TRIGGER-LESS READOUT AND UNBIASED DATA QUALITY MONITORING OF THE CMS DRIFT TUBES MUON DETECTOR

Matteo Migliorini[1,2], Jacopo Pazzini[1,2], Andrea Triossi[1,2], Marco Zanetti[1,2], Alberto Zucchetta[2]
on behalf of the CMS Collaboration

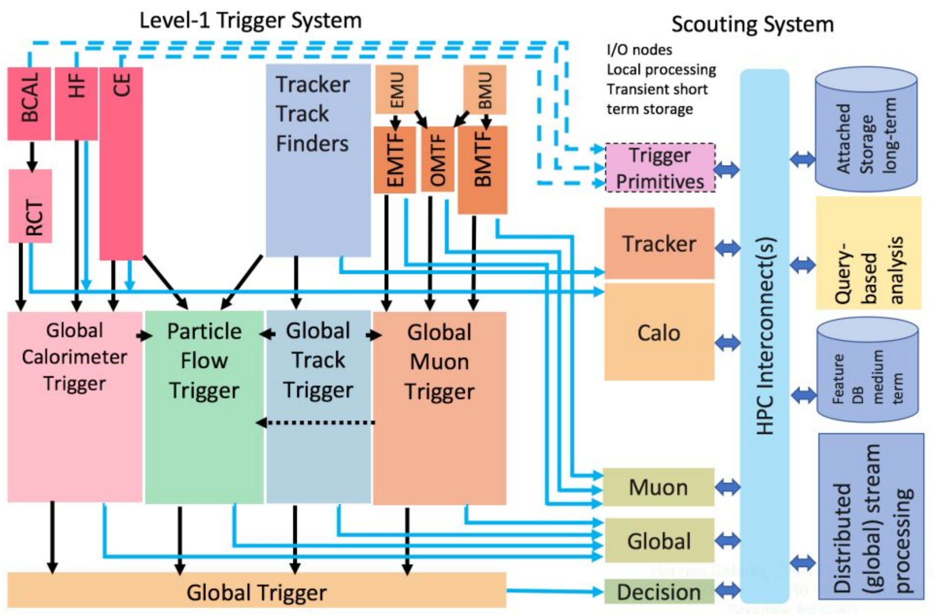
[1] Padova University — [2] INFN Padova Division

TRIGGER AND DETECTOR-DATA SCOUTING AT CMS

The High-Luminosity phase of the Large Hadron Collider will see an increase in the instantaneous luminosity delivered to the experiments to approximately $7.5 \cdot 10^{34}/\mathrm{cm}^2/\mathrm{s}$, resulting in an expected pile-up of up to 200 inelastic collisions per event [1].

To cope with the increased luminosity and pileup, the CMS experiment [2] will undergo a major upgrade (referred to as Phase-2 Upgrade), including:

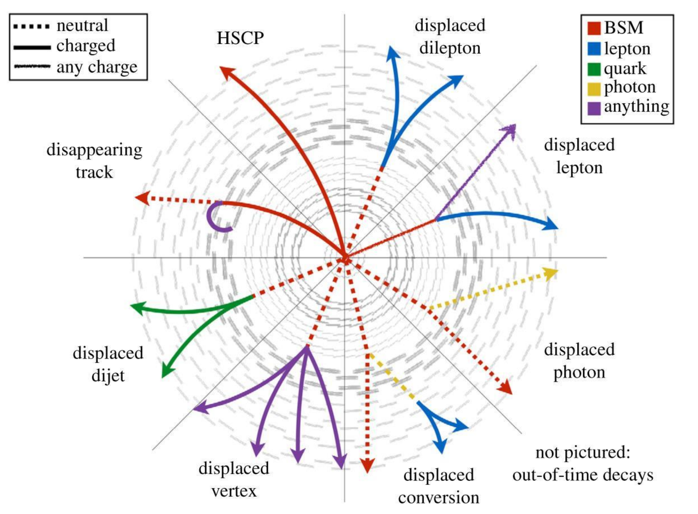
- the replacement of the tracking system, including a track-finder for Level-1 trigger
- the installation of a new timing detector and new high-granularity calorimeter in the endcap region
- the replacement of the Front-End (FE) detector electronics to improve radiation hardness and readout bandwidth
- a new Level-1 trigger based on state-of-the-art FPGAs capable of accepting events at a rate of 750 kHz



