The LHCb experiment at the LHC will upgrade its detector to take data at luminosities of above $10^{33}$ cm$^{-2}$s$^{-1}$ from 2018 onwards. The fully upgraded detector will use a new front-end electronics to readout collision events at 40MHz. A new fully software based trigger will be implemented and will increase the trigger efficiencies, specially in the hadronic decay channels. The silicon vertex detector baseline design will use pixel technology. The upgraded detector will accumulate in excess of 50fb$^{-1}$ of data, giving radiation doses of $4 \times 10^{15}$ n$_{eq}$ and above at the edge of the silicon. This requires an efficient cooling system to avoid thermal runaway. The existing cooling system based on evaporative CO$_2$ will be adapted. A strong R&D program has started on sensor material and radiation hardness, thickness, guard ring structures and the final module layout. In parallel the design of a new radiation hard ASIC for pixel readout, namely VELOpix, is progressing well. The main challenge will be the large multi Gbps data rates on and off this chip. In terms of precision, pattern recognition and impact parameter resolution, the goal is to achieve the same or better performance than the current vertex detector at the LHCb experiment.
The LHCb vertex detector upgrade

1. The LHCb upgrade: scenario and strategy

The LHCb experiment [1] at the LHC is currently running with excellent performance and it is expected that the experiment will collect an integrated luminosity in proton-proton collisions of around 1 \( \text{fb}^{-1} \) by the end of 2011. It is planned to continue operating the experiment in its present layout until 2018, accumulating in excess of 5\( \text{fb}^{-1} \) and allowing the experiment to cover its primary physics goals, namely the search for New Physics via the measurement of CP asymmetries and rare decays of b and c quarks. The time to double statistics then becomes too long to continue operating the experiment without a major upgrade. Even though LHCb is already operating beyond its design luminosity with \( \mathcal{L} = 2 \times 10^{32} \text{cm}^{-2} \text{s}^{-1} \), the LHC is already capable of delivering much higher luminosity to the experiment. LHCb plans to take advantage of this with an upgrade planned for 2018 [2]. The experiment will boost its data taking capabilities by reading out the entire detector at 40 MHz, allowing it to run at higher luminosity (\( \mathcal{L} > 20^{33} \text{cm}^{-2} \text{s}^{-1} \)) and to increase the trigger efficiencies, specially in decays to hadronic final states. The upgrade will allow LHCb to acquire an integrated luminosity of > 50 \( \text{fb}^{-1} \) (5 \( \text{fb}^{-1} \)/year) at a higher center of mass energy of \( \sqrt{s} = 14 \text{ TeV} \). The increase of luminosity is achieved by changing the beam parameters at the interaction point and does not depend on any LHC machine upgrade, hence the LHCb upgrade is independent of (but compatible with) the plans for the HL-LHC. The large data samples of b and c hadron decays will allow for more precise measurements in the flavour sector. These measurements have a wide-ranging sensitivity to New Physics effects and are a powerful method to both search for New Physics and to characterize its nature when found. The scientific goals of LHCb, however, extend far beyond quark-flavour physics. The upgraded experiment can be regarded as a multi-purpose detector in the forward direction. The unique acceptance, coupled with the flexible trigger, will enable LHCb to make measurements that are either complementary to, or of higher sensitivity than, those which are done at the LHC general purpose experiments (ATLAS/CMS).

At the present time LHCb trigger system [3, 1] has two levels: Level-0 (L0) is a hardware trigger while the High Level Trigger (HLT) consists of a software application which runs on a CPU-farm. The purpose of L0 is to reduce the rate of crossings with interactions to a maximal event rate of 1MHz (determined by the front-end electronics) so the HLT can process them in the CPU-farm. This hardware trigger uses information from objects of high transverse energy in the calorimeters and the muon chambers. Increasing the collision rate by running at higher luminosity requires that the thresholds of the L0 trigger should be raised. This leads, in particular for the hadronic channels, to a drop efficiency and even a lower total event yield. The time available to this first level hardware trigger is too short to try to implement a more complex decision, even if extra information could be made available. It was decided to increase the readout rate to 40MHz and implement a flexible event filtering in software on a larger CPU-farm. This implies that all the front-end electronics must be redesigned or adapted to transfer data of all the crossings to this CPU-farm and to zero suppress the data to reduce the required bandwidth. A low level hardware trigger (LLT) with similar functionality to L0 will be implemented. The LLT will enrich the selected sample with interesting events (not simply pre-scale to a rate acceptable by the DAQ and the CPU-farm like the L0), will protect against occupancy fluctuations that prevent full readout, and will have a tunable output rate to cope with the insufficient CPU-farm power in the beginning. The high level algorithms running in the CPU-farm will have all the event information available. It is
anticipated that a factor up to 2 in efficiency can be reached for the hadronic modes, which can be combined with the gain of running at higher luminosity. This trigger architecture will have an output rate to storage of $\sim$20kHz.

