Development of a New First Level Trigger for the Surface Array in the Pierre Auger Observatory based on the Cyclone™ Altera® FPGA

Z. Szadkowski*, Karl-Heinz Becker*, and Karl-Heinz Kampert*

*Bergische Universität Wuppertal, Fachbereich C - Physik, D-42097 Wuppertal, Germany

The development of a new first level trigger for the surface array in the Pierre Auger Observatory is described. It utilizes the new Altera® family Cyclone™ which is available since March 2003. Different from previous and the currently used design based on the APEX™ and ACEX® families, respectively, it offers much higher capacity of logic registers and memories, a simplified board design, lower power consumption, lower price, and lower cost. The paper describes the board design, internal structures of the programming routines, various trigger implementations and test results obtained in the laboratory.

1. Introduction

The Pierre Auger Observatory [1] surface array will contain 1600 surface detector stations distributed over 3000 km² on each Southern Site (under construction in Argentina) and Northern Site (foreseen to start in 2006 in USA). The large area covered, the limited power budget due to the solar power supply and the variation of the environmental temperature impose special constraints on the surface detector trigger system. The preliminary phase of the Pierre Auger Observatory was the Engineering Array containing 40 water Cherenkov detectors. It has been built to verify and improve several variants of the hardware and software designs.

The Cherenkov light of each Pierre Auger water tank is read out by three 9-inch photo-multiplier tubes (PMTs). The signals from the anodes (low-gain channel) and last dynodes (high-gain channel) are transmitted via equal-length shielded cables to a Front-End Board, attached to a Station Controller. In the Front-End Board, the 6 signals of a water tank are filtered by anti-aliasing 5-pole Bessel filters with a cut-off frequency of 20 MHz. Sampling is continuously done at 40 MHz by 10-bit AD9203 Flash ADCs [2]. With 5 bits of overlap between the two ADCs of a PMT, a dynamic range of 15 bits is achieved. The outputs of the six ADCs are processed by Programmable Logic Devices (PLD) Altera® [3] working as trigger/memory circuitry (TMC). They are supported by additional Dual-Port Random Access Memory (RAM) serving as a temporary buffer. The TMC evaluates the ADC outputs for interesting trigger patterns, stores the data in buffer memory, and informs the detector station micro-controller in case a trigger occurs. The station controller sends trigger packets and, when requested, event data to the observatory campus via a wireless network. A hierarchical event trigger is used to select events of interest and reject uninteresting events in order not to exceed the rate constraints imposed by the station micro-controller, the communications link bandwidth, and the central data acquisition system.

The advantages of programmable logic devices as compared to the classical Middle Scale of Integration (MSI) or Application Specified Integrated Circuit (ASIC) approach is the possibility to implement complex algorithms and keeping the flexibility of simple modifications if new features are needed due to experiences gained from collected data.

2. 1st generation - Engineering Array

The TMC, originally planned as an ASIC, contained all functions in a single chip, together with the memories for the fast (for
shower profiles investigations) and slow (for investigation of the muon component and calibration) channels [4]. The ASIC algorithm was tested in PLD chips EPF10k100ABC600-1 and in EPF10k200SBC600-1 from the Altera® FLEX® 10kA family to verify its functionality as well as the timing. The usage of PLDs allowed implementing sophisticated algorithms and avoiding possible risks, if some bugs in ASIC code were not found.

An alternative design, based on the PLD APEX™ chip, was developed in parallel to the ASIC and was treated in the beginning as a backup. The functionality of the APEX design (type of triggers generated, etc.), was as close as possible to the ASIC one, however, the implementation algorithm has been completely changed [5] and was based on the synchronous and pipelined approach applied in the second level trigger of the Fluorescence Detector in the Auger Observatory [6].

All tanks of the Engineering Array (a test segment consisting of 40 surface detectors) were equipped with Altera® APEX™ EP20k200RI240-2 chips [7,8]. The high capacity of resources and fast register performance made the APEX™ 20K and APEX™ 20KE chips a good solution for the surface detector trigger. However, because of its high price alternative low cost-effective designs were investigated in parallel.