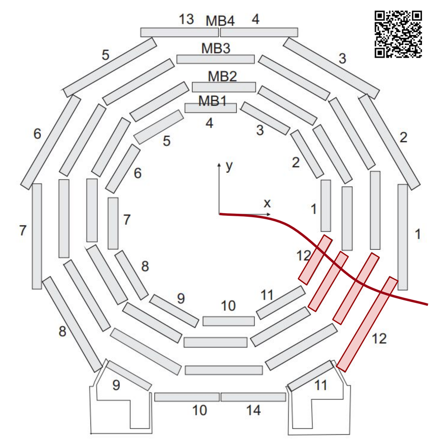
The upgraded electronics and Trigger systems will open the stage to the possibility of **collecting and analyzing Trigger data available at the collision rate of 40 MHz**. The 40MHz Level-1 Trigger scouting project [3] proposes to use spare optical outputs of Level-1 trigger boards, receiving trigger data as a complementary and parasitic system acquiring data in parallel as the standard CMS DAQ.

This can provide additional functionalities, for example enabling the online processing and data collection for physics channels lacking specific signatures compatible with the Level-1 Trigger system and not requiring full resolution, as in the case of long-lived particles or appearing or disappearing tracks.

The upgraded FE detector electronics, based on high-speed links, will enable to extend the 40MHz scouting approach directly to access the data stream from the detectors' FE, depending on the expected throughput, finally enabling the collection and reconstruction of full-resolution online data, as well as the monitoring FE based on the entirety of collision data.



CMS DRIFT TUBES HL-LHC UPGRADE DEMONSTRATOR

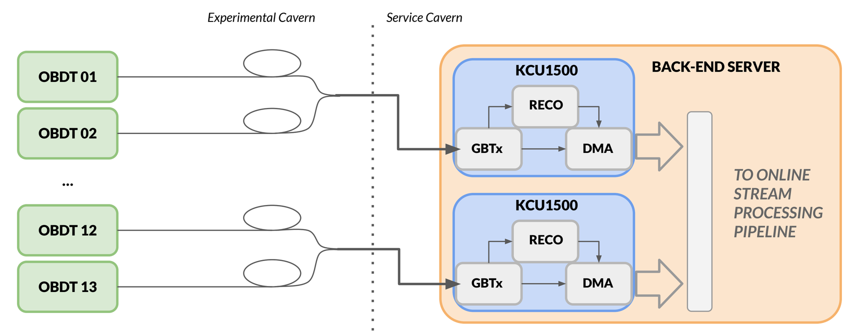


The CMS Drift Tube (DT) system [4] was chosen for the first implementation of the **40MHz raw-data scouting** approach in CMS, due to the limited expected data throughput from the DT detectors, and because a portion of the system is instrumented as a Phase-2 Upgrade demonstrator operating in parallel with the legacy readout during the LHC Run3.

One CMS DT sector is equipped with 13 prototype units of the Phase-2 upgrade On-Board DT readout boards (OBDT) [5], performing Time-to-Digital Conversions in FPGA with nanosecond resolution.

Each of the OBDT readout boards ships data to Back-End (BE) devices over high-speed optical links using the GBT protocol. The 40MHz DT scouting system collects the front-end signals parallel to both the legacy and Phase-2 Upgrade DT readout systems.

Xilinx KCU1500 development boards are used as BE devices to perform trigger-less readout of the OBDTs. Each BE board is capable of reading out 8 OBDT input links. The BE boards are installed in a dedicated server, where the data readout at the BE level is transferred to Memory based on a Direct Memory Access engine implemented in FPGA to avoid burdening the server CPU.



