovered Higgs boson. Hard physics problems aside, much can be gained from improvements to the data acquisition pipeline allowing for capturing a richer set of collision events, furthering scientific understanding.

The Large Hadron Collider (LHC) consists of a 27 km tunnel where superconducting magnets guide bunches of protons, circulating in opposite directions, which are then caused to collide at experimental sites (e.g. ATLAS and CMS) at a rate of 40 million times per second. The collision events emit various particles, which are tracked through a multitude of radiation-hardened detectors and fed into the L1 trigger system, which needs to reject >99% of the events within 10 microseconds due to detector cache constraints and available network capacity.

This data is further reduced by >99% in the High-Level Trigger (HLT) to conform to the current event analysis and simulation capacity. HEP experimentation is fundamentally stochastic, so without changing other factors, an increase in data collection throughput would allow for higher confidence in current results while increasing the likelihood of detecting novel particles in the current LHC setup. Furthermore, this capacity increase is absolutely needed for future LHC upgrades where each collision will have many more interesting events.

The interpretation of the LHC data relies on theoretical simulations of particle interactions in the Standard Model (SM) and in scenarios of new physics beyond the SM (BSM). The full exploitation of the immense HL-LHC datasets, and in perspective of the data from Future Colliders, will require radical improvements in the computing strategies of theory calculations, to increase their accuracy while

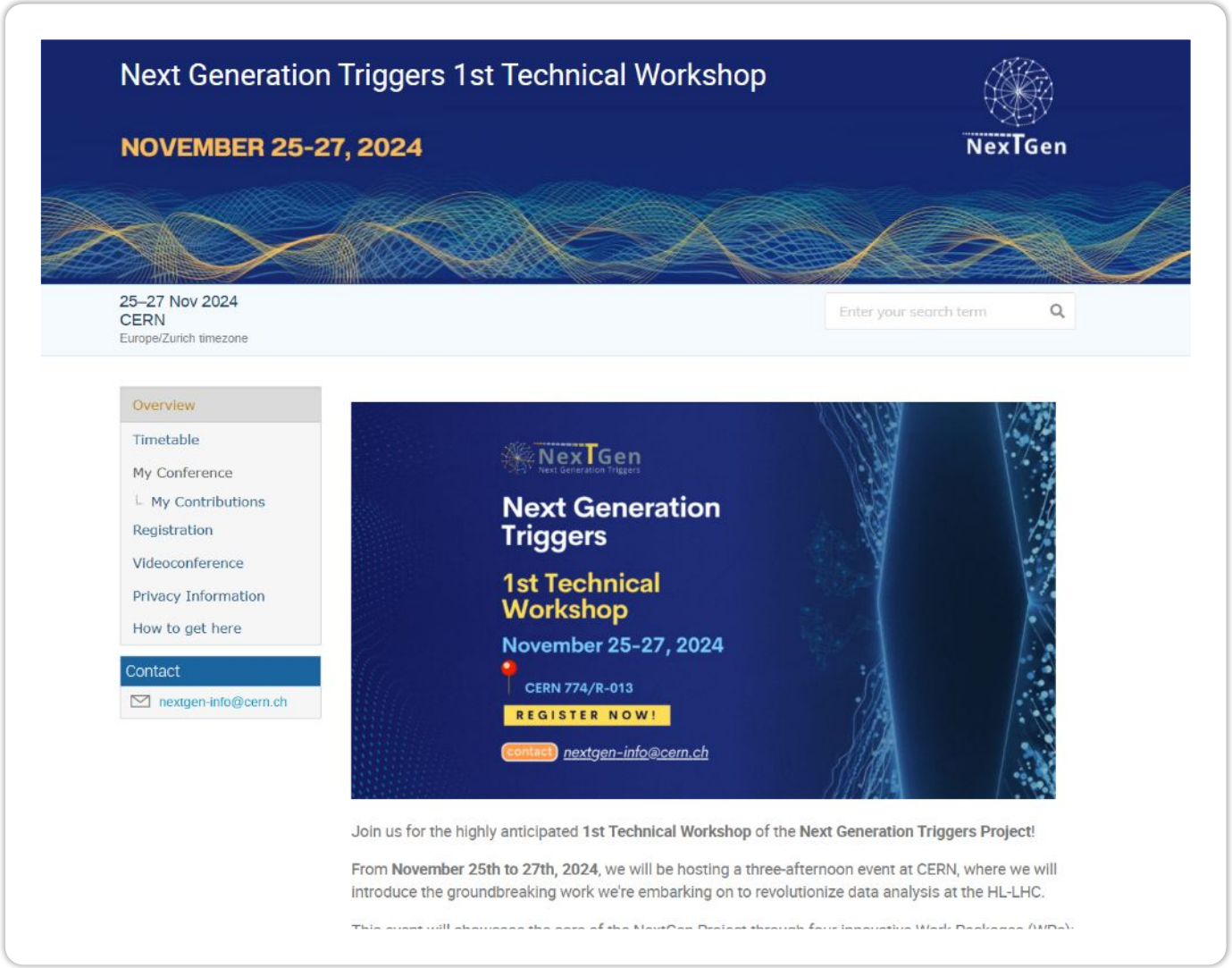
# THE YEARLY NEXTGEN WORKSHOP

The yearly NextGen Workshop is the main technical event.

