The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

An FPGA based track finder at L1 for CMS at the High Luminosity LHC

C. Amstutz, F. A. Ball, M. N. Balzer, J. Brooke, L. Calligaris, D. Cieri, E. J. Clement, G. Hall, T. R. Harbaum, K. Harder, P. R. Hobson, G. M. Iles, T. James, K. Manolopoulos, T. Matsushita, A. D. Morton, D. Newbold, S. Paramesvaran, M. Pesaresi, I. D. Reid, A. W. Rose, O. Sander, T. Schuh, C. Shepherd-Themistocleous, A. Shtipliyski, S. P. Summers, A. Tapper, I. Tomalin, K. Uchida, P. Vichoudis, M. Weber

Abstract

A new CMS Tracker is under development for operation at the High Luminosity LHC from 2025. It includes an outer tracker based on PT-modules which will construct tracker stubs, built by correlating clusters in two closely spaced sensor layers for the rejection of low transverse momentum track hits, and transmit them off-detector at 40MHz. If tracker data is to contribute to maintaining the Level-1 trigger rate under increased luminosity, a crucial component of the upgrade will be the ability to identify tracks with transverse momentum above 3GeV/c by building tracks out of stubs. A concept for an FPGA-based track finder using a fully time-multiplexed spatially pipelined architecture is presented, where track candidates are identified using a projective binning algorithm. Results from a hardware demonstrator system, where a slice of the track trigger will be constructed to help gauge the performance and requirements for a full system, will be included.

Presented at *IEEE-RT2016 IEEE-NPSS Real Time Conference (RT)*

An FPGA-Based Track Finder for the L1 Trigger of the CMS Experiment at the High Luminosity LHC

C. Amstutz, F. A. Ball, M. N. Balzer, J. Brooke, L. Calligaris, D. Cieri, E. J. Clement, G. Hall, T. R. Harbaum, K. Harder, P. R. Hobson, G. M. Iles, T. James, K. Manolopoulos, T. Matsushita, A. D. Morton, D. Newbold, S. Paramesvaran, M. Pesaresi, I. D. Reid, A. W. Rose, O. Sander, T. Schuh, C. Shepherd-Themistocleous, A. Shtipliyski, S. P. Summers, A. Tapper, I. Tomalin, K. Uchida, P. Vichoudis, M. Weber for the CMS Collaboration

Abstract—A new tracking system is under development for operation in the CMS experiment at the High Luminosity LHC. It includes an outer tracker which will construct stubs, built by correlating clusters in two closely spaced sensor layers for the rejection of hits from low transverse momentum tracks, and transmit them off-detector at 40 MHz. If tracker data is to contribute to keeping the Level-1 trigger rate at around 750 kHz under increased luminosity, a crucial component of the upgrade will be the ability to identify tracks with transverse momentum above 3 GeV/c by building tracks out of stubs. A concept for an FPGA-based track finder using a fully time-multiplexed architecture is presented, where track candidates are identified using a projective binning algorithm based on the Hough Transform. A hardware system based on the MP7 MicroTCA processing card has been assembled, demonstrating a realistic slice of the track finder in order to help gauge the performance and requirements for a full system. This paper outlines the system architecture and algorithms employed, highlighting some of the first results from the hardware demonstrator and discusses the prospects and performance of the completed track finder.

I. THE HIGH-LUMINOSITY LARGE HADRON COLLIDER

IN order to fully exploit the scientific potential of the Large Hadron Collider (LHC) [1], it is planned to operate the machine at a luminosity up to one order of magnitude above nominal design performance. The High-Luminosity LHC (HL-LHC) upgrade [2] is expected to take place during a 30 month shut-down around 2024, facilitating a peak luminosity of $5 - 7.5 \times 10^{34} \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$, corresponding to an average number of proton-proton interactions per 40 MHz bunch crossing, or
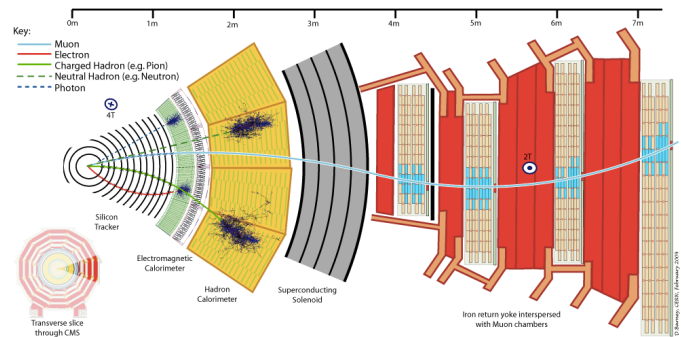
Fig. 1. Overview of the CMS detector, as a transverse slice through the barrel [3].

pileup (PU), of 140 to 200. With a targeted total integrated luminosity of $3000 \, \mathrm{fb}^{-1}$ the HL-LHC will enable precision Higgs measurements, searches for rare processes that may deviate from the Standard Model and further increase the high mass and low cross-section observation limits into the multi-TeV regime.

II. THE COMPACT MUON SOLENOID OUTER TRACKER UPGRADE

The Compact Muon Solenoid (CMS) is a large, general purpose particle detector at the LHC, designed to investigate a wide range of physics phenomena. It consists of a set of sub-detectors, including the tracking system, surrounding the interaction point, as shown in Fig. 1. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in [3].