3. 2nd generation - pre-production phase

The APEX design has been working perfectly in the Engineering Array, but the total cost of the TMC exceeded the assumed budget. The cost-effective Altera® ACEX® family offered inexpensive chips with sufficient resources of logic elements, but unfortunately insufficient internal memory. Two ACEX® chips were sufficient to perform all APEX™ tasks. Therefore, the signal path had to be split into two parallel paths performing synchronously [9]. However, the synchronous recording of data in both chips introduced additional complications; triggers were generated in the master ACEX_A chip and next were transferred to the slave ACEX_B. Such a dual chips design requires additional mechanism of synchronization. Some parts of the code had to be duplicated into the adjacent chip. Anyway, the ACEX® family allowed reducing the total price significantly without dismissing any functionality. The first level trigger based on ACEX® chips has been installed by now in more than 550 detectors [11].

4. 3rd generation - Cyclone™ design

In May 2003 the new Altera® FPGA family Cyclone™ appeared on the market. The new chip allows a significant simplification of the construction of the surface detector trigger as well as decreasing the total costs and improving parameters. Different from the currently used ACEX® chips, Cyclone™ chips contain much more internal memory avoiding the need to support the TMC by external memory. The TMC utilizes two types of channels: a fast one to investigate the shower profile and a slow one for calibration data. The memory in the fast channel requires 2 buffers · 768 words · 64 bits = 24 M4K RAM Blocks. The remaining 28 M4K RAM blocks (in the selected Cyclone™ chip - EP1C12Q240I7) can be used for the limited size memory in the slow channel (3584 words FIFO is estimated as sufficient). Nevertheless, we decided to keep the external larger memory rather than occupying internal memory blocks in order to allow implementing new trigger conditions.

The power consumption is at the same time reduced as the core is supplied by 1.5 Volts only. Furthermore, compared to the previous two-chip ACEX design, a single chip avoids problems with chip synchronization [12]. The registered performance indicated by the compiler is at the level of 130 MHz. Such a high internal speed would allow increasing the sampling frequency, thereby improving the time resolution of the triggers. Moreover, additional resources also allow implementing new kinds of triggers. The MegaCore® library offers DSP routines such as Fast Fourier Transform, which may be useful to recognize specific types of events.

Very recently, the family Cyclone II has appeared on the market.
4.1. Firmware Implementation

The structure of internal routines is fully pipelined. Some complex routines (mainly widebus adders) for reliable and high-speed performance requiring two clock cycles in ACEX® chips are implemented in the Cyclone® in a single clock cycle due to a newer Cyclone® architecture assuring much high register performance. Also, the address decoder has been redesigned to use the continuous address map, thereby freeing more than 4% of resources.

The Cyclone® chips offer sufficient resources and internal memory to implement the full algorithm into a single chip. However, such an approach would impose limits on future optimizations or on implementations of new ideas. As a compromise between simplicity and flexibility for future modifications we chose a two chip design; the Cyclone® combined with the DPRAM supporting slow memory as used in the previous designs. Remaining free logic elements (> 60%) and memory (> 50%) may be used for new trigger implementations, e.g. for trigger improvements in the observatory.

When developing the algorithms for signal processing in the PLD and DMA transfer, special care has been taken to provide sufficient safety margins for operation in environments of large temperature variations, such as in the Pampa Amarilla. This has been accomplished by designing complex functions such as comparators, coincidence logics, adders, multiplexers, shift registers, counters, internal memories (implemented as dual-port RAM) or DMA glue logic in pipeline stages with a precise control of the data flow. The internal embedded array blocks could be programmed as a simple FIFO. However, a dual-port implementation is chosen to assure higher registered performance.

