Pixel chip architecture optimization based on a simplified statistical and analytical model

E. Conti\textsuperscript{a,b,1}, J. Christiansen,\textsuperscript{c} P. Placidi\textsuperscript{a,b} and S. Marconi\textsuperscript{a,b,c}

\textsuperscript{a}Dipartimento di Ingegneria Elettronica e dell’Informazione, University of Perugia,
Via G. Duranti 93, I-06125 Perugia, Italy
\textsuperscript{b}INFN – section of Perugia,
Via Pascoli, I-06123 Perugia, Italy
\textsuperscript{c}CERN,
1211 Geneva, Switzerland

E-mail: elia.conti@cern.ch

ABSTRACT: The technical challenges related to increased collision rates of the LHC will significantly affect detector electronics design. Efficient hit processing is achieved in pixel detectors by grouping pixels in regions, which share buffering logic. We present an approach to determine an optimized sharing strategy between pixels, depending on the shape of clustered hits in the detector. Simple statistical models of such shapes have been developed with respect to the position in the detector. The buffering performance of different pixel region configurations has been compared, showing significant improvement from architectures that do not feature pixel grouping.

KEYWORDS: Pixelated detectors and associated VLSI electronics; Analysis and statistical methods; Front-end electronics for detector readout

\textsuperscript{1}Corresponding author.
1 Introduction

New hybrid pixel detectors with improved resolution capable of dealing with hit rates up to 2 GHz/cm$^2$ will be required for future High Energy Physics experiments in the Large Hadron Collider (LHC) at CERN [1]. In this scenario complex on-chip digital signal processing and buffering at the absolute lowest possible power consumption will be designed, requiring special design techniques to implement a system working reliably.

Particle hits in the pixel detector are naturally clustered, i.e. signal is produced in more than one pixel unit cell per particle track. A cluster is composed of a given number of hit pixels and its shape depends on a quantity of factors, e.g. position in the detector where it is produced, pixel size, sensor thickness, magnetic field component, radiation damage, discriminator threshold [2]. We describe an approach to determine an optimized sharing strategy of digital logic in pixel chip arrays, depending on the shapes of clustered hits in a detector. In order to correctly process clustered hits each hit pixel should associate hit information to the same cluster without errors. This can be accomplished efficiently by grouping pixels in so-called pixel regions (PR) where buffering logic is shared between them. Such a strategy has already been adopted in current readout chip architectures such as FE-I4 for the ATLAS experiment [3], Timepix3 and Velopix [4, 5].

The main goal of the study presented in this work is to determine a range of the most convenient PR shapes and sizes to adopt in the design of next generation pixel chips without taking into account implementation related issues (e.g. routing, occupation of area, power consumption). Assumptions on the shapes of clustered hits in a detector have been made and two simplified analytical/statistical models of cluster shapes have been elaborated. The performance of different PR
configurations, both square and rectangular (figure 1), has been evaluated by calculating the num-
ber of buffer locations and the total number of memory bits. The used cluster models have also
been compared to real physics data.

The paper is organized as follows: section 2 describes the analytical models created for the
cluster shapes and the statistical assumptions on the number of hit pixels inside such models; sec-
tion 3 illustrates the results of the study for a set of PR configurations with each cluster model,
describing in detail the adopted technique for comparing buffering requirements; section 4 ex-
plains the comparison of the study with real physics data. Finally, section 5 draws the conclusions
and presents further developments.

2 Cluster models

With reference to the cylindrical coordinate space inside the detector, we have focused on how
cluster shape varies depending on the position in the pixel detector. In the center of the barrel a
cluster is usually composed of a central hit pixel with additional signals generated in the periphery.
In the edges of the barrel, on the other hand, the shape is elongated because of the track angle
from the collision point. The number of hit pixels inside such shape is approximately the same
as its length in number of pixels. Two compliant analytical models have been developed from
these assumptions:

- Clusters related to the center of the barrel have been abstracted with a fixed square envelope
  of size 3 × 3 pixels (to be referred from now on as symmetrical cluster model). Inside such
  an envelope it is possible to model a cluster composed of 1 to 9 hit pixels, one of which is
  assumed to be in the center (figure 2(a)).

- Clusters related to the edges of the barrel, have been modeled using rectangular shapes with
  \( \phi = 1 \) and \( z = n_p \), where \( n_p \in [1, 16] \) is the number of hit pixels in the cluster (elongated
  cluster model, shown in figure 2(b)).