The complete replacement of the CMS tracker will be necessary during the shut-down preceding the HL-LHC, primarily due to the expected radiation damage of the silicon sensors following ~ 15 years of operation. The HL-LHC environment will additionally provide a significant challenge for the new tracker [4]. It must maintain a high track reconstruction efficiency and a low misidentification rate under increased pileup conditions. To achieve this the occupancy must be kept at or below the 1% level throughout, requiring an increase in granularity. As a result of increased exposure, the radiation hardness of the tracker must also be improved.
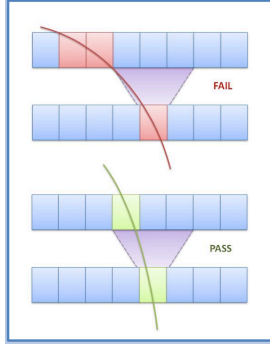
Fig. 2. Cluster matching in p_T-modules. Correlating closely spaced clusters between two sensor layers, separated by a few mm, allows discrimination of transverse momentum based on the particle bend in the CMS magnetic field. Only tracks with $p_T > 2 - 3\,\text{GeV/c}$ are transferred to the L1 trigger.
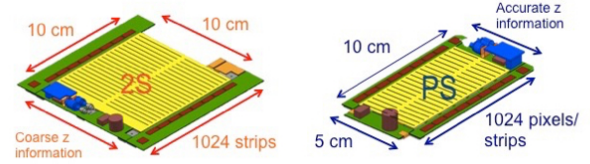

Fig. 3. The 2S module (left) and PS module (right), described in the text.
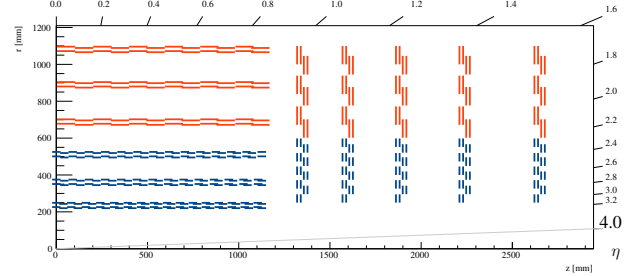

Fig. 4. One quadrant map of 2S (red) and PS (blue) module placement in the proposed outer tracker [4]. All results presented here assume this layout. However, in order to reduce construction costs and improve overall performance, the tracker is now expected to have tilted modules for the three PS barrel layers.

The Level-1 (L1) trigger is an event selection system based on custom electronics that uses coarse grained information from the calorimeter and muon sub-detectors to reject events that are not interesting for subsequent physics analysis. Under HL-LHC conditions, increasing the transverse momentum (p_T) or transverse energy (E_T) thresholds at the L1 trigger would not reduce the rate sufficiently without losses of potentially interesting events, unless some tracking information could be provided to the system. Track-based information would be able to reduce the trigger rate at L1 by validating trigger objects, for example in providing an improved p_T assignment to muon triggers which are a major cause of high background rates under increased pileup. However, it is not practical to transfer all tracking data to the L1 trigger. A novel design has therefore been proposed for the outer tracker upgrade, which allows a limited amount of tracking information to be sent to the L1 trigger.

The proposed solution [5], [6] utilises two sensors, closely separated (by order millimetres) in the track direction, to discriminate on track p_T based on its local bend within the 3.8 T magnetic field, see Fig. 2. Within these p_T-modules, charged particles will produce *stubs*, correlated pairs of clusters, if they are consistent with tracks of transverse momenta greater than a configurable threshold (typically 2-3 GeV/c). In a typical event approximately 98% of charged particles have a $p_T < 2\,\text{GeV/c}$ and these are not considered to be useful for event selection at L1. Therefore by transferring only the stubs to the L1 trigger it is expected that a rate reduction of ~ 10 is achievable [7], [8] enabling the use of lower bandwidth and lower power optical links for transmission off-detector.

Two p_T-modules are in development for the tracker upgrade, 2S *strip-strip* modules, and PS *pixel-strip* modules, see Fig. 3. The 2S modules are designed to be used at radii $r > 60\,\text{cm}$ from the beam axis, where the hit occupancies are lower. Both upper and lower sensors consist of $\sim 10\,\text{cm} \times 10\,\text{cm}$ silicon strip sensors, with a pitch of $90\,\mu\text{m}$ in r-ϕ and $5\,\text{cm}$ in z. The PS modules will be used at radii $20 < r < 60\,\text{cm}$ where the occupancies are highest. These consist of an upper silicon strip sensor and a lower pixelated sensor, both of dimension $\sim 5\,\text{cm} \times 10\,\text{cm}$, with a pitch of $100\,\mu\text{m}$ in r-ϕ. The pitch in z is $2.4\,\text{cm}$ for the strips, and $1.5\,\text{mm}$ for the pixels. The finer granularity afforded by the pixel sensors provides more accurate pointing resolution along the beam axis, which is crucial for identifying interaction vertices at L1 under high pileup conditions.

The stub finding correlation logic is executed by on-detector ASICs. Each 2S module contains 16 readout ASICs, known as CMS binary chips (CBCs) [9], [10] which each perform the stub finding logic on a 128 strip segment on upper and lower sensors, and transfer the stubs to concentrator ASICs on the corresponding half-module. The concentrator is capable of transmitting up to twelve stubs per eight 40 MHz bunch crossings. Data is expected to be optically transferred from 2S modules at 5.12 Gbps, depending on advancements in optical technology [11].

Fig. 4 depicts the proposed upgraded tracker layout, and the 2S and PS module placements. This includes six barrel layers, and five endcap disks on each side. The upgraded geometry extends the pseudorapidity range of the endcaps from $|\eta| = 2.5$ to $|\eta| = 3$; however, modules located at $\eta > 2.55$ are not expected to send data to L1.

An L1 track-trigger using fully reconstructed tracks would improve the performance of the L1 decision-making algorithms in a number of ways. Candidate tracks will be used to validate objects seen by the muon and calorimeter systems. Discrimination between genuine electron objects and $\pi_0 \rightarrow \gamma\gamma$ background will be made possible. Isolation of electrons, photons, muons and taus will be significantly improved, and the p_T resolution of muons at L1 will be greatly enhanced. Identification of primary vertices using tracks is expected to reduce the significant jet-based background rate [4].

In order to reconstruct tracks at L1, a number of challenges remain. The task can be broadly divided into three steps: data formatting and delivery, pattern recognition or track building, and fine track fitting; with an overall processing latency goal of $\sim 4\,\mu s$. Any feasible method will also need to take into account requirements from the first layer of off-detector tracker readout electronics known as the Data, Trigger and Control (DTC) [4] system, and detector cabling constraints. One approach to pattern recognition uses a fully time-multiplexed Hough Transform (HT) technique [12].

