1.1 Gigatracker (GTK)

1.1.1 Detector Requirements and Performance

The GigaTracKer (GTK) is a core part of the NA62 experiment. It is a spectrometer that provides precise measurements of momentum, time and angle of the incoming 75 GeV/c kaon beam. Beam measurement is essential for the selection of the $K^+ \to \pi^+ \nu\bar{\nu}$ decay and the background reduction will be achieved by the kinematical constraint based on the missing mass variable:

$$m^2_{miss} \approx m^2_K \left(1 - \frac{|P_\pi|}{|P_K|}\right) + m^2_\pi \left(1 - \frac{|P_K|}{|P_\pi|}\right) - |P_K||P_\pi|\theta^2_{\pi K}$$

In the equation above $|P_K|$ is entirely derived from GTK while the angle, $\theta_{\pi K}$, is derived from the measured directions of the decaying kaon given by the GTK and of the pion from the decay, which is measured with the downstream spectrometer, the STRAW detector, which also provides the measurement of $|P_\pi|$. This is symbolically indicated in the decay sketch in **Error! Reference source not found.** on page **Error! Bookmark not defined.**.

The Gigatracker is composed of three stations (GTK1, GTK2 and GTK3) mounted around four achromat magnets as shown in Figure 1.

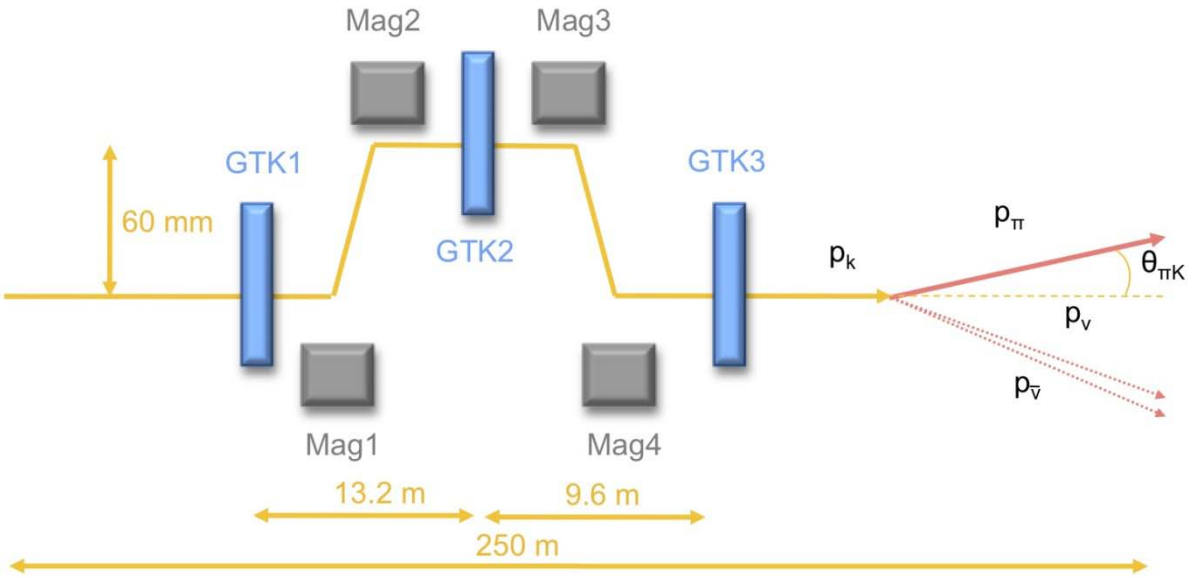

Figure 1 Layout of the Gigatracker stations and sketch of the decay.

This setup is placed along the beam line inside the vacuum tank, just before the fiducial region in the decay vacuum pipe. It has to sustain a high and non-uniform beam rate and has to survive in high radiation environment.

The overall simulation of the kaon decays in the detector (see section **Error! Reference source not found.**) has determined the beam track momentum and direction resolution requirements. From this study and taking into account the expected STRAW resolution, it has been derived that the Gigatracker has to measure the momentum with a relative resolution of $\sigma(p_K)/p_K \sim 0.2$ % and the

direction with a resolution of the order of 16 μrad. A pixel size of 300 μm x 300 μm is sufficient to achieve the required resolution.

Finally the beam spectrometer has to sustain a high and non-uniform beam rate of 0.75 GHz in total, hence the name Gigatracker, with a peak of 1.3 MHz/mm^2 around the centre.

To meet these resolution requirements the beam spectrometer is installed in vacuum with a minimal amount of material crossed by the beam to preserve the beam divergence and to limit beam hadronic interactions. The spectrometer is composed of three hybrid silicon pixel stations. Each station is made of one hybrid silicon pixel detector with a total size of 63.1 mm x 29.3 mm containing 18000 300 μm x 300 μm pixels arranged in a matrix of 90 x 200 elements. With this configuration the detector matches the expected beam dimensions of ~60 mm × 27 mm. The pixel dimensions and the distances between stations are adapted to deliver the required momentum and direction resolution.

The amount of material crossed by the beam at each station influences the angle measurement. The chosen sensor thickness of 200 μm corresponds to 0.22 % of a radiation length, X_0. The design efforts take into account to minimize the material as much as possible as the physics performance strongly depends on a low material budget. Giving an equivalent budget to the read-out and to the cooling the total amount of material per station has been required not to exceed 0.5% X_0.

Furthermore, due to the high intensity, the required time resolution on every single track using all three stations is 150 ps (rms).

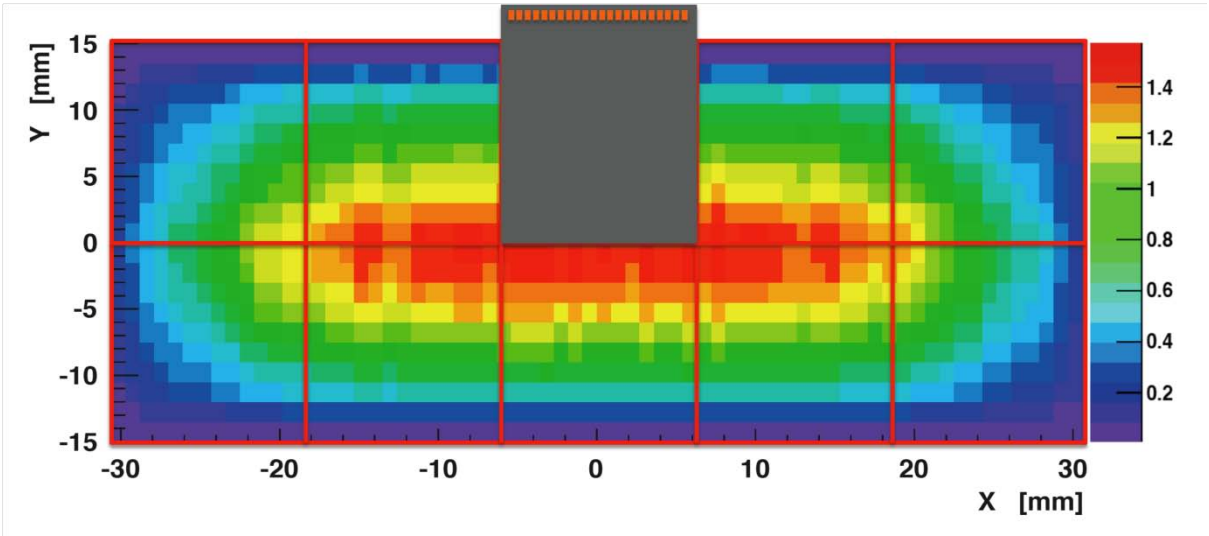

Figure 2 Beam intensity distribution over GTK station 3 (units are MHz/mm^2). One of the 2 x 5 read-out chip is drawn.

This is a novel requirement for a silicon pixel detector and none of the existing systems has such a capability.

The beam region is completely covered by one single, fully active silicon sensor bump-bonded to 2 x 5 read-out chips (ROCs), which each deliver a time resolution of better than 200 ps for each pixel hit, allowing to achieve a track resolution better than 150 ps. A description of the system is reported in an internal note (1). Figure 2 shows the intensity distribution in station 3.

The expected fluence for a typical 100 days run time year is 2×10^{14} 1 MeV n equivalent per cm^2 in the central region of the sensor. This value is comparable to those expected in the inner layers of the LHC trackers over ten years of operation and requires the sensors and the application specific integrated circuits (ASIC) to cope with the high radiation environment. Section 1.1.3 describes the measurements performed on sensors. In order to reduce the radiation induced leakage current on the relatively large sensor the operating temperature is set to 5^oC or below. The very low mass of the detector, the operation in vacuum and the need of limiting the leakage current increase due to radiation damage demands a very efficient and reliable cooling system.

Table 1 shows the principle detector specifications. These parameters are to be seen as acceptable values. The research efforts described in the following sections aim for improvements with respect to these values especially for the time resolution, the material budget and the efficiency and will result in a largely improved physics performance. In the following sections specifications derived from these parameters described.

Table 1 Detector performance requirements.

Material budget per station	≤ 0.5 % X_0
Overall efficiency per station	> 99 %
Dynamic range	0.6 -10 fC; 5'000 – 60'000 e^-
Pixel size	300 µm x 300 µm
Time resolution	< 200 ps rms

.

1.1.2 Detector Elements and Architecture

Each component of the hybrid silicon pixel detector is optimized in terms of material budget and performance. Hybrid pixel detector modules establish electrical connections between the silicon sensor and the read-out chips using Sn-Pb solder bumps. For the GTK, the sensor thickness is 200 µm and the thickness of the read-out chips is 100 µm. The stations are located in vacuum. For the on-detector system integration the following parameters are optimized.

- Material budget in active beam area.
- Read-out chip size and active matrix size.
- Accessibility to power and I/O connections on the chip.
- Particle rate per pixel.
- Heat dissipation and cooling.

In order to minimize material and maximize geometric efficiency in the active beam area any connections to the read-out chip are outside the beam area. The beam profile has been adapted in such a manner that two adjacent rows of read-out chips cover the beam area. Figure 2 illustrates the principle. The ASICs are accessed only from the top or the bottom of the assembly. Thus wire bond connections to the read-out chip are placed only on one side of the ASIC. The active pixel matrix

extends to the other 3 sides of the chip, allowing a maximum of efficiency in the beam centre with the highest particle flux.

Two options for the heat transfer are under investigation. One is based on a silicon substrate with micro cooling channels. The other is based on placing the module in a vessel with gaseous cooled nitrogen flowing. The module is placed in vacuum and the heat dissipated by the 10 read-out chips is estimated to be less than 2 W/cm^2 for a total of 32 W. The size of the active pixel matrix on each chip is 13.5 mm x 12 mm, corresponding to 45 rows x 40 columns of pixels. On the chip and outside the active pixel matrix the end-of-column circuitry performs data processing, control, configuration and off-chip data transfer.

An important design aspect of large pixel matrices, where power supply is provided from one side only, is the power drop along a column. The power supply metal plane layout has to be carefully considered. For chips in 0.13 µm CMOS technology and a column length of 13.5 mm with a power consumption of 2 W/cm^2 simulation studies indicate that powering from one side can be done.

In the configuration presented the material budget is uniform over the entire active area. Table 2 summarizes the material budget of the active components in the beam. The sensor and system parameters are discussed in sections 1.1.3 and 1.1.4.

Table 2 Material budget of sensor, read-out chip and bump bonds.

Component	**Material**	**Thickness [µm]**	**X_0 [%]**
Sensor	Si	200	0.22
Bump Bonds	Pb-Sn	≈25	0.001
Readout Chip	Si	100	0.11

Figure 3 shows the system block diagram. The GTK assembly (sensor and 10 pixel chips) is electrically connected via wire bonds to the GTK assembly carrier printed circuit board (PCB), which connects the ASICs to high/low voltage power, and to the serial control and to the status and data interfaces of each chip using dedicated signal traces. Special care must be taken with the layout as the power supplies (LV, HV up to 600 V), the common station ground plane, on-chip thermal sensor connection and up to 40 Gbit-serial links are in close proximity. The assembly carrier also acts as a vacuum feed-through. Inside the vacuum, only filtering capacitors, a temperature and humidity sensor but no active components are placed on the assembly carrier. In order to keep the stations electrically isolated from the control room and increase noise immunity, the data transmission to the off-detector read-out electronics is over optical fiber. Thus optical/electrical converters for the serial links to the control room are placed outside the vacuum. Local LV regulators are being considered. However, this will limit or complicate the remote control capabilities of the LV supply of the read-out ASICs. Depending on the radiation effects on the ASICs, compensation on the power supply voltage might be appropriate.

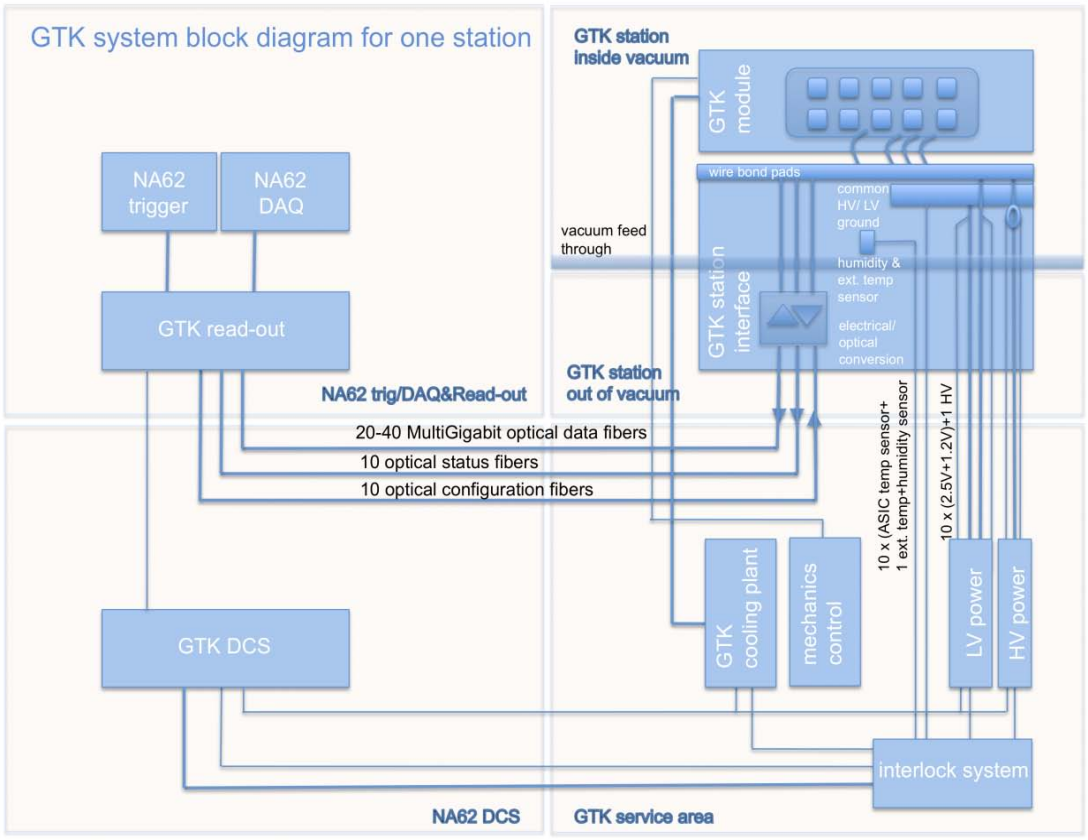

Figure 3 GTK system block diagram.

1.1.3 Silicon Sensor and Bump-Bonding

1.1.3.1 Silicon Sensor

The silicon sensor thickness has been studied in simulation, see section 1.1.4, and an overall thickness of 200 µm was found to be compatible with the signal requirements for precise timing information together with the material budget constraints.

A detailed study of the radiation damage effects was carried out using test-diodes produced by the manufacturer that will also deliver the final GTK sensor wafers (FBK-Trento, Italy). Irradiations were carried out up to a fluence of 2 x 10^{14} 1 MeV n_{eq} cm^{-2} and detailed annealing studies investigated the effects on leakage current and full depletion voltage. The results have been summarized in an internal NA62 note (2). The maximum acceptable operating temperature for the GTK stations has been set to 5°C which allows an operation time of approximately 50 days using standard p-in-n sensors as measurements on irradiated diodes indicate that the leakage current of a full size sensor, at an operating temperature of 5°C and 50 days of operation, will reach a maximum acceptable value of approximately 270 µA. However, efforts are being made to design the cooling system to operate at -20°C, which allows an operation of 100 days. Continuous monitoring of the sensor leakage current will allow the radiation induced leakage current increase to be followed and also provide an estimate of the overall fluence. A precision of 100 nA or better seems to be adequate.

The configuration of the experiment will permit the exchange of the Gigatracker stations at the end of their operation period. The baseline choice for the sensor material is currently p-in-n due to

simpler processing procedures, availability and experimental arrangement. However, the front-end design will allow the study of the connection of other sensor types, such as n-in-p or similar materials.

A mask design layout has been defined and a batch of p-in-n sensors has been delivered as 4" wafers. The layout of the wafer is shown in Figure 4, indicating the large silicon sensor for the full-size stations in the centre, surrounded by test sensors foreseen to be used with the prototype electronics and general test structures. A detailed list of the individual sensor parameters is given in Table 3.

The wafer substrate is n-type (phosphorous doped) with a resistivity of about 4-8 kΩcm. The wafer thickness is (200±10) μm and the bow (after processing) is less than 30 μm to comply with the flip chip bonding requirements. The wafer back side is metalized to allow wire bonding contacts for applying the reverse bias to the sensor.

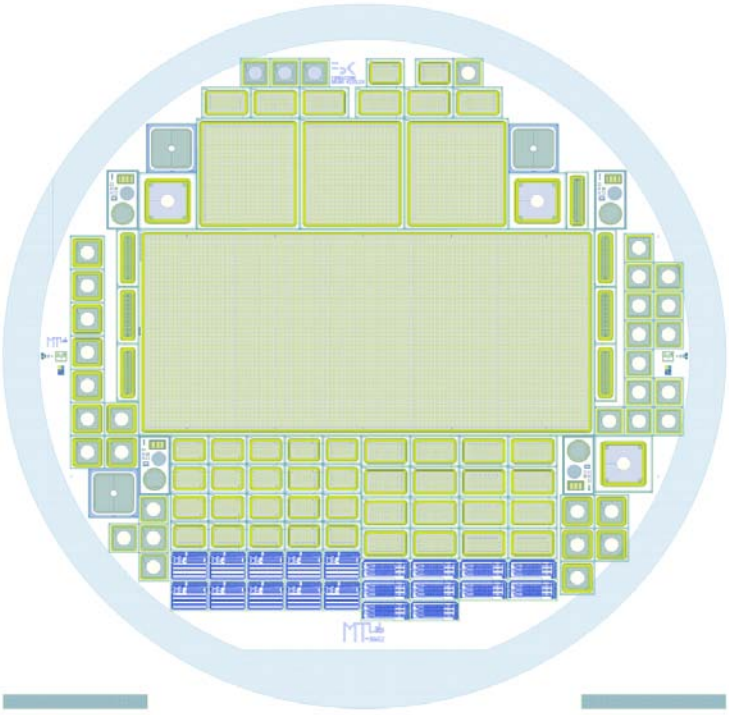

Figure 4 CAD drawing of sensor wafer.

Table 3 Main sensor elements on wafer

Number	Structure
1	**Full-size sensor:** • 18000 pixels (90 × 200 matrix) • 27.0 × 60.8 mm^2 active area • 29.3 × 63.1 mm^2 effective size on wafer

3	**Single-chip size sensor:**
	• 1800 pixels (45 × 40 matrix)
	• 3 types (see Figure 4):
	  • A: 300 µm × 300 µm pixels everywhere
	  • B and C: enlarged pixels in lateral columns (300 µm × 400 µm)
	• 13.5 × 12.0 (12.2 for B and C) mm^2 active area
	• 15.8 × 14.3 (14.5 for B and C) mm^2 effective size on wafer

22	**EOC prototype sensor:** • 60 pixels (6 × 10 matrix) • 1.8 × 3.0 mm^2 active area • 4.1 × 5.3 mm^2 effective size on wafer
22	**P-TDC prototype sensor:** • 105 pixels (7 × 15 2.1 × 4.5 mm^2 active area • 4.4 × 6.8 mm^2 effective size on wafer

In order to achieve the required timing precision it is necessary to operate the GTK sensors at high over-depletion. A target operating voltage of approximately 500 V has been defined and a multi-guard ring structure is implemented in every sensor in order to ensure stable operation at relatively high voltage.

Charge collection efficiency, speed and sensor capacitance influence the overall timing performance of the GTK station. Ideally the sensor characteristics should not degrade timing. However, under irradiation of the sensor the collected charge degrades. This degradation is non-uniform from chip to chip and pixel to pixel and may introduce non-uniform pixel time response.

1.1.3.2 Bump Bonding

Each pixel cell contains one bump pad to connect it with the corresponding cell in the front-end chip. In addition the guard ring of each sensor needs to be connected to the corresponding pads on the front-end chip using bump bonds. The bump-bonding pads have octagonal shape (26 µm with 20 µm opening in the passivation) and the bump pad centre is located at about 50 µm from the pixel edge. The bump-bonding pads are arranged in a mirrored scheme on neighboring columns (Figure 5) with enlarged pixel cells (300 µm x 400 µm) in the interchip region. This allows a fully sensitive area over the full sensor size to be maintained. A schematic drawing of the implemented geometry is shown in Figure 5 and Figure 6.

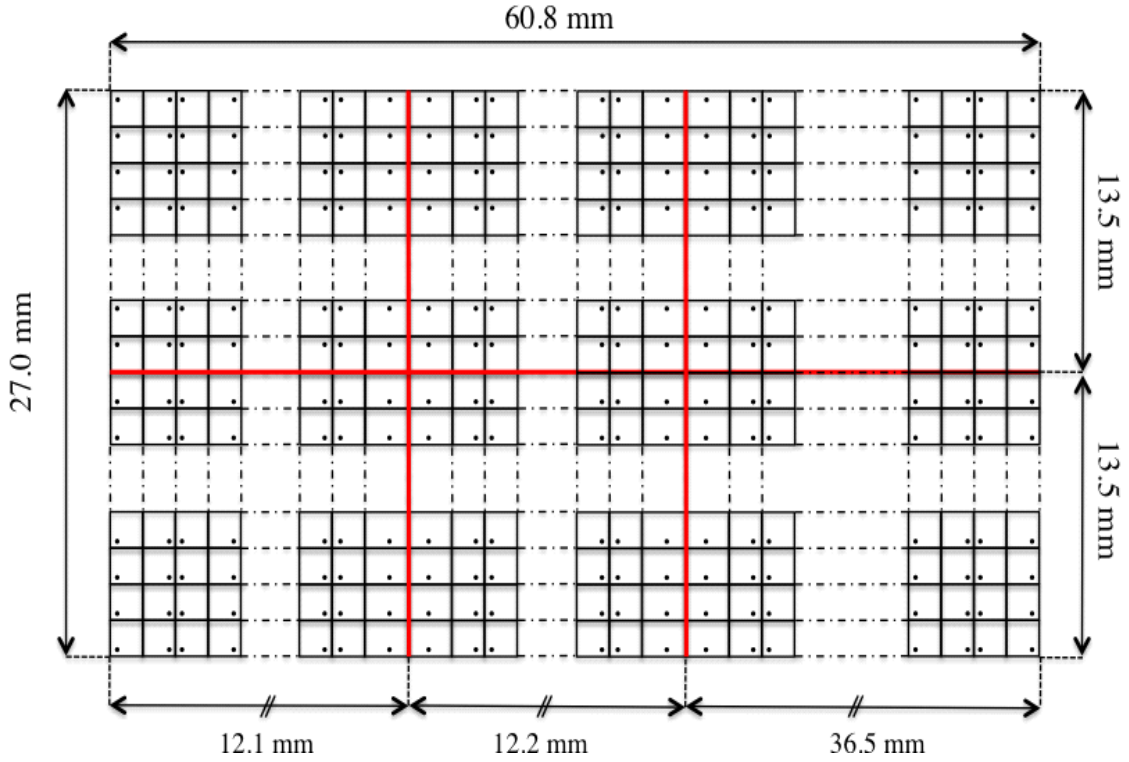

Figure 5 Bump pads arrangement scheme.

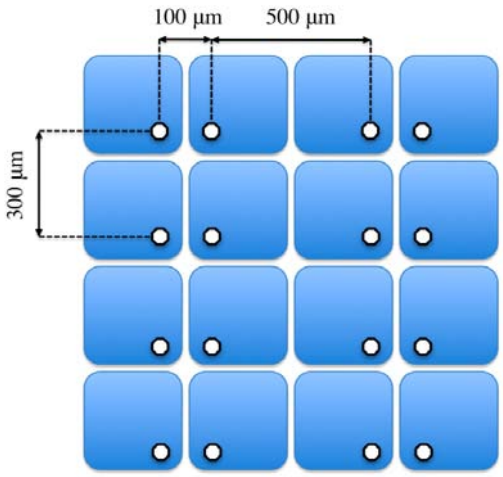

Figure 6 Schematic layout of the full-size sensor. Enlarged pixel cells (300 µm x 400 µm) are put in the border region between neighboring read-out chips.

Prototype assemblies with demonstrator ASICs containing only 1-2 columns and final GTK assemblies with ASICs containing the full matrix will differ in the way that the prototype chips will not be available on full size wafers. Generally the sensors will be made available on 4" wafers of 200 µm thickness. The prototype front-end chips will be available as single dies from multi project wafer runs whilst the final GTK front-end chips will be available on 200 mm diameter wafers. This will require two different technical solutions to build flip chip bonded assemblies for the prototype and for the final GTK planes.

A dedicated market survey was launched in July 2009 to identify companies capable of flip chip bonding the prototype assemblies and the final GTK assemblies. The two companies that fulfilled the

requirements defined the survey carried out at CERN (IZM Berlin and VTT Helsinki) have been invited to reply to a price enquiry for the prototype production. Following the price enquiry the contract for the production of prototype bump-bonded assemblies was awarded to IZM Berlin.

Specific emphasis is laid on the thinning of the final GTK readout chips in order to meet the material budget requirement. A final thickness of 100 µm or less will comply with the requirements, but will also demand the development of specific process steps during production.

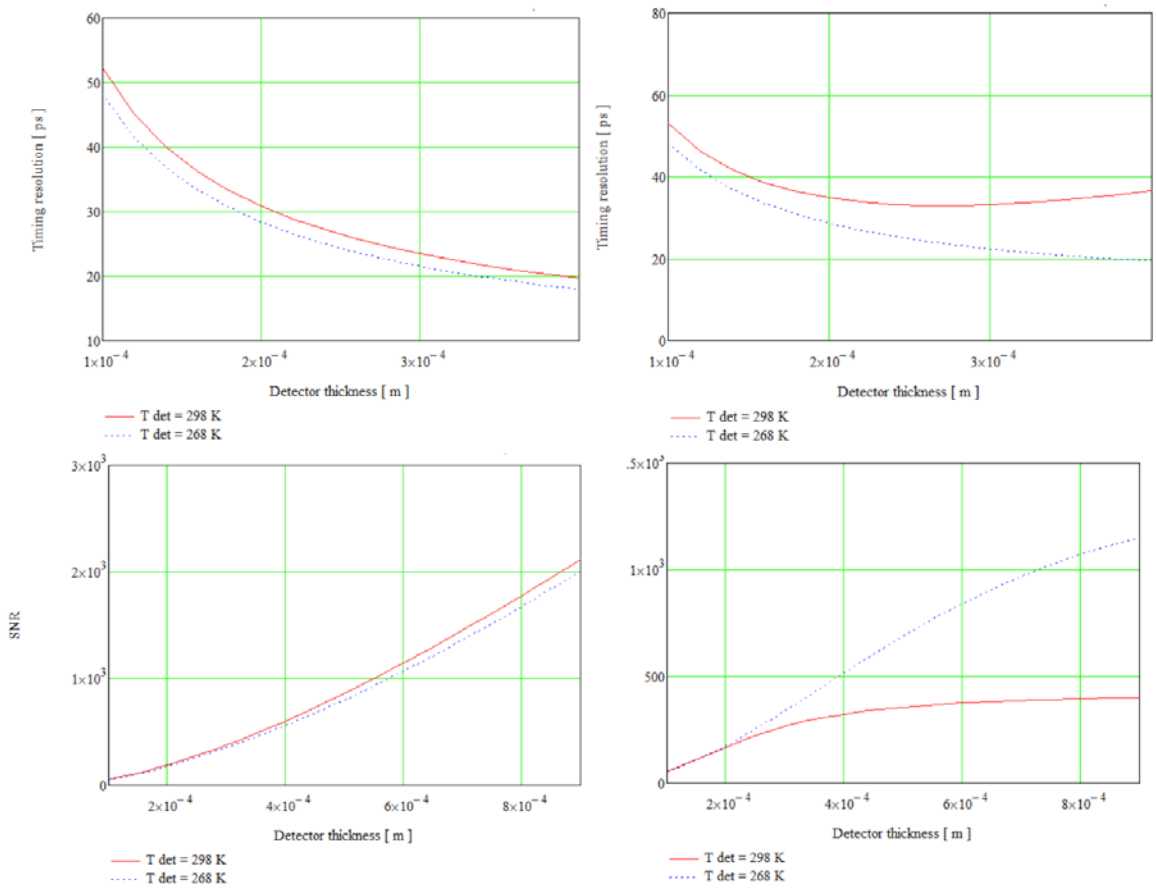

Figure 7 Simulation of time resolution and signal to noise ratio over sensor thickness.

