8   Luminosity and beam conditions

The main deliverables of the beam radiation instrumentation and luminosity (BRIL) system are the measurements of the online and offline luminosity, the beam-induced background, the beam losses and timing, and the radiation products in the CMS experimental cavern. The latter are used to compare with the values from corresponding simulations. The BRIL system delivers luminosity and beam-condition data in real time to the LHC and CMS control rooms. In this section, the instrumentation and technologies used to carry out these measurements are described. The software, computing, and monitoring aspects of the BRIL systems are also discussed. The precision measurement of the integrated luminosity for the different data-taking periods of Run 2 has been described in refs. [181–187]. For the 2016 data, an uncertainty of 1.2% was achieved [181]. For 2017, 2018, and Run 3, improved instrumentation and refined analysis techniques are expected to lead to further reduced uncertainties. Key ingredients of the precision are the availability of the various independently calibrated luminometers with different instrumental systematics and the in situ monitoring of their performance in terms of efficiency and linearity using short vdM-like transverse beam separation scans in physics fills.

For the online bunch-by-bunch luminosity measurement in Run 3, we rely on two fully dedicated luminosity detectors (luminometers): the pixel luminosity telescope (PLT) and the fast beam conditions monitor (BCM1F). In addition, the forward hadron calorimeter (HF) is used for the online luminosity measurement. The HF has a separate trigger-level readout dedicated to the luminosity measurement.

Both PLT and BCM1F were already operated successfully throughout Run 2. For operation during Run 3, improved versions of these detectors were constructed and installed during LS2. Other subsystems relevant to the luminosity measurement are the pixel detector and the muon drift tube system. These various luminometers provide redundancy and complementarity for the precise measurement of the online and offline luminosity and the online monitoring of the backgrounds.

The three independent online luminometers, PLT, BCM1F, and HF, are based on fundamentally different technologies, and each provide precise bunch-by-bunch measurements of the instantaneous luminosity and the beam-induced backgrounds (BIB). The BRIL online measurements are available at all times, in particular whenever there is beam in the LHC, independently of the status of other CMS subsystems and the central CMS DAQ system. The BRIL data-acquisition system (BRILDAQ) delivers data in real time every 1.458 s, corresponding to four so-called "lumi nibbles" (NB). The duration of one NB is 0.3645 s, corresponding to 2^{12} orbits, and 64 NB make up one luminosity section (LS) of about 23.3 s.

Data from the pixel detector and the muon systems contain important additional information that is used to ensure a precise calibration and long-term monitoring of the stability and linearity of the luminosity detectors. Tracker data are collected via the central CMS DAQ for offline analysis, and provide pixel cluster and vertex information per event during calibration runs, and bunch-by-bunch cluster counts per LS for data-taking periods with collisions. Muon data in Run 3 are collected via the BRILDAQ from three sources: (i) orbit-integrated muon track stubs from the muon barrel track finder; (ii) trigger primitives from the DT chambers that are equipped with prototype Phase 2 readout; and (iii) muon counts via the 40 MHz scouting system. The first source is the same as that used during Run 2, while the last two are new and were introduced for Run 3 in preparation for the Phase 2 upgrade.

To monitor the beam conditions, in addition to the dedicated luminometers, we also use the beam-halo monitor (BHM) and beam-condition monitor for beam losses (BCML).

A simulation and monitoring strategy is in place for the radiation background to evaluate and, if necessary, mitigate its effects.

8.1 Real-time bunch-by-bunch luminometers

8.1.1 Pixel luminosity telescope (PLT)

The PLT detector [188–192] is an independent system for the luminosity measurement. It consists of 48 silicon pixel sensors arranged on 16 "telescopes", eight on each side at 1.75 m from the interaction point (IP), close to the beam pipe, at a pseudorapidity $|\eta| \approx 4.2$. Each telescope is 7.5 cm long and contains three planes with individual silicon sensors. Their inner edges are as close as 4.7 cm from the beam line. Figure 110 shows a sketch of the layout of one end of the PLT (left) and a photograph of the actual detector showing the two half-frames that each side consists of (right).
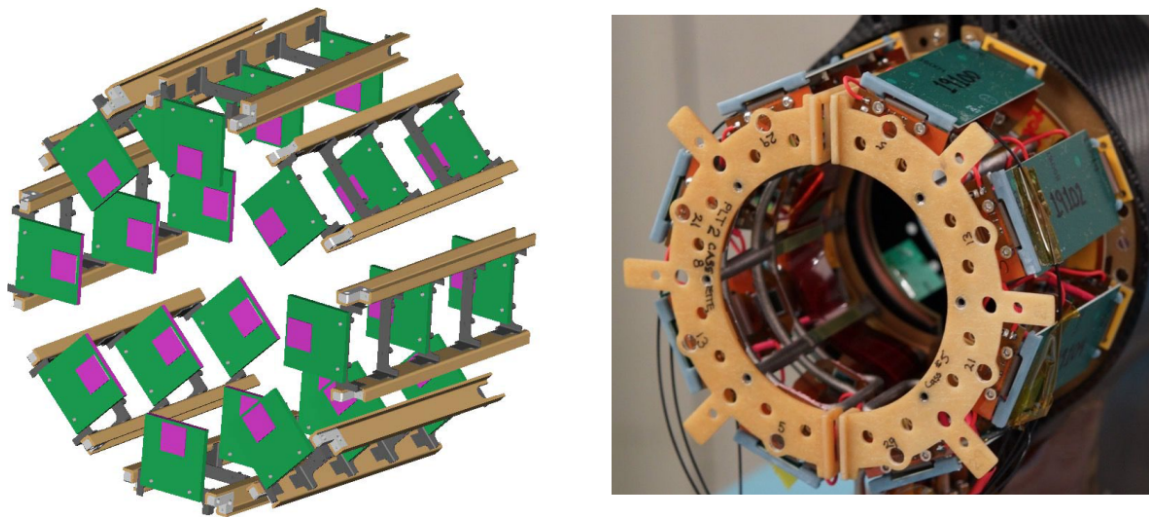
Figure 110. Left: sketch of the general PLT geometry. The sensors are indicated by the purple squares. Right: the actual PLT detector at one end of CMS, showing the arrangement of the eight telescopes around the beam pipe. Reproduced from [193]. CC BY 4.0.

The silicon sensors [194, 195] are n-in-n (45 sensors) and n-in-p (3 sensors) type and consist of 80 rows and 52 columns of pixels (or 26 double columns of 160 pixels each). The pixels are $150\,\mu$m wide and $100\,\mu$m high, and cover a total active area of $8 \times 8\,\text{mm}^2$ with an active thickness of 285 and $150\,\mu$m, respectively. During data taking, it is possible to set a smaller active area to decrease the contribution from accidental hits, which are triple coincidences of signals that do not actually come from a single particle track originating from the IP. The sensors are bump-bonded to a PSI46v2 readout chip (ROC) [196, 197]. Both the n-in-n silicon sensors and the PSI46v2 ROCs are inherited from the CMS Phase 0 pixel detector, while the telescope with n-in-p sensors of $150\,\mu$m thickness uses prototype CMS Phase 2 silicon sensors.

The PLT detector is usually operated in a "fast-or" mode in which the data are read out at the full bunch-crossing frequency of 40 MHz. The fast-or mode does not contain full information on hit

positions. In fast-or mode, if any pixels in a sensor register a hit over threshold, a single signal is produced. The ROCs also have a separate data path to read out the full pixel data (hit location and pulse height) at a lower rate (up to about 3.3 kHz) upon receipt of a level-1 accept (L1A) trigger signal. The full pixel information, read out at a lower rate, is used to measure calibration constants, corrections, and systematic uncertainties for the online and offline measurements.

The three ROCs for a single telescope are connected to a high-density interconnect (HDI) card, which contains a token bit manager (TBM) chip [198]. The TBM chip coordinates the readout of the three individual ROCs and produces a single readout for each telescope. Four telescopes are connected to a port card, which manages the communication and control signals for a single half-frame of the detector. The port card is in turn connected to the opto-motherboard (OMB), which converts the electrical signals into optical signals that are then sent by fibers to the CMS service cavern, where the backend readout electronics are located. Of the different kinds of application-specific integrated circuits (ASICs) used in the PLT, all but two (SlowHub and PLTdriver) were developed for the Phase 0 CMS pixel detector [1]. The SlowHub chip, on the OMB, uses blocks from the pixel TBM chip to extract the slow I^2C signals from the fast 40 MHz commands needed for the pixel readout, thus significantly reducing the total number of fibers needed. The PLTdriver chip manages the fast-or signal outputs from the individual ROCs.

The backend readout electronics comprise a single frontend controller (FEC) card, which issues commands to the ROCs, TBMs, and OMB, and three frontend driver (FED) cards. One FED is used to read out and decode the full pixel data from the ROCs, and is identical to the FEDs used by the Phase 0 pixel detector [199]. This data are then read out over an Slink [200] connection and saved to a dedicated PC. The other two are the fast-or FEDs, one for each side of the PLT. These read out the fast-or data from the ROCs and histogram the number of triple coincidences per channel and per bunch crossing (BX). These data are read out via a CAEN VME optical bridge to a separate dedicated PC. A dedicated trigger is used, in which events from colliding BXs are read out without a requirement for specific activity in the event, this way equally sampling all colliding BXs in the LHC orbit.

The readout hardware counts the number of triple coincidences, events where all three planes in a telescope register a signal, to determine the instantaneous luminosity. This fast-or readout allows the PLT to provide online per-bunch luminosity with excellent statistical precision, with the triple coincidence requirement providing a strong suppression of background from noise, residual radioactivity of the detector material (afterglow), and BIB sources from beam-gas interactions and beam halo. The luminosity measurement is obtained using a zero-counting technique, described, e.g., in ref. [181], in which the actual rate is inferred from the measured rate of no hits, thereby correcting for the overlap of signals.

8.1.2 Fast beam conditions monitor (BCM1F)

The BCM1F detector is a fast particle counter installed around the beam pipe, ±1.8 m away from the IP, at a radius of about 6 cm, corresponding to $|\eta| \approx 4.1$. The BCM1F provides a real-time measurement of both luminosity and beam-induced backgrounds.

The experience gained from BCM1F operation in Run 2 [201, 202] led to an update of the design for Run 3 [203]. For Run 3, the mix of silicon and poly-crystalline (pCVD) diamond sensors used in Run 2 was changed to an all-silicon sensor configuration with active cooling that has an

improved stability with increasing integrated fluence and an improved linear response of the sensors as a function of the instantaneous luminosity.

As in Run 2, the Run 3 detector consists of four "C-shaped" printed circuit boards (PCB), referred to as C-shapes (figure 111). A total of four C-shapes form two rings, one on each side of the IP. Each C-shape has six double-diode silicon sensors that are used to measure particle hits. The entire system comprises 48 identical channels. The C-shapes for Run 3 were designed for improved robustness and with additional space for active cooling contacts.
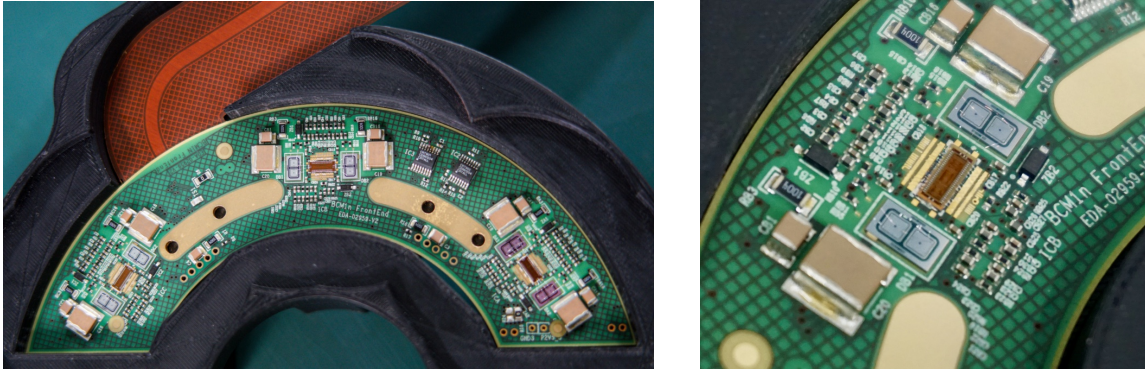
Figure 111. Photographs of the BCM1F detector used in Run 3. Reproduced with permission from [204]. Left: one full BCM1F C-shape printed circuit board. Right: a closeup of the frontend module, which includes the processing chip with two silicon double-diode sensors on each side.

Performance studies led to the choice of a sensor per-channel with an area of $1.7 \times 1.7\,\text{mm}^2$, balancing the occupancy at high interaction rates with the statistical precision at low luminosity. The sensors were designed as alternating current (AC)-coupled double-diodes and produced on the CMS Phase 2 outer tracker strip wafers [32] with a full depletion thickness of $290\,\mu\text{m}$.

Each sensor is directly connected via bond wires to one of the three fast JK16 [205] CMS ASIC chips. This preamplifier chip was designed to have a fast rise time of 7 ns, narrow pulses with 10 ns full width at half maximum, and a return to the baseline within 25 ns. The connections are established via AC coupling to prevent baseline drifts due to radiation-damage-induced leakage currents that would not be compensated by the preamplifier. The coupling capacitance is part of the sensor design, while a resistor on the PCB is used as the bias resistor.

Copper pads were added on the C-shapes for Run 3 to make the cooling contact with a new titanium 3D-printed cooling pipe, which is placed in series with the PLT cooling loop. This improves the thermal contact of the sensors to the cooling system and prevents thermal runaway induced by radiation damage, thereby ensuring the expected lifetime of the BCM1F silicon sensors exceeds the duration of Run 3.

Using analog opto-hybrids [206], the AC signals from the C-shapes are propagated to the backend electronics that includes CAEN V895 discriminators, real-time histogramming units (RHU) [207], and a CAEN V1721 flash ADC, all of which use a VME bus. Parallel μTCA electronics consist of an 8-bit ADC mezzanine card with a 1.25 GB/s sampling rate on a gigabit link interface board (GLIB) carrier board [208, 209], using an FPGA-based derivative peak-finder algorithm [210]. At the RHU, which was already used for Run 2, the data are filled into histograms with 6.25 ns granularity. This

time interval corresponds to the time of flight for relativistic particles arriving from the IP that is a quarter of the time between two collisions.

The BCM1F detector, situated in this location and delivering a fast response with zero dead time, allows for the bunch-by-bunch separation of collision signals from backgrounds. The luminosity is measured bunch-by-bunch using the zero-counting method mentioned above, followed by a correction for out-of-time hits and the application of a calibration constant. The relative contribution from background can be determined using the first bunch in the train or a noncolliding bunch in the beginning of the orbit, where the background from out-of-time hits due to activation from previous collisions is small. The background measurement is of crucial importance for the operation of the CMS detector since it is used by the silicon strip and pixel detectors as an automatic switch-on semaphore.

8.1.3 The forward hadron calorimeter (HF)

A comprehensive description of the HF is provided in section 5. Here, we highlight the aspects that are relevant for the luminosity measurement.

Like the PLT and BCM1F, the HF reports luminosity measurements at all times, independently of whether the main CMS DAQ system is in operation. The outputs from the circuits used to digitize the signals from the HF photomultipliers (PMTs) are routed to FPGAs that are part of the HF readout [60]. The FPGAs tap into the primary readout path in a noninvasive way and collect channel-occupancy (OC) and transverse-energy-sum data in histograms that have one bin for each of the 3564 BX time windows of one beam orbit. The histograms are periodically transmitted to a central BRILDAQ processor node. With this configuration, the HF delivers a continuous and precise measurement of the bunch-by-bunch luminosity at all times.

While all HF channels are available for use in the BRIL backend electronics, Monte Carlo simulation studies indicate that the best linearity of the luminosity measurement is obtained when using only the η rings 31 and 32, corresponding to a pseudorapidity range $3.15 < |\eta| < 3.5$, where the long-term radiation damage is predicted to be the smallest.

The HF luminosity system allows the extraction of a real-time instantaneous luminosity relying on two algorithms. The first (HFOC) is based on zero counting, in which the average fraction of below threshold towers is used to infer the mean number of interactions per BX. The second (HFET) exploits the linear relationship between the average transverse energy per tower and the luminosity. Both the HFOC and HFET methods exhibit excellent linearity, based on the experience acquired during Run 2 [181–183, 183, 186] and in the beginning of Run 3 [211].

8.2 Additional luminometers

8.2.1 Tracker

The CMS silicon tracker, described in section 3, is characterized by a low occupancy in the silicon pixel detector. This feature is exploited by the pixel cluster counting (PCC) method to measure the luminosity offline with excellent precision. The most recent PCC luminosity measurement during Run 2 [183] was based on data collected by zero-bias (colliding bunches only) and random triggers at recording rates of approximately 2 kHz and 400 Hz, respectively. The zero-bias data were used for the luminosity calculation, while random-triggered events served for the determination and subtraction of the afterglow backgrounds present in the zero-bias data. The method achieved a

statistical precision of about 0.2% per luminosity section during physics runs, with excellent stability after removing problematic sensor modules.

We applied a similar strategy for Run 3, with improvements in the data processing and recording to reduce and streamline the data. The cluster reconstruction and counting is implemented in the HLT software, which reduces the data size by three orders of magnitude. Potentially unstable modules are removed using an iterative method in which modules with large fluctuations relative to the average cluster count are removed. The implementation of a new processing path is in progress during Run 3, in preparation for the Phase 2 upgrade of the luminosity instrumentation for the HL-LHC [212]. In this path, the data from the HLT are transferred directly to the BRILDAQ.

An alternative method using information from the silicon tracker involves counting the number of primary-interaction vertices [181]. This method is used during van-der-Meer (vdM) calibration fills as a tool in the measurement of beam-dependent parameters. Vertices are also analyzed to derive the bunch density distributions of the beams, and to calibrate the distance by which the steering magnets displace the beams in the transverse direction.

8.2.2 Muon system

The BRIL system also makes use of level-1 trigger information from the muon barrel (MB) drift tube (DT) detectors, which deliver counts of the number of orbit-integrated muon stubs. While the rate of muons in the barrel is low, the stability and robustness of the muon system makes the DT a valuable source for luminosity monitoring. The muon detectors and trigger are described in sections 6 and 10.2, respectively. The muon information used by BRIL is aggregated by the barrel muon track-finder algorithm (BMTF) [213] of the level-1 muon trigger. The BRILDAQ system receives the data from the BMTF via a dedicated readout system.

In addition to the orbit-integrated data, during Run 3 a new histogramming firmware module was added to the backend of the DT slice test that uses Phase 2 readout technology to test how trigger primitives (muon segments per DT chamber) can be used for bunch-by-bunch luminosity measurements. This system serves as a demonstration of how a luminosity measurement can be made with per-bunch granularity for the CMS Phase 2 upgrade.

Figure 112 illustrates the readout diagram of the DT slice test. A detailed description is provided in ref. [88]. The MB1 and MB2 chambers are read out using the upgraded electronics. Data duplication was established for the MB3 and MB4 layers, thus allowing us to evaluate the performance of the upgraded electronics with respect to the older system. The BRIL histogramming module is placed on the so-called AB7 backend boards. Each AB7 board produces a single histogram. Each of the MB1, MB2, and MB3 chambers are read out by a single AB7, whereas two AB7 boards are needed to read out the MB4 chamber.

In addition, the 40 MHz scouting system in Run 3, described in detail in section 10.6, provides an alternative means to use the bunch-by-bunch muon rate from the level-1 muon trigger for the luminosity measurements. The concept of the 40 MHz scouting system is to capture part or all of the trigger data streams from different trigger layers or their subsystems using spare optical outputs. A demonstration of the scouting system was already tested in Run 2 based on the output from the global muon trigger algorithm μGMT described in section 10.2. Due to the inclusion of the layer-2 calorimeter trigger and the BMTF data, the Run 3 scouting demonstration offers additional objects with more sophisticated customizable selection possibilities. The histogramming for measuring
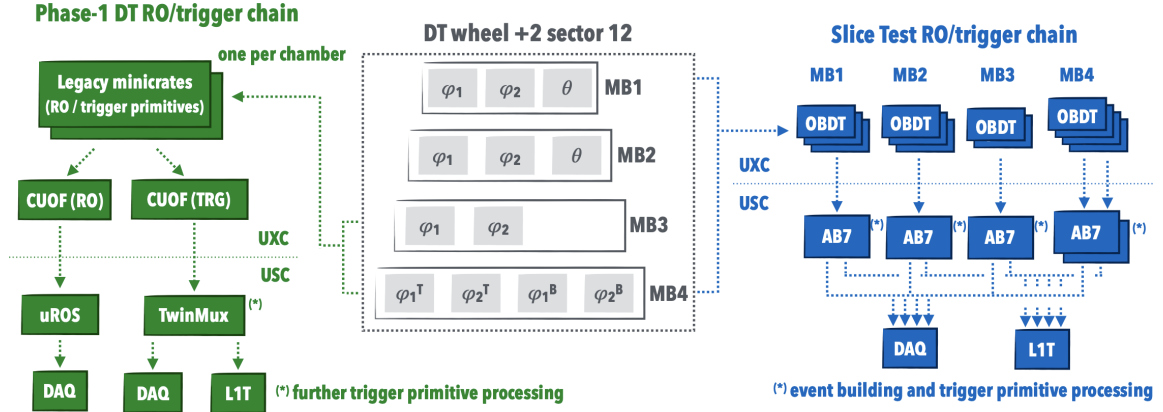
Figure 112. Diagram comparing the DT readout (RO) and trigger chains in the Phase 1 system used for data taking in Run 3 and in the DT slice test that acts as a demonstrator of the Phase 2 system in Run 3. Reproduced from [88]. CC BY 4.0. The central part of the figure indicates the number of Phase 2 on-board DT (OBDT) boards installed in each DT chamber. The left part of the figure shows how the Phase 1 readout and trigger-primitive generation are performed by the legacy on-board DT electronics (minicrates). Information, transmitted by optical fibers from the detector balconies to the counting room, is further processed independently for the readout (μROS) and trigger (TwinMux). The right part of the figure illustrates how the slice test TDC data are streamed by each OBDT to AB7 boards hosted in the counting room, which are used for event building and trigger-primitive generation.

the luminosity is performed per object type using the already mentioned generic histogramming firmware loaded on the scouting board.

The histogramming firmware is a generic module developed by BRIL that targets luminosity measurements using various, primarily Phase 2, CMS subsystems. The DT muon detector and the 40 MHz scouting system are the first to use this new module, and multiple instances of the module are placed on backend FPGAs. The module comes in two flavors: a synchronous type is adapted to the bunch-by-bunch luminosity measurements, when the counted objects arrive synchronously with the bunch clock, as is the case of the DT system. An asynchronous flavor is used for the 40 MHz scouting system in which preprocessing of the trigger primitives is performed at a rate of 250 MHz, providing histogram input approximately every six clock periods.

The module is implemented for the Xilinx 7-series and newer FPGAs. The readout of the histograms can be performed using any suitable data transfer protocol. In particular, the IPbus and DMA readout interfaces have been implemented for the DT and 40 MHz scouting systems, respectively.

8.2.3 Z boson counting

Due to a clean signature and a relatively high cross section, the $Z \rightarrow \mu^+\mu^-$ process is of particular interest for the luminosity measurement. It provides an alternative and complementary method to transfer integrated luminosity measurements between data sets, as discussed in refs. [214, 215]. In Run 3, Z boson counting is implemented as part of the data reproduction chain with the capability of providing precise results within about one week of the data taking. A full initial analysis of the 2022 data has been performed [216] and used to confirm the early calibration of the luminometers

by comparison with the predicted cross section [217]. The Z boson rates are also used as an additional method to directly compare the integrated luminosities delivered to the ATLAS and CMS experiments [218, 219].

8.3 Beam monitoring instrumentation

8.3.1 Beam-halo monitor (BHM)

To monitor the BIB at high radius, an efficient detection system was installed during LS1 and operated in 2015–2016. It was then recommissioned for Run 3. The BHM monitors [220] are composed of a total of forty cylindrical quartz Cherenkov detectors, twenty on each end of the CMS detector, mounted on the rotating shield. The rotating shield is situated at a distance of about 20 m from the interaction point. The exact location was chosen to maximize the time difference between the arrival of BIB and particles from the pp collisions. The BHMs are mounted on the outer radius of the shield, at a distance of 1.83 m from the beam line. The units are uniformly distributed in ϕ, starting from $\phi = -30^\circ$ and continuing through $\phi = 210^\circ$.
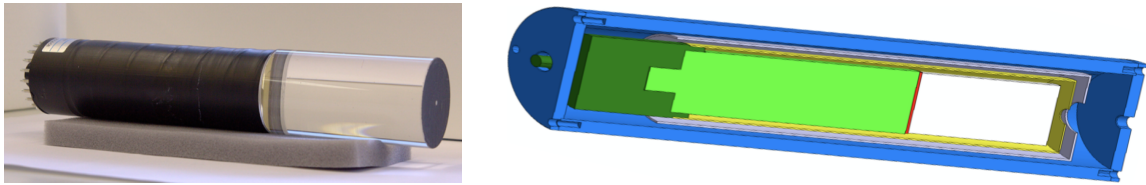
Figure 113. Left: photograph of the active element of one BHM detector unit, which is a 100 mm-long by 52 mm-diameter cylindrical quartz radiator, connected to a Hamamatsu R2059 photomultiplier. Right: shielding of a BHM detector unit, consisting of a low-carbon steel tube shown in blue, a mu-metal tube in gray, and permalloy in yellow. The quartz radiator, photomultiplier tube, and socket are shown in white, light green, and dark green, respectively. Reproduced from [220]. © CERN 2015. CC BY 3.0.

The quartz radiators, shown in figure 113 (left), are optically coupled to a fast ultraviolet-sensitive PMT on one end and painted black on the other. These detector units use the directional nature and timing of the BIB background and the pp collision products. In order to use the directional property, the PMT side of the quartz detector faces the center of the CMS detector. When a BIB muon arrives with the incoming beam and produces Cherenkov radiation in the quartz, the light propagates forward and is collected by the PMT. In contrast, pp collision products arrive from the opposite direction, and the produced light is absorbed by the black paint.

The shielding of each BHM unit consists of three coaxially arranged metallic tubes, as illustrated in figure 113 (right). The outer tube is a 10 mm-thick, 38 cm-long tube made of mild steel with endcaps on both sides composed of soft iron (ARMCO) with the same thickness. The intermediate layer is a 30 cm-long, 1.5 mm-thick mu-metal tube, and the innermost tube is 27 cm long and 0.8 mm thick made of permalloy. The shielding is designed to protect the PMT tubes from the 17 mT residual magnetic field and from radiation that mainly originates from low-energy electrons and positrons produced in the pp collisions, which otherwise would form the dominant contribution to the BIB.

The readout of the BHM detector makes use of many components developed for the HCAL Phase 1 electronics upgrade [60], with dedicated firmware and readout adapted to the beam-monitoring requirements [221]. The PMT signal is digitized by a charge-integrating ASIC (QIE10),

providing both the signal rise time and charge integrated over one bunch crossing. It ensures dead-timeless readout of the signal amplitude and edge time information with 500 ps resolution. As a part of the readout from the backend electronics, the occupancy histograms are obtained in time bins of a configurable duration. The histogram bins are typically 6.25 ns wide and are read out every 4 NB (about 1.458 s). In special running periods, detailed self-triggered data containing event-level information about amplitude and time-of-arrival of each hit can be also collected.

A calibration and monitoring system was also installed for the BHMs [222] to evaluate possible changes in performance of the PMTs and quartz radiator due to aging and radiation damage. The system uses a light signal produced by UV-emitting pulsed LEDs, which is sent to each detector unit through quartz optical fibers and optical splitters. The monitoring system can be set to send a laser pulse periodically in an empty orbit or in periods when beams are not present.

8.3.2 Beam-condition monitor for beam losses (BCML)

The BCML detector [223–225] is linked to the abort system of the LHC, and protects the CMS silicon tracker from beam-loss events. The detectors simultaneously measure the integrated currents over 12 different time intervals and produce an unmaskable beam-abort trigger if one of them exceeds a threshold, which is configured to be at least three orders of magnitude less than the amount estimated to cause damage. The BCML system is situated in four different locations within the CMS detector and includes the BCML1 detector located at $z = \pm 1.8$ m and the BCML2 detector located at $z = \pm 14.4$ m from the interaction point. A total of 16 channels of the BCML detector are actively used in the beam-abort system, with four channels in each location, providing CMS with monitoring redundancy.

The BCML system primarily uses pCVD diamond sensors, most of which were replaced in LS2. In addition, BCML2 has sapphire sensors installed, as described below. Diamond detectors are typically used as robust beam monitors in locations where the radiation levels are very high. Diamond is radiation hard and does not require active cooling.

There are four pCVD diamond sensors installed at each end of both the BCML1 (figure 114) and the BCML2 (figure 115) detectors, 16 sensors in total, each having an active volume of $10 \times 10 \times 0.4$ mm^3. The current created by ionization in the sensor is proportional to the ionizing particle flux through the active detector material, and it thus provides a good observable to determine the amount of radiation that a sensor receives. The BCML sensor signals are read out using the LHC beam-loss monitor (BLM) electronics [226–228].

If a beam loss signal, i.e., an integrated current over threshold, is detected in BCML1, the beam can be dumped within two to three LHC turns. The threshold is defined for two different integration time intervals. A short-duration (40 μs) threshold protects the silicon tracker from potentially damaging amounts of energy deposition in the electronics, and a longer-term (≈ 83.3 s) integration threshold protects against high beam background conditions that can result in problematic conditions for data taking and reconstruction, as well as an increased dose rate to the inner detector region. A machine beam interlock system [229] ensures that the BCML system is fully operational before beam injection into the LHC can begin. The BCML systems are powered independently of CMS from an uninterruptible power system, so that the LHC can run even if the local CMS power distribution is not functional.

New optical grade single-crystal sapphire sensors were installed in BCML2 during LS2 in six stacks. These are produced industrially via the Czochralski process at low cost [230]. Similar to
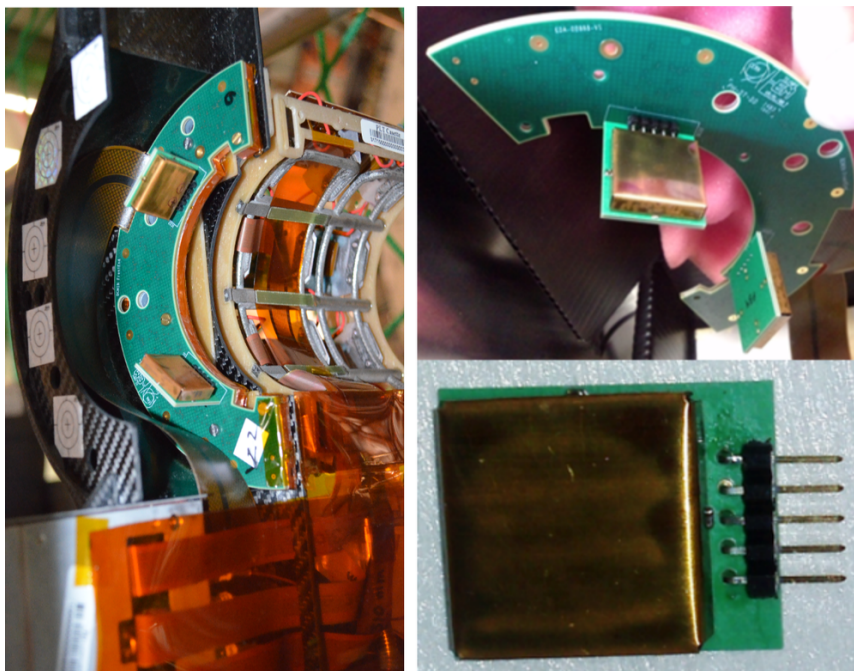
Figure 114. Mechanics for the BCML1 detector system. Left: a BCML1 mounted on the BCM1F C-shape PCB, attached to the PLT support structure. Right: a BCML1 mounted on a C-shape PCB (upper right), and a single sensor in a Faraday cage (lower right).
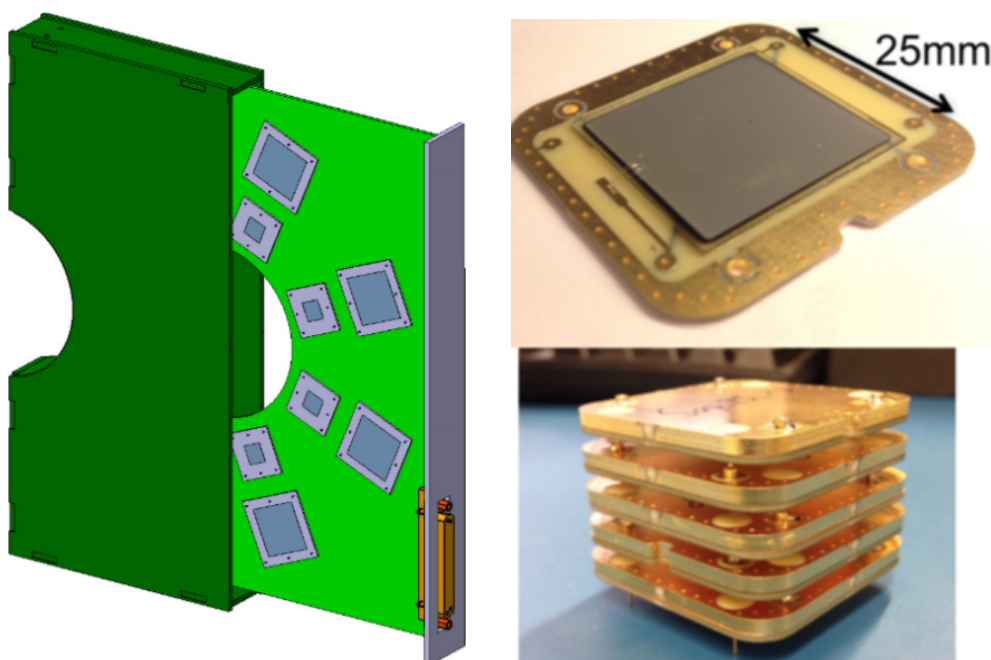
Figure 115. Mechanics for the BCML2 detector system. Left: the support structure of the BCML2 detectors, which also serves as mechanical protection. This structure is compatible with the Phase 2 beam pipe. Right: single-sensor base plate PCB for the large sensors used in BCML2 (upper right), and a complete sensor box assembly, with a stack of five sensors (lower right).

diamond, sapphire, chemically aluminum oxide Al_2O_3, behaves as a wide-band-gap semiconductor, with a band gap of 9.9 eV . First measurements indicated a very high tolerance against radiation damage [225, 231], although the charge collection efficiency is low ($\approx$1–10%) [231]. The sapphire Phase 2 prototype sensors used in BCML2 have an active volume of $25 \times 25 \times 0.5\,mm^3$ and can be installed in units ranging from a single sensor to a stack of up to five sensors, as shown in figure 115. In the stack option, the sensors are electrically mounted in parallel such that the observed current is the sum of the currents in each sensor.

To normalize the BCML2 readings and thereby monitor nonlinearities and instabilities in the response from the diamond and sapphire sensors, in addition to the solid-state detectors, two ionization chambers (one per end) of the LHC BLM type [226] were placed on top of the table originally designed to support the Castor subdetector. This is located at about $z = 1450$ cm, directly behind BCML2, and the chambers are connected to the same readout electronics as the BCML2 system. These detectors provide a good dynamic range, demonstrate linearity with a deviation less than 1% [232], and no radiation degradation is expected. The proximity to the BCML2 detectors, and the fact that these ionization chambers are in the same readout system as BCML2, makes them ideal for normalizing the BCML2 readings and thereby monitoring the nonlinearities and instabilities in the response from the diamond and sapphire sensors. Since the response of these ionization chambers is also well known for a mixed radiation field, the detectors can also be used to benchmark the radiation simulations.

8.4 Radiation instrumentation and simulation

Radiation background can cause unwanted signals in detectors, activation of materials, damage to electronics, and detector degradation. A radiation simulation and monitoring strategy is in place such that background radiation can be estimated, detected, analyzed, and lowered where necessary.

8.4.1 Radiation monitoring

The radiation monitoring strategy for Run 3 includes the use of the following systems:

- The HF RadMons [233] are gas-filled proportional counters situated inside the HF regions. There are eight in use during Run 3, four on either end of the CMS detector. The detection principle is based on neutron capture by boron with the emission of an α particle in the argon-filled proportional counter. The incident neutrons are slowed by a surrounding polyethylene moderator with a 3-inch radius. The HF RadMons provide information on the total neutron flux for benchmarking the simulation estimates.

- As the part of a wider radiation and environment monitoring unified supervision (REMUS) system under the responsibility of the CERN radiation protection group (HSE/RP), during Run 3 there are 12 inducted activity monitors (IAM) in the CMS cavern, read out by the radiation monitoring system for the environment and safety (RAMSES) [234, 235]. Ten of these are in the same place as in Run 2, and two additional ones were installed in the same location as the existing monitors on the HF, to support luminosity stability studies. The primary function of the IAMs is to measure the ambient-dose equivalent rates, $H^*(10)$, associated with the residual radiation field when the LHC is not operational. However, their

large dynamic range means that they can also provide reliable measurements of the prompt radiation field during collisions. They are used to monitor any upgrades to shielding in the rotating shield region and provide benchmark data for simulations. This system is also used by BRIL for luminosity stability and linearity studies.

- The LHC RadMon system [236, 237], which is maintained by the CERN EN/STI group, is designed to monitor the radiation that is related to electronics damage. There are currently 29 LHC RadMon units situated inside the experimental cavern, and it is foreseen to add ten more during Run 3. Each unit, upgraded to "Type v6", has 13 detectors, which includes two radiation-sensitive field effect transistors (RadFETs) to measure the total ionizing dose, three Pin Diodes (in series) to measure the 1 MeV-equivalent fluence in silicon, and eight static random access memory (SRAM) modules. The latter measure the cumulative fluence of high-energy hadrons via single-event effects and thermal neutrons using a different voltage setting. The BRIL group also uses the LHC RadMon data for radiation simulation benchmarking studies.

In addition to the monitoring systems listed above, several established methods are used to provide information about the radiation field. These include activation samples that are placed in the outer cavern with subsequent measurements performed by the HSE/RP group; monitoring of CMS detector degradation; and the two LHC beam-loss monitors that are used as part of the BCML setup, described in the previous section. Radiation monitoring is complemented by dedicated simulation studies, used to predict and understand in more detail the background radiation in CMS.

8.4.2 Radiation simulation

Monte Carlo simulations to predict radiation levels in the experimental cavern and the CMS detector are typically performed with the CERN FLUKA radiation transport code [238, 238]. The maintenance of the FLUKA geometry models is the responsibility of the BRIL radiation simulation team, as well as the dissemination of the simulation results to the CMS subdetector teams via a web-based tool. The baseline pp collision simulation using a geometry model that reflects the current CMS Run 3 configuration is tagged v.5.0.0.0. Relative to the Run 2 FLUKA geometry, this includes the implementation of the LS2 upgrades, the replacement of the central beam pipe from the IP to $z \approx 16.7\,\mathrm{m}$, modifications of material within the rotating shield regions including reinforcements to fill existing gaps, as well as modifications to the installation of the LHC vacuum assembly for experimental area (VAX) equipment in LS3. The radiation levels per integrated luminosity (or collision) in the central detectors for Run 3 are expected to be higher than in Run 2 by up to a factor of 1.5, depending on the location and type of particle, owing to changes in the shape of the central beam pipe. Changes in the predicted radiation levels in the outer detectors for Run 3 are influenced by the increased showering in the central regions (resulting in a reduction downstream) and modifications inside and outside the rotating shield region where a lot of the secondary radiation is generated and leaks into the cavern. The lighter aluminum beam pipe results in significantly lower activation levels for a given irradiation pattern. The FLUKA predictions of the radiation environment in Run 3 are illustrated in figure 116 for the CMS cavern and detector. The effects of an additional new forward shield, expected to be installed in 2024 around the existing rotating shield and included in geometry model v.5.1.0.2, are shown in figure 116 (lower).
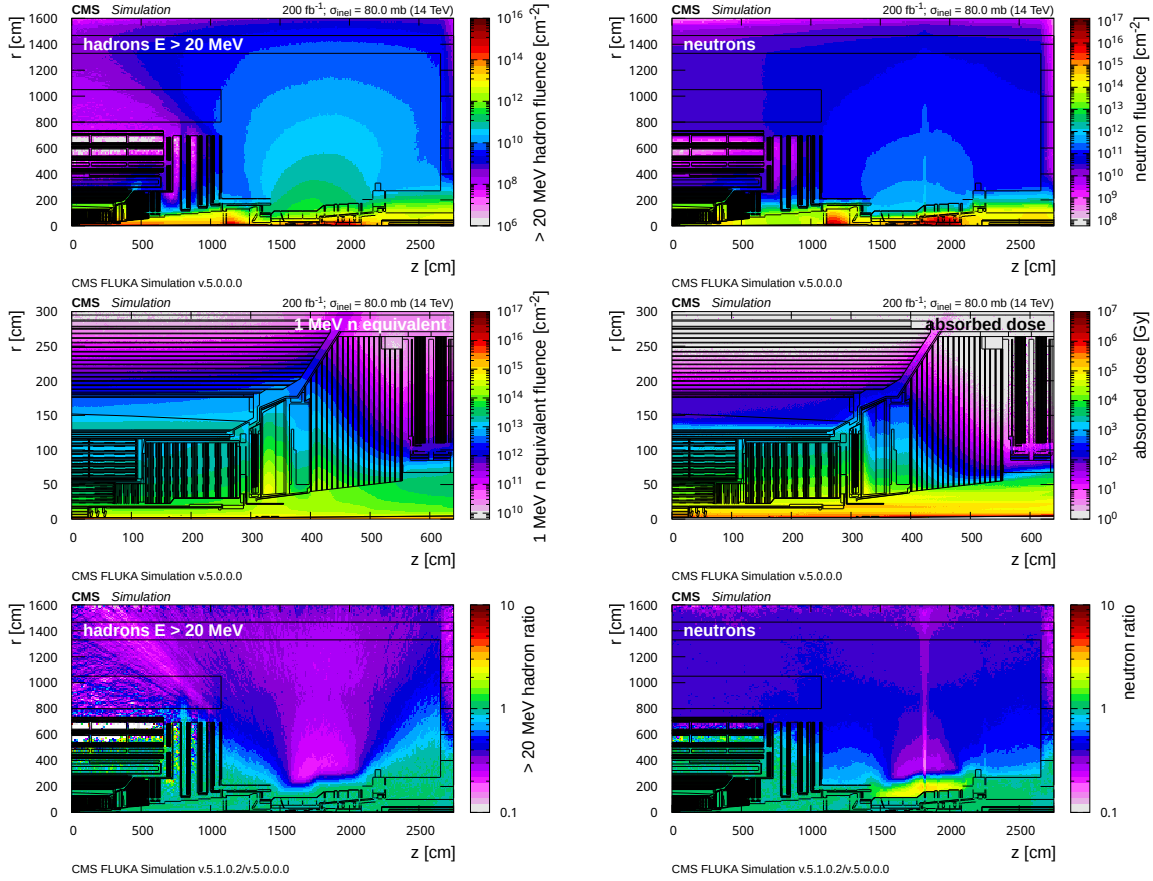
Figure 116. The FLUKA predictions for the expected Run 3 fluence of hadrons with energies greater than 20 MeV (upper left) and neutron fluence (upper right) normalized to an integrated luminosity of 200 fb^{-1} at 7 TeV per beam are shown for the CMS cavern and detector. For the central part of CMS, 1 MeV-neutron-equivalent fluence (middle left), and absorbed-dose (middle right) are also presented. The lower two plots show the expected effect of the new forward shield as the ratio of hadron (left) and neutron fluences (right) in the CMS cavern comparing the Run 3 FLUKA simulation results of v5.1.0.2 with the Run 3 baseline of v5.0.0.0.

8.5 The BRIL online data acquisition and monitoring

The BRIL data acquisition (BRILDAQ) system facilitates the real-time recording, processing, and monitoring of luminosity and beam background data at a latency of only a few seconds. Essential BRIL data are delivered in real-time to the CMS and LHC control rooms, allowing for prompt and efficient operation of the LHC machine and CMS detector. BRILDAQ is completely separate from the central CMS DAQ system, such that BRIL can operate continuously and independently of the CMS data-taking status even when all other CMS detector components are switched off.

Similar to the central DAQ system of CMS, BRILDAQ manages distributed and heterogeneous subsystems using a homogeneous distributed software architecture, where reliability and redundancy are important to prevent down time. The BRIL data from different subsystems can be viewed together and correlated in real time, and new subsystems can easily be integrated.

The core software is based on the CMS online framework XDAQ [239]. This event-driven and message-oriented architecture enables large numbers of loosely coupled software components

and services to exchange information in near real time. Messages are exchanged via the publisher/subscriber model (eventing) where the publishers send messages that can be received by any number of subscribers. Communication between BRILDAQ and the LHC is based on the data interchange protocol (DIP) [240], which also uses a publisher/subscriber model.

To fully exploit the redundant online luminosity and beam background monitoring systems, the histogramming of the BRIL subsystem frontends is synchronized using common timing signals distributed by the CMS TCDS system (section 9), defining the hit count integration interval boundaries. Long command counters are also distributed for additional intersystem synchronization. Using such a technique facilitates uniform accounting of the delivered luminosity and the downstream data handling by BRILDAQ.

Typically, a subsystem provides one or several source instances, which read the raw histograms from the hardware, and one single processor instance for data aggregation and calibration. Both kinds of applications are stateless to ensure the availability of the data, regardless of the status of the LHC beams and the running state of CMS. In Run 3, the participating subsystems are PLT, including the fast-or and the Slink data, BCM1F, with both RHU and μTCA readouts, HF, DT, and REMUS for luminosity measurements, and other BRIL measurements from subsystems such as the beam-pickup system for timing measurements (BPTX) [212], HF-RadMon gas-filled proportional chambers, BHM, and BCML. High-level components include the storage manager, the vdM monitor providing the luminometer calibration constants in real time after a beam-separation scan, the luminosity monitor, and the best-luminosity selector. The latter selects the value of the instantaneous luminosity that is delivered to the LHC. The DIP-related components use the service provided centrally by the CMS DAQ team. In addition, it is foreseen to process PCC data out of the HLT in near-real-time fashion, as mentioned in section 3.

The run control system, responsible for controlling and tracking the configurations of the BRIL applications, is based on the CMS run control framework (RCMS) [73] and its configuration database. Each BRIL process is controlled by a function manager (FM) that manages a simple finite-state machine.

The run control web frontend allows the operator to control the life cycle of BRILDAQ applications and to manage the configurations, which are versioned and can be retrieved and stored in the database via the web. A set of other web tools completes the system by providing useful functionalities such as displaying logs of the processes and showing results of the data analysis.

The BRIL online web monitor is a single-page application built on the Angular framework [241] where the website interacts with the user by dynamically rewriting the current web page with new data instead of reloading the entire page. It displays real-time and historical charts of online quantities. Monitoring data are published by BRILDAQ applications, then stored in the ElasticSearch [242] backend database, which responds to requests from the web client. The monitoring system is nonintrusive and flexible for easy integration of new data and charts.

The raw and calibrated data are stored on local disks, then transferred to an offline storage area, and eventually moved to tape. Summary data are loaded into the BRIL database. All types of data can be reprocessed, and multiple versions can coexist on disk and in the databases.

To provide high-quality luminosity measurements for use in physics analysis, an application toolkit is provided to the physics community, centered around the so-called "brilcalc" tool for luminosity calculation. It provides delivered and recorded luminosity with different calibration sets,

using a "normtag" that defines the best detectors and calibrations to use for each time period. This strategy has proven to be successful and remains unchanged in Run 3.

9 Data acquisition system

This section describes the scope, design choices, and implementation of the experiment's central data acquisition (DAQ) system. The present implementation for Run 3 is described in detail along with a discussion of its performance. We also give an overview of the evolution of the system over the lifetime of the CMS experiment.

9.1 Scope

The CMS online event selection is performed using two trigger levels: the level-one (L1) trigger, described in section 10, implemented in custom electronics, which selects approximately 100 kHz of events based on coarse information from the calorimeters and the muon detectors; and the high-level trigger (HLT), described in section 11, which runs on a farm of commercial computer nodes integrated with the DAQ data flow. The HLT processes fully assembled events, applying algorithms similar to those used in offline reconstruction, and selects a few kHz of events for storage on disk.

While this two-level approach with full event building after the first level greatly simplifies the overall system design compared to approaches with more trigger levels and/or partial event building, the resulting requirements on the data acquisition system are demanding: the DAQ system needs to read out approximately 700 detector backend boards at a rate of ≈ 100 kHz and perform event building and distribution with a throughput of about 100 GB/s— a challenging task when the DAQ system was first implemented [1] for Run 1 with the hardware available in the late 2000s.

The DAQ system is also responsible for collecting events selected by the HLT, buffering them at the experiment site, and transferring them to the Tier 0 computing center at CERN. The general structure of the DAQ system has remained fairly constant from its original implementation for Run 1 up to its current implementation (illustrated in figure 117). Across implementations, the main components of the system are:

- A trigger throttling system (TTS), consisting of custom electronics modules collecting fast status information from all the backend boards in order to throttle the trigger to avoid buffer overflows.

- Since Run 2, the trigger control and distribution system (TCDS), implementing the trigger control logic, reacting to fast status signals from the TTS or collected directly from upgraded FEDs, and distributing timing and trigger signals to the entire experiment.

- Custom electronics modules, the frontend readout (optical) link boards that receive data over a custom link from the detector backend boards (in CMS called frontend driver or FED boards), and send data over a commercial link.