III. THE HOUGH TRANSFORM

A well known method of detecting geometric features, such as straight lines, in digital images, the Hough Transform [13], can be applied to the task of identifying tracks from tracker stubs. In the case of CMS, a Hough Transform can be used to find primary charged particles with $p_\mathrm{T} > 3\,\mathrm{GeV/c}$ in the r-ϕ plane of the outer tracker. The trajectories of these particles are bent in this plane due to the 3.8 T homogeneous magnetic field in the z direction (longitudinal to the beam) provided by the solenoid. Therefore, within the tracking volume, the radius of curvature R (in m) of a charged particle is a function of its transverse momentum p_T (in GeV) and charge q (in electronic charge),

$$R = \frac{p_\mathrm{T}}{1.14\,q} . \tag{1}$$

To simplify the equation describing the path of the particle, the radius of curvature of particles with $p_\mathrm{T} > 3\,\mathrm{GeV/c}$ can be considered constant, i.e. any energy loss of the particle, for example through multiple scattering, can be neglected to first order. The trajectory $\vec{r}$ may therefore be described by a circle equation,

$$R^2 = \left(\vec{r} - \vec{M}\right)^2 , \tag{2}$$

where $\vec{M}$ points to the circle centre.

As a second simplification, track finding can be restricted to particles from the interaction point (anywhere along the beam axis) since these primary tracks are most relevant for the L1 trigger. Therefore the circle centre is given by

$$\vec{M} = R\,\hat{\phi}_{\phi_0} \quad \text{with} \quad \hat{\phi}_{\phi_0} = \begin{pmatrix} -\sin\phi_0 \\ \cos\phi_0 \end{pmatrix} . \tag{3}$$

Only two parameters are necessary to describe a track in the r-ϕ plane, for example R and the initial azimuth angle of the track ϕ_0. Inserting this information, and a single measured position (r, ϕ) of a particle in (2) leads to

$$R^2 = r^2 - 2rR\hat{r}_\phi \cdot \hat{\phi}_{\phi_0} + R^2 \quad \text{with} \quad \hat{r}_\phi = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}. \tag{4}$$

In fact, an infinite number of different circles can be drawn between the origin and this measured position, however the corresponding track parameters are correlated

$$\frac{r}{2R} = \sin\left(\phi - \phi_0\right) . \tag{5}$$
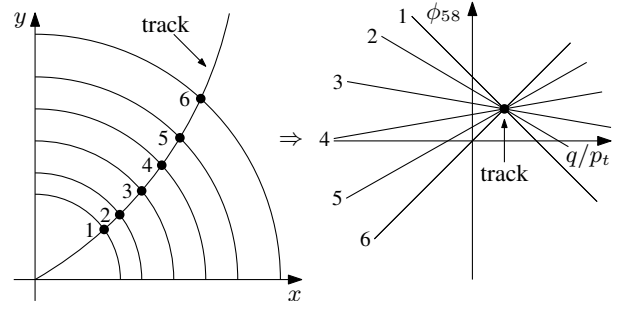

Fig. 5. Illustration of the Hough Transform. On the left-hand side is a sketch of one quarter of the tracker barrel in r-ϕ. The track of a single particle is drawn and the stubs from each of the six detector layers are shown as dots. On the right-hand side is the track parameter plane where the six corresponding Hough-transformed stubs are drawn as lines and their intersection identifies the track and its parameters $(q/p_\mathrm{T}, \phi_{58})$.

For high p_T, i.e. large R, the sine term may be linearised. This leads to the formula,

$$\phi_0 = \phi - \frac{0.57\,q}{p_\mathrm{T}} \cdot r . \tag{6}$$

This result shows how stub positions correspond to straight lines in the track parameter plane q/p_T-ϕ_0. As described by the Hough Transform, an intersection of those straight lines in the track parameter plane would identify a circle in the r-ϕ plane, consistent with the origin and all participating stubs. Fig. 5 visualises the use of the Hough Transform for six stubs produced by a primary track.

Since the radius of curvature of a particle with $p_\mathrm{T} > 3\,\mathrm{GeV/c}$ is greater than the outer radius of the tracker ($r = 1.2\,\mathrm{m}$), all such particles will reach the outermost extent of the tracker traversing a minimum number of six detector layers. A circle that is consistent with at least one stub from each layer should be identified as a track candidate, however to allow for detector inefficiencies, the threshold is set at a minimum of five detector layers.

In track parameter space, the gradients of the straight lines are given by the radius of the stubs, and are therefore always positive. To optimise the distribution of lines in track parameter space, the stub radius is transformed to $r_{58} = r - 58\,\mathrm{cm}$. As a consequence it is necessary to use ϕ_{58} as a track parameter, which is the ϕ coordinate of the track at a radius of 58 cm, to retain the form of (6):

$$\phi_{58} = \phi - \frac{0.57\,q}{p_\mathrm{T}} \cdot r_{58} . \tag{7}$$

Algorithmically, the Hough Transform in track parameter space can be achieved by constructing an array with a certain granularity, bounded by the full range of 2π in ϕ_{58} and all $|q/p_\mathrm{T}|$ above 3 GeV/c. The granularity should be as fine as possible to deliver the most precise track candidates, but course enough to take into account the simplifications described above (for example, misalignments due to multiple scattering) and the finite hit resolution of the tracker modules. Therefore the granularity of the track parameter plane for the entire tracker is given by 1024 rows in ϕ_{58} and 32 columns in q/p_T.

To find tracks, the stub positions in (r_{58}, ϕ) need to be transformed into straight lines with (7), so that they can be
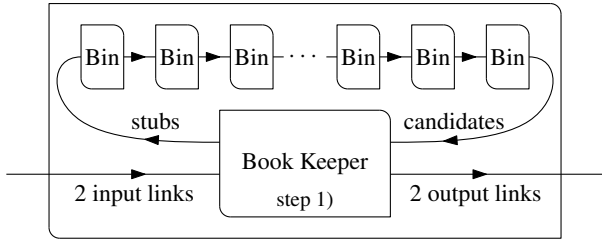
Fig. 6. Overview of the Hough Segment. Components are shown as boxes and data paths as lines, where arrows indicate the direction of communication.

filled accordingly into the track parameter array. One feature of the stub is that it contains a bend measurement, given by the difference in strips between the clusters on the upper and lower sensors in the module, which provides a rough estimate of the p_T of the particle that produced it. Therefore a stub only needs to be binned into the subset of q/p_T columns consistent with the stub bend. Valid track candidates correspond to elements in the track parameter array containing stubs from at least five different detector layers.