1.1.4 Readout ASICs

The design of the pixel read-out chip poses unprecedented challenges in the combination of the required pixel cell size of 300 x 300 µm^2, the particle arrival time measurement resolution of better than 200 ps and the material budget of less than 0.5 % X_0 of radiation length per station. The low material budget translates to a constraint on the power consumption of less than 2 W/cm^2 as the cooling system needs to be kept to a minimum. A trade-off between signal amplitude, signal speed and acceptable material in the active area defines the thickness of the sensor to 200 µm.

Figure 7 gives the theoretically achievable time resolution and the signal to noise ratio over the sensor of the expected detector signal and the front-end electronics thickness for the most probable input signal of 2.4 fC for two leakage current values representing the sensor state before and after

irradiation. The study (3) takes measured noise values from a 0.25 µm ASIC technology into account, but similar results were obtained with the selected 0.13 µm ASIC technology. The shaping time is assumed to be identical to the charge collection time of the respective detector thickness. The power consumption in the first input transistor is assumed to be 325 µW.

The time resolution can be expressed as

$$\sigma_{\mathrm{t}} \propto \frac{\mathrm{ENC} \cdot \tau_p}{\mathrm{Q_{in}}}$$

where ENC is the Electronic Noise Charge, τ_p represents the speed of the preamplifier and Q_{in} is the input signal charge.

For low leakage currents the white noise dominates, which decreases when the shaping time is increased with thicker sensors (Figure top left). As the signal and speed change approximately in the same way, time resolution is determined mainly by the noise. Thus the time resolution decreases with sensor thickness. For high leakage currents after irradiation and at room temperature increasing the shaping time does not decrease the noise anymore sufficiently and the parallel noise, which increases with shaping time becomes dominant (Figure top right). But cooling the detector reduces the parallel noise and brings us back to a situation similar to the pre-irradiation one. The two plots in the bottom of the figure show the signal to noise ratio (ratio of input signal divided by the rms noise values) before irradiation (left) and after (right), when increasing detector thickness the input signal is increased and thus the signal to noise ratio. When the parallel noise due to the increase of the leakage current after irradiation becomes dominant (Figure 7 bottom right plot), the signal to noise ratio still increases for thicker detectors, but at a slower pace. In conclusion a 200 µm thick detector is a good compromise between achievable time resolution and low material budget. The study also shows that cooling irradiated detector improves the time resolution.

Figure 8 shows the expected charge distribution (Landau) with the most probable value around 2.4 fC. The available signal charge (5000 to 60000 electrons or 0.8 to 10 fC) and the requirements on time resolution and jitter need analogue front-end time-walk compensation.

The integration of the analogue front-end together with time-to-digital converters and the high speed on chip data acquisition and read-out circuits need to be done with the utmost care in order not to disturb the analogue performance. Two different approaches have been considered to tackle this problem. One is to integrate the time-to-digital converter (TDC) together with the analogue front-end directly in the pixel cell. The TDC will be active only during a hit in that particular cell. A reference clock needs to be distributed over the entire pixel matrix posing a possible noise source to the analogue front-end electronics. The other approach is to keep the pixel cell as simple as possible, containing only the analogue front-end, and send the discriminator signal to the end-of-column area of the pixel matrix inside the ASIC where the TDC evaluates the arrival time. In this configuration the pixel cell stays simple, no clock needs to be distributed, but the discriminator signal needs to traverse the pixel matrix to be evaluated in the end-of-column.

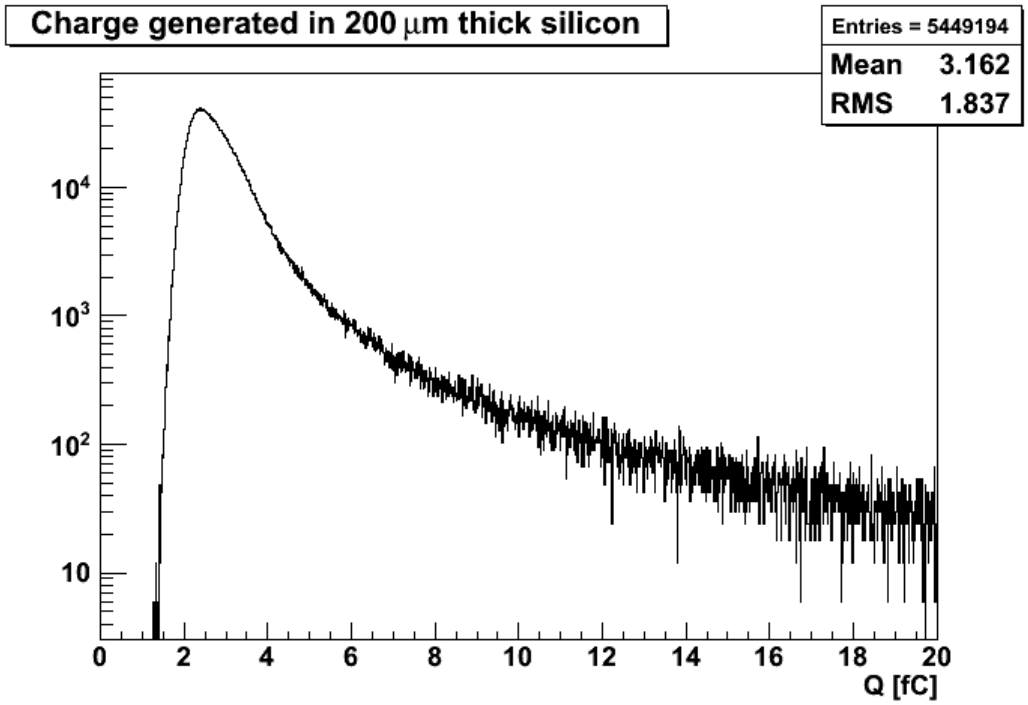

Figure 8 Landau distribution of input charge with most probable value around 2.4 fC.

The specified time resolution requires efficient time walk compensation. Two time walk compensation architectures have been investigated. a) Constant Fraction Discriminator (CFD), which directly gives a compensated discriminator signal, but uses sensitive analogue circuitry and b) Time-over-Threshold (ToT) discriminator, which delivers two trigger events to be time tagged but uses simpler and less sensitive circuitry. Compensation by sampling the signal shape with a Flash ADC has been excluded for size and power consumption reasons.

As simulations cannot show the performance and drawbacks of the architecture and time-walk compensation schemes in real life, it was decided to produce test structures for both architectures and compensation schemes.

1.1.4.1 ASIC Specification

The GTK pixel electronics is very demanding and requires performance close to the limits of the 130 nm CMOS technology used. The circuit is optimized in view of trade-offs between acceptable power dissipation of 2 W/cm^2, timing precision of less than 200 ps, centre pixel particle rate of 124 kHz/pixel and an output data rate per readout chip of up to 6 Gb/s including data transmission overhead. Table 4 presents a summary of the specification.

The beam particle rate is 750 MHz. It corresponds to the following rates for the centre chip. The hit rate in the centre pixel is 114 kHz. The hit rate of the centre column is 2.7 MHz or 0.66 MHz/mm^2. The average hit rate per pixel on the centre chip is 58 kHz with a chip hit rate of 105 MHz. Table 4 summarizes the rates.

Assuming a 32 bit word per hit, and considering the average chip hit rate, a data rate of more than 4 Gb/s needs to processed and sent off each read-out chip. On-chip triggering is ruled out as the earliest trigger signal arrives 1 ms or later after the event, accumulating 4 Mbit of data during the trigger latency. The area of silicon needed to store the data and process the trigger on chip would be too large. Consequently the data are shipped off the chip in a triggerless, data driven architecture. In order to reduce digital noise and connectivity a low number (2-4) of Gbit serial links are used for each chip.

Table 4 GTK ASIC specification

General	Max. material budget per station	0.5 % X_0
	Overall efficiency per station	> 99 %
	Pixel size	300 µm x 300 µm
Architecture/geometry	Chip active area	13.5 mm x 12 mm
	Chip size	19.5 mm x 12 mm
	Number of pixels per chip	40 columns x 45 rows = 1800
	Size of pixels	300 µm x 300 µm
	Thickness of read-out chip	100 µm
Pixel analogue front end	Input dynamic range	0.6 -10 fC; 5000 – 60000 e^-
	Electronic noise in pixel input with sensor	200 e- rms
	Time resolution of the discriminator	100 ps rms
	Discriminator time walk after correction	80 ps rms
	Overload detection threshold	10 fC; 60000 electrons
	Maximum dark current per pixel tolerance	1 µA
	Power dissipation per pixel	1 mW max
	TDC time binning	100 ps
Data rate	Beam rate	750 MHz
	Average centre chip hit rate	105 MHz
	Average particle rate per pixel in station	42 kHz
	Average particle rate per pixel in centre chip	58 kHz
	Peak rate per pixel /beam centre	114 kHz
	Average centre column hit rate	2.7 MHz or 66 MHz/cm^2
	Centre chip data rate (40 bit per hit word, no contingency, no encoding)	4.3 Gb/s
	Center chip data rate (incl. 8b10b encoding, 30% contingency)	6.4 Gb/s
	Serializer data rate (4 used per chip)	1.6 Gb/s
	Read-out efficiency	99% (98% in beam centre)
	Operating temperature in vacuum	-20 ^oC to +5 ^oC
Radiation	Centre pixel total dose in 1 year	2.10^{14} particles/cm^2, $6*10^4$ Gy
Trigger	Number of levels	1 (L0)
	Latency	~ 1 ms initially, future upgrade several ms.
Off chip trigger	Trigger window	75 ns

In order to identify high charge deposit in the sensor and the ASIC provides information if a charge of higher than 10 fC has been deposited.

1.1.4.2 ASIC Architecture

The active pixel matrix of 13.5 mm x 12 mm in size arranged in 45 rows x 40 columns of pixels with dimensions of 300 µm x 300 µm, results in 1800 pixels per chip. The chip area outside the active beam area used for data processing and control is estimated to be in the order of 4.5 to 6 mm deep resulting in a total chip size of 18 to 19.5 x 12 mm, see Figure 2. The communication to and from the pixel read-out chip is done via high speed differential signals. This reduces the number of connections and thus increases module reliability. The number of I/O pads is limited, as the connections to the module are done on one side of the chip only, see Figure 2. A high number of single ended data lines would increase the digital system noise. The clock signal and the serial control signals are connected on two high speed serial input lines. Configuration data and settings are transmitted on one slower serial input port and one parallel input port. The pixel hit data output ports consist of two to four multi gigabit serial ports for the data and one slower serial port for status information. For test purposes a number of single ended outputs are foreseen. Figure 9 illustrates the I/Os of the chip.

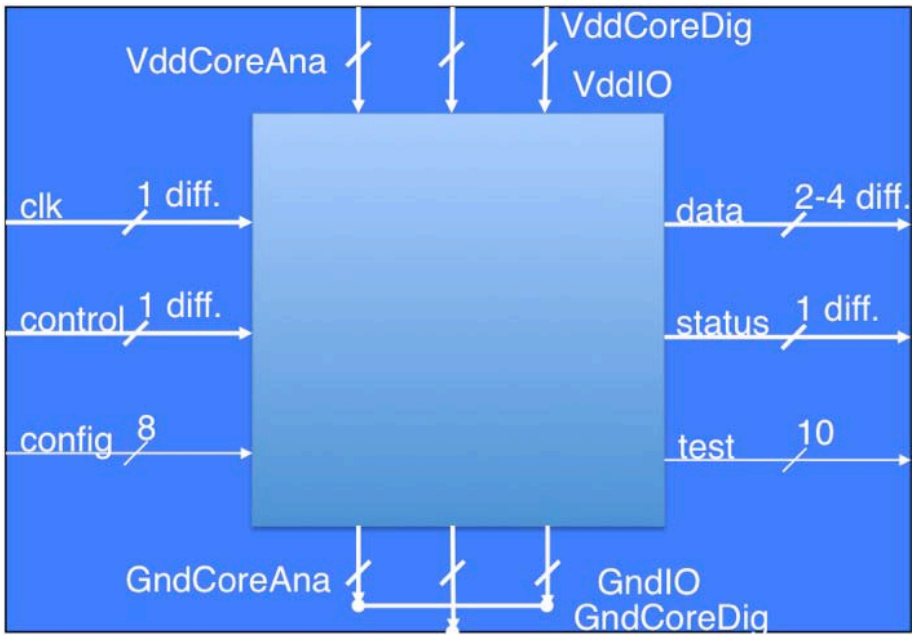

Figure 9 I/O connections of the GTK pixel read-out chip.

The design and layout of the digital processing units must be done so that the digital noise has little effect on the analogue front end. The implementation of the pixel read-out chip GTK is done in 0.13 µm CMOS technology. This technology is fully available and design tools have been provided. First studies indicate that the technology might be sufficiently radiation hard by itself for the expected average total dose of 10^5 Gy/year. However, a large variation of dose is expected between chips of a GTK station and between chip edge and centre. This dose gradient might introduce variation of

signal response between pixels. To circumvent this effect, the trimming in each pixel used for process mismatch can also partially compensate for the dose gradient.

Critical registers and nodes are protected against single event upset effects. The 130 nm has a relatively high substrate resistivity when compared to 250 nm technology. This makes it easier to isolate digital and analogue blocks by means of guard-rings. However, the high sensitivity and large bandwidth of the analogue front-end require the use of architectures to reduce the digital noise such as differential analogue and digital circuits, which tend to reduce both propagation of ground and substrate noise and improve immunity of the analogue front-end.

1.1.4.3 Design Options

Both architectures described above and time-walk compensation techniques have been developed and integrated into two demonstrator ASICs. In one the CFD and the TDC are implemented inside the pixel cell – TDC per pixel (P-TDC). In the other chip each pixel contains a time-over-threshold discriminator and a transmission line driver to send the discriminator signal to the end-of-column (EOC). Arrival time of leading and trailing edge of the hit signal are time tagged in delay locked loop (DLL) based TDCs in the EOC.

Both demonstrator chips contain at least one full, folded column containing the final analogue front-end, the full TDC functionality but reduced read-out capability compared to the final pixel chip design. The results of the qualifications are described in sections 1.1.4.5.4 and 0. The characterization of the demonstrator front-end chips in view of building the final full pixel matrix and selecting the final building blocks and architecture will address the following questions in general:

a) time resolution; b) charge collection efficiency (sensor/ASIC assembly); c) read-out efficiency.

The following paragraphs outline the measurements in detail. If applicable all tests are done with and without sensor. For the silicon assembly (sensor/ASIC) the parameters also are measured with respect to varying high voltage and position of charge injection on the pixel. Furthermore the absolute charge collection gain is determined.

- Arrival time resolution
  (for 2.5 ns charge injection duration and read-out clock activated)
  - Full chain time walk or compensation error over input charge and folded over Landau charge distribution.
  - Full chain jitter over input charge and folded over Landau charge distribution.
  - Influence of read-out clock on full chain resolution/jitter - digital clock/noise separation
  - Behaviour of long transmission lines (clock or signal)
  - Influence of signal charge injection duration on full chain resolution/jitter
- Analogue frontend behavior
  (for 2.5 ns charge injection duration and read-out clock activated)
  - analogue fronted time walk or compensation error over input charge and folded over Landau charge distribution
  - analogue frontend jitter over input charge and folded over Landau charge distribution
  - Front-end noise
  - Gain distribution

  - Offset distribution before trimming
- TDC behavior
  - TDC resolution
  - TDC INL/DNL
  - TDC jitter
- Information about deposited charge in the sensor
- System robustness
- Simulated efficiency for full matrix chip
- Simulated power of individual blocks and measured power as far as possible
- Size of individual blocks
- Temperature stability
- Estimated and possibly measured SEU rate of individual blocks

The P-TDC option uses a constant fraction discriminator in the pixel already compensating for the time walk. In the EOC column option a time-over-threshold discriminator is used which provides signal amplitude information by measuring the pulse width allowing the compensation of the time walk.

In the P-TDC option, a dual slope Wilkinson TDC in each pixel converts the discriminator edge into a time stamp synchronized with a master clock distributed to the entire pixel array. In the pixel a buffer and derandomizer is used to increase the read-out efficiency before the data are sent to the column bus and subsequently off the chip. In the EOC pixel circuit only a driver follows the discriminator stage to send the hit signal onto the column bus. There the hits are time tagged using a delay locked loop based TDC before being sent off the chip.

The features of the two options are summarized below. In the P-TDC, noise might be injected in the analogue circuits in the pixel array due to the digital activity in the pixel cell. The clock needs to be distributed over the entire matrix with a minimum of jitter to avoid non recoverable errors on the time tagging. In the EOC option each individual discriminated pixel signal needs to be sent to the EOC, however, the transmission lines are only active once a hit has occurred. The ToT compensation scheme imposes the measurement of two time tags increasing the output data volume and rate. The ToT discriminator is better adapted to measure the signal amplitude to detect high energy transfer in the sensor. The CFD discriminator would need to be modified to include large signal detection.

Table 5 summarizes the 2 architecture blocks. Figure 10 and Figure 11 show the basic architectures of the P-TDC and the EOC respectively.

Table 5 Summary of the design option for the p-TDC and EOC architectures.s

Chip Architecture	Time walk compensation	Pixel Circuit	Column Bus	EOC	Backend
P- TDC	CDF discriminator	TDC, Buffer, derandomizer	Synchronous bus	Buffer Derandomizer, Coarse counter drivers, read-out control	Serialiser FIFO logic
EOC	ToT discriminator	Hit driver	Asynchronous bus	EOC TDC, buffer derandomiser	Serialiser FIFO logic

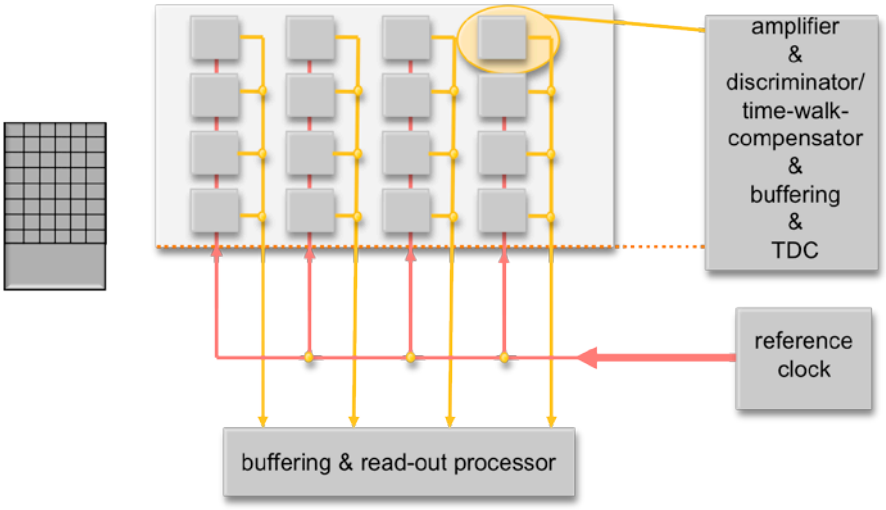

Figure 10 Block diagram Pixel in TDC option (pTDC).

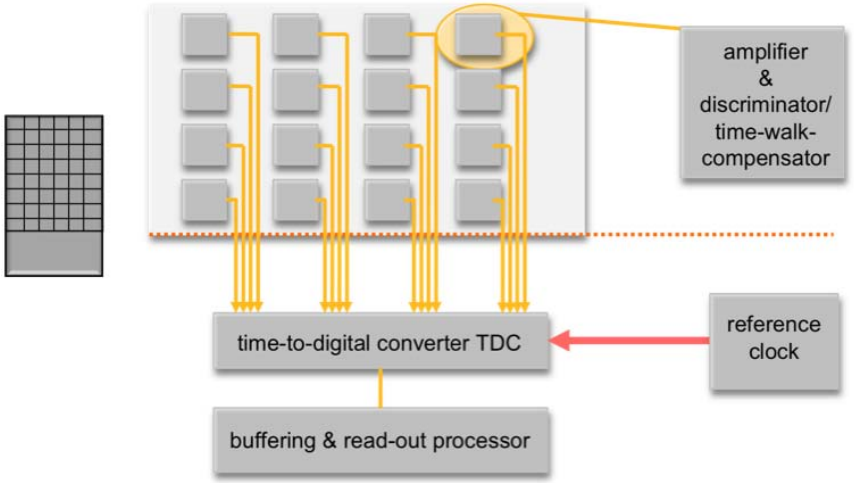

Figure 11 Block diagram End-of-column option.

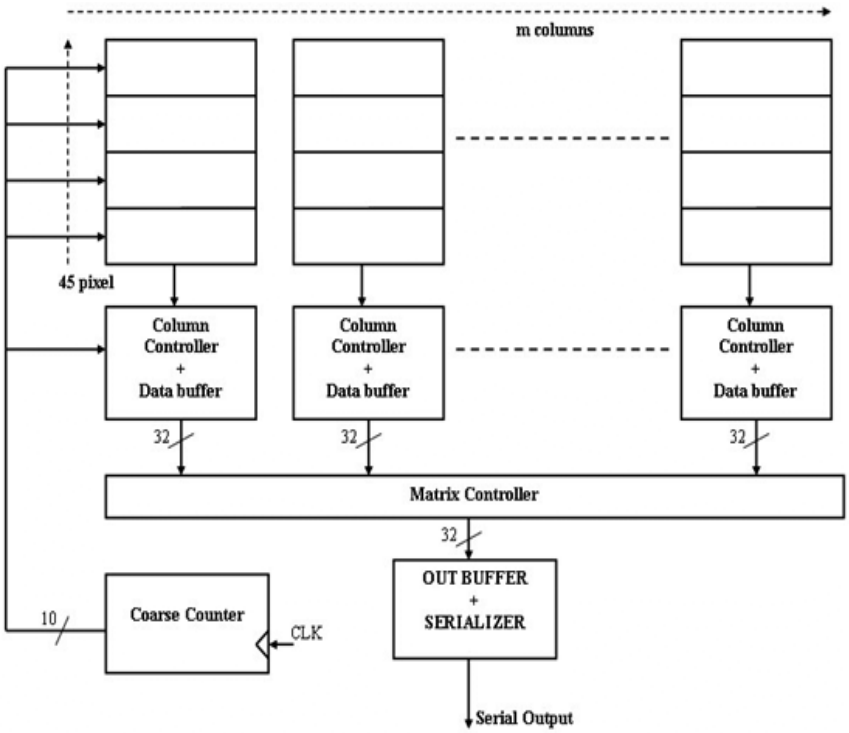

Figure 12 Block diagram of the P-TDC option.

1.1.4.4 TDC per Pixel (P-TDC) Architecture Overview

In the P-TDC most of the signal processing, including time to digital conversion and multi-event buffering, is performed inside each pixel. Figure 12 shows a simplified block diagram of the P-TDC architecture, while a scheme of the pixel cell is shown in Figure 13. Each column, 45 pixels, is read-out by a dedicated controller. The data are then merged and prepared for the output serializer. In this approach only digital signals are exchanged between the pixels and the periphery. The most critical signal is the clock, which must be distributed with high precision to the whole matrix. A dedicated transmission line for each pair of column is used for this purpose.

The time walk error is corrected at the analogue level with a Constant Fraction Discriminator (CFD). The master clock of the chip has a design frequency of 160 MHz. The clock pulses are counted and the resulting word is Gray-encoded and sent to all pixels. When a hit is flagged by the CFD, the value present on time-stamp bus is latched into local registers, providing a first coarse estimation of the event time. In each pixel, a TDC based on a Time to Amplitude Converter (TAC) generates a voltage proportional to the interval between the hit and the next clock edge by charging a capacitor with a constant current source. The resulting voltage is then digitized, providing the required fine time measurement. Figure 13 shows a block diagram of the P-TDC pixel cell.

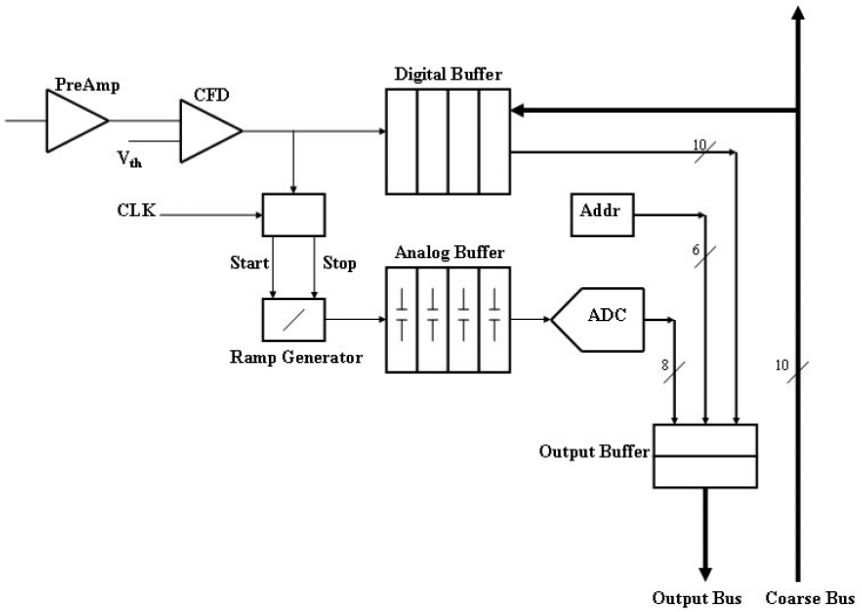

Figure 13 Block diagram of the P-TDC pixel cell.

1.1.4.4.1 P-TDC Preamplifier and Constant Fraction Discriminator

The analogue section can be divided into three main blocks: the front-end amplifier, the CFD and the TDC. The front-end has been designed to be compatible with sensors of either polarity. In order to minimize the risk of digital noise pick-up a fully differential topology is used for both the front-end amplifier and the CFD. The preamplifier employs a telescopic cascode with a class AB output stage. The latter is needed to drive the CFD with good power efficiency. A low-frequency feed-back, implemented with an RC low-pass filter followed by a slow op-amp, allows the compensation of the sensor leakage current. Dark currents up to +/- 200 nA can be accommodated while keeping the resulting output offset below +/- 10 mV. The input stage of the Constant Fraction Discriminator is formed by a passive RC filter, in which both the delayed copy of the signal and the signal fraction are generated. The filter is followed by a high gain differential amplifier, made by cascading two differential cells with resistive load. This amplifier drives a third stage, composed by two differential pairs sharing the same load. One pair is driven by the signal while the other one is used to apply the threshold. A hysteresis is introduced at this level to avoid excessive noise triggering. Finally, a differential to single ended converter generates a logic pulse of 1.2 Volt, suitable to drive the digital part.