For the chosen design it is found that the performance of a single global reset node connecting simultaneously both to the asynchronous clear inputs of the D-Flip-Flops (DFFs) and to the synchronous reset in the routines from the Library...
of Parameterized Functions (LPM) is poor. The compiler indicates significant limitations in a distribution of such a structure. To improve the registered performance, the global reset for the system is extended to the additional redundant DFFs working parallel and driving blocks, encapsulating routines performing all processes in a particular clock cycle in the pipeline chain. The \( \sim \text{Rst} \) node drives globally all asynchronous DFF inputs. However, the redundant registers \( \sim \text{SRst}_A, \sim \text{SRst}_B, \) etc. drive synchronous reset inputs (sclr) in the routines relating to the successive clock cycles to reduce the fan-out for non-global nodes and to reduce the propagation time. The \( \sim \text{SRst} \) registers are located close to the final destined routines and significantly improve the registered performance of the entire system.

Classical microprocessor based tri-state gates connected to the common internal data bus may produce glitches due to bus overloading. Thus, they were replaced by pipelined multiplexers. The interlaced (APEX\textsuperscript{TM}) [5] or non-interlaced (ACEX\textsuperscript{R⃝}) [9] mode for DMA transfer required 5 or 4 empty reads to prepare internal data for the transfer through the pipelined multiplexers. In the Cyclone design the non-interlaced mode has been selected in order to avoid significant changes in the Station Controller software.

4.2. Triggers

Two channels process continuously monitored data: the fast one records the shower profile and the slow one is used for the self-calibration by means of through-going muons through the water tanks (see Ref. [1,5]).

For the fast channel the following triggers are implemented: 1) A single bin trigger is generated, when the input signal is above a certain threshold. To investigate the signal profile, 256 words are recorded before the trigger and 512 words after the trigger. The input signal for the trigger generation can be taken on various levels of coincidence for separate ADC signals, as well as for the sum of signals from all three data busses. 2) A time over threshold (ToT) trigger requires 12 FADC bins with signals larger than a preset threshold in a sliding window of 3 \( \mu \)s (120 time bins). All data from the bin-comparator are put into a 256-bin shift register. The number of fired bins moving through the shift register is compared in the occupancy comparator with a fixed value written into a register. If the number of fired bins is above the occupancy threshold, the ToT trigger is generated. The time interval (the length of the shift register) tracking the fired bins can be changed by selecting one from the 256 possible inputs via a MUX.

For the slow channel only one type of trigger is implemented, however in two instances: a single bin trigger is generated, when the input signal is above a given threshold. As in the fast channel, the possibility of different coincidence levels and the sum has been implemented. When a trigger appears, data from 3 time bins before trigger and of 20 time bins after trigger are stored. The threshold is then lowered in order to allow investigating the tails of showers. More data are written if they are above the new threshold. To synchronize data from the slow and fast channels, time stamps are written together with slow channel data. They are inserted into the data stream before valid data. In cases where no place is left to insert additional stamps, they are neglected. This is done in order to avoid losing real data, for example when the “low” threshold expires and data are still above the “high” threshold. However, in such cases the change of the threshold change is recorded.

4.3. PLL adjustment and DMA transfer optimisation

ACEX\textsuperscript{R⃝} industrial version chips used in the 2nd generation design support neither a PLL 40 MHz clock with a multiplication factor \( \times 2 \) nor any precise shift tuning. In the Acex design, a PLL was not used at all. Cyclone\textsuperscript{TM} offer phase locked loops and a global clock network for clock management solutions. Cyclone\textsuperscript{TM} PLLs offer clock multiplication and division, phase shifting, programmable duty cycles and external clock outputs, allowing system-level clock management and skew control. The Altera\textsuperscript{R⃝} Quartus\textsuperscript{R⃝} II software enables Cyclone\textsuperscript{TM} PLLs and their features without using any external devices [13]. That feature was used for an optimization of the DMA transfer.
Both fast and slow channel are double buffered to reduce dead time, which occurs when both buffers are full and data from one of the buffers is not yet completely transferred to the Station Controller. The dead time depends on the data transfer speed (DMA mode) from the PLD to the Station Controller.

Zero wait-state (WS) modes provide maximum speed. Laboratory temperature tests confirmed the reliability of that mode. However, to extend the safety margin for real temperature extremes inherent in an uncontrolled environment, a 1 WS mode has been chosen for DMA transfer for both channels. However, the additional wait state without a possibility of the internal clock shift can be used to extend either the setup or hold time of data transferred in DMA. Cyclone™ provides a precious internal clock tuning with a 10 ps grid. In the chosen design 6020 ps shift of PLL clock in comparison to the external global clock assures optimal stable data position related to the falling edge of ~CAS3 edge responsible for the DMA data transfer (see Fig. 2).