- A data concentrator network based on commercial network technology.
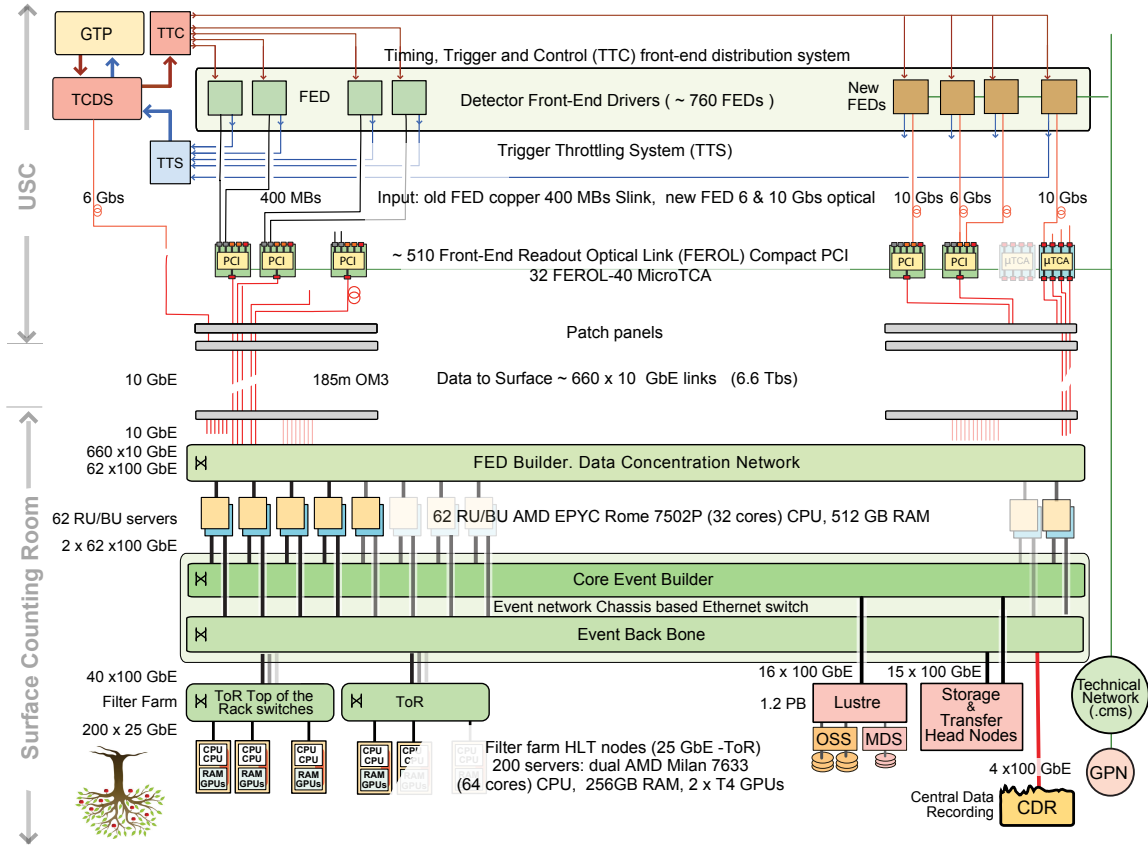
Figure 117. Diagram of the Run 3 DAQ system. The total numbers of cabled elements are given, including elements used in the MiniDAQ systems described in section 9.11 and elements installed as hot spares or for contingency. In typical global data-taking configurations, a subset is used, as described in the text.

- Readout unit (RU) servers that aggregate data from $O(10)$ FEDs into super-fragments and buffer the data.

- An event builder network based on commercial network technology.

- Builder unit (BU) servers that build full events from super-fragments received from all the RUs. In Run 3, the RUs and BUs run on the same set of servers (RU/BU nodes).

- An event backbone network, based on commercial network technology, connecting the BUs to the filter unit (FU) servers.

- The FU servers that execute the HLT algorithms.

- A storage and transfer system (STS), merging events selected by the HLT into larger files, buffering them locally, and transferring them to the Tier 0 at CERN via the central data-recording (CDR) network.

The DAQ system comprises software for data handling and controlling custom electronics, a hierarchical control system to manage most of the aforementioned components, a monitoring system that collects metrics, and several monitoring clients that interpret these metrics.

9.2 Evolution

Since its first implementation for Run 1, the CMS DAQ system has undergone multiple upgrades to keep pace with the evolving needs of the experiment. Towards the end of Run 1, it became clear that during Run 2 the LHC would reach higher instantaneous luminosities than originally anticipated, resulting in increased pileup and larger event sizes. In response, several subdetectors implemented upgrades and replaced VME-based backend electronics with new systems based on the μTCA standard. A new optical readout link with a higher bandwidth of up to 10 GB/s was developed for these new backend systems. The custom electronics of the DAQ system were upgraded to support this optical link and to support a new network technology for the data-concentrator network, from Run 2 onwards.

The bulk of the DAQ system downstream of the custom readout electronics consists of commercial computing and networking equipment, for which a regular replacement cycle must be observed. This circumstance provided the opportunity to reimplement large parts of the DAQ system for Run 2, and again for Run 3, benefiting from advances in technology to achieve a much more compact design while doubling the event building bandwidth from Run 2 onwards. The choice of networking technology has been adapted following trends in the supercomputing industry. An overview of the three generations of the CMS DAQ system is provided in table 12.

The HLT computing capacity was scaled up according to the experiment's needs on a yearly basis over Run 1 and Run 2. The Run 3 system includes general purpose graphics processing units (GPUs), providing cost-effective computing acceleration. This has required a major effort in porting the HLT code. Major changes in the software enabled better decoupling of the DAQ and HLT processes since Run 2. Various automated features and automatic diagnostics have been added to the control system to maximize the uptime of the DAQ system. Driven by the need for more detector partitions than the original system could support, and to host μTCA backends, a new trigger control and distribution system (TCDS) was developed for Run 2, absorbing the functionality of the original trigger control system (TCS), and TTC system, and providing optical inputs for the TTS signals from the upgraded backends. In the following sections, the upgrades to the DAQ system are described in more detail with an emphasis on the implementation for Run 3.

9.3 Subdetector readout interface

During Run 1, the CMS subdetectors were read out exclusively through the S-LINK64 [200] DAQ readout link, an LVDS-based copper link capable of transferring up to 400 MB/s. One or two such links are received by the frontend readout link (FRL) [246], a custom Compact-PCI card. The FRL forwarded the data to a commercial Myrinet network interface card (NIC) via an internal PCI-64 bus at 66 MHz. Super-fragments were built on the Myrinet NICs using custom firmware [246]. Assembled super-fragments were then transferred into the memory of a readout unit (RU) server via DMA.

At the start of Run 2, the Myrinet NIC on the FRL was replaced by a custom-developed PCI-X card, the frontend readout optical link (FEROL) [247]. The FEROL acts as a 10 Gb/s Ethernet NIC, sending data to the event builder. It receives data either via the PCI-X interface from the FRL (from the bulk of the subsystems that continue to use S-LINK64) or via optical SlinkExpress inputs from new or upgraded subsystems. Up to two SlinkExpress inputs at 6 Gb/s or one input at 10 Gb/s are supported. While S-LINK64 senders are mezzanine cards plugged onto the backend boards of the subdetectors, requiring a significant footprint, the SlinkExpress sender is a firmware IP-core that

Table 12. Key parameters of the CMS DAQ system in Run 1 [1, 243], Run 2 [244, 245], and Run 3.

	Run 1	Run 2	Run 3
Event building rate pp[a]	100 kHz	100 kHz	100 kHz
Event size pp[a]	1 MB	2 MB	2 MB
Read-out links S-LINK64 (copper) 400 MB/s[b]	636[c]	575[c]–532[d]	528[c]
Read-out links optical[e] 6 Gb/s[b]	—	55[d,f]	55[c,f]
Read-out links optical[e] 10 Gb/s	—	60[c]–167[d]	176[c]
FED builder network technology	Myrinet	Ethernet	Ethernet
FED builder network speed	2 rails of 2.5 Gb/s	10 & 40 Gb/s	10 & 100 Gb/s
Event builder # of readout units	640	108[d]	57[g]
Event builder network technology	Ethernet	Infiniband	Ethernet RoCE v2[h]
Event builder link speed	1–3 rails of 1 Gb/s	56 Gb/s	100 Gb/s
Event builder parallel slices	8	1	1
Event builder network throughput	1.0 Tb/s	1.6 Tb/s	1.6 Tb/s
Event builder # of builder units	1260[i]	73[d]	57[g]
BU RAM disk buffer	none	16 TB	10 TB
HLT # of filter units	720[c,i]–1260[d,i]	900[c]–1084[d]	200
HLT # of cores	5.8k[c]–13k[d]	16k[c]–31k[d]	26k[j]
HLT computing power (MHS06)	0.05[c]–0.20[d]	0.34[c]–0.72[d]	0.65[j]
HLT # of NVIDIA T4 GPUs	—	—	400
Storage system technology	16 SAN[k] systems	1 cluster file system	1 cluster file system
Storage system bandwidth write + read	2 GB/s	9 GB/s	30 GB/s
Storage system capacity	300 TB	500 TB	1.2 PB
Transfer system to Tier 0 speed	2×10 Gb/s	4×40 Gb/s	4×100 Gb/s

[a]Design value. [b]Main data-taking configuration, excluding links from partition managers used for partitioned running. [c]At the beginning of the run. [d]At the end of the run. [e]SlinkExpress. [f]54 links from mezzanine cards with optical SlinkExpress. [g]Readout and builder unit running on the same server ("folded event builder"). [h]Remote DMA over Converged Ethernet. [i]Filter and builder units running on the same server. [j]Not including the GPU computing power. [k]Storage-area network.

can be included in the FPGAs of the backend boards. Only the footprint of an optical transceiver cage is needed, facilitating the move to smaller form factors such as μTCA. The SlinkExpress works with 8b/10b encoding at up to 6.3 Gb/s or 64b/66b encoding at 10.3 Gb/s, resulting in an effective bandwidth of up to 5.0 or 10.0 Gb/s, respectively. The data format (definition of headers and trailers) is identical to that of S-LINK64. The SlinkExpress is packet-based and supports re-transmission at the packet level. Packets have a variable size of at most 4096 bytes. Fragments up to 4096 bytes are transferred into individual packets, while larger fragments are split across multiple packets.

At the output side, the FEROL sends data via TCP/IP, employing a custom TCP/IP engine implemented in the FPGA logic. This was achieved through a simplification of the TCP/IP protocol for unidirectional use, which reduced the number of states from 11 to 3 [248]. The TCP/IP streams (one per input) are sent via an Ethernet network to a standard Ethernet NIC in the readout unit (RU) server where they can be received with the standard Linux TCP/IP stack. Using performance tuning as described in ref. [247], a sustained point-to-point throughput of 9.7 Gb/s has been achieved for fragments larger than 1 kB.

For subdetectors that were potentially limited by the copper S-LINK64's bandwidth, but did not upgrade their backend electronics, a new type of mezzanine card with optical transmission was developed. It plugs onto the subdetector readout electronics in place of the original mezzanine cardand transmits data using the 6 Gb/s SlinkExpress, increasing the bandwidth to 625 MB/s. A version using 10 Gb/s SlinkExpress is also available. This mezzanine card has been deployed for the ECAL subdetector with 6 Gb/s SlinkExpress links.

In 2017, a new pixel detector was installed with new backend electronics requiring readout through 108 links at 10 Gb/s. Instead of producing more FEROL boards that are powered and controlled through PCI-X because of the legacy interface to the FRL, a new readout board was implemented, based on the μTCA standard, incorporating the features of four FEROL boards. The new FEROL-40 board [249] receives up to four channels at 10 Gb/s using the SlinkExpress and sends the data on four links of 10 Gb/s Ethernet using TCP/IP.

Table 13 shows the readout parameters of the CMS experiment at the end of Run 1 and Run 2, and at the start of Run 3. Starting in Run 3, the HCAL barrel partitions are read out with higher segmentation, as described in section 5.2.8, resulting in a significant increase in data size. Nine additional readout links were added for the HCAL. The dependence of the data size on the pileup (PU) and the resulting number of vertices has been studied for each data source. Data sizes observed at peak PU in Run 2 resulted in a total event size of 1.4 MB. The event size in Run 3 at the planned peak luminosity of $\mathcal{L} = 2 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ ($\langle\text{PU}\rangle = 56$) is 1.6 MB. An extrapolation, using a polynomial fit, to $\mathcal{L} = 3 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$, $\langle\text{PU}\rangle = 85$, results in an estimated event size of 2.0 MB. This is still within the design value of the DAQ system, which is therefore considered capable of handling the readout bandwidth at conditions beyond the planned peak luminosity in Run 3.

For heavy ion runs, where the L1 trigger rate will be 50 kHz, the expected event size based on the latest heavy ion run in 2018, taking into account the upgrade of the HCAL readout, is 3.3 MB. The overall throughput at the input to the DAQ system will thus be similar to that in pp runs. The FED sizes (FED can indicate both the subdetector electronic interface and, in this context, the data payload that is read out from that interface) is, however, distributed in a different way. In many cases, a special optimization of zero-suppression and/or selective readout algorithms at the level of the FEDs is applied in order not to be limited by the bandwidth of individual FEDs.

9.4 Event builder

The CMS experiment uses a two-stage event builder (EVB) system responsible for assembling event fragments retrieved from around 760 detector backend boards into a single event payload and delivering the built events to the HLT. The first stage, called the FED builder, reads fragments from the underground FEROL and FEROL-40 boards and, using the switched network, aggregates them in the RU/BU nodes. In the second stage, the core EVB, all the RU/BU nodes transfer fragments from each event into one destination node, assigned on a per-event basis. With the Run 3 DAQ system, each node in this way handles around 2 kHz of fully-built events. On each node, events are written into a large 200 GB RAM buffer and made available to the HLT via a dedicated data network.

Network interconnect technologies are a main driver for the EVB design and scaling. The Run 1 FED builder system [243] was based on the Myrinet 2.5 Gb/s network, comprising 640 RU nodes and split into 8 parallel slices for performance. Prior to Run 2, Infiniband and 40 Gb/s Ethernet became established mainstream standards in high-performance computing (HPC). They

Table 13. Subdetector readout configuration.

Subdetector	End of Run 1		End of Run 2 (start of Run 3)					Data size [kB]	
	# FED	# FRL	# FED	# FRL/FEROL copper	6 Gb/s	10 Gb/s	# F40[a]	Run 2[b] ⟨PU⟩ = 56	Run 3[c] ⟨PU⟩ = 56
Tracker pixel	40	40	108	—	—	—	32	259	
Tracker strips	440	250	440	250	—	—	—	731	
Preshower	40	28	40	26	—	—	—	54	
ECAL	54	54	54	—	54[d]	—	—	74	
HCAL	32	32	32 (41)	—	—	32 (41)	—	170	+221
Muons CSC	8	8	36	18	—	—	—	40	
Muons RPC	3	3	3	3	—	—	—	0.3	
Muons DT	10	10	9	—	—	9	—	22	
Trigger	5	5	14	—	—	14	—	41	
CASTOR	4	4	4 (0)	4 (0)	—	—	—	—	
TCDS	—	—	1	—	1	—	—	1.0	
CTPPS	—	—	11	9	—	2	—	2.2	
Muons GEM	—	—	2	—	—	2	—	—[e]	+22
Total	636	434	754 (759)	310 (306)	55	59 (68)	32	1.39 MB	1.63 MB

[a]FEROL-40 (4×10 Gb/s). [b]Observed at Run 2 peak luminosity. [c]Only the increment in data size with respect to Run 2 at the same peak luminosity is shown. [d]Mezzanine card with optical SlinkExpress. [e]Not included for regular data taking.

were evaluated [244, 250] and ultimately adopted [245, 251] as technologies of choice for the Run 2 DAQ upgrade. This allowed the system to be scaled down by an order of magnitude and implemented in a single slice, while increasing the overall EVB throughput [252].

For Run 3, the Run 2 equipment, which had reached the end of its vendor support, had to be retired. The design choice was taken to implement the EVB with 100 Gb/s Ethernet for the Run 3 system. This system, described in the following sections, satisfies nearly identical performance requirements as the Run 2 DAQ, and, owing to hardware evolution, is reduced in scale by more than a factor of two. Some design details of the preceding Run 2 system are also outlined.

9.4.1 FED builder

The FED builder transports data fragments over 10 Gb/s links from the underground cavern, over a distance of around 200 m, to data concentration network Ethernet switches located on the surface. A total of 557 links from the FEROL and FEROL-40 boards are used in a typical Run 3 data-taking configuration, as detailed in table 13.

The data concentration network is implemented using a Juniper QFX10016 [253] chassis-based 100 Gb/s Ethernet switch, providing a flat network, which allows full flexibility in routing the fragment traffic between readout nodes. For inbound traffic from the FEROLs, seven QFX10000-36Q line-cards are installed, capable of supporting a total of 672 10 Gb/s links, split from 24×40 Gb/s ports per line-card. Three QFX10000-30C cards installed in the switch provide a capacity of up to ninety

100 Gb/s ports for connections to the RU/BU nodes. A pause frame [254] flow control mechanism is used to achieve a lossless low-latency network and ensure optimal buffer occupancy on the FEROLs. The system is configured to support Ethernet jumbo frames of up to 9 kB to improve performance.

9.4.2 Core event builder

The core EVB has three distinct functional units, the RU, BU, and event manager (EVM). They are implemented as software applications and run on a single set of RU/BU computers, interconnected using a dedicated event-building data network. These applications are implemented within the XDAQ framework described in section 9.10.

The RUs unpack and merge fragments received via TCP streams from the FEROLs over the data concentration network. Each fragment checksum is verified before being merged with others into a super-fragment, i.e., a collection of fragments from the same event aggregated by a particular RU, and buffered in memory. The EVM, which runs on a single EVB node, orchestrates the event building process by performing a destination assignment, i.e., allocating a given event to a specific BU and sending a message to the RUs to send super-fragments to that destination. When the BUs receive a super-fragment from each RU the event is completely built. Completed events are written into ramdisk, a memory buffer structured as a file system, as described in section 9.5. This file system is exported through the event backbone network to the FUs. An FU group of typically three or four FUs is assigned to read the event data from each RU/BU node.

Starting in Run 2, EVB senders and receivers were implemented using the Infiniband Verbs-API, described in section 9.10, for communication over the network using remote direct memory access (RDMA) capabilities of network-interface cards and switches. This technology facilitates the bypass of the operating system networking stack and delivers the payloads directly into the application-accessible memory, avoiding CPU and memory overheads associated with the handling of high bandwidth and packet rate networking in software.

For the Run 3 EVB, the cost-effective 100 Gb/s Ethernet technology was chosen over the native Infiniband used in Run 2. A second Juniper QFX10016 chassis-based switch is used as the core EVB network backbone (equipped with nine QFX10000-30C line-cards having a total of 270 100 Gb/s ports). The switch supports a lossless Ethernet network, a prerequisite for RDMA over Converged Ethernet (RoCE) v2 protocol, providing encapsulated Infiniband protocol over Ethernet hardware. Importantly, like native Infiniband, it facilitates direct memory access between communicating nodes using offloading by NICs. A benefit of using RoCE v2 is that the Verbs-API-based EVB applications developed in Run 2 can be reused with minor adaptations.

In addition to the core EVB network, the switch runs the event backbone network, which handles the traffic between the EVB and HLT farm. The same chassis switch supports also the storage and transfer system (including transfers to Tier 0), traffic from the DAQ to the online data quality monitoring system, and running the online cloud on some of the legacy Run 2 HLT nodes. The 62 RU/BU nodes are connected to the EVB network over 100 Gb/s optical links, including 57 nodes as part of the nominal data-taking configuration and five as hot spares. In addition, there are seven cold spare nodes without network interface cards.

As pointed out above, the Run 3 system also introduced a folded core EVB setup, where the RU/BU nodes serve as both RU and BU functional units (RU/BU). This allows the bidirectional utilization of the network links, nearly halving the number of needed RU/BU nodes and the network

bandwidth compared to a nonfolded setup, thus reducing the overall size and cost of the system for the same throughput requirement. On the other hand, this design is demanding on the I/O and memory performance of the nodes, which in such a system need to receive and merge fragments, exchange super-fragments with other RU/BU nodes, build and serve full events, and pass-through the HLT output to the STS, all over multiple 100 Gb/s interfaces. Therefore, a modern server architecture was required for the task. Dell R7515 servers, with the 32-core AMD EPYC Rome 7502P CPU running at 2.5 GHz, equipped with 512 GB of DDR4 RAM, were selected, with all memory channels populated for the maximum memory performance. Each server is equipped with two Mellanox ConnectX6 dual-100 Gb/s PCIe Gen4 NICs. One card has both links used for the respective connections to the FED builder and event backbone networks, while the second card uses one link for connection to the core EVB network.

9.4.3 Performance

Detailed studies were done comparing the EVB performance of both the Intel Xeon Skylake dual-socket servers [144] and comparable AMD EPYC Rome single-socket platforms. The AMD platform was ultimately chosen for RU/BU nodes for Run 3 due to better performance and a simplified memory architecture. The AMD 7502P CPU is internally assembled from multiple 4-core silicon dies (CCDs), each having a separate L3 CPU cache and an interconnect fabric providing eight memory controllers, as well as multiple PCIe Gen4 interfaces. Due to these characteristics, the CPU internally resembles, to an extent, the nonuniform memory access (NUMA) memory architecture. It can be configured in a 4-, 2-, and single-node NUMA mode with respect to the CPU die and memory controller topology.

The single-node NUMA mode uses interleaved access to the RAM controllers with improved maximum bandwidth, with the drawback of potentially higher memory and I/O latency. In evaluations, this mode was found to be a well-balanced setup for folded EVB requirements, avoiding the delicate tuning of thread affinities to CPU cores, which would have been needed to optimize performance. The setup maintains CPU affinity settings for groups of threads running similar tasks to the same or adjacent CPU cores and placing them closer to the corresponding NIC PCIe lines in the interconnect topology.

A memory-based file system based on Linux tmpfs [255] is used as the EVB event output ramdisk. To achieve high read/write performance, 2 MB huge-page support for tmpfs is used, available in recent Linux kernel versions. It significantly improves the ramdisk throughput since it avoids bottlenecks when the CPU is handling a large number of 4 kB memory pages.

Measurement of the DAQ event building performance with emulated data generated by the RUs and discarded after event building, is shown in figure 118. This demonstrates that the EVB system is capable of handling the nominal 100 kHz event rate from the L1 trigger with event sizes up to ≈2.5 MB. A throughput of approximately 10 GB/s is achieved per RU/BU node at the plateau, which is 80% of the available network bandwidth.

The performance for input generated in the FEROLs and either discarded after the EVB or passed through the full-chain DAQ system (including the HLT and STS) is shown in figure 119. Two setups of a smaller scale than the data-taking setup were used for this measurement, comprising either four or 19 RU/BU servers in a folded setup, with an additional machine running an EVM application. Each RU application was concentrating 20 fragments from the FEROLs (two fragment streams per 10 Gb/s FEROL link, or a total of 100 Gb/s of the input network bandwidth per RU). A total of either 16 or 76 FUs, four assigned to each RU/BU, were used. These servers were Run 2
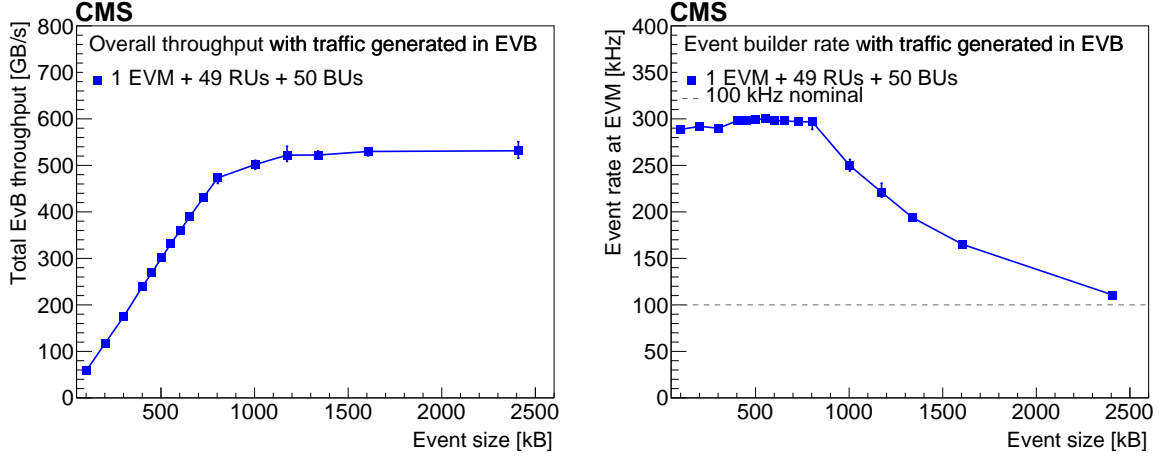
Figure 118. Core event building throughput (left) and event rate (right) of the full-scale RU/BU EVB setup shown for a range of built event sizes. Emulated input data are generated in the RU/BUs and discarded after the event building stage.

FU nodes equipped with 10 Gb/s Ethernet NICs. Thus, the total available bandwidth was 40 Gb/s over the event backbone network to the FU group. With each system, a plateau throughput of about 9 GB/s was achieved per RU (BU), with discarding the data after the EVB, amounting to over 70% of the available network bandwidth from the FEROLs. About 4 GB/s per server was achieved with the full-chain DAQ, which is about 80% of the available network bandwidth to the FUs.
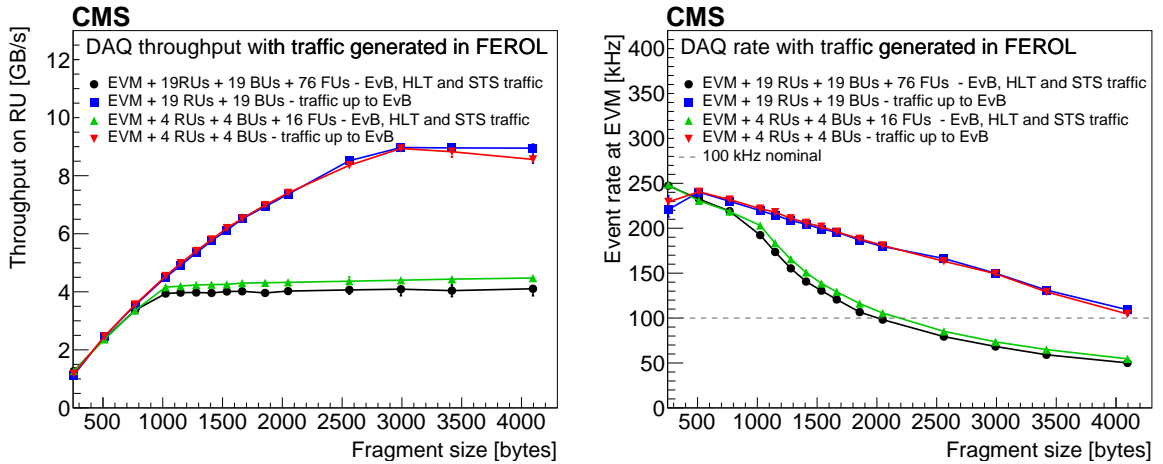
Figure 119. The DAQ throughput (left) and event rate (right) per RU/BU node for a range of uniform fragment sizes using a mode with the data discarded after the EVB and with the traffic flow through the HLT and STS. Emulated input data are generated at the FEROLs with 20 fragments concentrated per RU/BU.

This result demonstrates that the system is capable of handling traffic for fragment sizes larger than 2000 bytes at rates above 100 kHz. When extrapolated to the larger system of 50 RU/BUs (nearly the scale of the system used in Run 3), a fragment range around 2000 bytes corresponds to the total event size of over 2 MB, which is above the size expected in Run 3. With the fixed number of FUs per RU/BU, the achieved bandwidth approximately scales from the small- to the medium-size

system and gives confidence that the system comprising over 50 RU/BU nodes can robustly scale and handle the Run 3 requirements. Furthermore, while tests were performed using a 10 Gb/s network bandwidth per FU, the Run 3 FUs have more bandwidth available from the 25 Gb/s Ethernet links, as mentioned in section 9.5.5, and, therefore, higher full-chain throughput capacity.

9.4.4 Event builder load balancing

The EVB protocol implements a global dynamic re-balancing mechanism of the event building throughput on the RU/BU nodes. The assignment algorithm, running on the EVM, periodically determines the load status of each node, depending on factors such as the local output buffer occupancy, as well as the number and type of HLT computers attached to a particular RU/BU. From these, it determines the rate at which each node can build events to fairly distribute the load and equalize the processing latency at the EVB and HLT across the cluster. A local throttling mechanisms on the RU/BUs can also get triggered, based on occupancy thresholds and latency information reported by the HLT infrastructure software. This results in the BU application temporarily blocking or throttling the event building at the corresponding RU/BU.

9.5 Event filter

The HLT performs the second stage of event filtering, analyzing at a rate of approximately 100 kHz, the events selected by the L1 trigger, and accepting events at a rate on the order of a few kHz. It runs on a large cluster of multicore servers (FUs) that perform the reconstruction and filtering in software, implemented within the CMS software framework (CMSSW) [256].

By using the same framework and code base for the online and offline reconstruction, most of the algorithms developed in CMSSW for the offline reconstruction can also be employed in the HLT. This enables the reuse of the same data structures, supports a unified approach to the use of detector condition data, and allows for rapid deployment of both newly developed and established offline algorithms.

Starting from the first year of Run 2, the CMSSW and HLT implemented task-based multithreaded event processing. This feature was instrumental in reducing the memory requirements of the HLT: by running a smaller number of processes on the FUs, each using multiple threads, common data such as detector conditions and calibrations can be shared among threads, decreasing the overall memory footprint on the servers. As a result, this made it possible to fully exploit all the logical cores of the CPUs, leading to increased overall computing efficiency. At the beginning of Run 3, support for offloading to accelerators, such as GPUs, was also added. In addition, reducing the number of running processes per server was found to be beneficial in reducing the memory footprint on the GPU.

The framework architecture is described in more detail in section 12. The HLT algorithms and offloading are detailed in section 11. A dedicated software infrastructure couples the CMSSW workflow with the EVB and the storage and transfer system. It handles the data input and output for all these processes, as well as the bookkeeping of event processing and monitoring.

9.5.1 File-based filter farm

The HLT interface to the DAQ, running in the same process as the HLT algorithms, is implemented in ordinary CMSSW modules and services, which facilitate the communication and reading of input data from the RU/BU nodes. The data transfer into the HLT starts from a ramdisk file system serving

as a data buffer on the RU/BU [257]. An NFS (network file system) v4 server provides access to the file system over the network. The ramdisk is sized to buffer the data during the HLT startup time of 60–90 s, dominated by the loading of the software libraries, calibrations, and conditions, as well as to absorb fluctuations in the HLT event processing time.

The BU application writes fully built events into the ramdisk, using a custom binary file format with a header prepended to each event describing its total length and providing information such as an identification number and a checksum. The files created by the BU application, each containing around 100 events, are distributed among CMSSW processes running on several FUs assigned to the RU/BU node. A dedicated application running on the RU/BU nodes, called the file broker, is responsible for the unique assignment of files to the processes. The requests for files are placed with the file broker using an HTTP-based protocol, while the standard file system access API is used to read files from the NFS. All input-related communication and file access is implemented in the CMSSW input source module written in C++. This module operates several dedicated threads for pre-buffering of the input data to minimize stalling the CPU cores while waiting for input. The input source parses the assigned input files, extracts and verifies the event payload against stored fragment checksums, and hands it over to the framework, which schedules the HLT modules performing the event reconstruction, analysis, and filtering.

9.5.2 The HLT software infrastructure

The HLT daemon is a system service orchestrating the HLT operation in the RU/BUs and FUs. Implemented as a Python application, it relies on a file notification API of the Linux kernel, inotify [258], which provides low-overhead monitoring of file system events. The start of a run is triggered by the appearance of raw event data files in the ramdisk. Similarly, a run completes when all files have been processed by the HLT. The service runs on all RU/BUs, where it is responsible for contacting the attached FU nodes to process a run, as well as the FUs, where it is responsible for managing the life-cycle of the locally running CMSSW processes. The HLT daemon starts the CMSSW applications as forked child processes, monitors their status, and can restart them in case of unexpected process termination. Each child CMSSW process is allocated a number of CPU cores or hyperthreads (in the case of using simultaneous multithreading (SMT), a technique allowing separate tasks to run on the same CPU core), also taking into account the number of parallel threads configured in the CMSSW processes, as also described in section 11. By using such an allocation strategy, servers can be used at their maximum computing capacity.

9.5.3 The HLT menu and output data streams

The HLT configuration, also referred to as the "HLT menu" and described in section 11, defines a set of paths in which physics objects are reconstructed and events are filtered based on specific physics requirements. Events that are accepted through HLT paths for similar physics processes are stored in primary data sets, defined also in the HLT menu. The primary data sets are defined such that the total event rate of each data set is kept within the limits imposed by the offline data processing, while minimizing the overlap among different data sets. In turn, the primary data sets are grouped into streams, corresponding to the actual files that are output by the HLT processes and transferred to the Tier 0 for offline processing. In addition to those used to collect events for physics analyses, the HLT defines dedicated paths and streams to collect data for detector calibrations and for online

data quality monitoring (DQM). Finally, special paths and streams are used to collect statistical information about the HLT processes themselves, such as the individual trigger rates and CPU usage.

Events selected by the HLT are serialized using ROOT [259], compressed using gzip [260], LZMA [261], or Zstandard [262], and written to files on the local partition. Lossless compression used for the HLT output, applied to the entire serialized event payload, reduces the total event size to about 70% of the input event size for proton runs (specified in table 13), as observed in Run 2 and the first year of Run 3. For proton runs, it is preferable to minimize the CPU usage of the compression in order to preserve resources for the HLT reconstruction. The gzip tool, which was used in Run 2 and in 2022, typically used up to on the order of 1% of the CPU in the HLT during Run 2 and 2022.

The application of Zstandard was evaluated after the first year of Run 3 and shown to have improved performance for a compression factor comparable to gzip, with an estimated use of less than 0.5% of the CPU. The HLT began using this algorithm for compression in 2023 for proton runs. For the heavy ion run, the compressed size is estimated at about 3 MB for centrally colliding collision events and 1.2 MB for minimum-bias events, based on the 2018 heavy ion run. The LZMA algorithm, which can typically improve the compression factor by an additional 10–15%, but is an order of magnitude more CPU demanding than gzip or Zstandard, is being considered in order to maximize the HLT output rate within the DAQ bandwidth limits. The approaches to size reduction of heavy-ion events are discussed in section 11.6.

Each output stream is written into a separate file and includes the event raw data, a selection or a summary of the HLT objects, and the trigger decision for each path specific to a stream. The streamer file format used for output allows files belonging to the same stream, but produced by different processes and nodes, to be trivially concatenated. This step, which is performed separately for each stream, is called merging. In the first merging stage, streams are merged from local CMSSW processes and copied to the NFS-mounted output partition on the RU/BU. Subsequent merging stages are handled by the storage and transfer system (STS), as described in section 9.6.

The CMSSW input, output, and merging operate on sets of data and metadata files created separately for each stream at a regular time interval called the luminosity section (LS), spanning 2^{18} LHC orbits, which takes approximately 23.3 s as explained in section 8. Metadata, written by the BU, by the CMSSW processes, and by the HLT output merging at LS intervals, specify information such as event counts, checksums, and total output size. Bookkeeping (completion) and integrity checks are performed at the granularity of an LS by comparing input and output metadata. In the case of a failed integrity check, alerts are raised to notify the shift crew, and the affected data are discarded and marked as missing in the bookkeeping.

The HLT daemon services handle all the local output collection and merging tasks on the FUs. They also report the latency in event processing and merging to the BU application. The BU uses this feedback to throttle event building in case of high delays and can activate throttling in case of high ramdisk occupancy.

9.5.4 Monitoring

The file-based filter farm (F3) monitoring system was designed from the ground up in Run 2 [242] around the Elasticsearch NoSQL database [263]. Structured JSON-formatted [264] monitoring data are injected into an Elasticsearch cluster from the HLT daemon service, STS, and other sources. This includes bookkeeping such as input and output event counts, event file size, bandwidth, CPU

usage, HLT information, and the hardware state at the granularity of each FU node. Textual logs and error reporting from the HLT processes are also handled. Information is stored for an extended time of up to several years, allowing real-time as well as follow-up analysis of performance and faults in the system. Several web-based tools have been developed as frontends to the system, visualizing the information for experts and the shift crew. Monitoring of the HLT performance is also collected using non-event streams, such as histograms, which are merged and shipped to the DQM system, and trigger statistics, which are merged and stored in a database.

9.5.5 Evolution of the HLT farm

The F3 was gradually expanded almost every year during Run 2, as the HLT processing requirements increased due to evolving LHC and detector conditions. The FU nodes are typically replaced with a new generation of hardware after the end of their 5-year warranty periods, with the old nodes being assigned to the online cloud, as described in section 9.9.2. Table 14 summarizes the composition of the HLT farm at the end of Run 2 and at the beginning of Run 3, along with the computing power estimates based on the HEPSPEC 2006 [265] (HS06) benchmark measurements. Prior to the start of Run 3 data taking, the entire farm was replaced with new AMD CPU nodes, each equipped with two NVIDIA T4 GPUs [266]. In this configuration the average processing time of the HLT can be up to 500 ms per event, at the nominal L1 rate of 100 kHz.

Table 14. Summary of the HLT filter farm unit specifications, thermal design power, and performance based on HS06 in the final year of Run 2 and first year of Run 3.

Run	Run 2 (2018)			Run 3 (2022)
Architecture	Intel Haswell	Intel Broadwell	Intel Skylake	AMD Milan
CPU model	dual E5-2680v3	dual E5-2680v4	dual Gold 6130	dual 7763
CPU cores	2×12	2×14	2×16	2×64
Nominal freq. [GHz]	2.5	2.4	2.1	2.45
Turbo freq. [GHz]	3.3	3.3	3.7	3.5
TDP [W]	120	120	125	280
Memory [GB]	64	64	96	256
Nodes	360	324	400	200
CPU cores (total)	8640	9072	12800	25600
HS[a]/node	538	659	773	3224
TDP [W]/kHS[a]	223	182	162	87
kHS[a]	194	214	309	645
GPU card	—	—	—	2×NVIDIA T4

[a]HS06 measurements only take into account the CPU performance. The precision is around 1%.

Specific HLT reconstruction algorithms, such as those used in data scouting, described in section 11.4, can be offloaded to these GPUs, reducing the average HLT event processing time by over 40%, as shown in figure 134 and discussed in section 11. In comparison to an HLT farm

equipped only with CPUs, this corresponds to a reduction in the farm's overall cost by approximately 15% and its power consumption by 30%.

On F3 nodes, CMSSW jobs are divided into two groups, and, for each group, the CPU and memory affinity is pinned to a single CPU socket NUMA domain on dual-socket AMD nodes. Each GPU is attached to PCIe lanes on an individual CPU socket and assigned to the group of processes running on the same socket. The NVIDIA multiprocess daemon is used to schedule GPU access between processes, providing a small performance enhancement over direct access. A total of eight jobs with 32 threads and 24 streams each (i.e., parallel event processing pipelines in CMSSW) is used per node, since this configuration was found to fit within the memory capacity of the nodes and GPUs. It was determined that increasing the number of threads per process in CMSSW adds a small additional overhead, and thus it was not further increased to keep the optimal processing capacity of the F3.

To support NFS data transfers in Run 3, a Juniper QFX5120 top-of-the-rack (ToR) switch is employed, which uses 8×100 Gb/s up-links to the event backbone network from each rack. The network contains approximately 40 FU nodes with 25 Gb/s Ethernet connections to the ToR switch. A flat interconnect network is provided, enabling any FU to communicate with any RU/BU node.

9.6 Storage and transfer system

The storage and transfer system is the last stage of the DAQ data flow. It collects the HLT output from each RU/BU and associated FU group and writes it to a cluster file system. It later transfers the data to Tier 0 for repacking from the streamer format into the ROOT format and to permanent storage on disk and tape servers.

The HLT daemon services on the FUs within the FU group concatenate output streamer files, analyze metadata, and copy the corresponding files into a dedicated RU/BU output partition. This was a spinning disk RAID array in Run 2, and is a 200 GB ramdisk partition in Run 3 to support higher throughput. The merger service running on the RU/BU nodes periodically polls the streamer and JSON metadata files in this area, written by each FU for a particular stream and luminosity section. The JSON metadata files are used to verify the completion of per-LS output. Once all FUs have copied a complete set of files, the service distributes tasks to several worker threads to read the output of all streamer files and append them into a single file location in the distributed file system. For most streams and a majority of the output bandwidth, this step is done by simultaneously writing into a single file at a different offset. This technique is used to efficiently merge data into the final file object, one per stream and LS. This file can be transferred by a single copy operation to Tier 0, instead of requiring an additional read and write operation to perform such a concatenation.

The final stage of merging is performed on a set of dedicated STS nodes, with each handling a subset of streams. For most streams this amounts to verification of the metadata and checking the completion of files written by the merger service from each RU/BU. Fully merged streamer files, residing in a cluster file system, are handled and transferred to their destination by the transfer service, also running on the same nodes as the merger service. For data destined for the Tier 0, a pool of threads starts the file copy jobs to transfer data to the EOS disk system [267] using a high-speed link to the central data recording (CDR). The transferred file size is kept below 16 GB for optimal transfer throughput performance. The network infrastructure used to transfer to the Tier 0 is detailed in section 9.9. Streams are also delivered to several other destinations, such as the DQM and the

calibration cluster, or are parsed to extract the HLT and L1 trigger rate monitoring information and inject it into the relational database.

Extensive bookkeeping is required to track all the files passing through the system. This information is injected and visualized in the F3 monitoring system. Metadata relevant for the transfers are also written to an SQL database for bookkeeping and to provide transfer metadata to the Tier 0.

For Run 2, Lustre [268] was selected as the cluster file system, after initial evaluation in the DAQ integration system. The production system consisted of two disk servers (OSS) and a metadata server (MDS), providing on the order of 9 GB/s of total read and write throughput, with around 500 TB of storage space provided by a redundant hard drive setup.

In Run 3, the storage system requirements are driven mainly by the heavy ion running, which, apart from centrally colliding lead ion events, aims to collect a large amount of minimum-bias events. An estimated throughput of 17 GB/s is required in a scenario that includes a trigger selection of around 1 kHz of centrally colliding heavy ion events and a minimum-bias rate of 10 kHz at the HLT output, as described in section 9.5.3.

To cover these requirements, a hardware refresh was pursued in Run 3, retaining the same file system technology while significantly improving the bandwidth capability compared to Run 2. A new system comprising two DDN EXAScaler [269] SFA7990X data servers and a single SFA400NVX metadata server was acquired for the task. Each data server consists of 124 SAS 8 TB 7.2k RPM hard disks, organized in a RAID6 array. The metadata server comprises 23 SSDs, each of 1.8 TB capacity, in a RAID6 array. The system is connected via the chassis-based Ethernet switch also used for the EVB and HLT. Seven STS nodes were additionally installed and are connected to the chassis-based Ethernet switch. Together with the RU/BU nodes, these nodes are set up as Lustre clients and the file system made accessible through mount points. The Run 3 Lustre system storage space was scaled to provide several days of storage space for proton LHC runs or up to a day for heavy ion runs should the Tier 0 connection fail. The usable disk capacity is 1.2 PB, and the system is capable of simultaneously writing 24 GB/s and reading 11 GB/s using standard file system benchmarking tools. The TCP/IP protocol was used for communication in these tests. The system is capable of temporarily prioritizing the write bandwidth at the expense of the read bandwidth, up to the limit of the available disk space. This is particularly useful for runs with high peak bandwidth, such as heavy ion runs, where the read rate recovers and allows draining the accumulated files towards the end of the LHC fill and in interfill periods.

9.7 Trigger throttling system

The trigger throttling system (TTS) collects fast readiness signals from all FEDs, merges them per TTC partition with a priority logic, and makes them available to the trigger control logic to avoid buffer overflows by inhibiting triggers or drive recovery actions. While in Run 1 this trigger control logic was implemented in the trigger control system (TCS), since Run 2 it is part of the trigger control and distribution system (TCDS), described in section 9.8. A FED may signal the following main TTS states: Ready (to accept triggers), Warning/Busy (buffer fill level above high-water mark), Out-Of-Sync (synchronization loss), and Error (other error situation), in ascending order of priority. The TCDS reacts to the highest-priority TTS state by inhibiting triggers, executing a re-synchronization sequence, or executing a reset-sequence.

For the legacy FEDs, TTS signals are sent using four LVDS pairs, which are merged using compact-PCI based fast merging modules (FMMs) [246] that combine the TTS states of up to 20 FEDs using a priority logic. For partitions with more than 20 FEDs, the FMMs are arranged in a tree structure. For upgraded FEDs, TTS signals, transmitted over optical fiber, are merged with a similar logic by μTCA-based TCDS partition interface (PI) modules, as described in section 9.8.

9.8 Trigger control and distribution system

At the start of Run 2, a new TCDS [270] was introduced, replacing the Run 1 trigger control system that was integrated into the L1 trigger system, and the Run 1 TTC system. The TCDS provides support for an enlarged range of detector partitions and for detector backends sending their trigger throttling signals over optical fiber. These were both needed to integrate the additional and upgraded subdetectors with μTCA backends during Run 2. The TCDS distributes timing and control (synchronization) data that are flowing to the detector backends and frontends and receives back status information related to the readiness of the detector systems to handle more triggers. The clock reference that is distributed along with the fast control information is synchronous with the beams in the LHC, which is required to keep the data taking in step across the various detector systems. The TCDS is implemented in the μTCA architecture.

As illustrated in figure 120, a central crate contains a central partition manager (CPM) board and up to twelve local partition manager (LPM) boards. The CPM receives the LHC clock from the TTC machine interface and the L1 accept signal from the global trigger, as described in section 10.3). Each of the LPM boards contains eight independent integrated CMS interface (iCI) logic blocks that are able to control a detector partition for local running. The iCI blocks translate the generic TCDS synchronization commands to subdetector specific commands. The iCI block also contains a partition-specific emulator of the APV25 tracker readout chip buffer levels that inhibits triggers that would lead to overflows of these buffers. Each LPM board also contains two partition manager (PM) blocks that can orchestrate combined runs with sets of partitions in the same LPM. The CPM contains one PM block orchestrating global runs with all partitions. The PM blocks provide the following functionality:

- Trigger throttling, taking into account:

  - the TTS state of partitions;

  - trigger rules, suppressing bursts of triggers by limiting the number of triggers in certain windows of bunch crossings, as required by the subdetector frontend electronics;

  - protection against overflows in the pre-shower frontend ASIC (in a similar way to the protection against overflows in the APV25);

  - resonant trigger protection, i.e., protection against triggers arriving at regular intervals over prolonged periods, which could give rise to resonant vibrations that might damage systems (such as wire bonds),

  - the DAQ back pressure to the PM's readout link.

- Bunch mask trigger veto that can be used to inhibit triggers (such as prefiring triggers) in LHC bunch crossings that are not filled.

- Generation of random triggers.

- Generation of (periodic) calibration triggers to trigger the readout of calibration events (interleaved with physics events).

- Definition of sequences of synchronization commands such as run start, run stop, and detector backend recovery.

- Automatic triggering of sequences based on the trigger throttling state, for example, for global re-synchronization or subsystem-specific recovery actions.

- Monitoring of trigger rates, sequences, and dead times.

A secondary set of local and central partition manager hardware, connected in parallel to the primary partition managers, is available for final validation of firmware and/or software updates, or as hot spares.

Figure 120. Overview of the CMS trigger control and distribution system (TCDS).

Each of the iCI blocks is connected to a partition interface (PI) board, fanning out triggers and synchronization commands to a single detector partition and merging the throttling signals from the partition. These boards are located in separate μTCA crates. For the bulk of the backend boards, the PI board receives the TTS signal from an FMM, as described in section 9.7, while triggers and synchronization commands are sent via backwards-compatible TTC fibers and the legacy TTC hardware. For upgraded subsystems, the PI sends triggers and synchronization commands to up to

10 backend boards via bidirectional optical links. In the reverse direction, it receives TTS signals from the backends and merges these signals with a priority logic similar to the FMM.

9.9 Networking and computing infrastructure

The operation of the CMS experiment is supported by the CMS service network, a high-performance distributed network connecting all the computers directly related to the operation of the experiment. While isolated from the CERN general purpose network (GPN), connections to the CDR and LHC technical network (LHC-TN), as well as the GPN, are allowed for use in specific cases. The network follows a redundant design established in the early years of the experiment. Currently, routers provide 1 Gb/s Ethernet connections to the servers. Available bandwidth is commonly used for user access, control, and monitoring. The majority of the DAQ computers are located in racks, to which the service network is provided using ToR switches.

The DAQ data networks have been significantly redesigned and upgraded throughout the lifetime of the experiment, as described previously in this section. To facilitate sending data taken by the experiment in Run 3 to the Tier 0, four 100 Gb/s links from the event backbone network to the CDR are employed. This facility is also used by the online cloud for accessing external data stores and services, as described in section 9.9.2.

A network-attached storage (NAS) system is used by CMS for core storage needs. This includes home directories and subsystem storage, exposed via standard remote file system protocols to nodes in the CMS network. The system acquired for Run 3 provides 1.3 PB of storage space. The CMS online computing system runs the CERN-supported version of Linux, CERN CentOS 7 (CC7) [271], as well as Red Hat Enterprise Linux 8 (RHEL8) [272]. Both of these operating systems are also used for the DAQ operation. Machines are initially installed (or reinstalled) using a network preboot execution environment (PXE) boot installation service. After Run 1, the Quattor [273] system was replaced with the Puppet [274] software configuration management tool, which handles the OS installation, configuration, and deployment of the online software in distributed and reproducible fashion on thousands of DAQ and subsystem computers.