2. The vertex detector upgrade

In the case of the vertex detector, the front-end electronics is strongly integrated with the silicon sensor. This implies that for the upgrade the detector modules have to be completely redesigned and rebuilt. Furthermore, they have to cope with the increased radiation environment (a factor 10) which goes beyond the current design. The baseline choice for the replacement of the present vertex detector (Vertex Locator, VELO [4]) is a pixel based device [2, 5], based on an evolution of the Timepix/Medipix family of chips [6]. This consists of a $256 \times 256$ array of square pixels, which brings advantages in the pattern recognition in comparison to the $R$ and $\phi$ strip geometry used in the current VELO with the absence of fake combinatorial hits. Furthermore the equal precision in both orientations will halve the required number of measuring planes and therefore the total material budget.

It is intended to reuse many systems of the existing VELO detector, such as the mechanics and motion system, the vacuum systems, the power systems (LV/HV) and the evaporative CO$_2$ cooling system. The upgraded VELO will keep the concept of retractable detector halves around the beam axis, but the silicon modules and foil will be redesigned to meet the detector challenges outlined below: the high irradiation environment that will demand a efficient and low-mass cooling system to avoid thermal runaway of the silicon, the enormous on- and off-chip data rates, the low total material budget required to achieve or improve the excellent hit resolution ($\sim 4$-10$\mu$m) of the current detector [7].

2.1 Environment: Irradiation dose, data rates and occupancies

In both the current and upgraded detector the trigger algorithms of LHCb rely on the impact parameter cut. Therefore the impact parameter resolution is a very important performance number for the detector. This resolution, among other parameters, depends strongly on the radius of the first measured point [8]. It is therefore crucial to place the sensor as close as possible to the beam line during the proton-proton collisions. The sensor edges are placed at 7 mm from the beam. After 50 fb$^{-1}$ of integrated luminosity they will accumulate a dose of up to $0.41 \times 10^{16}$ n$_{eq}$ cm$^{-2}$ or 185Mrad, see figure 1. Recent studies [9] have shown that 300$\mu$m thick silicon irradiated to these levels still delivers a signal of 10ke$^{-}$/MIP. That dose will induce leakage currents in 200$\mu$m thick silicon sensors of $I_{\text{leak}} \simeq O(100) \mu$A/cm$^2$ at -15$^\circ$C and a bias voltage of $\sim 900$V at the inner most radius. The corresponding heat dissipation, $I_{\text{leak}}$V heats the silicon, which in turn results in larger leakage current. An efficient heat removal is mandatory to avoid thermal runaway at the edge of the sensor from bulk current and heat injected by the readout ASIC. This has to be implemented without introducing too much material in the acceptance. The readout electronics, apart from having a very low noise to cope with the reduced signal, has to be extremely radiation hard and single event upset tolerant.