Each cluster model has been modeled based also on statistical information on \( n_p \). Four typical
statistical distributions of \( n_p \) have been constructed for each model in order to represent different
Figure 3. Statistical distributions of \( n_p \) constructed for the symmetrical 3 \( \times \) 3 cluster model: (a) “Single”; (b) “Average 4.22”; (c) “Average 6.57”; (d) “Maximum”.

situations (e.g. distributions for the symmetrical model are shown in figure 3): i) clusters with \( n_p = 1 \), ii) clusters with the maximum value of \( n_p \), iii) two intermediate cases with an average value of \( n_p \) of 4.22 and 6.57 that fit well typical cases seen in pixel detectors.

A statistical cluster model with an envelope of 3 \( \times \) 4 has also been developed but it is not presented here, as it gives very similar results to the fully symmetrical 3 \( \times \) 3 model.

3 Results of the study

3.1 Average number of PRs occupied by a cluster

A quantity that is needed for calculating the data buffering during the trigger latency is the average number of pixel regions occupied by a cluster (\( \overline{PR} \)). It depends on the size and shape of the PR and it is related to the probability of the cluster to occupy a given number of PRs. This probability needs to be calculated: i) for each possible position of the cluster in the PR and ii) for each value of \( n_p \). A devoted procedure has been elaborated for each cluster model.

For the symmetrical cluster model, it is possible to identify in a PR different categories of pixel cells that produce the same results when the 3 \( \times \) 3 envelope is placed with its center on them (i.e. corner cells, edge cells, middle cells, as shown in figure 4): therefore it is just needed to compute the probability mentioned above for each cell category. Averaging the results with respect to the number of PRs and \( n_p \), using weights from the statistical distribution of \( n_p \), produces the cell type contribution to \( \overline{PR} \). The value of \( \overline{PR} \) for a given PR configuration is finally obtained by combining the number of cells of a given type in the region with their own contribution:

\[
\overline{PR} = \frac{\sum_j N_j \cdot \overline{PR}_j}{N_{PR}}
\]  

(3.1)
where $N_j$ is the number of cells of type $j$ in the PR, $PR_j$ is the contribution of cells of type $j$ to $PR$ and $N_{PR}$ is the total number of pixel cells in the PR.

For the elongated cluster model, on the other hand, the average number of PRs involved does not depend on the PR height because the $\phi$ component is unitary. The probability to occupy a given number of PRs has thus been calculated in all the possible positions of the cluster along the $z$ direction: the obtained results have shown a relation between $PR$, $n_p$, and the PR width $w$ which can be described by the following empirical formula:

$$PR = \sum_{n_p=1}^{16} \frac{w+n_p-1}{w} \cdot D(n_p)$$

(3.2)

where $D(n_p)$ is the chosen statistical distribution of $n_p$.

### 3.2 Algorithm for determining the required number of PR buffer locations

The performance of square and rectangular PR configurations can be evaluated using different quantities. First the number of buffer locations (needed to store data during trigger latency) was calculated: this quantity is defined as the required number of memory locations for keeping the overflow probability below a given value; for HEP applications an acceptable value of such probability is typically $10^{-2}$ at the absolute maximum hit rate.

In order to determine the overflow probability, the PR buffer has been modeled as an ideal array of finite memory locations where hit packets arrive at a rate that depends on the track rate in the detector ($R_T$, equal to cluster rate) and they are serviced by a trigger accept/reject after a known latency time. If the number of memory locations is $h$, an overflow is assumed to happen when the number of hit packets that arrive during the latency is greater than $h$. With the acceptable hypothesis of uncorrelated interactions taking place at each bunch crossing it is possible to calculate the probability of the arrival of more than $h$ hit packets during the trigger latency by applying the binomial distribution formula:

$$P_{\text{overflow}}(h) = 1 - \sum_{i=0}^{h} \binom{i}{L/BX} p^i (1-p)^{L/BX-i}$$

(3.3)

where $L$ is the latency time, $BX$ is the bunch crossing period and $p$ is the probability of arrival of a hit packet during a bunch crossing period. $p$ is determined by first calculating the pixel region rate.
Figure 5. Required number of PR buffer locations per pixel in order to keep the overflow probability below $10^{-2}$ for (a) square and (b) rectangular PRs using for both the “Average 4.22” statistical distribution of $n_p$.

$R_{PR}$, which is the product of the track rate in the detector with the average number of PRs occupied by a cluster:

$$R_{PR} = R_T \cdot PR.$$  \hfill (3.4)

Then $R_{PR}$ is multiplied by the pixel region area and the bunch crossing period:

$$p = R_{PR} \cdot A_p \cdot N_{PR} \cdot BX$$  \hfill (3.5)

where $A_p$ is the area of a single pixel cell and $N_{PR}$ is the number of pixel cells in the pixel region.