4.4. Power consumption

The surface detectors are powered by solar panels, so the power budget is one of the most critical parameter in the system. The total power consumption of the trigger should not be greater than 1 W. The Cyclone™ chips consume much less power as compared to the APEX™ and ACEX® families due to a new architecture and technology. The power saving allows replacing the formerly used switching voltage regulator by the linear one on the PCB, thereby reducing the digital noise on the board.

4.5. Temperature tests and long-term pattern comparison tests

Preliminary tests at a frequency of 62.5 MHz (for the design system clock of 40 MHz) performed in the full temperature range (−20°C...+70°C) confirmed that the internal structure of the PLD chip works perfectly even under such extremely conditions [10].

The interval between the rising edge of the clock and the falling edge of the ~CAS3 signal line (output hold time) may vary (for 40 MHz)
from 3 to 9 ns according to the Power PC specification. Due to the 4 ns grid values of the signal generator, only 24 instead of 25 ns cycles could be realized. The data transfer has been tested for 4 and 8 ns output hold time. The design has also been tested for much higher frequency of 62.5 MHz (i.e. 16 ns time bins) to verify the safety margins. For such high frequencies the output hold time can be extrapolated to be in the range of 2 to 6 ns. Because of the signal generator grid, the only value of the output hold time being available was 4 ns. When the the Fast/Slow buffer is filled, the PLD generates ~EVTCLKF/S interrupts that request the DMA transfer. The writing of data into memory and the reading via DMA are independent processes.

For the slow channel, due to small capacity of the employed logic analyzer, the DMA starts while the external memory is still being filled. The external memory is a static dual-port RAM and data are written into the left port and are read simultaneously and independently from the right port. The starting point of the DMA data transfer has been adjusted in such a way to not read from addresses, to which no data have been written yet. For 1 WS, reading is much slower (2 clock cycles per word) than writing, so the starting DMA point has been chosen close to the trigger. For 0 WS, the speed of reading and writing is the same. However, writing is a not a continuous (as DMA) process. For chosen parameters of the internal pattern generator (thresholds, muon gate length, noise-like patterns) data are being written into the external memory with gaps, when the current values of data generated by the internal pattern generator are below thresholds. As a consequence, the DMA reading process is generally faster than writing. To avoid the case of reading data from addresses that have not been written yet, the starting DMA point has been adjusted

Figure 3. Measured current consumption in the digital sections of the ACEX and Cyclone boards as a function of trigger rate. The ACEX® chips were supplied by a 2.5 V switching regulator and the Cyclone™ by a 1.5 V linear one.

Figure 4. Measured error probabilities for input setup (−22.5, −21.25, and −20.0 ns) and input hold (+25.0 ns) times as a function of temperature for the fast channel (1 WS mode). Note, the required input setup and hold times are 8 ns and 10 ns, respectively. For the range −18.75 ns to +23.75 ns, the measured error probability is below 10⁻¹². Data were sampled at a frequency of 800 MHz. The temperature dependence of the slow channel is similar.
based on simulations.

The internal noise-like signal generator has generated data which next were compared with patterns obtained from simulations. Details can be found in [5]. Patterns for both channels were prepared from simulated data and were uploaded as binary files to the Station Controller. The tests compared data with patterns in 24-hour cycles of varying temperature in the full range from -20°C up to 70°C for 5 days. All data have been transferred to the Station Controller via DMA with 0 WS. Data received from PLD were compared in the Power PC word by word. The test procedure allowed finding any bit error in the processed data. For the fast channel all words were compared, for the slow channel the time stamps were omitted. This was done because simulated stamps do not agree with the stamps in a real time system. Therefore, time stamps could not be tested with such kind of set-up. The Front End Board with the Station Controller worked in the climate chamber autonomously. Data were collected in the memory of the station controller and were downloaded to the external computer via a RS232 interface with handshaking connection.

