Fast on-detector integrated signal processing status and perspectives

V. Lindenstruth*, L. Musa

CERN, CH-1211 Geneva 23, Switzerland

For the ALICE Collaboration

Abstract

The large and increasing channel count of modern detectors requires the use of microelectronics. The data rate and signal integrity requirements drive complex electronics to be mounted close to or directly on the detectors, possibly even integrating the complete first-level trigger stage. The latest silicon road maps indicate that the integration density of microelectronics will continue to increase during the next decade. However, there are several constraints to be taken into account that cause ramifications with respect to on-detector electronics. For instance, the core voltage will be reduced to below 500 mV, the clock rates will exceed GHz, and the power density will increase further. This article outlines two examples of trigger and readout systems, the ALICE TPC and TRD, which are completely integrated in microchips. The article expands on the expected impact future silicon processes may have on the on-detector integrated signal processing.

PACS: 07.05.Hd.

Keywords: Signal processing

1. Introduction

Modern particle physics detectors implement of the order of $10^5$–$10^6$ electronic channels, which drives the need for ever higher integration densities of the on-detector electronics. This is particularly true for trigger systems, where the readout infrastructure is to be complemented with a fast processing system, which will determine the trigger decision in real time.

2. The ALICE TPC detector electronics

The ALICE TPC [1] is a large-scale detector that implements 570,000 electronic channels. It is operated with a maximum trigger rate of 200 Hz in heavy ion running and 1 kHz in proton–proton running.

2.1. TPC electronics chain overview

The overall architecture of the TPC readout system is sketched in Fig. 1. The charge-sensitive preamplifier’s noise is determined by its input capacity, therefore requiring its proximity to the TPCs pad planes. The preamplifier also implements first-level shaping and tail cancellation.
functionality. Here, the greatest design effort is directed towards noise reduction. The differential PASA outputs are digitized with 10 bits at 10 MSPS.

In the next step the signal is further processed using now the advantages of the digitized information. The first stage implements the baseline correction. Its main task is to prepare the signal for the tail cancellation by removing low frequency perturbations and systematic effects. The next processing block is a 4-exponential tail cancellation filter. The filter is able to suppress the tail of the pulses within 1 μs after the peak, with the accuracy of 1 LSB. Since the filter coefficients for each channel are fully programmable and reconfigurable, the circuit is able to cancel a wide range of signal tail shapes. This also allows maintaining a constant quality of the output signal regardless of ageing effects on the detector and/or channel-to-channel fluctuations. The subsequent processing block applies a baseline correction scheme based on a moving average filter. This scheme removes non-systematic perturbations of the baseline that are superimposed on the signal. At the output of this block, the signal baseline is constant with an accuracy of 1 LSB. The resulting data is zero suppressed and stored in a multi-event buffer.

The entirety of digitization, filtering and data storage is performed during the 88 μs drift time of the detector. Upon receipt of a L2 trigger accept, the readout is performed out of the on-detector event buffers. This architecture permits utilization of the optical readout links of more than 90%, without creating unnecessary dead time. The entire TPC electronics, as sketched in Fig. 1, is integrated into two chips: the preamplifier (PASA), and the digitization and readout chip (ALTRO) [2, 3]. They are mounted as separate 2 × 8 chips on the TPC front-end cards.

2.2. TPC electronics performance

One major concern of the simultaneous ADC and digital processing integration is the potential noise imposed by the digital switching on the same die. Fig. 2 shows a measurement of the ADC performance on the ALTRO chip in comparison with several other stand-alone 10-Bit ADCs [2, 3].

The ALTRO chip integrates a commercially available ADC, which allows a direct comparison of its integrated and stand-alone performance. The resolution of the integrated ADC exceeds the stand-alone operation. However, in order to achieve this result, care had to be taken in designing the various power paths on the chip. Further, the ADC and digital clocks were adjusted with respect to their mutual phases. There are several possible reasons for the improved behavior of the ADC in situ. For instance, the single-ended

---

1 TS 1001.
outputs do not have to drive large capacities, thus feeding random currents into the ADCs substrate, when operated in situ.

3. The ALICE TRD detector electronics

The ALICE TRD [4] implements 1.2 million analog channels, which are digitized during the $2\mu s$ drift time. Unlike in case of the TPC, the TRD also implements an on-line trigger, which is capable of tracking all of the up to 16,000 charged particles within the six detector layers. This trigger has a very tight time budget of $6\mu s$ for all digitization and processing [5,6].

3.1. TRD electronics chain overview

The overall TRD electronics chain is sketched in Fig. 3. There are many similarities to the TPC readout chain. Both detectors implement a separate analog preamplifier/shaper chip (PASA). In this particular case, both chips are relatively similar.