IV. IMPLEMENTATION

The Hough Transform track finder described above has been implemented in firmware for an FPGA. The design can be divided into two steps, the filling of the track parameter array with stubs, followed by the readout of stubs from the array cells that have been identified as track candidates. Each step has been implemented as a pipelined firmware, which can process one stub per clock cycle at 240 MHz. Since CMS is expected to produce on average about 12 k stubs per bunch crossing (pileup of 140) at 40 MHz, many parallel working arrays are necessary.

To achieve this parallelism the tracker can be subdivided into 288 segments (32 in ϕ and 9 in η). The regional segmentation requires duplication of stubs across segments leading to an expected 18 k stubs per bunch crossing for the entire tracker, i.e. an average of $\sim$60 stubs per segment.

Each segment will be processed by an independent unit, called a Hough Segment, which contains its own track parameter array. Fig. 6 gives an overview of a Hough Segment. Due to the regional segmentation in ϕ, each Hough Segment only needs to cover a sub-range in ϕ_{58} meaning that the track parameter array is 32 rows in ϕ_{58} and 32 columns in q/p_T. One Hough Segment consists of a Book Keeper and 32 Bins, where each Bin corresponds to a q/p_T column.

A. Input pipeline

As described above, the Hough Transform is implemented as two pipelines. The input pipeline describes the processing steps from incoming stubs to the creation of track candidates and starts with the Book Keeper.

1) Book Keeper: The Book Keeper connects the track parameter array with the segment I/O. It receives data, one stub per clock cycle, which is promptly stored within a 36 Kb (1 Kb = 1024 bits) block memory. The block memory address pointer of the stub is sent to the first Bin of the Hough
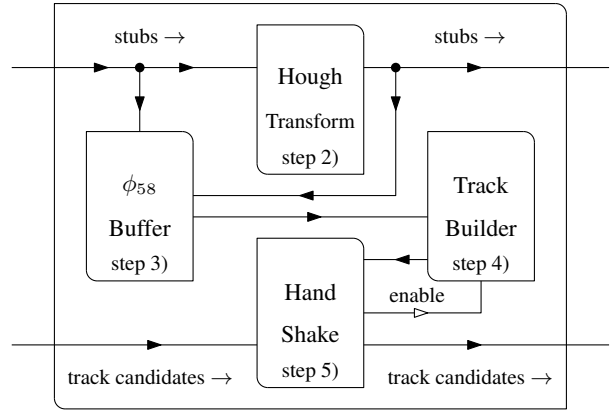

Fig. 7. Overview of one Bin, the component corresponding to one q/p_T column in the Hough Transform. All 32 Bins are daisy-chained together, starting and ending with the Book Keeper.

Segment, along with the (r_{58}, ϕ) coordinates of the stub, a layer identifier and the range of compatible q/p_T, according to the bend information.

The array itself is implemented as 32 daisy-chained Bins. On each clock cycle a stub propagates from Bin n to Bin $n + 1$. The components of a Bin are shown in Fig. 7. The stub propagation from Bin to Bin enables an iterative version of (7), where ϕ_{58} at the right boundary of the n_{th} Bin is given by:

$$\phi_{58}(n) = \phi_{58}(n-1) + \Delta\frac{q}{p_t} \cdot r_{58} , \qquad (8)$$

where $\Delta\frac{q}{p_T}$ is the constant width of an q/p_T column. The start value $\phi_{58}(0)$ is given by the stub coordinate ϕ. The calculation described in (8) is carried out in the next component called the Hough Transform.

2) Hough Transform: To minimise the number of calculations within the array, the r_{58} values are already expressed in units of the width of an q/p_T column. Equation (8) can therefore be implemented using a fabric addition, which produces the result within a clock cycle. Furthermore this unit can check if the stub is consistent with this particular q/p_T column. Since the range of columns compatible with the stub bend is pre-calculated, this check is a simple comparison of the range with a constant, given by the Bin number.

The ϕ_{58} rows of the array, for this column, are implemented in the Track Builder using memory structures. One advantage of using memory in this way is that the design can be made extremely compact. However, a straight line in the track parameter plane may cross up to two ϕ_{58} rows within a single q/p_T column with the array granularity defined above. Given that the Track Builder can only handle a single row value per clock cycle, it is necessary to duplicate and buffer a stub, if it belongs to two ϕ_{58} rows. This takes place in the ϕ_{58} Buffer component.

3) ϕ_{58} Buffer: The ϕ_{58} Buffer receives a stub and two ϕ_{58} values on each clock cycle, one at the left boundary and one at the right boundary of the q/p_T column. If a stub belongs to a single ϕ_{58} row only, it will be sent directly to the Track Builder. If however both values are valid but correspond to different ϕ_{58} rows, the stub and second row number will be

stored into a 18 Kb block memory FIFO for later processing. In the case that there is no valid input (i.e. during a gap in the data stream), the ϕ_{58} Buffer will send a stub from the FIFO to the Track Builder to maximise processing throughput.

4) Track Builder: The Track Builder sorts the stubs into 32 ϕ_{58} values using a segmented memory. The memory contains 64 pages, where one half is reserved for even events and the other half for odd events. Each page corresponds to a ϕ_{58} row, and has the capacity to store up to 32 stub pointers, such that the 64 pages fits into one 18 Kb block memory. The necessary book keeping, in particular the count of stored stubs per page, is achieved using distributed memory.

The Track Builder must also identify track candidates across any of the ϕ_{58} rows. The identification of pages as track candidates is again based on the use of distributed memory. For each row, a pattern of activated detector layers is stored as an 8-bit word, with one bit for each possible detector layer. On every clock cycle a valid stub arrives, a '1' will be written to the corresponding layer bit. If the word contains at least five '1's the corresponding ϕ_{58} row will be marked for readout.

The stream of incoming stubs ends inside this segmented memory, which therefore represents the end of the input pipeline.

B. Output pipeline

The goal is to create a pipeline of track candidate stubs, which propagate from the first Bin to the Book Keeper. The track candidate stubs are stored in a segmented memory inside each Track Builder. The readout of pages, which have been marked as track candidates, is controlled by the Hand Shake.