9.9.1 Virtualization

Virtualization allows the reuse of physical computers by sharing them for multiple services that are running in virtual-machine (VM) instances, while simultaneously providing an isolated execution environment for each service. A virtual-machine infrastructure has been set up and used to run the CMS online services based on the oVirt [275] open-source virtualization management platform on the CentOS 7 OS. The VMs running on this infrastructure are installed from OS boot images provided on the network and using the Puppet configuration management system described above. The VMs are accessible on the service network via dedicated network names assigned by the DNS.

9.9.2 Online cloud

The HLT farm consists of a large number of multicore nodes, representing significant computing power. During interfill periods, week-long technical stops, and longer LHC and detector upgrade periods, this capacity is mostly unused for the HLT tasks. Thus, starting in Run 2, infrastructure was developed [276] to run an Openstack-based cloud overlay to use the computing resources for

offline production jobs. In order to not overlap with the HLT-based workflows, the FUs are able to run virtual machines (VM) that provide the necessary software environment for the worldwide LHC computing grid (WLCG) CMS jobs. These jobs run tasks such as simulation and reconstruction in full isolation from the HLT. In section 12, more details on the WLCG and workflows are provided. An API is implemented in the HLT daemon to facilitate the automatic suspension and startup of the cloud mode or resumption of the HLT. The cloud images can be suspended to disk and later resumed, allowing the quick save of unfinished jobs when, for example, the cloud needs to be suspended based on the LHC status. The system can also perform the switch automatically, reacting to the LHC state or, in another mode, monitoring the HLT CPU usage and dynamically re-allocating unused fractions of the HLT to the cloud. About 5% of the farm remains in HLT mode at all times to provide an operable HLT for cosmic ray data taking, commissioning, and tests.

Openstack VMs use a virtual local area network (VLAN) for access to the CERN services such as EOS, where the job input data, as well as the destination of job results, are located. The VLAN is implemented in the event backbone network and routed through CDR links to the CERN IT infrastructure.

The overlay cloud has been extensively used since Run 2 and successfully expands the CMS computing resources through opportunistic re-purposing of the hardware. Furthermore, it consists not only of the active HLT cluster, but also computing nodes that were retired from the HLT. They are kept in cloud operation as long as the computing infrastructure support is possible. Overall, the online cloud, which is classified as a Tier 2 CMS site, as described in section 12, is one of the major contributors to CMS offline computing, comparable to the largest Tier 1 sites in the amount of CMS production workload.

9.10 Software, control, and monitoring of the DAQ

Two software frameworks have been developed within CMS to implement the bulk of the experiment's online software: the C++ based XDAQ framework [239], used to implement hardware access and data transport, and the Java based run control and monitoring system (RCMS) [73, 277], used to implement the hierarchical control structure and main user interface. These two frameworks, which have been adopted by the central DAQ system and by all subdetectors, are maintained and enhanced according to the evolving requirements of the experiment and continue to be used in Run 3.

The XDAQ software is a platform designed specifically for the implementation of distributed DAQ systems. It has a layered middleware structure, providing support for communication, a web user interface, high-speed networking, hardware access, multithreading, performance tuning, monitoring, error reporting, and logging. The XDAQ system builds upon industry standards, open protocols, and libraries, e.g., TCP, HTTP, XML, and Apache Xerces. Notable enhancements to XDAQ include the development of new "peer transport" plug-ins to support new network technologies, such as RDMA using Infiniband Verbs [278, 279], and a new service-based approach to the configuration of the built-in monitoring and alarming infrastructure [280]. New XDAQ-based applications have been added to control new types of custom hardware and existing applications, such as the event builder enhanced with features such as load balancing and fault tolerance.

The RCMS is a framework based on web applications running inside container instances (Apache Tomcat), which provides the building blocks to compose a distributed hierarchy of nodes to control and monitor the state of XDAQ applications and other online applications used during data

taking. Control nodes are based on state machines, with system-specific control logic implemented in Java. They are steered and monitored through web user interfaces. The RCMS includes database schemas to hold the configuration of all software components, define hardware configurations, and manage the complex interconnects required for the two-stage event builder. Extensive tools for configuration management are available. In the area of the RCMS, the development of guidance systems for operators and automation have helped to make the operation of the experiment less error-prone and less reliant on the knowledge of experts. The overall configuration of the experiment can be automatically selected based on the state of the LHC, actions needed in response to certain LHC state changes, and actions like high-voltage ramping in the detectors that are performed automatically [277]. Recovery from regular single-event upsets and from other typical data-taking problems is fully automated [281, 282]. Configuration management tools have been enhanced according to new requirements in Run 2 and Run 3 supporting, for example, fine-grained data-flow optimization according to the network topology.

The filter farm and the storage and transfer system are controlled and monitored by online software based on Python and Elasticsearch, as described in sections 9.5.2 and 9.5.4.

A number of monitoring clients transform the raw monitoring data from both the XDAQ monitoring system and the Elasticsearch-based monitoring system into web-based graphical and textual monitoring displays used by the shift crew. The monitoring clients typically display instant data with a latency of a few seconds and can also be used to browse historic data to facilitate post-mortem analysis. The Java-based DAQ expert tool [283–285] detects all common data-taking problems by evaluating rules encapsulated in logic modules using snapshots of monitoring data. This helps the shift-crew with the sometimes difficult task of pin-pointing the cause of data-flow problems. With problems for which a recovery is known, the tool can drive completely automatic recovery actions. It consists of several micro-services responsible for reasoning, notification, and control of the recovery.

A switch monitoring system with a web-based graphical representation was developed to monitor the link status and performance metrics of the DAQ data networks and assist experts in diagnosing network-related failures.

In addition to specific aggregation and presentation tools for the DAQ and HLT, a general service is provided to the subsystems and collaboration at large to aggregate and present online monitoring data stored in the different databases. The online monitoring service (OMS) is a new Run 3 software tool replacing a set of web-based monitoring tools (WBM [286]) used in Run 1 and Run 2 to provide unified remote access to the monitoring data. The OMS uses a generic relational database model (Data Warehouse) and interface, and a web-based presentation framework by which information across heterogeneous data sources and formats is aggregated and presented. The presentation is organized in a structure of folders and pages that contain portlets typically displaying information in the form of tables and graphs, showing, for example, run and fill details, trigger rates, or subdetector monitoring.

9.11 MiniDAQ

In addition to the global DAQ system, self-service DAQ systems, called MiniDAQs, are provided for most of the CMS subdetectors. These setups can be used at any time by the subdetector groups for calibration runs, tests, and debugging using detector partitions that are not participating in the

global data taking. These setups have proven extremely useful since they allow independent testing under almost the same conditions as in global data taking. Trigger control for MiniDAQ systems is provided by one of the PM blocks in a TCDS LPM. During Run 1 and Run 2, these MiniDAQ systems ran on dedicated RU, BU, and FU servers and provided limited bandwidth with respect to the global DAQ system. In Run 3, MiniDAQ systems share the RU/BU and FU servers with the global system and provide a full bandwidth to each subdetector. The configuration of the MiniDAQ systems is dynamically updated to follow any changes in the global system.

10   Level-1 trigger

The level-1 (L1) trigger is implemented in custom hardware processors. It comprises calorimeter and muon trigger systems that provide jets, e/γ, hadronic τ, and muon candidates, along with calculations of energy sums, to the global trigger (GT). At the GT, the trigger decision is generated, based on the multiplicity and kinematic information of the various candidate trigger objects. The trigger configuration is implemented in a trigger "menu" comprised of several hundred "seed" algorithms. Upon a positive GT decision, the full detector data are read out for further filtering in the higher-level trigger (HLT). During LS1, in 2013–2014, the L1 trigger hardware was entirely upgraded, and has subsequently been operated successfully since 2016. A detailed report on this Phase 1 L1 trigger upgrade and performance with Run 2 data is given in ref. [5].

For Run 3, although no major trigger hardware upgrade was performed, new capabilities have become available already through new algorithmic approaches, some of which are based on machine learning (ML) techniques. Software such as HLS4ML [287] facilitates the use of ML techniques in FPGAs. Developments for Run 3 within the L1 trigger mostly focus on broadening the physics reach of CMS through the addition of dedicated triggers for long-lived particle (LLP) signatures, improving object measurement and calibration, utilizing the upgraded calorimeter trigger primitives (TPs) and additional muon TPs from the new GEM muon detector, and implementing additional calculations in the global trigger to provide greater flexibility in the design of L1 trigger algorithms.

The addition of a 40 MHz scouting system, commissioned in the early stages of Run 3, that receives data from both the calorimeter and muon L1 trigger subsystems, has the potential to further broaden the physics reach of CMS. It enables the readout of unfiltered data, reconstructed in situ at limited precision but at full bunch-crossing rate, and provides unprecedented monitoring capabilities. The following sections describe the Run 3 developments specific to each of the L1 trigger subsystems.

10.1   Calorimeter trigger

10.1.1   Calorimeter layer 1 trigger

The calorimeter layer 1 trigger receives TPs from ECAL, HCAL, and HF, calibrates them, combines the ECAL and HCAL TPs into single trigger towers (TTs), and transmits the TTs to layer 2 for further processing. Calorimeter TPs for triggered events are readout to DAQ, and used in the data quality monitoring (DQM) system, where they provide the input to the software emulator, such that online and emulated data can be compared in real time for monitoring purposes. For Run 3, layer 1 receives updated HCAL TPs (section 5.3), which improves the mitigation of out-of-time pileup, and updated ECAL TPs (section 4.5) with improved rejection of spikes caused by particles striking the avalanche photodiodes.

The TT energies are calibrated to account for energy losses due to inactive material in front of the calorimeters. The calibration is performed separately for each calorimeter. Since the inactive-material map is symmetric in ϕ, the calibration is performed as a function of η and transverse energy, E_T, only. Whereas for the ECAL TP calibrations, the scale factors vary by less than 20% across the E_T range, the HCAL and HF calibration scale factors are much more E_T dependent, varying by about 50%. Both ECAL and HCAL scale factors vary by about 20% across the η range.

Due to the large volume of data produced by the TPs and limited DAQ bandwidth, only TPs for triggered events are read out by the DAQ system. To study possible trigger bias, for approximately every 100th event that passes the full L1 trigger selection, validation data are read out, that also contain the TT information. In Run 3, validation events contain the ECAL TP data for five bunch crossings, including the two bunch crossings before and after the triggered bunch crossing. This information is useful for studying unexpected detector effects, and can be used to study ECAL prefiring (section 4.5), both using the DQM and offline analysis. This additional ECAL TP data can also be used for various optimization studies, such as monitoring the timing of the ECAL TPs through the L1 trigger path during commissioning phases.

10.1.2 Calorimeter layer 2 trigger

Calorimeter layer 2 receives calibrated TTs from layer 1, reconstructs jet, e/γ, and τ candidates, and computes energy sums. The energies of jet, e/γ, and τ candidates are calibrated as a function of p_T and η, and isolation and ID criteria are applied to e/γ and τ candidates. Pileup mitigation is applied to all objects to reduce the rates while maintaining high efficiencies. The layer 2 hardware remains the same as for Run 2. Ten main processor cards each process data from the entire calorimeter for a single bunch crossing in a time-multiplexed configuration. A single demultiplexer processor receives data from the main processors, performs the final calculation of the energy sums, and forwards the object collections to the global trigger. More details can be found in ref. [5].

A range of improvements to the layer 2 algorithms are being investigated for Run 3, most of which involve utilizing the updated ECAL and HCAL calorimeter TPs as discussed in sections 4.5 and 5.3, respectively. In particular, the ability to trigger on LLP signatures by identifying displaced jets using the additional HCAL timing and depth information available for Run 3 is being pursued to help broaden the LLP physics program of CMS.

Jets are reconstructed by summing the energies of a 9×9 window of TTs centered on a jet seed that must have an energy greater than 4 GeV, which corresponds to approximately the same jet size as jets reconstructed offline with $\Delta R = 0.4$ within the barrel calorimeter. The energy contribution due to pileup is estimated by summing the three lowest energy out of the four 3×9 regions on the boundaries of the jet and subtracting this pileup estimate from the jet energy, which is then calibrated.

For Run 3, an LLP jet identification algorithm is implemented that uses the HCAL timing and depth information. Each TT has an HCAL feature bit set in layer 1, which compresses six feature bits received from the HCAL backend, containing timing and depth information, into one feature bit. When the jets are reconstructed from TTs at layer 2, an LLP jet ID bit is set if the jet contains more than a configurable number of TTs with the HCAL feature bit set. Additional LLP jet algorithms have been added to the GT menu that require the LLP jet ID bit to be true. In addition to tagging LLPs, the ability to tag boosted jets with substructure using a pattern-matching technique, and the use of ML techniques to calibrate the jet energy and perform pileup subtraction are being investigated for Run 3.

The missing transverse momentum (p_T^{miss}) is calculated as the vector sum of the TT energies across the full calorimeter. During Run 2, a significant increase in the instantaneous luminosity and thus the pileup per bunch crossing revealed a nonlinear relation between pileup and L1 p_T^{miss} trigger rate, leading to a large increase of rate at large pileup. To maintain the L1 p_T^{miss} trigger thresholds while keeping the rates manageable, a pileup mitigation procedure was implemented using a lookup table (LUT) to exclude TTs below a configurable energy threshold from the p_T^{miss} calculation, based on the η of the TT, and an estimate of the pileup of the event. The TT energy thresholds were rederived for Run 3, using Monte Carlo (MC) simulation samples containing inclusive pp events with a pileup distribution reflecting that of Run 3.

A comparison of the L1 p_T^{miss} trigger efficiencies between 2018 and Run 3, for different L1 p_T^{miss} trigger thresholds that provide the same L1 trigger rate, is shown in figure 121. Compared to the 2018 LUT, the updated Run 3 LUTs show a significant improvement of the efficiency relative to the true p_T^{miss} in the event. As the LHC conditions evolve, the pileup mitigation will be updated to ensure optimal performance. In the current tuning, the energy threshold below which TTs are excluded is set such that 99.5% of TTs are excluded for inclusive pp events.
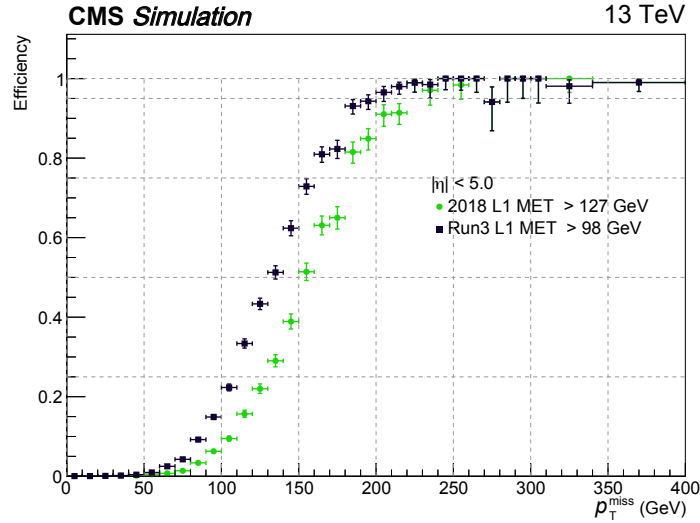
Figure 121. Comparison of the L1 p_T^{miss} trigger efficiency using pileup mitigation in 2018 (circles) and in Run 3 (squares) for thresholds that provide a rate of 4.3 kHz, for Z → μμ events.

The e/γ and τ candidates are constructed by clustering TTs containing energy deposits greater than 1 GeV around a cluster seed tower with an energy deposit of at least 2 GeV. This clustering is done dynamically using the available tower-level information. The cluster is trimmed by removing towers, and a veto based on the trimmed cluster shape is applied, to reject pileup and reduce background rates. A fine-grain veto is applied in the barrel calorimeter that quantifies the compactness of the electromagnetic shower within the seed tower to reject hadron-induced showers. A veto is also applied that requires a low HCAL-to-ECAL energy ratio (H/E) in the seed tower, with different thresholds used in the barrel and the endcap regions. Isolation requirements are applied to set an isolation bit and provide isolated candidates to the GT. Merged τ candidates are constructed by merging nearby clusters that pass a set of proximity conditions, to capture multiprong hadronic τ decays. Energy calibrations for e/γ candidates use the cluster shape and those for τ use H/E and the

presence of merged clusters, in addition to the p_T and η. While no significant changes to either the e/γ or τ algorithms have been implemented for Run 3, the existing calibrations and isolation working points have been and will be rederived throughout Run 3 to reflect updated detector conditions and calorimeter TP algorithms. Methods to improve the isolation working points of e/γ and τ candidates utilizing ML techniques are being implemented.

10.2 Muon trigger

The L1 muon trigger for Run 3 receives TPs from four partially overlapping muon subdetectors: DT, CSC, RPC, and GEM. As described in detail in section 6, three of these subdetectors, DT, CSC, and RPC, were operated during Run 2, while the GEM detector was added as part of the Phase 1 upgrade and is used for the first time in Run 3. The L1 muon trigger system reconstructs muon tracks and provides measurements of muon track parameters using TPs which provide position, timing, and quality information from detector hits. In the barrel, accurate directional information is also provided. The geometrical arrangement of the muon subdetectors, including the new GE1/1 detector in front of ME1/1, is shown in figure 39. In this section, the changes to the muon track finders are discussed in detail.

The L1 muon trigger system in Run 3 comprises the same overall design as in Run 2. Three muon track finders (TFs) reconstruct muon tracks in three distinct pseudorapidity regions using TPs from muon detectors. The barrel muon track finder (BMTF) receives inputs from DT and RPC in the barrel ($|\eta| < 0.83$), the overlap muon track finder (OMTF) uses DT, CSC, and RPC in the overlap between barrel and endcap ($0.83 < |\eta| < 1.2$), while the endcap muon track finder (EMTF) takes inputs from CSC, RPC, and GEM in the endcap ($1.2 < |\eta| < 2.4$). All three muon track finders transmit up to 36 muons each per bunch crossing to the global muon trigger (μGMT), which resolves duplicates and transmits a maximum of eight muon tracks per bunch crossing to the GT, similar to Run 2.

In Run 3, all three muon track finders additionally provide measurements of parameters for muon tracks that are displaced from the primary interaction point. The beamspot constraint requiring the track to originate from the interaction point is removed. The newly available track parameters are used in the GT to provide L1 muon trigger seeds targeting displaced-muon signatures that could originate from LLPs. Additionally, in Run 3, the EMTF receives TPs also from the GEM detectors, and this information can be used to improve both prompt and displaced muon triggering.

The new displaced-muon algorithms provide a p_T, measured without the beamspot constraint, and a transverse displacement, d_{xy}, from the beam line, for muon tracks obtained from propagating back to just the first muon station. Algorithms optimized for prompt muons typically underestimate the p_T of highly displaced muons, as the displacement is mistaken as increased track curvature due to the beamspot constraint. The new displaced-muon algorithms improve the p_T estimation for displaced muons, hence improving the efficiencies for these triggers. The displaced TF algorithms in general do not affect the muon track building, but provide additional measurements of unconstrained quantities. Due to this approach, the trigger efficiency is increased significantly for muon $p_T > 10$ GeV when the displacements are larger than 20 cm, while in the case of lower p_T muons, the displaced muon algorithms perform similarly to the prompt algorithms. This is due to the fact that all the prompt TF algorithms have a minimum p_T assignment value (2–4 GeV depending on the TF), and the underestimation of p_T for displaced muons becomes less important at low p_T values.

10.2.1 Barrel muon track finder (BMTF)

In the BMTF, the kBMTF algorithm reconstructs muons in the barrel region using TPs received from the DTs and RPCs via the TwinMux concentrator cards, and has been used online since 2018 [5]. It is the successor of the original BMTF algorithm, used between 2016 and 2018, and based on an approximate Kalman filter algorithm in which muon tracks are reconstructed from detector hits starting from the outermost muon station and propagating inwards while updating the track parameters.

In Run 3, the updated version of the kBMTF algorithm mainly improves the displaced-muon triggering performance. The new algorithm retains high efficiency for prompt muon tracks up to $d_{xy} \approx 50$ cm, while the displaced-kBMTF algorithm increases the efficiency of muon tracks with displacements above 50 cm. The expected performance for Run 3 was evaluated using a Run 2 cosmic ray muon data sample in which the displaced algorithm shows efficiencies around 80% for L1 $p_T > 10$ GeV, up to displacements of 100 cm (figure 122).
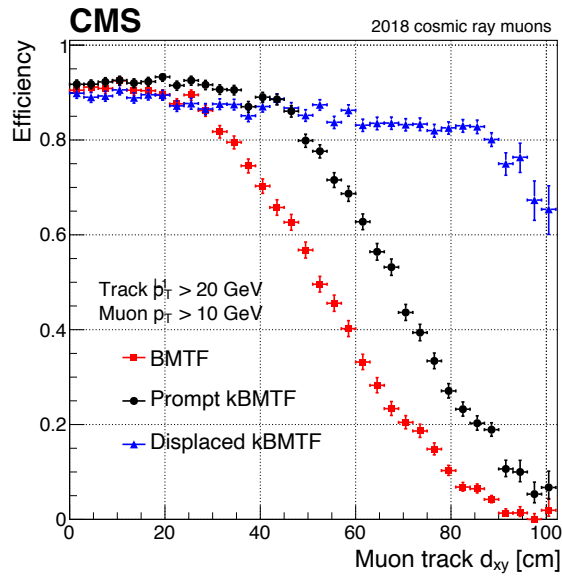
Figure 122. Displaced (blue) and prompt (black) kBMTF trigger efficiencies compared to the prompt BMTF (red) trigger efficiency with respect to the muon track d_{xy}, obtained using a sample of cosmic ray muons from 2018 data. The efficiencies are measured using muon candidates with $p_T > 10$ GeV. The prompt kBMTF improves BMTF efficiencies up to about 90% for up to 50 cm displacements, while displaced kBMTF retains efficiencies above 80% for up to 90 cm displacements.

10.2.2 Overlap muon track finder (OMTF)

The OMTF builds muon tracks using the TPs from the DTs and RPCs in the barrel and CSCs and RPCs in the endcap. The algorithm for Run 3, which is mostly identical to that of Run 2, uses a Bayes classifier algorithm based on precomputed patterns generated from simulated events to associate hits in each station with the reference hit in the pattern. The patterns contain information about muon track propagation and the probability density function of the hit distribution in ϕ with respect to the reference hit. There are 26 patterns for each muon charge, corresponding to p_T values between 2 and 140 GeV, that are then used to estimate the p_T of the muon track based on

the likelihood that the track matches a pattern. In Run 3, the OMTF includes additional patterns to improve displaced-muon triggering. Although the general structure of the algorithm remains the same, the updated Run 3 algorithm now finds the best matching prompt and displaced patterns for a given muon track. There are 22 displaced patterns for each muon charge corresponding to different d_{xy} values which are valid for high-momentum tracks ($p_T > 30$ GeV).

In Run 3, the OMTF uses prompt and displaced patterns in parallel for each track. The algorithm identifies prompt-muon tracks and estimates their p_T using prompt patterns, and at the same time it estimates the d_{xy} of the displaced muon using the displaced patterns. The expected performance for Run 3 was evaluated using a displaced-muon gun simulation sample, in which just muons are simulated with zero pileup, in order to compare ideal efficiencies, in which the displaced-OMTF algorithm shows efficiencies around 80% up to displacements of 200 cm (figure 123).
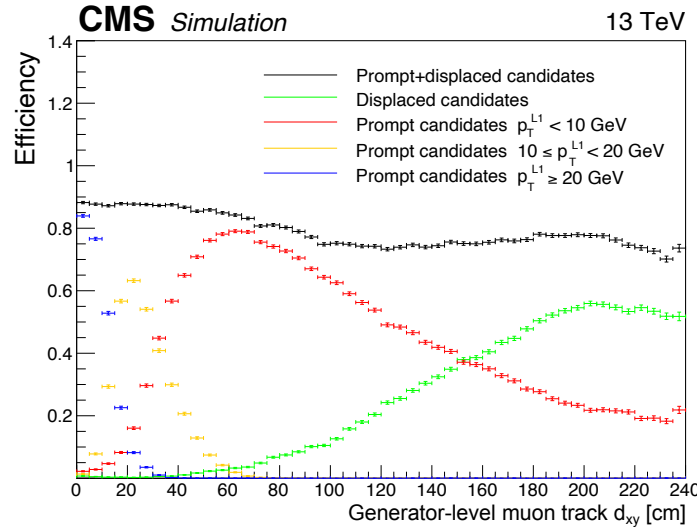
Figure 123. The OMTF trigger efficiencies for displaced and prompt algorithms with respect to muon track d_{xy} obtained using a displaced-muon gun sample. The efficiency curves are plotted for different values of the p_T estimate from the prompt algorithm (red, yellow, and blue), for the displaced algorithm (green), and for the combination (black). The prompt algorithm underestimates the p_T of displaced tracks, causing most of the tracks to have $p_T < 10$ GeV. The displaced algorithm can recover these tracks and improve the efficiencies to be around 80% for up to 200 cm displacements.

10.2.3 Endcap muon track finder (EMTF)

The EMTF builds muon tracks using TPs from CSCs, RPCs, and GEMs in both endcaps. The Run 3 EMTF algorithm includes GEM inputs for the first time. Already in Run 2, the algorithm used one TP per endcap muon station, with CSCs having priority over RPCs, to build muon tracks based on a set of predefined patterns. After the patterns were found, the differences in ϕ and θ angles between hits in different stations were used to estimate the p_T of the muon track using a boosted decision tree (BDT). The BDT was trained using simulated single-muon events and the output values of the BDT were stored in LUTs for fast evaluation. Since the beginning of Run 3, the EMTF algorithm also receives the new GEM TPs in station 1, as well as the updated CSC TPs in all stations, which provide better position and bending resolution. These TPs can be used both at the track building stage, as

well as the p_T estimation stage to improve the performance of the EMTF algorithm. Additionally, the EMTF uses a neural network (NN), implemented in FPGA logic, to estimate the p_T and d_{xy} of displaced-muon tracks, and forwards data to a hadronic shower trigger that uses the multiplicity of hits in the endcap CSCs to trigger on LLPs producing showers as they enter the endcap muon systems.

The first of the GEM detectors (GE1/1) is included in the EMTF algorithm in station 1 of the CMS endcap muon system. The Run 3 EMTF algorithm can use GE1/1 hits in conjunction with ME1/1 hits to improve prompt and displaced trigger efficiencies and reduce rates caused by mismeasured muons. Due to the placement of the GE1/1 and ME1/1 chambers in the CMS endcaps, the strong magnetic field in this region causes a larger bending of the muon track. This bending information between GE1/1 and ME1/1 is foreseen to be used to improve the p_T assignment for both prompt and displaced-muon tracks.

Similar to the other track finders, the EMTF also includes a new algorithm to improve displaced-muon triggering. The EMTF for Run 3 includes a NN-based p_T and d_{xy} assignment algorithm, which runs in parallel to the prompt algorithm. The NN has been directly incorporated into the EMTF firmware and estimates the p_T and d_{xy} of muon tracks that are built by the EMTF track building algorithm. The EMTF performance for prompt muons originating from the primary vertex remains identical. As shown in figure 124, the displaced-EMTF algorithm (NN-EMTF) shows an improved efficiency up to a d_{xy} of about 100 cm, while the prompt EMTF algorithm retains a high efficiency for prompt muon tracks up to about 25 cm. The expected performance for Run 3 was evaluated using a displaced-muon gun simulation sample with zero pileup. Zero-pileup samples were used since they are more useful in optimizing the NN and comparing ideal efficiencies. For L1 $p_T > 10$ GeV, the NN-EMTF algorithm shows efficiencies above 80% for $1.2 < |\eta| < 1.6$ and up to 100 cm displacements, while efficiencies for $1.6 < |\eta| < 2.1$ and $2.1 < |\eta| < 2.5$ at 60 cm displacement are around 20 and 5%, respectively (figure 124).

Finally, the EMTF for Run 3 forwards CSC hit information to provide a standalone method for triggering on hadronic showers occurring in the CSC detectors. An LLP decaying to hadronic particles within or slightly before the endcap muon systems can cause a shower of charged particles hitting the muon detectors, which are then recognized through a high hit multiplicity. In Run 3, the CSC detector sends information on whether a high multiplicity is found in any given chamber by comparing the measured multiplicity to a set of predetermined thresholds, individually for each CSC station and ring combination. The EMTF processes this information to decide whether there was a hadronic shower of a given quality in at least one of the CSC chambers in any given sector. The expected performance of this algorithm for Run 3 was evaluated using multiple physics simulation samples containing LLPs and found to provide efficiencies around 30%.

10.3 Global trigger

The hardware for the present global trigger system, the μGT, was installed as part of the Phase 1 upgrade, and was used for most of LHC Run 2 and is used for Run 3. The flexibility of the L1 trigger system has allowed for the addition of a μGT test crate containing the same hardware to be added to further extend the global trigger capabilities. It is used for testing and development purposes, for example, to test new or experimental trigger menus, or to test Phase 2 algorithms using ML or autoencoding, as described in ref. [88]. It receives the same optical inputs as the production crate from a passive optical splitter panel. For Run 3, this crate has been included as an optional component
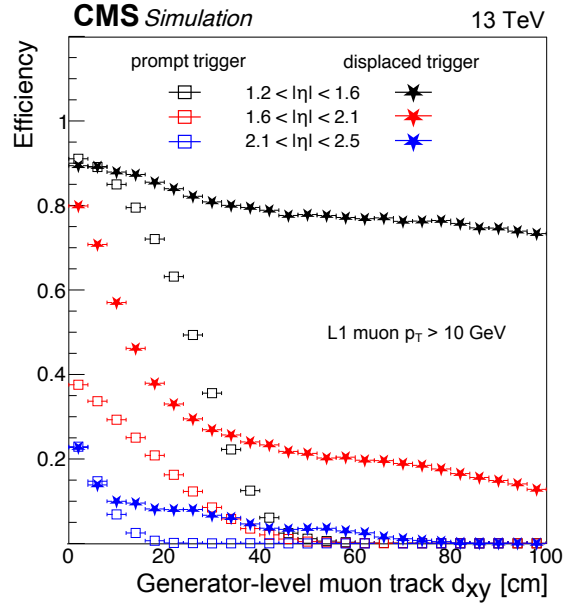
Figure 124. The EMTF trigger efficiencies for prompt and displaced-muon algorithms for L1 $p_T > 10\,\text{GeV}$ with respect to muon track d_{xy} obtained using a displaced-muon gun sample. The solid stars show displaced NN performance while hollow squares show the prompt BDT performance. The different colors show different η regions: $1.2 < |\eta| < 1.6$ (black), $1.6 < |\eta| < 2.1$ (red), and $2.1 < |\eta| < 2.5$ (blue).

into the CMS data acquisition system. While the test crate does not issue triggers itself, for events triggered by the production crate, the full test crate data (inputs and results of calculations) is available in the CMS data stream for offline analysis. The production and test crates are shown in figure 125.

By running the test crate with the same firmware, and thus the same trigger algorithms, as the production crate, it is possible to carry out consistency and hardware checks. More importantly, one can also run with different firmware in the test crate. This allows trigger developers to test new trigger menus with actual data. During normal data taking, a stream of zero-bias data is included, where the only trigger requirement is that the bunches for that crossing are filled. The zero-bias trigger has a high prescale applied, such that only 1 in N events are recorded, where N is the prescale value. This makes it possible to investigate the performance of new triggers in the test crate menu. A test crate algorithm that is more restrictive than a specific algorithm in the production system can also be studied by using the latter algorithm as a reference, thus benefiting from a larger data sample. The fact that the test crate can be included in data acquisition during normal running at no additional cost means that large amounts of realistic data can be collected, thus allowing for accurate tests even of very restrictive algorithms that only rarely provide a trigger. By normalizing the offline data to the number of zero-bias triggers taken or by recording the online monitoring data, one can also check the total trigger rate of a new menu. This is very useful, since it is not trivial to accurately estimate the total menu rate offline using the rate of individual algorithms, since multiple algorithms can fire the same event simultaneously. This is important in order to guarantee that when running with the new menu the L1 trigger rate remains within the total bandwidth, so that deadtime is minimized.

The upgraded monitoring backend is configured to collect and store data from the test crate in order to provide integration with the central monitoring services. Trigger menu rates are stored in the cen-
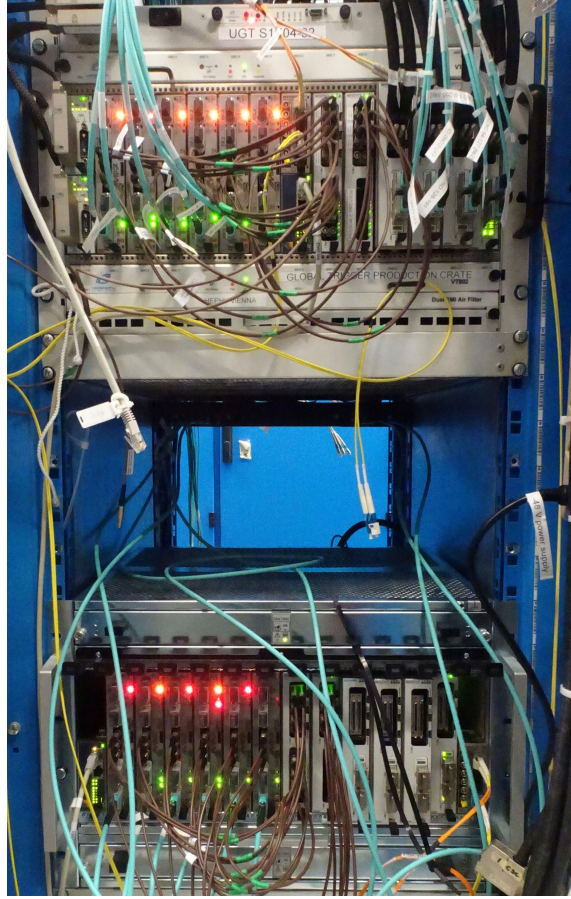
Figure 125. Production (upper) and the new test crate (lower) of the μGT.

tral Prometheus monitoring database, described in section 10.5, together with those from the production crate for online monitoring and prompt offline studies. Data are also expected to be sent to offline computing systems for long-term storage. Alerting services are configured to warn shifters and experts with relaxed severity compared to the production system, as the system is not critical for data taking.

10.4 Trigger menu

Expanding the CMS physics reach significantly beyond that explored already during Run 2 [5] requires important changes to the Run 3 trigger menu, and relies mostly on improvements to the L1 trigger algorithms. The LHC beam conditions and filling scheme, as well as other aspects like luminosity leveling, described in more detail in section 7.5, should be taken into account in order for the menu to be efficient and robust within the L1 bandwidth limit of about 100 kHz.

The updates to the calorimeter and muon trigger systems discussed earlier aim to retain the Run 2 physics coverage described in ref. [5], while providing additional access to signatures that do not originate at the primary vertex, such as displaced muons or displaced jets. Furthermore, the GT is able to perform two new kinematic computations that can be utilized to increase the physics coverage: the three-body invariant mass and the di-object ratio between invariant mass and the $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$ between the objects. Special L1 trigger menus are available in addition to

the standard menus to account for various data-taking scenarios, as was done during Run 2, e.g., targeting signatures related to the physics of bottom quarks using the so-called "B parking" [288], recording additional data for B physics and other studies, as further described in section 11.5.

10.4.1 Trigger seeds for displaced muons

Physics signatures involving LLPs were not well represented within the L1 trigger menu during Run 2, since specific algorithms capable of triggering on displaced objects within the detector were not yet implemented. The changes for Run 3 provide the ability to trigger on displaced muons with higher efficiency compared to Run 2. Updates to the muon algorithms in all three muon track finders provide unconstrained p_T and d_{xy} measurements for displaced muons, as discussed in section 10.2.

To benefit from the displaced muon algorithms, the L1 trigger menu of Run 3 extends that of Run 2 by offering seeds that use the unconstrained p_T and d_{xy} measurements. These new seed features are available for muon objects with $p_T > 10\,\text{GeV}$, and possible additional selections are defined according to the planned physics program. The chosen p_T threshold is based on the expected gain in the trigger efficiency described in section 10.2, concerning only muons with $p_T > 10\,\text{GeV}$ and displacements larger than approximately 20 cm. Since the displaced muon algorithms do not perform better compared to the prompt algorithms in the case of low p_T muons, no new seed is created if the leading muon in the considered seed has a p_T below 10 GeV.

10.4.2 Trigger seeds using new kinematic variables

Low-mass resonance searches, for example those used to study the physics of bottom quarks, are based on targeting a final state with low p_T objects. While the final states targeted in B-physics searches are often expected in the barrel region, many other interesting physics scenarios predict a wide pseudorapidity distribution for the final state objects. Ideal seeds with a minimum object selection with low object thresholds would result in a high trigger rate that is unsustainable, exceeding the available trigger bandwidth under any LHC data taking conditions, and provide low purity.

During Run 2, the most used unprescaled seeds relying on a single object had relatively high p_T selections to keep trigger rates manageable, e.g., the single-muon trigger with $p_T > 22\,\text{GeV}$. Special B parking triggers, used for events that are written to disk storage without full online event reconstruction at HLT, were developed to lower the L1 trigger thresholds while keeping within the HLT trigger bandwidth. To provide reasonable rates, typical B parking seeds in general restrict pseudorapidity to the barrel and overlap regions, e.g., the double-muon trigger with no p_T selection but with $|\eta| < 1.5$ and $\Delta R_{\mu\mu} < 1.4$. However, the HLT thresholds used for the B parking seeds, generally above 5 GeV, would often significantly reduce the total acceptance of interesting physics signals.

The substantial L1 trigger rates related to low p_T trigger thresholds can alternatively be managed by using kinematic variables that are optimized to increase the acceptance of predicted signal events. During Run 3, the GT is able to perform two new kinematic variable computations: the three-body invariant mass, and the di-object ratio between invariant mass and the ΔR. These allow the rate of seeds with low p_T thresholds to be reduced. For example, the three-body invariant mass can be harnessed to target the decays of a τ into three muons, a final state with low-p_T final objects with wide pseudorapidity distributions. Alternatively, the di-object ratio between invariant mass and the ΔR could benefit a low-mass dimuon resonance search providing a dark photon interpretation.

10.4.3 Run 3 trigger rates

The baseline L1 trigger menu in Run 3 is identical to that of Run 2 detailed in ref. [5]. Since the beam and detector conditions, and upgraded TPs are different from those during Run 2, the baseline trigger menu has been reviewed using Run 3 simulation samples to understand if slight modifications to the existing seeds are necessary to respect the total rate budget of 100 kHz. Two preliminary rate studies have been performed by reweighting the pileup distribution in Run 3 simulation samples, which has an unrealistic flat shape between 30 and 80. Firstly, by reweighting the Run 3 MC simulation samples to the Run 2 pileup distribution, it was confirmed that the rates using Run 3 simulations with updated TPs provide the same rate as obtained from Run 2 data, for the same pileup conditions. Secondly, the expected rates for realistic Run 3 pileup conditions were estimated by reweighting using a pileup distribution determined from the expected instantaneous luminosity during luminosity leveling. This approach takes into account various beam conditions, such as the bunch-to-bunch variation in pileup, which can have a significant impact on the rates. The rate allocation for single- and multi-object triggers and cross triggers is shown in figure 126. These results can be compared to a similar figure for the L1 trigger rate allocation under the Run 2 conditions, provided in ref. [5]. Additional seeds, relying on the aforementioned new features, are implemented on top of the baseline L1 trigger menu. This menu is then tuned according to the desired physics program of CMS for Run 3.
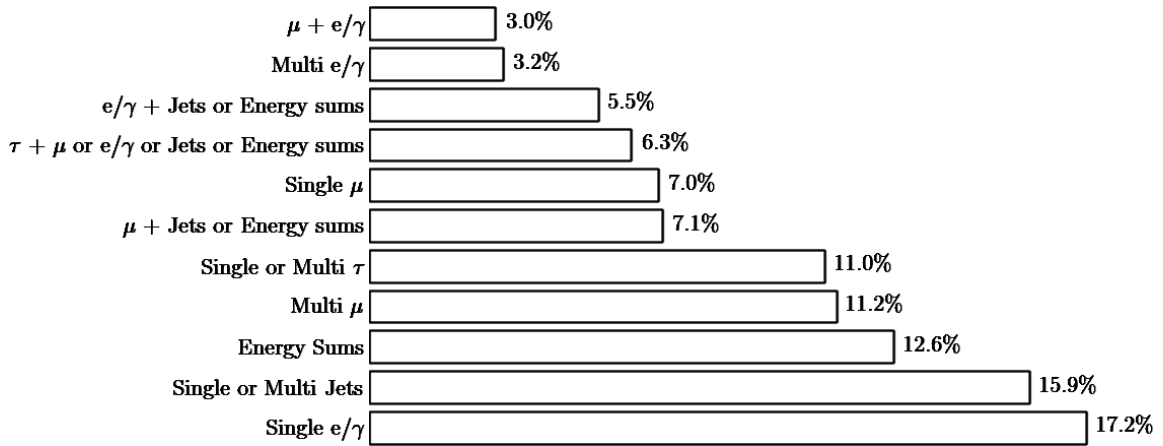
Figure 126. Fractions of the 100 kHz rate allocation for single- and multi-object triggers and cross triggers in the baseline Run 3 menu, calculated using Run 3 Monte Carlo simulation samples of inclusive pp events with appropriate pileup.

10.5 Online software and monitoring

During LS2, a number of improvements to the development workflow for the L1 trigger online software have been used. Practices from the well-established methodologies DevOps [289] and Agile [290] have been introduced in the Run 3 online software to simplify the development and deployment of upgrades and enable continuous improvement. The online software is built in centrally-provided Docker images that provide a replicable development and test environment. Every online software project employs continuous integration (CI) pipelines in GitLab to automatically build and test new commits. In order to minimise maintenance efforts and facilitate the development

of CI pipelines, an in-house Auto-DevOps platform was developed, taking inspiration from the GitLab Auto-DevOps project. Auto-DevOps is employed to centrally provide a set of configurable and generic CI jobs that represent typical tasks in the online software development, such as building RPM package manager packages and Docker images. Feedback from introducing standard CI pipelines in all software projects was excellent: given the large number of software components and teams that the online software consists of, this procedure established standard workflows that greatly increased the safety of new deployments. New automated tests are constantly developed and integrated into the CI whenever new incidents occur, reinforcing the importance of feedback from operations to minimise disruptions of service. Work is ongoing to develop a testing cluster managed by RedHat OpenShift running on CERN computing services where test deployments can be performed from CI pipelines. The goal is to provide a complete environment where software checks on the entire infrastructure can be run before deploying on the CMS computing resources.

The Run 3 monitoring backend is based on industry-standard tools and is centered around a Prometheus monitoring database instance collecting monitoring data from all trigger subsystems [291]. The database is expected to collect around 300 000 metrics every 20 s and correlate monitoring information coming from different sources to provide a powerful monitoring and alerting system. Alerts are handled by the Prometheus Alertmanager service [291]. The software enables binding alerts to external software based on conditions to page shifters and escalate to experts if necessary. Actions to automatically mitigate or solve problems can also be configured to minimise downtime. Monitoring dashboards are built using Grafana and can provide powerful inspection tools for online and prompt post-run analysis (figure 127) [292]. Prometheus is not designed for providing long-term storage of metrics, therefore integration with offline computing systems is under consideration.
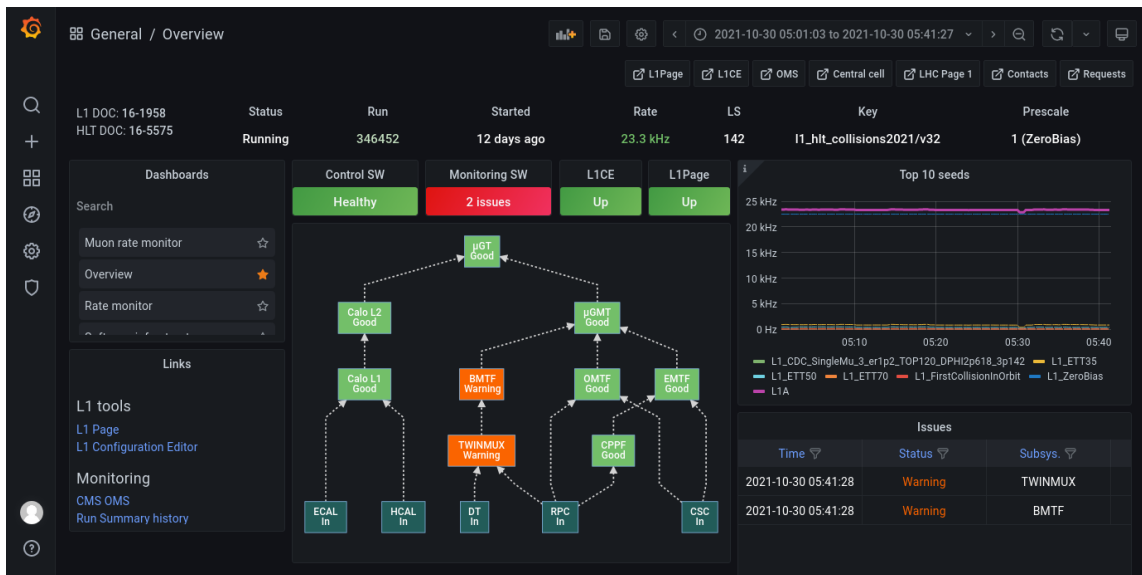
Figure 127. Screenshot of the Grafana L1 trigger monitoring dashboard.

The Run 3 trigger configuration editor and main monitoring page are designed in the Vue.js JavaScript framework, which superseded the outdated and slower Polymer framework. The Run 3 main monitoring page has also been updated to present data taken from the updated monitoring backend.

The significantly improved L1 online software and monitoring enables users to spot problems much more efficiently, and react more quickly compared to Run 2. It also facilitates more efficient diagnosis of problems after they have occurred, since more historical information is available and is easy to access.

10.6 The L1 scouting system

For Run 3, the L1 trigger project added dedicated hardware for the triggerless recording of objects reconstructed in the level-1 trigger, at a rate of 40 MHz without trigger or filtering. The data are received by dedicated FPGA-based processing cards, either housed in powerful servers (I/O nodes) or operating standalone in custom enclosures. Once received and pre-processed, the information is provided to a computing farm via Ethernet. Triggerless data recording, referred to in CMS as "40 MHz scouting" or "L1 scouting", facilitates improved precision of measurements such as the luminosity, as well as unprecedented levels of trigger monitoring. The term "scouting" is also used in the context of HLT, described in section 11.4, where it refers to the selection and high-rate recording of small-size analysis data sets.

The L1 scouting system deployed for Run 3 is meant as the first large-scale demonstration of a larger system planned for the Phase 2 L1 trigger upgrade [88]. Besides providing a testing ground for the technical implementation of the readout, it provides the opportunity to study solutions for large-scale distributed processing of high-rate data, the correlation of multi-bunch-crossing signals, as well as first studies of possible physics applications using limited resolution trigger data. The measurement of a range of physics processes, such as LLPs with displaced muons and flavor anomalies in τ physics, could potentially benefit from L1 scouting.

10.6.1 Architecture of the L1 scouting system

The L1 scouting system receives data from the level-1 trigger via spare output links and processes them quasi-online in a dedicated computing farm. The system operates largely independently from the standard CMS trigger and data acquisition chain. In the first test system in Run 2, the data sent from the μGMT to the μGT was duplicated, and final muon objects, as well as intermediate muon candidates derived from the BMTF inputs, were transmitted over eight 10 Gb/s optical links to the L1 scouting system.

For Run 3, an additional set of duplicated μGMT outputs supply a L1 scouting processor dedicated to luminosity monitoring, as described in section 10.6.2. Each of the twelve BMTF processors dedicates two 10 Gb/s links to L1 scouting information. L1 scouting data from the calorimeter layer 2 system also mirrors the trigger objects provided to the μGT. Each μGT processor transmits 512 bits per bunch crossing over three 10 Gb/s links, indicating which of the trigger algorithms has fired in a given bunch crossing.

The architecture of the initial Run 3 L1 scouting system is a scaled up version of the Run 2 system, and consists of I/O nodes housing FPGA-based input boards that receive up to eight 10 Gb/s links using the L1 trigger link protocol. Data are transmitted unidirectionally with no back-pressure to the trigger. The L1 scouting system therefore does not interfere with the standard trigger system in any way. The FPGA logic performs both zero suppression and preprocessing of the data such as reformatting or recalibration. The use of fast NN algorithms, implemented within FPGA resources to improve recalibration performance has been demonstrated (see section 10.6.2). Data are transferred

by direct memory access (DMA) via a Gen-3 PCIe x16 bus into the memory of the I/O nodes, from where they are sent to dedicated processing units in the surface data center via 100 Gb/s Ethernet over coarse wavelength division multiplexing CWDM4 single mode optical infrastructure. Even after full zero-suppression, the long-term storage of the huge amount of raw data produced by the trigger processors, in view of a subsequent "classic" multitiered offline analysis and reduction, does not represent a viable approach. Data taken in the early months of LHC Run 3 are being used to investigate various methods for a real-time analysis.

In the Run 2 demonstrator system, a KCU1500 Xilinx development kit, equipped with a KU15P FPGA, was used to capture the μGMT inputs. Additional I/O nodes for Run 3 are equipped with more powerful boards, such as the Micron SB-852, using a large Xilinx Ultrascale+ VU9P FPGA and also providing access to the Micron Deep Learning Accelerator (MDLA) [293], which is a proprietary compiler that translates pre-trained ML networks into instructions for an FPGA-based hardware implementation. Monitoring of the readout board is being developed using the AXI-lite interface provided by the Xilinx xDMA core. Custom software exposes access to monitoring and control registers via a RESTful interface and monitoring and diagnostic data are exported to the Prometheus server described in section 10.5.