Like the TPC, the TRD implements a multi-channel ADC with all the required digital back-end processing on a second chip. The remainder of the TRD electronics chain implements a short 64-word single event buffer plus a tracklet processor, which identifies potential high-$p_T$ track candidates for further processing.

In the case of the TPC, the readout is performed out of the multi-event buffers and therefore outside the dead time, thereby allowing the implementation of a bus. In the case of the TRD, the readout is performed in two stages, first during the trigger processing, where all tracklet candidates are shipped within 600 ns from the 65664 TRAP chips, to the global tracking unit (GTU) for merging of the six detector layers. Later the event buffer is readout in case of an accept of the event.

The larger number of chips and the high aggregate bandwidth requirements of 270 GB/s result in a multi-stage 4:1 readout tree. The appropriate merging of the ingress data streams is performed by the tracklet merger (TM in Fig. 3), which is also integrated on the TRAP chip for efficiency reasons. It occupies 0.3 mm$^2$ of silicon.
real estate. The high-speed TRD readout is performed with 1080 2.5 GBit optical links. The appropriate serializer chips are designed in full custom for on-detector operation and are attached to the TRAP chips. They form the interface between the on-detector electronics and the off-detector GTU.

The GTU inspects the tracklets for global high-$p_T$ tracks and issues an appropriate trigger signal. In case of an accept decision, it also receives the entire RAW data from the on-detector event buffers and stores the event in a multi-event buffer for later readout in case of a L2 accept.

3.2. Trap architecture

The architecture of the TRD tracklet processor (TRAP) chip is sketched in Fig. 4.

Since the TRD is operating under very tight latency and power constraints, the commercially available ADC, used for the TPC, was not very applicable here. A custom ADC was designed from scratch\(^2\) in the UMC 0.18 $\mu$m process. This 10-Bit, 10 MSPS ADC requires 6 mW of power, has a conversion latency of 1.5 clocks and uses $0.1 \mu$m\(^2\) of silicon real estate.

The next stage of the signal processing chain implements digital filters, compensating for the detector’s gain variations and cross-talk. It also implements an appropriate two-exponential $1/t$ tail cancellation filter. It should be noted that under normal stand-by operation only the ADCs and digital filters, running at 10 MHz, are operational. During acquisition, after a pre-trigger, only the event buffers and preprocessor functionality are enabled, which are also running at 10 MHz. After completion of the digitization during the drift time, the MIMD processor is started, executing at 120 MHz. Here, up to four tracklets are processed in parallel by the four RISC processors. The CPUs implement an especially optimized arithmetic unit. Benchmarks computing integer sums, such as $\sum i^2$, when compared with a 1.6 GHz Athlon processor, perform faster by a factor of 2.5 per MHz clock rate.

In order to assist the multiprocessor synchronization, the four RISC processors implement a global register file and a global four-port memory, which was designed in full custom.

In order to simplify the overall architecture, the required readout functionality is also integrated on the TRAP chip, because it requires little silicon real estate. In cases where it is not needed in full, the appropriate parts are disabled. This is particularly true for the LVDS outputs, which require 7 mW per channel. The LVDS cells were designed such that a receiver receives a defined zero, even when the corresponding transmitter is powered off. This design feature allows the power cycling of a transmitter without the receiver detecting any glitches. The LVDS receiver requires minimal stand-by power.

3.3. Reliability, radiation tolerance

In general, there is a fundamental design choice to be made beforehand. In the case of very high radiation levels, such as those encountered in the area of virtex detectors, the electronics have to be radiation hard. There are special design techniques, for instance, using only enclosed transistor structures and appropriate standard cells [7]. However, although they are proven to work, those design techniques typically require a factor four more silicon real estate with the corresponding additional cost and power consumption.

In the case of the ALICE TPC and TRD, the expected radiation levels are rather moderate, corresponding to $10^8$ protons/a cm\(^2\). In such a radiation environment, latch-up effects have to be avoided by design, which is a standard technique and single event upsets have to be handled properly. In the case of TPC and TRD, appropriate care was taken, such as implementing error correction on all memories, including the state registers of all major state machines. The event buffers are only protected by parity, as error detection is sufficient here.

3.4. TRD detector— electronics integration

Due to the layered structure of the TRD, all electronics are located within the active detector area and therefore contribute to the material budget.

\(^2\)University Kaiserslautern.
Fig. 5 shows a picture of a TRD MCM, implementing both chips PASA and TRAP on one carrier. The PASA analog output is directly bonded to the TRAP ADC inputs. The chosen technology (chip-on-board) is very low cost. The PCB is designed as ball grid array and directly soldered to the detectors readout PCB, thereby avoiding any connector cost. The only production cost items are the two chips, the PCB, and the bonding of the chips.