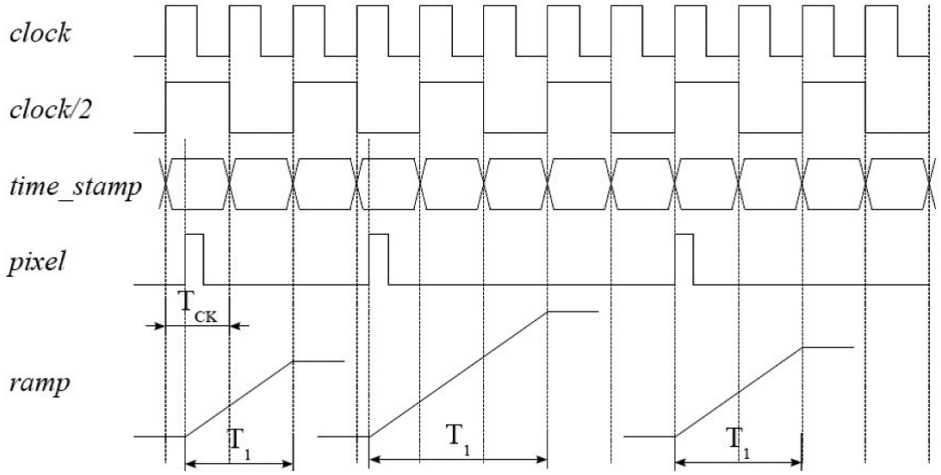

Figure 14 Generation of the analogue ramp for the TDC.

1.1.4.4.2 P-TDC Time-to-Digital-Converter

In each pixel, a TDC based on a Time to Amplitude Converter (TAC) generates a voltage proportional to the interval between the hit and one of the clock edges by charging a capacitor with a constant current source. The resulting voltage is then digitized, providing the required fine time measurement. The relevant waveforms are sketched in Figure 14. One must take into account potential misalignments between the coarse and the fine time logic. For this reason, the latter works at half the clock rate. With such an arrangement, the LSB of the coarse counter overlaps the MSB of the TDC, thus making digital error correction possible. Furthermore, the analogue ramp is stopped on the first rising edge following a falling edge. This introduces a time offset, which allows a proper settling of the circuits generating the ramp. The ADC employs a Wilkinson topology: the capacitor of the TAC is discharged by a small current and the clock pulses necessary to restore the baseline are counted. The time needed for the fine time measurement may reach up to 1 μs in the worst case. Since the maximum expected rate per pixel is 140 kHz, a multi-buffering scheme is mandatory and it has been implemented with two FIFOs, one analogue and one digital. The digital FIFO allows the storage of up to four values of the coarse counter, while the analogue one, formed by four capacitors, accommodates the corresponding fine time measurements.

The P-TDC approach allows an optimal use of the relatively large pixel area and a simplified design of the End of Column logic (EoC). This can be sized according to the average event rate since data are already derandomized in the pixel cell.

Figure 15 shows a simplified schematic of the TDC. It is formed by two PMOS current sources, four TAC cells (which implement the analogue FIFO), a capacitor and a synchronous comparator.

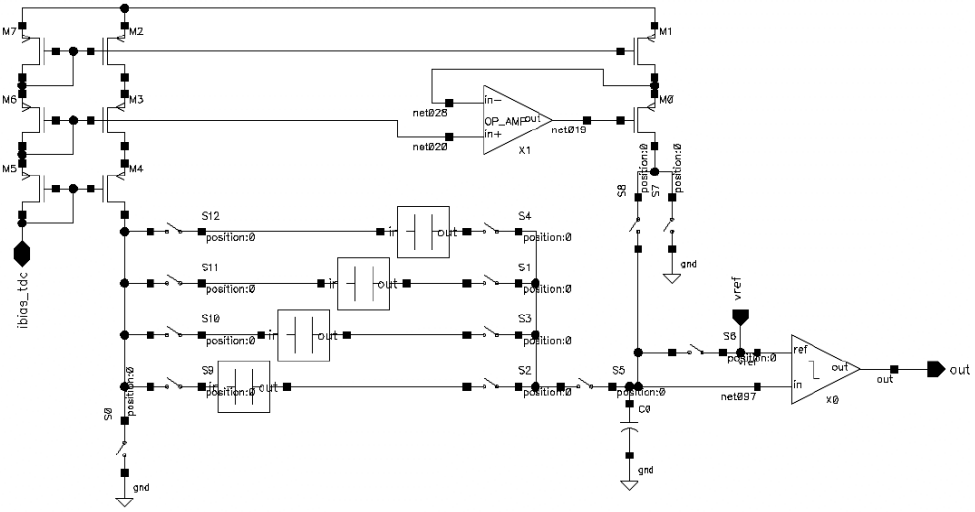

Figure 15 Simplified schematic of the TDC.

The TDC works as follows. When the CFD fires, the current source formed by M2-M4 is connected to one of the TAC cells, as shown in Figure 15. The capacitor is put in the feed-back path of a single-ended cascode amplifier. The purpose of this circuit is two-fold. First, it keeps the drain of M4 at virtual ground, thus avoiding a modulation of the current through the finite output conductance of the transistors that would introduce nonlinearity. Second, it provides an inversion of the signal, which is stored on capacitor C9. In other words, while the TAC capacitor is *charged*, C9 is *discharged*. Therefore, the baseline value at the input of the comparator can be restored with a current having the same polarity as the one delivered by M2-M4. Both current sources are hence implemented with the same type of transistors, which is very important because the matching between them is critical. Additionally, the value of C0 is four times the one of the TAC capacitors. In this way, to have a 7 bit resolution, the current provided by M1-M0 only needs to be 32 times the one given by M2-M4, resulting in a much more compact circuit Figure 16 shows the schematic of the TAC.

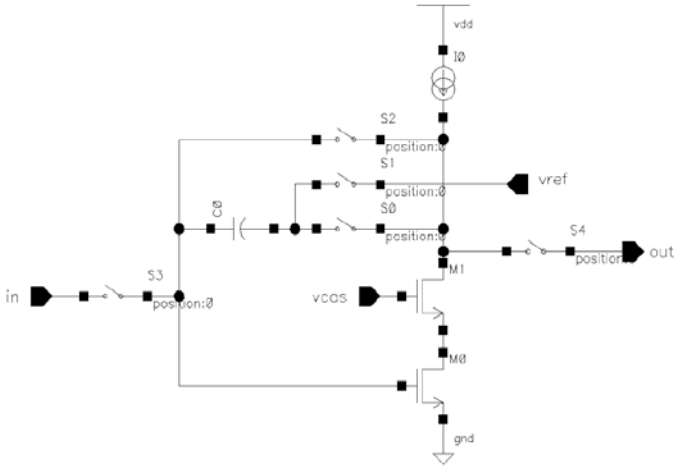

Figure 16 Schematic of the Time to Amplitude Converter. Four of these circuits form the analogue FIFO.

1.1.4.4.3 P-TDC Digital Section

The digital section of the pixel is composed by the digital FIFO that stores the coarse counter values, the logic controlling the TAC, the TDC and the interface to the End of Column readout. One of the possible drawbacks of the P-TDC option is that the many digital gates in the pixel are directly on the beam trajectory. The protection against Single Event Upset is therefore a primary concern and is addressed through the systematic use of Hamming encoding, as detailed in one of the next sections.

1.1.4.4.4 P-TDC - End of Column Logic (EoC)

The schematic of the column controller is shown in Figure 17. A state machine sends a read enable signal to the pixels, using a token ring mechanism. The read enable is then passed from one pixel to the next. When the read enable is high each pixel writes out the data content and its empty FIFO flag. A busy signal alerts the End of Column FSM if there are data available in at least one of the 45 pixels. In order to avoid ambiguities in the time of reconstruction the events belonging to the same cycle of the time stamp counter value are grouped together. Adding an extra bit to the time stamp does this rearrangement. The additional bit is used to tag events of type 0 and events of type 1 in two different FIFOs. A merger circuit puts the data frame in the right order, adding a header and a trailer. The header contains the information of the frame counter value. To separate the frames, the circuit uses the information provided by the pixel logic. Each pixel sets an “old-data flag” if the frame changes when it still contains data.

This flag is propagated through the whole column with a fast or. The depth of the FIFO has been chosen to be 32. VHDL simulations have shown that at the maximum nominal rate of 3.3 MHz per column, the maximum FIFO occupancy is 20. The layout area of the EoC is 1500 μm x 300 μm.

1.1.4.4.5 Single Event Upset Protection

Given the high particle flux, protection against Single Event Upset (SEU) is a serious issue for the GTK front-end electronics. In the P-TDC option the following solutions have been adopted:

- Pixel cell: Hamming encoded state machines and registers with auto-correction (single error correction and double error detection).
- End of Column: Hamming encoded state machines with auto-correction (single error correction and double error detection).

Concerning the EoC FIFOs the following considerations have been made. The particle flux in the EoC area is 1.3 10^{6} /cm^{2} per second assuming that in the EoC area the flux is 1 % of the peak flux. The SEU cross section in silicon is 10^{-8} cm^{2}/bit (4) and the collision probability is: $P(x) = e^{-x/L} = 3.3\ 10^{-6}$, where x is the thickness of the active silicon containing transistors (1 μm) and L is the mean free path in silicon (30 cm). The flux of particles giving SEU is therefore 1.3 10^{6} x 3.3 10^{-6} = 4.3 particles per square centimetre per second. One FIFO of 32 words with 32 bits per word has ~1200 flip-flops including control logic and hamming code protection. The number of SEU per second is then 1200 x 10^{-8} x 4.4=5.14 10^{-5}. The data stay in the FIFOs for at most 12 ms, so the upset probability is 6 10^{-10} for single error and 10^{-19} for double errors. For these reasons the FIFOs have a Hamming encoder at the input and a Hamming decoder with single error correction and double errors detection at the output registers only and are protected by triple mode redundancy (TMR).

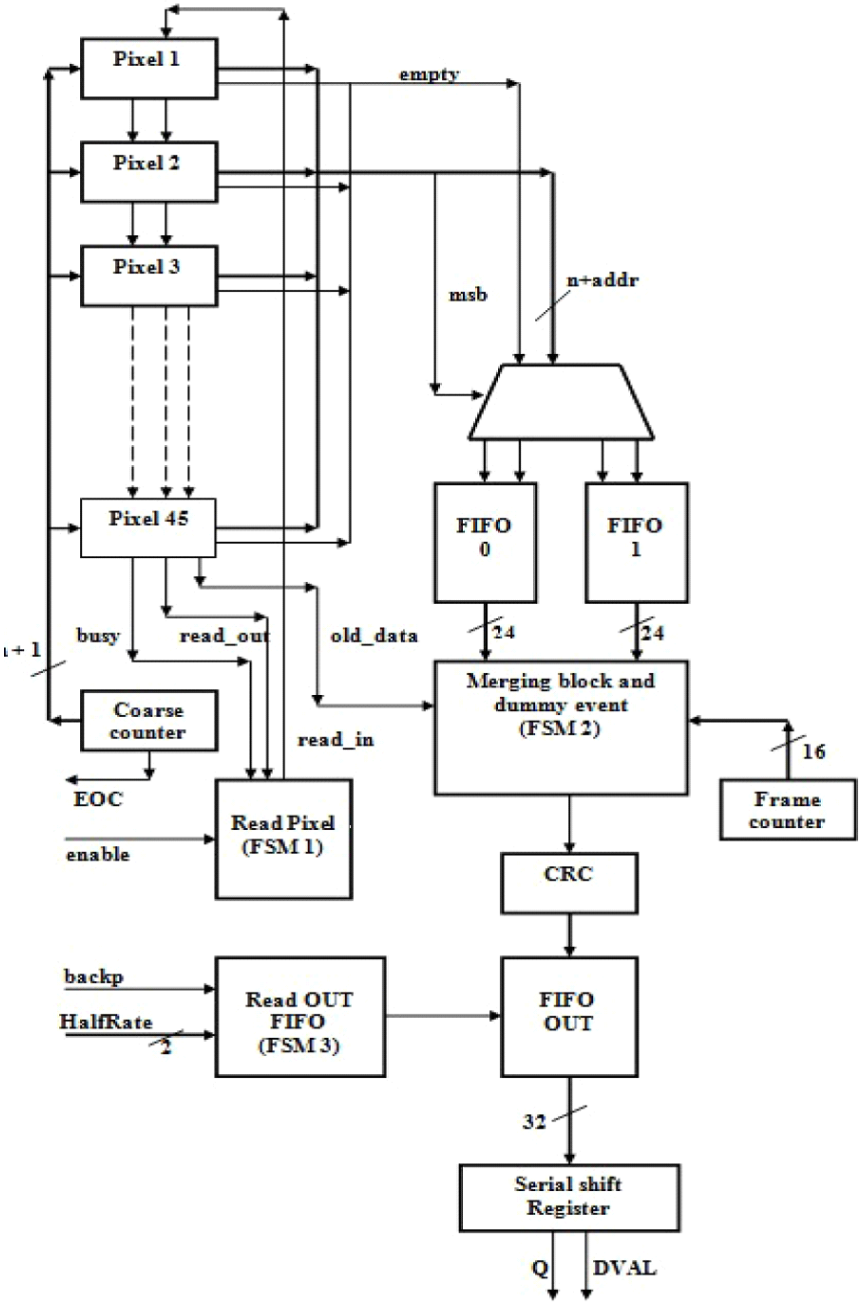

Figure 17 Block scheme of the End of Column logic.

1.1.4.4.6 Demonstrator for the P-TDC Option

Prototype implementation.

The key goal of the design was to have a chip as complete as possible, so that all the critical issues associated to this option could be evaluated and a future extension to a larger, final ASIC could be done with a minimum of risk. Therefore, all the blocks contained in the pixel cell (front-end

amplifier, constant fraction discriminator, DAC for threshold setting and full in-pixel logic) are implemented in their final form. The only circuit, which is missing in the pixel cell is a “watch-dog” to flag hits that deliver a very high amount of charge in the sensor. However, due to the moderate resolution required (2-3 bits) the implementation of such a circuit does not entail critical issues. The design of the EoC logic is final and includes the SEU protection strategies discussed in the previous section.

The digital circuitry that merges the data coming from different columns and generates the serial output stream has not been included in this version. The prototype ASIC contains two long columns with the final number of pixels (45) and one short column with 15 pixels. Two spare pixels are provided for debugging purposes. In order to have an acceptable form factor, the long columns have been folded in three segments. One of the critical aspects of the P-TDC option is the quality of the clock distributed to the pixels. To assess this point the clock and the digital signals are propagated to the long columns with folded lines, which have a length slightly higher than the one expected for the final layout (14 mm). Figure 18 shows a picture of the prototype ASIC.

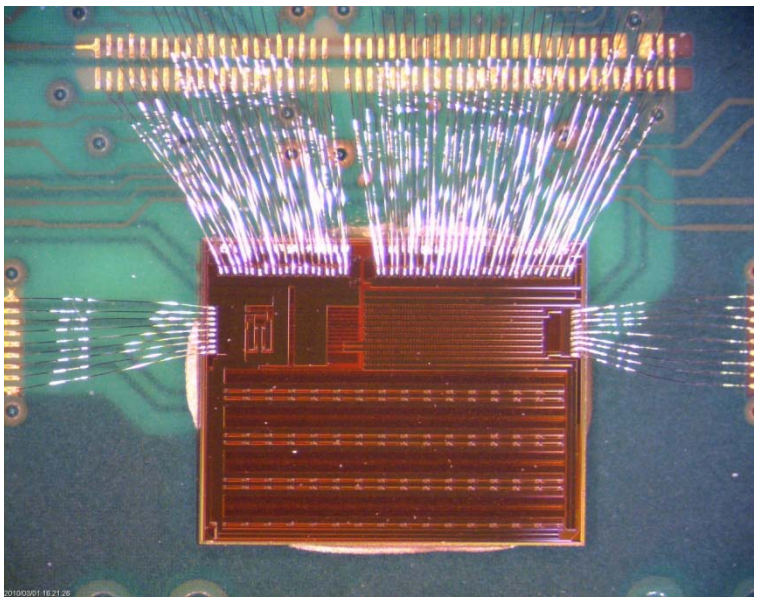

Figure 18 Picture of the p-TDC ASIC prototype wire bonded to the test printed circuit board

Test set-up

The bare die is wire bonded to a custom designed PCB. A combined data-pattern generator/logic state analyzer unit is used as the main testing instrument to provide the test vectors and read-back the chip response. A computer running routines in lab-view controls the test equipment and is employed for data analysis.

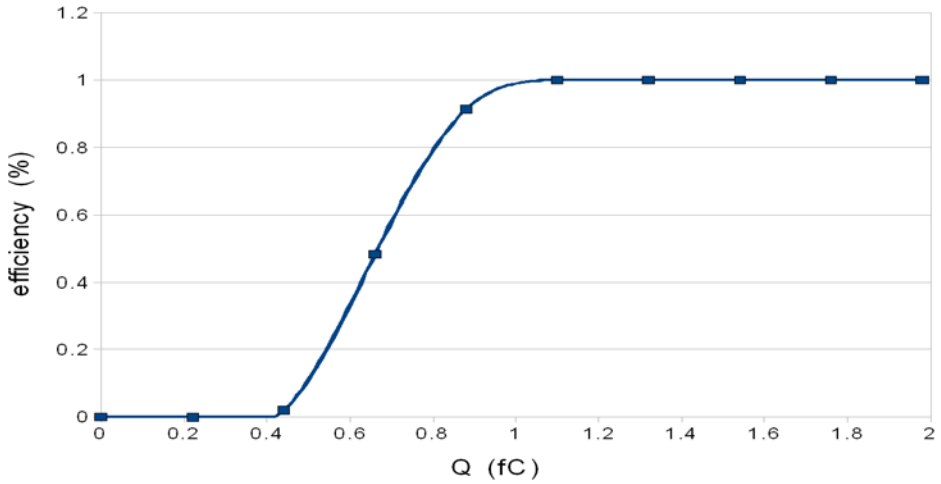

Figure 19 Efficiency curve for one pixel.

Test results

For testing purposes, a calibration signal can be injected through a 22 fF nominal capacitor. One leg of the capacitor is connected to the preamplifier input. An external trigger signal switches the other leg between the preamplifier virtual ground and a voltage level controlled by an external DAC. The difference between these voltages multiplied by the capacitor value gives the amount of injected charge. The switching leg of the capacitor is also connected to a low-pass filter. The purpose of this circuit is to smooth the input signal, providing a shape as similar as possible to the one that will be given by a real detector. Figure 19shows a typical S-curve obtained for one pixel.

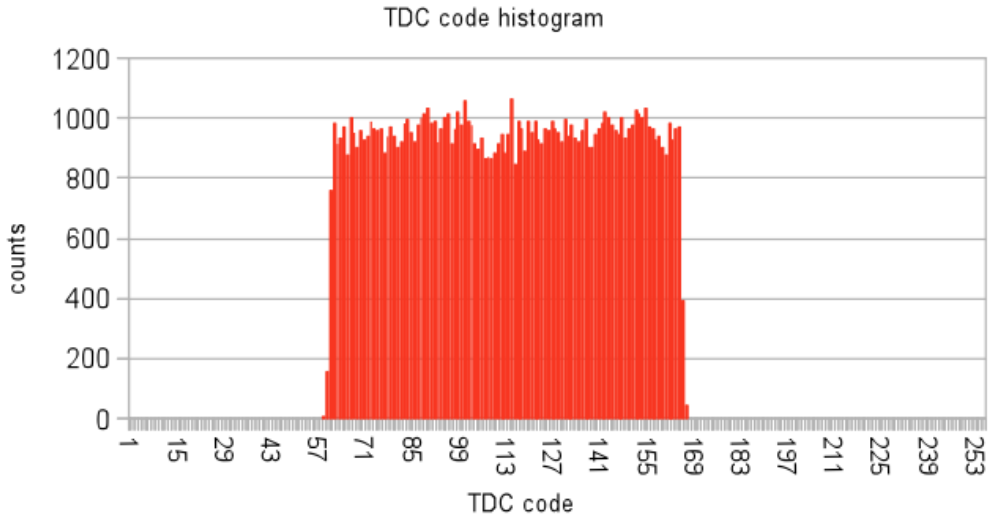

Figure 20 Example of raw code distribution used in the calculation of TDC DNL and INL from a data sample of 100.000 events.

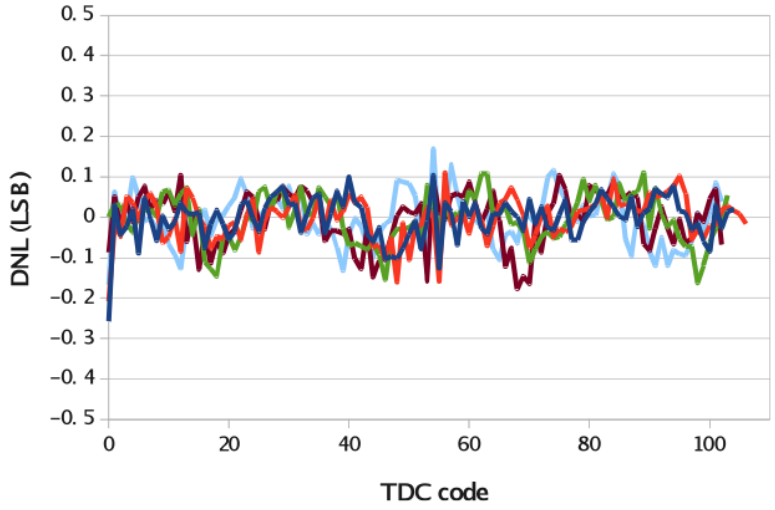

Figure 21 Differential non-linearity of five different TDCs from the matrix.events.

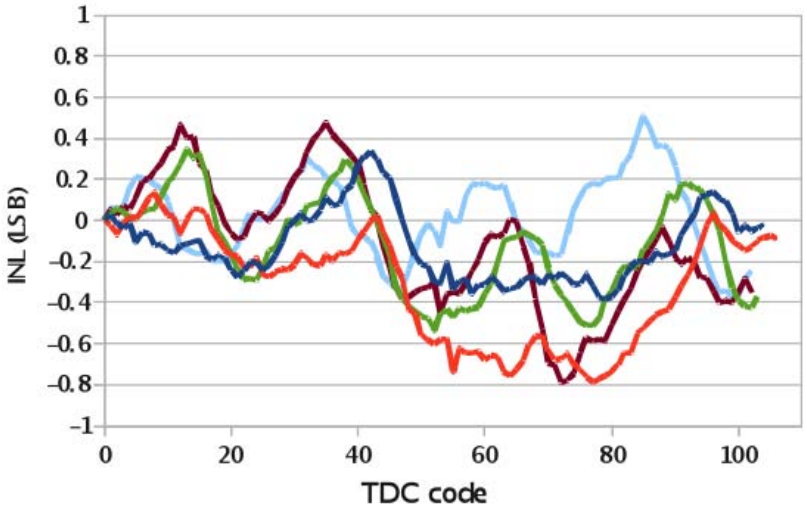

Figure 22 Integral non-linearity of five different TDCs from the matrix.

The TDC can be tested independently from the front-end part. In fact, it is possible to trigger the TDC either by an external digital pulse or by an internally generated one. The latter can have a length either of one or three clock cycles. In this way it is possible to perform code density tests (with the external pulse) and to calibrate the full-scale range of the TDC (with the internal one). Due to an undersized buffer in the EoC logic the maximum working frequency is at present 128 MHz. Therefore, the expected time bin is 122 ps instead of the design value of 97 ps. The experimental average measured value is 140 ps, with the extra 18 ps coming from a slightly higher value of the discharged current due to a systematic mismatch in the current source. While both effects that lead to the increased bin sized are well understood and can be easily fixed in a final version, a bin value of 140 ps gives a quantization noise of 40 ps rms, so its impact on the overall time resolution would be still acceptable. Figure 25 shows a raw data plot used to calculate the DNL and the INL of the TDC. This plot is obtained by feeding to the TDC an external test pulse with a frequency uncorrelated with the one of the clock. Figure 26 and Figure 27 report the DNL and INL profile, respectively. For clarity, the results of five different TDCs randomly selected from the three different columns have been superimposed in these plots.

Figure 28 shows a measurement of the residual time walk of the CFD taken at different clock frequencies (blue squares=no clock, yellow triangles=160 MHz, red circles=320 MHz). Figure 29 shows the jitter measured in the range 1.5-10 fC. The measurement was done at the clock frequency of 128 MHz.

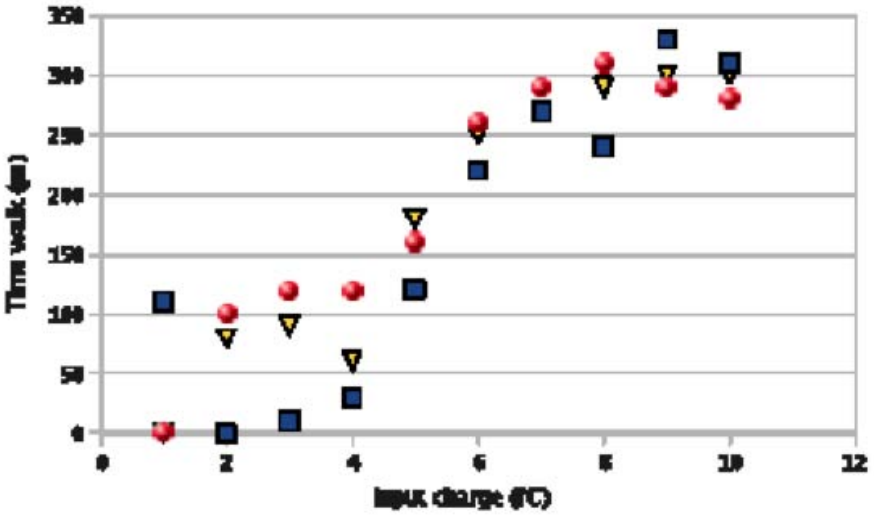

Figure 23 Residual CFD time walk at different clock frequencies.

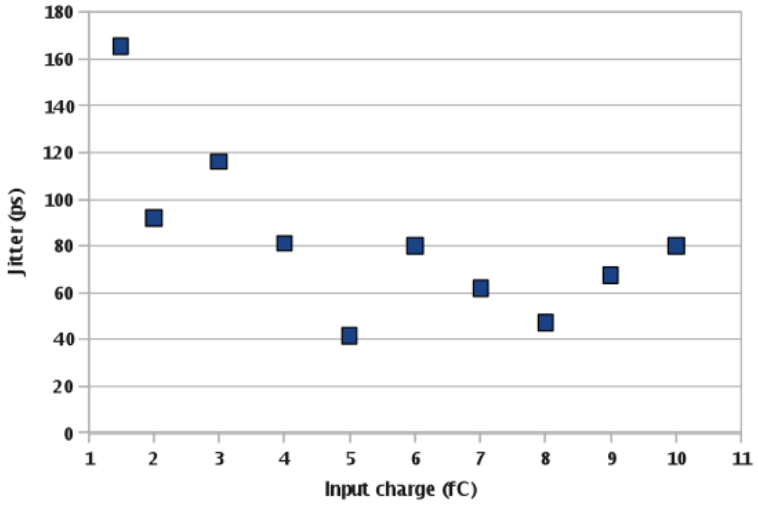

Figure 24 System jitter at 128 MHz clock.

The accuracy of the present test results is limited by the statistic, which can be achieved in a reasonable measuring time with the set-up. At the time of writing a new system based on a state-of-the art FPGA in under commissioning.

The most critical point of the P-TDC option is the risk of coupling digital noise to the amplifier inputs, since a lot of logic is resident in the pixel. This risk may be enhanced by the presence of the detector capacitance, so a full and realistic assessment of the performance of this option can be done only when the prototypes bump bonded to a sensor will be available.