In addition to the I/O nodes with dedicated receiver boards, Xilinx VCU128 development kits [294] are used as standalone receivers for L1 scouting data. These boards, which are housed in a custom enclosure with PCIe extender buses for control and monitoring, are equipped with a VU37P FPGA with 8 GB of high bandwidth memory. The VU35P, with a very similar architecture, is planned for use on the DAQ-800 board currently being designed for the CMS Phase 2 data acquisition [295], and is the anticipated readout board for the L1 scouting system of the Phase 2 upgrade. The VCU128 boards in Run 3 thus allow a realistic test of the Phase 2 design, albeit with reduced input bandwidth. Equipped with an additional FMC mezzanine to provide up to 32 10 Gb/s input links, they transmit L1 scouting data directly over TCP/IP to the surface computing farm, providing more efficient utilization of the bandwidth available on the 100 Gb/s Ethernet links. The Run 3 scouting architecture is illustrated in figure 128.

10.6.2 Applications of the L1 scouting system

The use of ML algorithms to improve the physics potential of L1 trigger objects captured by the triggerless recording of trigger data at 40 MHz has been investigated. An example of this is the use of deep NNs, optimized for a throughput of around 1 MHz, to recalibrate L1 muon objects in real time, such that the accuracy of the p_{T}, η, and ϕ parameters of the muons can be improved over the standard L1 trigger reconstruction, which is optimized for efficiency at threshold, not for a full physics analysis. Neural networks have been trained on zero-bias triggered data from LHC Run 2, to predict the offline fully reconstructed muon parameters of matched L1 muon objects. This recalibration has been shown to improve the accuracy of the L1 muon parameters when compared to the standard L1 reconstruction, both for μGMT and BMTF L1 muon parameters [296, 297]. Additionally, neural networks that are trained to reject pairs of L1 muon candidates that do not correspond to matched fully reconstructed muons, and networks for anomaly detection have also been implemented in the Xilinx VU9P FPGA with a throughput of about 1 MHz, using the MDLA.

A copy of the μGMT scouting system will be used to provide direct per-bunch measurements of the L1 muon multiplicity to the CMS BRIL system described in section 8. FPGA firmware
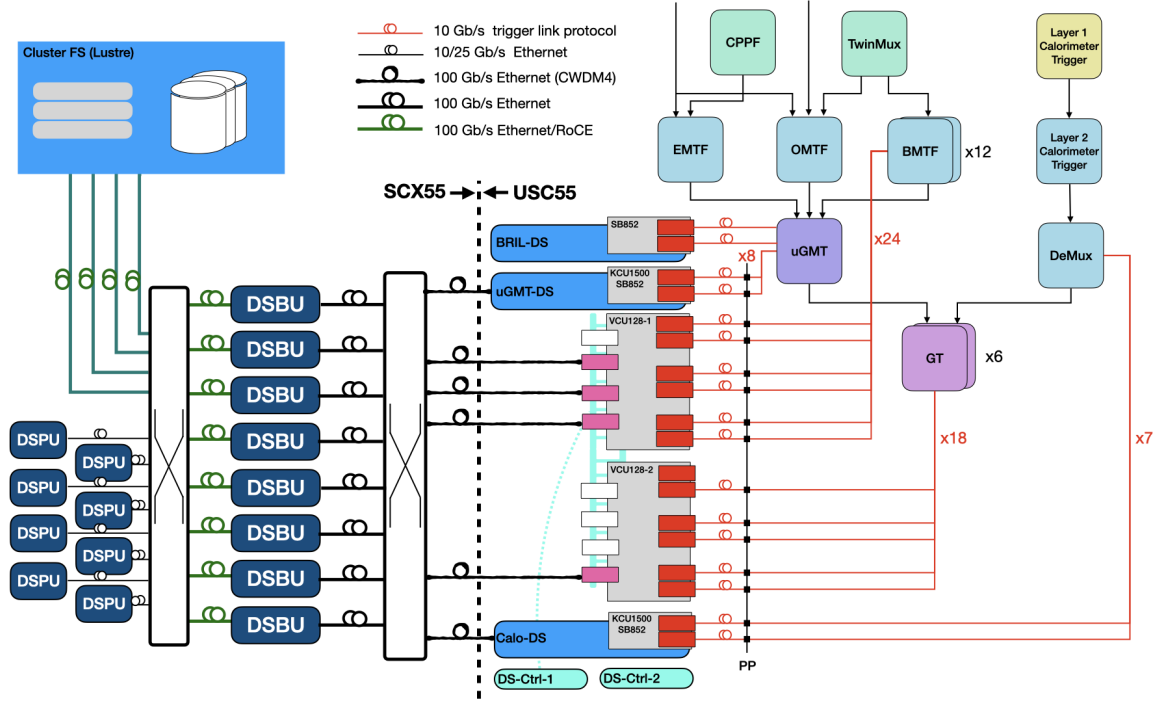
Figure 128. Architecture of the Run 3 L1 scouting prototype system. The input system, located in the experiment service cavern, consists of a combination of different types of FPGA receiver boards, hosted in the PCIe bus of I/O server nodes or extender. The boards receive and pre-process data from the different trigger systems. Two of the boards (KCU1500 and SB852) use DMA to move their data to host memory, from where they are transmitted to the surface data center over 100 Gb/s Ethernet links. The VCU128 implements a TCP/IP core in the FPGA and directly trasmits data to the surface system. In the surface data center, links from the L1 scouting input system are connected to a switched network. Data streams are received through the said network by L1 scouting buffer servers (DSBU) and buffered in files on large RAMdisks. The L1 scouting processing units (DSPU) access buffered data from the DSBUs to perform data reduction and analysis. The processed data are finally moved to a Lustre cluster file system for long-term storage.

has been developed and implemented in both the KCU1500 and SB-852 boards to histogram the muon multiplicity per bunch over a period of four lumi nibbles, corresponding to 1.458 s. These histograms are read out to the host PC in real time via the PCIe DMA implementation.

Cosmic muon data taken with the L1 scouting system in special runs during LS2 have been used to analyse the performance of the new kBMTF algorithm. The L1 track finding algorithms assume that muons always originate from the interaction region. Therefore, a cosmic muon traversing the full detector appears as two back-to-back muons, usually separated in time by one or two bunch crossings. The data were taken when the magnet was off. In the absence of a magnetic field, muons traverse the detector in a straight line, leaving two muon tracks in the barrel drift tubes, typically two bunch-crossings apart (ΔBX = 2). Figures 129–130 show the relationship between the L1 reconstructed impact parameters, d_{xy}, and the angle ($\phi_{\text{in}} - \phi_{\text{out}}$) of the two corresponding L1 tracks.

The Run 3 demonstrator system was used to systematically collect and store data during collision runs throughout 2022 from the μGMT and calorimeter layer 2. The offline analysis of these data provided invaluable insight into the reliability of the acquisition system and in particular of the hardware,
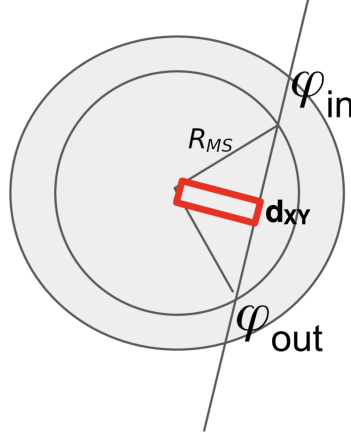
Figure 129. Illustration of the correlation between the impact parameter d_{xy}, highlighted in red, and the difference between the angles measured for the incoming and outgoing legs of a cosmic muon.
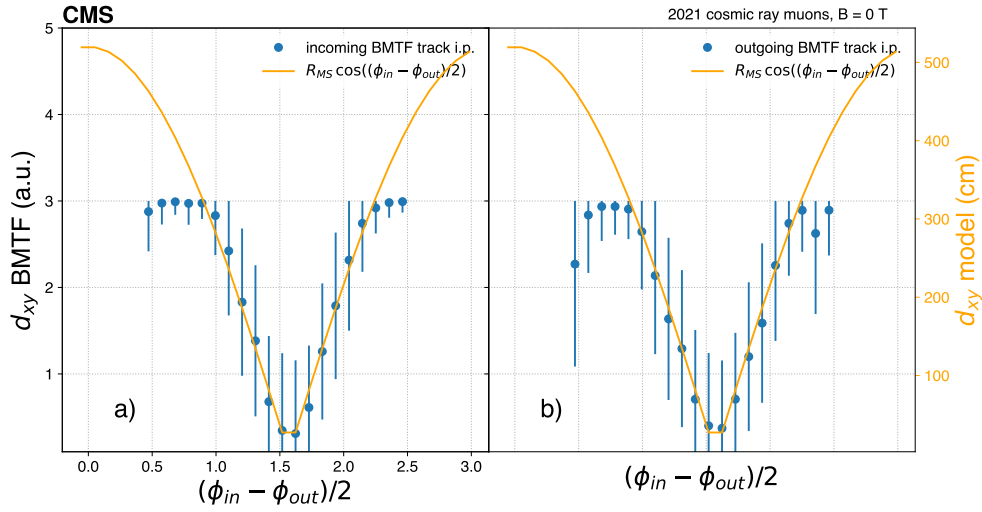
Figure 130. The impact parameter d_{xy} of the incoming (left) and outgoing (right) cosmic ray muon tracks as measured by the BMTF as a function of the difference of azimuthal coordinates of the incoming and outgoing legs. The BMTF firmware encodes the impact parameter in two bits, hence the range of values on the left y axis. The orange curves model this dependence of the actual impact parameter as $R_{MS} \cos(\phi_{\text{in}} - \phi_{\text{out}})/2$, where R_{MS} is the radius at which the BMTF measures the ϕ coordinate of the track. The right hand side y axis shows the d_{xy} values (in cm) as predicted by this model, which exhibits remarkable consistency with the measurement (for the values within the range), if one assumes one unit of the left axis to correspond to an impact parameter of about 100 cm. Reprinted from [298], Copyright (2023), with permission from Elsevier.

firmware and software. The results of this work have been and are being used to guide the design of further extensions of the demonstrator and the design of the Phase 2 systems. The quality of the data was studied in view of assessing their usability for trigger diagnostics, luminosity measurement, and physics studies. This work will continue throughout the rest of Run 3, along with the completion of the demonstrator with the commissioning of the VCU128 acquisition boards, collecting data from the uGT and BMTF systems, which represents a major step forward in validating the Phase 2 design.

11 High-level trigger

11.1 Overview

To select events of potential physics interest, the CMS trigger system divides the processing into two levels: the level-1 (L1) trigger, implemented in custom hardware as described earlier in section 10, and the high-level trigger (HLT), implemented in software, running on a farm of commercial computers. The HLT refines the purity of the physics objects selected by the L1 trigger, with a maximum input rate increased to 110 kHz in 2023 from 100 kHz earlier. In 2022, for standard pp collisions at an average instantaneous luminosity of $1.5 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$, the average rate to offline storage for promptly reconstructed physics events was approximately 1.7 kHz. In 2023, for a peak luminosity of $2.0 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$, the rate of promptly reconstructed physics events was about 2.6 kHz. Additional data streams for calibration purposes or "HLT data scouting" are stored at higher rates with a smaller event content. The HLT data scouting differs from the L1 scouting (described in section 10.6) in that events must still satisfy a subset of L1 triggers before being written to disk at a high rate.

The HLT runs on a cluster of 200 nodes, each equipped with two AMD EPYC "Milan" 7763 CPUs, two NVIDIA T4 GPUs, and 256 GB of memory, running Red Hat Enterprise Linux 8. The HLT farm is described in section 9.5, with the details specific to Run 3 highlighted in section 9.5.5.

The HLT data processing uses the concept of "paths" to structure its workflow. These paths are sequences of algorithmic steps designed to reconstruct physics objects and make selections based on specific physics requirements. Steps within a path are typically organized in ascending order of complexity, reconstruction refinement, and physics sophistication. For example, the resource-intensive track reconstruction process is usually carried out after completing a series of initial reconstruction and selection steps involving the data from the calorimeters and muon detectors.

The reconstruction modules and selection filters of the HLT use the same software framework that is also used for offline simulation, reconstruction, and analysis (CMSSW [256]). As noted in section 9.5.3, HLT paths selecting similar physics object topologies are grouped into primary data sets for subsequent offline processing, and collections of primary data sets are organized into streams for efficient handling.

In preparation for Run 3, the HLT software was adapted to make use of heterogeneous computing architectures, and several reconstruction modules were developed to take advantage of that to meet the challenges of processing data at ever increasing luminosity and pileup. Algorithms implemented to run on both CPUs and GPUs, are automatically directed to run on a GPU if a GPU is available; otherwise, the CPU-based version of the algorithm is executed. The Patatrack project [299] has created parallelized versions of pixel track and vertex reconstruction algorithms that can run on an NVIDIA GPU and were written using the NVIDIA CUDA language. The data structures are optimized for GPU, and the entire reconstruction chain is executed on the GPU to minimize time-consuming data transformations and transfers. A subset of ECAL and HCAL local reconstruction algorithms have also been ported to GPU, also using CUDA. Based on these efforts a reduction in overall event processing time of about 40% has been achieved. More details can be found in section 11.3. To make full use of the gain, CMS deployed and commissioned a filter farm composed of nodes comprising two GPUs in addition to two CPUs, as noted in section 9.5. More information about multithreaded processing and GPU offloading is given in section 12.5.

The HLT selects data for storage through the application of a trigger "menu", in which the collection of individual HLT paths is configured. The trigger path definitions, physics object thresholds, and rate allocations are set to meet the physics objectives of the experiment. In 2022, the HLT menus for pp data taking typically contained around 600 paths. This includes the primary HLT paths for analysis as well as paths for calibration and efficiency measurements that are typically looser than the primary paths. These latter HLT paths are often "prescaled", i.e., only a fraction of the events that pass the requirements are actually accepted, to limit processing time and storage rate. Different trigger menus are used for the recording of heavy-ion collision data. The rates, physics breakdown, and CPU timing of the pp menu are described further in section 11.3.

11.2 HLT reconstruction

The HLT paths in the menu depend on the modules that produce the physics objects from the all-silicon inner tracker and from the crystal electromagnetic and brass-scintillator hadron calorimeters, operating inside a 3.8 T superconducting solenoid, together with data from the gas-ionization muon detectors embedded in the flux-return yoke outside the solenoid. A foundation to many of the specific object reconstructions is the particle-flow algorithm [300], which uses information from these systems to identify candidates for charged and neutral hadrons, electrons, photons, and muons. The main features of the HLT physics object reconstruction and improvements for Run 3 are described in the following subsections.

11.2.1 Tracking

Tracking using the hits recorded by the pixel and strip trackers is generally performed iteratively using a combinatorial Kalman filter, starting with tight requirements for the track seeds that become looser for each subsequent iteration. Hits in the tracking detectors that have already been used in a track are removed at the beginning of the next iteration. For Run 2, initially, the track reconstruction in the HLT consisted of three iterations. The first two iterations required four consecutive hits in the pixel detector to seed the tracking. These iterations target first higher p_T tracks and then lower p_T ones, and use the full volume of the pixel detector. The third iteration relaxes the requirement on the number of hits in the pixel detector to three, and is restricted to the vicinity of jet candidates identified from calorimeter information and the tracks reconstructed in the two previous iterations. The track reconstruction is limited to tracks that are consistent with the leading vertices reconstructed with the pixel detector (those vertices with the largest summed p_T^2 of pixel tracks).

In 2017, several issues with the Phase 1 pixel detector were identified that led to a nonnegligible fraction of inactive pixel modules (section 3.1.1). During the 2017–2018 year-end technical stop, the performance of the pixel detector was restored. Nevertheless, an additional recovery iteration was added to the tracking, to safeguard against a recurrence of this or possible other detector failures. In particular, track seeds consisting of just two pixel hits were allowed to be reconstructed in regions of the detector where two inactive modules overlap.

For Run 3 the seeding and tracking were significantly revised. Tracking is now performed using only a single global iteration, and is seeded by a loose selection of the pixel tracks reconstructed by the Patatrack algorithm, which offers improved performance over the four-hit pixel tracking used for data taking in 2018 [299]. To be used as track seeds, Patatrack pixel tracks must satisfy a loose selection requiring three or more pixel hits, $p_T > 0.3$ GeV, and be compatible with a primary

vertex candidate. Despite fewer iterations and less CPU time required, the Run 3 tracking has improved performance over that used for Run 2. The tracking efficiency and fake rate are measured in simulated $t\bar{t}$ events with an average pileup of 63. Figure 131 shows the Run 3 tracking efficiency and fake rate as determined from Monte Carlo (MC) simulation. The tracking efficiency is defined as the fraction of simulated particles from the signal interaction within the considered p_T and η regions, and longitudinal (transverse) impact parameters <35 (70) cm that are matched to a reconstructed track. The fake rate is defined as the fraction of reconstructed tracks that could not be matched to a simulated particle. The Run 3 tracking efficiency is higher than that of Run 2 for $p_T > 0.7$ GeV, mostly in the central tracking region and the overall fake rate is lower, particularly around the transition region between the barrel and the endcap pixel detectors ($0.9 < |\eta| < 2.1$).

11.2.2 Muons

Tracking algorithms are also deployed to identify and reconstruct muons measured in the muon detectors and in combination with the pixel and strip trackers. While the algorithms used during Run 2 are described in more detail in ref. [301], a brief summary is given here along with the changes implemented for Run 3.

Muon track reconstruction at the HLT takes place in two steps: first using hits only in the muon system (L2 reconstruction), followed by a combination with hits in the inner tracking system (L3 reconstruction). The L2 reconstruction is equivalent to the offline standalone muon reconstruction. The L3 reconstruction is seeded by an L2 muon and follows an iterative track reconstruction similar to that described in the previous section in a region around the seed starting from the outer tracking layers and working inward ("outside-in") or from the inner tracking layers working out ("inside-out"). The inside-out approach can also be seeded directly by L1 trigger muons. The L3 reconstruction is essentially 100% efficient with respect to that in the L1 trigger, while consuming only about 30% of the overall HLT CPU time. After the muon track reconstruction, identification criteria are applied, as well as isolation criteria for the isolated muon category. The isolation is based on the sum of p_T from additional tracks associated with the primary vertex and calorimeter energy deposits that are clustered using a particle-flow algorithm in a cone of radius $\Delta R = 0.3$ around the muon. The estimated contribution from pileup to the energy deposits in the calorimeter is subtracted.

For Run 3, several modifications to the muon reconstruction were made to improve the HLT performance, in particular with respect to CPU timing. The L2 muon reconstruction was extended to include hits from the GEM detectors, described in section 6.4. At L3, the tracking was adapted to make use of the Patatrack pixel track seeds, followed by a single iteration with the full tracker. The efficiency for low p_T muons was improved by optimizing the search regions around seeds in which muon tracks are reconstructed in the tracker, achieving an efficiency above 80% for $p_T > 2$ GeV. Additionally, machine-learning algorithms were incorporated: a boosted decision tree (BDT) classifier for the search algorithm and a neural network (NN) algorithm for seeding strategy. The inside-out algorithm uses the BDT classifier to consider only the seeds with high quality. The NN algorithm was developed to choose the best seeding strategy for the outside-in reconstruction from the L2 muon and thus limit the total number of seeds considered. With these changes, the L3 efficiency remains unchanged, while the CPU time was reduced by 15%.
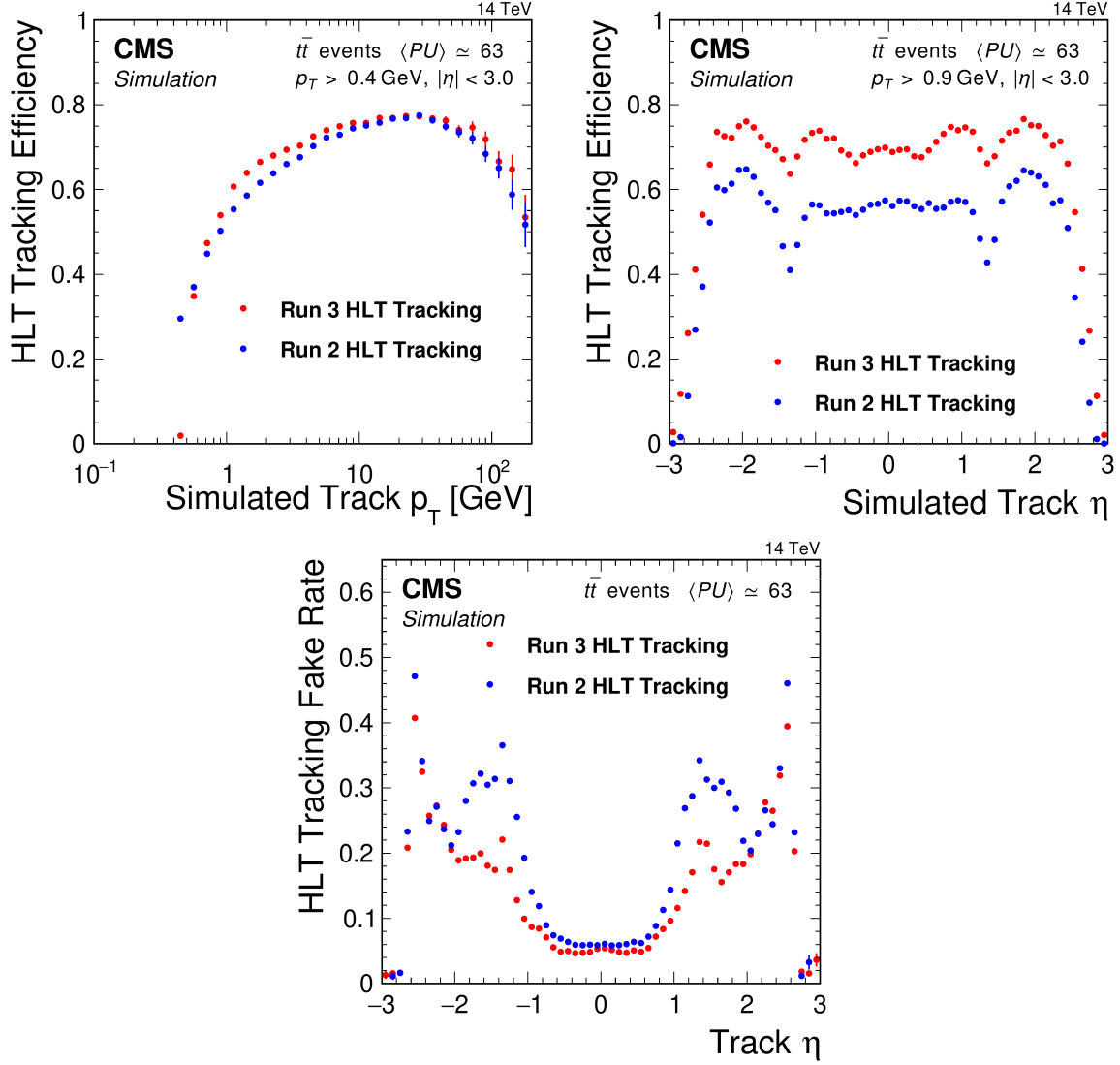
Figure 131. Tracking efficiency for the Run 2 HLT tracking (blue) and the Run 3 HLT single-iteration tracking (red) as a function of the simulated track p_T (upper left) and track η (upper right). Only simulated tracks with $|\eta| < 3.0$ are considered in the efficiency measurement, with $p_T > 0.4$ (0.9) GeV required for the upper left (right) plots. The tracking fake rate (lower) is shown as a function of the reconstructed track η for the Run 2 HLT tracking (blue) and the Run 3 HLT single-iteration tracking (red).

11.2.3 Electrons and photons

The HLT electron and photon identification uses the L1 calorimeter trigger candidates as a starting point to perform a "regional" reconstruction of the energies deposited in the ECAL crystals around them. Subsequently, superclusters (clusters of ECAL deposits within a certain geometric area around the seed cluster) are built using the same reconstruction algorithm as used offline [7]. The energy correction applied to HLT superclusters is simpler than the one used offline in that it uses ECAL information only. Requirements are then imposed on the minimal energy, as well as other properties of the energy deposits in the ECAL and HCAL subdetectors.

For electrons, the ECAL supercluster is associated to a reconstructed track with a direction compatible with its location. The first step is a match with pixel detector hits. Since 2017, the pixel matching algorithm requires three pixel hits, to maximize early background rejection, while a hit doublet is accepted only if the trajectory passes through a maximum of three active modules. If the supercluster is successfully matched with the pixel hit seeds, the electron track is reconstructed using a Gaussian sum filter (GSF) tracking algorithm [302].

Variables to enhance the identification of true electrons and photons are applied based on the shower shape in the ECAL, the energy deposition in the HCAL, and, in the case of electrons, the matching between the track and the ECAL supercluster, as well as the quality of the GSF track. As with the other leptons, isolation criteria are generally applied, except for some specific paths, to electrons and photons based on the calorimeter energy deposits in a cone of radius $\Delta R = 0.3$ around the electron or photon and the sum of p_T from additional tracks associated with the primary vertex. Several HLT paths with different isolation criteria and p_T thresholds are defined to provide a range of efficiencies and rates for specific physics analyses.

11.2.4 Tau leptons

The reconstruction of hadronic tau-lepton decays (τ_h) at the HLT is also of crucial importance for the physics program. During Run 2, it was performed in three steps. The first step, the L2 reconstruction, is seeded by L1 τ_h candidates. The energy depositions in the calorimeter towers around the candidates within a cone of radius 0.8 are clustered, and L2 τ_h candidates are reconstructed by using the anti-k_T algorithm [303, 304] with a distance parameter of 0.2.

In the second step, known as L2.5, a charged particle isolation criterion, based on pixel detector information, is implemented. Pixel tracks are reconstructed around L2 τ_h candidates with $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$ in a region of $\Delta\eta \times \Delta\phi = 0.5 \times 0.5$. Tracks originating from the primary vertex with a transverse impact parameter $d_{xy} < 0.2\,\mathrm{cm}$, at least three hits, and a trajectory in an isolation cone of $0.15 < \Delta R < 0.4$ around an L2 τ_h candidate, are considered for the isolation sum. An L2 τ_h candidate is considered isolated if the scalar sum of the p_T of the associated pixel tracks is less than 4.5 GeV.

The final step, the L3 reconstruction, includes track reconstruction using the full tracker. For Run 2, tracking used a reduced number of iterations to fit into CPU time budget. Moreover, the track reconstruction was performed regionally around the L2 τ_h candidates. Until mid 2018, the L3 reconstruction was performed using a cone-based algorithm. It was then upgraded to the hadrons-plus-strips (HPS) algorithm [305] that is also used in offline reconstruction. Both algorithms start with jets reconstructed by the anti-k_T algorithm with a distance parameter of 0.4.

In Run 3, the L2 and L2.5 sequences were replaced by a convolutional neural network, in which the pixel tracks from the Patatrack algorithm and the calorimeter candidates are used as input. This improved the rejection rate at L2 by about a factor of 2 for similar efficiency. Additionally, the efficiency of the L3 reconstruction was increased by introducing a neural network, DeepTau [306], adapted from the offline reconstruction such that it matched HLT requirements for speed and performance.

11.2.5 Jets and global energy sums

Jets are reconstructed at the HLT using the anti-k_T clustering algorithm [303, 304] with a nominal distance parameter of 0.4, or of 0.8 in the case of wide jets used in boosted topologies and multijet triggers. Inputs to the jet algorithm are usually particle-flow candidates, or alternatively calorimeter

towers. Depending on detector and beam conditions, corrections are applied to the jet energy scale, as well as the measured particle-flow hadron and average pileup energies.

Triggers using jets and global energy sums are used in CMS across a wide spectrum of physics analyses. Multijet trigger paths, for example, are key to select vector boson fusion event candidates which contain two very forward jets in opposite endcaps, with a large angular separation and a large dijet invariant mass. Final states with boosted multijet signatures can also be identified using dedicated jet substructure techniques such as the soft drop approach [307]. Signatures with many jets in the final state can also be triggered using H_T, the transverse energy sum of all jets, or S_T that combines jets with leptons.

Paths based on missing transverse momentum $\vec{p}_T^{\mathrm{miss}}$, defined as the negative vector sum of the p_T of input objects, also exist. For these paths, the accounting for noise and beam-induced backgrounds is especially important in order to keep rates and resolutions under control. Trigger paths based on p_T^{miss} alone, or in combination with jets, leptons, or photons in the event, are also used, e.g., to search for weakly interacting particles.

11.2.6 b jet tagging

The identification of b jets at the HLT is essential in order to enhance the fraction of events containing heavy flavor jets from processes like vector-boson associated Higgs boson production where the Higgs boson decays into a pair of b quarks. Such processes would otherwise be unlikely to pass the standard thresholds for leptons, jets, or missing transverse momentum.

Since b tagging relies on the measurement of tracks that are displaced with respect to the primary vertex, both the pixel and the silicon strip trackers are used to improve the spatial and momentum resolutions of such tracks. In 2016, the combined secondary vertex algorithm CSVv2 was used. Subsequently, in 2017 and 2018, the multiclassifier neural network DeepCSV was implemented [308]. With DeepCSV, the b jet tagging efficiency was improved by 5–15% at constant gluon or light-quark misidentification rates.

For Run 3, two new neural network taggers, DeepJet [309] and ParticleNet [310], were deployed in 2022, with further improved performance. In addition to tracks, the DeepJet algorithm also uses information from neutral and charged particle-flow jet constituents. The ParticleNet algorithm provides multiclass jet-flavor classification for categories of b, c, and light quarks, gluons, and hadronically decaying tau leptons. For use in HLT, the Run 3 tagging algorithms were trained on dedicated HLT-reconstructed simulation samples.

Figure 132 shows the light-flavor jet misidentification rate versus the b jet efficiency for the different tagging algorithms, evaluated on simulated top-quark pair production events with an HLT jet selection of $p_T > 30\,\mathrm{GeV}$ and $|\eta| < 2.5$. Compared to the performance of the DeepCSV algorithm used during Run 2, the DeepJet algorithm trained using HLT quantities has a light-flavor jet misidentification rate that is lower by about a factor of 3 (up to efficiencies of about 75%). The ParticleNet algorithm reduces the misidentification rate by another factor of 2.5.

11.2.7 New HLT paths for long-lived particles

In addition to the improved reconstruction algorithms discussed in section 11.2, the Run 3 HLT menu has been significantly expanded to explore new and unconventional physics signatures. The Run 3 HLT menu includes new dedicated triggers targeting long-lived particle (LLP) signatures, such as
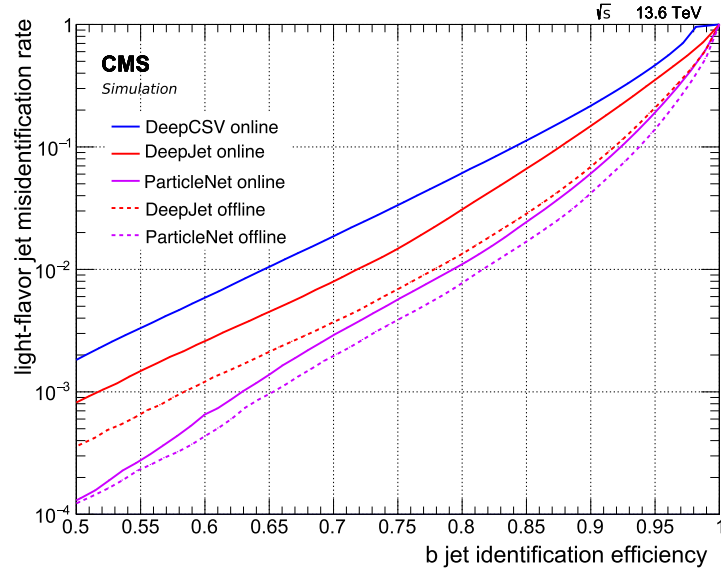
Figure 132. Light-flavor jet misidentification rate versus the b jet efficiency for the various b tagging algorithms. The solid curves show the performance of the DeepCSV (blue), DeepJet (red), and ParticleNet (magenta) algorithms in the HLT. The dashed curves show the corresponding offline performance for DeepJet (red) and ParticleNet (magenta) taggers using offline reconstruction and training.

displaced jets, displaced leptons and photons, and delayed jets using timing information from ECAL and HCAL, as well as HCAL depth information, as detailed in section 5.3. Furthermore, it features high-multiplicity trigger (HMT) paths that trigger on hadronic showers in the muon system, and new triggers for B physics, as described in section 11.5. In 2023, further resources were deployed to enhance the sensitivity for demanding event topologies, such as the production of Higgs boson pairs decaying with at least two b jets, vector boson fusion (VBF) processes, and events with LLP signatures. Several of these new HLT paths are also seeded by new dedicated L1 trigger candidates (section 10.4), further improving the signal acceptance for their target searches. The new triggers for B and VBF physics will be described in more detail in section 11.5.

There are now several flavors of displaced and delayed jet paths available at the HLT in Run 3. First, there is a suite of displaced dijet paths that were already available to a certain extent in Run 2, but that have gone through major improvements over time. These displaced dijet paths are either inclusive, in that they select events with calorimeter $H_T > 650$ GeV and two jets with less than two prompt tracks each, or they are more exclusive and require events with calorimeter $H_T > 430$ GeV and two jets with less than two prompt tracks and at least one displaced track. For Run 3, the displaced jet selections have been improved at the HLT. In particular, the selection on the number of prompt tracks was tightened in order to reduce the rate, while at the same time the definitions of prompt and displaced tracks were loosened in order to improve the signal efficiency for low-mass LLPs. Furthermore, additional L1 trigger seeds have been added. An L1 seed that requires a single muon in addition to a small amount of H_T helps to reduce the displaced dijet H_T threshold at the HLT. In addition, the new HCAL timing and depth seeds, described in section 5.3, bring improved efficiency for LLPs with $c\tau > 0.5$ m. All of these improvements to the displaced dijet paths provide better efficiency to trigger on low-mass LLPs, especially those with heavy-flavor decays.

In addition to delayed jet paths that use HCAL timing, there are also new HLT paths that exploit the ECAL timing. For LLPs that produce jets with delays of about 1 ns or more, the signal efficiency is improved by an order of magnitude, with respect to the MET triggers that were available for this analysis in Run 2. In particular, there are two different kinds of delayed jet triggers that use ECAL timing at the HLT; there are paths that are seeded by H_T, and there are paths that are seeded by L1 τ objects. For both seeds, different requirements are made at the HLT, namely one or two jets, independently of whether those jets are trackless or not, and the amount of timing delay. The paths seeded by H_T improve the sensitivity to low-mass LLPs with respect to the MET paths, and the paths that are seeded by L1 τ objects increase the efficiency to trigger on Higgs boson decays to long-lived scalars that decay to four b jets as well as to four τ leptons.

Neutral LLPs with particularly long lifetimes could decay hadronically beyond the calorimeters, creating a high-multiplicity shower in the muon system. Such showers are expected to consist of hundreds of hits, but no tracks or jets reconstructed in the inner detectors. Essentially, the CMS muon system would act as a sampling calorimeter. As mentioned in section 10.2.3, new L1 seeds have been developed to collect these high-multiplicity events in the CSCs. The high-multiplicity triggers (HMTs) at L1 are used to seed several HLT paths. At the HLT, a clustering of hits in the muon system is performed using the Cambridge-Aachen algorithm [311, 312]. The first HMT HLT path reconstructs a single CSC cluster, with stricter cluster requirements than at L1 in order to control the rate. The second HMT HLT path reconstructs a CSC cluster as at L1, and then additionally requires a cluster in the DTs with at least 50 hits. The last available HMT HLT path reconstructs a single DT cluster with 50 hits, makes no requirements on CSC clusters, and is seeded by MET triggers at the L1 trigger. As compared with the MET triggers that were available in Run 2, the trigger efficiency for these unique signals is improved by factors of 3 to 20 depending on the path.

Paths for displaced muons at the HLT have also been improved in Run 3. Displaced dimuon paths are seeded by two L1 muons with low p_T thresholds and by new displaced kBMTF double muon seeds with unconstrained p_T and d_{xy}, as described in section 10.2.1. These seeds feed into several types of displaced dimuon paths at the HLT. There are L2 double-muon paths that require an in-time collision based on the beam pickup timing device with a veto on prompt muons, complemented by paths that make use of a seed developed for cosmic ray muons. These two paths require displacements of at least 1 cm. Lastly, there are L3 double-muon paths that require displacements of at least $100\,\mu$m. This suite of HLT paths covers a wide range of displacements and improves the signal efficiency over that of Run 2. The efficiency, measured in a cosmic ray muon sample recorded in 2022, was measured to be 100% for displacements $d_{xy} < 100$ cm for the L2 pp seed + prompt-veto path, and 90% for $d_{xy} < 350$ cm for the L2 cosmic seed + prompt veto path. The efficiency of the L3 displaced dimuon path is measured to be 85% in the data sample of non-prompt J/ψ events. At the same time, the background efficiency is small: it is measured to be <1% in Drell-Yan data events for all three displaced dimuon HLT paths.

11.3 Run 3 HLT menu composition, rates, and timing

While the complete list of HLT paths in the Run 3 HLT menu is too long to be listed here, a representative sample of some standard triggers with their HLT thresholds and rates is provided in table 15.

Table 15. HLT thresholds and rates of some generic triggers in the Run 3 HLT menu. The rates were obtained from measurements during an LHC fill in November 2022 and have been scaled to a luminosity of $2.0 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

HLT algorithm	Rate
Isolated muon with $p_\mathrm{T} > 24\,\mathrm{GeV}$	250 Hz
Isolated electron with $E_\mathrm{T} > 32\,\mathrm{GeV}$	182 Hz
Particle-flow based $p_\mathrm{T}^\mathrm{miss} > 110\,\mathrm{GeV}$	81 Hz
4 PF jets with $p_\mathrm{T} > 70, 50, 40$, and $35\,\mathrm{GeV}$ with two b tags	57 Hz
Two isolated tau leptons with $p_\mathrm{T} > 35\,\mathrm{GeV}$	54 Hz
Muon with $p_\mathrm{T} > 50\,\mathrm{GeV}$	51 Hz
Two electrons with $E_\mathrm{T} > 25\,\mathrm{GeV}$	21 Hz
AK4 PF jet with $p_\mathrm{T} > 500\,\mathrm{GeV}$	16 Hz
Two same-sign muons with $p_\mathrm{T} > 18$ and $9\,\mathrm{GeV}$	10 Hz
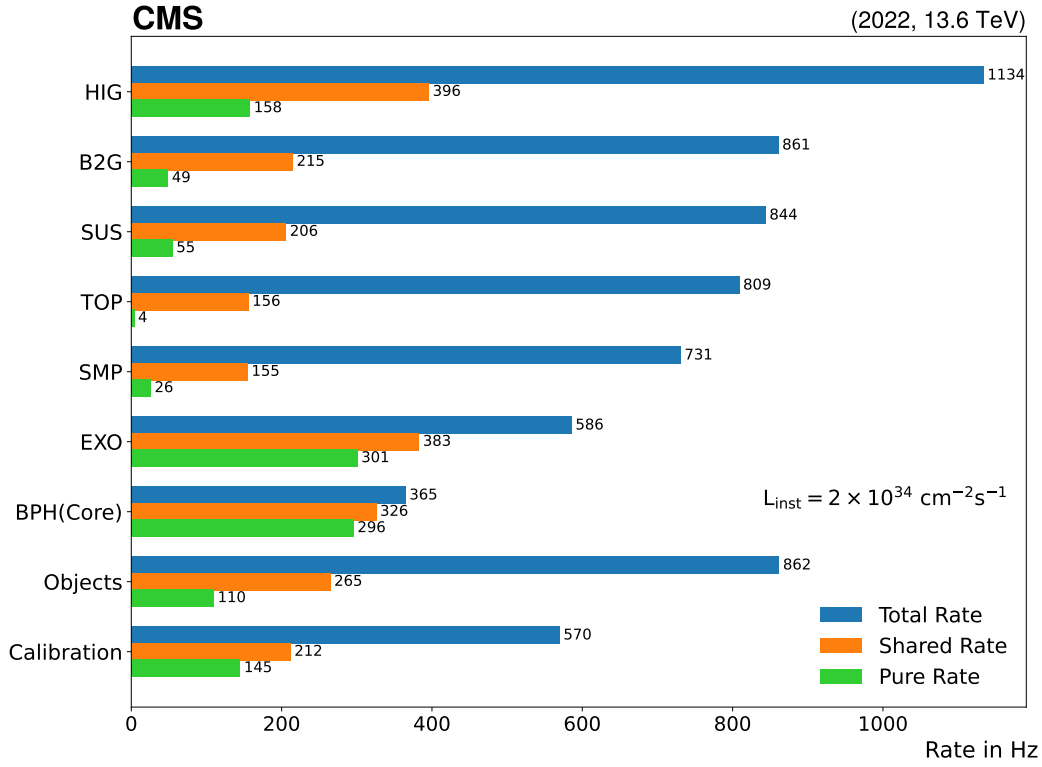
Figure 133. The HLT rate allocation by physics group for the Run 3 menu deployed in November 2022, scaled to a luminosity of $2.0 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The total rate (blue bar) is the inclusive rate of all triggers used by a physics group and the pure rate (green bar) is the exclusive rate of all triggers unique to that group. The shared rate (orange bar) is the rate calculated by dividing the rate of each trigger equally among all physics groups that use it, before summing the total group rate.

Figure 133 shows the "standard physics" HLT rates consumed by each CMS physics group, estimated from a fraction of events recorded by CMS from a fill taken in November 2022. "Standard physics" refers to the collection of triggers included in the primary data sets whose reconstruction starts within 48 hours after the data were recorded. This does not include data scouting at the HLT or parking triggers, as detailed in sections 11.4–11.5. The average delivered instantaneous luminosity during this fill was $1.8 \times 10^{34} \, \mathrm{cm^{-2} \, s^{-1}}$, but the measured rates have been scaled to correspond to a luminosity of $2.0 \times 10^{34} \, \mathrm{cm^{-2} \, s^{-1}}$. For the rate measurement, events are assigned to a physics group if the group uses at least one of the HLT algorithms that triggered the event. The group categories correspond to the physics analysis working groups: Higgs boson physics (HIG), searches for new physics in boosted signatures (B2G), searches for new physics in final states with imbalanced transverse momentum (SUS), top quark physics (TOP), standard model physics (SMP), B physics (BPH), and searches for exotica (EXO). The calibration category corresponds to all HLT algorithms used for subdetector alignment and calibration purposes. The "objects" category corresponds to HLT algorithms used for monitoring and calibration by the so-called physics object groups.

What can be seen from figure 133 is that roughly one third of the menu rate has been devoted to standard model physics processes (HIG, SMP, TOP), one third to searches for physics processes beyond the standard model (EXO, SUS, B2G), and the remaining one third to B physics processes, physics objects (for monitoring and calibration purposes for example), and subdetector calibration. The trigger selection for B physics and LLP (included in EXO) are largely unique to those groups.

The distribution of HLT time spent processing the data is shown in figure 134. The processing time running only on CPUs (left) is compared to that when part of the reconstruction is offloaded to GPUs (right). These results were obtained for an HLT configuration representative of the 2022 conditions, running over a sample of 64 000 pp collision events with an average pileup of 56 collisions. The measurements were performed on a machine identical to those used in the HLT farm, as described in section 9.4, equipped with 2× AMD EPYC Milan 7763 CPUs and 2× NVIDIA T4 GPUs. The node was configured identically to the HLT farm, with simultaneous multithreading (SMT) enabled, NVIDIA multiprocess server (MPS) enabled, and running eight jobs in parallel with 32 CPU threads and 24 concurrent events each. The average processing time per event is 690 ms when running only on CPUs and 384 ms when offloading part of the reconstruction to GPUs, corresponding to a speedup of over 40%. The maximum processing time per event for the initial Run 3 event filter farm configuration is 500 ms, as noted in section 9.5.5.

11.4 Data scouting at the HLT

A limiting factor for the data-acquisition rate is the bandwidth of the data to record on disk (a few GB/s), not the event rate per se. Thus, if the size of the data per event is reduced, a higher rate of events can be recorded, i.e., using significantly lower trigger thresholds. In the so-called "HLT data scouting", only the most relevant physics information is stored, as reconstructed by the HLT, and not the complete set of raw data. Scouting was first implemented in Run 1 [313], and was developed further during Run 2, for selected physics objects, such as jets [314] and dimuons [315]. For instance, the H_T threshold in the HLT scouting data was reduced from 800 to 410 GeV, and this data was used for a search for three-jet resonances [316].

A further benefit of HLT data scouting is that events are reconstructed only once, using the resources of the HLT, and thus do not require further computing resources to perform the offline
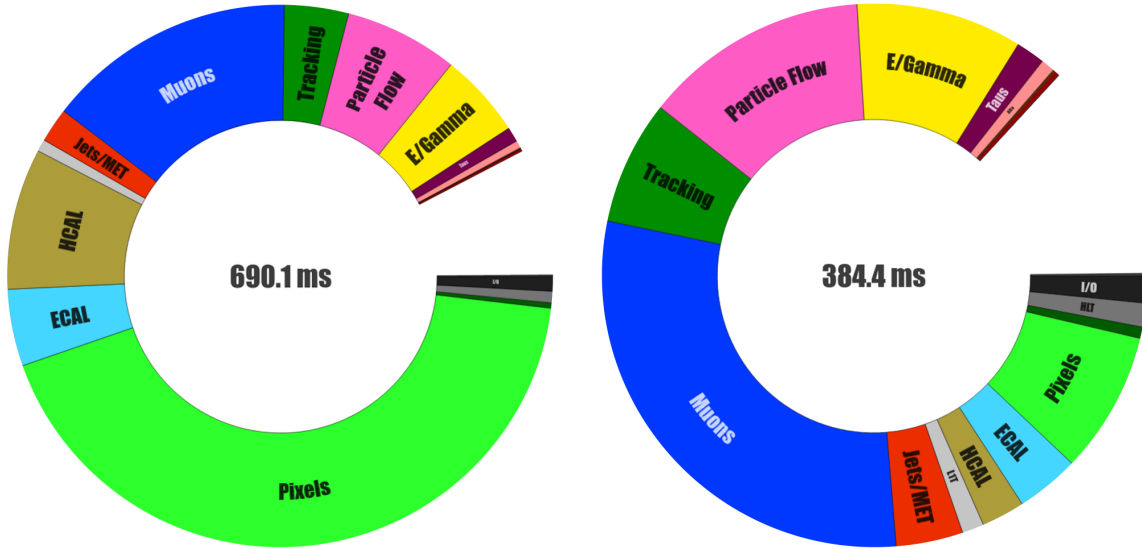
Figure 134. Pie chart distributions of the processing time for the HLT reconstruction running only on CPUs (left) and offloading part of the reconstruction to GPUs (right). The slices represent the time spent in different physics object or detector reconstruction modules. The empty slice indicates the time spent outside of the individual algorithms.

reconstruction step, which is described in section 12. On the other hand, a complete reprocessing of HLT scouting data not possible. The quality of the scouting data depends on the calibrations and alignments used in the HLT, but these must anyway be precise to maintain good trigger performance. Furthermore, the CPU processing time budget per event is limited at the HLT (less than about 300 ms per CPU core during Run 2) and scouting requires extra processing, which must fit in the constraints.

During the beginning of Run 3, HLT scouting data was recorded with a rate of up to 30 (22) kHz for the 2022 (2023) data-taking periods, respectively, and an event size of about 13 kB, compared to the full raw data event size of about 1 MB. Events are reconstructed in the HLT scouting scheme if they are accepted by an array of L1 triggers targeting one or two electrons, muons or jets or a moderate amount of H_{T} with thresholds lower than for other HLT paths. For Run 3, a special version of the particle-flow reconstruction algorithm using pixel tracks reconstructed with Patatrack (section 11.2.1) was deployed. This special version allowed offloading to GPUs at the expense of a slightly degraded parameter resolution in comparison to tracks reconstructed using the full tracker information. The reduced reconstruction time made it possible to use particle-flow reconstruction on a larger fraction of input events.

In addition to the objects from the particle-flow reconstruction [300] that were already stored during Run 2 (muons, jets, and particle-flow candidates), HLT scouting in Run 3 includes the reconstruction and storage of electrons, photons, and tracks. At the beginning of the 2023 data-taking period, the electron reconstruction was further optimized by loosening the L1 seeding requirement, resulting in an increased efficiency for low p_{T} electron reconstruction. No event selection is applied after the reconstruction. Reconstructed objects are stored if they fulfill relatively loose criteria, such as $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 3$ for jets, and $p_{\mathrm{T}} > 0.6\,\mathrm{GeV}$ and $|\eta| < 3$ for particle-flow candidates. In addition to kinematic quantities like p_{T} and η, other information to facilitate offline analyses is also

stored. In the case of charged particle-flow candidates, this includes parameters of associated tracks to facilitate, e.g., the training of jet tagging using machine-learning algorithms. Muons are stored not only with track parameters but also with information about the hits in the tracker, allowing the possibility to refit the dimuon vertices.

Since HLT scouting is able to access low-momentum objects with higher rates than conventional HLT trigger paths, it is well suited for analyses targeting low momenta and low-mass particles. Current studies include, for example, analyses of low-mass dimuons, diphotons, and dielectrons as well as a $H \rightarrow b\bar{b}$ analysis that benefits from the decreased threshold on H_T.

11.5 Data parking

One limiting factor for the HLT output rate is the bandwidth of the prompt event reconstruction at Tier 0. An alternative approach to increase the amount of data available for physics analysis is to increase the storage rate on disk, while delaying the reconstruction of the data until a later time, when the necessary computing resources are available. The reconstruction can be scheduled during a year-end technical stop or a long shutdown, for instance. This concept, known as "data parking", was already implemented in Run 1 [6] and Run 2 to record additional data for B physics and other studies. During an LHC fill, as the luminosity decreases, the bandwidth to trigger additional events increases, and the trigger thresholds for data-parking events are gradually relaxed to record parking data.