5) Hand Shake: The Hand Shake first shifts the track candidate stubs from the previous Bin to the next Bin along until there are no more stubs in the pipeline. Then it enables the readout of its Track Builder, such that a contiguous block of track candidate stubs will be created. The ϕ_{58} row and the Bin number, i.e. the q/p_T column, will be attached to the stub pointer by the Track Builder.

6) Book Keeper: The Book Keeper initiates the readout at the end of an input packet of stubs from an event, and after 32 clock cycles receives a stream of track candidate stubs for this event from the last Bin in the daisy-chain. A track candidate stub consists of a ϕ_{58} value, a q/p_T value, and a stub pointer. The stub pointer corresponds to the position in the 36 Kb memory where the full stub information was stored at the beginning of the input pipeline, which is then extracted by the Book Keeper. As a last action, the Book Keeper formats the full stub information and track parameters and transfers the data to the output of the Hough Segment.

Table I shows the resource utilisation of one Bin and Table II shows the utilisation of one Hough Segment, including LUTs (Look Up Tables), LUTRAMs (distributed RAMs), FF (Flip Flops) and BRAM (Block RAMs), for the Xilinx Virtex-7 XC7VX690T FPGA [14]. Through the smart use of common memory structures, it is possible to map the complex Hough Transform array into the FPGA in an extremely compact way. Division of the array into daisy-chained Bins is particularly

TABLE I
RESOURCE UTILISATION OF ONE BIN BASED ON THE XILINX VIRTEX-7 XC7VX690T.

Resource	Number Used	in %
LUT	188	0.04
LUTRAM	26	0.01
FF	204	0.02
BRAM	1	0.07

TABLE II
RESOURCE UTILISATION OF ONE SEGMENT BASED ON THE XILINX VIRTEX-7 XC7VX690T.

Resource	Number Used	in %
LUT	6014	1.39
LUTRAM	836	0.48
FF	6718	0.78
BRAM	33	2.24

advantageous, as it enables highly flexible placement and routing possibilities.

V. THE HARDWARE DEMONSTRATOR

The purpose of the demonstrator system is to run Monte Carlo simulated physics samples, with pileup, under HL-LHC conditions for an upgrade tracker geometry, through hardware and ensure matching with emulation software, in order to validate expected performance within latency constraints. A scalable slice of the track finder has been designed to allow the demonstration of the concept using currently available technology. One demonstrator slice can process at a time 1/8 of the tracker in ϕ, all of tracker in η, and one in every 36 bunch crossings. One can sequentially run data for all eight ϕ-octants through the demonstrator hardware, allowing results to be obtained for the entire tracker.

The hardware demonstrator, located at the CERN Tracker Integration Facility (TIF) is installed in a standard LHC rack, which provides power and cooling. The demonstrator consists of one custom Schroff dual-star MicroTCA crate, equipped with a commercial NAT MicroTCA Carrier Hub (MCH) for Gigabit Ethernet communication via the backplane, and a CMS specific auxiliary card known as the AMC13 [15] for synchronisation, timing and control. The Schroff crate is also equipped with eleven Imperial Master Processor, Virtex-7, Extended Edition (MP7-XE) double width AMC cards [16], which act as the processing boards for the track finder demonstrator. The MP7 was originally developed for the CMS L1 calorimeter trigger upgrade, which was installed in CMS in 2015. Each MP7 is equipped with a Xilinx Virtex-7 XC7VX690T FPGA, and 12 Avago Technologies MiniPOD optical transmitters/receivers, each providing 12 optical links running at up to 10.3 Gbps, for a total optical bandwidth of 0.74 Tbps in each direction.

The MP7 development group also provides a generic infrastructure firmware, which segregates core tasks such as transceiver buffering, I/O formatting, external communication, and configuration, from the algorithm itself. This generic core allows a system such as the demonstrator to be built up of
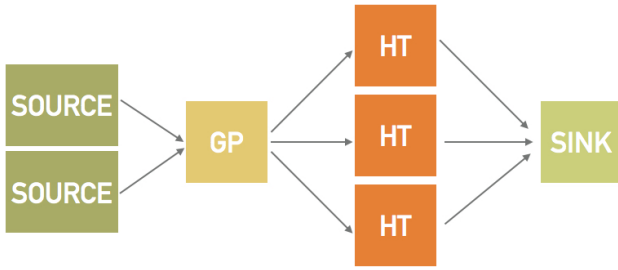
Fig. 8. The demonstrator system consists of four layers of MP7s; source, Geometric Processor (GP), Hough Transform (HT) and sink. The data moves downstream on optical links. Thirty-six optical links connect each source board to the GP. Twenty-four links connect the GP to each HT board. Each HT board is also connected to the sink with 24 optical links.

Fig. 9. The demonstrator crate is equipped with 11 MP7-XE boards, an AMC13, MCH and the required optics.

firmware blocks, each residing on a single MP7-XE, daisy-chained together with high speed optics. Using multiple MP7s in this way, one can easily extrapolate to the FPGA resources that may be available in a future processing card, meaning that final system performance can be estimated with currently available technology. This also allows firmware tasks to be easily divided between personnel, if common I/O formats between the firmware blocks are defined. These firmware blocks and the connections between them are shown in Fig. 8 and described below.

Seven MP7-XE boards are currently used for the demonstrator chain. Two of these boards, named sources, contain large buffers that can store up to 30 events of stub data for a single detector octant. The sources represent data from a set of up to 72 DTCs. The stub data from the DTCs is injected into the large buffers on the source boards via IPBus [17], and is already pre-formatted in a 48-bit global coordinate scheme. Each source provides a stream of data to the downstream board on 36 links, equivalent to the DTCs that make up adjacent detector octants. Input data from two adjacent octants is required, to be able to handle tracks that traverse the regional boundary.

The Geometric Processor (GP) board must format the 48-bit stubs into the 64-bit stubs required by the Hough Transform (HT) board, and assign each stub to one of four sub-sectors in ϕ, and nine sub-sectors in η, duplicating across sector boundaries when necessary. In order to simplify the HT firmware, the GP assigns stubs a layer ID (0-5 in barrel, 0-4 in endcaps), and a minimum/maximum viable p_T column in the HT, based on the stub bend. The duplication rules across ϕ are tuned for a $p_T = 3$ GeV/c particle, and utilize both the ϕ coordinate of the stub, and the stub bend information, to keep the duplication rate below 25% without efficiency loss.