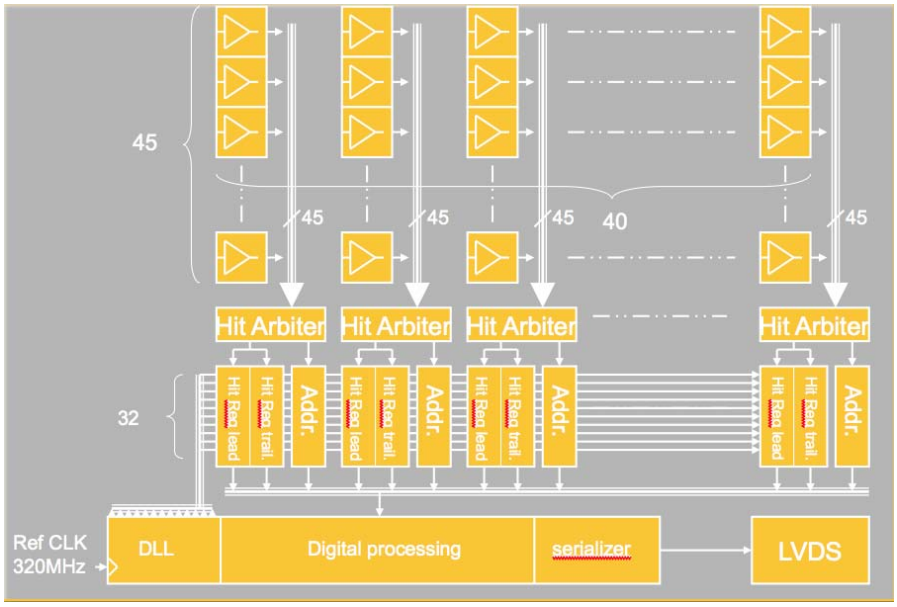

Figure 25 EOC architecture block diagram.

1.1.4.5 End-of-Column (EOC) Architecture Overview

The time walk compensation is based on time over threshold correction (ToT) measurements. Thus the rising and the falling edges of the discriminator signal are time tagged. The leading edge time stamp provides information of the hit arrival time and the trailing edge provides input charge amplitude information used to correct time walk. At the time of the leading edge, a synchronous coarse clock counter value is latched to increase the dynamic range of the TDC to 12.8 µs. The double time stamp with coarse counter and address information are stored in a buffer, before they are serialized and sent off the chip.

Figure 25 shows a block diagram of the final EOC chip architecture (5)(6). The end-of-column (EOC) option employs simple pixel cells containing the amplifier, time-over-threshold (ToT) discriminator and a transmission line driver per pixel cell, which sends the discriminated signal to the end-of-column region (EOC) of the chip. There a delay locked loop (DLL) based TDC (7)(8) time tags the rising and falling edge of the discriminator signal with a time bin of 100 ps. In the chip, 5 not-neighbouring pixels in a column are multiplexed together and connected to one TDC using a combinatorial hitArbiter circuit. Simulation shows that the hit efficiency for the worst-case centre column stays above 99.5 %. Pixel hit addresses and rising and falling time tag words are sent to a buffer FIFO before the data are serialized and sent out of the chip. The time walk compensation using the two time tags is performed off-line allowing the evaluation of the input signal charge for each hit. In this architecture no clock signals are distributed over the clock matrix, reducing possible interference with the analogue electronics and data are only transmitted to the EOC circuits once a hit has been registered. The digital processing and transmission units are geographically separated from the analogue front-end electronics.

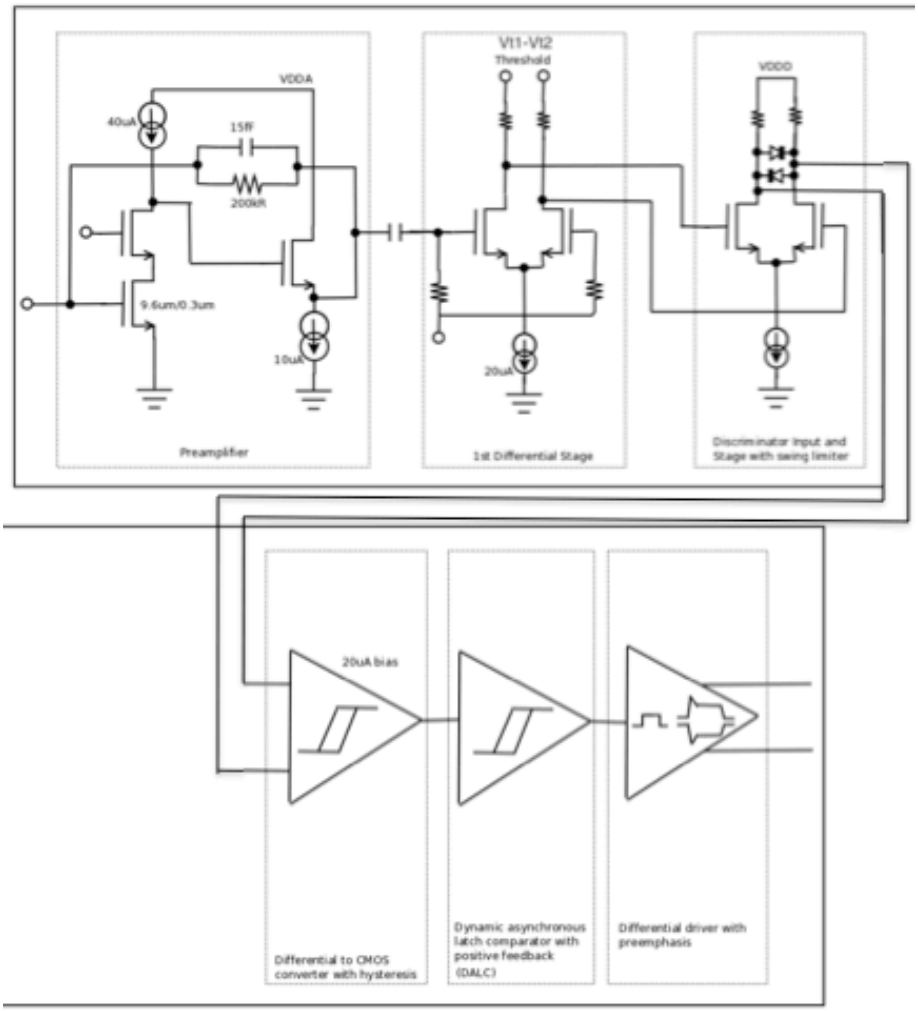

Figure 26 Block diagram of EOC pixel cell.

1.1.4.5.1 EOC Analogue Pixel Frontend

The pixel cell contains 7 functional blocks:

- a transimpedance preamplifier;
- a differential post amplifier;
- a first stage of the discriminator;
- a second stage of the discriminator with hysteresis;
- a dynamic asynchronous latch comparator with transitional positive feedback;
- a differential transmission line current driver with pre-emphasis; and
- a coplanar transmission line.

These functional blocks are shown in Figure 26 with the exception of the transmission lines. The preamplifier uses a cascode amplifier with an NMOS input transistor biased at 40 µA of dimensions 9.6 µm width and 300 nm length with a simple resistive (200 kΩ) and capacitive feedback defining the pulse gain at a level of 30 mV/fC. The dimensions of the input transistor were optimized for the detector capacitance of 250 fF. The low value of the input capacitance allows the use of a simple cascode stage offering very high bandwidth (1 GHz gain bandwidth product) at very affordable power consumption (60 µW). The open loop gain is in the range of 45 dB, the input impedance of the preamplifier stays in the range of 1 to 2 kΩ for 1 GHz bandwidth, which is sufficiently low to provide efficient charge collection from the detector. Crosstalk between neighbouring pixels is expected to be less than 4 %.

The first differential pair applies the external differential threshold voltage $V_{T1} - V_{T2}$ to the input signal. The fully differential structure of the comparator provides very good rejection of common mode noise from the digital power supply and good threshold uniformity. The entire preamplifier-shaper circuit has a gain of 70 mV/fC with CR-RC3 characterics and a peaking time of around 5.5 ns.

The following three stages of the comparator provide high sensitivity together with very high speed at a power of 50 µW. The simulated time jitter (transient noise simulation) of the full front end stimulated with 3 fC signal is in the order of 30 ps RMS. The predicted ENC values are shown in Figure 27. These simulation results show that for a detector capacitance of 250 fF the predicted ENC is below 200 e^- .

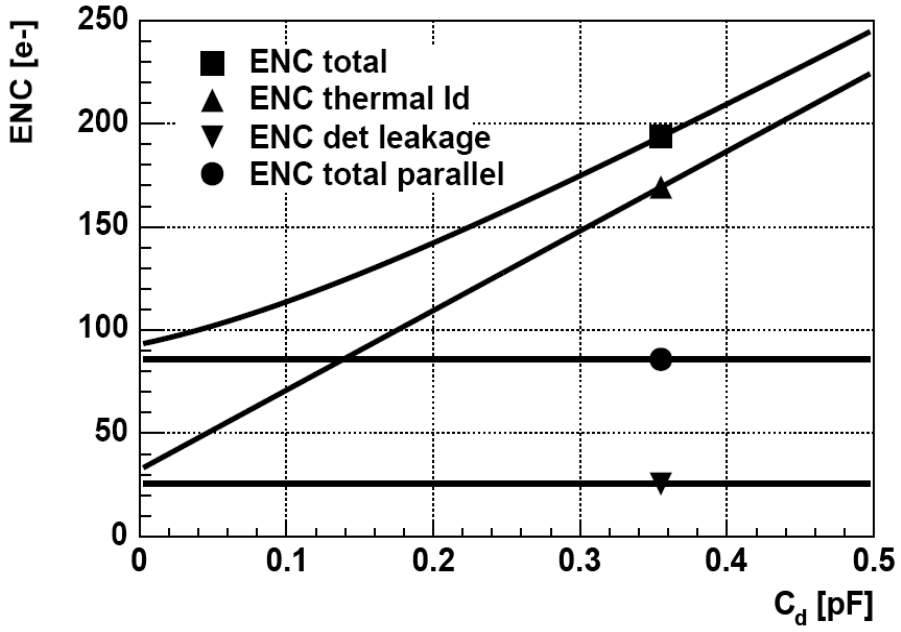

Figure 27 Noise analysis for 0.13 µm CMOS technology.

Figure 28 shows simulations of the time walk and Figure 29 of the time over threshold pulse width with a detector capacitance value of ~ 250 fF, giving a time walk of ~2.1 ns for an injected charge of 5.25 fC. The ToT discriminator pulse width simulation gives values from ~8.45 ns for an input injected charge of 1 fC to ~14.4 ns for an injected charge of 5 fC.

The input of each pre-amplifier is connected to the sensor bump bond pad and a calibration capacitance of 20 fF. The other terminal of the capacitance is connected to a calibration line accessible outside the chip. A voltage step on that line injects charge into the preamplifier. The rise time of this voltage step controls the lengths of the charge injection duration.

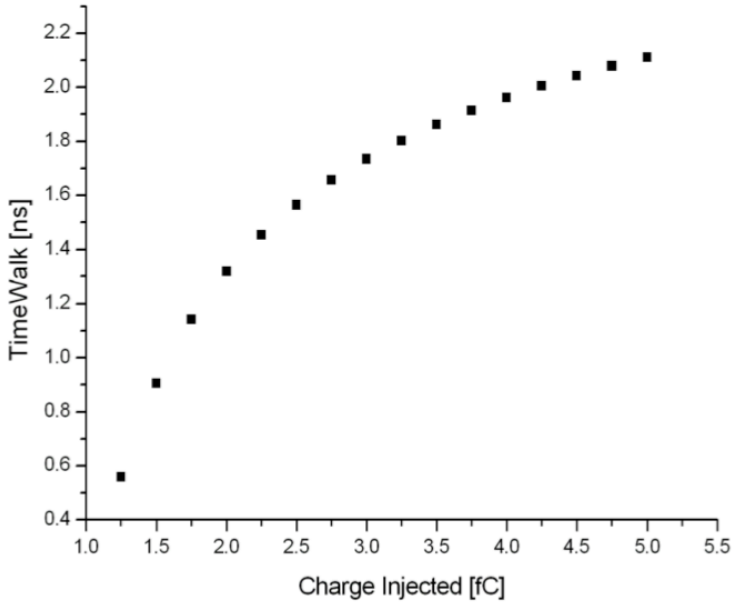

Figure 28 Simulated time walk for a threshold of 0.7 fC, reference is 1 fC. Reference where time walk is 0 if for 1 fC injected charge.

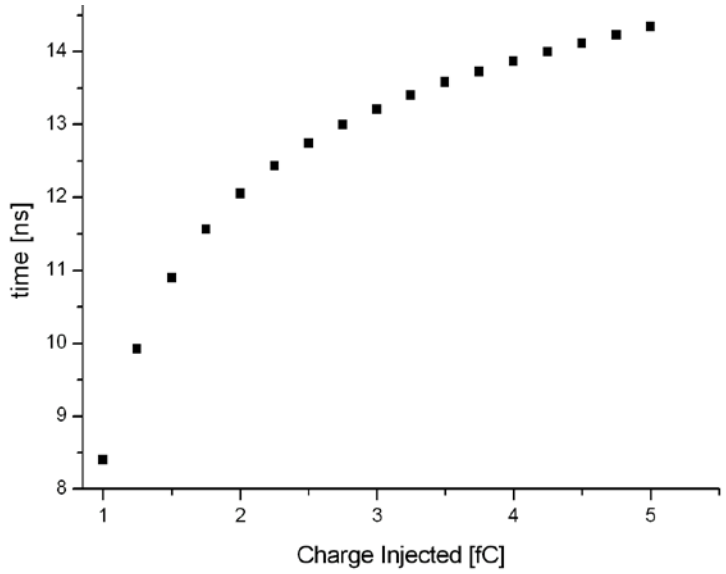

Figure 29 Simulated time over threshold pulse width of the ToT discriminator for a threshold of 0.7 fC.

1.1.4.5.2 EOC Bus System

During the transmission of the discriminator signal to the EOC region the steep rising and falling edges must be preserved to keep jitter values low. A current driver in the pixel cell is connected to differential transmission lines. The traditional use of CMOS inverters as repeater stages is avoided because they introduce delays that scale linearly with the wire capacitance and hence the wire length. Also the big CMOS swing limits the transmission speed. The use of a transmission line limits the resistive energy losses, where the high frequency components of the signal travel more quickly than the low frequency (9)(10).

Each pixel discriminator is connected to a differential driver, which switches a current source with a value of 100 µA when a hit occurs. A second DC current source equalizes the DC current during steady state time. The current signal swing defining the logic level in coplanar waveguides has been specified to ±100 µA, with a differential swing of 240 mV. The series resistance of the line varies for pixels close to the EOC and pixels at the other side of the matrix (distance=13.5 mm). However, this effect does not disturb signal discrimination because the far end amplitude does not change with the pixel to receiver distance thanks to the current drive operation. The high series resistance of lossy transmission lines in the CMOS process has an impact to the signal integrity of the bus. It degrades signal edge in slowing down rise time signal after 13.5 mm signal propagation. To circumvent this effect a current pre-emphasis is implemented in the line driver. Injecting a fast current pulse on the edges of the hit pulses does the signal pre-emphasis, achieving a rise time of 1 ns for the farthest pixels.

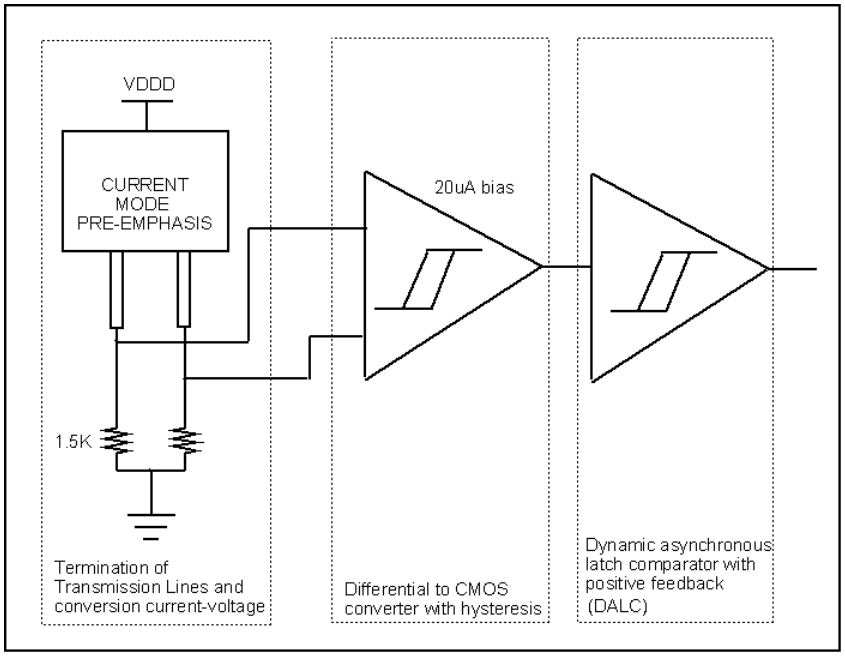

Figure 30 Schematic of the transmission line receiver.

The transmission line receiver consists of 2 different parts. The first one is a differential to CMOS converter with hysteresis and a 20 µA bias current. The second part is a dynamic asynchronous latch comparator with positive feedback (DALC). The DALC generates the fast transition signals without consuming static power by using positive feedback during transition to increase transition speed. A circuit that senses the differential current with a differential common gate cascode stage has been

applied. The block diagram of the receiver is presented in Figure 30. The receiver input stage biasing is provided directly by the static current of the transmission line of about 150 µA. The output of the receiver input stage is sensed with a broadband differential to single ended amplifier stage whose output is converted into a fast digital CMOS level signal by a DALC stage which generates pulses edges of 50 ps to drive the TDC inputs.

1.1.4.5.3 End of Column Logic

In the EOC architecture the arrival time is measured using delay locked loop (DLL) based TDCs, where buffers connected in series are used as base elements. The TDC circuits used in the end of column logic are based on previous developments done in 0.25 µm CMOS technology (7) and in 0.13 µm CMOS technology (8). The propagation speed of these buffers is controlled by a voltage. The base frequency is sent to the first buffer and the phase relationship between the input of the first buffer and the output of the last buffer is monitored. If the phase difference is positive (the output is too slow) the control circuit of the DLL will increase the delay cell control voltage and thus the speed of the buffers and inversely decrease it in case of a negative phase difference. The phase detector continues to measure the phase of the input and output clock signals throughout the operation of the system continuously correcting the speed of the buffers. Then the individual buffers divide the base clock period by the number of buffers in the chain. Upon arrival of the hit the hit registers capture the state of these buffers and the state of a synchronous base frequency clock counter. Figure 31 shows the basic principle. For the EOC architecture a base frequency of 320 MHz is applied and 32 buffers divide the period of 3.125 ns in bins of 98 ps.

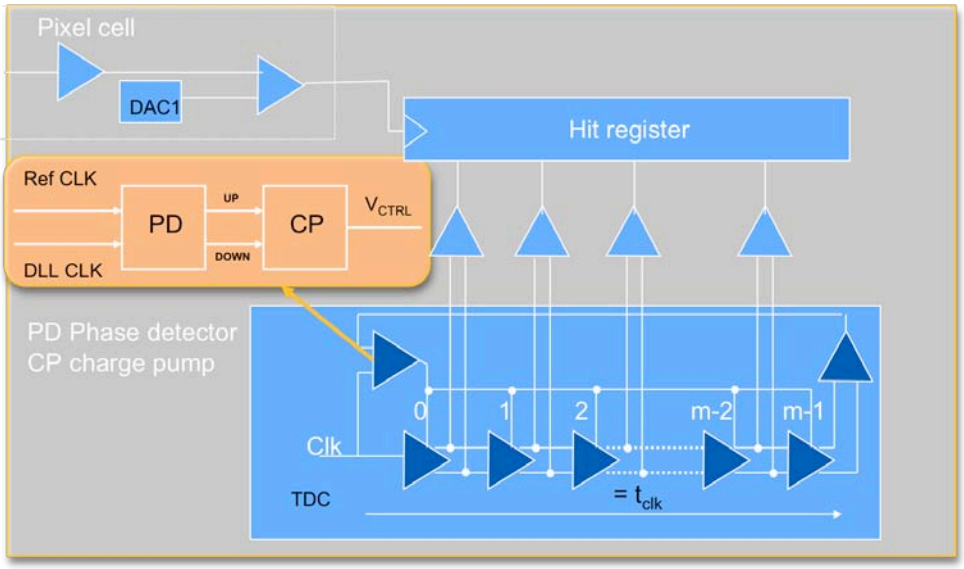

Figure 31 DLL based TDC principle.

It is not required that each pixel cell has its own hit register. An efficiency for the highest rate centre column of 99.5 % can be achieved by multiplexing 5 pixels to one hit register. A digital arbiter circuitry has been designed which sends hit pulses to the TDC banks and masks the rare simultaneous signals from the same group of 5 pixels. These overlaps are indicated with the address of the masked pixel in the output data stream. In order to process clusters and pixel charge sharing, the pixel group feeding each hit arbiter is formed from pixels that are not adjacent, see Figure 32.

The 45 pixels of each column are grouped in 9 sets of 5. Correspondingly 9 rising edge hit registers and 9 falling edge hit registers are used for each column.

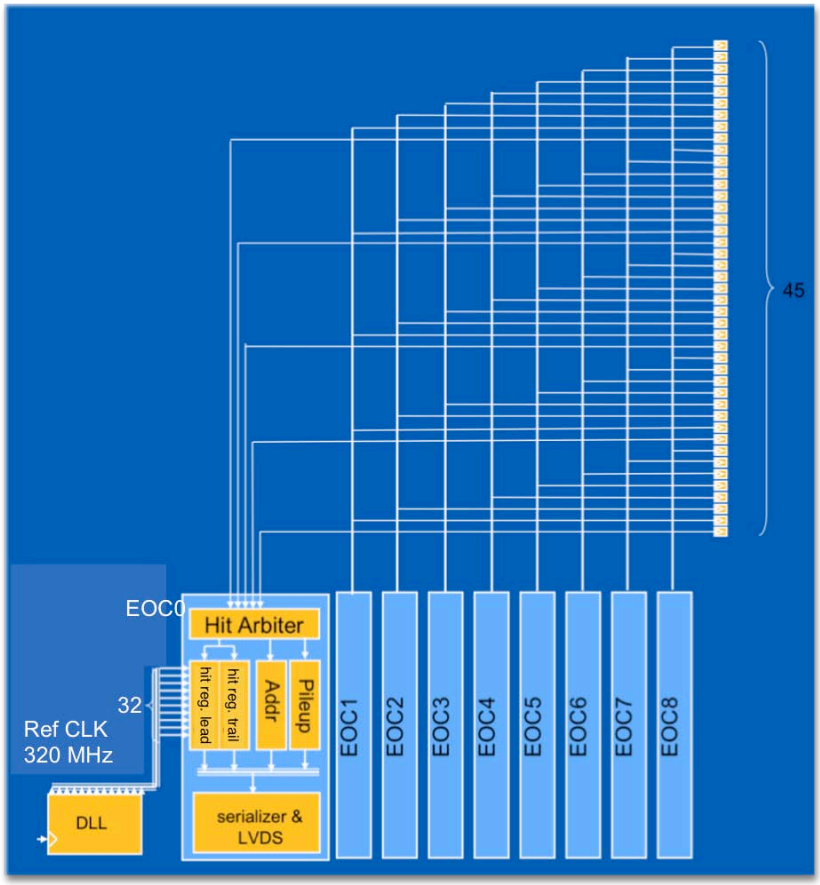

Figure 32 Block diagram of one column with transmission line scheme.

The GTK read-out chip will contain a number of DLLs and the buffered DLL signals are distributed to the hit registers. The output of the DLL buffer is differential and is converted to single-ended CMOS at the input buffer of each TDC bank. The hit registers in the TDC are built using 32 D-type flip-flops with rising edge trigger. The output of the receiver cell provides a rising edge trigger for both leading and trailing edge hit registers.

The sensitive analogue circuits are in the pixels and the digital power supply domain is geographically well separated from the analogue supply, reducing noise influence on the pixel cells. EOC circuits have been designed so that the receiver bank and the hit register bank for a full column and the digital column read-out circuits fit inside the 300 µm wide envelope dictated by the pixel size.

Figure 33 shows the block diagram for the full chip with the EOC architecture.

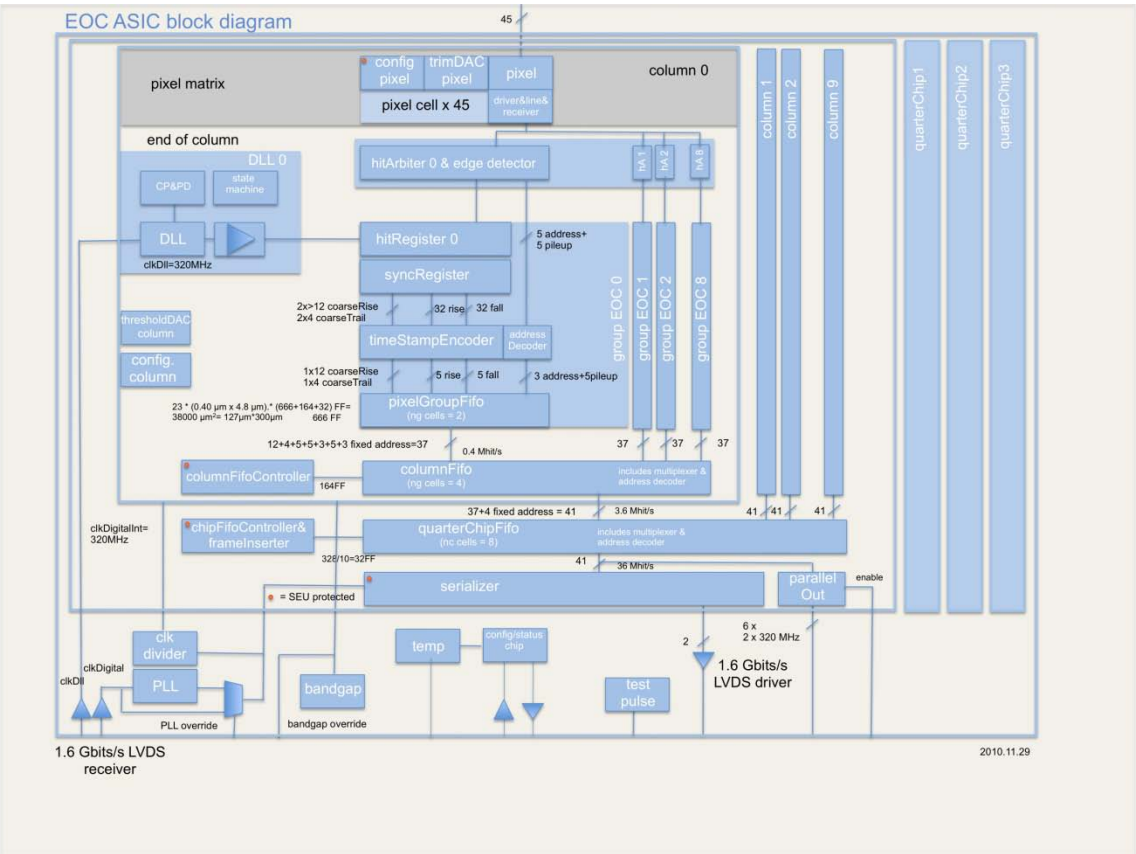

Figure 33 Full chip EOC block diagram.