In 2018, for example, the collection of $b\bar{b}$ events was enhanced by tagging and storing events containing at least one displaced muon, e.g., from a semileptonic B decay. The p_T threshold at the HLT for a single isolated muon was 24 GeV for the standard physics menu. For the parked B data the p_T threshold was as low as 7 GeV and the HLT output rate reached 5 kHz at the end of fills, enabling CMS to accumulate about 10^{10} b hadrons [288]. In Run 3, data parking still targets B physics, but it also includes a rich set of other physics data. As of the end of 2022, the parking streams record events with at least one muon candidate with $p_T > 12$ GeV and a transverse impact parameter significance larger than 6; events with two muons with $p_T > 4$ and 3 GeV with an invariant mass less than 8.5 GeV; and events with two electrons with $|\eta| < 1.22$, $p_T < 4$ GeV, and an invariant mass less than 6 GeV. With the addition of the parking streams, the total HLT reaches peak output rates of 6 kHz. This can be seen in figure 135, which shows (separately) the HLT output rates for promptly reconstructed events and for parked data for an LHC fill recorded in 2023 with a peak levelled luminosity of about $2 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$.

In 2023 the parking strategy was extended to improve the signal acceptance for critical Higgs boson measurements and searches. By dropping the single-muon parking approach, which was limited to instantaneous luminosities lower than $1.7 \times 10^{34} \, \text{cm}^{-2} \, \text{s}^{-1}$, and by improving the purity of the dielectron triggers, CMS is now able to dedicate bandwidth for events targeting final states with two b-tagged jets, the VBF production mechanism, and LLP signatures. The two b-tagged jet criterion, mainly designed for the production and decay of HH into four b quarks, relies on the presence of four jets with $p_T > 30$ GeV, two loose b-tagged jets using the ParticleNet tagger, and an aggressive threshold $H_T > 280$ GeV. Figure 136 (left) shows the trigger efficiencies for prompt and parking data. A clear improvement of more than 20% is observed with respect to the trigger efficiency of Run 2. The measured efficiency in a single-muon dataset recorded in 2023 confirms a plateau efficiency of more than 90% (figure 136, right).
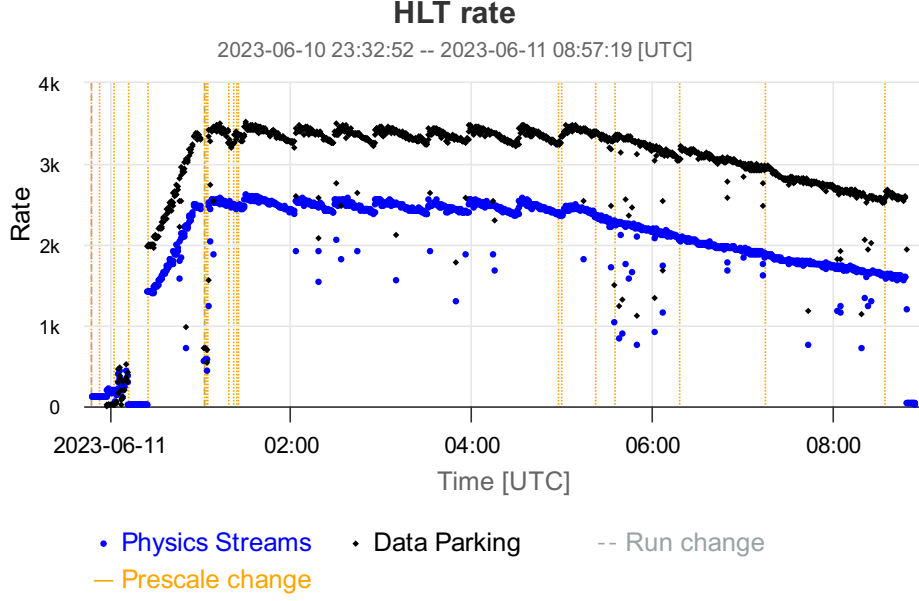
Figure 135. The HLT rates for promptly reconstructed data streams (blue) and parked data (black) as a function of time during an LHC fill in 2023.
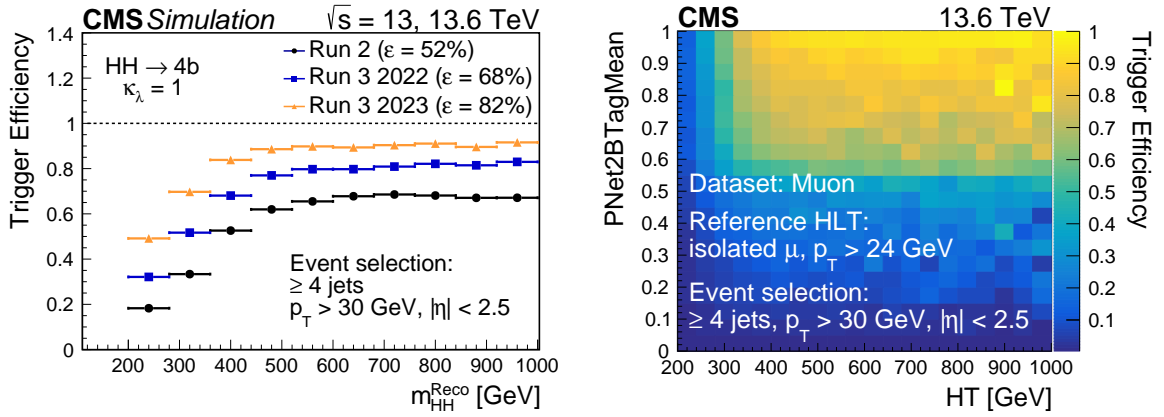
Figure 136. Left: comparison of the trigger efficiency of the HH $\rightarrow$ b$\bar{\text{b}}$ trigger among the three different strategies used in Run 2 (black), 2022 (blue), and 2023 (orange) using the signal MC sample. Right: trigger efficiency of the HH $\rightarrow$ b$\bar{\text{b}}$ trigger using events collected by the single muon trigger in 2023.

For VBF event candidates a twofold strategy, one inclusive and one exclusive, is pursued. The inclusive approach relies only on the invariant mass of two forward jets and applies a tight selection of $m_{\text{jj}} > 1000$ GeV, while in the exclusive approach a looser threshold on the invariant mass is affordable because there are additional requirements on the central objects in the events. Table 16 summarizes the details of the two complementary approaches.

11.6 Heavy ion physics

Heavy-ion collisions impose unique challenges on the DAQ and HLT systems, as also noted in section 9. A completely custom HLT menu was developed with almost no overlap in content

Table 16. HLT thresholds and rates of the VBF triggers, as obtained from measurements during an LHC fill in June 2023 at an average instantaneous luminosity of $2.0 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, corresponding to a pileup of 61.

HLT algorithm	Rate [Hz]
2 jets with $p_\mathrm{T} > 105/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 1000\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 3.5$	720
2 jets with $p_\mathrm{T} > 70/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 600\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, 2 central jets $p_\mathrm{T} > 60\,\mathrm{GeV}$	430
2 jets with $p_\mathrm{T} > 105/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 1000\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 3.5$, 3 central jets	120
2 jets with $p_\mathrm{T} > 90/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 600\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, isolated muon $p_\mathrm{T} > 3\,\mathrm{GeV}$	110
2 jets with $p_\mathrm{T} > 75/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 500\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, $p_\mathrm{T}^\mathrm{miss} > 85\,\mathrm{GeV}$	110
2 jets with $p_\mathrm{T} > 45\,\mathrm{GeV}$, $m_\mathrm{jj} > 500\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, tau $p_\mathrm{T} > 45\,\mathrm{GeV}$	40
2 jets with $p_\mathrm{T} > 90/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 600\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, isol. μ $p_\mathrm{T} > 3\,\mathrm{GeV}$, 3 jets	12
2 jets with $p_\mathrm{T} > 45\,\mathrm{GeV}$, $m_\mathrm{jj} > 500\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, electron $p_\mathrm{T} > 12\,\mathrm{GeV}$	5
2 jets with $p_\mathrm{T} > 75/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 500\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, $p_\mathrm{T}^\mathrm{miss} > 85\,\mathrm{GeV}$, 3 jets	5
2 jets with $p_\mathrm{T} > 70/40\,\mathrm{GeV}$, $m_\mathrm{jj} > 600\,\mathrm{GeV}$, $\Delta\eta_\mathrm{jj} > 2.5$, 2 jets $p_\mathrm{T} > 60\,\mathrm{GeV}$	3

and paths with the pp menu. Additionally, because of the dense environment, the physics object reconstruction algorithms used by the HLT are generally customized for heavy-ion running. For example, while jet reconstruction remains based on the anti-k_T algorithm, the underlying event energy subtraction differs from that used to handle pileup in pp running.

In preparation for heavy-ion data taking in Run 3, significant effort is devoted to increasing the available L1 bandwidth. In 2018 up to 30 kHz was achieved, whereas we plan up to 50 kHz in Run 3 conditions. The most critical components to study include the ECAL, pixel, and tracker detectors. During a heavy ion test run in 2022, a scan of different ECAL readout settings was performed to study the readout size and potential impact on physics objects. A size reduction is crucial to achieve the target L1 rate. In order to fully utilize the L1 bandwidth, fractional prescale factors will be used to operate the trigger at the optimal point between rate and dead time.

The target for the PbPb collision run in 2023 is to record about 5×10^9 events. The HLT will be operated at an output rate above 10 kHz. About 10 kHz are "minimum bias" events, collision events with only a minimal L1 selection. The budget also contains 1–2 kHz of triggers of selected physics objects, such as muons, electrons, photons, jets, and track multiplicity conditions in central collisions. A significant fraction of the latter events will also be contained in the minimum bias data set.

To use the bandwidth between the HLT and the Tier 0 center in an optimal way, all triggers are collected into a single "HIPhysics" data set. Only later, during the offline processing step, which is described in section 12, will they be split into secondary data sets. Tools have been developed to automatically configure the splitting of large data sets into multiple outputs.

In addition to hadronic collision events, a suite of ultra-peripheral collision triggers will also be deployed. These target, for example, processes where the two ultrarelativistic nuclei do not collide directly, but their electromagnetic fields interact with each other. Such events are typically very clean, with only a few physics objects in the detector and no other visible activity. The size of the event data is minimal and contributes little to the bandwidth.

With the aging of the detector, the efficiency to trigger on some of the more peripheral collisions drops. Therefore a task force was formed to understand how the decrease in efficiency can be

avoided, including, for example, combining information with other detector parts like the Zero Degree Calorimeter. The heavy ion event size for hadronic collisions is comparable to that for pp collisions, and a sustained throughput of 17 GB/s is anticipated. Even though the bandwidth is much higher than that in Run 2, given the event size, the target number of recorded minimum bias events will be hard to achieve. To increase the event rate further, a new approach has been developed in which one of the most significant components of the event record, the tracker information, is removed. Instead, for each strip cluster, only summary information is written out. This approach is expected to reduce the event size by about 30%. The algorithm was commissioned during the 2022 heavy-ion test run, where both event format contents (the reduced format, "RawPrime", and the complete information) were written out.

12 Offline software and computing

12.1 Overview

CMS offline computing has evolved over the past 15 years to support the ever-growing needs to trigger, filter, store, transfer, calibrate, reconstruct, and analyze the recorded and simulated data of the experiment. The offline system receives a subset of the real-time detector information from the data acquisition system, once it is filtered by the high-level trigger at the experimental site, ensures safe curation of the raw data, and produces data for physics analysis. Major activities are also the production and distribution of Monte Carlo (MC) simulation data, as well as the processing of conditions and calibration information, and other nonevent data. The data input and output layer of the CMS data processing software is provided by the ROOT framework [259].

Key components of the offline computing system, described in the sections that follow, include an event data model and corresponding application framework, the processing chain and data tiers, computing centers, referred to as Tier 1, Tier 2, and Tier 3, which provide storage and processing resources all connected through a distributed world-wide computing grid, and a set of computing services that provide tools to transfer, locate, and process the data.

A timeline of the major data processing and computing software improvements put into production over the past decade or so is shown in figure 137. CMS developed a highly flexible computing model, with the goal of using efficiently all of the resources available to the experiment while minimizing both hardware and personnel needs and maximizing overall throughput. The innovations led to an increase in efficiency in the use of computing resources, and enabled the experiment to use additional new and diverse types of computing resources that were not available a decade ago.

12.2 Detector simulation

The CMS detector simulation is based on the GEANT4 [317] toolkit. It is augmented with computationally efficient techniques, such as shower libraries for the forward calorimeters, and specific identification criteria per particle type and detector region for neutrons [318], which guarantee high fidelity of the simulation. The "Full MC" simulation chain includes execution of the standard reconstruction chain. A subset of physics analyses, such as scans of new-physics signatures over large parameter space, use the "Fast MC" chain or "Fast Simulation" application. This does not rely on GEANT4, but rather a parameterized approach [319] with a simplified CMS geometry. The

	2010	Particle-flow adopted for reconstruction
64-bit binaries in production	2010	Particle-flow adopted for reconstruction
Invention and adoption of the VDT math library	2011	
	2012	First algorithm vectorised: vertexing
	2013	
Framework support for multithreading	2014	AAA and xRootD federation in production
HTCondor pools and GlideinWMS full adoption		Forward shower libraries for simulation
Dynamo DDM in production	2015	MiniAOD used for most analyses in production
Multicore GlideinWMS pilots		Russian roulette in production
Premixing for PU simulation in production	2016	Unified (e.g. automatic tape handling, block-by-block processing) in production
VecGeom in production	2017	First HLT algorithm running on GPU
Beginning of NanoAOD adoption	2018	Framework support for offload on accelerators
20% of the HLT ported to CUDA	2019	CMSSW offload to cloud accelerators demonstrated
First CMSSW portability library evaluation	2020	Rucio and CTA in production
Geant4 10.7 integrated	2021	DD4hep integrated

Figure 137. Timeline of the major data processing and computing software improvements put in production since 2010.

Fast MC chain also uses a faster version of the reconstruction step. The performance of the CMS simulation is detailed in ref. [320].

12.3 Event reconstruction

Event reconstruction is the processing step that transforms the event information contained in the raw data, i.e., packed detector readout data, into high-level physics objects, such as electrons, muons, photons, or jets, which are used in physics data analysis. In CMS, the event reconstruction is logically partitioned into several steps, starting from the local reconstruction in which the data are processed individually by a single detector component, and ending with global particle-flow reconstruction [300] and object identification. The particle-flow algorithm aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector.

A great deal of attention is dedicated to the computing performance of the event reconstruction, both for pp and heavy-ion collision data. The run-time optimization is a priority for CMS, and is achieved through algorithmic and technical improvements. Examples of technical improvements are the adoption of a more recent compiler version, utilization of special compilation flags and other code optimizations, without impact on physics performance. The timing is benchmarked continuously and in great detail in order to identify and remove performance degrading patterns as early as possible in the release integration process [321].

Run-time performance improvements must be paired with precise physics validation of the software, delivered swiftly to the code authors. Physics validation happens at multiple levels during the software integration process, mainly through automatic comparison of data quality monitoring (DQM) histograms of relevant distributions and physics quantities. Before new reconstruction code is merged into the central software repository, a limited number of events are processed and the

quality of the reconstruction is assessed. During twice-daily integration builds of the CMS software, a growing level of detail is added to the automatic validation, and a larger number of events are processed to obtain the DQM histograms.

Every major release, which may be intended for data processing or mass production of MC simulation samples, is preceded by weekly or biweekly pre-releases. The validation of pre-releases involves confirmation by detector and physics object experts who assess the quality of the results obtained with the software compared to actual Monte Carlo and data processing campaigns of a few million events. Final validation, referred to as release validation [322], is required before new CMS software releases are put into production.

12.4 Computer architectures and platforms

CMS strives to use efficiently all of the computing resources at the experiment's disposal. To this end, the data processing software supports a generic mechanism, described in section 12.5, to offload work onto accelerators such as GPUs, described in section 12.5.2.

Our builds support three different CPU architectures, x86_64, ARM, and IBM Power, driven by the opportunities to obtain allocations at HPC centers where these architectures are available (section 12.7.1). We perform integration and unit testing of the CMS code, and create regular releases of the whole CMS software suite CMSSW [256] for these architectures [323]. ARM support began in 2012 and Power in 2016. Installations of non-x86_64 releases are performed since 2014 on the file system used as a vector for CMS software, CVMFS, an aggressively cached distributed read-only file system [324, 325].

CMS not only builds, tests, and runs code regularly on different CPU architectures, but also employs two different compilers, GCC and Clang, as well as different operating systems. The combination of a CPU architecture, a compiler, and an operating system is commonly referred to as a "platform". The ability to build, test, and run CMSSW and perform integration tests on several platforms significantly contributes to achieving top code quality, a capability which CMS plans to preserve for the entire Run 3 and beyond. There is an increased risk of bugs and unstable algorithms silently altering the results of computations without being noticed if the code is tested and executed only on one platform.

12.5 Application framework

12.5.1 Multithreading

For LHC Run 2, the application memory usage was foreseen to increase beyond the 2 GB-per-CPU core limit of the computing grid worker nodes, due to increased pileup compared with Run 1. In order to reduce memory usage per CPU core, the CMS application framework was enhanced to support parallelism through a multithreading paradigm [256, 326]. The multithreaded framework follows a task-parallel paradigm implemented with Intel's oneTBB [327] library, which expresses concurrent units of work as tasks and passes them to oneTBB's task scheduler to run. The multithreaded framework was introduced for production jobs in 2015 for Tier 1 computing resources, and during 2016 for the majority of other resources. Currently, production jobs use eight threads by default, which is a good compromise between application memory usage and CPU efficiency. In addition, 8-core slots have been agreed as a standard job size on the shared worldwide LHC computing grid (WLCG) [328] resources.

The initial version of the multithreaded framework processed only separate events concurrently. The framework had to synchronize between the worker threads at specific stages of the data processing, and combined with large variability between the processing times of different collision events, this led to noticeable inefficiencies in the CPU utilization. The threading efficiency has been gradually improved [329], and currently many levels of concurrency are exploited. Within one event, independent modules are run simultaneously whenever possible according to their data dependencies. Modules can use oneTBB's parallel constructs in their internal work including so-called event setup modules that process conditions data. Events from multiple luminosity sections (section 9) and interval-of-validity ranges of the conditions data can be processed in parallel. The framework supports physics modules with different levels of thread-safety guarantee, which are also associated with their threading efficiency or memory usage.

12.5.2 Offloading to accelerators

The framework has generic support for offloading computations from the CPU worker threads, which allows those worker threads to continue to work on other computations [330]. A module that offloads computations has its event processing function split into two stages, where the first function is to offload the computations, and the second function is called when the offloaded computations have been completed.

Support for specific offloading technologies is implemented on top of the framework's generic mechanisms. The Nvidia GPUs on the same computing node are supported with the CUDA API [330]. Utility classes help with asynchronous execution, sharing resources between modules via the event, and minimizing data movements. At the time of writing, pixel local reconstruction, pixel track and vertex reconstruction [299], ECAL unpacking and local reconstruction, and HCAL local reconstruction have GPU implementations. Offloading them to a GPU reduces the HLT CPU usage by about 40% [295].

Historically, the accelerator vendors have provided their own APIs. However, developing and maintaining separate versions of algorithms for each platform is unsustainable and, therefore, CMS investigated ways to achieve performance portability with a single code base. A performance portability framework makes it possible to have one single code base and to build libraries for different classes of hardware, such as CPUs or different vendors of GPUs and accelerators. The suitability of Alpaka [331] and Kokkos [332] libraries for the CMS data processing model and software were explored in detail [333, 334], with the conclusion that Alpaka was better suited for CMS for LHC Run 3.

The ability to use accelerators on a remote computing node, or in separate processes on the same node, could allow more flexible use of accelerators. Using remote accelerators for machine learning (ML) inference has already been demonstrated from the CMS application framework [335–337] using inference servers such as Nvidia Triton [338]. In general, ML algorithms are expected to be easily portable between various accelerators.

12.5.3 Geometry

Geometry information is fundamental for several CMS applications, such as simulation, reconstruction, and event visualization. For the start of Run 3, the in-house developed geometry description tool used during Run 1 and Run 2, DDD [339], was replaced with the community tool, DD4Hep [340],

Table 17. Description of the data tiers regularly produced by centrally managed workflows. The ROOT framework is used to write and read the data.

Name	Description
GEN	Intermediate and outgoing stable ($c\tau \gtrsim 1$ cm) particles from the collision simulation. May include Les Houches accord event (LHE) data from the matrix-element generator, if applicable.
SIM	Detailed description of energy deposits left by stable outgoing particles in the detector material. Two options are available: a highly-accurate Geant4-based application (Full MC); and a parametric fast simulation application (Fast MC), which trades accuracy for a 100-fold decrease in detector simulation time or 10-fold decrease in total CPU time per simulated event. The level of inaccuracy introduced by Fast MC is typically a difference of less than 10% in final analysis observables.
DIGI	Digitized detector readout or simulation thereof. In simulation, the effect of additional collision events (pileup) is folded into the event description in this step. In Run 2, a "premixing" technique was introduced, where the additional events are summed in a separate processing step and then applied to the simulated primary event.
RAW	Packed detector readout data.
RECO	Detailed description of calibrated detector hits and low-level physics objects.
AOD	Reduced description of calibrated detector hits and low-level physics objects, uncalibrated high-level physics objects.
MiniAOD	Reduced low-level physics objects and calibrated high-level physics objects. A truncated floating-point representation is used for most object attributes. Introduced for Run 2 to reduce the number of analyses requiring AOD inputs.
NanoAOD	Compact data format containing only high-level physics object attributes stored as (arrays of) primitive data types. Introduced during Run 2 to reduce the number of analyses requiring MiniAOD inputs.

which is also used by the LHCb experiment and others. The selection of DD4Hep was made because its library is well behaved in multithreaded environments and because the replacement of an in-house solution with a community-supported tool improves the sustainability of our software stack. The migration of the geometry for Run 3 took place during LS2 and was an opportunity to review the entire description of the CMS detector, even improving it in some respects. The migration to DD4Hep of the Phase 2 geometry was completed during 2022.

12.6 Data formats and processing

Data and simulation processing workflows are broken into several steps, each defined by the output data structures per event it produces, referred to as a "data tier". The data tiers in use for centrally produced simulation and reconstruction workflows are listed in table 17. No changes in the data formats were made between Run 2 and Run 3. A single executable process may produce multiple outputs corresponding to different data tiers, which reduces I/O operations when the necessary data structures are already in process memory.

The smaller-size analysis formats are key to reducing both the overall amount of data stored by CMS and analysis processing time per event. In Run 2, approximate event sizes in each format are 400 kB for AOD (Analysis Object Data), 40 kB for MiniAOD, and 1–2 kB for NanoAOD. The goal is for 50% of CMS analyses to use only NanoAOD data sets as input before the end of Run 3. Although this format is not suitable for every analysis, the increase in usability and speed should appeal to many users.

Nonevent data are used to interpret and reconstruct events [1]. Four types of nonevent data remain in use: construction data, generated during the construction of the detector; equipment management data; configuration data, comprising programmable parameters related to detector operation; and conditions data, including calibrations, alignments, and detector status information. A procedure for deriving a selection of online calibration constants has been in place since Run 1. This has since been consolidated and the ability to create calibration constants from data collected across multiple runs was developed along with new calibration workflows.

A typical Run 3 simulation workflow will be a so-called "step-chain" job composing GEN-SIM-DIGI/PU MIX-RAW-RECO-AOD steps and the reduction to MiniAOD and NanoAOD formats. For detector data, only the RECO-AOD step is performed, followed by the reduction to Mini and NanoAOD. For a given data-taking period, AOD is produced 1–3 times in large-scale processing campaigns, while MiniAOD [341] and NanoAOD [342] data tiers are reproduced more frequently as high-level physics object calibrations are updated.

In the CMS workflow management system, a processing step applied to a given set of inputs (or requested number of events in the case of GEN) forms a task. Tasks are chained together to form a complete workflow, with intermediate tasks writing their output to site-local storage, and optionally registering the output in the data management system, as discussed in section 12.8.1. To reduce data transfer, task chains are converted to step-chains when possible, with intermediate output kept only on job-local scratch disk and all subsequent steps executed in a single job. As discussed in section 12.5, the steps are typically executed multithreaded, using up to eight cores per executable, which reduces the workflow management overhead. The GEN step uses software tools from the HEP theory community, and although some tools may force a step to be executed single-threaded, work is ongoing to improve per-executable parallelism in this context.

12.6.1 Premixing

In Run 2 the increase in pileup events resulted in a more I/O and computing-intensive pileup simulation. A number of individual minimum-bias events comparable to the pileup level had to be read from local or remote disk pools and superimposed to the hard scatter event. For this reason, a "premixing" simulation method [318, 343] was introduced to drastically reduce the I/O by summarising all the parasitic pileup collisions in one single "pileup-only" event. Hard-scatter events are generated and simulated without pileup; separately, a sample of MC pileup-only events is also produced, using the pileup distribution for a certain running period and considering in-time and out-of-time interactions in all subdetectors. A selected premixed set of pileup events is then overlaid on the hard scatter events. This approach reduces I/O operations by 90% compared with the previous method of overlaying individual minimum-bias events. On average, digitization and reconstruction of simulated events with pileup is twice as fast [343]. The size of the premix library is proportional to the number of simulated events and contained about 200 million events in Run 2. For Run 3, the integrated luminosity is expected to be higher, and larger premixed samples will be generated.

12.7 Computing centers

The organization of the distributed computing infrastructure used by CMS was initially based on a model described by MONARC [344], where the sites were organized in "Tiers", pledging CPU, disk storage, and tape resources, proportional to the commitment of each funding agency within the CMS Collaboration. The roles of the various Tiers were established as follows:

- A Tier 0 center close to the experiment (CERN) to execute a first calibration and reconstruction pass, the so-called "prompt reconstruction", as well as to maintain a custodial copy of all RAW data on tape; 24h/7 support is guaranteed;

- Tier 1 regional centers (six in use by CMS), which maintain a second distributed custodial copy of the RAW data on tape, and provide CPU for re-reconstruction and MC simulation; 24h/7 support is guaranteed;

- Tier 2 local centers (about 50 used by CMS), providing support for analysis activity and MC simulation; guaranteed support only during working hours.

The Tier 0 is the largest of the CMS sites. It provides computing capacity, disk and tape storage and is hosted by CERN. At the Tier 0, prompt reconstruction starts 48 hours after the data are acquired. This delay is necessary for an initial set of detector calibrations to be bootstrapped, starting from the execution of alignment and calibration processing sequences on the newly acquired data — the so-called "express" reconstruction. One such workflow is described in ref. [345]. Once the calibrations are derived, they are consolidated in a payload written to a database that is then read by prompt reconstruction at a later stage. The flexibility of the CMS central job submission infrastructure makes it possible to produce substantial amounts of simulated samples at the Tier 0.

France, Germany, Italy, the US, the U.K., and Spain provide the six Tier 1 sites of CMS. These sites have a primary role in the computing model, offering the precious combination of computing capacity, disk storage and tape archival space, all together with high bandwidth connectivity to CERN and Tier 1 s through the LHCOPN private IP network, as well as to other sites through the LHCOne network. Tier 1 s are used to store the active copy of the RAW data on tape, to perform re-reconstruction passes when data processing algorithms and calibrations improve so much to require it. The Tier 1 centers have been adapted to support the Tier 0 in the task of prompt data processing, whenever needed. In the current computing model a job can now run, in principle, wherever free CPU is available while accessing input data through the WAN (figure 138). The total amount of computing resources pledged by the largest and smallest of the Tier 1 centers of CMS differs by an order of magnitude, still these special sites are equally critical for the support of the CMS physics program.

The computing model used in Run 2 and Run 3 is a significant evolution of the MONARC hierarchical model. The roles of the Tier 1/2s have become more similar in order to optimize the usage of the resources with high efficiency, and following collaboration-wide priorities. The presence of high-speed wide-area network (WAN) connections, the development of advanced data federation [346, 347] and caching [348] technologies and the creation of optimized data tiers such as MiniAOD and NanoAOD, allow distribution of the tasks among the Tier 1's and Tier 2's following CMS priorities, with minimal net distinction between analysis, MC production, and data (re-)processing. However, Tier 1 sites still tend to be assigned a larger proportion of central production jobs.
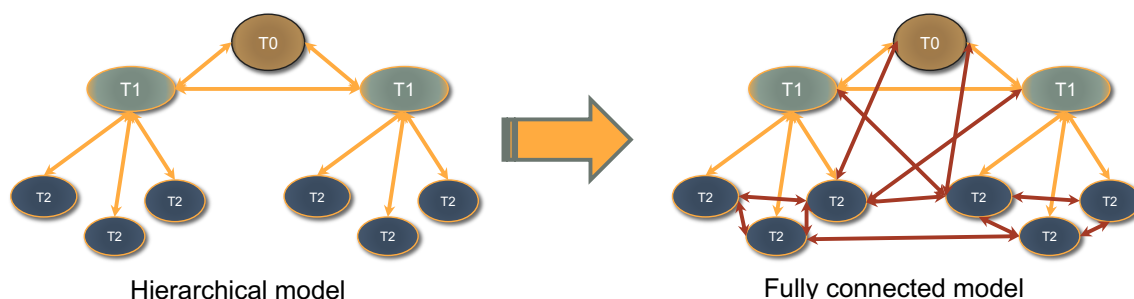
Figure 138. The evolution of the CMS computing model from a hierarchical (left) to fully connected structure (right).

Tier 3 sites are those which do not pledge resources, but nevertheless provide CPU and storage in varying amounts. These could be sites with batch farms for mainly local use, or those that predominantly support another experiment, but allow CMS to take CPU slots opportunistically when they are available.

An important aspect of the computing model in CMS is the continuous monitoring of the distributed data centers, both via recording of CPU and data metrics of the workloads, and with automatic tests checking periodically the health and reliability of the various services. These tests allow the production system to temporarily disable sites with transient issues and to re-enable them when working again.

In order to easily run computing jobs on diverse resources, the environment was standardized during Run 2 via the use of virtualization and containerization in Singularity [349].

12.7.1 High performance computing (HPC)

National and supranational bodies are investing substantial resources in supercomputers. These machines are part of our computing infrastructure and, given the advantage they represent for many scientific and industrial applications, they are here to stay. In the exascale era, HEP might be compelled to count on these resources for the bulk of its computing capacity. CMS has already started to take advantage of HPCs, adapting its software and computing tools. However, the integration of individually highly optimized HPC systems with the CMS experiment computing environment is not a trivial task. From a technical perspective, HPC centers differ in a variety of specialized hardware setups and strict usage policies can apply, particularly related to security. As a consequence, a variety of conditions can be encountered when attempting to integrate these centers, spanning from less common operating systems, absence of local scratch disk space on computing nodes, low memory availability per core, limited or even absent outbound network connectivity, and, of course, distinct architectures. Given the unique nature of HPC systems, CMS has invested a substantial amount of effort to develop solutions case by case.

CMS has deployed several solutions which can be classified broadly into two categories: a transparent site-extension model where HPC computing resources are seamlessly integrated into an existing CMS grid site, and the HEPCloud model in which an additional layer of job submission infrastructure sits between the experiment's workflow management system and the HPC computing resources. In both cases, the storage of existing CMS grid sites is used through remote access from the HPC, complemented by on-site caches or local storage at some HPC centers.

Integration of HPCs that are transparent to central computing operations have been performed in several countries in Europe, e.g., Italy, Switzerland, Germany, and Spain, for example expanding existing Tier 1 or Tier 2 sites elastically into these machines [350–352], or by deploying so-called overlay batch systems [353]. It is also through this model that a validation of the physics performance of CMS software on the PowerPC architecture could be performed [354].

The HPC resources located in the US are integrated into the CMS systems via HEPCloud [355], a portal to an ecosystem of diverse computing resources, commercial or academic, hosted at FNAL. Since 2020, HEPCloud has provisioned resources from seven different HPC centers located for example at the NERSC, PSC, TACC, or SDSC computing centers.

12.7.2 Data archive at CERN

CMS migrated to the CERN tape archive (CTA) [356] system in December 2020. The CTA system provides the tape backend to the CERN EOS disk system [267], and together EOS and CTA replace CASTOR. This upgrade was necessary to prepare for the higher data rates of Run 3 and beyond, as well as to provide a uniform API for disk and tape operations, with support for newer protocols and authentication methods. It is difficult to achieve high throughput using a traditional buffer made up of spinning disks, due to contention on the drives from multiple simultaneous streams. Therefore a new approach was taken, with a small, fast SSD buffer in front of the physical tapes. This allows the tape drives to operate close to their nominal speed of 400 MB/s.

12.8 Computing services

12.8.1 Data management

During Run 1 and Run 2, CMS used PhEDEx [357] and Dynamo [358] as data management tools. However, for Run 3 and beyond it was necessary to adopt a more scalable, flexible, and powerful system to increasingly automate the data management, and include the possibility to scale up transfers to around 100 petabytes per day by the late 2020s for the start of the High-Luminosity LHC (HL-LHC) [359]. The new software should also support future technologies such as token authorization and non-FTS (File Transfer Service) [360] transfers. The Rucio system [361], a data management project for scientific communities, was adopted by CMS at the end of 2020. It can perform all of the functions of PhEDEx and Dynamo, and also make higher-level decisions about data placement. Rucio is run centrally, without an agent at every grid site as required by PhEDEx.

The CMS Rucio infrastructure is based on Helm [362], Kubernetes [363], and Docker [364], which are industry standards. All of the Rucio services are built into a single Kubernetes cluster which can be brought up from scratch in under an hour. Rucio removes data as additional space is needed at a site. Only data that is not held in place by one or more rules is eligible to be removed. To make this decision, Rucio uses the last access time of the data. This information is taken from job reports, CMSSW file reads, and monitoring of the AAA (Any data, Anywhere, Any time) system, which is described in the following subsection.

CMS developed a mechanism to check the consistency of the Rucio database with the actual state of files on disk. Remote disks are scanned using XRootD tools, comparing the results with the state of the database, and generating lists of missing and unneeded files. These lists are passed to additional Rucio components for file removal or re-transfer.

12.8.2 Data transfer protocols

During Run 1 and Run 2, CMS transferred data among sites via protocols based on the grid security infrastructure (GSI) [365, 366], with GSIFTP the most common protocol used and GridFTP its most common implementation. When the end of support for GSI was announced, the LHC community started looking for a replacement both for the protocol and an authentication mechanism. The WebDAV protocol [367], an extension of HTTP which supports third party copy (TPC) transfers and tokens for authentication, was selected. The adoption of the WebDAV protocol started early in 2020 and was completed in 2022.

The "Any Data, Anywhere, Anytime" (AAA) data federation [346] was introduced in 2014–2015 during LS1. This is a model for effective federation of distributed storage resources via an XRootD cluster at each computing center, allowing for remote access to any file within the CMS namespace. The XRootD framework supports partial reads of files and is commonly used by analysis jobs. The combination of AAA and caching [348] technologies allows nonlocality between data and CPU, and is expected to lead to entirely storage-free computing centers in Run 3. These are the first steps towards a "Data Lakes" architecture [368], a centralized data repository, as envisioned for the HL-LHC era.

12.8.3 Central processing and production

Large-scale MC sample production and data event reconstruction activities are performed in a distributed computing infrastructure, coupled to a specialized workload management system (WMS), described in section 12.8.4. A global batch queue manages the distribution of production and analysis jobs to the CMS distributed computing system in an optimized and flexible way. The submission infrastructure (SI) employs GlideinWMS [369] and HTCondor [370] software suites in order to build and manage a "Global Pool" [371] of computing resources where the majority of the CMS tasks are executed.

The SI comprises multiple interconnected HTCondor pools [372], as shown in figure 139, redundantly deployed at CERN and FNAL in order to ensure a high-availability service. The main component of the SI, the Global Pool, obtains the majority of its resources via the submission of pilot jobs to WLCG [328] and open science grid (OSG) sites. However, locally instantiated processing nodes via, for example, DODAS [373] or BOINC [374], as well as opportunistic resources, such as the HLT filter farm [276] when not in use for data taking, can also be employed. As described in section 12.7.1, the SI computing capacity has recently expanded into HPC facilities which are integrated as part of the Global or HEPCloud pools. CERN on-site CMS resources, along with opportunistic local (BEER [375]) and cloud [376, 377] computing slots, are organized into a third HTCondor pool, built on a dedicated set of hosts to isolate it from potential issues in the main Global Pool, given its critical role in supporting Tier 0 tasks during data-taking periods. Specialized nodes known as "schedds" control workload submission. While being primarily attached to one "pool", these schedds can interact with other federated pools, requesting additional resources when demands are not covered in the primary pool, as indicated in figure 139.

The SI can leverage any grid, cloud, HPC, or opportunistic resource available to CMS, which in total have tripled in size over the last five years, currently amounting to around 400k CPU cores. The Global Pool matches diverse CMS workloads, which include single-core as well as multicore
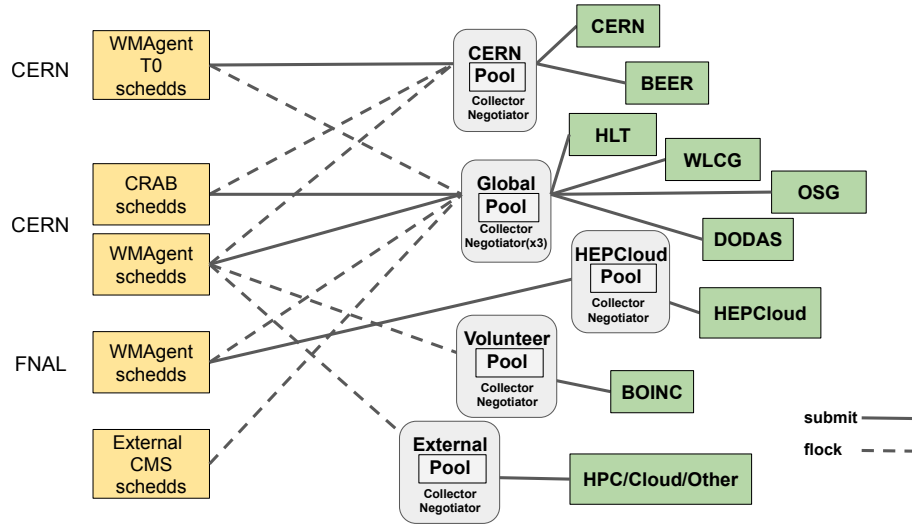
Figure 139. Schematic diagram of the submission infrastructure, including multiple distributed central processing and production (WMAgent) and analysis (CRAB) job submission agents (schedds). Reproduced from [372]. CC BY 4.0. Computing resources allocated from diverse origins (green boxes) are grouped into HTCondor pools (gray boxes), federated via workload flocking. The collector and negotiator agents (yellow boxes) keep the state of each pool and perform the workload-to-resource matchmaking.

requests [378] to these resources. A successful scheduling is achieved by simultaneously ensuring that all available resources are efficiently used [379], a fair share of resources between users is reached, and the completion of CMS tasks follows their prioritization, minimizing job failures and manual intervention. While the SI typically manages 100k to 150k simultaneously executing tasks, recent scalability tests [380] have demonstrated the capacity of the infrastructure to sustain in excess of half a million concurrently running jobs.

12.8.4 Workload management system

The CMS production WMS, based on the WMCore framework, comprises many services, including WMAgent, the most important and complex system in this architecture. It is a stateful software, responsible for the work splitting, event and data bookkeeping, job management and submission, while interfacing with the data management tools for global data bookkeeping and grid data transfers. The WMAgent system is tightly coupled to the schedds in the SI, enabling it to make use of the Global Pool resources, handling thousands of requests in parallel and enforcing the workflow prioritization. Other components in the production WMS include the WMStats system, responsible for the workflow and job monitoring, and providing logging and recovery features, and the request manager, which manages and stores workflow descriptions, submitting workloads for processing on grid resources. A suite of "Microservices" has been developed to perform very specific tasks such as automatic input and output data placement, monitoring, and cleanup of the data used in central production. The introduction of Microservices has resulted in multiple features of the system being automated, thus reducing the required operational effort.

12.8.5 Distributed analysis

For CMS scientists, a dedicated tool known as CMS remote analysis builder (CRAB) has been available since the start of Run 1, to manage the submission on the grid of data analysis applications or small-scale productions of data samples that are not centrally managed. CRAB offers the same functionalities as originally planned [1]. Users interact with a thin client running inside the CMSSW environment, which packs the user analysis environment into a sandbox and uploads it together with a configuration file to a central, database-centric service hosted in the CMSWeb services framework described in section 12.8.7. Centralized components run either as Docker containers, or under the control of HTCondor [370] or DAGMan [381]. These components handle preparation, execution, bookkeeping, error recovery, monitoring, and output delivery for the user application. CRAB has been continuously improved to provide increased automation and scaling, leaner operations, better error recovery, access to data archived on tape and integration with the Rucio data management system, while keeping the same basic interface for the user. For users needing a simpler, more interactive service to access resources from the Global Pool to make histograms, for example, the CMS Connect [382] service is available as an alternative to CRAB.

12.8.6 DBS database

The data bookkeeping service (DBS) is a catalog holding event metadata for all simulated and experimental data processed and stored by the CMS experiment. DBS contains all necessary information for tracking the data sets, including provenance information, processing history and parentage relationships between files and data sets. DBS is used by all computing jobs in CMS. The current version (version 3) was completely redesigned and reimplemented following a project review in 2009, in order to better fit the evolved CMS data processing model, to better integrate with the evolved CMS data management and workflow management (DMWM) projects, and to ensure a higher scalability. DBS3 is implemented as a Python-based web service using a standardized architecture provided by the CMS DMWM project. Client-server communication in combination with thin client APIs using the JSON format as a lightweight replacement of XML-RPC led to better scalability. In addition, the database schema has been streamlined and fine-tuned in order to avoid both excessive table joins, reduce query latency, and improve server stability, which enhance data accessibility.

The data aggregation system (DAS) [383] provides access to distributed CMS metadata via a common query language (QL). Users may access DBS, Rucio and other metadata sources by placing simple QL queries without prior knowledge of service APIs, database schemas, and implementation or location of the services.

12.8.7 Web services and security

In order to guarantee successful data taking, CMS uses a series of web applications, which perform various tasks related to job submission, job monitoring and bookkeeping, and the location of the data sets. These services are provided and maintained by several teams and centrally deployed under the umbrella CMSWeb.

To achieve optimal resource utilization and high availability, the services have been containerized and are now being deployed in a Kubernetes cluster [384]. Detailed monitoring, for all the services and the nodes they run on, has been built on top, using the central CMS computing monitoring

tools. Authorization and authentication (A&A) is also being offered as a service in the CMSWeb cluster. Currently, A&A is using X509 certificates, but is rapidly migrating to authentication tokens, following the shift to the OAuth standard initiated by the WLCG. This migration should take place by 2024, according to current plans [385].

Security is a very important aspect of web services. A great effort is required to stay up-to-date with the latest security recommendations. The web services and security teams work closely with the CERN IT security team, following up vulnerabilities. In a joint effort, proactive penetration tests are organized several times a year to discover potential issues ahead of time and to ensure that the infrastructure remains safe. As part of improving security, the logs of all the CMSWeb services are now pushed to the security operations center (SOC), which analyses them and identifies possible threats.

12.8.8 Monitoring and analytics

A scalable and reliable monitoring system is required to ensure efficient operation of the CMS distributed computing services, and to provide comprehensive measurements of the system performance. Metrics such as wall-time consumption of computing resources, memory, CPU, storage usage, and data access patterns are monitored to study the evolution of the performance over time and allow for in-depth analyses of the main system parameters.

The main components of the CMS monitoring infrastructure are presented in ref. [386]. CMS relies on central CERN [387] and dedicated CMS monitoring infrastructures. The former, supported by the CERN IT department, is used extensively to store data from CMS computing subsystems, such as HTCondor, submission infrastructure, CMSWeb user activities, analysis and production workflows coming from the WMAgent and CRAB job submission tools, and data-transfer and storage information from Rucio. The status of all CMS computing systems is monitored in real time using predefined views and dashboards.

The dedicated CMS monitoring infrastructure is composed of several Kubernetes clusters. It is mostly used to monitor computing nodes and services, and to provide additional features such as a sophisticated alert and notification system, data aggregation, annotations, and automation of several workflows. Several hundred nodes and services are monitored, and a fault-tolerant infrastructure is provided with minimal operational and maintenance effort.

13   Summary

Since the beginning of the LHC operation in 2009, the CMS detector has undergone a number of changes and upgrades, adapting the experiment to operating conditions at luminosities well beyond the original design. In 2022, the LHC Run 3 began and CMS successfully recorded its first $40\,\mathrm{fb}^{-1}$ of proton-proton data at a center-of-mass energy of 13.6 TeV with an operation efficiency of 92%. This paper describes the modifications, as installed and commissioned for LHC Run 3.

The upgraded pixel tracking detector was installed in early 2017. In the new detector, the number of barrel layers was increased from three to four, and the number of disks in each endcap from two to three, whereas the material budget was reduced, leading to a better tracking performance up to an absolute pseudorapidity of 3.0 for pixel tracks. The upgrade also involved a new readout chip enabling increased hit detection efficiencies at higher occupancy.

In the electromagnetic calorimeter, measures were taken to improve the monitoring and calibration of effects caused by irradiation, leading to a loss in the $PbWO_4$ crystal transparency and an increase of the leakage current in the avalanche photodiodes. In the second long shutdown (LS2) the calibration algorithms were refined to identify and remove spurious spike signals and to determine time-dependent correction factors for the laser monitoring system.

The upgrade of the hadron calorimeter included new readout electronics with finer granularity, leading to an increase in the number of channels and longitudinal segmentation. The previous generation of photosensors was replaced by silicon photomultipliers, which measure the scintillator light output with a better signal-to-noise ratio.

In the muon system, a gas electron multiplier (GEM) detector, consisting of four gas gaps separated by three GEM foils, was added in the endcaps. The other subsystems, drift tubes (DT), cathode strip chambers (CSC), and resistive-plate chambers (RPC), underwent several upgrades. In the DT, the muon trigger logic was replaced by a new data concentrator based on μTCA architecture. The top of CMS was covered with a neutron shield to reduce the background in the top external DT chambers. An outer ring of CSCs (ME4/2) was added in LS1, and in view of the High-Luminosity LHC, the bulk of the CSC electronics upgrades that required chamber access were performed already during LS2. An outer rings of the RPC chambers in station four (RE4/2 and RE4/3) were added as well. The endcap muon track finder of the L1 trigger was upgraded to utilize GEM-CSC joint track segments to optimize the final track reconstruction and resolution at the trigger level.

The precision proton spectrometer was upgraded significantly. Its tracker radiation-damaged sensors and chips were replaced. The mechanics of the detector, as well as the front-end electronics, were completely redesigned to add a novel internal motion system designed to mitigate the effects of radiation damage. In the timing system a second station was installed in each arm. All detector modules were replaced by new double-diamond modules with the aim of further improving the timing resolution.

In the beam radiation instrumentation and luminosity system, new versions of the pixel luminosity telescope (PLT), the fast beam conditions monitor (BCM1F), and the beam conditions monitor for losses (BCML) were installed for Run 3.

To cope with increasing instantaneous luminosities, the CMS data acquisition (DAQ) system underwent multiple upgrades. The backend technology was gradually moved to the more powerful μTCA standard. A new optical readout link with a higher bandwidth of 10 Gb/s was developed. The bulk of the DAQ system downstream from the custom readout benefited from advances in technology to achieve a much more compact design, while doubling the event building bandwidth.

The first level (L1) trigger, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of up to 110 kHz within a fixed latency of about 4 μs. The developments in the L1 trigger mostly focused on the addition of dedicated triggers that became possible due to enhanced capabilities of the global trigger logic and increased trigger information delivered by the calorimeters and muon systems. Among other applications, new triggers for long-lived particle signatures were implemented. The addition of a 40 MHz scouting system that receives data from both the calorimeter and muon systems, further broadens the physics reach of CMS.

The high-level trigger (HLT) performs the second stage of event filtering and accepts events at a sustained rate of the order of 5 kHz. Since Run 3 began, an additional 30 kHz of HLT scouting data is recorded. Since 2016, the HLT has been operated using multithreaded event processing software,

minimizing memory requirements through reduction of the number of processes concurrently running. For Run 3, GPUs were successfully deployed in the HLT.

Substantial improvements were achieved in the physics performance and speed of the software, as well as in the computing infrastructure. Some of the major changes are: support for multithreaded processes and utilization of GPUs; direct remote data access; and usage of high-performance computing centers. New tools such as Rucio for data management were adopted with future data rates in mind. Considerable effort was put into the automation of the workflows and the validation of the software. Physics analyses have been moved to smaller and smaller formats for centrally produced and experiment-wide shared data samples, the most recent of which is the NanoAOD.

The development of the CMS detector, as described in this paper, constitutes a solid basis for future data taking.

Acknowledgments

References

[1] CMS collaboration, *The CMS Experiment at the CERN LHC*, 2008 *JINST* **3** S08004.

[2] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [arXiv:1207.7235].

[3] CMS collaboration, *Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **06** (2013) 081 [arXiv:1303.4571].

[4] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [arXiv:1207.7214].

[5] CMS collaboration, *Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV*, 2020 *JINST* **15** P10017 [arXiv:2006.10165].

[6] CMS collaboration, *The CMS trigger system*, 2017 *JINST* **12** P01020 [arXiv:1609.02366].

[7] CMS collaboration, *Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, 2015 *JINST* **10** P06005 [arXiv:1502.02701].

[8] CMS collaboration, *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV*, 2018 *JINST* **13** P06015 [arXiv:1804.04528].

[9] CMS collaboration, *Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{(s)} = 8$ TeV*, 2015 *JINST* **10** P08010 [arXiv:1502.02702].

[10] CMS collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, 2014 *JINST* **9** P10009 [arXiv:1405.6569].

[11] CMS collaboration, *Cutaway diagrams of CMS detector*, CMS Collection CMS-OUTREACH-2019-001 (2019).

[12] CMS and TOTEM collaborations, *Proton reconstruction with the CMS-TOTEM Precision Proton Spectrometer*, 2023 *JINST* **18** P09009 [arXiv:2210.05854].

[13] M. Hoch, *CMS detector and cavern (open view)*, CMS Collection CMS-PHO-GEN-2008-019 CERN, Geneva, Switzerland (2008).