Although two MP7-XEs contain the FPGA logic resources required to run the HT for a ϕ-octant, three demonstrator boards are allocated, to allow for future optimisations and additions to the firmware.

Downstream of the HT is the sink board. The sink runs identical firmware to the two sources, and can buffer the HT output from about 30 simulated physics events, before being read-out with IPBus. Additional boards are also installed in the demonstrator crate, w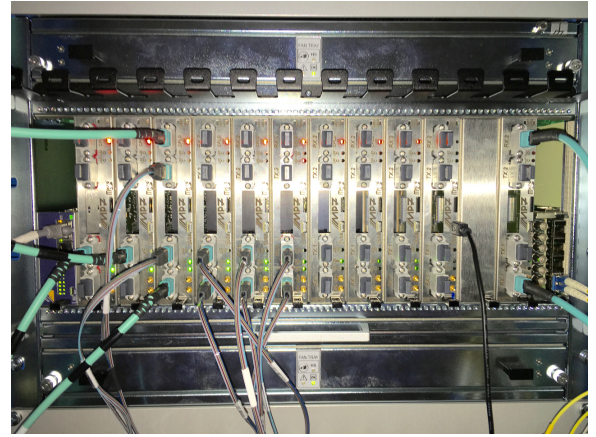hich allow testing of individual firmware blocks, and single board data taking in parallel with full demonstrator operation. The demonstrator crate is shown in Fig. 9.

In addition to the primary demonstrator at the CERN TIF, there are smaller single or dual-MP7 set-ups at Rutherford Appleton Laboratory (RAL), Imperial College London, and CERN. These allow for development and validation of individual firmware blocks, before they are integrated in the full demonstrator.

VI. Hardware Demonstrator Results

Simulated physics events up to a pileup of 200 interactions per bunch crossing have been run through the demonstrator system. Studies so far have focused on $\mu^+\mu^-$ and $t\bar{t}$ events at pileup of 0, 140 and 200. The demonstrator software framework allows simulated physics samples generated in the official CMS Software (CMSSW) to be converted into text files which are then injected into the hardware demonstrator via IPBus. The output of the hardware is then converted back into a CMSSW format, and is compared with the results of emulation software running on the same simulated physics event. This way it is possible to compare the results of the hardware and the emulation software, validating any simulation results of track finder performance. Two different versions of the emulation software have been developed. One version uses integer precision, and is clock-cycle accurate. It is written using the CIrcuit DAta Flow (CIDAF) framework [18]. The other version of the emulator is simpler, faster, and also uses integer precision. It is, however, not designed to be clock cycle accurate. All results in this paper use this latter version of the emulator.

In Fig. 10, 11 and 12 hardware results are plotted as black points, alongside software emulation (red lines). All plots were generated with a dataset of 1000 bunch crossings with the specified physics conditions.

Excellent matching between hardware and software has been measured with both $t\bar{t}$ and $\mu^+\mu^-$ signals, at a pileup of 0, 140 and 200. Fig. 10 show the average track rates for $\mu^+\mu^-$ at pileup of 140, and $t\bar{t}$ at pileup of 200, respectively. An average hardware/software matching of 99.5% is observed
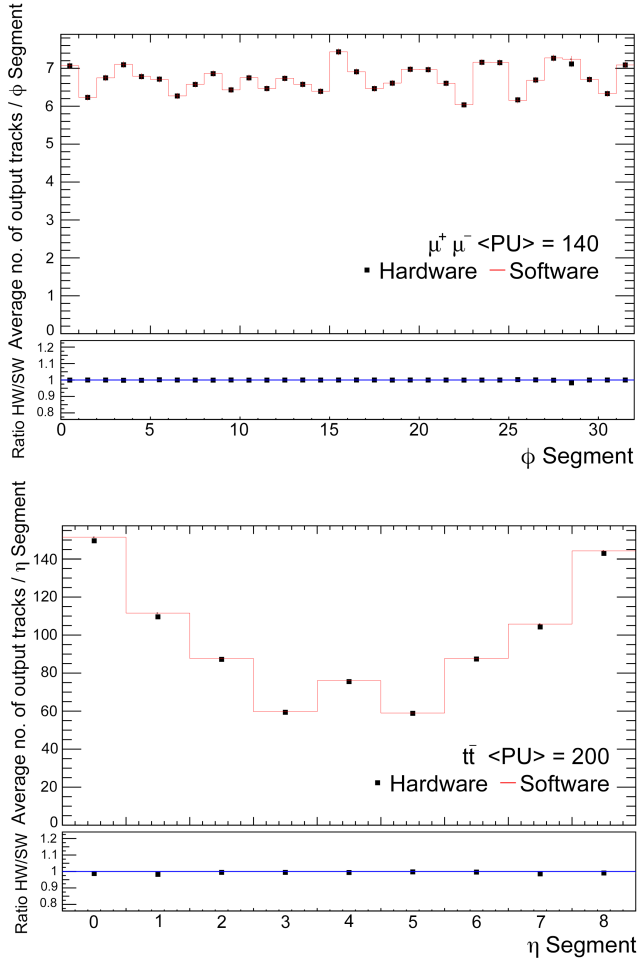
Fig. 10. Average track rate vs ϕ segment (0-31) for hardware demonstrator and emulation software, $\mu^+\mu^-$ at pileup 140 (above). Average track rate vs η segment (0-8) for hardware demonstrator and emulation software, $t\bar{t}$ at pileup 200 (below). Ratio plots are also provided below the rate plots, where the ratio of hardware rate over software rate (black point) is shown alongside unity (blue line).

for $t\bar{t}$ inclusive tracks at pileup of 200. At up to 200 pileup, average $\mu^+\mu^-$ matching is greater than 99.9%.

Fig. 11 and 12 show efficiency as measured in both hardware and emulation software for finding $\mu^+\mu^-$ and $t\bar{t}$ inclusive tracks at pileup of 200. Tracking efficiency here is measured relative to Monte Carlo simulation truth tracks, with $p_T > 3\,\mathrm{GeV/c}$, and stubs in a minimum of five tracker layers. An overall tracking efficiency of 99.79% is observed for $\mu^+\mu^-$ signals in both hardware and software. For $t\bar{t}$ inclusive tracks, an overall tracking efficiency of 97.58% and 98.28% are observed in hardware and software respectively. This small discrepancy of 0.7% between hardware and software efficiency appears primarily at low $p_T (< 10\,\mathrm{GeV/c})$ and high $|\eta|(> 1.5)$.