1.1.4.5.4 EOC Demonstrator

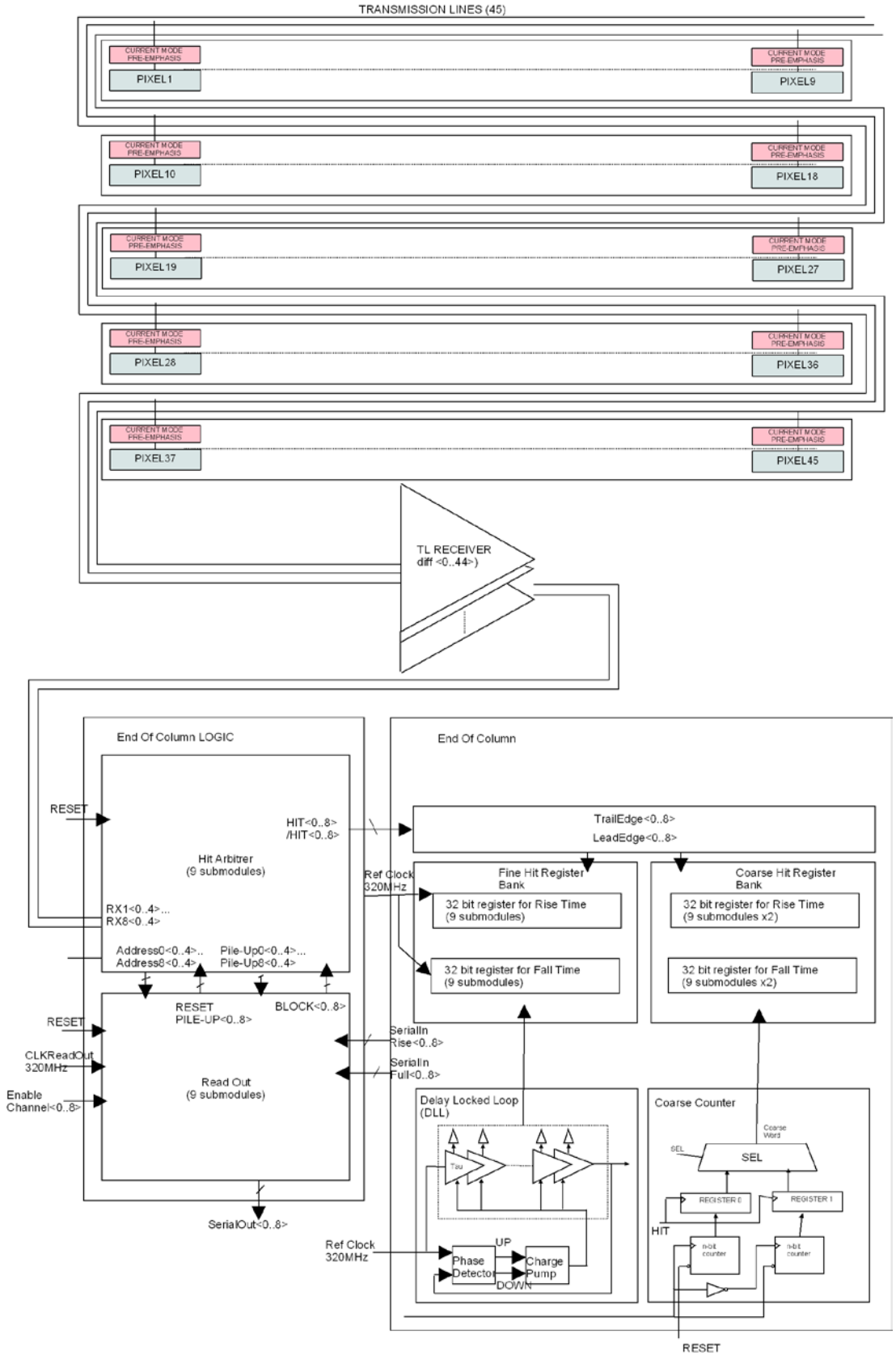

Figure 34 Block diagram of the 45 pixel demonstrator.

A demonstrator ASIC for the EOC column architecture with the ToT time walk compensation technique has been developed. Figure 34 shows a block diagram of the demonstrator. The demonstrator contains one full 45 pixel column folded back on itself to save silicon area. In addition to this column, test structures with 15 pixels have been included for the characterization of the analogue pixel cell. This demonstrator serves experimentally as proof of principle for the analogue front-end, the ToT time walk compensation technique, the low-jitter transmission of the signal along the column and the time tagging of the rising and falling edge maintaining a time walk compensated arrival time measurement resolution of less than 200 ps. The homogeneity of the pixel matrix will be evaluated. The entire chain from analogue pixel input, via the amplification stages and discriminator to the transmission line amplifier, transmission line and receiver, the hit arbiter logic and the TDC has been fully realized as for the final chip. The layout of the EOC circuits largely follows the 300 μm pitch defined by the width of the column. The data encoding and read-out from the TDC has been simplified as the implementation is straight forward.

The demonstrator is organized in such a way that the 45 pixel column feeds a set of 9 hit arbiters and 18 hit registers connected to one DLL. Rising and falling edge hit registers of the 9 pixel groups are read out using 9 x 320 MHz serializers independently. In order to asses the influence of digital noise on the DLL and analogue parts, the read-out clock can be switched on and off for each of the 9 streams independently. A set of 9 test pixels is connected to a second instance of the EOC column logic and operates independently. Test outputs along the processing chain allow qualification of individual blocks. Figure 35 shows the layout of the prototype.

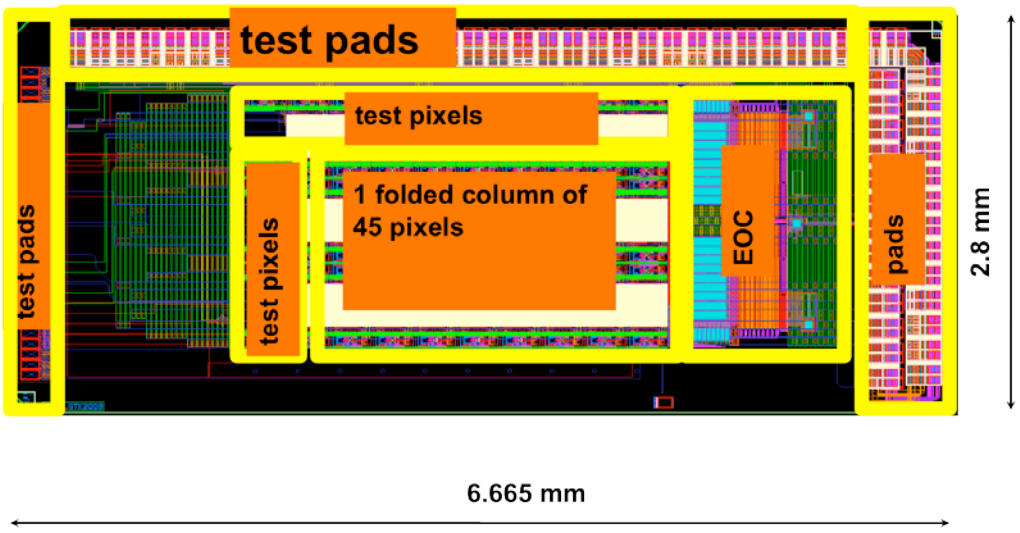

Figure 35 EOC column demonstrator layout.

The qualification of the demonstrator includes the testing of each individual block, the full processing chain, the full 45 pixel column and its uniformity using programmable test pulses on each individual pixel.

After the bonding of the demonstrator chip to the sensor, laser and beam tests were made to qualify the assembly of sensor and read-out chip (see section 1.1.5 on page 53).

Test setup

A test system was built up around the EOC GTK demonstrator ASIC. Figure 36 shows a schematic of the test bench containing the most important components. This system is based on an Altera Stratix III FPGA development card. A daughter card was designed to host the EOC ASIC and interface to the development card. This services card provides power, current and voltage references, analogue and digital signal buffering and SMA/SMC connectors to bring in and out test signals. Control of references is done through a single I2C bus upon which there are 8 voltage output DACs, each one providing a signal between 0 V and 1.25 V to 12 bit precision. Where necessary, current source and sink conversion is done using transistors. An EEPROM provides non-volatile digitally accessible identification of the card. A temperature sensor gives a local and potentially remote measurement of the temperature to about 1^{o}C precision. The test setup includes a Tektronix AFG3252 pulse generator and a LeCroy WaveRunner or WavePro oscilloscope. The test system uses a private gigabit Ethernet network to connect the various components together. A C++ software framework provides control and configuration of all the components of the setup, allowing automated sweeps to be performed in a detailed manner over all pertinent parameters. The collection and handling of the data output from the setup is also dealt with in the framework using a run oriented approach, in which a unique run number is generated for each instance of a measurement. This permits a given measurement to be repeated as many times as desired without overwriting any pre-existing data. The high precision DLL clock is provided by the clock generator (SRS CG635), whereas the readout clock comes directly from the FPGA

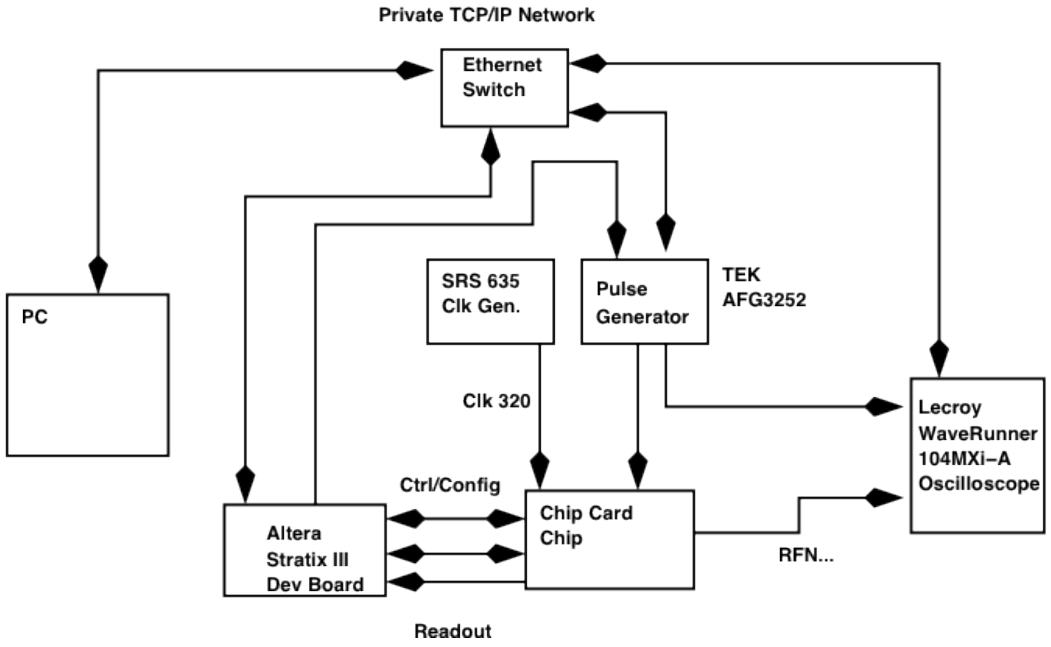

Figure 36 A Schematic view of the automatic measurement test setup.

Test results

Test principle:

All tests have been done with the nominal 320 MHz read-out and TDC reference clock frequency, corresponding to a TDC time bin of 97.7 ps.

Figure 37 shows the definition of particle arrival time t_0, the leading edge discriminator output time t_1, the trailing edge discriminator output time t_2 and the time-over-threshold pulse width t_{tot}.

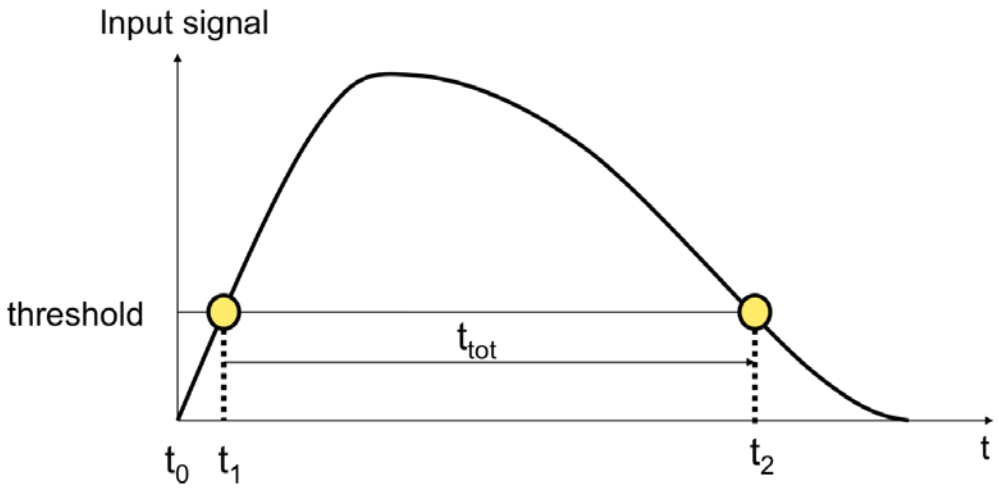

Figure 37 Definition of signal times: t_0 particle arrival time, t_1 leading edge discriminator output time, t_2 trailing edge discriminator output time and t_{tot} time-over-threshold pulse width.

The principle behind the testing is to present charge in varying amounts at the input of the front end amplifier. In the absence of a detector this is achieved by the inclusion of a capacitor in the pixel, one terminal of which is connected to the front end amplifier, the other being connected via an isolation circuit to a global, low resistivity, calibration net. The capacitance of this capacitor is expected to be approximately 20 fF, with an uncertainty of about 10 % on a given chip and a channel to channel spread of about 1-2 %. When a given pixel is selected, a voltage step present on the global calibration net causes a charge to be injected at the front end of that pixel. The orientation of the voltage step decides the polarity of the charge, and the rise time determines the duration of the charge pulse; the current being the time derivative of the voltage. Since the expected detector technology is p-in-n, the results presented are acquired with a positive going voltage step. The duration of the test pulse (rise time of injected charge) is varied from 2.5 ns to 6.5 ns in order to emulate the expected elongation of the current pulse from the detector during its lifetime resulting from the accumulated radiation damage. The total amount of charge injected is varied over the expected range of 1-10 fC. This requires a voltage step on the calibration line of between 50 mV and 500 mV. Once the charge injection pulse is complete, the signal level is returned to the starting value over a period of time of greater than 1 microsecond. This is long compared to the front end amplifier peaking time, and as such, avoids the opposite polarity pulse occurring at the output of this amplifier. This behaviour is what would be expected with the detector. The repetition rate of this pulse can be controlled via the test setup software. It may be run continuously, triggered by the chip controller card, or by software. Additional synchronisation with the DLL clock is possible to the level of around 15 ps (RMS).

In the following when referring to full-chain measurements an electrical test pulse has been injected into the pre-amplifier and the digital time stamp of the TDC bank is evaluated, including all components along the chain – pre-amplifier, ToT discriminator, transmission line driver, transmission line, transmission line receiver, TDC and hit registers. Analogue front-end measurements evaluated the pre-amplifier and ToT discriminator, whereas TDC measurements are done by injecting digital

input pulses directly to the TDC bank and reading the digital time stamp. Figure 38 shows graphically in the block of the EOC processing chain and the scope of the measurements performed.

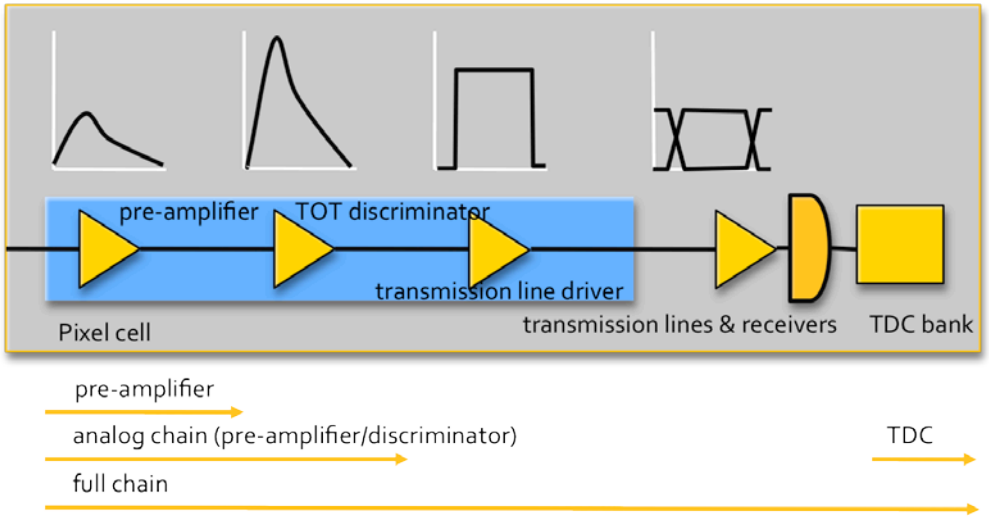

Figure 38 Block diagram of EOC processing chain.

Analogue frontend behavior

Unless stated otherwise, the charge injection duration is 2.5 ns and the read-out clock is activated. All measurements concerning the analogue frontend behaviour in this section are taken from the discriminated test pixel output.

- Analogue frontend time walk and jitter over input charge

Figure 39 shows the analogue front-end t_1 and t_2 responses over input charge and several discriminator thresholds. For a threshold of 0.7 fC the full input charge range time walk of t_1 is 2.2 ns and the ToT-pulse width is 7 ns. From this plot we also can see that the relationship between t_1 and time over threshold (t_2-t_1) is monotonic which required for a time walk correction.

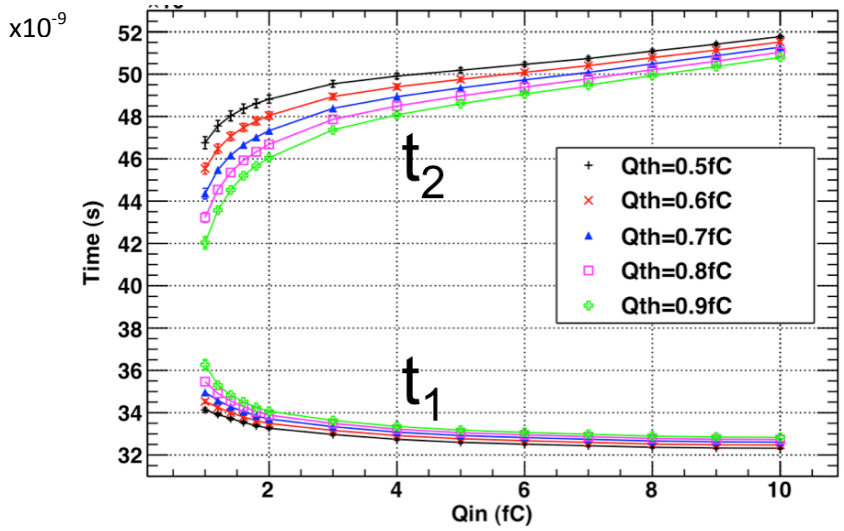

Figure 39 Analogue front end response for t_1 and t_2 over input charge for 2.5 ns input charge charge injection duration.

Figure 40 and Figure 41 show the analogue front-end t_1 and t_2 rms jitter respectively over input charge for several threshold values and input signal charge injection duration of 2.5 ns. These data are taken with 10000 traces per point, reducing the statistical error to a negligible level on the scale of interest. The behaviour shows that as the charge increases, the jitter appears to approach an asymptote. As the charge nears the threshold level, the jitter grows sharply. This can be explained as the noise on the signal being translated into the time domain by the finite gradient of the waveform as it crosses the threshold. A smaller gradient leads to a longer lever arm for this action.

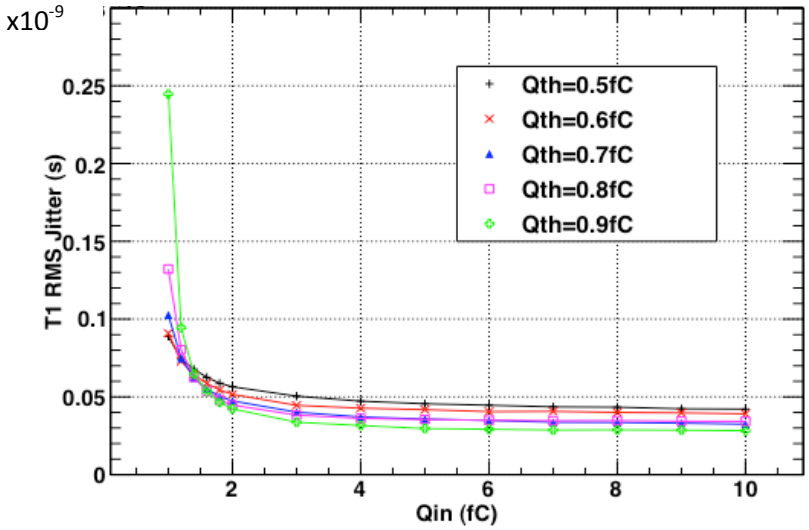

Figure 40 The t_1 rms jitter[ns] versus input charge for several thresholds and input signal charge injection duration of 2.5 ns.

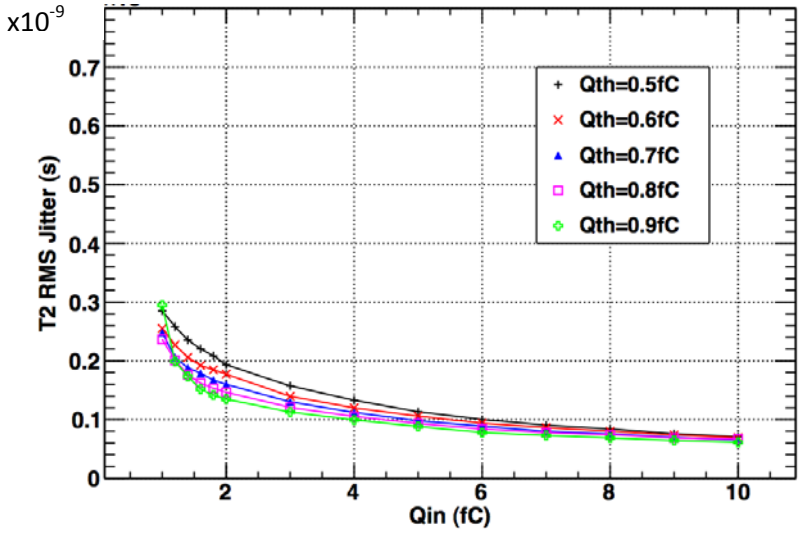

Figure 41 Analogue front-end t_2 rms jitter [ns] over input charge for several threshold values and input signal charge injection duration of 2.5 ns.

The t_2 jitter exhibits the same features, although the values are approximately a factor of 3 higher. This is consistent with the ratio of the gradients of the leading and trailing edges of the signal. The jitter of t_1 around the most probable charge value of 2.4 fC is around 55 ps. The jitter of t_2 for that charge is 150 ps. These jitter values do not contribute equally to the time walk compensated jitter,

but the t_2 jitter contribution needs to be weighted by the gradient of the time over threshold t_2-t_1 and t_1, which is around 1/3.3. The jitter depends on the level of the threshold. The preamplifier output is steepest in the mid section. Thus for high thresholds and charges higher than 2 fC the performance is consistently better. At lower charges, the higher threshold settings are closer to the pulse peak where the lower slew rate causes a corresponding degradation in performance. Thus the curves cross over each other, which can be seen at approximately 1.3 fC. This indicates that the threshold can be optimised for the charge spectrum issued from the detector.

- Behaviour of pixel at design rate:

Figure 42 shows how the t_1 RMS might be expected to vary with the frequency at which the pixel is firing for the nominal charge threshold of 0.7 fC and nominal charge injection duration of 2.5 ns. During this measurement, fixed frequencies between 10 kHz and 150 kHz, in steps of 10 kHz, were used over the standard 1-10 fC input charge range. The behaviour is largely flat with some small systematic variations occurring at high charges. These are almost certainly due to signal components with longer time constants appearing as the input charge becomes large, causing baseline fluctuations that contribute to the noise on the subsequent pulse. These results indicate that the pixel will indeed operate in a well behaved manner up to the maximum required by the experiment.

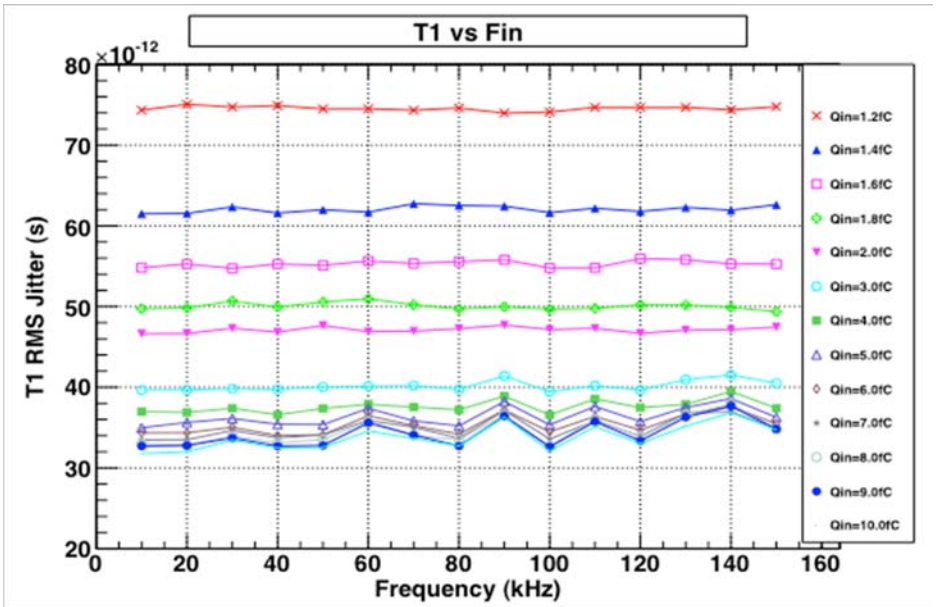

Figure 42 The t_1 rms jitter over pixel hit frequency for a threshold of 0.7 fC and various values on input charge.

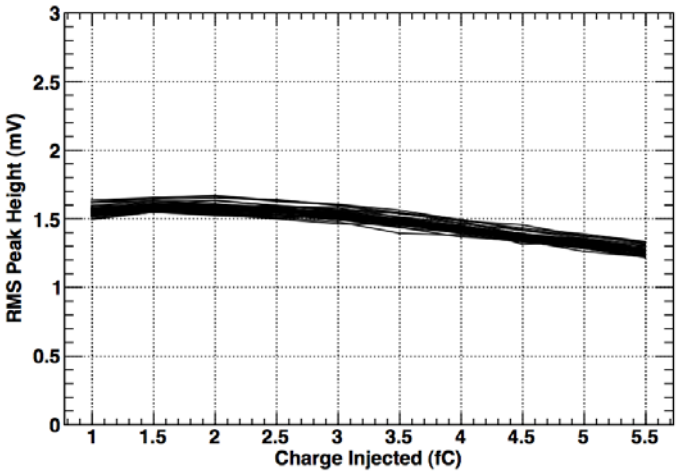

Figure 43 Rms peak height over injected charge.

Noise front-end

Figure 43 shows the rms of the peak height as a function of charge. With a gain of 70 mV/fC this corresponds to approximately 130 electrons. The measurement has been performed by varying the threshold for a given charge injection and evaluating the discriminator response (measuring the s-curve). The same noise value can be obtained by evaluation of the preamplifier output pulse without input signal. Figure 44 shows the rms of the base line to be 576 µV. This value corresponds with a to approximately 132 electrons (the gain of 70 mV/fC on the ASIC together with the external instrumentation amplifier yields a gain of 27 mv/fC).

Figure 44 Noise level on preamplifier output without input signal.

TDC behavior

TDC INL/DNL

The TDC has a digital trigger input to permit the evaluation of its performance in a manner decoupled from that of the front end and the transmission lines. The linearities shown in Figure 45 and Figure 46 were measured by applying a large number of triggers (28 million) from a flat distribution on the time scale of a clock cycle (3.125 ns). An ideal TDC would exhibit an equal probability for a given trigger to fall in any of the bins, thus by estimating the probability each bin has of receiving a trigger from the ratio of the bin content to the total trigger count, an estimator for the bin width can be obtained. This estimated bin width is then used to calculate the differential and integral non-linearities with the values of 0.17 and 0.27 of an LSB of 97 ps. The non linearity will corrected for by means of a look up table off line.

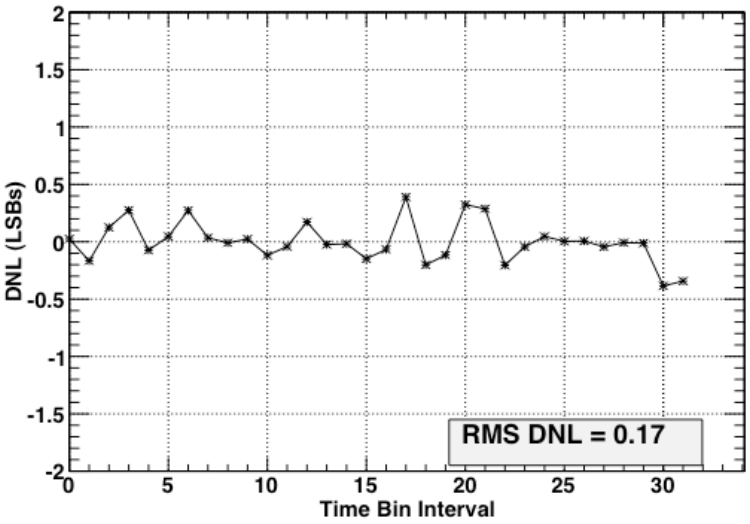

Figure 45 TDC differential non-linearity.