The average particle density per crossing closest to the beam will be $\sim 5$ particles/cm$^2$/crossing and decrease with radial distance from the beam as $r^{-1.9}$, see figure 2. Taking the model of a
256×256 matrix pixel detector and assuming an average cluster size of 2 leads to the conclusion that for the innermost readout chip the data rate would be $\sim 13\text{Gbit/s}$ with a total data rate of around 3000 Gbit/s for the entire detector. This would require a minimum of 940 data links running at 3.2 Gbit/s. Furthermore, extracting the data to the pixel periphery becomes non-negligible at these rates, and fundamental design changes to the front-end readout chip are needed. These will be discussed later in the article. The above particle density give a rather low occupancy per event ($\sim 10^{-4}$) for a pixel size of $55\times 55\mu m^2$.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{irradiation_dose}
\caption{Irradiation dose as a function of radius.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{particle_occupancy}
\caption{Particle occupancy per event (simulation).}
\end{figure}

\subsection{2.2 Pixel module design and detector layout}

The pixel module design is depicted in figure 3. Four 3-ASIC pixel tiles are assembled in a L-shaped layout on alternating sides of a High Thermal Conductivity (HTC) 400µm substrate. The L-shaped design ensures that the pixel columns are oriented in a way that the maximum rate is experienced at the end of the column and not along its full length, and the bandwidth is shared among the columns. The pixel tiles, (Fig. 4 top) consist of three readout ASICs, with dimensions $14mm \times 18mm$, bump bonded on a common silicon sensor of a size of $43\times 15mm$ and with a single guard ring of $500\mu m$. Between ASICs, the use of elongated pixels avoids loss of detection efficiency. Past experience indicates that a balanced 2-sided design keeps the structure planar and eases the construction minimizing the distortions due to temperature variations both during construction and operation. The four sensor tiles are arranged on opposite sides of the substrate such that they overlap in the transverse plane minimizing dead areas caused by guard rings or peripheral circuitry of the ASIC (Fig. 4). For the innermost ASICs, metal traces on the substrate or kapton bring the signals to miniaturized connectors. The HTC substrate is made of pCVD diamond. It provides the necessary mechanical stability and drains the heat to the cooling channel. This substrate extends about 1cm beyond the edge of the tiles and allows the attachment of a cooling channel and the installation of the above mentioned miniaturized connectors. The outer part of the module where the material is less critical, consists of TPG (Thermal Pyrolytic Graphite) frame which acts as a support for the cooling and all the readout infrastructure. So far in this design, the innermost edge of the silicon is a 7mm from the beam axis and the active area starts at 7.5mm,
The LHCb vertex detector upgrade

Abraham Gallas

Figure 3: Layout of the upgraded VELO pixel station. A station consists of two L-shaped modules around the beam axis.

Figure 4: Top: Sensor tile. Bottom: Module cross section.

given a guard ring of 500\textmu m. Any possible reduction in either the distance to the beam or the size of the guard ring would improve the impact parameter resolution performance of the detector as it was mentioned in the previous section. A R&D program is being pursued to achieve that goal.

The detector layout (Fig. 5) comprises two movable detector halves on either side of the beam line (dashed red line in Fig. 5) as in the current VELO. Each of these halves consists of 26 modules with varying spacing along the beam axis. The minimal pitch is 24mm. The two detector halves have a relative offset to each other along the beam line. Furthermore, modules from opposite halves are positioned such the sensitive areas are overlapping in some regions for tracks emerging from the interaction point. This results very helpful for the precise relative alignment of the detector halves. This layout has a active area of \( \sim 100\% \) with the exception of some small gaps. Using this layout for the VELO, the decay \( B_s \rightarrow D_s K \) has been simulated in the LHCb detector acceptance (15mrad-250mrad×300mrad). The efficiency for reconstructing the four tracks in the LHCb acceptance has

Figure 5: Detector layout of the upgraded VELO
been found to be 98.3%(99.7%) if 4(3) hits are required per track. For this detector layout, the shape of the foil that separates the secondary vacuum of the VELO from the primary LHC vacuum must be modified to accommodate the L-shaped structure.

2.3 The R&D on sensors

2.3.1 Reduction of the guard ring size

A smaller guard ring will allow the sensitive area to be closer to the beam improving the impact parameter resolution as already discussed in section 2.1. The gaps in the module design will be as well reduced due to reduction of the inactive sensor material. Testing structures and prototypes are being produced at the time of the writing at the IMB-CNM (Barcelona). The goal is to achieve dicing distance from the sensitive area of the sensor down to 250µm or less. The guard ring structures should withstand ~ 900V after irradiation.