Two alternative approaches are proposed and tested for the transmission and processing of the trigger-less DT raw data stream.

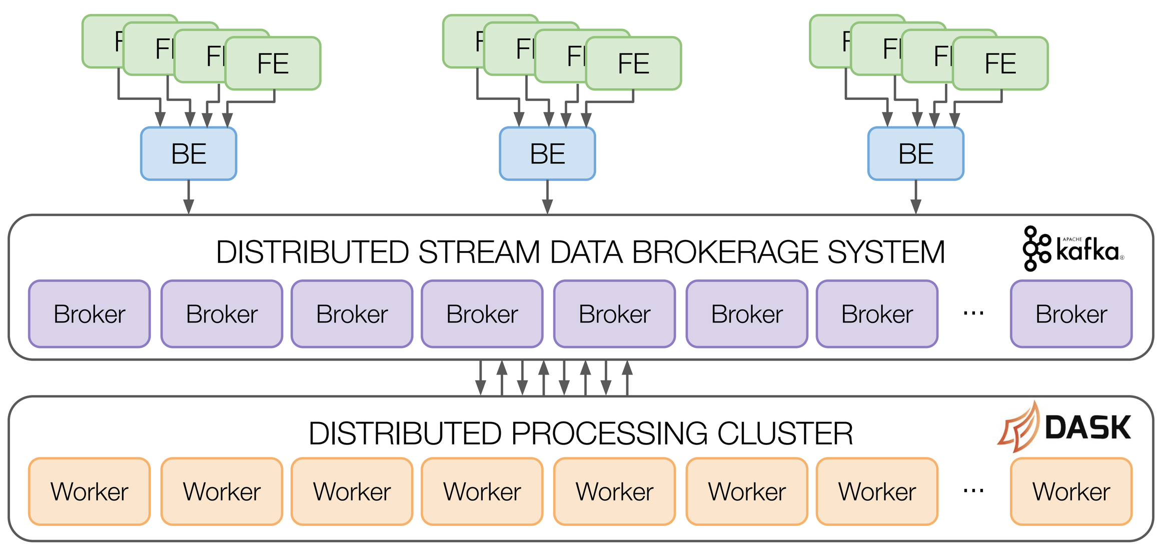
DISTRIBUTED PROCESSING AND MESSAGING SYSTEMS

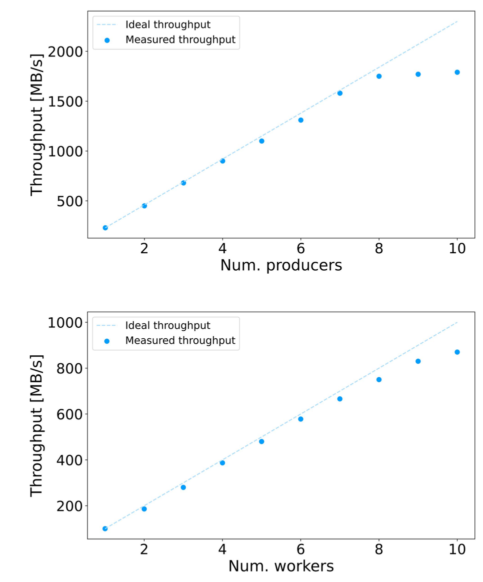
The unbounded stream of data is processed through a **distributed computing framework**, deployed either on bare-metal or on virtualized hardware resources deployed in a cloud infrastructure [6].

Each BE unit directly sends the data stream to a remote data buffer and messaging system, implemented with Apache Kafka [7], a distributed event streaming platform based on a pub-sub messaging model. Each BE unit acts as a data producer, publishing data on a specific topic, hosted in several partitions on the remote brokers.

Dask [8] is used as the distributed online processing engine to schedule the real-time data reconstruction workloads. A cluster of Dask workers, instantiated via the Kubernetes orchestrator, continuously reads from the Kafka partitions to collect raw detector hits and to create distributed data structures.

The intermediate Kafka data brokerage stage fully decouples the readout and processing stages.