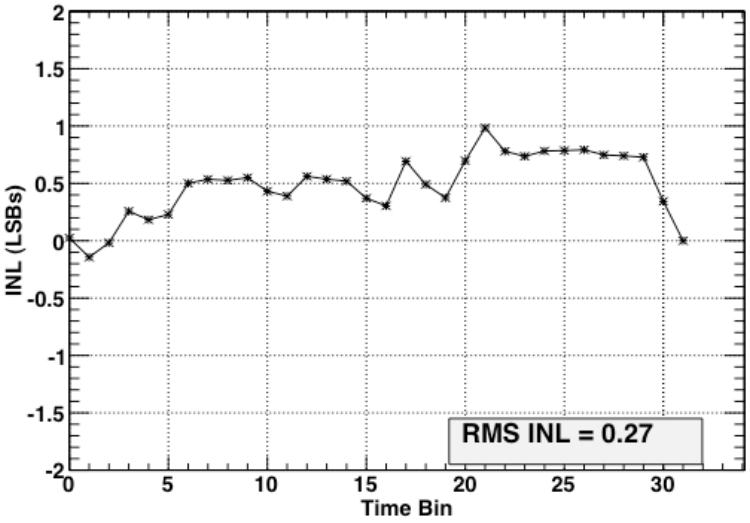

Figure 46 TDC integral non-linearity.

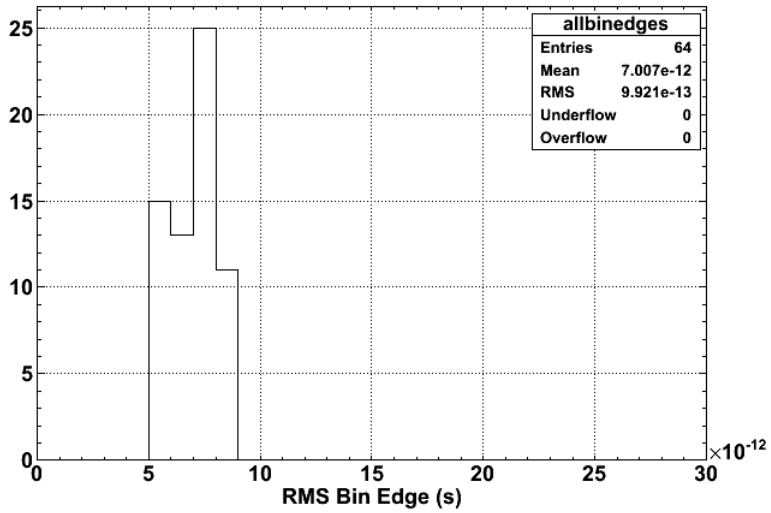

Figure 47 TDC jitter measurement.

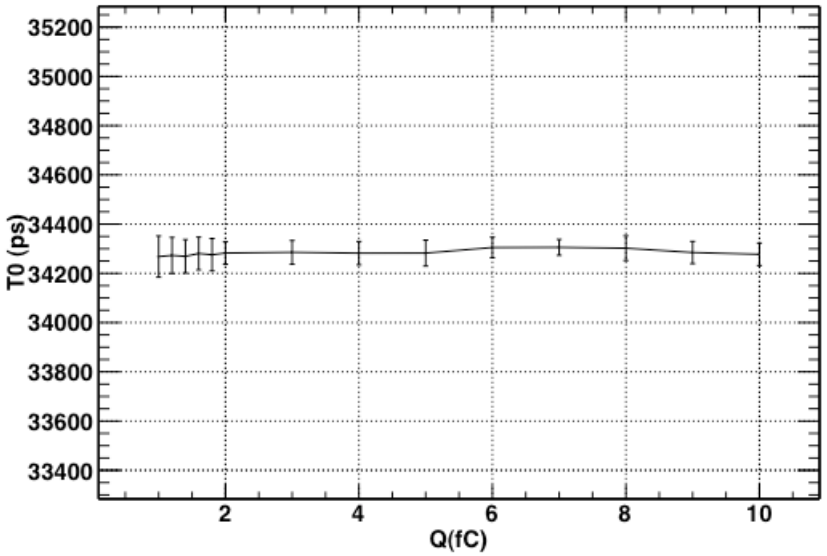

Figure 48 The t_0 - full chain time walk after compensation.

TDC jitter

By measuring the phase of the TDC digital trigger input to the DLL clock at the input to the chip on a case by case basis, two measurements of the same event are recorded: one from the oscilloscope and the other from the TDC. Taking the time given by measurements on the oscilloscope data as being the "true" time of the event, and histogramming these as a function of the TDC fine time bin, the tails on the distributions can be used to give an estimate of the jitter introduced by the TDC: i.e. when the TDC reported a given time code but should have reported the preceding (or subsequent) one. The jitter of the TDC has been measured to be 7 ps rms on average (see Figure 47).

The low jitter value shows the appropriate design and stability of the DLL control loop and the delay buffers. The performance of the TDC is dominated by the quantisation error.

Full chain arrival time resolution

Unless stated otherwise, the charge injection duration is 2.5 ns and the read-out clock is activated.

- Full chain time walk or compensation error and jitter over input charge.

The resolution of the time walk compensated time measurement is comprised of the systematic error given by the time walk compensation and the statistical error given by the jitter of the electronic components. The time-over-threshold (ToT) method gives in addition to the signal arrival time (t_1) the amplitude of the input signal by evaluation of the ToT-pulse width (t_2-t_1).

The systematic error of the time walk compensation can be eliminated. As the TDC transfer curve is known, the non-linearities can be compensated for each time code even improving the time measurement. Figure 48 shows the time walk after compensation, t_0, for a selected time offset of 1000 ps after the rising edge of the TDC clock. The error bars on the corrected time of the input charge correspond to the weighted sum of the full chain jitter of t_1 and ToT pulse width. Figure 49 gives a histogram of the t_0 residuals averaged over all phases and charges. The RMS width of the distribution is 10 ps corresponding to the uncertainty added by the time walk compensation. Figure

50 shows the jitter on t_0 from actual measurements with electrical test pulses. The mean error on the t_0 measurements for the most probable charge of 2.4 fC is around 60 ps.

Figure 51 and Figure 52 shows the rms t_1 and ToT jitter for all 45 pixels overlaid over input charge. The t_1 jitter for the most probable input charge of 2.4 fC is 55 ps. This shows the uniformity across all 45 pixels.

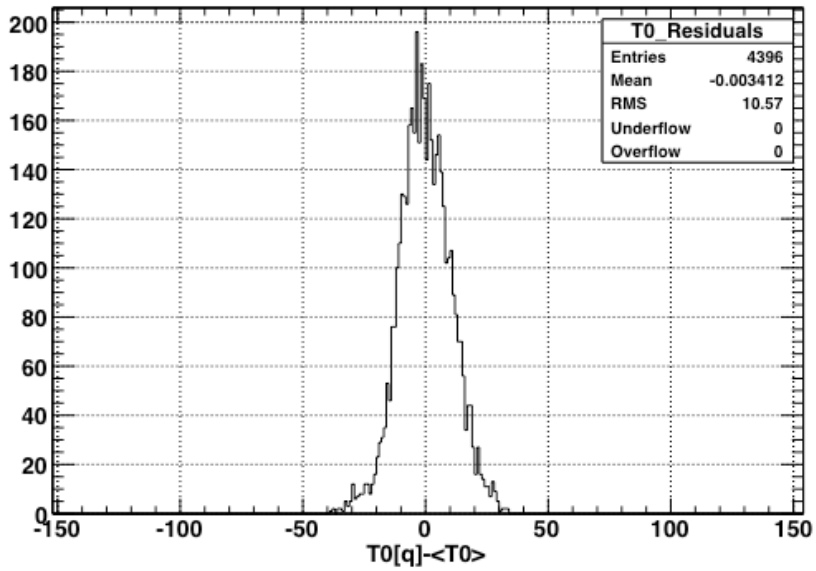

Figure 49 The t_0 residuals for all phase and all charge measurement points.

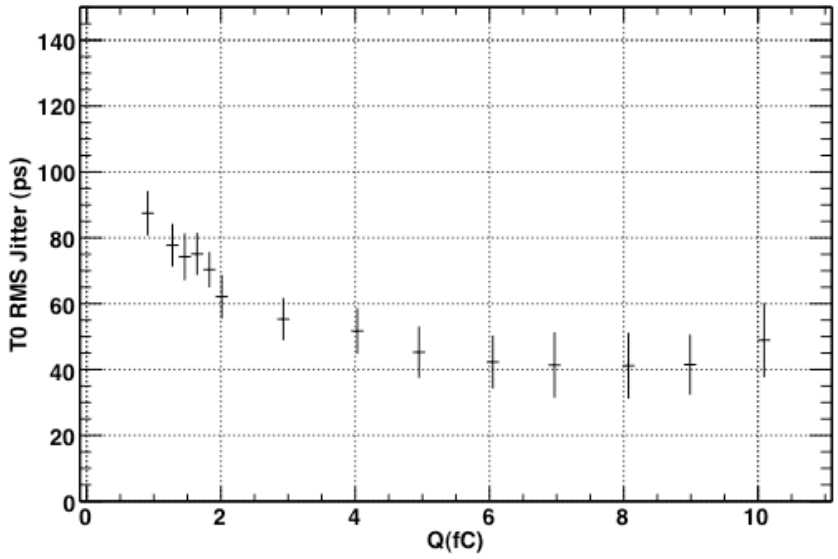

Figure 50 Jitter t_0 or time walk correction error measured with electrical test pulses.

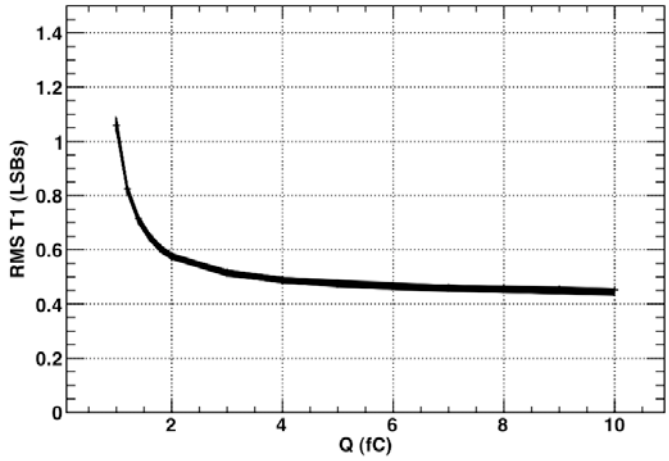

Figure 51 Full chain t_1 jitter over input charge for 2.5 ns input charge charge injection duration.

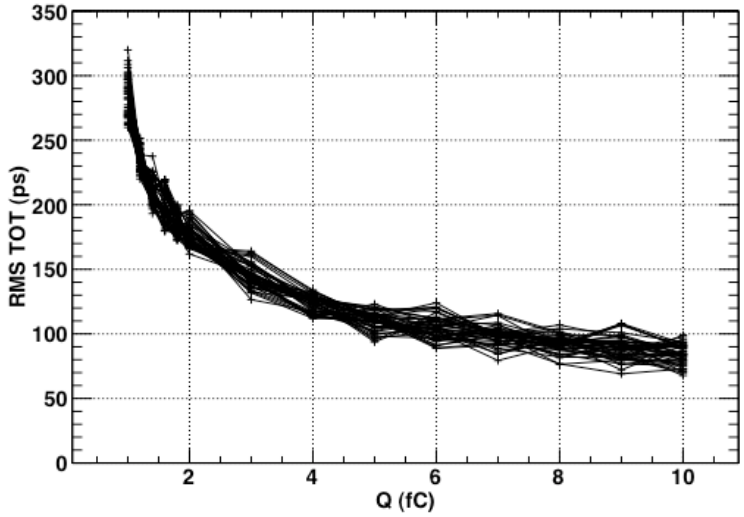

Figure 52 ToT pulse width rms jitter over input charge for 0.7 fC threshold.

Figure 53 Pre-amplifier signal output trace.

Influence of read-out clock on full chain resolution/jitter - digital clock/noise separation

During the measurements presented above the read-out clock was switched on. In order to get a measure of the influence of the digital activity on the analogueue front-end and to show a possible coupling of the DLL and the 320 MHz EOC read-out clock to the front-end a charge of 2.4 fC was injected to the front-end of the test pixel, once with the clock on and once off. It should be noted that the test pixel is at the far end of the EOC area. Figure 53 shows the signal trace after the pre-amplifier. No clock influence can be seen.

S-curves - gain and offset distribution

Control over the global thresholds is given by a 12 bit DAC producing an output from 0-1.25 V, corresponding to a step size of 305 µV. With the design gain of the front end amplifier being 70 mV/fC, this represents a step of approximately 0.00436 fC or 27 electrons. Stepping the threshold across a (large) fixed number of pulses generated from a given input charge and measuring the turn-on curve permits the front end amplifier offset, gain and noise to be determined, though a contribution to the noise value from the pulse generator is included. The integrated noise profile of the pulse can be accurately fitted using a complementary error function or cumulative frequency curve. The mean and standard deviation of the underlying Gaussian give the point where 50 % of the injected charges evoke a response from the discriminator and an estimate of the front end amplifier noise respectively. Repeating this procedure for charges in the range from 1.0 fC to 5.5 fC in steps of 0.5 fC results in the family of curves shown in Figure 54. From these measurements, the gains of all the pixels in the main array have been estimated using a straight line fit to the region from 1.0 fC to 3.0 fC. Figure 55 shows the distribution of the pixel gains for one full column with a mean of 72 mV/fC – the simulation value is 70 mV/fC.

Projecting the gain curves back to the axis permits the offset for each pixel to be estimated. The distribution of pixel offsets for the main pixel array is shown in Figure 55. This gives crucial input into the design of the threshold trim (see section Trim DAC requirement on page 52).

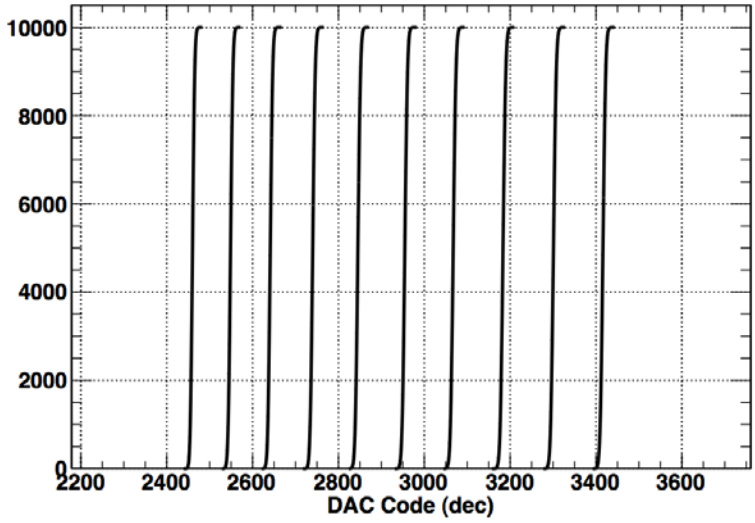

Figure 54 S-curves for several thresholds and 2.5 ns input signal charge injection duration.

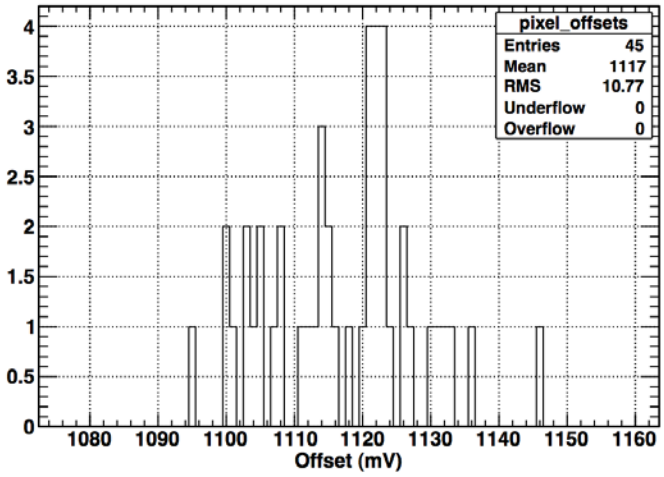

Figure 55 Distribution of pixel offsets for a full column.

Figure 55 shows the peak to peak distribution of the pixel offset for a full column before trimming of ~60 mV.

Behaviour of long transmission lines

In the EOC architecture the signals from the ToT discriminator are sent from the pixel cells over differential transmission lines to the end-of-column area. Depending on the position of the pixel this transmission length ranges from 0 to 14 mm. The maximum lengths exceed the value of 13.5 mm (45 * 0.3 mm) for the final columns as the column in the demonstrator has been folded and the curves of the folded structure add 0.5 mm. Figure 56 shows full chain rms jitter t_1 pulse width over the distance from the end-of-column region. They show that there is no dependency of the position of the pixel with respect to clock and it can be concluded that the transmission does not add to the measurement error.

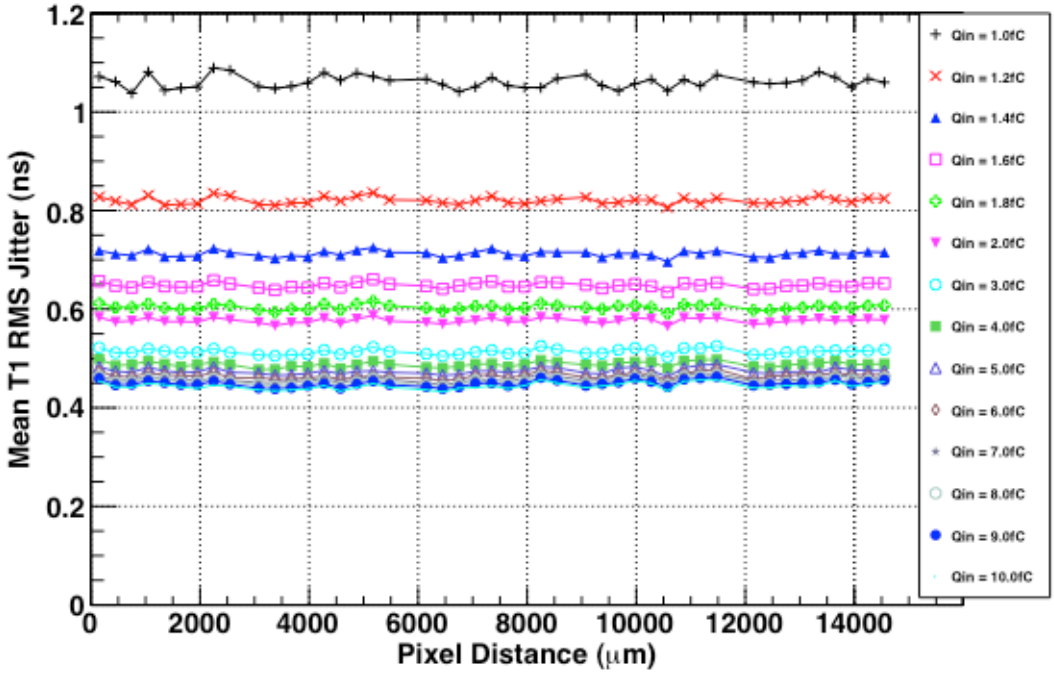

Figure 56 Full chain t_1 jitter over distance from end-of-column

Influence full chain resolution/jitter on signal charge injection duration

As the sensor signal shape will change with radiation, a study on the influence has been done. The nominal test pulse charge injection duration of 2.5 ns was varied up to 6.5 ns.

A summary of the gain as a function of the charge injection duration can be seen in Figure 57; as the distance between the curves for different charge injections do not change the plot also shows that the offset is preserved. The lower effective gain results in a lower signal gradient at the output of the front end amplifier, which in turn causes a degradation in the jitter performance. This can be seen in Figure 58, which shows the t_1 jitter behaviour as a function of injected charge for the same range of charge injection durations. Note that the threshold here is set to 0.7 fC, including an adjustment for the new effective gain.

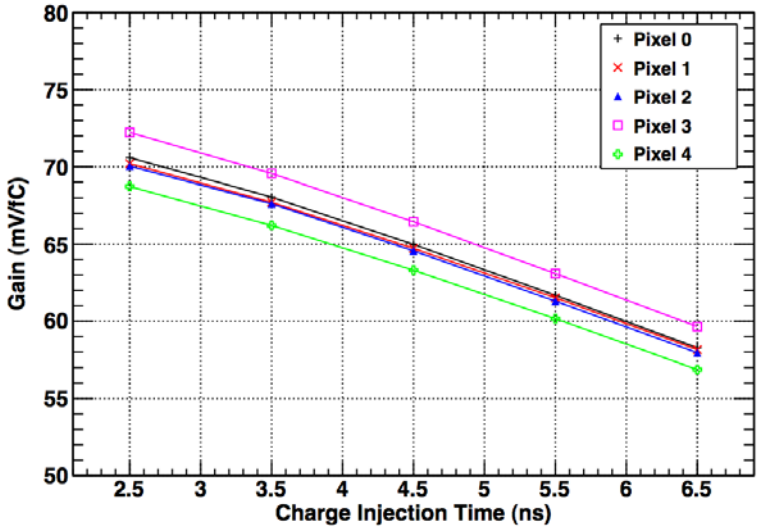

Figure 57 Pixel gain versus charge injection time.

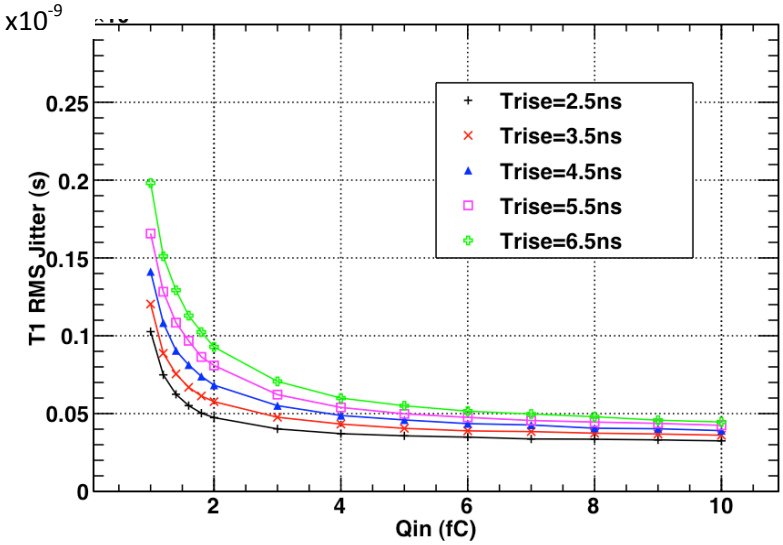

Figure 58 shows t1 jitter over signal charge injection duration.

Figure 58 and Figure 59 show the t_1 jitter and ToT pulse width of the discriminator over input signal charge injection duration. It is assumed that the time walk compensation is calibrated for each of the signal charge injection duration values. As the silicon sensor accumulates total dose over weeks of

operation, the signal shape and duration will change. The ToT scheme allows recalibration to changing sensor signal shapes.

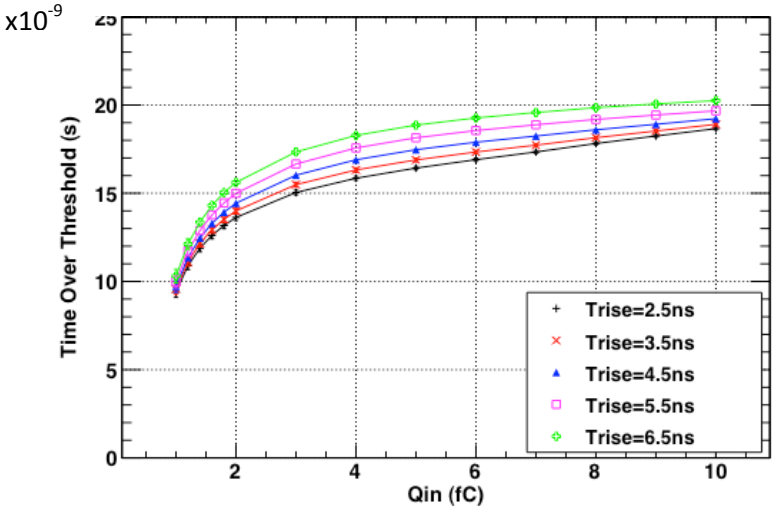

Figure 59 ToT pulse width versus input charge for different charge injection durations.

First measurement of the silicon sensor bonded to the read-out ASIC (silicon assembly)

The silicon assemblies where the sensor is bump bonded to the read-out chip have been produced and delivered in July 2010. Test with radioactive source indicate that the bump bonding process and the charge collection is effective.

Noise analysis with the silicon assembly has been conducted. Figure 60 shows the equivalent noise charge ENC over high voltage for different operation modes for the test pixel with analogue output. It shows that the noise for a high voltage value of 50 V is ~ 180 e-. The values have been obtained by evaluation of the pre-amplifier output signal. Analysis of the s-curve, see Figure 61 and Figure 62, yields a value of ~ 180 e- as well.

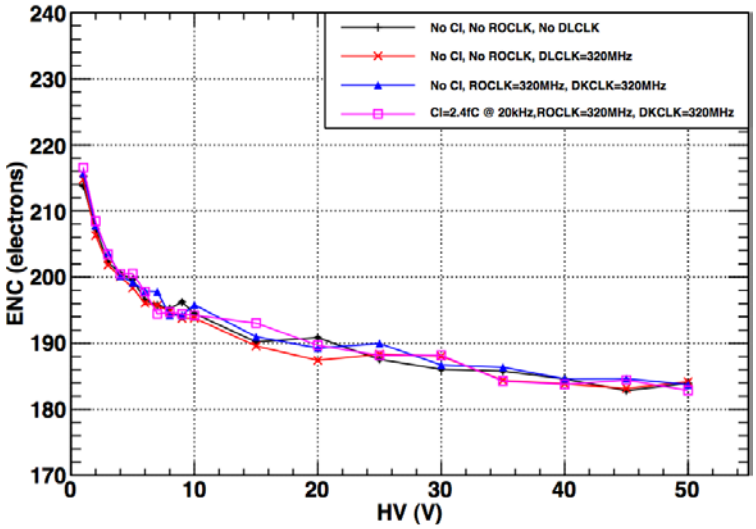

Figure 60 Equivalent noise charge versus high voltage (CI=charge input, ROCLK=read-out clock, DLCLK=digital clock).

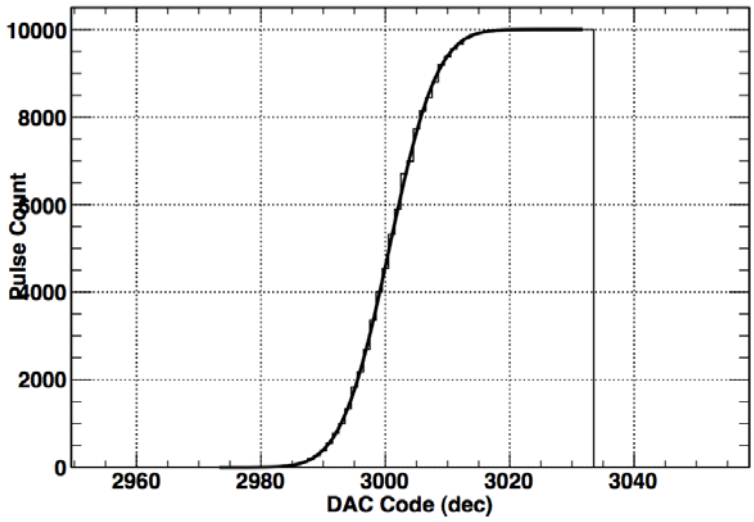

Figure 61 S-curve for the silicon assembly (DAC point = 0.305 mV).

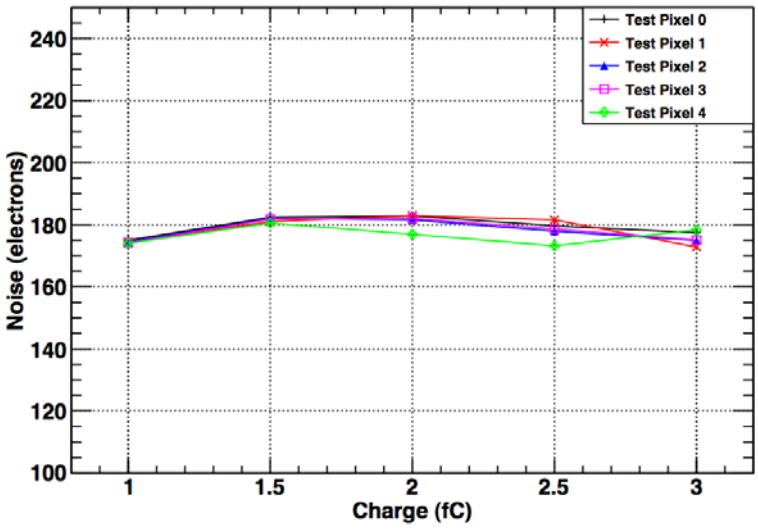

Figure 62 Noise versus electrical charge. HV=50V.