2.3.2 Reduction of sensor thickness

In order to reduce material in the active area, both the ASICs and the sensors can be thinned after manufacturing. It has been proved at CNM that it is possible to thin silicon down to 150µm, or to 100µm for the sensors and afterwards bump bonded them on to ASICs. The thinning of the ASICs can be performed beyond that. The use of a supporting structure is a must, however bowing and handling difficulties may limit what can be achieved here.

2.3.3 Sensor options

Different options are actively being evaluated in test beams regarding their charge collection efficiency, leakage currents, charge sharing characteristics and position resolution at the required radiation dose.

**Planar Silicon** A significant amount of experimental data has been acquired in the last years showing a surprisingly large charge collection in severely irradiated silicon detectors. A 140µm planar silicon sensor can provide a signal of 6000e− for a MIP particle even after doses as high as $1 \times 10^{16} n_{eq} \text{cm}^{-2}$ [11]. After this levels of irradiation, thin detectors (140µm) also provide a higher signal charge than thicker (300µm) at the same bias voltage, because of the a charge multiplication effect at high electrical field [11]. Several sensor wafers n-in-p are being produced at different manufacturers, with different thicknesses (from 50µm to 300µm), different guard ring structures and dimensions (up to 3x1 ASIC tiles). These sensors will be used in the R&D above mentioned and in the prototyping of a first sensor tile with two Timepix ROC [6]. The charge sharing is expected to diminish with irradiation, this as well as the different design of guard ring structures will be investigated during the next irradiation and test beam campaigns.

**3D pixel detectors** 3D pixel sensors have demonstrated the largest charge collected for any given fluency of any known silicon technology. In contrast to the planar sensors, 3D detectors deplete laterally. The distance between the electrodes is very small resulting in high electric fields and high drift velocities with relative low bias voltages, reducing as well the trapping probability. The greater signal charge is due to the faster collection times that make the devices radiation hard. Moderate bias voltages, 150V, are necessary for 100% charge collection for fluencies up to $10^{15} n_{eq} \text{cm}^{-2}$ [12].
Higher radiation doses reduce the collection efficiency but for $10^{16} neq \text{cm}^{-2}$ the collected charge at 150V is 44% (10k electrons) and charge multiplication is also observed at higher bias voltage [13]. The reduced bias voltage also lower the risk for thermal runaway.

These devices are difficult to produce, nevertheless the University of Glasgow in collaboration with IMB-CNM Barcelona, have developed a new type of 3D detector known as double-sided 3D sensors [14], which eases the fabrication difficulties and has been demonstrated with high yield in several productions runs in the last years. Furthermore this new design improves the areas of inefficiency of the detector. An additional advantage of the 3D technology is that the self-shielding geometry and the additional processing can allow devices to be active within 10$\mu$m of the physical edge.

**p-CVD diamond**  For the radiation fluencies envisaged for LHCb upgrade, the Chemical Vapor Deposition (CVD) diamond has some appealing properties as a sensor element. It has been demonstrated to be extremely radiation tolerant. Besides, the leakage currents remain low after irradiation so there is not risk of thermal runaway. In addition its large thermal conductivity may be exploited for intelligent integrated cooling concepts so the diamond substrate acts as a sensor and at the same time as a thermal path. Given the modest total sensor surface of $\sim 1340 \text{ cm}^2$, the diamond sensor option could be financially affordable for the VELO upgrade. Some pCVD diamond samples suitable for bump bonding with Timepix ASICs were acquired. Collection lengths between 200 and 250$\mu$m were measured with a $^{90}$Sr source in the 5 samples bought. These measurements and other properties such efficiency and spatial resolution are under way in tests with beams.

**Strip option**  In case material budget or power of the pixel option goes beyond acceptable levels, a strip solution is also under development, with finer pitch, higher granularity and lower mass than the current VELO. New sensors have been designed and sent for production at Hamamatsu and are expected by the end of this year. For this option a 40MHz radiation hard ROC has to be designed. This ASIC development would be in common with any ASIC developed for a possible silicon strip readout for the upgraded Silicon Tracker.