3.5. TRD electronics performance

The testing and qualification of the TRD electronics is on-going. The complete digital circuitry is tested successfully. The ADC overall performance lies at 8.9 ENOBs, which is less than expected due to two understood issues with the design, which were corrected in the last MPW run.

The most critical aspect here is the overall noise imposed by the two chips operating in close proximity. Fig. 6 shows a preliminary plot of a baseline noise spectrum and a PASA step function response digitized by the TRAP chip.

4. Perspectives

The overall performance of silicon processing is improving constantly, giving rise to the question of what this may mean for next generation on-detector readout and processing electronics.

4.1. IRTS road maps

The target features and extrapolations of silicon processing capabilities are monitored and published by the International Technology Roadmap
for Semiconductors (ITRS) consortium [8]. For example, the extrapolations of this group predict that for the year 2016, it will be possible to produce chips with gate lengths between 9 and 11 nm, operating at clock rates exceeding 25 GHz. More than 3 billion transistors are projected on one high-performance chip, consuming more than 280 W.

Fig. 7 shows one example of the extrapolations of the effective oxide thickness (EOT) as a function of time. Note that in 2016, transistors with effective oxide thicknesses of less than 0.5 nm are projected.

A direct consequence of this development is the reduction of core operating voltage, in order to avoid a break-down of the gate oxide. This development is already visible today, where quarter micron technology operates at 2.5 V, while 180 nm technology is limited to core supply voltages of 1.8 V. In order to allow interfacing with other devices, modern processes implement a second kind of high-voltage transistor type with a thick gate oxide layer in order to operate in a more high-voltage environment (here 3.3 V). The core supply voltage in 2016 is projected to be between 0.4 and 0.6 V. Note the corresponding supply currents for a projected 280 W chip.

Also note the increase in gate leakage current, increasing by two orders of magnitude in the near future as a function of decreasing feature size.

Another consequence of reduced gate area is associated with the statistics of the limited number of dopant atoms in the gate area, which becomes relevant for gate lengths below 100 nm [9]. Therefore, the transistors’ threshold voltage \( V_T \) starts to fluctuate from transistor to transistor.

However, the thin gate oxides have the advantage of an inherent radiation tolerance. On one hand, the geometric cross-section is decreasing and on the other hand, any ionization trapped within the gate oxide, which would cause a \( V_T \) shift, is more likely to drift out of the active area, particularly if a supply voltage is applied, for thinner gate oxide thicknesses. The additional cost for higher integration of radiation hard layout techniques may be offset by the smaller feature size.

Another factor to be considered is the cost of production. Even the largest architectures discussed still have a significant cost contribution by the non-recurring costs, such as the mask set for production. Unfortunately the mask cost increases exponentially as the feature size decreases and is approaching 1 million dollars per mask set, making even multi-project wafer runs prohibitively expensive.

4.2. Applicability to detector readout and processing

The increasing number of transistors, possibly being integrated on one die, is not the driving factor in detector electronics. Even the presented
designs are close to being PAD limited, where a smaller structure size would not reduce the chip area further.

The corresponding decrease in supply voltage reduces the available signal-to-noise. Even in the TRD design, the first ADC stage is implemented using the high-voltage 3.3 V transistors. The increasing gate leakage is an added complication to analog designs. Further statistical \( V_T \) fluctuations make analog transistor matching impossible.

Reduced feature sizes, which grant higher clock rates, allow the reduction of silicon real estate per function by operating multiple operations sequentially at a high rate. However, this technique would have been applicable already for the 21 digital filters on the TRAP chip and was not used in order to avoid high clock rates. In contrast, all digital filters are specifically operated at the ADC digitization rate with a clock phase being adjusted to match the ADCs least sensitive time slot towards the end of its conversion cycle.

5. Conclusions

Microelectronics have become an integral part of modern particle detectors, particularly TRDs. However, not all features of modern Si technology scaling will be applicable to this field. Feature sizes below 100 nm do not seem to be very useful here. The electronics in such processes can be designed with appropriate but area consuming techniques to be radiation hard, supporting radiation doses of kGy, or can be designed with commercial grade technology to be radiation tolerant, using fault tolerance techniques. The power distribution becomes increasingly important. Low-noise radiation hard step-down regulators, which are capable of operating in strong magnetic fields, are required to avoid the distribution of large currents. The increasing non-recurring engineering costs drive more generic multi-purpose chip developments.

References