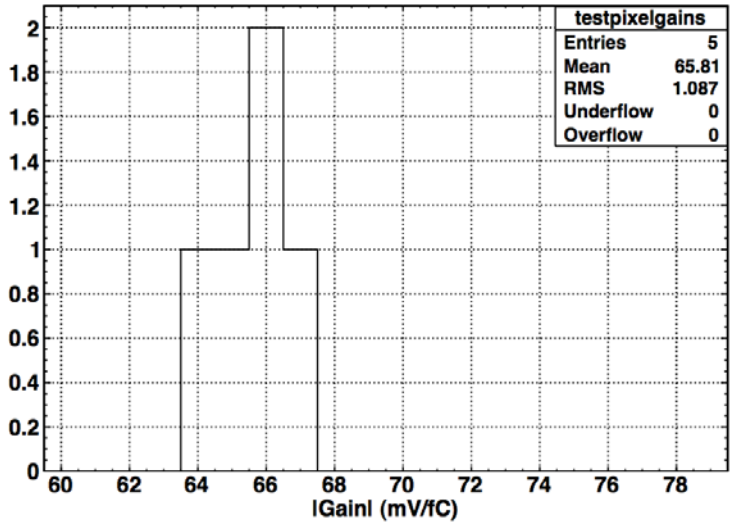

Figure 63 Analogue test pixel gain distribution. HV=50V.

As expected and in coherence with simulation the gain is 66 mV/fC. The reduction of the gain compared to the gain without sensor (70 mV/fC) is due to the additional capacitive load on the input. Figure 63 shows the gain distribution for the test pixels with analogue output.

Trim DAC requirement

In the final implementation the threshold of each pixel cell will be set by on-chip DACs. In-pixel DACs located in each pixel cell are used to compensate the offset in each pixel to equalize all pixels. A global DAC in each column will set the operating threshold for the offset compensated full column. The DACs are also used to scan the threshold during a s-curve scan and measure the turn-on behaviour.

The noise of the analogue front-end has been measured to be 130 electrons. Assuming a worst-case scenario the noise with the sensor is estimated to be below 300 electrons, corresponding to 0.05 fC. Providing a 5 σ separation from that noise estimate the minimum threshold is set to 0.25 fC. The maximum threshold is set to 4 fC in order to permit accurate determination of offset and gain by threshold scanning across known charge values (s-curve). The offset distribution earlier in this document in Figure 55 shows a peak-peak distribution of 50 mV for a 45 pixel column. This and an appropriate contingency is the range, which the in-pixel DAC needs to cover. The bin size of the in-pixel DAC is defined by the need to scan the s-curve over the turn-on region. The width of the turn-on region corresponds to the front-end noise. It is planned to apply a 5 bit in pixel DAC with programmable LSB 1.5 to 4 mV with 8 discrete values. With a minimum binning of 1.5 mV and a conversion factor of 70 mV/fC, the corresponding LSB equivalent value is 0.021 fC.

In order to scan the turn-on region of the s-curve with at least 6 points the LSB of the global 8 bit DAC has to be less than 1 sigma of the noise. For this reason a value of 1.2 mV was chosen and the data supporting this can be seen in Figure 61. The corresponding full range is 300 mV or 4.4 fC.

Table 6 shows the specifications for the DAC in the final chip. The minimum values in the paragraphs above have been adapted to accommodate tolerances.

Table 6 TRIM DAC specifications

	Global DAC	**In-pixel DAC**
# of bits	8	5
LSB	1.2 mV / 0.017 fC	1.5 - 4 mV / 0.021 - 0.057 fC
Min. value	0	0
Max. value	307 mV / 4.3 fC	45 - 128 mV / 0.064 – 0.183 fC

1.1.5 Gigatracker Test Results

1.1.6 Laser test

The EOC sensor/ASIC assembly was tested in a dedicated laser setup. The laser setup consists of a 1064 nm laser head mounted on an automated xy-stage with a focusing point of 7 µm (Gaussian sigma). The electronic test setup was identical to the setup used for the electrical charge injection tests described earlier with the exception that the test pulse trigger signal was sent to the laser, allowing full timing qualification. The laser energy injected into the sensor was measured by means of an optical splitter.

Using ^{109}Cd and ^{241}Am sources the laser setup was calibrated and a conversion table between laser energy and injected charge was built up. Figure 64 shows the energy pulse-height spectrum obtained with the Americium source. The output of the analog test pixel was histogrammed.

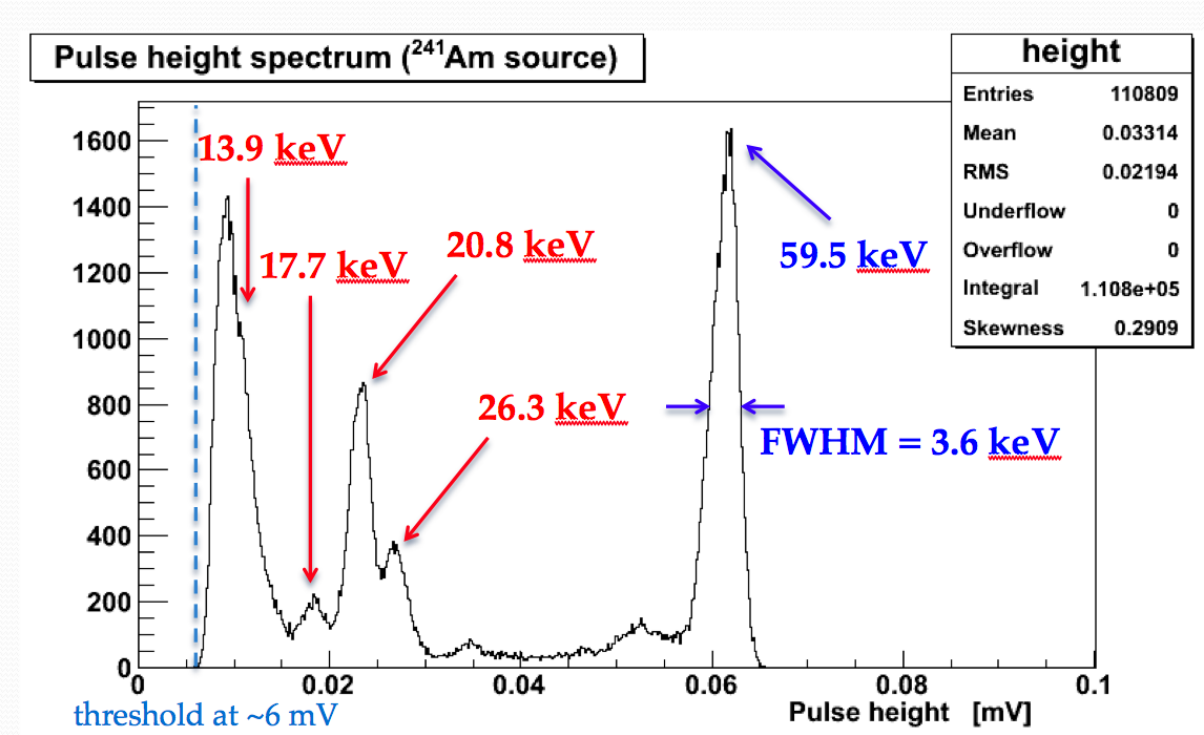

Figure 64 Pulse height spectrum for a ^{241}Americium source.

The laser setup was used to determine the timing behaviour of the full processing chain. Figure 65 shows the t_0 jitter for a pixel at the far end and close end of the column. The time walk compensated jitter t_0 for the most probable charge input of 2.4 fC is around 70 ps. For this measurement the laser light was centred in the pixel cell. Further measurements will be conducted to evaluate possible degradation when the charge is deposited off centre. The difference of the jitter between far and close pixel is very small. This again is an indication that neither the signal transmission of some 14 mm to the EOC logic nor the close proximity of the digital logic in the EOC contributes significantly to the jitter.

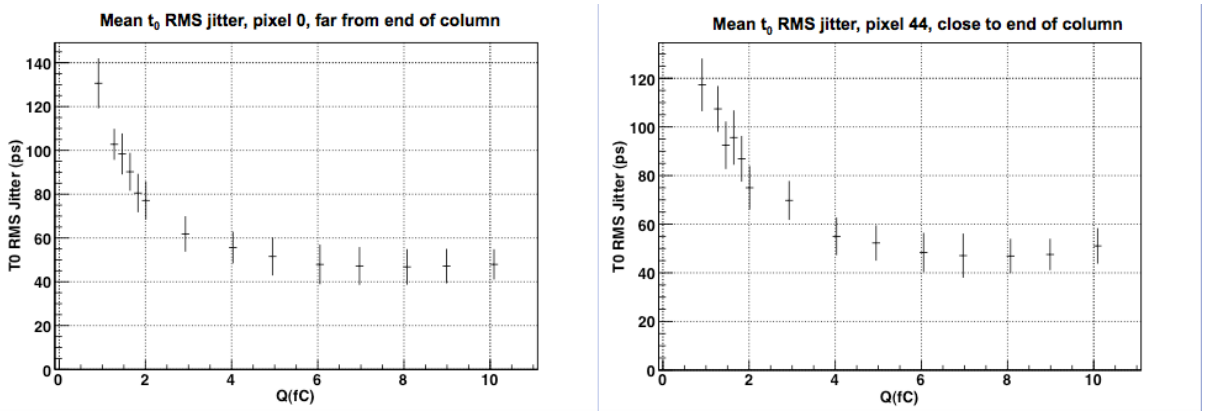

Figure 65 t_0 rms jitter for pixel 0 (far from EOC, left) and pixel 44 (close to EOC, right).

1.1.7 Gigatracker test-beam

The demonstrator bump-bonded assemblies were characterized in a dedicated test-beam. The setup of the experiment run in T9 (PS East Hall) is shown in Figure 66: a low rate 10 GeV/c hadron beam (mainly π+ and p) traversed consecutive GTK assemblies (four EOC and four p-TDC) installed on precisely aligned mechanical supports and scintillators with ~40 ps time resolution as time reference.

The assemblies were continuously operated at 300 V bias voltage for three weeks and both the assemblies and the DAQ showed very stable performance.

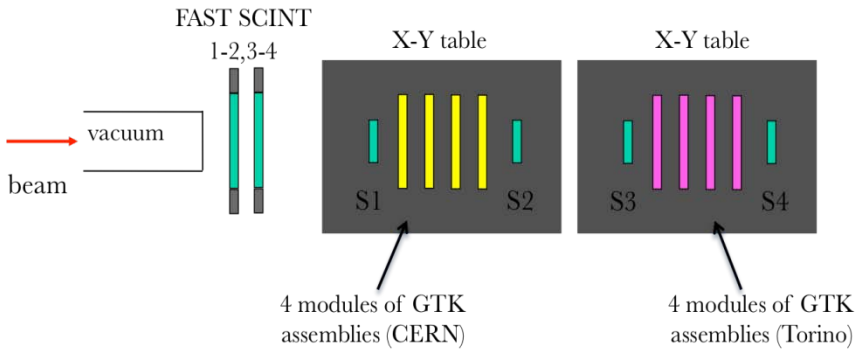

Figure 66 Schematic layout of the test beam setup in the T9 beam line.

The test beam data analysis was at the writing of this document ongoing. A first crude analysis with several data cuts applied is described below.

The time-walk correction is applied using the fast scintillators data, which provide the reference time. The initial simple analysis corrects the time walk but cuts the tails of the Time over Threshold distribution and rejects consecutive hits from pixels in a given 5-pixel group, as the absolute delay of these hits is different. Due to these cuts more than 20 % of the data was rejected. The cuts were applied to simplify and speed up the analysis; the final analysis will take all hits into account and applies a correction for the different delays.

After these cuts and the time-walk correction, the distribution of time differences between hits in the same pixel of two consecutive GTK stations is shown in Figure 67. The width of the distribution is ~250 ps rms, which corresponds to ~175 ps for a single GTK station. The analysed data on the EOC architecture indicate also a very low level of noise, as no pixel hits were recorded out-of-spill period.

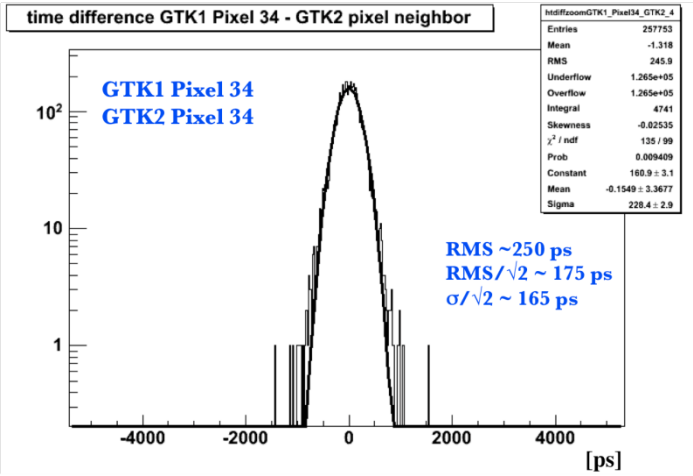

Figure 67 Plot of time correlation between hits in two consecutive GTK stations (EOC architecture).

1.1.8 Cooling

The cooling design is driven by the following GTK requirements:

- Material budget:
  - radiation length in active beam area ≤ 0.5% of X_0,
  - minimum amount of material in the sensor area (60x27mm) plus a 10 mm safety zone,
  - outside that zone no material budget constraints.
- Detector operating temperature 5°C or lower;
- Best possible uniformity of the temperature across the sensor area;
- Power by one GTK station 32 W;
- System operating in vacuum.

Due to the low mass of the detector module thermal run-away and destruction needs to be avoided. Furthermore the cooling system for the Gigatracker reduces the radiation damage of the sensor and consequently increases the module life time. Although the upper limit of the operation temperature has been set to 5°C, a lower operation temperature is highly desirable, as the module life time would increase from 50 days to 100 days when cooled at -20 °C. The material budget and the operation in vacuum are main integration challenges of the project. Mechanics and cooling integration can profit from the fact that no material budget restrictions exist for the area 10 mm outside the beam. From a cooling performance point of view the anticipated 2 W/cm^2 power dissipation by the active electronics are not difficult to meet. However, the cooling and mechanics support architecture must introduce only a minimum of material into the beam area and at the same time assure mechanical stability of the module and allow access to high speed electrical connections. At this time of the project the working group is pursuing two independent options: 1) a micro-channel structure; 2) a gas cooling vessel with mylar walls. **Error! Reference source not found.**

1.1.8.1 Micro-Channel Cooling

1.1.8.1.1 Motivation and Goals

One viable option to obtain a local cooling system satisfying the stringent requirements mentioned is the development of a thin silicon micro-channel cooling plate circulating liquid C_6F_{14}. The development aims for an operating temperature of -20°C to -30°C. Figure 68 shows a schematic of the hydraulic layout of the micro channels: the coolant will enter and exit the straight channels via manifolds positioned on top and bottom. The channels, distribution manifold and openings for the inlet and outlet connectors are etched into a silicon wafer, which is then coupled to a second wafer closing the hydraulic circuit. The final goal is to have both wafers in silicon bonded together by fusion bonding to produce a monolithic cooling element (11).

Recent results obtained in two different fields of development provide guidelines to solve the issues in engineering this cooling device, but dedicated R&D is nevertheless unavoidable for the specific application under study. On the one hand, micro-channel cooling devices have started to be actively studied for future applications for high power computing chips or 3D architectures (12), (13). For these applications, where the power densities are extreme, the mass of the device (hence its material budget) is an irrelevant parameter. On the other hand, thin and light micro-fluidic devices in silicon are in development for bio-chemical applications (14), but the typical values of the flow rate

and pressure are much lower. Furthermore the presence of a low temperature fluid and of a high radiation level(15) is unique to the application in a HEP detector.

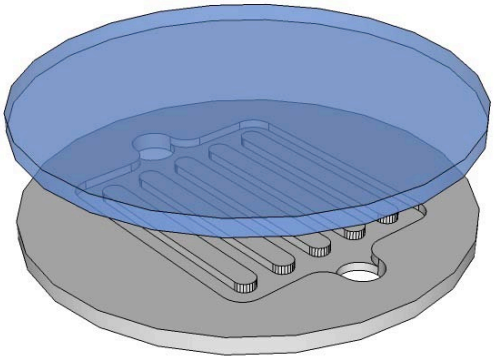

Figure 68 - Schematic of the micro-channel cooling device.

The procedure followed to tackle the different challenges and to converge in a limited time to a single device satisfying all the requirements is to move in parallel along different lines of R&D in a "matrix" approach, where the intermediate results of one line are used to steer the parallel developments. In this way, the study of the optimal fabrication technique provides information on which geometrical layouts are reliably attainable and which process is best suited for a particular layout. Analytical calculations and numerical CFD simulations are performed in order to explore a number of possible layouts and define optimal targets from the point of view of the thermo-hydraulic performance. Numerical structural simulations set the pressure limits to be considered for each channel dimension and wafer thickness. Experimental tests are conducted to validate the numerical model and to allow for a safe extrapolation of the numerical forecast to new configurations.

The maximum temperature difference of the coolant between the inlet and outlet has been set to $\Delta T=5$ K to guarantee a relatively uniform temperature distribution and minimize all thermal stress and related deformations. Also, to limit the mechanical stress in the micro channel cooling plate, the maximum acceptable pressure drop inside the cooling plate has been set to $\Delta p=5$ bar. Based on the above criteria the micro-channel dimensions could be as follows:

- 90 μm deep channels etched into a silicon wafer of 120 μm thickness
- Covered by a fusion-bonded silicon wafer with 30 μm thickness over the active area.
- Outside this area, the thickness of both silicon wafers is 520 μm, allowing for 400 μm deep outlet and inlet manifolds as well as good structural stabilization of the whole device

This design leads to an additional material budget 0.16% of X_0.

An alternative design, in case of technical difficulties with the fusion bonding process, relies on a flat Pyrex cover 50 μm thick anodic-bonded to the silicon wafer carrying the hydraulic circuit (16). On top of this flat plate, an additional silicon frame (surrounding the beam area) will again be anodic-bonded. In this way the global structure of the cooling wafer will be symmetric, the effects of differences in the coefficient of thermal expansion (CTE) between silicon and Pyrex will be minimized and the same resistance to pressure and manipulation as in the baseline case will be attained.

For technological reasons and for optical access to the channels in the test and development phase, the first prototypes consist at the moment of a silicon wafer housing the channels and a closing Pyrex wafer joined to it by anodic bonding. In the following sections, details about the status of each R&D line will be given and the results leading to the proposed design will be discussed.

1.1.8.1.2 Device Fabrication

The fabrication of the cooling devices is performed at EPFL CMI[1] (Center of MicroNanoTechnology) in collaboration with the Microsystems Laboratory[2]. The micro-fabrication process is shown in Figure 69. It starts with a Czochralski silicon wafer polished on both sides (4″ diameter, 380 µm thick, 0.1-0.5 ohm-cm p-type). A layer of 1 µm of oxide (SiO_2) is grown on both sides of the wafer (Figure 69 (a)). Clariant AZ-1512HS photoresist is spin coated on one side of the wafer at 2000 rpm and lithography is performed to obtain an image of the channels in the photoresist (Figure 69 (b)). The chrome masks for the lithography used in this process are fabricated at CMI using the Heidelberg DWL200 LASER lithography system. The patterned photoresist is baked at 115°C for 50 s on a hotplate. Dry etching of the top layer oxide is used to transfer the micro-channels pattern (Figure 69 (c)). This structured oxide layer will subsequently be used as a mask to transfer the pattern of the channels in the Si wafer by deep reactive ion etching (DRIE). A second lithography is performed with frontside alignment to image two fluid transfer holes, 1.4 mm diameter, for fluid injection and collection from the two manifolds (Figure 69 (d)). DRIE is used to etch the access holes partially down to 280 µm (Figure 69 (e)). The photoresist is stripped in Microposit Remover 1165 at 70°C (Figure 69 (f)) and DRIE is used to etch anisotropical 100 µm deep channels separated by 25 µm wide structures in silicon (Figure 69 (g)). Subsequently the oxide layers are removed by wet etching in BHF 7:1 for 20 min at 20°C (Figure 69 (h)).

At present, the processed Si wafer and an unprocessed Pyrex wafer (4″ diameter and 525 µm thick) are then cleaned in a Piranha bath (H_2SO_4 + H_2O_2) at 100°C and anodic bonding is performed to close the channels with the Pyrex wafer (Figure 69 (i)). The bonding is performed at ambient pressure and the temperature is raised to 350°C before being lowered to 320°C. At this stage a constant voltage of 800 V is applied between the Si and Pyrex wafer (Figure 70). The current is monitored and the bonding process is stopped when it reaches 10% of the initial value, typically in the order of 1 mA.

In the final production both the processed and the unprocessed wafers will be of 525 µm thick silicon. At this stage of the fabrication the two wafers will be dispatched to a specialized fusion-bonding laboratory after the execution of the specific surface preparation required. The bonded wafer will then be received back at CMI for the final processing. The related preliminary discussions with two specialized laboratories at present being pursued.

The basic fabrication process-flow described above is used to produce prototype devices to validate CFD thermo-fluid dynamics simulations and ANSYS structural calculations. The results provided with such validated numerical simulations allow predicting the behaviour of thinner devices with

[1] http://cmi.epfl.ch/

[2] http://lmis4.epfl.ch/

improved design to be predicted, thereby reducing the number of tests and the types of samples to be produced. A more complex micro-fabrication process is at present in preparation. This includes an additional step of photolithography to allow for different heights of manifolds and a final local etching to be applied to obtain a thinner region in the beam acceptance area.

Following any additional final processing (e.g. local thinning over the minimum material area), the bonded wafer is then diced according to alignment marks previously etched in Si to obtain a cooling plate with precise external references for integration into the electromechanical assembly.

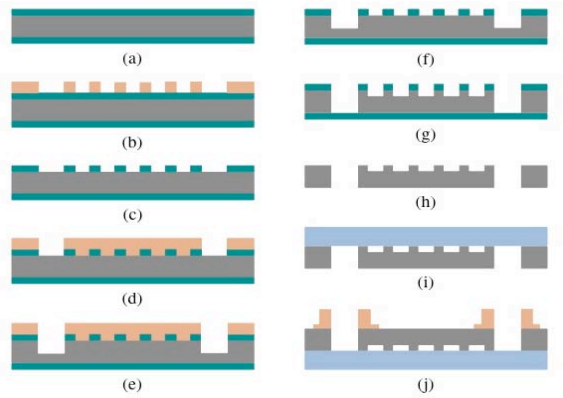

Figure 69 - Fabrication process-flow.

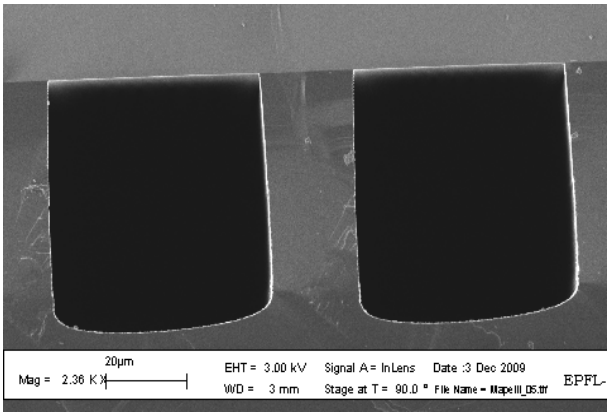

Figure 70 - Scanning Electron Microscope image of the cross-section of micro-channels etched in silicon bonded to a Pyrex wafer.

Finally, PEEK connectors (NanoPort assemblies from Upchurch Scientific) are aligned, with a gasket and with a preformed adhesive ring to the inlet and outlet on the silicon before clamping. They undergo a thermal treatment at 180°C for 2 hours to develop a complete bond between the NanoPort and the silicon substrate. The supplier guarantees correctly attached NanoPort connectors to withstand pressures of 34.5 bars. However we tested them repeatedly up to pressures of 70 bars before failure.

The choice of the cooling fluid circulating in the micro-channels has naturally been oriented towards perfluorocarbon fluids (C_nF_{2n+2}), which are widely used as coolant medium in LHC detectors. In particular C_3F_8 and C_6F_{14} are used in the inner tracking detectors of ATLAS and CMS. They exhibit interesting properties for cooling applications in high radiation environment such as thermal and chemical stability, non-flammability and good dielectric behaviour. Radiation resistance studies have been performed (17) by irradiating samples of C_6F_{14} with gammas from a ^{60}Co source up to an accumulated dose of about 5.6×10^4 Gy. At the location of the GTK stations, accumulated doses in the order of 6×10^4 Gy are expected over one year of operation. At such doses the sample C_6F_{14} “Flutec PP1” (iso-perfluorohexane from F2 CHEMICALS) shows the best radiation hardness and chemical stability. It has thus been selected as the cooling fluid for the micro-channel devices.

1.1.8.1.3 Thermo-Fluid Dynamics Design, Simulation and Validation of Micro-Channels

The relevant properties of C6F14 for thermo-fluid dynamics calculations are summarized in Table 7.

Table 7 C_6F_{14} properties.

Properties	**C_6F_{14} @ -25°C**
Density r [kg/m^3]	1805
Viscosity n [10^{-7} m^2/s]	8.2
Heat capacity c_p [J/(kg K)]	975
Thermal conductivity l [10^{-2} W/(m K)]	6.275

The required C_6F_{14} mass flow to extract the 32 W dissipated by the readout chips with the assumed temperature difference of 5 K is 7.325 x 10^{-3} kg/s. As a consequence, the Hagen-Poiseuille-law applies for the analytical pressure drop calculations, as the flow inside the channels is laminar in all configurations considered (Reynolds-Number between 100 and 300).

In all calculations the heat source is considered uniformly distributed over a surface of 60 mm x 40 mm, corresponding to the area of the readout chips. The hydrodynamic and thermodynamic length of the channels are considered coincident, L = l = 40 mm. The wall thickness between the channels has been set to half the channel width, to withstand mechanical stress due to the pressure in the channels and to provide enough bonding surface.

Analytical design

Analytical calculations were performed to obtain an initial estimate of the feasible values for the micro channel dimensions. For a given pressure drop, Figure 71 plots on the vertical axis the temperature difference between inlet and outlet as a function of the channel geometry (height and width). The continuous surface represents the ΔT in the coolant and the mesh surface the ΔT in the channel wall. It appears necessary to adjust both channel dimensions in order to keep a temperature difference of 5K. The reason is that in channels with a rectangular cross section, the smaller of the two dimensions dominates the pressure drop equation with a cubic power dependence (18). For a fixed pressure drop, the ratio between height and width should typically be kept between 0.5 and 2. It is also visible from the plot that the temperature difference between the fluid and the wall, represented by the vertical separation of the two surfaces, rises with an increase of the channel size due to a higher line load per channel and a decrease in heat transfer surface. This has to be taken into account when selecting the channel cross section, as it will have an impact on the gap between the operating temperature of the coolant and the minimal temperature attained on the surface of the sensor.

It has to be considered that the channel height translates into a requirement for the thickness of the silicon wafer, where the channel will be etched. Therefore, channel height values larger than 120 μm cannot be considered for the present application. Figure 72 relates the reachable mass flow to the channel width, for a fixed pressure difference of 2 bar and a fixed channel height of 90 μm. One can see that the required nominal mass flow is reached for a channel width of approximately 120 μm. It

is also visible that the gain in mass flow is favourable up to an increase in width to around 150 µm. For wider channels the width-to-height ratio becomes too large and the channel height starts to be dominant for the mass flow.