Results are extremely satisfactory. All data (47 GB) generated in the internal pattern generator, processed through the trigger, temporarily stored in the internal memory and finally transferred via the DMA channel to the Station Controller perfectly agreed with expected patterns gotten from simulations. No any bit has been distorted. Thus, the error rate for the tests can be calculated to be less than $2.5 \cdot 10^{-12}$.

Setup and hold times for the DMA transfer depend on temperature insensibly (see Fig. 4). The safety margin is sufficiently large for the whole wide temperature range.

5. New triggers

The currently used Time over Threshold (ToT) trigger performs simple a one bit analysis in order to count the number of FADC samples exceeding a fixed threshold within a sliding time window of consecutive samples. Due to the low threshold applied, equivalent to only 20% of a passing vertical muon, special attention has to be paid to avoid variations caused by pedestal and temperature fluctuations. New powerful FPGA chips allow implementing much more sophisticated algorithms in order to improve the stability, reliability and noise suppression. Particularly, making use of the full 10-bit information of the FADC traces enables selecting events with a much higher precision.

A natural estimator as an extension of the ToT idea is an integral of all FADC values in a sliding window. The integral of the time bins corresponds to the charge from PMTs and to the Cherenkov flux collected in the tank. A second estimator is the sum of the squared time bins in a sliding window. Here, the idea is to distinguish a contribution of a signal from significantly high-level time bins from low-level ones close to noise. A third estimator evaluated is the on-line calculated Area/Peak ratio.

5.1. Area over Threshold (AoT) trigger

The 1st and 2nd generation of the First Level Surface Detector Trigger based on the APEX™ EP20k200RI240-2 and in particular ACEX® EP1k100QI208-2 Altera® chips did not have sufficiently many logic registers and internal memory to implement more sophisticated triggers other than a simple threshold- and the ToT trigger.

The observed temperature dependent long and short time drifts of the pedestal call for additional routines trying to compensate for that effect. Practically, we opted the well known sigma-delta algorithm. An 18-bit counter controlling an effective pedestal is incremented/decremented by a single count depending on a comparison of the ADC value and the current pedestal. The 18-bit output value from the digital filter corresponds to the compensated 10-bit integer and 8-bit fractional ADC value. The performance of the full 18-bit bus significantly increases the accuracy of a final trigger.

The running sum of the 18 bit compensated ADC values in a sliding window of 128 consecutive samples is calculated by means of two adders, one of which is driven by a 128 word shift register (see Fig. 5). Next, the sum is compared with a preset threshold to produce the AoT trigger.
Figure 5. Structure of the AoT trigger routine. Pipeline registers are not shown. The length of the shift register (128 time bins = 3.2 $\mu s$) corresponds approximately to the currently used size of the sliding window in the ToT trigger (3 $\mu s$).

An implementation of the shift register with a length $= 2^k$, (where $k$ is an integer number) is most effective. However, this preliminary value is subject for optimization with real data.

$AoT : \sum_{i=0}^{len_i} (A_i - ped_i) > Thr_A$

A shift register is implemented as altshift_taps LPM routine utilizing the internal memory [14]. Contrary to the ToT trigger, where the length of the sliding window could be tuned dynamically, the length of the delay chain is fixed in the AoT. However, the ToT chain was implemented in logic elements (LEs) as a 1-bit shift register. Implementing 18-bit shift registers in LEs would consume an unacceptable amount of resources. However, shift registers can be implemented in internal memory blocks without access to the intermediate stages.

5.2. Power over Threshold (PoT) trigger

The digitized FADC values are composed of the pedestal and charge collected from the PMTs. In other words, the FADC traces can be considered as some rare signal on a noisy background. The power of the signal can be calculated by

$PoT : \sum_{i=0}^{len} (A_i - ped_i)^2 > Thr_P$

Sometimes, the ToT trigger can be generated by rather accidental conditions, since the shape of the FADC traces is ignored. For example, a relatively narrow but high bump (not sufficient to generate a simple threshold trigger) may be too narrow to exceed the occupancy threshold and could not be registered by ToT trigger. On the other hand, the occupancy threshold may be exceeded by some tail contribution of a low level signal, since the ToT threshold, needed to fire time bins, is relatively low.