### 2.4 Data readout rates

As pointed out in section 2.1, the particles rate will be very high close to the beam, averaging 200 MHz per ASIC. Assuming an average cluster size of 2 pixels and 32 data bits per pixel hit leads to the conclusion that from the innermost chip the data rate would be 13 Gbit/s. The 32 data bits are composed of 4 bit for ToT (‘time-over-threshold’) value, 12 bit for bxo (‘bunch crossing’) identification and 16 bit for the pixel identity. A 30% reduction can be obtained by clustering the data of simultaneous and nearby hits, since the bxo and some pixel address bits can be shared. An architecture has already been developed to minimize the data bandwidth from the pixel matrix [15]. This is based on grouping the digitized data from a $4 \times 4$ array (a super-pixel) and transmitting this information as a formatted and encoded cluster. In the ASIC the digital logic of this group of $4 \times 4$ pixels will be concentrated in a single area and the analog part of the pixels is put in either side of the digital area (Fig. 6). This brings many advantages, such as space saving (some digital blocks can be shared), a more efficient power and global signal routing, better isolation between analog and digital sections and the possibility of synthesize the column logic from a standard library rather
than custom cells. A disadvantage is the possible digital noise feedback into the bonding pads on top of the digital area.

![Figure 6: 4×4 superpixel](image)

**Figure 6:** 4×4 superpixel

**Figure 7:** The main requirements of the VELOpix chip

### 2.5 ASIC design: the VELOpix

Specific simulation studies have demonstrated that the pixel charge has to be digitized to 4 or more bits to obtain the required spatial resolution by charge interpolation. A time-over-threshold ADC conversion is adequate, since the maximal pixel hit rate is at most ∼6 kHz. In fact, a maximal ToT of 400 ns will cause a dead time < 0.5%.

The VELOpix will be developed as a follow-up to the Timepix3 [16], which is being designed in 130 nm CMOS process. The Timepix3 is itself a follow-up to the currently existing Timepix chip [6], and shares many design features with the current existing 130 nm CMOS Medipix3 chip. The ASIC contains an array of 256×256 square 55μm pixels. It is 3-side buttable and measures ∼14 mm on each side. It will provide simultaneous time-over-threshold and time stamping information, and will provide sparse, data-driven readout on high speed links. Timepix3 will be radiation hard to the levels required at the LHCb upgrade. The analog functionality of Timepix3 is identical to the requirements of the VELOpix front-end, and is listed in table 7. The modifications needed for the final VELOpix chip will be changes to the cluster formatting and buffering, to cope with the LHCb upgrade data rates, SEU protection, and the addition of multi-Gbit/s output links (4 × 3.2 Gbit/s) as was discussed in section 2.1. As already mentioned in section 2.7 the total ASIC power consumption should stay below 3W, to avoid causing thermal runaway in the sensor.

### 2.6 The RF foil R&D

The main function of the RF foil is to act as a *de facto* beam pipe, i.e. it must: separate the primary (accelerator) and secondary (detector) vacua, carry the image currents of the particle beams, allow for the close-in VELO sensor geometry with overlap, shield against RF noise pick-up in the VELO front-end electronics, withstand the heat load and the high radiation levels of the beam. This imposes very severe requirements in the design and construction of such RF foil. It should be vacuum tight (< 10⁻⁹ mbar l/s), extremely radiation hard, conductive and thermally stable. As discussed in section 2.2, it is necessary to redesign a new RF foil to house the new
L-shape modules. Furthermore, it has been decided to reduce even more its mass (a factor of two) [2], since it is the largest single contributor to the total material thickness of the VELO detector (∼50%), adding significant Coulomb scattering to the vertex and tracking errors [8]. A conceptual view of the upgraded RF foil is shown in Fig. 8.

The technologies under study include either developing a new material involving several plies of carbon fiber polymer (CFRP) reinforced with nano-particles and coated with aluminum or using a workable metal alloy similar to AlMg3. For the first technology a proof of principle has been performed with very encouraging results, however some issues need to be improved. In the second technology, a first prototype with a simpler geometry was produced by direct milling of the foil from a solid metal block with a 3-axis milling machine. The foil thickness was measured to be 300µm and the object has been tested to be leak tight to < 10⁻⁹ mbar l/s. Right now, a more complex RF foil with the L-shape is being produced using a 5-axis milling machine. It will be evaluated for vacuum tightness and thickness uniformity.