[14] S. Yammine, G. Le Godec, and H. Thiesen, *Study report for the new free wheeling thyristors system for the upgrade of the CMS solenoid power converter*, Tech. Rep. EDMS Document 1845659 v3.2, CERN, Geneva (2023).

[15] CMS collaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, Tech. Rep. CERN-LHCC-2012-016 CERN, Geneva (2012).

[16] CMS Tracker Group collaboration, *The CMS Phase-1 Pixel Detector Upgrade*, 2021 *JINST* **16** P02027 [arXiv:2012.14304].

[17] ROSE collaboration, *Radiation hard silicon detectors developments by the RD48 (ROSE) Collaboration*, *Nucl. Instrum. Meth. A* **466** (2001) 308.

[18] T. Rohe et al., *Fluence dependence of charge collection of irradiated pixel sensors*, *Nucl. Instrum. Meth. A* **552** (2005) 232 [physics/0411214].

[19] H.C. Kästli, *Frontend electronics development for the CMS pixel detector upgrade*, *Nucl. Instrum. Meth. A* **731** (2013) 88.

[20] R. Horisberger, *Readout architectures for pixel detectors*, *Nucl. Instrum. Meth. A* **465** (2000) 148.

[21] CMS TRACKER GROUP collaboration, *The DAQ and control system for the CMS Phase-1 pixel detector upgrade*, 2019 *JINST* **14** P10017.

[22] A. Marchioro and P. Moreira, *QPLL: a Quartz Crystal Based PLL for Jitter Filtering Applications in LHC*, in the proceedings of the *9th Workshop on Electronics for LHC Experiments*, Amsterdam, The Netherlands, 29 September–3 October 2003, pp. 280–284 [DOI:10.5170/CERN-2003-006.280].

[23] C. Nägeli, *Analysis of the rare decay $B_s^0 \to \mu^+\mu^-$ using the compact muon solenoid experiment at CERN's Large Hadron Collider*, Ph.D. thesis, ETH, Zürich, Switzerland (2013).

[24] M. Pesaresi et al., *The FC7 AMC for generic DAQ & control applications in CMS*, 2015 *JINST* **10** C03036.

[25] E. Hazen et al., *The AMC13XG: a new generation clock/timing/DAQ module for CMS MicroTCA*, 2013 *JINST* **8** C12036.

[26] S. Michelis et al., *DC-DC converters in 0.35 μm CMOS technology*, 2012 *JINST* **7** C01072.

[27] L. Feld et al., *The DC-DC conversion power system of the CMS Phase-1 pixel upgrade*, 2015 *JINST* **10** C01052.

[28] L. Feld et al., *Experience from design, prototyping and production of a DC-DC conversion powering scheme for the CMS Phase-1 Pixel Upgrade*, 2016 *JINST* **11** C02033.

[29] C. Paillard, A. Marchioro and C. Ljuslin, *The CCU25: a network oriented communication and control unit integrated circuit in a 0.25 μm CMOS technology*, in the proceedings of the *8th Workshop on Electronics for LHC Experiments*, Colmar, France, 9–13 September 2002, pp. 174–178 [DOI:10.5170/CERN-2002-003.174].

[30] F. Faccio, S. Michelis, and G. Ripamonti, *Summary of measurements on FEAST2 modules to understand the failures observed in the CMS pixel system*, Tech. Rep., CERN, Geneva (2018), https://espace.cern.ch/project-DCDC-new/Shared%20Documents/SummaryMeasurements18.pdf.

[31] P. Tropea et al., *Advancements and plans for the LHC upgrade detector thermal management with CO_2 evaporative cooling*, *Nucl. Instrum. Meth. A* **936** (2019) 644.

[32] CMS collaboration, *The Phase-2 Upgrade of the CMS Tracker*, Tech. Rep. CERN-LHCC-2017-009, CERN, Geneva (2017) [DOI:10.17181/CERN.QZ28.FLHW].

[33] M.J. French et al., *Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker*, *Nucl. Instrum. Meth. A* **466** (2001) 359.

[34] CMS collaboration, *Operation and performance of the CMS silicon strip tracker with proton-proton collisions at the CERN LHC*, to be submitted to *JINST* (2023).

[35] CMS collaboration, *Silicon Strip Tracker Performance results 2018*, Tech. Rep. CMS-DP-2018-052 CERN, Geneva (2018).

[36] C. Barth, *Performance of the CMS Tracker under Irradiation*, Ph.D. thesis, KIT, Karlsruhe, Germany (2013).

[37] K. Deiters et al., *Properties of the most recent avalanche photodiodes for the CMS electromagnetic calorimeter*, *Nucl. Instrum. Meth. A* **442** (2000) 193.

[38] F. Addesa and F. Cavallari, *Performance prospects for the CMS electromagnetic calorimeter barrel avalanche photodiodes for LHC phase I and phase II: Radiation hardness and longevity*, *Nucl. Instrum. Meth. A* **787** (2015) 114.

[39] CMS ECAL collaboration, *Radiation hardness qualification of $PbWO_4$ scintillation crystals for the CMS Electromagnetic Calorimeter*, 2010 *JINST* **5** P03010 [arXiv:0912.4300].

[40] R. Benetta et al., *The CMS ECAL readout architecture and the clock and control system*, in the proceedings of the *11ᵗʰ International Conference on Calorimetry in High-Energy Physics*, Perugia, Italy, 29 March–2 April 2004, p. 162–172 [`DOI:10.1142/9789812701978_0021`].

[41] CMS collaboration, *Reconstruction of signal amplitudes in the CMS electromagnetic calorimeter in the presence of overlapping proton-proton interactions*, 2020 *JINST* **15** P10002 [`arXiv:2006.14359`].

[42] CMS ECAL collaboration, *Reconstruction of the signal amplitude of the CMS electromagnetic calorimeter*, *Eur. Phys. J. C* **46S1** (2006) 23.

[43] CMS collaboration, *Time Reconstruction and Performance of the CMS Electromagnetic Calorimeter*, 2010 *JINST* **5** T03011 [`arXiv:0911.4044`].

[44] P. Paganini, *CMS electromagnetic trigger commissioning and first operation experiences*, *J. Phys. Conf. Ser.* **160** (2009) 012062.

[45] M. Hansen, *The new readout architecture for the CMS ECAL*, in the proceedings of the *9ᵗʰ Workshop on Electronics for LHC Experiments*, Amsterdam, Netherlands, 29 September–3 October 2003, pp. 78–82 [`DOI:10.5170/CERN-2003-006.78`].

[46] D.A. Petyt, *Anomalous APD signals in the CMS Electromagnetic Calorimeter*, *Nucl. Instrum. Meth. A* **695** (2012) 293.

[47] CMS collaboration, *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s}$ = 7 TeV*, 2013 *JINST* **8** P09009 [`arXiv:1306.2016`].

[48] F. Thiant et al., *New development in the CMS ECAL Level-1 trigger system to meet the challenges of LHC Run 2*, *PoS* **TWEPP2018** (2019) 052.

[49] CMS collaboration, *Search for long-lived particles using delayed photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Phys. Rev. D* **100** (2019) 112003 [`arXiv:1909.06166`].

[50] CMS collaboration, *The hadron calorimeter*, Tech. Rep. CERN-LHCC-97-031, CERN, Geneva (1997).

[51] CMS HCAL collaboration, *Design, performance, and calibration of CMS hadron-barrel calorimeter wedges*, *Eur. Phys. J. C* **55** (2008) 159.

[52] CMS HCAL collaboration, *Design, performance and calibration of the CMS forward calorimeter wedges*, *Eur. Phys. J. C* **53** (2008) 139.

[53] CMS HCAL collaboration, *Design, performance, and calibration of the CMS Hadron-outer calorimeter*, *Eur. Phys. J. C* **57** (2008) 653.

[54] CMS collaboration, *Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s}$ = 13 TeV*, 2020 *JINST* **15** P05002 [`arXiv:1910.00079`].

[55] CMS collaboration, *Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s}$ = 13 TeV*, 2020 *JINST* **15** P05002 [`arXiv:1910.00079`].

[56] T. Zimmerman and J.R. Hoff, *The Design of a charge integrating, modified floating point ADC chip*, *IEEE J. Solid State Circuits* **39** (2004) 895.

[57] A. Baumbaugh et al., *QIE10: a new front-end custom integrated circuit for high-rate experiments*, 2014 *JINST* **9** C01062.

[58] T. Roy et al., *QIE: Performance Studies of the Next Generation Charge Integrator*, 2015 *JINST* **10** C02009.

[59] D. Hare et al., *First large volume characterization of the QIE10/11 custom front-end integrated circuits*, 2016 *JINST* **11** C02052.

[60] CMS collaboration, *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*, Tech. Rep. CERN-LHCC-2012-015 CERN, Geneva (2012).

[61] CMS collaboration, *The Phase-2 Upgrade of the CMS Barrel Calorimeters*, Tech. Rep. CERN-LHCC-2017-011, CERN, Geneva (2017).

[62] CMS collaboration, *Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter*, 2010 *JINST* **5** T03014 [`arXiv:0911.4881`].

[63] R.A. Shukla et al., *Microscopic characterisation of photodetectors used in the hadron calorimeter of the Compact Muon Solenoid experiment*, *Rev. Sci. Instrum.* **90** (2019) 023303.

[64] CMS collaboration, *Measurements with silicon photomultipliers of dose-rate effects in the radiation damage of plastic scintillator tiles in the CMS hadron endcap calorimeter*, 2020 *JINST* **15** P06009 [`arXiv:2001.06553`].

[65] P. Cushman, A. Heering and A. Ronzhin, *Custom HPD readout for the CMS HCAL*, *Nucl. Instrum. Meth. A* **442** (2000) 289.

[66] A. Heering et al., *Parameters of the preproduction series SiPMs for the CMS HCAL phase I upgrade*, *Nucl. Instrum. Meth. A* **824** (2016) 115.

[67] Y. Musienko et al., *Radiation damage studies of silicon photomultipliers for the CMS HCAL phase I upgrade*, *Nucl. Instrum. Meth. A* **787** (2015) 319.

[68] F. Vasey et al., *The Versatile Link common project: feasibility report*, 2012 *JINST* **7** C01075.

[69] G. Cummings, *CMS HCAL VTRx-induced communication loss and mitigation*, 2022 *JINST* **17** C05020.

[70] P. Moreira et al., *The GBT-SerDes ASIC prototype*, 2010 *JINST* **5** C11022.

[71] J. Gutleber and L. Orsini, *Software Architecture for Processing Clusters Based on I_2O*, *Clust. Comput.* **5** (2002) 55.

[72] C. Ghabrous Larrea et al., *IPbus: a flexible Ethernet-based control system for xTCA hardware*, 2015 *JINST* **10** C02019 [CMS-CR-2014-334].

[73] V. Brigljevic et al., *Run control and monitor system for the CMS experiment*, *eConf* **C0303241** (2003) THGT002 [`cs/0306110`].

[74] D. Charousset, R. Hiesgen and T.C. Schmidt, *Revisiting Actor Programming in C++*, *Comput. Lang. Syst. Struct.* **45** (2016) 105 [`arXiv:1505.07368`].

[75] JSON-RPC Working Group, *JSON-RPC 2.0 specification*, 2013, https://www.jsonrpc.org/specification.

[76] I. Fette and A. Melnikov, *The WebSocket protocol*, RFC Proposed Standard 6455, 2011 [`DOI:10.17487/RFC6455`].

[77] B. Bilki, *Review of Scalar Meson Production at $\sqrt{s} = 7$ TeV in CMS, U(1)$'$ Gauge Extensions of the MSSM and Calorimetry for Future Colliders*, Ph.D. thesis, University of Iowa, Iowa City, IA, U.S.A. (2011), CERN-THESIS-2011-229.

[78] CMS HCAL collaboration, *Study of Various Photomultiplier Tubes with Muon Beams And Cerenkov Light Produced in Electron Showers*, 2010 *JINST* **5** P06002.

[79] CMS ECAL and HCAL collaborations, *The CMS barrel calorimeter response to particle beams from 2 to 350-GeV/c*, *Eur. Phys. J. C* **60** (2009) 359.

[80] CMS collaboration, *Performance of CMS Hadron Calorimeter Timing and Synchronization using Test Beam, Cosmic Ray, and LHC Beam Data*, 2010 *JINST* **5** T03013 [arXiv:0911.4877].

[81] J. Lawhorn, *New method of out-of-time energy subtraction for the CMS hadronic calorimeter*, *J. Phys. Conf. Ser.* **1162** (2019) 012036.

[82] CMS collaboration, *Performance of the local reconstruction algorithms for the CMS hadron calorimeter with Run 2 data*, arXiv:2306.10355.

[83] CMS collaboration, *Noise in Phase 1 HF detector in 2017*, Tech. Rep. CMS-DP-2017-034 CERN, Geneva (2017).

[84] CMS collaboration, *Search for invisible decays of the Higgs boson produced via vector boson fusion in proton-proton collisions at s=13 TeV*, *Phys. Rev. D* **105** (2022) 092007 [arXiv:2201.11585].

[85] CMS collaboration, *The Performance of the CMS Muon Detector in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV at the LHC*, 2013 *JINST* **8** P11002 [arXiv:1306.6905].

[86] CMS collaboration, *Calibration of the CMS Drift Tube Chambers and Measurement of the Drift Velocity with Cosmic Rays*, 2010 *JINST* **5** T03016 [arXiv:0911.4895].

[87] CMS collaboration, *Technical proposal for the upgrade of the CMS detector through 2020*, Tech. Rep. CERN-LHCC-2011-006 CERN, Geneva (2011).

[88] CMS collaboration, *The Phase-2 Upgrade of the CMS Level-1 Trigger*, Tech. Rep. CERN-LHCC-2020-004, CERN, Geneva (2020).

[89] Á. Navarro-Tobar and C. Fernández-Bedoya, *CMS DT upgrade: the sector collector relocation*, 2016 *JINST* **11** C02046.

[90] Á. Navarro-Tobar, C. Fernández-Bedoya and I. Redondo, *Low-cost, high-precision propagation delay measurement of 12-fibre MPO cables for the CMS DT electronics upgrade*, 2013 *JINST* **8** C02001.

[91] CMS collaboration, *CMS Technical Design Report for the Level-1 Trigger Upgrade*, Tech. Rep. CERN-LHCC-2013-011 CERN, Geneva (2013).

[92] A. Triossi et al., *A New Data Concentrator for the CMS Muon Barrel Track Finder*, *PoS* **TIPP2014** (2014) 412.

[93] Xilinx Inc., *7 series FPGAs GTX/GTH transceivers*, User Guide UG476 v1.12.1, 2018, https://docs.xilinx.com/v/u/en-US/ug476_7Series_Transceivers.

[94] P. Moreira et al., *A radiation tolerant gigabit serializer for LHC data transmission*, in the proceedings of the *7th Workshop on Electronics for LHC Experiments*, Stockholm, Sweden, 10–14 September 2001, pp. 145–149, http://cds.cern.ch/record/588665 [DOI:10.5170/CERN-2001-005.145].

[95] CMS collaboration, *Performance of the CMS TwinMux Algorithm in late 2016 pp collision runs*, Tech. Rep. CMS-DP-2016-074 CERN, Geneva (2016).

[96] Á. Navarro-Tobar et al., *Phase 1 upgrade of the CMS drift tubes read-out system*, 2017 *JINST* **12** C03070.

[97] CMS collaboration, *Efficiency of the CMS Drift Tubes at LHC in 2017*, Tech. Rep. CMS-DP-2018-016 CERN, Geneva (2018).

[98] CMS collaboration, *Performance of the CMS Drift Tubes at the end of LHC Run 2*, Tech. Rep. CMS-DP-2019-008 CERN, Geneva (2019).

[99] E. Conti and F. Gasparini, *Test of the wire ageing induced by radiation for the CMS barrel muon chambers*, *Nucl. Instrum. Meth. A* **465** (2001) 472.

[100] CMS collaboration, *The Phase-2 Upgrade of the CMS Muon Detectors*, Tech. Rep. CERN-LHCC-2017-012, CERN, Geneva (2017).

[101] G. Altenhöfer et al., *The drift velocity monitoring system of the CMS barrel muon chambers*, *Nucl. Instrum. Meth. A* **888** (2018) 1.

[102] CMS collaboration, *Background measurements in the CMS DT chambers during LHC Run 2*, Tech. Rep. CMS-DP-2020-011 CERN, Geneva (2020).

[103] CMS collaboration, *Local reconstruction algorithms in the cathode strip chambers of CMS*, *EPJ Web Conf.* **214** (2019) 02014.

[104] J. Heilman, *Work on CMS muon detector (CSCs) during long shutdown 1 (LS1)*, CMS Collection CMS-PHO-MUON-2015-001 (2014).

[105] CMS collaboration, *Upgrade of the CSC Muon System for the CMS Detector at the HL-LHC*, *PoS* **EPS-HEP2021** (2022) 837.

[106] CAEN S.p.A., *Mod. A1733-A1833-A1733B-A1833B HV boards*, Technical Information Manual rev. 10, 2013, https://www.caen.it/download/.

[107] D. Pfeiffer et al., *The radiation field in the Gamma Irradiation Facility GIF++ at CERN*, *Nucl. Instrum. Meth. A* **866** (2017) 91 [arXiv:1611.00299].

[108] B. Bylsma et al., *Radiation Testing of Electronics for the CMS Endcap Muon System*, *Nucl. Instrum. Meth. A* **698** (2013) 242 [arXiv:1208.4051].

[109] F. Ferrarese, *Statistical analysis of total ionizing dose response in 25-nm NAND flash memory*, Ph.D. thesis, Università degli Studi di Padova, Padova, Italy (2014), http://tesi.cab.unipd.it/45696/.

[110] K. Bunkowski, *Optimization, Synchronization, Calibration and Diagnostic of the RPC PAC Muon Trigger System for the CMS detector*, Ph.D. thesis, Warsaw University, Warsaw, Poland (2009).

[111] E.P. Corrin, *Development of Digital Readout Electronics for the CMS Tracker*, Ph.D. thesis, Imperial College, London, U.K. (2002).

[112] S. Colafranceschi et al., *Resistive plate chambers for 2013-2014 muon upgrade in CMS at LHC*, 2014 *JINST* **9** C10033.

[113] CMS collaboration, *CMS technical design report, volume II: Physics performance*, *J. Phys. G* **34** (2007) 995.

[114] L.-B. Cheng, P.-C. Cao, J.-Z. Zhao and Z.-A. Liu, *Design of online control and monitoring software for the CPPF system in the CMS Level-1 trigger upgrade*, *Nucl. Sci. Tech.* **29** (2018) 166.

[115] A. Samalan et al., *RPC system in the CMS Level-1 Muon Trigger*, 2020 *JINST* **15** C10007.

[116] M. Abbrescia et al., *Resistive plate chambers performances at cosmic rays fluxes*, *Nucl. Instrum. Meth. A* **359** (1995) 603.

[117] S. Colafranceschi et al., *Performance of the Gas Gain Monitoring system of the CMS RPC muon detector and effective working point fine tuning*, 2012 *JINST* **7** P12004 [arXiv:1209.3893].

[118] J. Goh et al., *CMS RPC tracker muon reconstruction*, 2014 *JINST* **9** C10027.

[119] M.I. Pedraza-Morales, M.A. Shah and M. Shopova, *First results of CMS RPC performance at 13 TeV*, 2016 *JINST* **11** C12003 [arXiv:1605.09521].

[120] M.A. Shah et al., *Experiences from the RPC data taking during the CMS RUN-2*, 2020 *JINST* **15** C10027 [arXiv:2005.12532].

[121] M.A. Shah et al., *The CMS RPC Detector Performance and Stability during LHC RUN-2*, 2019 *JINST* **14** C11012 [arXiv:1808.10488].

[122] A. Gelmi, R. Guida and B. Mandelli, *Gas mixture quality studies for the CMS RPC detectors during LHC Run 2*, 2021 *JINST* **16** C04004.

[123] M. Abbrescia et al., *HF production in CMS-Resistive Plate Chambers*, *Nucl. Phys. B Proc. Suppl.* **158** (2006) 30.

[124] CMS collaboration, *The CMS muon project: Technical Design Report*, Tech. Rep. CERN-LHCC-97-032, CERN, Geneva (1997).

[125] A. Gelmi et al., *Longevity studies on the CMS-RPC system*, 2019 *JINST* **14** C05012.

[126] R. Guida, *GIF++: A new CERN Irradiation Facility to test large-area detectors for the HL-LHC program*, *PoS* **ICHEP2016** (2016) 260.

[127] R. Aly et al., *Aging Study on Resistive Plate Chambers of the CMS Muon Detector for HL-LHC*, 2020 *JINST* **15** C11002 [arXiv:2005.11397].

[128] S. Costantini et al., *Radiation background with the CMS RPCs at the LHC*, 2015 *JINST* **10** C05031 [arXiv:1406.2859].

[129] CMS Muon collaboration, *Latest results of Longevity studies on the present CMS RPC system for HL-LHC phase*, *Nucl. Instrum. Meth. A* **1055** (2023) 168452.

[130] G. Carboni et al., *An extensive aging study of bakelite resistive plate chambers*, *Nucl. Instrum. Meth. A* **518** (2004) 82.

[131] F. Thyssen, *Performance of the resistive plate chambers in the CMS experiment*, 2012 *JINST* **7** C01104.

[132] G. Pugliese et al., *Aging studies for Resistive Plate Chambers of the CMS muon trigger detector*, *Nucl. Instrum. Meth. A* **515** (2003) 342.

[133] R. Guida and B. Mandelli, *R&D strategies for optimizing greenhouse gases usage in the LHC particle detection systems*, *Nucl. Instrum. Meth. A* **958** (2020) 162135.

[134] R. Guida, B. Mandelli and G. Rigoletti, *Studies on alternative eco-friendly gas mixtures and development of gas recuperation plant for RPC detectors*, *Nucl. Instrum. Meth. A* **1039** (2022) 167045.

[135] F. Sauli, *GEM: A new concept for electron amplification in gas detectors*, *Nucl. Instrum. Meth. A* **386** (1997) 531.

[136] CMS Muon collaboration, *Layout and Assembly Technique of the GEM Chambers for the Upgrade of the CMS First Muon Endcap Station*, *Nucl. Instrum. Meth. A* **918** (2019) 67 [arXiv:1812.00411].

[137] C. Calabria, *Large-size triple GEM detectors for the CMS forward muon upgrade*, *Nucl. Part. Phys. Proc.* **273-275** (2016) 1042.

[138] CMS collaboration, *Benchmarking LHC background particle simulation with the CMS triple-GEM detector*, 2021 *JINST* **16** P12026 [arXiv:2107.03621].

[139] P. Aspell et al., *VFAT3: A Trigger and Tracking Front-end ASIC for the Binary Readout of Gaseous and Silicon Sensors*, in the proceedings of the *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Sydney, NSW, Australia, 10–17 November 2018, [DOI:10.1109/NSSMIC.2018.8824655].

[140] P. Aspell et al., *Development of a GEM Electronic Board (GEB) for triple-GEM detectors*, 2014 *JINST* **9** C12030.

[141] D. Abbaneo, *CMS muon system phase 2 upgrade with triple-GEM detectors*, in the proceedings of the *IEEE Nuclear Science Symposium and Medical Imaging Conference*, San Diego, CA, U.S.A., 31 October 2015–7 November 2015, [DOI:10.1109/NSSMIC.2015.7581803].

[142] T. Lenzi, *A micro-TCA based data acquisition system for the Triple-GEM detectors for the upgrade of the CMS forward muon spectrometer*, 2017 *JINST* **12** C01058.

[143] A. Svetek et al., *The Calorimeter Trigger Processor Card: the next generation of high speed algorithmic data processing at CMS*, 2016 *JINST* **11** C02011.

[144] R.K. Mommsen et al., *The CMS event-builder system for LHC run 3 (2021-23)*, *EPJ Web Conf.* **214** (2019) 01006.

[145] P. Aspell et al., *VFAT2: A front-end "system on chip" providing fast trigger information and digitized data storage for the charge sensitive readout of multi-channel silicon and gas particle detectors.*, in the proceedings of the *IEEE Nuclear Science Symposium and Medical Imaging Conference and 16*[th] *International Workshop on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors*, Dresden, Germany, 19–25 October 2008, pp. 1489–1494 [DOI:10.1109/NSSMIC.2008.4774696].

[146] D. Abbaneo et al., *Design of a constant fraction discriminator for the VFAT3 front-end ASIC of the CMS GEM detector*, 2016 *JINST* **11** C01023.

[147] P. Moreira et al., *The GBT Project*, in the proceedings of the *Topical Workshop on Electronics for Particle Physics*, Paris, France, 21–25 September 2009, pp. 342–346 [DOI:10.5170/CERN-2009-006.342].

[148] P. Alfke, *Xilinx Virtex-6 and Spartan-6 FPGA families*, in the proceedings of the *IEEE Hot Chips 21 Symposium*, Stanford CA, U.S.A., 23–25 August 2009, pp. 1–20.

[149] D. Abbaneo et al., *The Triple-GEM Project for the Phase 2 Upgrade of the CMS Muon System*, *PoS* **TIPP2014** (2014) 065.

[150] K. Ecklund et al., *Upgrade of the cathode strip chamber level 1 trigger optical links at CMS*, 2012 *JINST* **7** C11011.

[151] D. Acosta, *Boosted Decision Trees in the Level-1 Muon Endcap Trigger at CMS*, *J. Phys. Conf. Ser.* **1085** (2018) 042042.

[152] P. Golonka et al., *FwWebViewPlus: integration of web technologies into WinCC OA based Human-Machine Interfaces at CERN*, *J. Phys. Conf. Ser.* **513** (2014) 012009.

[153] M. Abbas et al., *Detector Control System for the GE1/1 slice test*, 2020 *JINST* **15** P05023.

[154] R. Venditti, *Production, Quality Control and Performance of GE1/1 Detectors for the CMS Upgrade*, *J. Phys. Conf. Ser.* **1498** (2020) 012055.

[155] CMS Muon collaboration, *Performance of prototype GE1/1 chambers for the CMS muon spectrometer upgrade*, *Nucl. Instrum. Meth. A* **972** (2020) 164104 [arXiv:1903.02186].

[156] CMS and TOTEM collaborations, *CMS-TOTEM Precision Proton Spectrometer*, Tech. Rep. CERN-LHCC-2014-021 CERN, Geneva (2014).

[157] TOTEM collaboration, *Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report*, Tech. Rep. CERN-LHCC-2004-002, CERN, Geneva (2004).

[158] M.G. Albrow, *The CMS-TOTEM Precision Proton Spectrometer: CT-PPS*, *PoS* **DIS2015** (2015) 064.

[159] V. Vacek, G.D. Hallewell, S. Ilie and S. Lindsay, *Perfluorocarbons and their use in Cooling Systems for Semiconductor Particle Detectors*, *Fluid Phase Equilib.* **174** (2000) 191.

[160] G. Ruggiero et al., *Characteristics of edgeless silicon detectors for the Roman Pots of the TOTEM experiment at the LHC*, *Nucl. Instrum. Meth. A* **604** (2009) 242.

[161] TOTEM collaboration, *Performance of the TOTEM Detectors at the LHC*, *Int. J. Mod. Phys. A* **28** (2013) 1330046 [`arXiv:1310.2908`].

[162] F. Ravera, *The CT-PPS tracking system with 3D pixel detectors*, 2016 *JINST* **11** C11027.

[163] F. Ravera, *3D silicon pixel detectors for the CT-PPS tracking system*, Ph.D. thesis, Università degli Studi di Torino, Torino, Italy (2017), CERN-THESIS-2017-473.

[164] D.M.S. Sultan et al., *First Production of New Thin 3D Sensors for HL-LHC at FBK*, 2017 *JINST* **12** C01022 [`arXiv:1612.00638`].

[165] G. Pellegrini et al., *3D double sided detector fabrication at IMB-CNM*, *Nucl. Instrum. Meth. A* **699** (2013) 27.

[166] G.-F. Dalla Betta et al., *Small pitch 3D devices*, *PoS* **Vertex2016** (2017) 028.

[167] D. Hits and A. Starodumov, *The CMS Pixel Readout Chip for the Phase 1 Upgrade*, 2015 *JINST* **10** C05029.

[168] J. Hoss et al., *Radiation tolerance of the readout chip for the Phase I upgrade of the CMS pixel detector*, 2016 *JINST* **11** C01003.

[169] TOTEM collaboration, *Diamond Detectors for the TOTEM Timing Upgrade*, 2017 *JINST* **12** P03007 [`arXiv:1701.05227`].

[170] E. Bossini, *The CMS Precision Proton Spectrometer timing system: performance in Run 2, future upgrades and sensor radiation hardness studies*, 2020 *JINST* **15** C05054 [`arXiv:2004.11068`].

[171] E. Bossini and N. Minafra, *Diamond Detectors for Timing Measurements in High Energy Physics*, *Front. in Phys.* **8** (2020) 248 [`arXiv:2020.00248`].

[172] CMS collaboration, *Time resolution of the diamond sensors used in the Precision Proton Spectrometer*, Tech. Rep. CMS-DP-2019-034 CERN, Geneva (2019).

[173] E. Bossini, D.M. Figueiredo, L.M.C. Forthomme and F.I. Garcia Fuentes, *Test beam results of irradiated single-crystal CVD diamond detectors at DESY-II*, Tech. Rep. CMS-NOTE-2020-007, CERN, Geneva (2020).

[174] F. Anghinolfi et al., *NINO: An ultra-fast and low-power front-end amplifier/discriminator ASIC designed for the multigap resistive plate chamber*, *Nucl. Instrum. Meth. A* **533** (2004) 183.

[175] J. Christiansen, *HPTDC: High performance time to digital converter. version 2.2 for HPTDC version 1.3*, Tech. Rep., https://cds.cern.ch/record/1067476/, CERN, Geneva (2000).

[176] C. Royon, *SAMPIC: a readout chip for fast timing detectors in particle physics and medical imaging*, *J. Phys. Conf. Ser.* **620** (2015) 012008 [`arXiv:1503.04625`].

[177] TOTEM collaboration, *Timing Measurements in the Vertical Roman Pots of the TOTEM Experiment*, Tech. Rep. CERN-LHCC-2014-020 CERN, Geneva (2014).

[178] M. Bousonville and J. Rausch, *Universal picosecond timing system for the Facility for Antiproton and Ion Research*, *Phys. Rev. ST Accel. Beams* **12** (2009) 042801.

[179] P. Moritz and B. Zipfel, *Recent Progress on the Technical Realization of the Bunch Phase Timing System BuTiS*, *Conf. Proc. C* **110904** (2011) 418.

[180] M. Quinto, F. Cafagna, A. Fiergolski and E. Radicioni, *Upgrade of the TOTEM DAQ using the Scalable Readout System (SRS)*, 2013 *JINST* **8** C11006.

[181] CMS collaboration, *Precision luminosity measurement in proton-proton collisions at $\sqrt{s}$ = 13 TeV in 2015 and 2016 at CMS*, *Eur. Phys. J. C* **81** (2021) 800 [`arXiv:2104.01927`].

[182] CMS collaboration, *CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s}$ = 13 TeV*, Tech. Rep. CMS-PAS-LUM-17-004, CERN, Geneva (2018).

[183] CMS collaboration, *CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s}$ = 13 TeV*, Tech. Rep. CMS-PAS-LUM-18-002, CERN, Geneva (2019).

[184] CMS collaboration, *CMS luminosity measurement using nucleus-nucleus collisions at $\sqrt{s_{NN}}$ = 5.02 TeV in 2018*, Tech. Rep. CMS-PAS-LUM-18-001, CERN, Geneva (2022).

[185] CMS collaboration, *CMS Luminosity Calibration for the pp Reference Run at $\sqrt{s}$ = 5.02 TeV*, Tech. Rep. CMS-PAS-LUM-16-001, CERN, Geneva (2016).

[186] CMS collaboration, *Luminosity measurement in proton-proton collisions at 5.02 TeV in 2017 at CMS*, Tech. Rep. CMS-PAS-LUM-19-001, CERN, Geneva (2021).

[187] CMS collaboration, *CMS luminosity measurement using 2016 proton-nucleus collisions at nucleon-nucleon center-of-mass energy of* 8.16 *TeV*, Tech. Rep. CMS-PAS-LUM-17-002, CERN, Geneva (2018).

[188] A. Kornmayer, *The CMS pixel luminosity telescope*, *Nucl. Instrum. Meth. A* **824** (2016) 304.

[189] K. Rose, *The new pixel luminosity telescope of CMS at the LHC*, in the proceedings of the *IEEE Nuclear Science Symposium and Medical Imaging Conference*, San Diego CA, U.S.A., 31 October–7 November 2015 [`DOI:10.1109/NSSMIC.2015.7581875`].

[190] P. Lujan, *Performance of the Pixel Luminosity Telescope for Luminosity Measurement at CMS during Run2*, *PoS* **EPS-HEP2017** (2017) 504.

[191] CMS BRIL collaboration, *The Pixel Luminosity Telescope: a detector for luminosity measurement at CMS using silicon pixel sensors*, *Eur. Phys. J. C* **83** (2023) 673 [`arXiv:2206.08870`].

[192] N. Karunarathna, *Run 3 luminosity measurements with the Pixel Luminosity Telescope*, *PoS* **ICHEP2022** (2022) 936.

[193] CMS collaboration, *Pixel luminosity telescope (PLT) calibration and first measurements*, CMS Detector Performance Note CMS-DP-2015-025 (2015).

[194] G. Bolla et al., *Sensor development for the CMS pixel detector*, *Nucl. Instrum. Meth. A* **485** (2002) 89.

[195] Y. Allkofer et al., *Design and performance of the silicon sensors for the CMS barrel pixel detector*, *Nucl. Instrum. Meth. A* **584** (2008) 25 [`physics/0702092`].

[196] H.C. Kastli et al., *Design and performance of the CMS pixel detector readout chip*, *Nucl. Instrum. Meth. A* **565** (2006) 188 [`physics/0511166`].

[197] M. Barbero, *Development of a radiation-hard pixel read out chip with trigger capability*, Ph.D. thesis, Universität Basel, Basel, Switzerland (2003).

[198] E. Bartz, *The 0.25um Token Bit Manager Chip for the CMS Pixel Readout*, in the proceedings of the *11th Workshop on Electronics for LHC and Future Experiments*, Heidelberg, Germany, 12–16 September 2005, pp. 153–157, http://cds.cern.ch/record/920426 [`DOI:10.5170/CERN-2005-011.153`].

[199] M. Pernicka et al., *The CMS Pixel FED*, in the proceedings of the *Topical Workshop on Electronics for Particle Physics*, Prague, Czech Republic, 3–7 September 2007, pp. 487–491, http://cds.cern.ch/record/1091743 [`DOI:10.5170/CERN-2007-007.487`].

[200] E. van der Bij, R.A. McLaren, O. Boyle and G. Rubin, *S-LINK, a data link interface specification for the LHC era*, *IEEE Trans. Nucl. Sci.* **44** (1997) 398.

[201] A.A. Zagozdzinska et al., *New Fast Beam Conditions Monitoring (BCM1F) system for CMS*, 2016 *JINST* **11** C01088.

[202] M. Guthoff, *The new Fast Beam Condition Monitor using poly-crystalline diamond sensors for luminosity measurement at CMS*, *Nucl. Instrum. Meth. A* **936** (2019) 717.

[203] J. Wańczyk, *Upgraded CMS Fast Beam Condition Monitor for LHC Run 3 Online Luminosity and Beam Induced Background Measurements*, *JACoW* **IBIC2022** (2022) 540.

[204] A.G. Delannoy Sotomayor and J. Wańczyk, *CMS BRIL BCM1F c-shape*, CMS Collection CMS-PHO-GEN-2022-001 (2022).

[205] D. Przyborowski, J. Kaplon and P. Rymaszewski, *Design and Performance of the BCM1F Front End ASIC for the Beam Condition Monitoring System at the CMS Experiment*, *IEEE Trans. Nucl. Sci.* **63** (2016) 2300.

[206] M. Friedl, *Analog optohybrids, CMS tracker TOB/TEC*, Tech. Rep. CMS-TK-CS-0002, CERN, Geneva (2004).

[207] A. Dabrowski, *A real-time histogramming unit for luminosity measurements of each bunch crossing at CMS*, talk given at the *Topical Workshop on Electronics for Particle Physics*, Perugia, Italy, 23–27 September 2013 [DESY-2013-00940].

[208] A.A. Zagozdzinska et al., *The fast beam condition monitor BCM1F backend electronics upgraded MicroTCA based architecture*, *Proc. SPIE* **9290** (2014) 92902L.

[209] A.A. Zagozdzinska, *The CMS fast beams condition monitor back-end electronics based on MicroTCA technology*, in the proceedings of the *IEEE Nuclear Science Symposium and Medical Imaging Conference*, San Diego CA, U.S.A., 31 October–7 November 2015 [`DOI:10.1109/NSSMIC.2015.7581987`].

[210] A.J. Rüde, *New Peak Finding Algorithm for the BCM1F Detector of the CMS Experiment at CERN*, M.Sc. thesis, Ernst-Abbe-Hochschule, Jena, Germany (2018), CERN-THESIS-2018-021.

[211] CMS collaboration, *BRIL luminosity performance plots: Cross-detector stability in early Run 3 data*, Tech. Rep. CMS-DP-2022-038 CERN, Geneva (2022).

[212] CMS collaboration, *The Phase-2 Upgrade of the CMS Beam Radiation, Instrumentation, and Luminosity Detectors: Conceptual Design*, Tech. Rep. CMS-NOTE-2019-008, CERN, Geneva (2020).

[213] A. Triossi, *The CMS Barrel Muon trigger upgrade*, 2017 *JINST* **12** C01095.

[214] J. Salfeld-Nebgen and D. Marlow, *Data-Driven Precision Luminosity Measurements with Z Bosons at the LHC and HL-LHC*, 2018 *JINST* **13** P12016 [`arXiv:1806.02184`].

[215] CMS collaboration, *Luminosity determination using Z boson production at the CMS experiment*, `arXiv:2309.01008`.

[216] CMS collaboration, *Luminosity monitoring with Z counting in early 2022 data*, Tech. Rep. CMS-DP-2023-003 CERN, Geneva (2023).

[217] CMS collaboration, *First measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s}$ = 13.6 TeV*, *JHEP* **08** (2023) 204 [`arXiv:2303.10680`].

[218] CMS collaboration, *CMS Z(μμ) yields for comparisons with ATLAS*, Tech. Rep. CMS-DP-2012-014 CERN, Geneva (2012).

[219] C. Schwick and B. Petersen, *LPC's view on Run 2*, in the proceedings of the *9th Evian Workshop on LHC beam operation*, Evian Les Bains, France, 30 January–1 February 2019, pp. 27–33, http://cds.cern.ch/record/2750274.

[220] S. Orfanelli et al., *A novel Beam Halo Monitor for the CMS experiment at the LHC*, 2015 *JINST* **10** P11011.

[221] N. Tosi et al., *Electronics and Calibration system for the CMS Beam Halo Monitor*, *PoS* **TIPP2014 (2014) 411**.

[222] N. Tosi, *The new Beam Halo Monitor for the CMS experiment at the LHC*, Ph.D. thesis, Università di Bologna, Bologna, Italy (2015), [CERN-THESIS-2015-283] [DOI:10.6092/unibo/amsdottorato/6767].

[223] S. Müller, *The Beam Condition Monitor 2 and the Radiation Environment of the CMS Detector at the LHC*, Ph.D. thesis, KIT, Karlsruhe, Germany (2011) [CERN-THESIS-2011-085] [DOI:10.5445/IR/1000022456].

[224] M. Guthoff, *Radiation Damage to the diamond-based Beam Condition Monitor of the CMS Detector at the LHC*, Ph.D. thesis, KIT, Karlsruhe, Germany (2014) [CERN-THESIS-2014-216] [DOI:10.5445/IR/1000044731].

[225] F.R. Kassel, *The Rate Dependent Radiation Induced Signal Degradation of Diamond Detectors*, Ph.D. thesis, KIT, Karlsruhe, Germany (2017) [CERN-THESIS-2017-071] [DOI:10.5445/IR/1000071068].

[226] B. Dehning et al., *LHC beam loss monitor system design*, *AIP Conf. Proc.* **648** (2002) 229.

[227] J. Emery et al., *Functional and Linearity test system for the LHC Beam Loss Monitoring data acquisition card*, in the proceedings of the *12th Workshop on Electronics for LHC and Future Experiments*, Valencia, Spain, 25–29 September 2006, pp. 447–451, http://cds.cern.ch/record/1057135 [DOI:10.5170/CERN-2007-001.447].

[228] C. Zamantzas, *The Real-Time Data Analysis and Decision System for Particle Flux Detection in the LHC Accelerator at CERN*, Ph.D. thesis, Brunel University, London, U.K. (2006), http://cds.cern.ch/record/976628CERN-THESIS-2006-037.

[229] B. Todd, A. Dinius, C. Martin and B. Puccio, *User interface to the beam interlock system*, Tech. Rep. EDMS Document 636589 v1.5 CERN, Geneva (2011).

[230] V.N. Kurlov, *Sapphire: Properties, growth, and applications*, in *Reference module: Material science and materials engineering*, Elsevier, Amsterdam, Netherlands (2016), ISBN 978-0128035818 [DOI:10.1016/B978-0-12-803581-8.03681-X]

[231] O. Karacheban et al., *Investigation of a direction sensitive sapphire detector stack at the 5 GeV electron beam at DESY-II*, 2015 *JINST* **10** P08008 [arXiv:1504.04023].

[232] CMS collaboration, *FLUKA Run 2 simulation benchmark with beam loss monitors in the CMS forward region*, Tech. Rep. CMS-DP-2021-008 CERN, Geneva (2021).

[233] A.M. Gribushin et al., *A neutron field monitoring system for collider experiments*, *Instrum. Exp. Tech.* **60** (2017) 167.

[234] G. Segura Millan, D. Perrin, and L. Scibile, , in the proceedings of the *10th International Conference*

on Accelerator and Large Experimental Physics Control Systems, Geneva, Switzerland, 10–15 October 2005 [*Conf. Proc. C* **051010** (2005) TH3B.1-3O].

[235] A. Ledeul et al., *CERN Supervision, Control and Data Acquisition System for Radiation and Environmental Protection*, in the proceedings of the *12ᵗʰ International Workshop on Personal Computers and Particle Accelerator Controls*, Hsinchu, Taiwan, 16–19 October 2018, pp. 248–252 [DOI:10.18429/JACoW-PCaPAC2018-FRCC3].

[236] G. Spiezia et al., *The LHC Radiation Monitoring System — RadMon*, *PoS* **RD11** (2011) 024.

[237] C. Martinella, *High energy hadrons fluence measurements in the LHC during 2015, 2016 and 2017 proton physics operations*, Tech. Rep. CERN-ACC-NOTE-2018-088, CERN, Geneva (2018).

[238] T.T. Böhlen et al., *The FLUKA Code: Developments and Challenges for High Energy and Medical Applications*, *Nucl. Data Sheets* **120** (2014) 211.

[239] J. Gutleber, S. Murray and L. Orsini, *Towards a homogeneous architecture for high-energy physics data acquisition systems*, *Comput. Phys. Commun.* **153** (2003) 155.

[240] B. Copy, E. Mandilara, I. Prieto Barreiro and F. Varela, *Monitoring of CERN's Data Interchange Protocol (DIP) system*, in the proceedings of the *16ᵗʰ International Conference on Accelerator and Large Experimental Physics Control Systems*, Barcelona, Spain, 8–13 October 2017, pp. 1797–1800 [DOI:10.18429/JACoW-ICALEPCS2017-THPHA162].

[241] Y. Fain and A. Moiseev, *Angular 2 development with TypeScript*, Manning Publications, Shelter Island, NY, U.S.A. (2016), ISBN 978-1617293122.

[242] J.M. Andre et al., *A Scalable Monitoring for the CMS Filter Farm Based on Elasticsearch*, *J. Phys. Conf. Ser.* **664** (2015) 082036.

[243] G. Bauer et al., *The CMS Event Builder and Storage System*, *J. Phys. Conf. Ser.* **219** (2010) 022038.

[244] G. Bauer et al., *The new CMS DAQ system for LHC operation after 2014 (DAQ2)*, *J. Phys. Conf. Ser.* **513** (2014) 012014.

[245] T. Bawej et al., *The New CMS DAQ System for Run-2 of the LHC*, *IEEE Trans. Nucl. Sci.* **62** (2015) 1099.

[246] G. Bauer et al., *The Terabit/s Super-Fragment Builder and Trigger Throttling System for the Compact Muon Solenoid Experiment at CERN*, *IEEE Trans. Nucl. Sci.* **55** (2008) 190.

[247] G. Bauer et al., *10 Gbps TCP/IP streams from the FPGA for the CMS DAQ eventbuilder network*, 2013 *JINST* **8** C12039.

[248] G. Bauer et al., *10 Gbps TCP/IP streams from the FPGA for High Energy Physics*, *J. Phys. Conf. Ser.* **513** (2014) 012042.

[249] D. Gigi et al., *The FEROL40, a microTCA card interfacing custom point-to-point links and standard TCP/IP*, *PoS* **TWEPP-17** (2017) 075.

[250] G. Bauer et al., *Upgrade of the CMS Event Builder*, *J. Phys. Conf. Ser.* **396** (2012) 012039.

[251] J.-M. André et al., *Performance of the new DAQ system of the CMS experiment for run-2*, in the proceedings of the *20ᵗʰ IEEE-NPSS Real Time Conference*, Padua, Italy, 5–10 June 2016 [DOI:10.1109/RTC.2016.7543164].

[252] J.M. Andre et al., *Performance of the CMS Event Builder*, *J. Phys. Conf. Ser.* **898** (2017) 032020.

[253] Juniper Networks, Inc., *QFX10000 modular Ethernet switches*, Datasheet 1000529-022-EN, 2021, https://www.juniper.net/us/en/products/switches/qfx-series/qfx10000-modular-ethernet-switches-datasheet.html.

[254] IEEE Standards Association, *IEEE 802.3-2012, Standard for Ethernet*, 2012, https://standards.ieee.org/ieee/802.3/5084/.

[255] *tmpfs*, Software available in the Linux kernel, 2021, https://man7.org/linux/man-pages/man5/tmpfs.5.html.

[256] C.D. Jones et al., *Using the CMS Threaded Framework In A Production Environment*, *J. Phys. Conf. Ser.* **664** (2015) 072026.

[257] J.M. Andre et al., *File-Based Data Flow in the CMS Filter Farm*, *J. Phys. Conf. Ser.* **664** (2015) 082033.

[258] *inotify*, software available in the Linux kernel, https://man7.org/linux/man-pages/man7/inotify.7.html.

[259] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.

[260] P. Deutsch and J.-L. Gailly, *ZLIB compressed data format specification version 3.3*, RFC Informational 1950, 1996 [`DOI:10.17487/RFC1950`].

[261] *XZ Utils* https://github.com/tukaani-project/xz. More information at https://tukaani.org/xz/

[262] Y. Collet and M. Kucherawy, *Zstandard compression and the 'application/zstd' media type*, RFC Informational 8878, 2021 [`DOI:10.17487/RFC8878`].

[263] *Elasticsearch*, https://github.com/elastic/elasticsearch. More information at https://www.elastic.co/elasticsearch/.

[264] T. Bray, *The JavaScript object notation (JSON) data interchange format*, RFC Proposed Standard 7159, 2014 [`DOI:10.17487/RFC7159`].

[265] M. Michelotto et al., *A comparison of HEP code with SPEC benchmarks on multi-core worker nodes*, *J. Phys. Conf. Ser.* **219** (2010) 052009.

[266] Nvidia Corporation, *NVIDIA T4 tensor core GPU*, Datasheet, 2019, https://www.nvidia.com/en-us/data-center/tesla-t4/.

[267] A.J. Peters, E.A. Sindrilaru and G. Adde, *EOS as the present and future solution for data storage at CERN*, *J. Phys. Conf. Ser.* **664** (2015) 042042.

[268] J.M. Andre et al., *Online Data Handling and Storage at the CMS Experiment*, *J. Phys. Conf. Ser.* **664** (2015) 082009.

[269] DataDirect Networks, *EXAScaler product family*, DDN Data Sheet, 2023, https://www.ddn.com/exascaler.

[270] J. Hegeman et al., *The CMS Timing and Control Distribution System*, in the proceedings of the *IEEE Nuclear Science Symposium and Medical Imaging Conference*, San Diego CA, U.S.A., 31 October–7 November 2015 [`DOI:10.1109/NSSMIC.2015.7581984`].

[271] CentOS Project, *CentOS-7 (2009)*, Release Notes, 2020, https://wiki.centos.org/action/show/Manuals/ReleaseNotes/CentOS7.2009.

[272] Red Hat, Inc., *Red Hat Enterprise Linux 8.8*, Release Notes, 2023, https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/8.

[273] R. García Leiva et al., *Quattor: Tools and techniques for the configuration, installation and management of large-scale grid computing fabrics*, *J. Grid Comput.* **2** (2004) 313.

[274] *Puppet*, https://github.com/puppetlabs/puppet. More information at https://www.puppet.com/.

[275] *oVirt*, https://github.com/ovirt/. More information at https://www.ovirt.org/.

[276] J.-M. Andre et al., *Experience with dynamic resource provisioning of the CMS online cluster using a cloud overlay*, *EPJ Web Conf.* **214** (2019) 07017.

[277] G. Bauer et al., *First operational experience with a high-energy physics run control system based on Web technologies*, *IEEE Trans. Nucl. Sci.* **59** (2012) 1597.

[278] G. Bauer et al., *A comprehensive zero-copy architecture for high performance distributed data acquisition over advanced network technologies for the CMS experiment*, in the proceedings of the *18th IEEE-NPSS Real-Time Conference*, Berkeley CA, U.S.A., 11–15 June 2012 [DOI:10.1109/RTC.2012.6418171].

[279] T. Bawej et al., *Achieving High Performance With TCP Over 40 GbE on NUMA Architectures for CMS Data Acquisition*, *IEEE Trans. Nucl. Sci.* **62** (2015) 1091.