With the above definition of the PoT, the sensitivity of a trigger to noise is reduced considerably. The contribution of time bins just above the pedestal is significantly suppressed in comparison to the relatively higher bumps.

The implementation of the algorithm is similar to the AoT case, only an additional multiplier is needed. However, a direct multiplication of 18-bit data utilizes a lot of resources. The Cyclone™
Figure 7. A prototype Cyclone™ Front End Board.

FPGA unfortunately does not contain embedded DSP blocks (such as the Stratix) and a multiplication routine has to be implemented as a soft core. To optimize the resources occupancy, let us notice, that the performance of the full 18-bit bus seems to be spendthrift. As in case of the AoT, a shift register has to be implemented in the internal memory. However, the width of such implemented bus should be a multiple of the Embedded System Block width of 8 bits. Thus, we have chosen 16 instead of 18 bits to not block memory for other routines. Neglecting the two least significant bits from the compensated ADC routine seems to be a good compromise between performance and resource utilization. Still, a 16-bit bus provides sufficient accuracy. Let us notice, that the digital filter is working independently and still compensates FADC traces with an additional 8-bit fractional counter.

Calculating a square by a direct multiplication of the same data is inefficient. Therefore, let us consider 16-bit data as a two-byte structure:

\[(256 \cdot x + y)^2 = (256)^2 \cdot x^2 + 512 \cdot x \cdot y + y^2 .\]

Moreover, multiplication by a constant value being a power of two is de facto a shift of data:

\[(ADC[15..0])^2 = (ADC[15..8])^2 \gg 16 + ADC[15..8] \cdot ADC[7..0] \gg 9 + (ADC[7..0])^2 .\]

The first and third term of the above formula correspond to the calculation of a square of 8-bit data. Results are 16-bit. The most effective way treats the ADC values as an 8-bit input address for a ROM with a 16-bit output. The ROM will contain square coefficients. However, since both terms require the same set of coefficients, it is more efficient to implement that coefficients into the Dual-Port RAM preloaded by coefficients during the FPGA configuration. The Dual-Port RAM would be working in read-only mode. Only the second term requires a real multiplier.
5.3. Area/Peak Ratio trigger

The area/peak ratio is a useful estimator calculated at present only offline. Preliminary results show that the values of the area/peak ratio are contained in a relative narrow dynamic range. That parameter could be used also for trigger purposes. Cyclone™ chips offer a lpm divide routine supporting the division of two integer numbers. The area is already calculated in the AoT routine. A parallel routine finds online a peak values in a sliding window.

The division routine is 9 clock cycles pipelined to assure sufficient registered performance. AoT, PoT and Area/Peak ratio triggers are preliminary implemented and will be tested under real conditions in the field. To optimize consumption of memory and logic registers, all new triggers are merged to a single routine to utilize the same shift registers as much as possible. All new triggers work in parallel simultaneously. Additional flags will indicate the type of triggers.

6. Conclusions

The currently used ACEX design for generating first level triggers in the Pierre Auger surface detectors offered the most cost-effective solution when starting the construction of the observatory. The Cyclone design presented here is being developed as part of a R&D work to increase the integrity of the system, to simplify it, as well as to allow implementing new features/ triggers impossible to implement in the ACEX design due to lack of resources. More resources allow better optimization the code and the reduced power consumption increases the long-term reliability. Higher registered performance allows to use faster data sampling to increase the time resolution.

7. Acknowledgments

The authors wish to thank K. Bowie for his assistance and T. Suomijärvi, J. Beatty and D. Nitz for their cooperation. One of us (Z. Szadkowski) acknowledges additional support from the Polish Committee of Scientific Researches (Grant no. 2 PO3D 011 24), as well as Collège de France, where the author started a preliminary development of the Cyclone design during December 2002 - July 2003. We also thank the anonymous referee for many valuable comments.

REFERENCES

13. “Using PLLs in Cyclone™ Devices”,
14. “Quartus II Development Software”,
http://www.altera.com/literature/lit-qts.jsp