2.7 Cooling studies

As explained in section 2.1, because of the large particle fluence the silicon closest to the beam will draw large currents, \( I_{\text{leak}} \approx O(100) \mu A/cm² \) at 900 V of biasing. Thermal runaway can be avoided if the silicon is kept sufficiently cool, below about -10°C to -15°C. ANSYS [10] thermal simulations of a pixel module have been done (Fig. 9). The model includes radial \((r^{-1.9})\) and temperature dependent leakage current in the silicon. The 400µm-thick pCVD diamond substrate, described in section 2.2, extends to the edge of the sensor. The cooling channel is placed at 1 cm from the outermost silicon edge. The simulation assumes that a two-phase CO₂ cooling, as in the current VELO, capable of maintaining the temperature of the cooling at -35°C is feasible. Nominal thicknesses and thermal conductivities were used. The thermal performance of the module has been studied varying different parameters and the results suggest in order of importance: (1) The cooling channel must be as close as possible to the silicon; (2) The ROC power must be minimized; (3) Thermal impedances between the tiles and the substrate and between the cooling channel and the substrate should be minimize.

For a maximum ROC power of ∼3.0-3.5 W/chip CO₂ cooling should be sufficient to avoid...
thermal runaway up to a dose $\sim 10^{16}$ n$_{eq}$/cm$^2$, giving a safety factor of two above the expected upgrade fluence. The verification of the simulation with thermal mockups are envisaged. Identification of the adhesives and the process to glue the sensor tiles onto the pCVD substrate using a very thin layer of epoxy are being investigated. The process should be repeatable and the epoxy layers should be capable of stand shear from CTE differences. Since the detector modules are operated in the secondary vacuum ($10^{-6}$ mbar), leakage from the cooling system has to be avoided by using a pre-engineered leak tight system and attaching the cooling channels during module installation. The development of such a low mass cooling channel compatible with the substrate CTE and adequate thermal performance is a priority.

2.8 Test beam results

The performance of the Timepix chip for charged particle tracking has been verified with the construction of a dedicated Timepix telescope and an extensive characterization of a series of hybrid pixel devices, including 300 and 150 $\mu$m thick planar sensors and irradiated and non irradiated 3D sensors [17, 18]. A first version of this telescope is described in [5]. Several improvements have been applied to this telescope. Nowadays, this telescope has a track rate capability greater than 5kHz with a resolution at the Device Under Test (DUT) down to 1.5 $\mu$m. It has implemented a TDC for the trigger scintillator fingers that in conjunction with a Timepix plane in ToA gives a track time stamping $\sim$ 1ns. It is able to provide and record synchronized triggers to 40MHz readout systems used in the LHC. A standard 300 $\mu$m sensor bump bonded to a Timepix has been evaluated in the telescope and has shown a resolution of 4 $\mu$m, together with very high efficiency and accurate charge calibration [17]. Fig. 10 shows a direct comparison of the resolution as a function of the incident track angle for 300 and 150 $\mu$m thick sensor in identical conditions. The change in optimum angle as well as the small degradation in resolution is clearly demonstrated for the thin sensor. In Fig. 11 it is shown for the 300 $\mu$m thick sensor the resolution as a function of the incident track angle for two different biasing voltages: 10 V (close to depletion) and 100V (over-depleted). There is a clear gain at lower voltage, due to the extra diffusion and charge sharing corresponding to the lower electric field, however it should be noted that this result is for the slower
peaking time of the current Timepix electronics (∼100 ns) and needs to be investigated for faster peaking times.

3. Conclusions

The LHCb upgraded detector will be installed in 2018. The electronics of all the sub-detectors has to be adapted or replaced to be able to read out events at 40MHz. The VELO detector modules will be rebuilt taking as a baseline a pixel sensor. The conceptual module and detector layout are quite advanced. The R&D on sensor technology, readout ASIC design, cooling system and RF foil enclosure are progressing well.

References

[9] A. Affolder et al., Collected charge of planar silicon detectors after pion and proton irradiations up to $2.2 \times 10^{16} \text{n}_{eq}\text{cm}^{-2}$, Nucl. Instr. and Meth. A 623, 2010, 177-179.