The overflow probability has thus been plotted with respect to $h$ in order to determine the required number of buffer locations for a PR ($h_{req}$). We have taken into account a case study related to next generation pixel chips with a track rate of 500 MHz/cm$^2$ in the detector, a trigger latency of 20 µs and a bunch crossing clock period of 25 ns [1]. Figure 5 shows the number of memory locations normalized per pixel (in order to compare the results between different configurations) for square and rectangular PR configurations and both symmetrical and elongated cluster models. It is possible to notice that the same results are obtained with both models for square PRs (figure 5(a)) and this is linked to the fact that the “Average 4.22” statistical distribution of $n_p$ has been taken into account for both of them; moreover, it is shown that less memory units per pixel are needed as the PR size increases. On the other hand, the adoption of rectangular PRs (figure 5(b)) does not bring any benefit to clusters in the center of the barrel. In the case of the end of the barrel the decrease of buffering locations for PR shapes elongated in the same direction of the $z$ of the cluster is lower (14.3%) than the increment (35.7%) for those elongated in the opposite direction.

### 3.3 Total number of memory bits

The most appropriate quantity to study is the total number of memory bits in the shared PR buffer ($b_{tot}$), as it holds information related to both the buffer locations and the memory organization. It is possible in fact to have different kinds of memory organization in a PR, depending on the format of the information that needs to be retrieved. For this study we have assumed a packet containing the bunch crossing ID (B-ID) (10 bits) and the hit amplitude (4 bits) in each pixel channel of the
Figure 6. Total number of PR memory bits per pixel in order to keep the overflow probability below $10^{-2}$ (solid line) and $10^{-3}$ (dashed line) for (a) square and (b) rectangular PRs using the “Average 4.22” statistical distribution of $n_p$.

region, with the amplitude expressed in terms of Time over Threshold (ToT). The product between the number of buffer locations and the number of bits of a single location gives the total number of memory bits in a PR. This quantity has been normalized per pixel in order to compare the results between different configurations.

The plots in figure 6(a) show the total number of required memory bits for the symmetrical and elongated cluster models taking into account square PRs. This quantity has an apparent optimum at $2 \times 2$ with a modest increase for $3 \times 3$ and $4 \times 4$ pixel regions that are the most suitable PR configurations. The total number of memory bits has also been plotted for the case of keeping the overflow probability below $10^{-3}$, requiring up to 30% more memory. The plots in figure 6(b) are related to rectangular PR and confirm the results that have been found in the previous subsection. We can thus deduce that the best solution is to adopt only square pixel region architectures. Choosing to implement one configuration for pixel chips in the central area of the barrel and another one for those in the edges is not feasible, as it makes the design significantly more complicated.

The final memory optimization will depend strongly on the type of memory used for a pixel chip implementation. If dedicated SRAM blocks are used for data buffering it is in general much more efficient, from a layout/area point of view, to have a few “large” memory blocks than many “small” memory blocks. This can make it more advantageous to have a relatively large pixel region (e.g. $4 \times 4$) instead of the indicated optimum of a $2 \times 2$ PR. The very hostile radiation environment for very high rate pixel detectors (e.g. $\sim 1$ Grad for ATLAS/CMS phase 2 pixels [1]) will dictate what kind of memory structure can be used.

4 Comparison of used cluster model to real physics data

In order to have an experimental confirmation of the assumptions underlying the elaborated cluster models, results have been carried out using also realistic statistical distributions of $n_p$ and they have been compared to those obtained with the typical ones.
Figure 7. Cluster footprint distributions related to (a) center and (b) edges of the barrel.

We have used cluster footprint distributions constructed with data coming from CMS LHC Run 200091 (figure 7(a) shows the distribution for the center of the barrel, while figure 7(b) shows that for the edges of the barrel). These distributions are related to all the detected clusters and therefore they include both those associated with tracks and the untracked ones, which are caused by loopers and other effects. Around 40% of the clusters are on tracks; nevertheless, we have chosen to use all the clusters rather than only the tracked ones, because the pixel detector should be able to work with all kinds of hits without problems, including the random ones which are potentially dangerous.

The value of the footprint distributions in point \((z, \phi)\) corresponds to the probability of the smallest rectangular envelope that encapsulates a clustered hit to have a size of \(z \times \phi\). Different strategies, described in the following subsections, have been adopted for each cluster model in order to carry out distributions of \(n_p\) from them.

4.1 Comparison for symmetrical cluster model

The symmetrical model describes all the possible cluster shapes inside a \(3 \times 3\) envelope. Such footprints are described by the distribution in figure 7(a) in the sub-domain \(z \in [1, 3]\) and \(\phi \in [1, 3]\). This distribution contains 84.2% of the total volume in the given sub-domain and this is therefore an acceptable simplification for the cluster model.