[280] G. Bauer et al., *Distributed Error and Alarm Processing in the CMS Data Acquisition System*, *J. Phys. Conf. Ser.* **396** (2012) 012038.

[281] G. Bauer et al., *Automating the CMS DAQ*, *J. Phys. Conf. Ser.* **513** (2014) 012031.

[282] J.M. Andre et al., *New Operator Assistance Features in the CMS Run Control System*, *J. Phys. Conf. Ser.* **898** (2017) 032028.

[283] J.-M. Andre et al., *DAQExpert — An expert system to increase CMS data-taking efficiency*, *J. Phys. Conf. Ser.* **1085** (2018) 032021.

[284] J.-M. Andre et al., *Operational experience with the new CMS DAQ-Expert*, *EPJ Web Conf.* **214** (2019) 01015.

[285] G. Badaro et al., *DAQExpert the service to increase CMS data-taking efficiency*, *EPJ Web Conf.* **245** (2020) 01028.

[286] J.A. Lopez-Perez et al., *The Web Based Monitoring project at the CMS experiment*, *J. Phys. Conf. Ser.* **898** (2017) 092040.

[287] J. Duarte et al., *Fast inference of deep neural networks in FPGAs for particle physics*, 2018 *JINST* **13** P07027 [arXiv:1804.06913].

[288] R. Bainbridge, *Recording and reconstructing 10 billion unbiased b hadron decays in CMS*, *EPJ Web Conf.* **245** (2020) 01025.

[289] F.M.A. Erich, C. Amrit and M. Daneva, *A qualitative study of DevOps usage in practice*, *J. Softw. Evol. Process* **29** (2017) e1885.

[290] T. Dingsøyr, S. Nerur, V. Balijepally, and N. B. Moe, *A decade of agile methodologies: Towards explaining agile software development*, *J. Syst. Softw.* **85** (2012) 1213.

[291] *Prometheus*, https://github.com/prometheus/prometheus. More information at https://prometheus.io/.

[292] S.H. Sunil Kumar and C. Saravanan, *A comprehensive study on data visualization tool: Grafana*, *JETIR* **8** (2021) f908.

[293] A.X. Ming Chang et al., *Deep neural networks compiler for a trace-based accelerator*, *J. Syst. Archit.* **102** (2020) 101659.

[294] Xilinx Inc., *VCU128 evaluation board*, User Guide UG1302 v1.2, 2022, https://docs.xilinx.com/r/en-US/ug1302-vcu128-eval-bd.

[295] CMS collaboration, *The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger*, Tech. Rep. CERN-LHCC-2021-007, CERN, Geneva (2021).

[296] D. Golubovic et al., *40 MHz Scouting with Deep Learning in CMS*, in the proceedings of the *6th Internation Workshop Connecting the Dots*, Princeton NJ, U.S.A., 20–30 April 2020 https://zenodo.org/record/5040675 [DOI:10.5281/ZENODO.5040675].

[297] CMS collaboration, *40 MHz Scouting with Deep Learning in CMS*, Tech. Rep. CMS-DP-2022-066 CERN, Geneva (2022).

[298] R. Ardino et al., *A 40 MHz Level-1 trigger scouting system for the CMS Phase-2 upgrade*, *Nucl. Instrum. Meth. A* **1047** (2023) 167805.

[299] A. Bocci et al., *Heterogeneous Reconstruction of Tracks and Primary Vertices With the CMS Pixel Tracker*, *Front. Big Data* **3** (2020) 601728 [arXiv:2008.13461].

[300] CMS collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, 2017 *JINST* **12** P10003 [arXiv:1706.04965].

[301] CMS collaboration, *Performance of the CMS muon trigger system in proton-proton collisions at* $\sqrt{s} = 13\,TeV$, 2021 *JINST* **16** P07001 [arXiv:2102.04790].

[302] W. Adam, R. Fruhwirth, A. Strandlie and T. Todorov, *Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC*, *J. Phys. G* **31** (2005) N9 [physics/0306087].

[303] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189].

[304] M. Cacciari, G.P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896 [arXiv:1111.6097].

[305] CMS collaboration, *Performance of tau-lepton reconstruction and identification in CMS*, 2012 *JINST* **7** P01001 [arXiv:1109.6034].

[306] CMS collaboration, *Identification of hadronic tau lepton decays using a deep neural network*, 2022 *JINST* **17** P07023 [arXiv:2201.08458].

[307] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146 [arXiv:1402.2657].

[308] CMS collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, 2018 *JINST* **13** P05011 [arXiv:1712.07158].

[309] E. Bols et al., *Jet Flavour Classification Using DeepJet*, 2020 *JINST* **15** P12012 [arXiv:2008.10519].

[310] H. Qu and L. Gouskos, *ParticleNet: Jet Tagging via Particle Clouds*, *Phys. Rev. D* **101** (2020) 056019 [arXiv:1902.08570].

[311] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323].

[312] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in the proceedings of the *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, Hamburg, Germany, 27–30 April 1998, p. 270–279 [hep-ph/9907280].

[313] CMS collaboration, *Search for Narrow Resonances and Quantum Black Holes in Inclusive and b-Tagged Dijet Mass Spectra from pp Collisions at* $\sqrt{s}$ = 7 *TeV*, *JHEP* **01** (2013) 013 [arXiv:1210.2387].

[314] CMS collaboration, *Search for narrow resonances in dijet final states at* $\sqrt{s}$ = 8 *TeV with the novel CMS technique of data scouting*, *Phys. Rev. Lett.* **117** (2016) 031802 [arXiv:1604.08907].

[315] CMS collaboration, *Search for a Narrow Resonance Lighter than 200 GeV Decaying to a Pair of Muons in Proton-Proton Collisions at* $\sqrt{s}$ = 13 *TeV*, *Phys. Rev. Lett.* **124** (2020) 131802 [arXiv:1912.04776].

[316] CMS collaboration, *Search for pair-produced three-jet resonances in proton-proton collisions at* $\sqrt{s}$ = 13 *TeV*, *Phys. Rev. D* **99** (2019) 012010 [arXiv:1810.10092].

[317] GEANT4 collaboration, *GEANT4 — a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[318] D.J. Lange, M. Hildreth, V.N. Ivantchenko and I. Osborne, *Upgrades for the CMS simulation*, *J. Phys. Conf. Ser.* **608** (2015) 012056.

[319] S. Sekmen, *Recent Developments in CMS Fast Simulation*, *PoS* **ICHEP2016** (2016) 181 [arXiv:1701.03850].

[320] K. Pedro, *Current and future performance of the CMS simulation*, *EPJ Web Conf.* **214** (2019) 02036.

[321] C. Caputo, *Enabling continuous speedup of CMS Event Reconstruction through continuous benchmarking*, in the proceedings of the *21*[th] *International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, Bari, Italy, 24–28 October 2022 [CMS-CR-2023-038].

[322] D. Piparo, *Automated quality monitoring and validation of the CMS reconstruction software*, *J. Phys. Conf. Ser.* **368** (2012) 012008.

[323] M.N. Rodozov and D. Lange, *Modernizing the CMS software stack*, https://zenodo.org/record/3599193 [DOI:10.5281/ZENODO.3599193].

[324] J. Blomer, C. Aguado Sanchez, P. Buncic and A. Harutyunyan, *Distributing LHC application software and conditions databases using the CernVM file system*, *J. Phys. Conf. Ser.* **331** (2011) 042003.

[325] J. Blomer et al., *The CernVM File System: v2.7.5*, https://zenodo.org/record/4114078 [DOI:10.5281/ZENODO.4114078].

[326] C.D. Jones and E. Sexton-Kennedy, *Stitched Together: Transitioning CMS to a Hierarchical Threaded Framework*, *J. Phys. Conf. Ser.* **513** (2014) 022034.

[327] *oneAPI Threading Building Blocks (oneTBB)*, https://github.com/oneapi-src/oneTBB. More information at https://oneapi-src.github.io/oneTBB/.

[328] I. Bird et al., *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, Tech. Rep. CERN-LHCC-2005-024, CERN, Geneva (2005).

[329] C.D. Jones, *CMS event processing multi-core efficiency status*, *J. Phys. Conf. Ser.* **898** (2017) 042008.

[330] A. Bocci et al., *Bringing heterogeneity to the CMS software framework*, *EPJ Web Conf.* **245** (2020) 05009 [arXiv:2004.04334].

[331] E. Zenker et al., *Alpaka — An Abstraction Library for Parallel Kernel Acceleration*, in the proceedings of the *IEEE International Parallel and Distributed Processing Symposium Workshops*, Chicago IL, U.S.A., 23–27 May 2016 [arXiv:1602.08477] [DOI:10.1109/IPDPSW.2016.50].

[332] H.C. Edwards, C.R. Trott and D. Sunderland, *Kokkos: Enabling manycore performance portability through polymorphic memory access patterns*, *J. Parallel Distributed Comput.* **74** (2014) 3202.

[333] M.J. Kortelainen et al., *Porting CMS Heterogeneous Pixel Reconstruction to Kokkos*, *EPJ Web Conf.* **251** (2021) 03034 [arXiv:2104.06573].

[334] A. Bocci et al., *Performance portability for the CMS Reconstruction with Alpaka*, *J. Phys. Conf. Ser.* **2438** (2023) 012058.

[335] J. Duarte et al., *FPGA-accelerated machine learning inference as a service for particle physics computing*, *Comput. Softw. Big Sci.* **3** (2019) 13 [arXiv:1904.08986].

[336] D.S. Rankin et al., *FPGAs-as-a-Service Toolkit (FaaST)*, in the proceedings of the *IEEE/ACM International Workshop on Heterogeneous High-Performance Reconfigurable Computing*, Atlanta GA, U.S.A., 13 November 2020 [arXiv:2010.08556] [DOI:10.1109/H2RC51942.2020.00010].

[337] J. Krupa et al., *GPU coprocessors as a service for deep learning inference in high energy physics*, *Mach. Learn. Sci. Tech.* **2** (2021) 035005 [arXiv:2007.10359].

[338] Nvidia Corporation, *Triton inference server*, Release Notes RN-08995-001 v23.04, 2023, https://docs.nvidia.com/deeplearning/triton-inference-server/pdf/Triton-Inference-Server-Release-Notes.pdf.

[339] M. Case, A.T.M. Aerts, A. Muhammad and M. Liendl, *CMS Detector Description: New Developments*, in the proceedings of the *14th International Conference on Computing in High-Enery and Nuclear Physics*, Interlaken, Switzerland, 27 September–1 October 2004, http://cds.cern.ch/record/865635 [DOI:10.5170/CERN-2005-002.498].

[340] M. Frank, F. Gaede, C. Grefe and P. Mato, *DD4hep: A Detector Description Toolkit for High Energy Physics Experiments*, *J. Phys. Conf. Ser.* **513** (2014) 022010.

[341] G. Petrucciani, A. Rizzi and C. Vuosalo, *Mini-AOD: A New Analysis Data Format for CMS*, *J. Phys. Conf. Ser.* **664** (2015) 7 [arXiv:1702.04685].

[342] A. Rizzi, G. Petrucciani and M. Peruzzi, *A further reduction in CMS event data for analysis: the NANOAOD format*, *EPJ Web Conf.* **214** (2019) 06021.

[343] M. Hildreth, V.N. Ivanchenko and D.J. Lange, *Upgrades for the CMS simulation*, *J. Phys. Conf. Ser.* **898** (2017) 042040.

[344] M. Aderholz et al., *Models of Networked Analysis at Regional Centres for LHC Experiments (MONARC), Phase 2 Report, 24th March 2000*, Tech Rep. CERN-LCB-2000-001, CERN, Geneva (2000).

[345] D. Futyan, *Commissioning the CMS alignment and calibration framework*, *J. Phys. Conf. Ser.* **219** (2010) 032041.

[346] K. Bloom et al., *Any data, any time, anywhere: Global data access for science*, in the proceedings of the *IEEE/ACM 2nd International Symposium on Big Data Computing*, Limassol, Cyprus, 7–10 December 2015 [DOI:10.1109/BDC.2015.33].

[347] L. Bauerdick et al., *Using Xrootd to Federate Regional Storage*, *J. Phys. Conf. Ser.* **396** (2012) 042009.

[348] L.A.T. Bauerdick et al., *XRootd, disk-based, caching proxy for optimization of data access, data placement and data replication*, *J. Phys. Conf. Ser.* **513** (2014) 042044.

[349] G.M. Kurtzer, V. Sochat and M.W. Bauer, *Singularity: Scientific containers for mobility of compute*, *PLOS One* **12** (2017) e0177459.

[350] T. Boccali et al., *Dynamic Distribution of High-Rate Data Processing from CERN to Remote HPC Data Centers*, *Comput. Softw. Big Sci.* **5** (2021) 7.

[351] T. Boccali et al., *Extension of the INFN Tier-1 on a HPC system*, *EPJ Web Conf.* **245** (2020) 09009 [arXiv:2006.14603].

[352] C. Acosta-Silva et al., *Exploitation of network-segregated CPU resources in CMS*, *EPJ Web Conf.* **251** (2021) 02020.

[353] M. Fischer et al., *Effective Dynamic Integration and Utilization of Heterogenous Compute Resources*, *EPJ Web Conf.* **245** (2020) 07038.

[354] T. Boccali et al., *Enabling CMS Experiment to the utilization of multiple hardware architectures: a Power9 Testbed at CINECA*, *J. Phys. Conf. Ser.* **2438** (2023) 012031.

[355] D. Hufnagel et al., *HPC resource integration into CMS Computing via HEPCloud*, *EPJ Web Conf.* **214** (2019) 03031.

[356] M.C. Davis et al., *CERN Tape Archive - from development to production deployment*, *EPJ Web Conf.* **214** (2019) 04015.

[357] J. Rehn et al., "PhEDEx high-throughput data transfer management system", in the proceedings of the *15th International Conference on Computing in High Energy and Nuclear Physics*, Mumbai, India, 13–17 February 2006, https://indico.cern.ch/event/408139/contributions/979918/attachments/815907/1118010/CHEP06_PhEDEx_overview.pdf

[358] Y. Iiyama et al., *Dynamo: Handling Scientific Data Across Sites and Storage Media*, *Comput. Softw. Big Sci.* **5** (2021) 11 [arXiv:2003.11409].

[359] I. Béjar Alonso et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, Tech. Rep. CERN-2020-010 CERN, Geneva (2020) [DOI:10.23731/CYRM-2020-0010].

[360] E. Karavakis et al., *FTS improvements for LHC Run-3 and beyond*, *EPJ Web Conf.* **245** (2020) 04016.

[361] M. Barisits et al., *Rucio - Scientific data management*, *Comput. Softw. Big Sci.* **3** (2019) 11 [arXiv:1902.09857].

[362] *Helm: the package manager for Kubernetes*, https://github.com/helm/helm. More information at https://helm.sh/.

[363] *Kubernetes (K8s)*, https://github.com/kubernetes/kubernetes. More information at https://kubernetes.io/.

[364] *Docker*, https://github.com/docker. More information at https://www.docker.com/.

[365] I. Foster, C. Kesselman, G. Tsudik and S. Tuecke, "A security architecture for computational grids", in the proceedings of the *5th ACM conference on Computer and Communications Security*, San Francisco CA, U.S.A., 2–5 November 1998, [DOI:10.1145/288090.288111].

[366] R. Butler et al., *A national-scale authentication infrastructure*, *Computer* **33** (2000) 60.

[367] L. Dusseault, *HTTP extensions for web distributed authoring and versioning (WebDAV)*, RFC Proposed Standard 4918, 2007 [DOI:10.17487/RFC4918].

[368] X. Espinal et al., *The Quest to solve the HL-LHC data access puzzle*, *EPJ Web Conf.* **245** (2020) 04027.

[369] I. Sfiligoi et al., *The pilot way to Grid resources using glideinWMS*, in the proceedings of the *WRI World Congress on Computer Science and Information Engineering* Los Angeles CA, U.S.A., 31 March–2 April 2009, pp. 428–432 [DOI:10.1109/CSIE.2009.950].

[370] D. Thain, T. Tannenbaum, and M. Livny, *Distributed computing in practice: the Condor experience*, *Concurr. Comput.* **17** (2005) 323.

[371] J. Balcas et al., *Using the GlideinWMS System as a Common Resource Provisioning Layer in CMS*, *J. Phys. Conf. Ser.* **664** (2015) 062031.

[372] A. Pérez-Calero Yzquierdo et al., *Evolution of the CMS Global Submission Infrastructure for the HL-LHC Era*, *EPJ Web Conf.* **245** (2020) 03016.

[373] D. Spiga et al., *Exploiting private and commercial clouds to generate on-demand CMS computing facilities with DODAS*, *EPJ Web Conf.* **214** (2019) 07027.

[374] D.P. Anderson, *BOINC: A platform for volunteer computing*, *J. Grid Comput.* **18** (2020) 99.

[375] D. Smith et al., *Sharing server nodes for storage and compute*, *EPJ Web Conf.* **214** (2019) 08025.

[376] D. Giordano et al., *Cern-It Evaluation Of Microsoft Azure Cloud Iaas*, https://zenodo.org/record/48495 [`DOI:10.5281/ZENODO.48495`].

[377] C. Cordeiro et al., *CERN computing in commercial clouds*, *J. Phys. Conf. Ser.* **898** (2017) 082030.

[378] A. Perez-Calero Yzquierdo et al., *CMS Readiness for Multi-Core Workload Scheduling*, *J. Phys. Conf. Ser.* **898** (2017) 052030.

[379] B.P. Bockelman et al., *Improving the Scheduling Efficiency of a Global Multi-Core HTCondor Pool in CMS*, *EPJ Web Conf.* **214** (2019) 03056.

[380] A. Pérez-Calero Yzquierdo et al., *Reaching new peaks for the future of the CMS HTCondor Global Pool*, *EPJ Web Conf.* **251** (2021) 02055.

[381] P. Couvares et al., *Workflow management in Condor*, in *Workflows for e-Science*, Springer Nature, London, U.K., 2007. ISBN 978-1846285196 [`DOI:10.1007/978-1-84628-757-2_22`].

[382] J. Balcas et al., *CMS Connect*, *J. Phys. Conf. Ser.* **898** (2017) 082032.

[383] V. Kuznetsov, D. Evans and S. Metson, *The CMS data aggregation system*, in the proceedings of the *10th International Conference on Computational Science*, Amsterdam, Netherlands, 31 May–2 June 2010 [`DOI:10.1016/j.procs.2010.04.172`].

[384] M. Imran et al., *Migration of CMSWEB cluster at CERN to Kubernetes: a comprehensive study*, *Clust. Comput.* **24** (2021) 3085.

[385] WLCG Authorization WG, *WLCG Token Transition Timeline*, https://zenodo.org/record/7014668 [`DOI:10.5281/ZENODO.7014668`].

[386] C. Ariza-Porras, V. Kuznetsov and F. Legger, *The CMS monitoring infrastructure and applications*, *Comput. Softw. Big Sci.* **5** (2021) 5 [`arXiv:2007.03630`].

[387] A. Aimar et al., *Unified Monitoring Architecture for IT and grid services*, *J. Phys. Conf. Ser.* **898** (2017) 092033.

The CMS collaboration

Yerevan Physics Institute, Yerevan, Armenia

A. Hayrapetyan, A. Tumasyan [1]

Institut für Hochenergiephysik, Vienna, Austria

W. Adam, J.W. Andrejkovic, B. Arnold, H. Bergauer, T. Bergauer, S. Chatterjee, K. Damanakis,
M. Dragicevic, A. Escalante Del Valle, P.S. Hussain, M. Jeitler [2], N. Krammer, D. Liko,
I. Mikulec, J. Schieck [2], R. Schöfbeck, D. Schwarz, M. Sonawane, S. Templ,
W. Waltenberger, C.-E. Wulz [2]

Universiteit Antwerpen, Antwerpen, Belgium

M.R. Darwish [3], T. Janssen, P. Van Mechelen

Vrije Universiteit Brussel, Brussel, Belgium

E.S. Bols, J. D'Hondt, S. Dansana, A. De Moor, M. Delcourt, H. El Faham, S. Lowette,
I. Makarenko, A. Morton, D. Müller, A.R. Sahasransu, S. Tavernier, M. Tytgat [4],
S. Van Putte, D. Vannerom

Université Libre de Bruxelles, Bruxelles, Belgium

B. Clerbaux, G. De Lentdecker, L. Favart, D. Hohov, J. Jaramillo, A. Khalilzadeh, K. Lee,
M. Mahdavikhorrami, A. Malara, S. Paredes, L. Pétré, N. Postiau, L. Thomas,
M. Vanden Bemden, C. Vander Velde, P. Vanlaer

Ghent University, Ghent, Belgium

M. De Coen, D. Dobur, J. Knolle, L. Lambrecht, G. Mestdach, C. Rendón, A. Samalan,
K. Skovpen, N. Van Den Bossche, B. Vermassen, L. Wezenbeek

Université Catholique de Louvain, Louvain-la-Neuve, Belgium

A. Benecke, G. Bruno, C. Caputo, C. Delaere, I.S. Donertas, A. Giammanco, K. Jaffel,
Sa. Jain, V. Lemaitre, J. Lidrych, P. Mastrapasqua, K. Mondal, T.T. Tran, S. Wertz

Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil

G.A. Alves, E. Coelho, C. Hensel, T. Menezes De Oliveira, A. Moraes, P. Rebello Teles,
M. Soeiro

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil

W.L. Aldá Júnior, M. Alves Gallo Pereira, M. Barroso Ferreira Filho, H. Brandao Malbouisson,
W. Carvalho, J. Chinellato [5], E.M. Da Costa, G.G. Da Silveira [6], D. De Jesus Damiao,
S. Fonseca De Souza, J. Martins [7], C. Mora Herrera, K. Mota Amarilo, L. Mundim,
H. Nogima, A. Santoro, S.M. Silva Do Amaral, A. Sznajder, M. Thiel, A. Vilela Pereira

Universidade Estadual Paulista, Universidade Federal do ABC, São Paulo, Brazil

C.A. Bernardes [6], L. Calligaris, T.R. Fernandez Perez Tomei, E.M. Gregores, P.G. Mercadante,
S.F. Novaes, B. Orzari, Sandra S. Padula

Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria

A. Aleksandrov [image], G. Antchev [image], R. Hadjiiska [image], P. Iaydjiev [image], M. Misheva [image], M. Shopova [image], G. Sultanov [image]

University of Sofia, Sofia, Bulgaria

A. Dimitrov [image], T. Ivanov [image], L. Litov [image], B. Pavlov [image], P. Petkov [image], A. Petrov [image], E. Shumka [image]

Instituto De Alta Investigación, Universidad de Tarapacá, Casilla 7 D, Arica, Chile

S. Keshri [image], S. Thakur [image]

Beihang University, Beijing, China

T. Cheng [image], Q. Guo, T. Javaid [image], M. Mittal [image], L. Yuan [image]

Department of Physics, Tsinghua University, Beijing, China

G. Bauer[8], Z. Hu [image], K. Yi [image][8,9]

Institute of High Energy Physics, Beijing, China

G.M. Chen [image][10], H.S. Chen [image][10], M. Chen [image][10], F. Iemmi [image], C.H. Jiang, A. Kapoor [image], H. Liao [image], Z.-A. Liu [image][11], F. Monti [image], R. Sharma [image], J.N. Song[11], J. Tao [image], J. Wang [image], H. Zhang [image]

State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China

A. Agapitos [image], Y. Ban [image], A. Levin [image], C. Li [image], Q. Li [image], X. Lyu, Y. Mao, S.J. Qian [image], X. Sun [image], D. Wang [image], H. Yang, C. Zhou [image]

Sun Yat-Sen University, Guangzhou, China

Z. You [image]

University of Science and Technology of China, Hefei, China

N. Lu [image]

Institute of Modern Physics and Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) - Fudan University, Shanghai, China

X. Gao [image][12], D. Leggat, H. Okawa [image], Y. Zhang [image]

Zhejiang University, Hangzhou, Zhejiang, China

Z. Lin [image], C. Lu [image], M. Xiao [image]

Universidad de Los Andes, Bogota, Colombia

C. Avila [image], D.A. Barbosa Trujillo, A. Cabrera [image], C. Florez [image], J. Fraga [image], J.A. Reyes Vega

Universidad de Antioquia, Medellin, Colombia

J. Mejia Guisao [image], F. Ramirez [image], M. Rodriguez [image], J.D. Ruiz Alvarez [image]

University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia

D. Giljanovic [image], N. Godinovic [image], D. Lelas [image], A. Sculac [image]

University of Split, Faculty of Science, Split, Croatia

M. Kovac [iD], T. Sculac [iD]

Institute Rudjer Boskovic, Zagreb, Croatia

P. Bargassa [iD], V. Brigljevic [iD], B.K. Chitroda [iD], D. Ferencek [iD], S. Mishra [iD], A. Starodumov[13], T. Susa [iD]

University of Cyprus, Nicosia, Cyprus

A. Attikis [iD], K. Christoforou [iD], S. Konstantinou [iD], J. Mousa [iD], C. Nicolaou, F. Ptochos [iD], P.A. Razis [iD], H. Rykaczewski, H. Saka [iD], A. Stepennov [iD]

Charles University, Prague, Czech Republic

M. Finger [iD], M. Finger Jr. [iD], A. Kveton [iD]

Escuela Politecnica Nacional, Quito, Ecuador

E. Ayala [iD]

Universidad San Francisco de Quito, Quito, Ecuador

A. Cardenas [iD], E. Carrera Jarrin [iD], D. Cazar Ramírez [iD], E.F. Mendez Garces, X. Riofrio [iD]

Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt

Y. Assran[14,15], S. Elgammal[15]

Center for High Energy Physics (CHEP-FU), Fayoum University, El-Fayoum, Egypt

M. Abdullah Al-Mashad [iD], M.A. Mahmoud [iD]

National Institute of Chemical Physics and Biophysics, Tallinn, Estonia

R.K. Dewanjee [iD][16], K. Ehataht [iD], M. Kadastik, T. Lange [iD], S. Nandan [iD], C. Nielsen [iD], J. Pata [iD], M. Raidal [iD], L. Tani [iD], C. Veelken [iD]

Department of Physics, University of Helsinki, Helsinki, Finland

H. Kirschenmann [iD], K. Osterberg [iD], M. Voutilainen [iD]

Helsinki Institute of Physics, Helsinki, Finland

S. Bharthuar [iD], E. Brücken [iD], F. Garcia [iD], J. Havukainen [iD], K.T.S. Kallonen [iD], M.S. Kim [iD], R. Kinnunen, P.P. Koponen, T. Lampén [iD], K. Lassila-Perini [iD], S. Lehti [iD], T. Lindén [iD], M. Lotti, L. Martikainen [iD], M. Myllymäki [iD], M.m. Rantanen [iD], H. Siikonen [iD], E. Tuominen [iD], J. Tuominiemi [iD], R. Turpeinen

Lappeenranta-Lahti University of Technology, Lappeenranta, Finland

P. Luukka [iD], H. Petrow [iD], T. Tuuva[†]

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

M. Besancon [iD], F. Couderc [iD], M. Dejardin [iD], D. Denegri, J.L. Faure, F. Ferri [iD], S. Ganjour [iD], P. Gras [iD], G. Hamel de Monchenault [iD], V. Lohezic [iD], J. Malcles [iD], J. Rander, A. Rosowsky [iD], M.Ö. Sahin [iD], A. Savoy-Navarro [iD][17], P. Simkina [iD], M. Titov [iD]

Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France

C. Baldenegro Barrera [image], F. Beaudette [image], A. Buchot Perraguin [image], P. Busson [image], A. Cappati [image], C. Charlot [image], A. Chiron, F. Damas [image], O. Davignon [image], G. Falmagne [image], B.A. Fontana Santos Alves [image], S. Ghosh [image], A. Gilbert [image], R. Granier de Cassagnac [image], A. Hakimi [image], B. Harikrishnan [image], L. Kalipoliti [image], G. Liu [image], J. Motta [image], M. Nguyen [image], C. Ochando [image], L. Portales [image], T. Romanteau, R. Salerno [image], U. Sarkar [image], J.B. Sauvan [image], Y. Sirois [image], A. Tarabini [image], E. Vernazza [image], A. Zabi [image], A. Zghiche [image]

Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France

J.-L. Agram [image][18], J. Andrea [image], D. Apparu [image], D. Bloch [image], J.-M. Brom [image], E.C. Chabert [image], L. Charles, C. Collard [image], S. Falke [image], U. Goerlach [image], C. Grimault, L. Gross, R. Haeberle [image], A.-C. Le Bihan [image], M.A. Sessini [image], P. Van Hove [image]

Institut de Physique des 2 Infinis de Lyon (IP2I), Villeurbanne, France

G. Baulieu [image], S. Beauceron [image], B. Blancon [image], G. Boudoul [image], N. Chanon [image], J. Choi [image], D. Contardo [image], P. Depasse [image], C. Dozen [image][19], H. El Mamouni, J. Fay [image], S. Gascon [image], M. Gouzevitch [image], C. Greenberg, G. Grenier [image], B. Ille [image], I.B. Laktineh, M. Lethuillier [image], N. Lumb, L. Mirabito, S. Perries, D. Pugnere [image], F. Schirra, M. Vander Donckt [image], P. Verdier [image], J. Xiao [image]

Georgian Technical University, Tbilisi, Georgia

G. Adamov, I. Bagaturia [image][20], D. Chokheli [image], O. Kemularia, A. Khvedelidze [image][13], D. Lomidze [image], I. Lomidze [image], A. Melkadze, T. Toriashvili [image][21], Z. Tsamalaidze [image][13]

RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany

C. Autermann [image], V. Botta [image], L. Feld [image], W. Karpinski, M.K. Kiesel, K. Klein [image], M. Lipinski [image], D. Meuser [image], A. Pauls [image], G. Pierschel, M.P. Rauch, N. Röwert [image], C. Schomakers, J. Schulz, M. Teroerde [image], M. Wlochal

RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany

S. Diekmann [image], A. Dodonova [image], N. Eich [image], D. Eliseev [image], F. Engelke [image], M. Erdmann [image], P. Fackeldey [image], B. Fischer [image], T. Hebbeker [image], K. Hoepfner [image], F. Ivone [image], A. Jung [image], M.y. Lee [image], L. Mastrolorenzo, M. Merschmeyer [image], A. Meyer [image], S. Mukherjee [image], D. Noll [image], A. Novak [image], F. Nowotny, A. Pozdnyakov [image], Y. Rath, W. Redjeb [image], F. Rehm, H. Reithler [image], V. Sarkisovi [image], A. Schmidt [image], S.C. Schuler, A. Sharma [image], A. Stein [image], F. Torres Da Silva De Araujo [image][22], L. Vigilante, S. Wiedenbeck [image], S. Zaleski

RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany

C. Dziwok [image], G. Flügge [image], W. Haj Ahmad [image][23], T. Kress [image], A. Nowack [image], O. Pooth [image], A. Stahl [image], T. Ziemons [image], A. Zotz [image]

Deutsches Elektronen-Synchrotron, Hamburg, Germany

H. Aarup Petersen [image], M. Aldaya Martin [image], J. Alimena [image], S. Amoroso, Y. An [image], S. Baxter [image], M. Bayatmakou [image], H. Becerril Gonzalez [image], O. Behnke [image], A. Belvedere [image], S. Bhattacharya [image], F. Blekman [image][24], K. Borras [image][25], D. Brunner [image], A. Campbell [image], A. Cardini [image], C. Cheng, F. Colombina [image], S. Consuegra Rodríguez [image], G. Correia Silva [image], M. De Silva [image], G. Eckerlin, D. Eckstein [image], L.I. Estevez Banos [image], O. Filatov [image], E. Gallo [image][24], A. Geiser [image], A. Giraldi [image], G. Greau, V. Guglielmi [image], M. Guthoff [image], A. Hinzmann [image], A. Jafari [image][26], L. Jeppe [image], N.Z. Jomhari [image], H. Jung [image], B. Kaech [image],

M. Kasemann, H. Kaveh, C. Kleinwort, R. Kogler, M. Komm, D. Krücker, W. Lange, D. Leyva Pernia, K. Lipka[27], W. Lohmann[28], R. Mankel, I.-A. Melzer-Pellmann, M. Mendizabal Morentin, J. Metwally, A.B. Meyer, G. Milella, M. Mormile, A. Mussgiller, A. Nürnberg, Y. Otarid, D. Pérez Adán, E. Ranken, A. Raspereza, B. Ribeiro Lopes, J. Rübenach, A. Saggio, M. Scham[29,25], V. Scheurer, S. Schnake[25], P. Schütze, C. Schwanenberger[24], M. Shchedrolosiev, R.E. Sosa Ricardo, L.P. Sreelatha Pramod, D. Stafford, F. Vazzoler, A. Ventura Barroso, R. Walsh, Q. Wang, Y. Wen, K. Wichmann, L. Wiens[25], C. Wissing, S. Wuchterl, Y. Yang, A. Zimermmane Castro Santos

University of Hamburg, Hamburg, Germany

A. Albrecht, S. Albrecht, M. Antonello, S. Bein, L. Benato, M. Bonanomi, P. Connor, M. Eich, K. El Morabit, Y. Fischer, A. Fröhlich, C. Garbers, E. Garutti, A. Grohsjean, M. Hajheidari, J. Haller, H.R. Jabusch, G. Kasieczka, P. Keicher, R. Klanner, W. Korcari, T. Kramer, V. Kutzner, F. Labe, J. Lange, A. Lobanov, C. Matthies, A. Mehta, L. Moureaux, M. Mrowietz, A. Nigamova, Y. Nissan, A. Paasch, K.J. Pena Rodriguez, T. Quadfasel, B. Raciti, M. Rieger, D. Savoiu, J. Schindler, P. Schleper, M. Schröder, J. Schwandt, M. Sommerhalder, H. Stadie, G. Steinbrück, A. Tews, M. Wolf

Karlsruher Institut fuer Technologie, Karlsruhe, Germany

S. Brommer, M. Burkart, E. Butz, T. Chwalek, A. Dierlamm, A. Droll, N. Faltermann, M. Giffels, A. Gottmann, F. Hartmann[30], M. Horzela, U. Husemann, M. Klute, R. Koppenhöfer, M. Link, A. Lintuluoto, S. Maier, S. Mitra, Th. Müller, M. Neukum, M. Oh, G. Quast, K. Rabbertz, I. Shvetsov, H.J. Simonis, N. Trevisani, R. Ulrich, J. van der Linden, R.F. Von Cube, M. Wassmer, S. Wieland, F. Wittig, R. Wolf, S. Wunsch, X. Zuo

Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece

G. Anagnostou, P. Assiouras, G. Daskalakis, A. Kyriakis, A. Papadopoulos[30], A. Stakia

National and Kapodistrian University of Athens, Athens, Greece

D. Karasavvas, P. Kontaxakis, G. Melachroinos, A. Panagiotou, I. Papavergou, I. Paraskevas, N. Saoulidou, K. Theofilatos, E. Tziaferi, K. Vellidis, I. Zisopoulos

National Technical University of Athens, Athens, Greece

G. Bakas, T. Chatzistavrou, G. Karapostoli, K. Kousouris, I. Papakrivopoulos, E. Siamarkou, G. Tsipolitis, A. Zacharopoulou

University of Ioánnina, Ioánnina, Greece

K. Adamidis, I. Bestintzanos, I. Evangelou, C. Foudas, P. Gianneios, C. Kamtsikis, P. Katsoulis, P. Kokkas, P.G. Kosmoglou Kioseoglou, N. Manthos, I. Papadopoulos, J. Strologas

MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary

M. Csanád, K. Farkas, M.M.A. Gadallah[31], Á. Kadlecsik, P. Major, K. Mandal, G. Pásztor, A.J. Rádl[32], O. Surányi, G.I. Veres

Wigner Research Centre for Physics, Budapest, Hungary

M. Bartók [33], C. Hajdu, D. Horvath [34,35], F. Sikler, V. Veszpremi

Institute of Nuclear Research ATOMKI, Debrecen, Hungary

G. Bencze, S. Czellar, J. Karancsi [33], J. Molnar, Z. Szillasi

Institute of Physics, University of Debrecen, Debrecen, Hungary

P. Raics, B. Ujvari [36], G. Zilizi

Karoly Robert Campus, MATE Institute of Technology, Gyongyos, Hungary

T. Csorgo [32], F. Nemes [32], T. Novak

Panjab University, Chandigarh, India

J. Babbar, S. Bansal, S.B. Beri, V. Bhatnagar, G. Chaudhary, S. Chauhan, N. Dhingra [37],
R. Gupta, A. Kaur, A. Kaur, H. Kaur, M. Kaur, S. Kumar, P. Kumari, M. Meena,
K. Sandeep, T. Sheokand, J.B. Singh [38], A. Singla

University of Delhi, Delhi, India

A. Ahmed, A. Bhardwaj, A. Chhetri, B.C. Choudhary, A. Kumar, M. Naimuddin,
K. Ranjan, S. Saumya

Saha Institute of Nuclear Physics, HBNI, Kolkata, India

S. Baradia, S. Barman [39], S. Bhattacharya, D. Bhowmik, S. Dutta, S. Dutta, B. Gomber [40],
P. Palit, G. Saha, B. Sahu [40], S. Sarkar

Indian Institute of Technology Madras, Madras, India

P.K. Behera, S.C. Behera, S. Chatterjee, P. Jana, P. Kalbhor, J.R. Komaragiri [41], D. Kumar [41],
M. Mohammad Mobassir Ameen, L. Panwar [41], R. Pradhan, P.R. Pujahari, N.R. Saha,
A. Sharma, A.K. Sikdar, S. Verma

Tata Institute of Fundamental Research-A, Mumbai, India

T. Aziz, I. Das, S. Dugad, M. Kumar, G.B. Mohanty, P. Suryadevara

Tata Institute of Fundamental Research-B, Mumbai, India

A. Bala, S. Banerjee, R.M. Chatterjee [42], M. Guchait, S. Karmakar, S. Kumar, G. Majumder,
K. Mazumdar, S. Mukherjee, A. Thachayath

National Institute of Science Education and Research, An OCC of Homi Bhabha National Institute, Bhubaneswar, Odisha, India

S. Bahinipati [43], A.K. Das, C. Kar, D. Maity [44], P. Mal, T. Mishra,
V.K. Muraleedharan Nair Bindhu [44], K. Naskar [44], A. Nayak [44], P. Sadangi, P. Saha, S.K. Swain,
S. Varghese [44], D. Vats [44]

Indian Institute of Science Education and Research (IISER), Pune, India

A. Alpana, S. Dube, B. Kansal, A. Laha, S. Pandey, A. Rastogi, S. Sharma

Isfahan University of Technology, Isfahan, Iran

H. Bakhshiansohi [45,46], E. Khazaie [46], M. Zeinali [47]

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

S. Chenarani [48], S.M. Etesami, M. Khakzad, M. Mohammadi Najafabadi

University College Dublin, Dublin, Ireland

M. Felcini, M. Grunewald

INFN Sezione di Bari^a, Università di Bari^b, Politecnico di Bari^c, Bari, Italy

M. Abbrescia ^{a,b}, R. Aly ^{a,c,49}, A. Colaleo ^{a}, D. Creanza ^{a,c}, B. D' Anzi ^{a,b}, N. De Filippis ^{a,c},
M. De Palma ^{a,b}, G. De Robertis ^{a}, A. Di Florio ^{a,c}, W. Elmetenawee ^{a,b}, L. Fiore ^{a}, M. Franco^a,
G. Iaselli ^{a,c}, F. Licciulli^a, F. Loddo ^{a}, G. Maggi ^{a,c}, M. Maggi ^{a}, I. Margjeka ^{a,b},
V. Mastrapasqua ^{a,b}, S. My ^{a,b}, S. Nuzzo ^{a,b}, A. Pellecchia ^{a,b}, A. Pompili ^{a,b}, G. Pugliese ^{a,c},
R. Radogna ^{a}, G. Ramirez-Sanchez ^{a,c}, D. Ramos ^{a}, A. Ranieri ^{a}, L. Silvestris ^{a},
F.M. Simone ^{a,b}, Ü. Sözbilir ^{a}, A. Stamerra ^{a}, R. Venditti ^{a}, P. Verwilligen ^{a}, A. Zaza ^{a,b}

INFN Sezione di Bologna^a, Università di Bologna^b, Bologna, Italy

G. Abbiendi ^{a}, C. Baldanza^a, C. Battilana ^{a,b}, D. Bonacorsi ^{a,b}, L. Borgonovi ^{a}, V.D. Cafaro^a,
P. Capiluppi ^{a,b}, A. Castro ^{a,b}, F.R. Cavallo ^{a}, A. Crupano ^{a}, M. Cuffiani ^{a,b}, G.M. Dallavalle ^{a},
T. Diotalevi ^{a,b}, F. Fabbri ^{a}, A. Fanfani ^{a,b}, D. Fasanella ^{a,b}, P. Giacomelli ^{a}, L. Giommi ^{a,b},
V. Giordano^a, C. Grandi ^{a}, C. Guandalini^a, L. Guiducci ^{a,b}, S. Lo Meo ^{a,50}, L. Lunerti ^{a,b},
S. Marcellini ^{a}, G. Masetti ^{a}, F.L. Navarria ^{a,b}, A. Perrotta ^{a}, F. Primavera ^{a,b}, A.M. Rossi ^{a,b},
T. Rovelli ^{a,b}, G.P. Siroli ^{a,b}

INFN Sezione di Catania^a, Università di Catania^b, Catania, Italy

S. Costa ^{a,b,51}, A. Di Mattia ^{a}, R. Potenza^{a,b}, A. Tricomi ^{a,b,51}, C. Tuve ^{a,b}

INFN Sezione di Firenze^a, Università di Firenze^b, Firenze, Italy

G. Barbagli ^{a}, G. Bardelli ^{a,b}, B. Camaiani ^{a,b}, A. Cassese ^{a}, R. Ceccarelli ^{a}, V. Ciulli ^{a,b},
C. Civinini ^{a}, R. D'Alessandro ^{a,b}, E. Focardi ^{a,b}, G. Latino ^{a,b}, P. Lenzi ^{a,b}, M. Lizzo ^{a,b},
M. Meschini ^{a}, S. Paoletti ^{a}, A. Papanastassiou^{a,b}, G. Sguazzoni ^{a}, L. Viliani ^{a}

INFN Laboratori Nazionali di Frascati, Frascati, Italy

L. Benussi, S. Bianco, M. Caponero [52], S. Meola [53], D. Piccolo, G. Saviano[54]

INFN Sezione di Genova^a, Università di Genova^b, Genova, Italy

S. Cerchi^a, P. Chatagnon ^{a}, F. Ferro ^{a}, R. Puppo^a, E. Robutti ^{a}, C. Rossi ^{a}, S. Tosi ^{a,b}, A. Trovato^a

INFN Sezione di Milano-Bicocca^a, Università di Milano-Bicocca^b, Milano, Italy

A. Benaglia ^{a}, G. Boldrini ^{a}, F. Brivio ^{a,b}, F. Cetorelli ^{a,b}, F. De Guio ^{a,b}, M.E. Dinardo ^{a,b},
P. Dini ^{a}, S. Gennai ^{a}, A. Ghezzi ^{a,b}, P. Govoni ^{a,b}, L. Guzzi ^{a,b}, M.T. Lucchini ^{a,b},
M. Malberti ^{a}, S. Malvezzi ^{a}, A. Massironi ^{a}, D. Menasce ^{a}, L. Moroni ^{a}, M. Paganoni ^{a,b},
D. Pedrini ^{a}, B.S. Pinolini^a, S. Ragazzi ^{a,b}, N. Redaelli ^{a}, T. Tabarelli de Fatis ^{a,b}, D. Zuolo ^{a,b}

INFN Sezione di Napoli[a], Università di Napoli 'Federico II'[b], Napoli, Italy; Università della Basilicata[c], Potenza, Italy; Università G. Marconi[d], Roma, Italy

S. Buontempo [a], A. Cagnotta [a,b], F. Carnevali[a,b], N. Cavallo [a,c], A. De Iorio [a,b], F. Fabozzi [a,c], A.O.M. Iorio [a,b], L. Lista [a,b,55], P. Paolucci [a,30], B. Rossi [a], C. Sciacca [a,b]

INFN Sezione di Padova[a], Università di Padova[b], Padova, Italy; Università di Trento[c], Trento, Italy

R. Ardino [a], P. Azzi [a], N. Bacchetta [a,56], L. Barcellan[a], M. Bellato [a], M. Benettoni [a], A. Bergnoli [a], L. Berti[a], M. Biasotto [a,57], D. Bisello [a,b], P. Bortignon [a], A. Bragagnolo [a,b], R. Carlin [a,b], L. Castellani[a], P. Checchia [a], L. Ciano[a], A. Colombo[a], D. Corti[a], A. Crescente[a], T. Dorigo [a], S. Fantinel [a], F. Fanzago [a], F. Gasparini [a,b], U. Gasparini [a,b], F. Gonella [a], A. Gozzelino [a], A. Griggio[a], G. Grosso[a], M. Gulmini [a,57], R. Isocrate[a], L. Layer[a,58], E. Lusiani [a], M. Margoni [a,b], G. Maron [a,57], A.T. Meneguzzo [a,b], M. Michelotto [a], M. Migliorini [a,b], L. Modenese[a], F. Montecassiano [a], M. Negrello[a], M. Passaseo [a], J. Pazzini [a,b], L. Ramina[a], M. Rampazzo[a], M. Rebeschini[a], P. Ronchese [a,b], R. Rossin [a,b], M. Sgaravatto [a], F. Simonetto [a,b], G. Strong [a], R. Temporin[a], M. Tessaro[a], M. Toffano[a], N. Toniolo [a], M. Tosi [a,b], A. Triossi [a,b], S. Ventura [a], H. Yarar[a,b], M. Zanetti [a,b], P.G. Zatti[a], P. Zotto [a,b], A. Zucchetta [a,b], G. Zumerle [a,b]

INFN Sezione di Pavia[a], Università di Pavia[b], Pavia, Italy

S. Abu Zeid [a,59], C. Aimè [a,b], A. Braghieri [a], S. Calzaferri [a,b], D. Fiorina [a,b], P. Montagna [a,b], V. Re [a], C. Riccardi [a,b], P. Salvini [a], I. Vai [a,b], P. Vitulo [a,b]

INFN Sezione di Perugia[a], Università di Perugia[b], Perugia, Italy

S. Ajmal [a,b], P. Asenov [a,60], G. Baldinelli [a], F. Bianchi [a], G.M. Bilei [a], S. Bizzaglia[a], B. Checcucci [a], D. Ciangottini [a,b], L. Fanò [a,b], L. Farnesini[a], M. Ionica [a], M. Magherini [a,b], G. Mantovani[a,b], V. Mariani [a,b], M. Menichelli [a], A. Morozzi [a,b], F. Moscatelli [a,60], D. Passeri [a,b], A. Piccinelli [a,b], P. Placidi [a,b], M. Presilla [a,b], A. Rossi [a,b], A. Santocchia [a,b], D. Spiga [a], T. Tedeschi [a,b], C. Turrioni [a]

INFN Sezione di Pisa[a], Università di Pisa[b], Scuola Normale Superiore di Pisa[c], Pisa, Italy; Università di Siena[d], Siena, Italy

P. Azzurri [a], G. Bagliesi [a], A. Basti [a,b], R. Bhattacharya [a], L. Bianchini [a,b], M. Bitossi [a], T. Boccali [a], L. Borrello [a], F. Bosi[a], E. Bossini [a], D. Bruschini [a,c], R. Castaldi [a], M.A. Ciocci [a,b], M. Cipriani [a,b], V. D'Amante [a,d], R. Dell'Orso [a], S. Donato [a], F. Fiori [a], A. Giassi [a], F. Ligabue [a,c], G. Magazzu [a], M. Massa [a], D. Matos Figueiredo [a], A. Messineo [a,b], A. Moggi [a], M. Musich [a,b], F. Palla [a], F. Palmonari[a], S. Parolia [a], F. Raffaelli [a], A. Rizzi [a,b], G. Rolandi [a,c], S. Roy Chowdhury [a], T. Sarkar [a], A. Scribano [a], P. Spagnolo [a], R. Tenchini [a,b], G. Tonelli [a,b], N. Turini [a,d], A. Venturi [a], P.G. Verdini [a]

INFN Sezione di Roma[a], Sapienza Università di Roma[b], Roma, Italy

P. Barria [a], C. Basile [a,b], R. BIANCO[a], M. Campana [a,b], F. Cavallari [a], L. Cunqueiro Mendez [a,b], I. Dafinei [a], D. Del Re [a,b], E. Di Marco [a], M. Diemoz [a], F. Errico [a,b], F. Failla[a], E. Longo [a,b], P. Meridiani [a], J. Mijuskovic [a,b], C.A. Nicolau[a], M. Nuccetelli[a], G. Organtini [a,b], F. Pandolfi [a], R. Paramatti [a,b], V. Pettinacci [a], C. Quaranta [a,b], S. Rahatlou [a,b], C. Rovelli [a], F. Santanastasio [a,b], L. Soffi [a], R. Tramontano [a,b]

INFN Sezione di Torino^a, Università di Torino^b, Torino, Italy; Università del Piemonte Orientale^c, Novara, Italy