A stub rate reduction factor of 9.95 has been measured in hardware at pileup of 140, see Table III. For the highly occupied conditions of $t\bar{t}$ events at pileup of 200, the rate reduction factor is measured to be 3.76.

Latency measurements of the demonstrator chain have been made, see Table IV. The latency of the HT stage was measured at 1092 ns, while the latency of the GP stage was
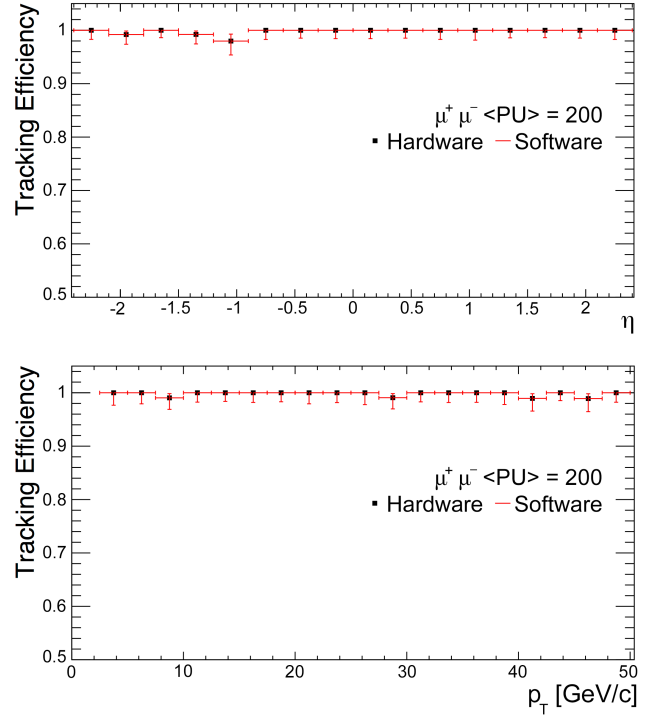

Fig. 11. Demonstrator results for $\mu^+\mu^-$ signals at pileup of 200. Tracking efficiency vs η for hardware demonstrator and emulation software (above). Tracking efficiency vs p_T for hardware demonstrator and emulation software (below). The sample size is 1000 bunch crossings. Tracking efficiency is measured relative to Monte Carlo simulation truth tracks, with $p_T > 3\,\mathrm{GeV/c}$, and stubs in a minimum of five tracker layers.

TABLE III
MEASURED RATE REDUCTION IN DEMONSTRATOR. NOTE THAT STUB SIMULATION CONDITIONS WERE SLIGHTLY DIFFERENT IN THE $t\bar{t}$ AND $\mu^+\mu^-$ SAMPLES.

Pileup	Signal	Stubs in	Stubs out	Tracks out
140	$\mu^+\mu^-$	15920	1599	215
140	$t\bar{t}$	16131	2483	322
200	$\mu^+\mu^-$	25382	5529	736
200	$t\bar{t}$	25081	6665	875

measured to be 280 ns, for a total algorithmic latency of 1372 ns. Measurements of the infrastructure latency have also been made, which includes optical link traversal time and serialisation/de-serialisation. The total infrastructure latency is 483 ns. When measured for the entire demonstrator chain of source to sink at once, or measured individually for each layer and summed, these results agree. The latency of the system is fixed, regardless of pileup, or segment occupancy.

VII. SUMMARY

The LHC luminosity will be increased by up to an order of magnitude from 2026 onwards, in order to maximise the potential physics reach of the collider. It will also be necessary to upgrade the CMS outer tracker at this time. The high pileup conditions expected at HL-LHC necessitate the incorporation of tracking information early in the triggering chain to maintain the L1 selection rate and physics performance.
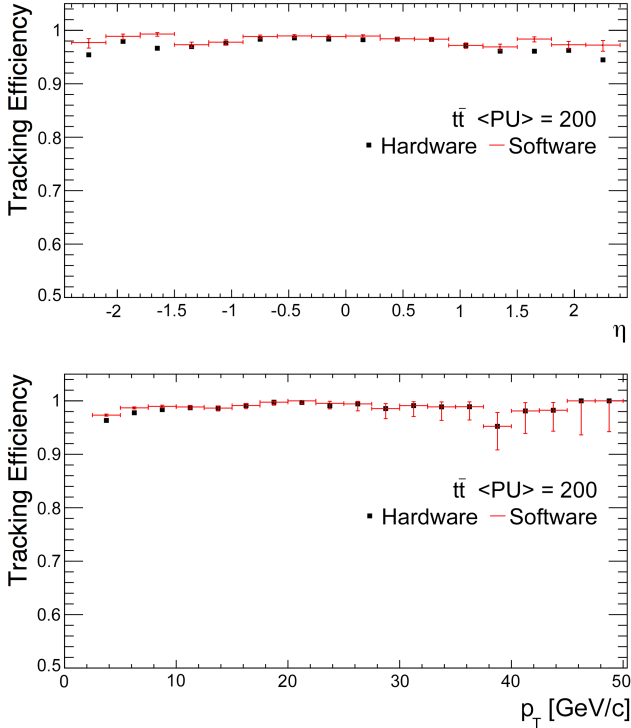
Fig. 12. Demonstrator results for $t\bar{t}$ signals at pileup of 200. Tracking efficiency vs η for hardware demonstrator and emulation software (above). Tracking efficiency vs p_T for hardware demonstrator and emulation software (below). The sample size is 1000 bunch crossings. Tracking efficiency is measured relative to Monte Carlo simulation truth tracks, with $p_T > 3\,\text{GeV/c}$, and stubs in a minimum of five tracker layers.