Open to any interested person in the community to be informed on activities and progress, exchange ideas, ask questions, make suggestions

The first workshop took place on Nov 25-27 with report from all 24 tasks, presentation available for anyone.

All material is available to anyone!



# THE YEARLY NEXTGEN WORKSHOP

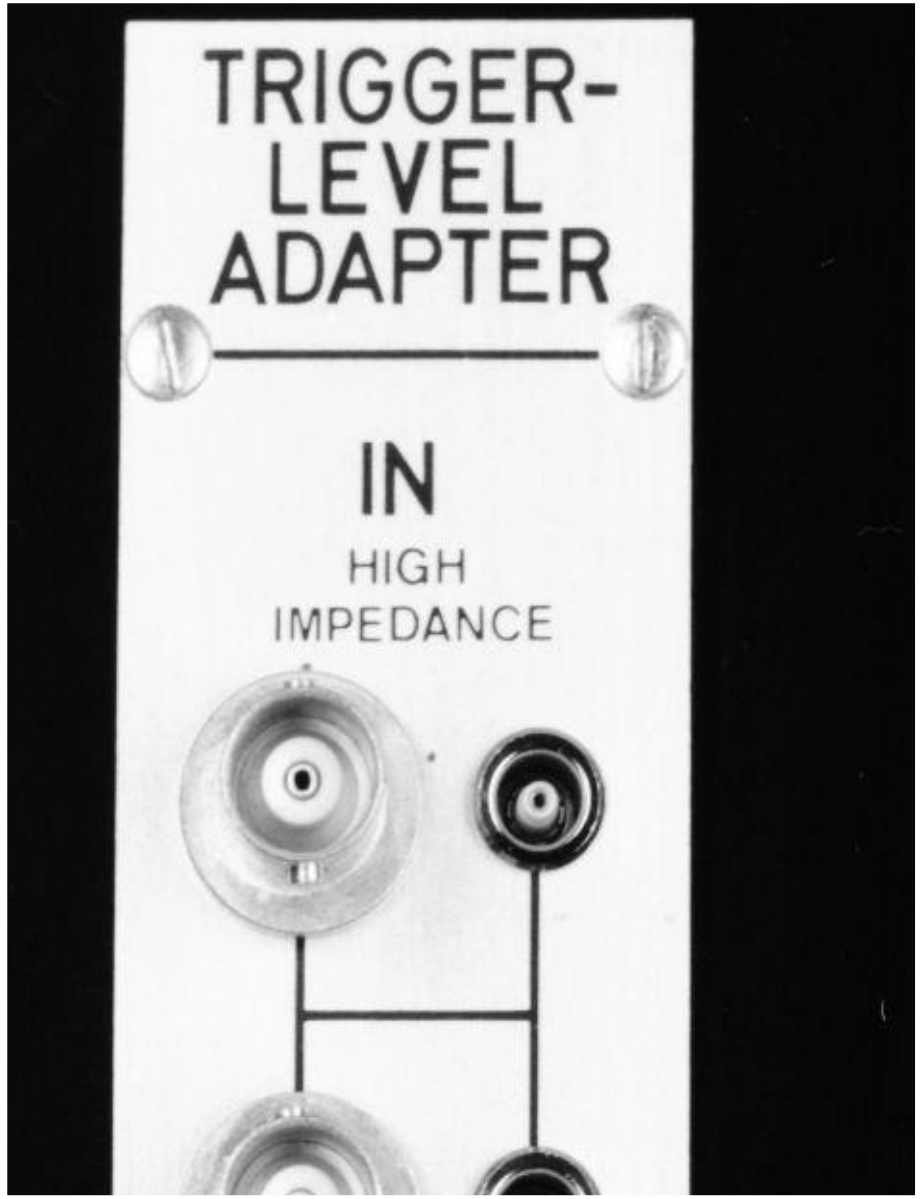
# Conclusion

## Introducing the NextGen Triggers Project

Fusing computing, algorithms, theory, ML: future solutions for the experiments

Embedded in the experiments, significant influx of community expertise: building with the community, for the community

Now advancing in full steam: technical workshop on Nov 25-27 for first results!