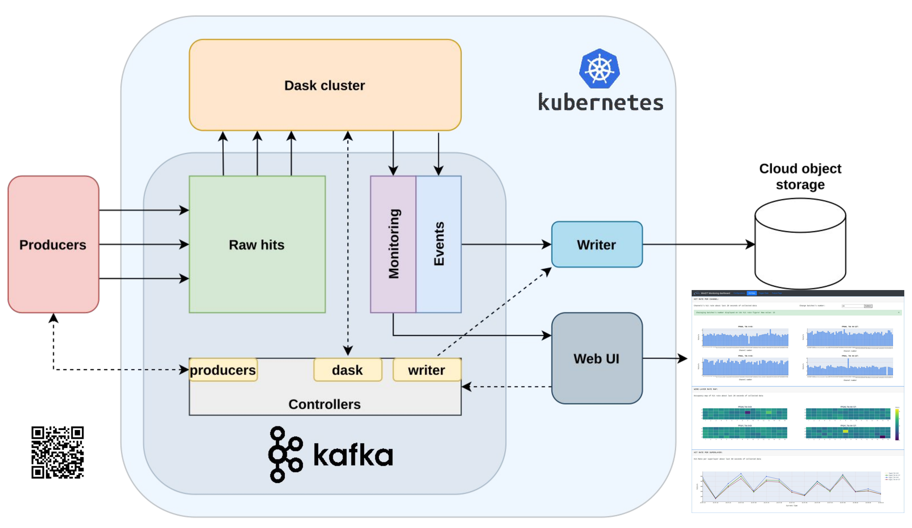




Simple aggregated quantities, useful for detector monitoring, are computed directly by exploiting the data-analytics functionalities of the distributed computing framework.

An event-building stage is performed in parallel by subdividing the data batches by the **LHC Orbit ID**.

Both the distributed data brokerage and the processing systems can scale efficiently with the number of nodes, and do not rely on intermediate event-building stages to aggregate or order the raw data.



MULTISTAGE ARCHITECTURE

As an alternative to the distributed cluster processing implementation, a **multi-stage data aggregation architecture** is deployed and currently under testing to assess its performance and robustness against high data throughput. The goal of the architecture is to aggregate individual event fragments dispatched from the FE boards to several BE devices during the data taking, building *events-like data structures* in the early stages of the data processing.

The **LHC Orbit ID** is used as the main "event" identifier, before performing the online analysis required to identify the individual collisions from the trigger-less data stream and complete the even reconstruction.

A set of Readout Units (RUs) is deployed directly in the servers hosting the BE boards. The RUs act as in-memory key-value stores by temporarily caching the collected DT hits into arrays indexed by LHC Orbit ID.

An Orbit Manager process (OBM) performs the bookkeeping of the LHC Orbits IDs available on the RUs. This is achieved by the RUs by sending a new message to the OBM every time a new Index is created into the RUs cache.

Another set of processes is in charge of merging all hits pertaining to a given LHC orbit and stored in the cache of multiple RUs, thus creating a single hit collection consistent with a given LHC Orbit ID. These multi-threaded builder processes, named Builder Units (BUs), interface directly with both the Orbit FIFO managed by the OBM, and with the RUs to fetch the cache records.