TABLE IV
DEMONSTRATOR LATENCY MEASUREMENTS

	Latency [ns]
Source event buffers $\rightarrow$ GP rx buffers	179
GP tx buffers $\rightarrow$ HT rx buffers	142
HT tx buffers $\rightarrow$ Sink event buffers	162
Infrastructure total	**483**
GP algorithm	280
HT algorithm	1092
Algorithm total	**1372**

CMS plans on selecting only hits from charged particles compatible with $p_T > 2 - 3\,\text{GeV/c}$ by correlating particle hits between stacked silicon sensors, which can then be read out to a L1 track-finder at an acceptable rate. This information will then be combined with muon and calorimeter objects before the final L1 decision is made.

A design for a L1 track-finder utilising Time Multiplexed 2D Hough transforms in FPGAs has been proposed. Simulation results demonstrate a high track-finding efficiency, and a rate reduction of order 10 at 140 pileup.

Firmware for the initial stages of this track-finder has been developed. By regionalising the detector data for independent parallel Hough Transform segments, the full pattern recognition algorithm can easily be scaled, distributing over several FPGAs if necessary. Together with a well balanced and

compact implementation of each segment, use of the Hough Transform to find tracks under HL-LHC conditions is feasible even in todays technology.

This design has been demonstrated in hardware, using a number of MP7 boards in a MicroTCA crate. The demonstrator slice can find tracks in all of η, 1/8 of the tracker in ϕ, and every 1/36 bunch crossings in time. However, each ϕ-octant can be processed in turn, to take data for the entire tracker. Fully simulated physics events including $t\bar{t}$ and $\mu^+\mu^-$ at pileup of 0, 140 and 200 have been processed by the demonstrator hardware. Results show excellent matching (99.5% in high occupancy physics conditions) between the hardware demonstrator and the emulation software, and tracking efficiencies in excess of 97.5%. Stub rate reduction with the Hough transform by a factor of 10 at 140 pileup has been demonstrated. Latency measurements made with hardware demonstrate a fixed algorithmic latency of 1372 ns, well within the allocated $4\,\mu s$ for track primitive generation. Downstream stages will soon be added to the demonstrator that will filter and fit the tracks in preparation for correlation with muon chambers and calorimeter information.

REFERENCES

[1] L. Evans and P. Bryant, *LHC Machine*, 2008 JINST 3 S08001, Available: http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[2] G. Apollinari, I. Bejar Alonso, O. Bruning, M. Lamont, L. Rossi *High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report*, CERN, Geneva, 2015, DOI: 10.5170/CERN-2015-005.

[3] The CMS Collaboration (S. Chatrchyan et al.), *The CMS experiment at the CERN LHC*, 2008 JINST 3 S08004, DOI: 10.1088/1748-0221/3/08/S08004.

[4] The CMS Collaboration, *Technical Proposal for the Phase-II Upgrade of the Compact Muon Solenoid*, CERN-LHCC-2015-010. June 2015, Available: https://cds.cern.ch/record/2020886.

[5] J. Jones, G. Hall, C. Foudas, A. Rose, *A Pixel Detector for Level-1 Triggering at SLHC*, 11th Workshop on Electronics for LHC Experiments, Heidelberg, September 2005, CERN Report CERN-2005-011(2005) 130-134.

[6] M. Pesaresi, *Development of a new Silicon Tracker for CMS at Super-LHC*, Imperial College thesis (2010), Available: https://workspace.imperial.ac.uk/highenergyphysics/Public/theses/.

[7] M. Pesaresi and G. Hall, *Simulating the performance of a pT tracking trigger for CMS*, 2010 JINST 5 C08003, DOI: 10.1088/1748-0221/5/08/C08003.

[8] G. Hall, M. Raymond and A. Rose, *2-D PT module concept for the SLHC CMS tracker*, 2010 JINST 5 C07012, DOI: 10.1088/1748-0221/5/07/C07012.

[9] M. Raymond *et al.,* The CMS binary chip for microstrip tracker readout at the SLHC*, 2012 JINST 7 C01033, DOI:10.1088/1748-0221/7/01/C01033.*

[10] *G. Hall, M. Pesaresi, M. Raymond, D. Braga, L. Jones, P. Murray, M. Prydderch et al., CBC2: a CMS microstrip readout ASIC with logic for track-trigger modules at HL-LHC, Nucl. Instrum. Meth. A **765** (2014) 214, DOI: 10.1088/1748-0221/7/10/C10003.*

[11] *F. Ravera for the CMS collaboration, CMS Tracker upgrade for HL-LHC: R&D plans, present status and perspectives, Nucl. Instrum. Meth. A **824** (2016) 455-458, Conference: C15-05-24.1 Proceedings, DOI: 10.1016/j.nima.2015.09.029.*

[12] *G. Hall, A Time-Multiplexed Track-Trigger for the CMS HL-LHC upgrade, Frontier Detectors for Frontier Physics: 13th Pisa Meeting on Advanced Detectors, La Biodola, Isola D'elba, Italy, 24 - 30 May 2015, CMS-CR-2015-109, Available: https://cds.cern.ch/record/2027712.*

[13] *P. V. C. Hough, Method and means for recognizing complex patterns, December 18th 1962, US Patent 3,069,654.*

[14] *Xilinx, 7 Series FPGAs Overview, Product Specification, DS180 (v1.17) May 27, 2015, Available: http://www.xilinx.com/support/documentation/data_sheets/ds180_7Series_Overview.pdf.*

[15] *E. Hazen, A. Heister, C. Hill, J. Rohlf, S.X. Wu and D.Zou, The AMC13XG: a new generation clock/timing/DAQ module for CMS MicroTCA, 2013 JINST 8 C12036, DOI: 10.1088/1748-0221/8/12/C12036.*

[16] *K. Compton, A. Rose et al., The MP7 and CTP-6: multi-hundred Gbps processing boards for calorimeter trigger upgrades at CMS, 2012 JINST 7 C12024, DOI: 10.1088/1748-0221/7/12/C12024.*

[17] *C. Ghabrous Larrea, K. Harder, D. Newbold, D. Sankey, A. Rose, A. Thea and T. Williams, IPbus: a flexible Ethernet-based control system for xTCA hardware, 2015 JINST 10 C02019, DOI:10.1088/1748-0221/10/02/C02019.*

[18] *C. Amstutz et al. Emulation of a prototype FPGA track finder for the CMS Phase-2 upgrade with the CIDAF emulation framework, IEEE Real Time Conference, 5 Jun 2016, Proceedings of this conference.*