The complete statistical/analytical study has been repeated for the symmetrical model using realistic statistical distributions of \(n_p\) that have been carried out from the depicted distribution with the following relationship:

\[
P(n_p = k) = \sum_{z=1}^{3} \sum_{\phi=1}^{3} P_{z\phi}(n_p = k) \cdot f(z, \phi)
\]  

(4.1)

where \(P_{z\phi}(n_p = k)\) is the contribution of the single point \((z, \phi)\) to \(P(n_p = k)\) and \(f(z, \phi)\) is the value of the distribution in figure 7(a) in point \((z, \phi)\).

Based on two opposite hypotheses on \(P_{z\phi}(n_p = k)\), two different distributions have been obtained. The first one takes into account all the possible clusters that can be encapsulated in an
envelope of smallest size $z \times \phi$ with the same probability (e.g. if $(z, \phi) = (2, 2)$ the corresponding $P_{22}(n_p = k)$ is the ratio between the number of occurrences of clusters made of $n_p = k$, where $k = 2, 3$ or 4, and the number of occurrences of all the possible clusters); this distribution is plotted in figure 8(a). The second one, on the other hand, considers only the cluster with the maximum value of $n_p$ inside the $z \times \phi$ envelope (e.g. if $(z, \phi) = (2, 2)$ we have that $P_{22}(n_p = k)$ is zero for $k = 2, 3$ and is 1 for $k = 4$); it is shown in figure 8(b).

The results of the study for both the realistic distributions, plotted in figure 9(a), are consistent with the ones carried out using the typical “Average 4.22” distribution. The maximum error between the plots related to the two different hypotheses is 8.3% (not taking into account the non-relevant $1 \times 1$ case). All the other assumptions that can be made on cluster probability produce results that are contained between those two plots.

### 4.2 Comparison for elongated cluster model

Because $n_p$ at the edges of the barrel is around the same as its $z$ component, in order to keep $\phi = 1$ for the elongated model the distribution in figure 7(b) has been intersected with planes parallel to the $z$ axis. Since the distribution contains 84% of the total volume in the sub-domain $\phi \in [1, 3]$ and $z \in [1, 16]$, the intersections of the distribution with planes $\phi = 1, \phi = 2$ and $\phi = 3$ have been taken into account. The realistic statistical distribution of $n_p$ has then been subsequently carried out by calculating the average of the three intersections, weighted by the area subtended by each of them (figure 8(c)).
The complete statistical/analytical study has been repeated for the model using this distribution and the results are compared with the initial ones (typical “Average 4.22” distribution) in figure 9(b). Consistency between plots validates the adopted cluster model.

5 Conclusion

A study of shared buffering performance has been conducted for a set of square and rectangular pixel regions. The study is based on two analytical and statistical cluster shape models related to the different positions in a detector; the goal is to get a perception of the most suitable configurations that can be taken into account for the design of next generation pixel chips which will operate at very high hit rates in the HL-LHC experiments upgrades.

The pixel region performance has been evaluated in terms of memory bits in regional shared buffers and the results show that square regions with size from $2 \times 2$ to $4 \times 4$ pixels seem the most suitable ones. The most appropriate configuration will be chosen in the indicated range and this will strongly depend on the physical chip implementation.

The statistics underlying the study have been validated using a set of real physics data and the results show consistency with the initial assumptions. A suitable simulation and verification framework is currently being developed using the hardware verification and description language SystemVerilog, with the goal of modeling the candidate pixel chip architectures with more detail. It is also planned to use the elaborated cluster models for the generation of input hits. This verification environment will then be used for further architecture optimizations and will finally be used for extensive design verification.

Acknowledgments

The authors would like to thank Professor Morris Swartz from John Hopkins University for determining and providing the cluster footprint distributions from CMS Run 200091 data.
References

[1] M. Garcia-Sciveres and J. Christiansen, *RD Collaboration Proposal: Development of pixel readout integrated circuits for extreme rate and radiation*, CERN-LHCC-2013-008 (2013).

[2] M. Swartz, *A detailed simulation of the CMS pixel sensor*, CMS Note 2002/027, July 2002. Available online: http://cds.cern.ch/record/687440/files/note02_027.pdf.

[3] D. Arutinov et al., *Digital architecture and interface of the new ATLAS pixel front-end IC for upgraded LHC luminosity*, IEEE Trans. Nucl. Sci., 56 (2008) 388.

[4] V. Gromov et al., *Development and applications of the Timepix3 readout chip*, PoS(Vertex 2011)046.

[5] T. Poikela et al., *Architectural modeling of pixel readout chips Velopix and Timepix3*, 2012 JINST 7 C01093.