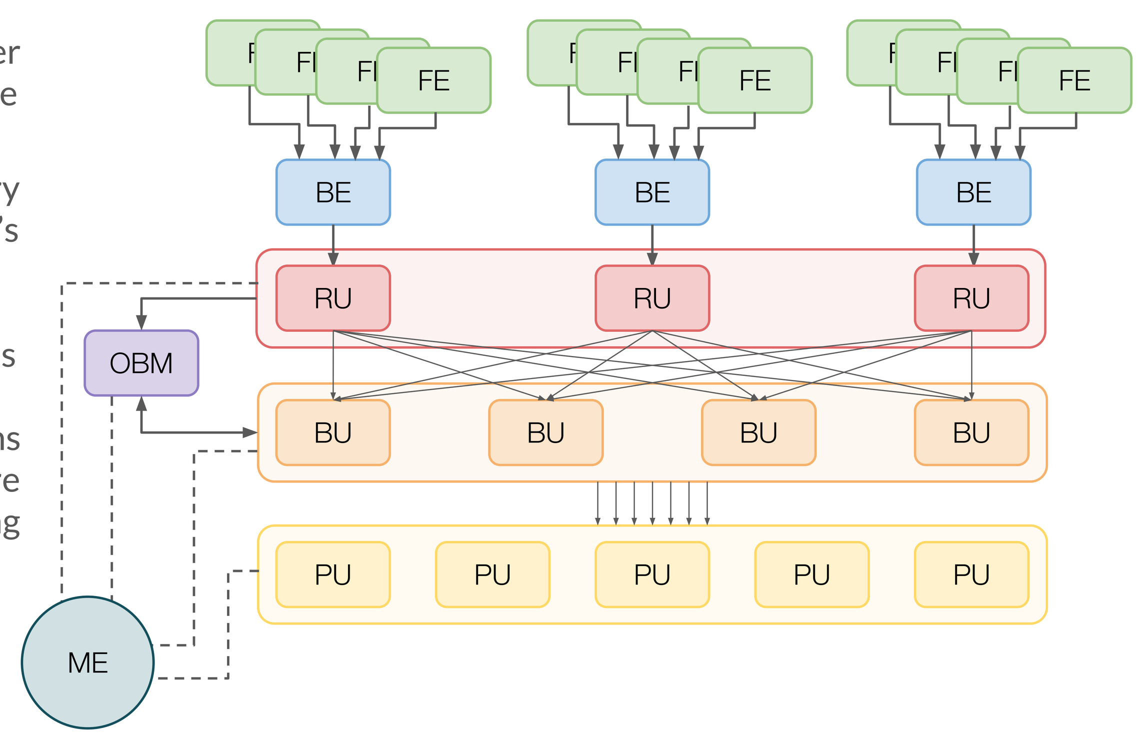
The BUs push the sets of hits aggregated by LHC Orbit ID to a number of Processing Units (PUs), each PU running tasks to cluster the hits and reconstruct the features corresponding to the passage of the muons through the detector.

A Master Entity (ME) controls all the other processes (RU, BU, PU, OBM), which are subscribed to it when first initialized.

The ME exposes an API acting as the entry point to interact with all the system's components, including:

- start/stop of a data taking
- collection of metrics from all processes

In this architecture, all communications both inter- and intra-processes are handled via the ZeroMQ [9] messaging library.