CERN
NexTGen
Next Generation Triggers

# A LITTLE
## BACKGROUND STORY

**2022**

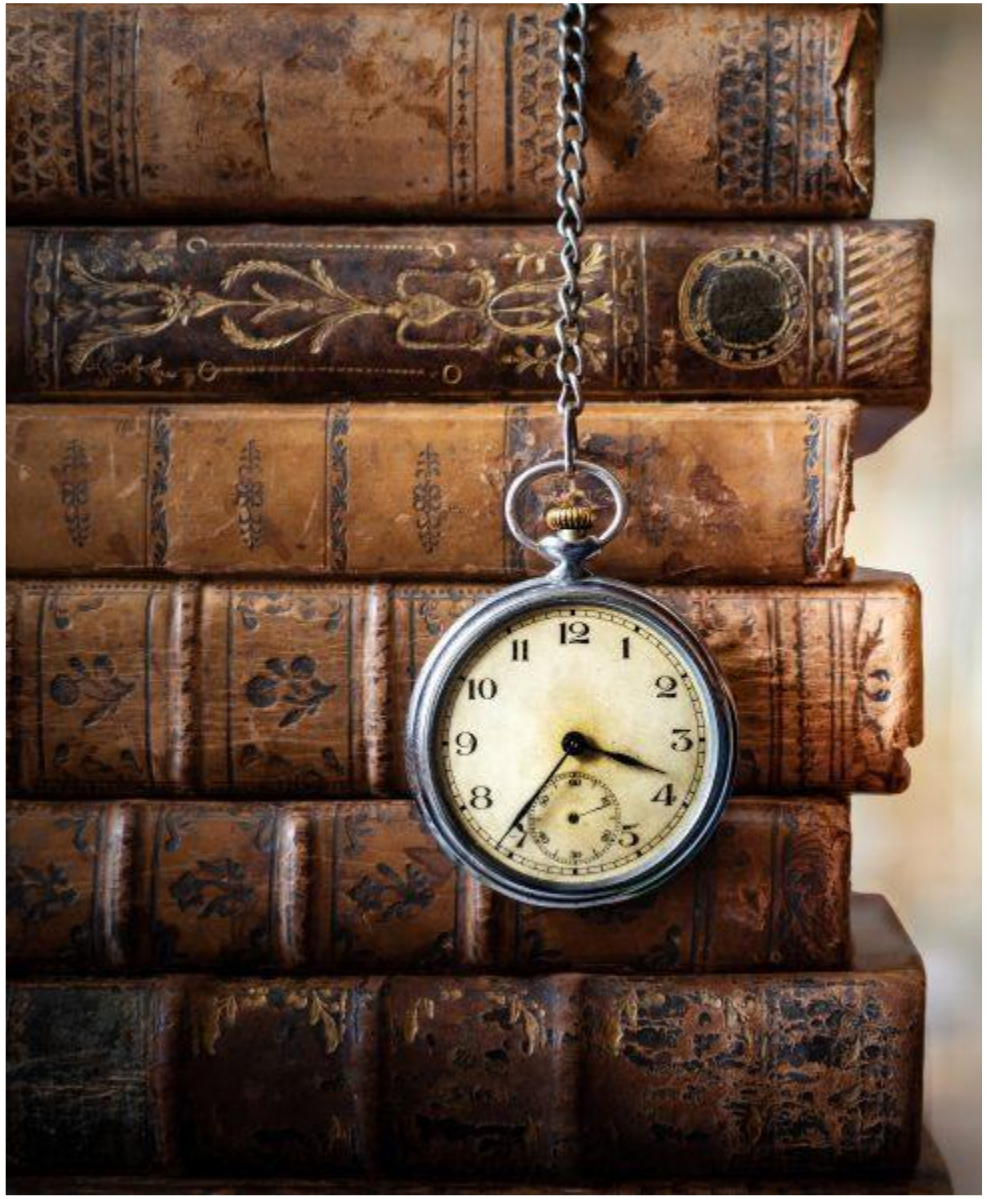
A group of private donors, visits CERN to know more about its missions and programmes.

**2023**

This first visit eventually evolves into an agreement with the Eric and Wendy Schmidt Fund for Strategic Innovation, approved by the CERN Council in October 2023, to fund a project to support advanced research for the future trigger systems at the HL-LHC and beyond.

**2024**

January kick-off. **NextGen Triggers was born!**



# THE SUPPORTING STRATEGY

## Of The Five-Year NextGen Triggers Projects

NextGen will collaborate with experts in academia and industry.

The work builds on CERN's Open Science and knowledge-sharing principles.

A unique multidisciplinary education program for NextGen researchers is included.

Targeted events and conferences for the wider scientific community will be organized.

Intellectual property from the NextGen Triggers project, owned by CERN, will be released and shared under open licenses, in compliance with the CERN Open Science Policy.