N. Amapane [ID]^{a,b}, R. Arcidiacono [ID]^{a,c}, S. Argiro [ID]^{a,b}, M. Arneodo [ID]^{a,c}, N. Bartosik [ID]^a, R. Bellan [ID]^{a,b}, A. Bellora [ID]^{a,b}, C. Biino [ID]^a, N. Cartiglia [ID]^a, S. Coli [ID]^a, M. Costa [ID]^{a,b}, R. Covarelli [ID]^{a,b}, D. Dattola^a, N. Demaria [ID]^a, L. Finco [ID]^a, M. Grippo [ID]^{a,b}, B. Kiani [ID]^{a,b}, F. Legger [ID]^a, F. Luongo [ID]^{a,b}, C. Mariotti [ID]^a, S. Maselli [ID]^a, A. Mecca [ID]^{a,b}, E. Migliore [ID]^{a,b}, M. Monteno [ID]^a, R. Mulargia [ID]^a, M.M. Obertino [ID]^{a,b}, G. Ortona [ID]^a, L. Pacher [ID]^{a,b}, N. Pastrone [ID]^a, M. Pelliccioni [ID]^a, M. Ruspa [ID]^{a,c}, F. Siviero [ID]^{a,b}, V. Sola [ID]^{a,b}, A. Solano [ID]^{a,b}, C. Tarricone [ID]^{a,b}, M. Tornago [ID]^{a,b}, D. Trocino [ID]^a, G. Umoret [ID]^{a,b}, A. Vagnerini [ID]^{a,b}, E. Vlasov [ID]^{a,b}

INFN Sezione di Trieste^a, Università di Trieste^b, Trieste, Italy

S. Belforte [ID]^a, V. Candelise [ID]^{a,b}, M. Casarsa [ID]^a, F. Cossutti [ID]^a, K. De Leo [ID]^{a,b}, G. Della Ricca [ID]^{a,b}

Kyungpook National University, Daegu, Korea

S. Dogra [ID], C. Huh [ID], B. Kim [ID], D.H. Kim [ID], J. Kim, J. Lee [ID], S.W. Lee [ID], C.S. Moon [ID], Y.D. Oh [ID], S.I. Pak [ID], M.S. Ryu [ID], S. Sekmen [ID], Y.C. Yang [ID]

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Korea

G. Bak [ID], P. Gwak [ID], H. Kim [ID], D.H. Moon [ID]

Hanyang University, Seoul, Korea

E. Asilar [ID], D. Kim [ID], T.J. Kim [ID], J.A. Merlin, J. Park [ID]

Korea University, Seoul, Korea

S. Choi [ID], S. Han, B. Hong [ID], K. Lee, K.S. Lee [ID], J. Park, S.K. Park, J. Yoo [ID]

Kyung Hee University, Department of Physics, Seoul, Korea

J. Goh [ID]

Sejong University, Seoul, Korea

H. S. Kim [ID], Y. Kim, S. Lee

Seoul National University, Seoul, Korea

J. Almond, J.H. Bhyun, J. Choi [ID], S. Jeon [ID], W. Jun [ID], J. Kim [ID], J.S. Kim, S. Ko [ID], H. Kwon [ID], H. Lee [ID], J. Lee [ID], S. Lee, B.H. Oh [ID], S.B. Oh [ID], H. Seo [ID], U.K. Yang, I. Yoon [ID]

University of Seoul, Seoul, Korea

W. Jang [ID], D.Y. Kang, Y. Kang [ID], S. Kim [ID], B. Ko, J.S.H. Lee [ID], Y. Lee [ID], I.C. Park [ID], Y. Roh, I.J. Watson [ID], S. Yang [ID]

Yonsei University, Department of Physics, Seoul, Korea

S. Ha [ID], H.D. Yoo [ID]

Sungkyunkwan University, Suwon, Korea

M. Choi [ID], M.R. Kim [ID], H. Lee, Y. Lee [ID], I. Yu [ID]

College of Engineering and Technology, American University of the Middle East (AUM), Dasman, Kuwait

T. Beyrouthy, Y. Maghrbi

Riga Technical University, Riga, Latvia

K. Dreimanis, A. Gaile, G. Pikurs, A. Potrebko, M. Seidel, V. Veckalns [61]

University of Latvia (LU), Riga, Latvia

N.R. Strautnieks

Vilnius University, Vilnius, Lithuania

M. Ambrozas, A. Juodagalvis, A. Rinkevicius, G. Tamulaitis

National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia

N. Bin Norjoharuddeen, I. Yusuff [62], Z. Zolkapli

Universidad de Sonora (UNISON), Hermosillo, Mexico

J.F. Benitez, A. Castaneda Hernandez, H.A. Encinas Acosta, L.G. Gallegos Maríñez, M. León Coello, J.A. Murillo Quijada, A. Sehrawat, L. Valencia Palomo

Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico

G. Ayala, H. Castilla-Valdez, E. De La Cruz-Burelo, I. Heredia-De La Cruz [63], R. Lopez-Fernandez, C.A. Mondragon Herrera, D.A. Perez Navarro, A. Sánchez Hernández

Universidad Iberoamericana, Mexico City, Mexico

C. Oropeza Barrera, M. Ramírez García

Benemerita Universidad Autonoma de Puebla, Puebla, Mexico

I. Bautista, I. Pedraza, H.A. Salazar Ibarguen, C. Uribe Estrada

University of Montenegro, Podgorica, Montenegro

I. Bubanja, N. Raicevic

University of Canterbury, Christchurch, New Zealand

P.H. Butler

National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan

A. Ahmad, M.I. Asghar, A. Awais, M.I.M. Awan, H.R. Hoorani, W.A. Khan

AGH University of Krakow, Faculty of Computer Science, Electronics and Telecommunications, Krakow, Poland

V. Avati, L. Grzanka, M. Malawski

National Centre for Nuclear Research, Swierk, Poland

H. Bialkowska, M. Bluj, B. Boimska, M. Górski, M. Kazana, M. Szleper, P. Zalewski

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland

K. Bunkowski ©, K. Doroba ©, A. Kalinowski ©, K. Kierzkowski ©, M. Konecki ©, J. Krolikowski ©, W. Oklinski

Warsaw University of Technology, Warsaw, Poland

K. Pozniak ©, W. Zabolotny ©

Laboratório de Instrumentação e Física Experimental de Partículas, Lisboa, Portugal

M. Araujo ©, D. Bastos ©, C. Beirão Da Cruz E Silva ©, A. Boletti ©, M. Bozzo ©, G. Da Molin ©, P. Faccioli ©, M. Gallinaro ©, J. Hollar ©, N. Leonardo ©, T. Niknejad ©, M. Pisano ©, J. Rasteiro Da Silva, J. Seixas ©, J. Varela ©, J.W. Wulff

Faculty of Physics, University of Belgrade, Belgrade, Serbia

P. Adzic ©, P. Milenovic ©

VINCA Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia

M. Dordevic ©, J. Milosevic ©, V. Rekovic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

M. Aguilar-Benitez, J. Alcaraz Maestre ©, M. Barrio Luna, Cristina F. Bedoya ©, M. Cepeda ©, M. Cerrada ©, N. Colino ©, J. Cuchillo Ortega, B. De La Cruz ©, C.I. De Lara Rodríguez, A. Delgado Peris ©, D. Fernández Del Val ©, J.P. Fernández Ramos ©, J. Flix ©, M.C. Fouz ©, O. Gonzalez Lopez ©, S. Goy Lopez ©, J.M. Hernandez ©, M.I. Josa ©, J. León Holgado ©, O. Manzanilla ©, D. Moran ©, C. M. Morcillo Perez ©, Á. Navarro Tobar ©, R. Paz Herrera, C. Perez Dengra ©, A. Pérez-Calero Yzquierdo ©, J. Puerta Pelayo ©, I. Redondo ©, D.D. Redondo Ferrero ©, L. Romero, S. Sánchez Navas ©, J. Sastre ©, L. Urda Gómez ©, J. Vazquez Escobar ©, C. Willmott

Universidad Autónoma de Madrid, Madrid, Spain

J.F. de Trocóniz ©

Universidad de Oviedo, Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias (ICTEA), Oviedo, Spain

B. Alvarez Gonzalez ©, J. Cuevas ©, J. Fernandez Menendez ©, S. Folgueras ©, I. Gonzalez Caballero ©, J.R. González Fernández ©, E. Palencia Cortezon ©, C. Ramón Álvarez ©, V. Rodríguez Bouza ©, A. Soto Rodríguez ©, A. Trapote ©, C. Vico Villalba ©, P. Vischia ©

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain

S. Bhowmik ©, S. Blanco Fernández ©, J.A. Brochero Cifuentes ©, I.J. Cabrillo ©, A. Calderon ©, J. Duarte Campderros ©, M. Fernandez ©, C. Fernandez Madrazo ©, G. Gomez ©, C. Lasaosa García ©, C. Martinez Rivero ©, P. Martinez Ruiz del Arbol ©, F. Matorras ©, P. Matorras Cuevas ©, E. Navarrete Ramos ©, J. Piedra Gomez ©, C. Prieels, L. Scodellaro ©, I. Vila ©, J.M. Vizan Garcia ©

University of Colombo, Colombo, Sri Lanka

M.K. Jayananda ©, B. Kailasapathy ©[64], D.U.J. Sonnadara ©, D.D.C. Wickramarathna ©

University of Ruhuna, Department of Physics, Matara, Sri Lanka

W.G.D. Dharmaratna [65], K. Liyanage, N. Perera, N. Wickramage

CERN, European Organization for Nuclear Research, Geneva, Switzerland

D. Abbaneo, C. Amendola, V. Amoiridis, E. Auffray, G. Auzinger, J. Baechler, D. Barney,
A. Bermúdez Martínez, M. Bianco, B. Bilin, A.A. Bin Anuar, A. Bocci, E. Brondolin,
P. Brummer, C. Caillol, T. Camporesi, G. Cerminara, N. Chernyavskaya, B. Curé,
D. d'Enterria, A. Dabrowski, A. David, A. De Roeck, M.M. Defranchis, M. Deile,
C. Deldicque, M. Dobson, A. Dvorak, F. Fallavollita[66], L. Forthomme, G. Franzoni, W. Funk,
S. Giani, D. Gigi, K. Gill, F. Glege, L. Gouskos, N. Gutic, M. Haranko, J. Hegeman,
V. Innocente, G. Izquierdo Moreno, T. James, P. Janot, J. Kieseler, C. Kishimoto Bisbe,
N. Kratochwil, S. Laurila, P. Lecoq, E. Leutgeb, C. Lourenço, B. Maier, L. Malgeri,
M. Mannelli, A.C. Marini, F. Meijers, S. Mersi, E. Meschi, V. Milosevic, R.K. Mommsen,
F. Moortgat, M. Mulders, S. Orfanelli, L. Orsini, F. Pantaleo, K. Peron, M. Peruzzi, A. Petrilli,
G. Petrucciani, A. Pfeiffer, M. Pierini, D. Piparo, A. Poupakis, H. Qu, D. Rabady, A. Racz,
G. Reales Gutiérrez, M. Rovere, H. Sakulin, S. Scarfi, C. Schwick, M. Selvaggi, A. Sharma,
K. Shchelina, P. Silva, P. Sphicas[67], A.G. Stahl Leiton, A. Steen, S. Summers, D. Treille,
P. Tropea, A. Tsirou, C. Vázquez, D. Walter, J. Wanczyk[68], K.A. Wozniak, P. Zehetner,
P. Zejdl, W.D. Zeuner

Paul Scherrer Institut, Villigen, Switzerland

T. Bevilacqua[69], L. Caminada[69], A. Ebrahimi, W. Erdmann, R. Horisberger, Q. Ingram,
H.C. Kaestli, D. Kotlinski, C. Lange, B. Meier, M. Missiroli[69], L. Noehte[69], T. Rohe, S. Streuli

ETH Zurich - Institute for Particle Physics and Astrophysics (IPA), Zurich, Switzerland

T.K. Aarrestad, K. Androsov[68], M. Backhaus, G. Bonomelli, S. Burkhalter[70], A. Calandri,
C. Cazzaniga, K. Datta, A. De Cosa, G. Dissertori, M. Dittmar, M. Donegà, F. Eble, T. Gadek,
M. Galli, K. Gedia, F. Glessgen, C. Grab, N. Härringer, T.G. Harte, D. Hits, W. Lustermann,
A.-M. Lyon, R.A. Manzoni, M. Marchegiani, L. Marchese, C. Martin Perez, A. Mascellani[68],
F. Nessi-Tedaldi, F. Pauss, V. Perovic, S. Pigazzini, M.G. Ratti, C. Reissel, T. Reitenspiess,
B. Ristic, F. Riti, D. Ruini, R. Seidita, K. Stachon, J. Steggemann[68], D. Valsecchi, R. Wallny

Universität Zürich, Zurich, Switzerland

C. Amsler[71], P. Bärtschi, K. Boesiger, C. Botta, D. Brzhechko, M.F. Canelli, K. Cormier,
A. De Wit, R. Del Burgo, M. Gienal, J.K. Heikkilä, D. Hernandez Garland, M. Huwiler, W. Jin,
A. Jofrehei, B. Kilminster, S. Leontsinis, S.P. Liechti, A. Macchiolo, R. Maier, P. Meiring,
V.M. Mikuni, U. Molinatti, B. Neuenschwander, I. Neutelings, G. Rauco, A. Reimers, P. Robmann,
D. Salerno, S. Sanchez Cruz, K. Schweiger, M. Senger, Y. Takahashi, S.A. Wiederkehr, D. Wolf

National Central University, Chung-Li, Taiwan

C. Adloff[72], C.M. Kuo, W. Lin, P.K. Rout, P.C. Tiwari[41], S.S. Yu

National Taiwan University (NTU), Taipei, Taiwan

L. Ceard, Y. Chao [iD], K.F. Chen [iD], P.s. Chen, Z.g. Chen, W.-S. Hou [iD], T.h. Hsu, Y.w. Kao, R. Khurana, G. Kole [iD], Y.y. Li [iD], R.-S. Lu [iD], E. Paganis [iD], A. Psallidas, X.f. Su, J. Thomas-Wilsker [iD], H.y. Wu, E. Yazgan [iD]

Chulalongkorn University, Faculty of Science, Department of Physics, Bangkok, Thailand

C. Asawatangtrakuldee [iD], N. Srimanobhas [iD], V. Wachirapusitanand [iD]

Çukurova University, Physics Department, Science and Art Faculty, Adana, Turkey

D. Agyel [iD], F. Boran [iD], Z.S. Demiroglu [iD], F. Dolek [iD], I. Dumanoglu [iD][73], E. Eskut [iD], Y. Guler [iD][74], E. Gurpinar Guler [iD][74], C. Isik [iD], O. Kara, A. Kayis Topaksu [iD], U. Kiminsu [iD], G. Onengut [iD], K. Ozdemir [iD][75], A. Polatoz [iD], B. Tali [iD][76], U.G. Tok [iD], S. Turkcapar [iD], E. Uslan [iD], I.S. Zorbakir [iD]

Middle East Technical University, Physics Department, Ankara, Turkey

K. Ocalan [iD][77], M. Yalvac [iD][78]

Bogazici University, Istanbul, Turkey

B. Akgun [iD], I.O. Atakisi [iD], E. Gülmez [iD], M. Kaya [iD][79], O. Kaya [iD][80], S. Tekten [iD][81], E.A. Yetkin [iD][82]

Istanbul Technical University, Istanbul, Turkey

A. Cakir [iD], K. Cankocak [iD][73], Y. Komurcu [iD], S. Sen [iD][83]

Istanbul University, Istanbul, Turkey

O. Aydilek [iD], S. Cerci [iD][76], V. Epshteyn [iD], B. Hacisahinoglu [iD], I. Hos [iD][84], B. Isildak [iD][85], B. Kaynak [iD], S. Ozkorucuklu [iD], H. Sert [iD], C. Simsek [iD], D. Sunar Cerci [iD][76], C. Zorbilmez [iD]

Yildiz Technical University, Istanbul, Turkey

T. Yetkin [iD]

Institute for Scintillation Materials of National Academy of Science of Ukraine, Kharkiv, Ukraine

A. Boyaryntsev [iD], B. Grynyov [iD]

National Science Centre, Kharkiv Institute of Physics and Technology, Kharkiv, Ukraine

L. Levchuk [iD]

University of Bristol, Bristol, United Kingdom

D. Anthony [iD], J.J. Brooke [iD], A. Bundock [iD], F. Bury [iD], E. Clement [iD], D. Cussans [iD], H. Flacher [iD], M. Glowacki, J. Goldstein [iD], H.F. Heath [iD], L. Kreczko [iD], B. Krikler, S. Paramesvaran [iD], S. Seif El Nasr-Storey, V.J. Smith [iD], N. Stylianou [iD][86], K. Walkingshaw Pass, R. White [iD]

Rutherford Appleton Laboratory, Didcot, United Kingdom

A.H. Ball, K.W. Bell [iD], A. Belyaev [iD][87], C. Brew [iD], R.M. Brown [iD], D.J.A. Cockerill [iD], C. Cooke [iD], K.V. Ellis, K. Harder [iD], S. Harper [iD], M.-L. Holmberg [iD][88], Sh. Jain [iD][89], J. Linacre [iD], K. Manolopoulos, D.M. Newbold [iD], E. Olaiya, D. Petyt [iD], T. Reis [iD], G. Salvi [iD], T. Schuh, C.H. Shepherd-Themistocleous [iD], I.R. Tomalin [iD], T. Williams [iD]

Imperial College, London, United Kingdom

R. Bainbridge ⓘ, P. Bloch ⓘ, C.E. Brown ⓘ, O. Buchmuller, V. Cacchio, C.A. Carrillo Montoya ⓘ,
G.S. Chahal ⓘ[90], D. Colling ⓘ, J.S. Dancu, P. Dauncey ⓘ, G. Davies ⓘ, J. Davies, M. Della Negra ⓘ,
S. Fayer, G. Fedi ⓘ, G. Hall ⓘ, M.H. Hassanshahi ⓘ, A. Howard, G. Iles ⓘ, M. Knight ⓘ, J. Langford ⓘ,
L. Lyons ⓘ, A.-M. Magnan ⓘ, S. Malik, A. Martelli ⓘ, M. Mieskolainen ⓘ, J. Nash ⓘ[91], M. Pesaresi,
B.C. Radburn-Smith ⓘ, A. Richards, A. Rose ⓘ, C. Seez ⓘ, R. Shukla ⓘ, A. Tapper ⓘ, K. Uchida ⓘ,
G.P. Uttley ⓘ, L.H. Vage, T. Virdee ⓘ[30], M. Vojinovic ⓘ, N. Wardle ⓘ, D. Winterbottom ⓘ

Brunel University, Uxbridge, United Kingdom

K. Coldham, J.E. Cole ⓘ, A. Khan, P. Kyberd ⓘ, I.D. Reid ⓘ

Baylor University, Waco, Texas, U.S.A.

S. Abdullin ⓘ, A. Brinkerhoff ⓘ, B. Caraway ⓘ, J. Dittmann ⓘ, K. Hatakeyama ⓘ, J. Hiltbrand ⓘ,
A.R. Kanuganti ⓘ, B. McMaster ⓘ, M. Saunders ⓘ, S. Sawant ⓘ, C. Sutantawibul ⓘ, M. Toms ⓘ,
J. Wilson ⓘ

Catholic University of America, Washington, DC, U.S.A.

R. Bartek ⓘ, A. Dominguez ⓘ, C. Huerta Escamilla, A.E. Simsek ⓘ, R. Uniyal ⓘ, A.M. Vargas Hernandez ⓘ

The University of Alabama, Tuscaloosa, Alabama, U.S.A.

R. Chudasama ⓘ, S.I. Cooper ⓘ, S.V. Gleyzer ⓘ, C.U. Perez ⓘ, P. Rumerio ⓘ[92], E. Usai ⓘ, C. West ⓘ

Boston University, Boston, Massachusetts, U.S.A.

A. Akpinar ⓘ, A. Albert ⓘ, D. Arcaro ⓘ, C. Cosby ⓘ, Z. Demiragli ⓘ, C. Erice ⓘ, E. Fontanesi ⓘ,
D. Gastler ⓘ, J. Rohlf ⓘ, K. Salyer ⓘ, D. Sperka ⓘ, D. Spitzbart ⓘ, I. Suarez ⓘ, A. Tsatsos, S. Yuan ⓘ

Brown University, Providence, Rhode Island, U.S.A.

G. Benelli ⓘ, X. Coubez[25], D. Cutts ⓘ, M. Hadley ⓘ, U. Heintz ⓘ, J.M. Hogan ⓘ[93], T. Kwon ⓘ,
G. Landsberg ⓘ, K.T. Lau ⓘ, D. Li ⓘ, J. Luo ⓘ, S. Mondal ⓘ, M. Narain ⓘ[†], N. Pervan ⓘ, S. Sagir ⓘ[94],
F. Simpson ⓘ, W.Y. Wong, X. Yan ⓘ, D. Yu ⓘ, W. Zhang

University of California, Davis, Davis, California, U.S.A.

S. Abbott ⓘ, J. Bonilla ⓘ, C. Brainerd ⓘ, R. Breedon ⓘ, M. Calderon De La Barca Sanchez ⓘ,
M. Chertok ⓘ, M. Citron ⓘ, J. Conway ⓘ, P.T. Cox ⓘ, R. Erbacher ⓘ, G. Haza ⓘ, F. Jensen ⓘ, O. Kukral ⓘ,
G. Mocellin ⓘ, M. Mulhearn ⓘ, D. Pellett ⓘ, B. Regnery ⓘ, W. Wei ⓘ, Y. Yao ⓘ, F. Zhang ⓘ

University of California, Los Angeles, California, U.S.A.

M. Bachtis ⓘ, R. Cousins ⓘ, A. Datta ⓘ, J. Hauser ⓘ, M. Ignatenko ⓘ, M.A. Iqbal, T. Lam ⓘ,
E. Manca ⓘ, W.A. Nash ⓘ, D. Saltzberg ⓘ, B. Stone ⓘ, V. Valuev ⓘ

University of California, Riverside, Riverside, California, U.S.A.

R. Clare ⓘ, M. Gordon, G. Hanson ⓘ, W. Si ⓘ, S. Wimpenny ⓘ[†]

University of California, San Diego, La Jolla, California, U.S.A.

J.G. Branson ⓘ, S. Cittolin ⓘ, S. Cooperstein ⓘ, D. Diaz ⓘ, J. Duarte ⓘ, R. Gerosa ⓘ, L. Giannini ⓘ,
J. Guiang ⓘ, R. Kansal ⓘ, V. Krutelyov ⓘ, R. Lee ⓘ, J. Letts ⓘ, M. Masciovecchio ⓘ, F. Mokhtar ⓘ,

S. Morovic , A. Petrucci , M. Pieri , M. Quinnan , B.V. Sathia Narayanan , V. Sharma ,
M. Tadel , E. Vourliotis , F. Würthwein , Y. Xiang , A. Yagil

University of California, Santa Barbara - Department of Physics, Santa Barbara, California, U.S.A.

L. Brennan , C. Campagnari , G. Collura , A. Dorsett , J. Incandela , M. Kilpatrick , J. Kim ,
A.J. Li , P. Masterson , H. Mei , M. Oshiro , J. Richman , U. Sarica , R. Schmitz , F. Setti ,
J. Sheplock , D. Stuart , S. Wang

California Institute of Technology, Pasadena, California, U.S.A.

J. Balcas, A. Bornheim , O. Cerri, A. Latorre, J.M. Lawhorn , J. Mao , H.B. Newman ,
T. Q. Nguyen , M. Spiropulu , J.R. Vlimant , C. Wang , S. Xie , R.Y. Zhu

Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.

J. Alison , S. An , M.B. Andrews , P. Bryant , V. Dutta , T. Ferguson , A. Harilal , C. Liu ,
T. Mudholkar , S. Murthy , M. Paulini , A. Roberts , A. Sanchez , W. Terrill

University of Colorado Boulder, Boulder, Colorado, U.S.A.

J.P. Cumalat , W.T. Ford , A. Hassani , G. Karathanasis , E. MacDonald, N. Manganelli ,
F. Marini , A. Perloff , C. Savard , N. Schonbeck , K. Stenson , K.A. Ulmer , S.R. Wagner ,
N. Zipper

Cornell University, Ithaca, New York, U.S.A.

J. Alexander , S. Bright-Thonney , X. Chen , D.J. Cranshaw , J. Fan , X. Fan , D. Gadkari ,
S. Hogan , J. Monroy , J.R. Patterson , J. Reichert , M. Reid , A. Ryd , J. Thom , P. Wittich ,
R. Zou

Fermi National Accelerator Laboratory, Batavia, Illinois, U.S.A.

J. Adelman-McCarthy , M. Albrow , M. Alyari , O. Amram , G. Apollinari , A. Apresyan ,
L.A.T. Bauerdick , D. Berry , J. Berryhill , P.C. Bhat , K. Burkett , J.N. Butler , A. Canepa ,
G.B. Cerati , H.W.K. Cheung , F. Chlebana , G. Cummings , W. Dagenhart , J. Dickinson ,
I. Dutta , V.D. Elvira , Y. Feng , J. Freeman , A. Gandrakota , P. Gartung , Z. Gecse , L. Gray ,
D. Green, S. Grünendahl , D. Guerrero , O. Gutsche , R.M. Harris , R. Heller , T.C. Herwig ,
J. Hirschauer , L. Horyn , D. Hufnagel , B. Jayatilaka , S. Jindariani , M. Johnson , C.D. Jones,
U. Joshi , T. Klijnsma , B. Klima , M.J. Kortelainen , K.H.M. Kwok , S. Lammel , D. Lincoln ,
R. Lipton , T. Liu , C. Madrid , K. Maeshima , C. Mantilla , D. Mason , P. McBride ,
P. Merkel , S. Mrenna , S. Nahn , J. Ngadiuba , D. Noonan , V. Papadimitriou , N. Pastika ,
K. Pedro , C. Pena [95], F. Ravera , A. Reinsvold Hall [96], L. Ristori , E. Sexton-Kennedy ,
N. Smith , A. Soha , L. Spiegel , S. Stoynev , L. Taylor , S. Tkaczyk , N.V. Tran , L. Uplegger ,
E.W. Vaandering , I. Zoi

University of Florida, Gainesville, Florida, U.S.A.

C. Aruta , P. Avery , D. Bourilkov , L. Cadamuro , P. Chang , V. Cherepanov , R.D. Field,
E. Koenig , M. Kolosova , J. Konigsberg , A. Korytov , K.H. Lo, K. Matchev , N. Menendez ,

G. Mitselmakher ©, A. Muthirakalayil Madhu ©, N. Rawal ©, D. Rosenzweig ©, S. Rosenzweig ©, K. Shi ©, J. Wang ©

Florida State University, Tallahassee, Florida, U.S.A.

T. Adams ©, A. Al Kadhim ©, A. Askew ©, N. Bower ©, R. Habibullah ©, V. Hagopian ©, R. Hashmi ©, R.S. Kim ©, S. Kim ©, T. Kolberg ©, G. Martinez, H. Prosper ©, P.R. Prova, O. Viazlo ©, M. Wulansatiti ©, R. Yohay ©, J. Zhang

Florida Institute of Technology, Melbourne, Florida, U.S.A.

B. Alsufyani, M.M. Baarmand ©, S. Butalla ©, T. Elkafrawy ©[59], M. Hohlmann ©, R. Kumar Verma ©, M. Rahmani, F. Yumiceva ©

University of Illinois Chicago, Chicago, U.S.A., Chicago, U.S.A.

M.R. Adams ©, C. Bennett, R. Cavanaugh ©, S. Dittmer ©, R. Escobar Franco ©, A. Evdokimov ©, O. Evdokimov ©, C.E. Gerber ©, D.J. Hofman ©, J.h. Lee ©, D. S. Lemos ©, A.H. Merrit ©, C. Mills ©, S. Nanda ©, G. Oh ©, B. Ozek ©, D. Pilipovic ©, T. Roy ©, S. Rudrabhatla ©, M.B. Tonjes ©, N. Varelas ©, X. Wang ©, Z. Ye ©, J. Yoo ©

The University of Iowa, Iowa City, Iowa, U.S.A.

M. Alhusseini ©, B. Bilki ©[97], D. Blend, P. Debbins ©, K. Dilsiz ©[98], L. Emediato ©, G. Karaman ©, O.K. Köseyan ©, J.-P. Merlo, A. Mestvirishvili ©[99], M.J. Miller, J. Nachtman ©, O. Neogi, H. Ogul ©[100], Y. Onel ©, A. Penzo ©, I. Schmidt, C. Snyder, D. Southwick ©, E. Tiras ©[101], J. Wetzel ©

Johns Hopkins University, Baltimore, Maryland, U.S.A.

B. Blumenfeld ©, L. Corcodilos ©, J. Davis ©, A.V. Gritsan ©, L. Kang ©, S. Kyriacou ©, P. Maksimovic ©, M. Roguljic ©, J. Roskes ©, S. Sekhar ©, M. Swartz ©, T.Á. Vámi ©

The University of Kansas, Lawrence, Kansas, U.S.A.

A. Abreu ©, L.F. Alcerro Alcerro ©, J. Anguiano ©, P. Baringer ©, A. Bean ©, Z. Flowers ©, D. Grove, J. King ©, G. Krintiras ©, M. Lazarovits ©, C. Le Mahieu ©, C. Lindsey, J. Marquez ©, N. Minafra ©, M. Murray ©, M. Nickel ©, M. Pitt ©, S. Popescu ©[102], C. Rogan ©, C. Royon ©, R. Salvatico ©, S. Sanders ©, C. Smith ©, Q. Wang ©, G. Wilson ©

Kansas State University, Manhattan, Kansas, U.S.A.

B. Allmond ©, A. Ivanov ©, K. Kaadze ©, A. Kalogeropoulos ©, D. Kim, Y. Maravin ©, K. Nam, J. Natoli ©, D. Roy ©, G. Sorrentino ©

Lawrence Livermore National Laboratory, Livermore, California, U.S.A.

F. Rebassoo ©, D. Wright ©

University of Maryland, College Park, Maryland, U.S.A.

E. Adams ©, A. Baden ©, O. Baron, A. Belloni ©, A. Bethani ©, Y.m. Chen ©, S.C. Eno ©, N.J. Hadley ©, S. Jabeen ©, R.G. Kellogg ©, T. Koeth ©, Y. Lai ©, S. Lascio ©, A.C. Mignerey ©, S. Nabili ©, C. Palmer ©, C. Papageorgakis ©, M.M. Paranjpe, L. Wang ©, K. Wong ©

Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

J. Bendavid ⓘ, W. Busza ⓘ, I.A. Cali ⓘ, Y. Chen ⓘ, M. D'Alfonso ⓘ, G.L. Darlea ⓘ, J. Eysermans ⓘ,
C. Freer ⓘ, G. Gomez-Ceballos ⓘ, M. Goncharov, P. Harris, D. Hoang, D. Kovalskyi ⓘ, J. Krupa ⓘ,
L. Lavezzo ⓘ, Y.-J. Lee ⓘ, K. Long ⓘ, C. Mironov ⓘ, C. Paus ⓘ, D. Rankin ⓘ, C. Roland ⓘ, G. Roland ⓘ,
S. Rothman ⓘ, Z. Shi ⓘ, G.S.F. Stephans ⓘ, J. Wang, Z. Wang ⓘ, B. Wyslouch ⓘ, T. J. Yang ⓘ

University of Minnesota, Minneapolis, Minnesota, U.S.A.

B. Crossman ⓘ, E. Frahm, B.M. Joshi ⓘ, C. Kapsiak ⓘ, M. Krohn ⓘ, D. Mahon ⓘ, J. Mans ⓘ,
M. Revering ⓘ, R. Rusack ⓘ, R. Saradhy ⓘ, N. Schroeder ⓘ, N. Strobbe ⓘ, M.A. Wadud ⓘ

University of Mississippi, Oxford, Mississippi, U.S.A.

L.M. Cremaldi ⓘ

University of Nebraska-Lincoln, Lincoln, Nebraska, U.S.A.

K. Bloom ⓘ, M. Bryson, D.R. Claes ⓘ, C. Fangmeier ⓘ, F. Golf ⓘ, C. Joo ⓘ, I. Kravchenko ⓘ, I. Reed ⓘ,
J.E. Siado ⓘ, G.R. Snow[†], W. Tabb ⓘ, A. Wightman ⓘ, F. Yan ⓘ, A.G. Zecchinelli ⓘ

State University of New York at Buffalo, Buffalo, New York, U.S.A.

G. Agarwal ⓘ, H. Bandyopadhyay ⓘ, L. Hay ⓘ, I. Iashvili ⓘ, A. Kharchilava ⓘ, C. McLean ⓘ, M. Morris ⓘ,
D. Nguyen ⓘ, J. Pekkanen ⓘ, S. Rappoccio ⓘ, H. Rejeb Sfar, A. Williams ⓘ

Northeastern University, Boston, Massachusetts, U.S.A.

G. Alverson ⓘ, E. Barberis ⓘ, J. Dervan, Y. Haddad ⓘ, Y. Han ⓘ, A. Krishna ⓘ, J. Li ⓘ, M. Lu ⓘ,
G. Madigan ⓘ, B. Marzocchi ⓘ, D.M. Morse ⓘ, V. Nguyen ⓘ, T. Orimoto ⓘ, A. Parker ⓘ, L. Skinnari ⓘ,
A. Tishelman-Charny ⓘ, B. Wang ⓘ, D. Wood ⓘ

Northwestern University, Evanston, Illinois, U.S.A.

S. Bhattacharya ⓘ, J. Bueghly, Z. Chen ⓘ, K.A. Hahn ⓘ, Y. Liu ⓘ, Y. Miao ⓘ, D.G. Monk ⓘ,
M.H. Schmitt ⓘ, A. Taliercio ⓘ, M. Velasco

University of Notre Dame, Notre Dame, Indiana, U.S.A.

R. Band ⓘ, R. Bucci, S. Castells ⓘ, M. Cremonesi, A. Das ⓘ, R. Goldouzian ⓘ, M. Hildreth ⓘ, K.W. Ho ⓘ,
K. Hurtado Anampa ⓘ, C. Jessop ⓘ, K. Lannon ⓘ, J. Lawrence ⓘ, N. Loukas ⓘ, L. Lutton, J. Mariano,
N. Marinelli, I. Mcalister, T. McCauley ⓘ, C. Mcgrady ⓘ, K. Mohrman ⓘ, C. Moore ⓘ, Y. Musienko[13],
H. Nelson ⓘ, M. Osherson ⓘ, R. Ruchti ⓘ, A. Townsend, M. Wayne ⓘ, H. Yockey, M. Zarucki ⓘ,
L. Zygala ⓘ

The Ohio State University, Columbus, Ohio, U.S.A.

A. Basnet ⓘ, B. Bylsma, M. Carrigan ⓘ, L.S. Durkin ⓘ, C. Hill ⓘ, M. Joyce ⓘ, A. Lesauvage ⓘ,
M. Nunez Ornelas ⓘ, K. Wei, B.L. Winer ⓘ, B. R. Yates ⓘ

Princeton University, Princeton, New Jersey, U.S.A.

F.M. Addesa ⓘ, H. Bouchamaoui ⓘ, P. Das ⓘ, G. Dezoort ⓘ, P. Elmer ⓘ, A. Frankenthal ⓘ, B. Greenberg ⓘ,
N. Haubrich ⓘ, S. Higginbotham ⓘ, G. Kopp ⓘ, S. Kwan ⓘ, D. Lange ⓘ, A. Loeliger, D. Marlow ⓘ,
I. Ojalvo ⓘ, J. Olsen ⓘ, D. Stickland ⓘ, C. Tully ⓘ

University of Puerto Rico, Mayaguez, Puerto Rico, U.S.A.

S. Malik

Purdue University, West Lafayette, Indiana, U.S.A.

A.S. Bakshi, V.E. Barnes, S. Chandra, R. Chawla, S. Das, A. Gu, L. Gutay, M. Jones, A.W. Jung, D. Kondratyev, A.M. Koshy, M. Liu, G. Negro, N. Neumeister, G. Paspalaki, S. Piperov, A. Purohit, J.F. Schulte, M. Stojanovic[17], J. Thieman, A. K. Virdi, F. Wang, A. Wildridge, W. Xie

Purdue University Northwest, Hammond, Indiana, U.S.A.

J. Dolen, N. Parashar, A. Pathak

Rice University, Houston, Texas, U.S.A.

D. Acosta, K. Banicz, A. Baty, U. Behrens, T. Carnahan, S. Dildick, K.M. Ecklund, P.J. Fernández Manteca, S. Freed, P. Gardner, F.J.M. Geurts, P. Kelling, R. Krawczyk, A. Kumar, W. Li, J.H. Liu, M. Matveev, O. Miguel Colin, T. Nussbaum, B.P. Padley, R. Redjimi, J. Rotter, E. Yigitbasi, Y. Zhang

University of Rochester, Rochester, New York, U.S.A.

A. Bodek, P. de Barbaro, R. Demina, J.L. Dulemba, C. Fallon, A. Garcia-Bellido, O. Hindrichs, A. Khukhunaishvili, P. Parygin, E. Popova, R. Taus, G.P. Van Onsem

The Rockefeller University, New York, New York, U.S.A.

K. Goulianos

Rutgers, The State University of New Jersey, Piscataway, New Jersey, U.S.A.

B. Chiarito, J.P. Chou, Y. Gershtein, E. Halkiadakis, A. Hart, M. Heindl, D. Jaroslawski, O. Karacheban[28], I. Laflotte, A. Lath, R. Montalvo, K. Nash, H. Routray, S. Salur, S. Schnetzer, S. Somalwar, R. Stone, S.A. Thayil, S. Thomas, J. Vora, H. Wang

University of Tennessee, Knoxville, Tennessee, U.S.A.

H. Acharya, D. Ally, A.G. Delannoy, S. Fiorendi, T. Holmes, N. Karunarathna, L. Lee, E. Nibigira, S. Spanier

Texas A&M University, College Station, Texas, U.S.A.

D. Aebi, M. Ahmad, T. Akhter, O. Bouhali[103], M. Dalchenko, R. Eusebi, J. Gilmore, T. Huang, T. Kamon[104], H. Kim, S. Luo, S. Malhotra, R. Mueller, D. Overton, D. Rathjens, A. Safonov

Texas Tech University, Lubbock, Texas, U.S.A.

N. Akchurin, J. Damgov, V. Hegde, A. Hussain, Y. Kazhykarim, K. Lamichhane, S.W. Lee, A. Mankel, T. Mengke, S. Muthumuni, T. Peltola, I. Volobouev, A. Whitbeck

Vanderbilt University, Nashville, Tennessee, U.S.A.

E. Appelt, S. Greene, A. Gurrola, W. Johns, R. Kunnawalkam Elayavalli, A. Melo, F. Romeo, P. Sheldon, S. Tuo, J. Velkovska, J. Viinikainen

University of Virginia, Charlottesville, Virginia, U.S.A.

B. Cardwell [ID], B. Cox [ID], S. Goadhouse [ID], J. Hakala [ID], R. Hirosky [ID], A. Ledovskoy [ID], A. Li [ID], C. Neu [ID], C.E. Perez Lara [ID]

Wayne State University, Detroit, Michigan, U.S.A.

P.E. Karchin [ID]

University of Wisconsin - Madison, Madison, Wisconsin, U.S.A.

A. Aravind, S. Banerjee [ID], K. Black [ID], T. Bose [ID], S. Dasu [ID], I. De Bruyn [ID], P. Everaerts [ID], C. Galloni, H. He [ID], M. Herndon [ID], A. Herve [ID], C.K. Koraka [ID], A. Lanaro, R. Loveless [ID], J. Madhusudanan Sreekala [ID], A. Mallampalli [ID], A. Mohammadi [ID], S. Mondal, G. Parida [ID], D. Pinna, A. Savin, V. Shang [ID], V. Sharma [ID], W.H. Smith [ID], D. Teague, H.F. Tsoi [ID], W. Vetens [ID], A. Warden [ID]

Authors affiliated with an institute or an international laboratory covered by a cooperation agreement with CERN

S. Afanasiev [ID], V. Andreev [ID], Yu. Andreev [ID], T. Aushev [ID], M. Azarkin [ID], A. Babaev [ID], A. Belyaev [ID], V. Blinov[105], E. Boos [ID], V. Borshch [ID], D. Budkouski [ID], M. Chadeeva [ID][105], V. Chekhovsky, A. Dermenev [ID], T. Dimova [ID][105], D. Druzhkin [ID][106], M. Dubinin [ID][95], L. Dudko [ID], A. Ershov [ID], G. Gavrilov [ID], V. Gavrilov [ID], S. Gninenko [ID], V. Golovtcov, N. Golubev [ID], I. Golutvin, I. Gorbunov [ID], A. Gribushin [ID], Y. Ivanov [ID], V. Kachanov [ID], L. Kardapoltsev [ID][105], V. Karjavine, A. Karneyeu [ID], L. Khein, V. Kim [ID][105], M. Kirakosyan, D. Kirpichnikov [ID], M. Kirsanov [ID], V. Klyukhin [ID], O. Kodolova [ID][107], D. Konstantinov [ID], V. Korenkov [ID], A. Kozyrev [ID][105], N. Krasnikov [ID], A. Lanev, P. Levchenko [ID][108], A. Litomin, O. Lukina [ID], N. Lychkovskaya [ID], A. Malakhov [ID], V. Matveev [ID][105], V. Murzin [ID], A. Nikitenko [ID][109,107], S. Obraztsov [ID], V. Oreshkin [ID], A. Oskin, V. Palichik [ID], V. Perelygin [ID], S. Petrushanko [ID], S. Polikarpov [ID][105], V. Popov, O. Radchenko [ID][105], V. Rusinov, M. Savina [ID], V. Savrin, D. Selivanova [ID], V. Shalaev [ID], S. Shmatov [ID], S. Shulha [ID], Y. Skovpen [ID][105], S. Slabospitskii [ID], V. Smirnov, D. Sosnov [ID], V. Sulimov [ID], E. Tcherniaev [ID], A. Terkulov [ID], O. Teryaev [ID], I. Tlisova [ID], A. Toropin [ID], L. Uvarov, A. Uzunian [ID], A. Vorobyev[†], N. Voytishin [ID], B.S. Yuldashev[110], A. Zarubin [ID], I. Zhizhin [ID], A. Zhokin [ID]

[†] *Deceased*

[1] *Also at Yerevan State University, Yerevan, Armenia*

[2] *Also at TU Wien, Vienna, Austria*

[3] *Also at Institute of Basic and Applied Sciences, Faculty of Engineering, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt*

[4] *Also at Ghent University, Ghent, Belgium*

[5] *Also at Universidade Estadual de Campinas, Campinas, Brazil*

[6] *Also at Federal University of Rio Grande do Sul, Porto Alegre, Brazil*

[7] *Also at UFMS, Nova Andradina, Brazil*

[8] *Also at Nanjing Normal University, Nanjing, China*

[9] *Now at The University of Iowa, Iowa City, Iowa, U.S.A.*

[10] *Also at University of Chinese Academy of Sciences, Beijing, China*

[11] *Also at University of Chinese Academy of Sciences, Beijing, China*

[12] *Also at Université Libre de Bruxelles, Bruxelles, Belgium*

[13] *Also at an institute or an international laboratory covered by a cooperation agreement with CERN*

[14] *Also at Suez University, Suez, Egypt*

[15] *Now at British University in Egypt, Cairo, Egypt*

[16] *Also at Birla Institute of Technology, Mesra, Mesra, India*

17 *Also at Purdue University, West Lafayette, Indiana, U.S.A.*

18 *Also at Université de Haute Alsace, Mulhouse, France*

19 *Also at Department of Physics, Tsinghua University, Beijing, China*

20 *Also at Ilia State University, Tbilisi, Georgia*

21 *Also at Tbilisi State University, Tbilisi, Georgia*

22 *Also at The University of the State of Amazonas, Manaus, Brazil*

23 *Also at Erzincan Binali Yildirim University, Erzincan, Turkey*

24 *Also at University of Hamburg, Hamburg, Germany*

25 *Also at RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany*

26 *Also at Isfahan University of Technology, Isfahan, Iran*

27 *Also at Bergische University Wuppertal (BUW), Wuppertal, Germany*

28 *Also at Brandenburg University of Technology, Cottbus, Germany*

29 *Also at Forschungszentrum Jülich, Juelich, Germany*

30 *Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland*

31 *Also at Physics Department, Faculty of Science, Assiut University, Assiut, Egypt*

32 *Also at Wigner Research Centre for Physics, Budapest, Hungary*

33 *Also at Institute of Physics, University of Debrecen, Debrecen, Hungary*

34 *Also at Institute of Nuclear Research ATOMKI, Debrecen, Hungary*

35 *Now at Universitatea Babes-Bolyai - Facultatea de Fizica, Cluj-Napoca, Romania*

36 *Also at Faculty of Informatics, University of Debrecen, Debrecen, Hungary*

37 *Also at Punjab Agricultural University, Ludhiana, India*

38 *Also at UPES - University of Petroleum and Energy Studies, Dehradun, India*

39 *Also at University of Visva-Bharati, Santiniketan, India*

40 *Also at University of Hyderabad, Hyderabad, India*

41 *Also at Indian Institute of Science (IISc), Bangalore, India*

42 *Also at University of Minnesota, Minneapolis, Minnesota, U.S.A.*

43 *Also at IIT Bhubaneswar, Bhubaneswar, India*

44 *Also at Institute of Physics, Bhubaneswar, India*

45 *Also at Deutsches Elektronen-Synchrotron, Hamburg, Germany*

46 *Now at Department of Physics, Isfahan University of Technology, Isfahan, Iran*

47 *Also at Sharif University of Technology, Tehran, Iran*

48 *Also at Department of Physics, University of Science and Technology of Mazandaran, Behshahr, Iran*

49 *Also at Helwan University, Cairo, Egypt*

50 *Also at Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Bologna, Italy*

51 *Also at Centro Siciliano di Fisica Nucleare e di Struttura Della Materia, Catania, Italy*

52 *Also at ENEA - Casaccia Research Center, S. Maria di Galeria, Italy*

53 *Also at Università degli Studi Guglielmo Marconi, Roma, Italy*

54 *Also at Facoltà Ingegneria, Università di Roma, Roma, Italy*

55 *Also at Scuola Superiore Meridionale, Università di Napoli 'Federico II', Napoli, Italy*

56 *Also at Fermi National Accelerator Laboratory, Batavia, Illinois, U.S.A.*

57 *Also at Laboratori Nazionali di Legnaro dell'INFN, Legnaro, Italy*

58 *Also at Università di Napoli 'Federico II', Napoli, Italy*

59 *Also at Ain Shams University, Cairo, Egypt*

60 *Also at Consiglio Nazionale delle Ricerche - Istituto Officina dei Materiali, Perugia, Italy*

61 *Also at Riga Technical University, Riga, Latvia*

62 *Also at Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia*

63 *Also at Consejo Nacional de Ciencia y Tecnología, Mexico City, Mexico*

64 *Also at Trincomalee Campus, Eastern University, Sri Lanka, Nilaveli, Sri Lanka*

65 *Also at Saegis Campus, Nugegoda, Sri Lanka*

66 *Also at INFN Sezione di Pavia, Università di Pavia, Pavia, Italy*

[67] *Also at National and Kapodistrian University of Athens, Athens, Greece*

[68] *Also at Ecole Polytechnique Fédérale Lausanne, Lausanne, Switzerland*

[69] *Also at Universität Zürich, Zurich, Switzerland*

[70] *Also at Paul Scherrer Institut, Villigen, Switzerland*

[71] *Also at Stefan Meyer Institute for Subatomic Physics, Vienna, Austria*

[72] *Also at Laboratoire d'Annecy-le-Vieux de Physique des Particules, IN2P3-CNRS, Annecy-le-Vieux, France*

[73] *Also at Near East University, Research Center of Experimental Health Science, Mersin, Turkey*

[74] *Also at Konya Technical University, Konya, Turkey*

[75] *Also at Izmir Bakircay University, Izmir, Turkey*

[76] *Also at Adiyaman University, Adiyaman, Turkey*

[77] *Also at Necmettin Erbakan University, Konya, Turkey*

[78] *Also at Bozok Universitetesi Rektörlügü, Yozgat, Turkey*

[79] *Also at Marmara University, Istanbul, Turkey*

[80] *Also at Milli Savunma University, Istanbul, Turkey*

[81] *Also at Kafkas University, Kars, Turkey*

[82] *Also at Istanbul Bilgi University, Istanbul, Turkey*

[83] *Also at Hacettepe University, Ankara, Turkey*

[84] *Also at Istanbul University - Cerrahpasa, Faculty of Engineering, Istanbul, Turkey*

[85] *Also at Ozyegin University, Istanbul, Turkey*

[86] *Also at Vrije Universiteit Brussel, Brussel, Belgium*

[87] *Also at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom*

[88] *Also at University of Bristol, Bristol, United Kingdom*

[89] *Now at Tata Institute of Fundamental Research-B, Mumbai, India*

[90] *Also at IPPP Durham University, Durham, United Kingdom*

[91] *Also at Monash University, Faculty of Science, Clayton, Australia*

[92] *Also at Università di Torino, Torino, Italy*

[93] *Also at Bethel University, St. Paul, Minnesota, U.S.A.*

[94] *Also at Karamanoğlu Mehmetbey University, Karaman, Turkey*

[95] *Also at California Institute of Technology, Pasadena, California, U.S.A.*

[96] *Also at United States Naval Academy, Annapolis, Maryland, U.S.A.*

[97] *Also at Beykent University, Istanbul, Turkey*

[98] *Also at Bingol University, Bingol, Turkey*

[99] *Also at Georgian Technical University, Tbilisi, Georgia*

[100] *Also at Sinop University, Sinop, Turkey*

[101] *Also at Erciyes University, Kayseri, Turkey*

[102] *Also at Horia Hulubei National Institute of Physics and Nuclear Engineering (IFIN-HH), Bucharest, Romania*

[103] *Also at Texas A&M University at Qatar, Doha, Qatar*

[104] *Also at Kyungpook National University, Daegu, Korea*

[105] *Also at another institute or international laboratory covered by a cooperation agreement with CERN*

[106] *Also at Universiteit Antwerpen, Antwerpen, Belgium*

[107] *Also at Yerevan Physics Institute, Yerevan, Armenia*

[108] *Also at Northeastern University, Boston, Massachusetts, U.S.A.*

[109] *Also at Imperial College, London, United Kingdom*

[110] *Also at Institute of Nuclear Physics of the Uzbekistan Academy of Sciences, Tashkent, Uzbekistan*