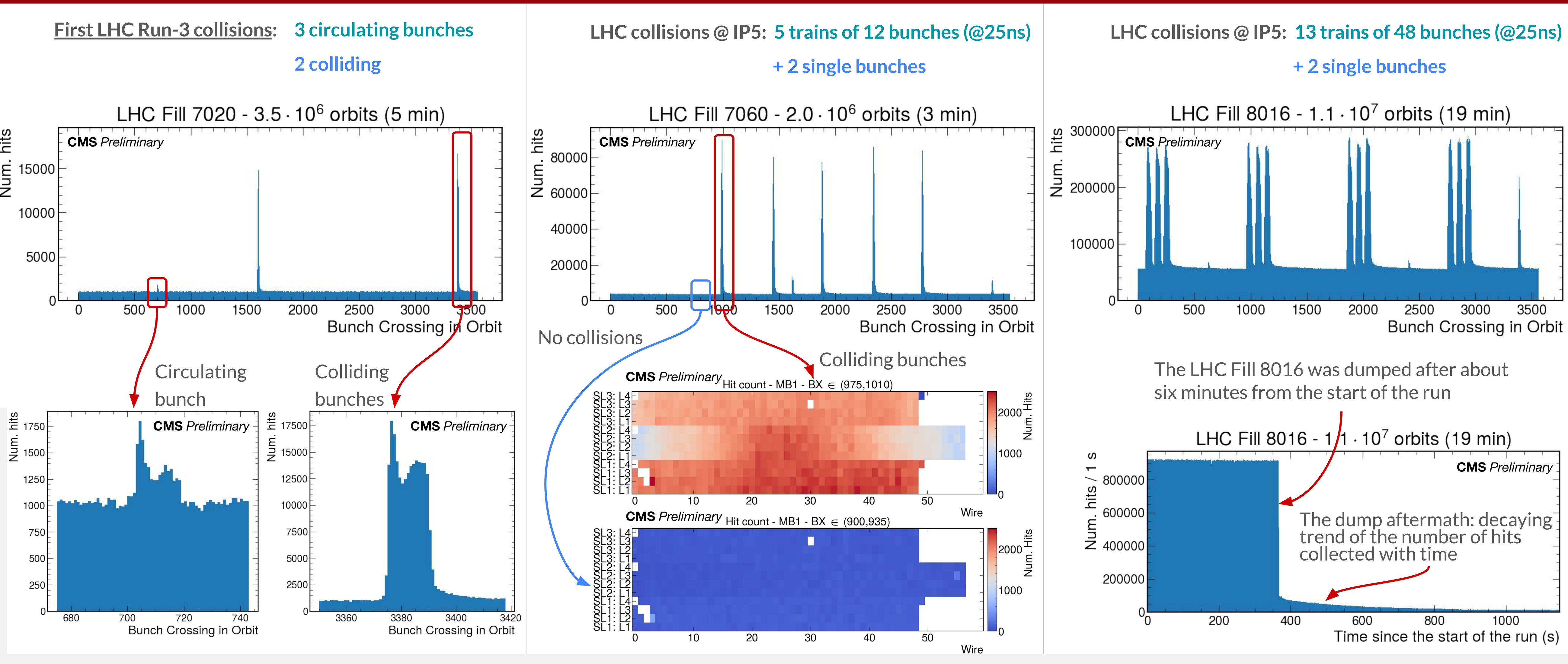
SELECTED PRELIMINARY RESULTS FROM THE FIRST RUNS COLLECTED WITH THE 40MHz DATA READOUT AT CMS DURING LHC RUN3

The trigger-less 40MHz data readout system is collecting the data produced in the CMS DT Phase-2 Upgrade demonstrator. The readout is completely opportunistic and data can be collected independently from the central DAQ of the experiment.

A full online data-reconstruction pipeline is under development, and physics-oriented data analyses will be implemented in the future.

The data collected with this trigger-less DAQ model open the stage for complementary Data Quality Monitoring studies, based on 100% of the hits produced by the detectors' FE, and completely unbiased thanks to the absence of dedicated Trigger selection logic.

- Detector occupancy both in-bunch and out-of-bunch
- Unbiased background noise rate, and its evolution with time
- High-statistics investigation of transient detector effects



[1] O. Aberle et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN Yellow Reports; doi:10.23731/CYRM-2020-0010
[2] S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST 3 (2008) S08004; doi:10.1088/1748-0221/3/08/S08004
[3] G. Badaro et al., *40 MHz Level-1 Trigger Scouting for CMS*, EPJ Web Conf. 245 (2020) 01032; doi:10.1051/epjconf/202024501032
[4] S. Chatrchyan et al., *Performance of the CMS Drift Tube Chambers with Cosmic Rays*, JINST 5 (2010) T03015; doi:10.1088/1748-0221/5/03/T03015
[5] A. Triossi et al., *Electronics Developments for Phase-2 Upgrade of CMS Drift Tubes*, PoS TWEPP2018 (2019) 035; doi:10.22323/1.343.0035
[6] M. Migliorini et al., *A horizontally scalable online processing system for trigger-less data acquisition*, Nucl.Instrum.Meth.A 1036 (2022) 166869; doi:10.1016/j.nima.2022.166869
[7] Apache Kafka; available at https://kafka.apache.org
[8] Dask: Library for dynamic task scheduling; available at https://dask.org
[9] ZeroMQ, an open-source universal messaging library; available at https://zeromq.org/