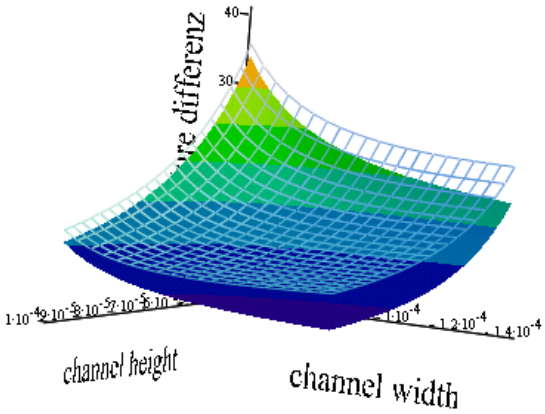

Figure 71 - Temperature differences between inlet and outlet in the coolant and in the channel wall.

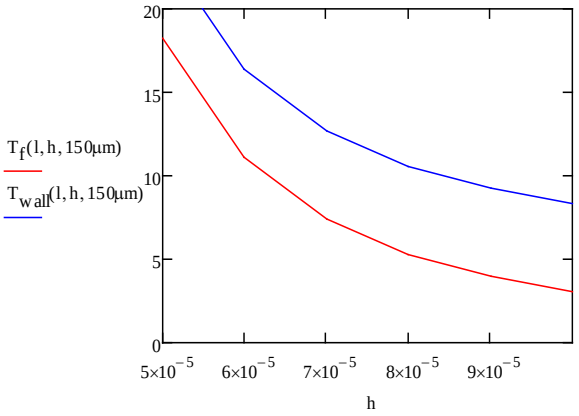

Figure 72 - Reachable mass flow related to the channel width for a fixed channel height.

The results from the analytical calculations performed indicate that the suitable ranges of the micro channel dimensions are the following:

- Width: between 100 µm and 150 µm
- Height: between 80 µm and 120 µm
- Fin width: between 25 µm and 75 µm
- Between 300 and 500 channels to cover the area

CFD simulations

Two manifolds will feed and drain the micro channels. The manifold design must aim at a small total pressure drop and a uniform distribution of the coolant over the channels. In particular, the total pressure drop gives the necessary pressure at the inlet of the cooling plate and is therefore an important factor for the mechanical stability of the design. A good hydraulic design of the manifolds requires the use of full 3D Computational Fluid Dynamic (CFD) simulations.

Different configuration and sizes of manifolds have been simulated in connection with channel geometries in the range identified above. The calculated total pressure drops span from 15 to 2 bar, mainly depending on the cross sectional area of the feeding manifold. The details of the best performing solution provided by the simulations are shown in the box of Figure 73: it is a wedge-shaped manifold, providing the advantage of a more uniform flow distribution inside the channels with respect to a simple straight geometry (19).

Figure 73 shows the pressure map for the following specifications and boundary conditions:

- Channel cross section 100 µm x 100 µm
- 2 NanoPort connectors placed on opposite sides of the inlet and outlet manifolds

- Wedged manifold with maximum width 1.6 mm on the inlet and 2.1 mm on the outlet
- Manifold depth 400 µm
- Inlet: mass flow 7.325 x 10-3 kg/s (ΔT 5 K)
- Outlet: pressure 1 bar

The pressure drop distribution across the channels is very uniform, thus indicating a uniform flow distribution of coolant across the device surface. The total pressure drop is 3.1 bar.

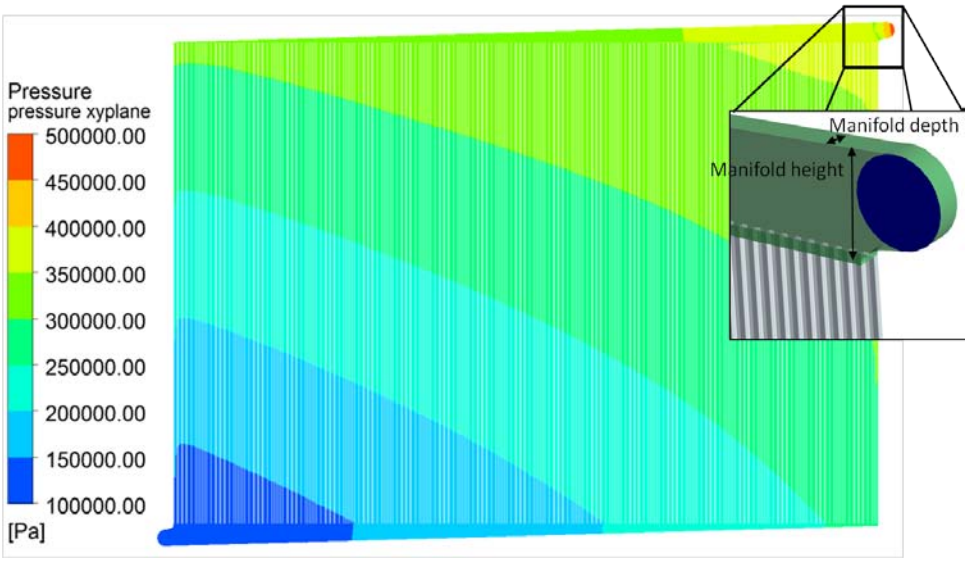

Figure 73 - Pressure [Pa] in micro channels and manifold with inlet and outlet connector on opposite sides.

Validation on full-scale prototypes

The analytical and computational simulations have been validated on full scale prototypes. A slightly simplified design has been used for these tests. It consists of an array of 480 channels, uniformly distributed over a 60 mm wide area. The channels have a cross section of 100 µm × 100 µm, are 48 mm long and are separated by 25 µm wide silicon walls. Two 60 × 1 mm rectangular distribution manifolds, 100 µm deep, connect all the micro-channels together at both ends. Injection and collection of fluids is performed through PEEK capillaries screwed into a NanoPort connector placed at the centre of each manifold. This geometry is not optimized for the overall pressure drop, but allows for a faster and less expensive production.

Figure 74 compares the pressure drop between inlet and outlet of the CFD simulation and the wafer in the test stand. One can see that the prediction of the simulation is very close to the experimental data. This validates the numerical model developed for the CFD simulations and lends reliability to the results obtained with this model for the detailed design of the micro channels and manifolds, thus largely reducing the number of samples to be produced for testing purposes.

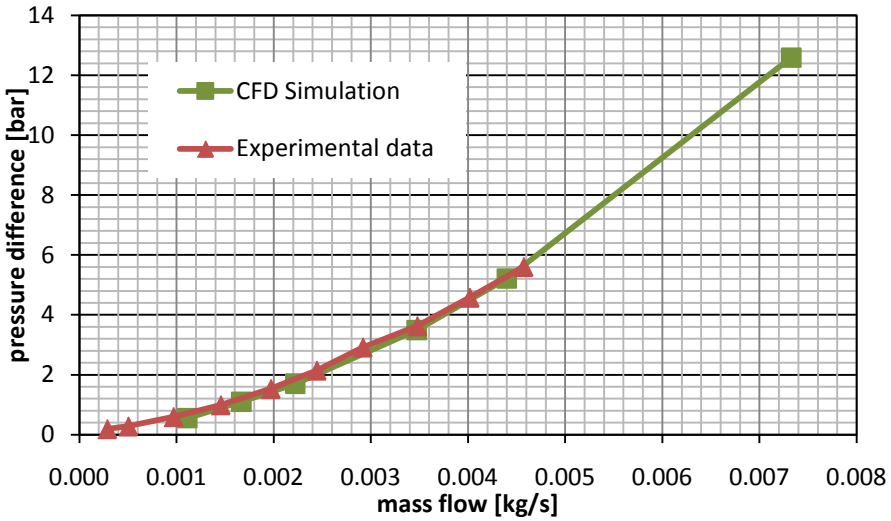

Figure 74 - Pressure drop comparison between simulation and tests.

Figure 75 shows a sample of full scale prototype wafer with a heating foil simulating the distributed thermal load glued on one side (left) and two NanoPort connectors bonded on the opposite side (middle). A typical mounting on the test stand for room temperature tests is also presented (right).

Figure 76 shows a cool down test of the wafer with the simplified geometry. The pictures are taken with a thermal camera. The light yellow rectangle on the left picture is the powered dummy load with approximately the same surface area as the read-out chips. The picture in the middle shows the moment when the cooling liquid enters the manifold and channels. The picture on the right shows the rapid temperature drop on the surface of the wafer and the dummy heater few seconds after the coolant circulation is established. A preference of the cooling liquid for the middle channels is clearly visible. The improved wedge-shaped manifold design will correct this effect.

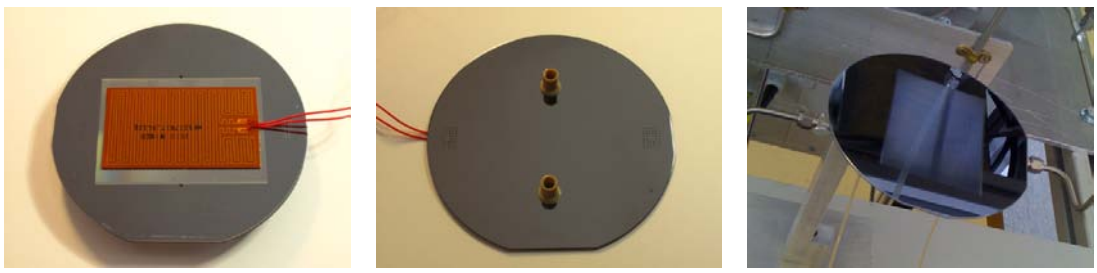

Figure 75 - Cool down test.

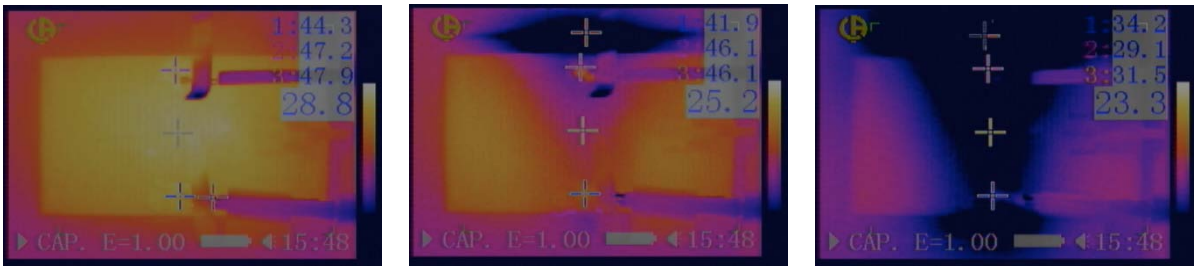

Figure 76 - Cool down test.

1.1.8.1.4 Pressure Tests and Structural Simulation

The maximum pressure safely attainable inside the cooling device without risk of collapse is a crucial parameter, as this determines the minimum amount of material needed for a given flow rate, to attain a given temperature uniformity on the detector surface. Accordingly, an experimental determination of this structural limit would require the production of an extremely large number of samples with different wall thickness. The problem has been approached in a similar way as for the analysis of the thermo-fluid dynamics properties: a numerical model has been developed and one set of samples has been produced for testing purposes. The comparison between the results provided by the simulation and the test performed allows the numerical model to be refined and validated; this can then be used to extrapolate the results obtained to different configurations.

Simplified experimental samples characterized by a 60 mm long rectangular manifold of different width w = 0.2, 0.25, 0.5, 1 and 2 mm have been produced by anodic bonding of a 380 μm thick silicon wafer and a 520 μm thick Pyrex wafer. The samples can be connected to a controlled pressure device through a NanoPort connector similar to those used for the complete cooling device. The pressure in the sample is raised until failure occurs and the limiting pressure is then determined from the recorded data while images from a high-speed camera help understanding the failure mechanism.

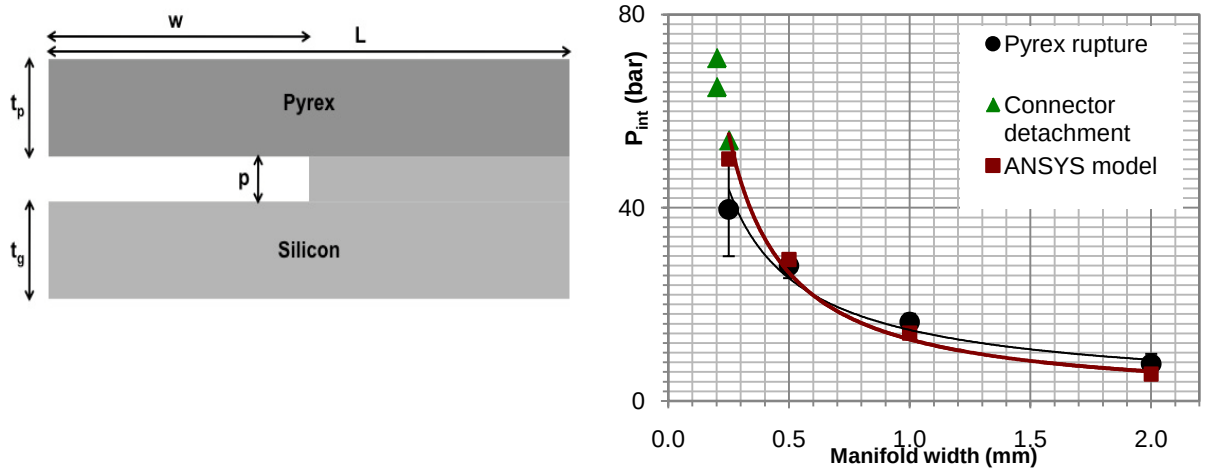

Figure 77 Left side: Geometry of the sample used in the simulations. Right side: Rupture pressure [bar] versus manifold thickness. Comparison between experimental results and simulations.

A 2D ANSYS parametric model simulating the geometry under test has been developed as shown in Figure 77: one layer of Pyrex (thickness T_p) bonded onto one layer of silicon (thickness T_g) in which a channel is engraved (width W and depth p). The total width of the sample is L. A symmetry boundary condition has been applied on the left edge and the right and bottom edges are fixed. A constant pressure P_{int} is applied in the channel.

For a defined Pyrex thickness t_p, the tensile strength was calculated in the Pyrex layer for different channel width W as a function of the applied internal pressure P_{int}. Delaminating of the bonded

interface was not taken into account in this study: indeed, as found in the literature, the Pyrex layer is identified as the weakest point in our sample. The maximum yield strength of Pyrex is 25 MPa[3] (165 MPa for Silicon). This value is then considered as the maximum stress admissible in the Pyrex layer before breakage.

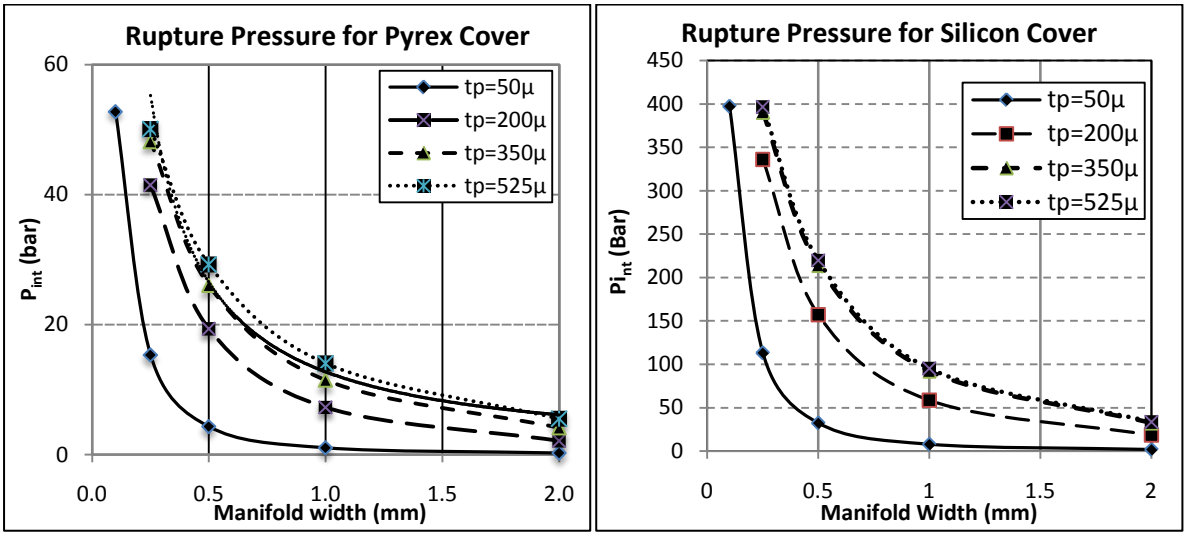

Figure 78 Simulation of the maximum sustainable pressure as functions of manifold width and thickness (left: for Pyrex; right for Silicon).

Figure 78 shows the pressure test results together with the calculated pressure for which the global stress (von Mises formulation) in the Pyrex layer is higher than 25 MPa for a Pyrex thickness of 525 µm and different values of W. The apparent good correlation between tests and simulation validates the ANSYS model and allows is to be used to forecast the behaviour of a Pyrex layer of different thickness. The slight divergence between calculated and measured data indicates a slight over-estimation for smaller values of W and under-estimation for larger values of W.

The simulated behaviour of various Pyrex thicknesses allows the maximum attainable pressure to be determined depending on the channel (or manifold) width and the Pyrex thickness desired.

Figure 77shows for the 100 µm wide channels in the beam area of the device, that even if one has a Pyrex thickness of 50 µm, the P_{max} allowed gives a large safety factor (more than 10) compared to the 3 to 4 bar pressure expected in this region. For the larger dimensions of the order of 1000 µm on the manifolds, the simulations indicate that if we want to keep a good safety factor and be able to withstand an absolute pressure of 5 bar to 7 bar, one should go to a thicker Pyrex layer. A 525 µm Pyrex thickness could hold a least 14 bar, thus providing a safety margin of 2 to 3 on the maximum allowed pressure. The manifolds must therefore be placed outside of the beam area, where a increase of the wafer thickness is allowed.

[3] As found in CES material database, www.grantadesign.com/products/ces.

Similar calculations are shown on the right hand plot of Figure 78 for a full silicon solution, for which the limiting value of the global stress considered is 165 MPa. The extremely high pressure values attained suggest that local de-bonding phenomena should not be entirely neglected in the calculations for this case. However, the higher mechanical properties of silicon clearly allow for a further reduction of the wall thickness to the desired 30 μm level. Should manipulation or production issues prevent this further reduction in wall thickness, the full silicon solution would safely allow for higher pressure levels corresponding to reduced channel dimensions. Thus preserving the target material budget might be preserved.

1.1.8.2 Gas Cooling

1.1.8.2.1 Configuration

The other cooling option is the construction of a vessel housing the GTK module and cooling it via a flow of cold gaseous nitrogen. The nitrogen will enter the vessel at a temperature of 100 K. The cold flow will keep the operation temperature of the module less than 5°C. In this design the GTK is installed between two cylindrical Kapton walls, 40 m thick, supported by an aluminum frame, see Figure 79. A second pair of thinner (10 μm) Kapton walls is inside the vessel surrounding the detector and chip assembly, providing a smoother flow over the detector surface. The total thickness crossed by the beam is (2 x 40) + (2 x 10) = 100 μm Kapton. The radiation length for this polymide film $(C_{22}\ H_{10}\ N_2\ O_5)_m$ is 28.6 cm[4] so the total material budget of the vessel is 0.035% X_0.

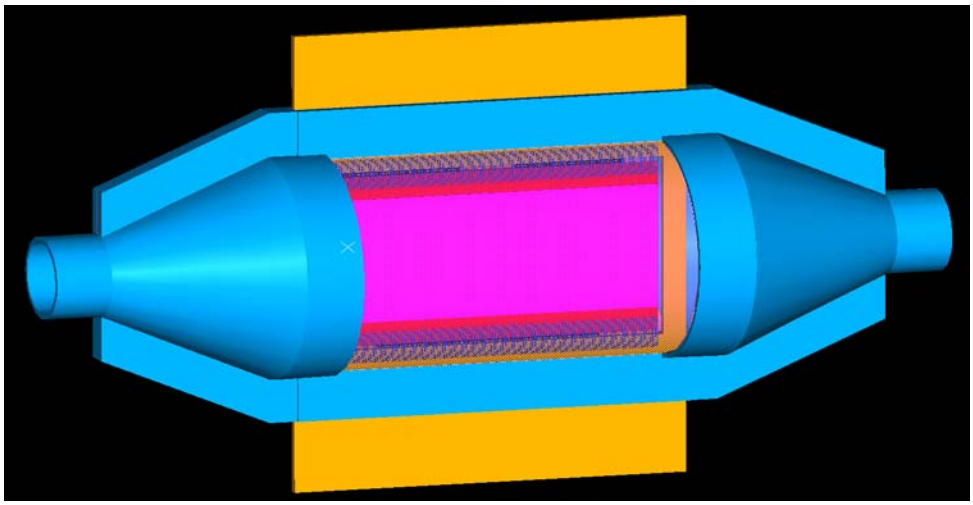

Figure 79 Cooling vessel for the GTK module.

The detector assembly is electrically connected to assembly carrier printed circuit board, which provides power and signal connections. However, the assembly is thermally decoupled from the assembly carrier board; the sliding support compensates for the different thermal expansion coefficients of the materials.

[4] Review of Particle Physics, http://pdg.lbl.gov/2010/AtomicNuclearProperties

The thermal analysis indicates a good uniformity of the temperature distribution across the detector area with a max value of 4 °C corresponding to a convection coefficient of 44 W/m^{2}K, see Figure 80.

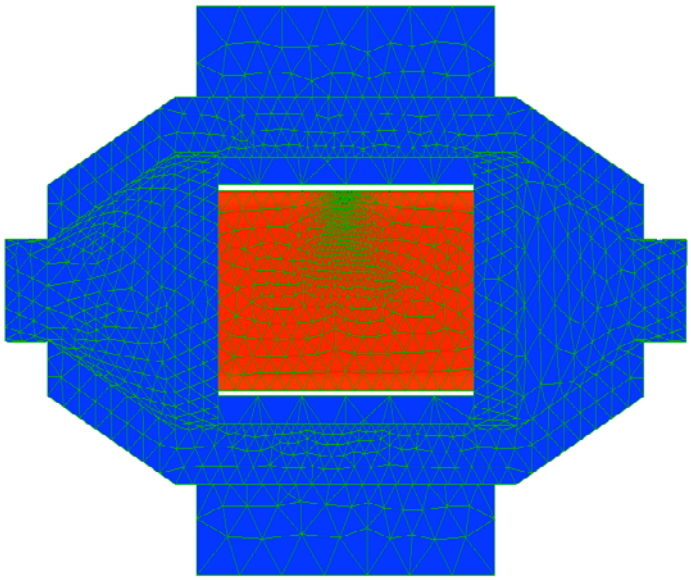

Figure 80 Temperature distribution as from simulation.

A structural analysis has been performed and takes the inner pressure due to the vacuum and the pumping system, the different elongation of the materials and the fixed constraints into account. It gives low mechanical stress values and displacements both for the vessel and the detector (< 2 MPa), see Figure 81.

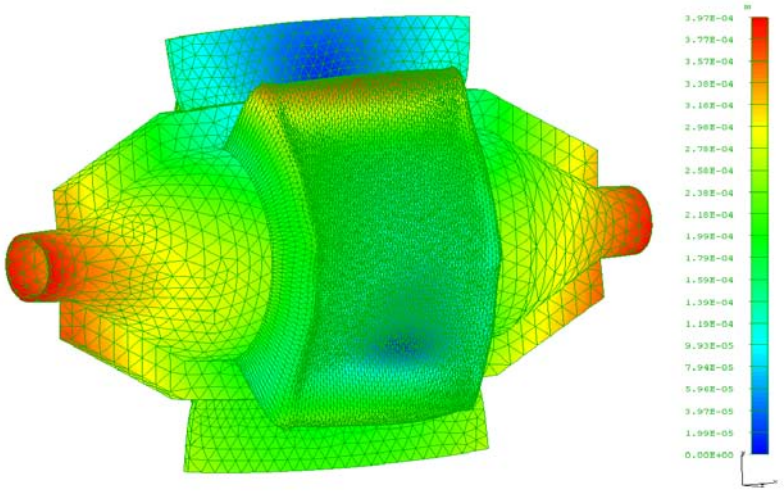

Figure 81 Map of thermal stress.

1.1.8.2.2 Design and Construction of a Full Scale Prototype

Following the results coming from the simulations, a full scale prototype has been designed and built in October 2009. Figure 82 shows the assembled prototype.

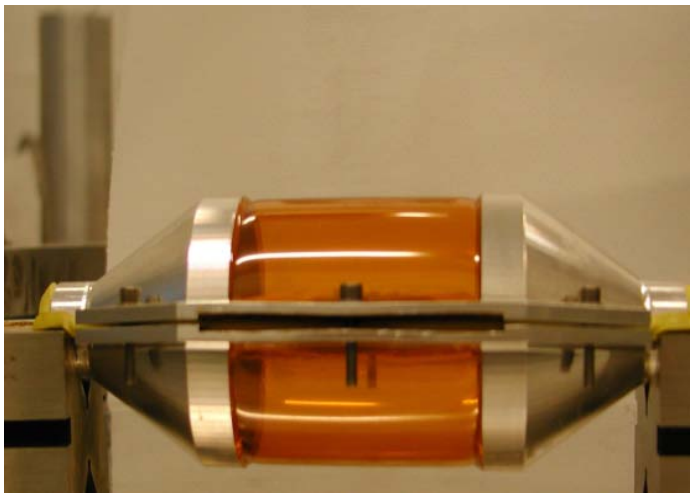

Figure 82 Full scale prototype.

The materials chosen for the construction are listed in Table 8.

Table 8 Material properties.

PART	MATERIAL	DENSITY (Kg/m^3)	THERMAL COND. (W/mK)	CTE (K^{-1})	ELASTIC MOD. (Pa)	POISSON'S RATIO
PCB	GLASS FIBER	2076	0,83	6,7E-6	45E+9	0,28
VESSEL WALL	KAPTON TYPE H	1420	0,16	45E-6	2,5E+9	0,34
VESSEL FRAME	ALUMINUM	2700	237	23,5E-6	70,6E+9	0,345
RESIN	EPOXY (ARALDITE 2012)	1180	0,22	30E-6	2,5E+9	0,34

The prototype was used to compare the results coming from the simulation to a full scale model working at the nominal operation parameters.

The following tests were done:

- structural test
- cooling test.

1.1.8.2.3 The Structural Test

The structural test concerned the pressure inside the vessel and the aluminum – kapton joint .

Kapton failure pressure test

The simulation and the standard calculations give a breaking pressure of 0,49 MPa (4,9 bar) for the 50 µm thickness kapton wall. A test, checking the failure pressure of the vessel, was set up, see Figure 83. The results of the test gave a breaking load of 0,52 MPa (5,2 bar) in a good agreement with the calculation. Considering the pressure inside the vessel of 2 bar given by the vacuum and the pumping system, the safety factor is 2.5.

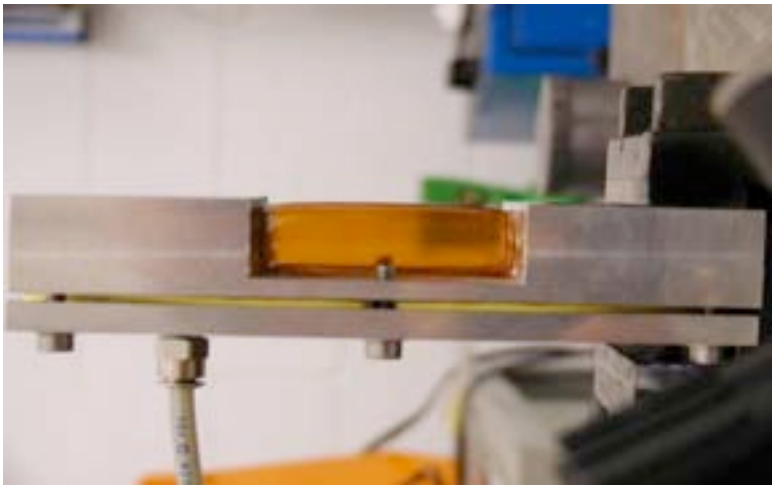

Figure 83 The Kapton failure pressure setup.

Quality of the joint Kapton – resin – aluminum

The overlap of the joint Kapton – aluminum has been designed to withstand a stress 2 times the breaking load of the kapton; the safety factor for the resin is 5. Fig. 6 shows the good behaviour of the joint after the breaking test.

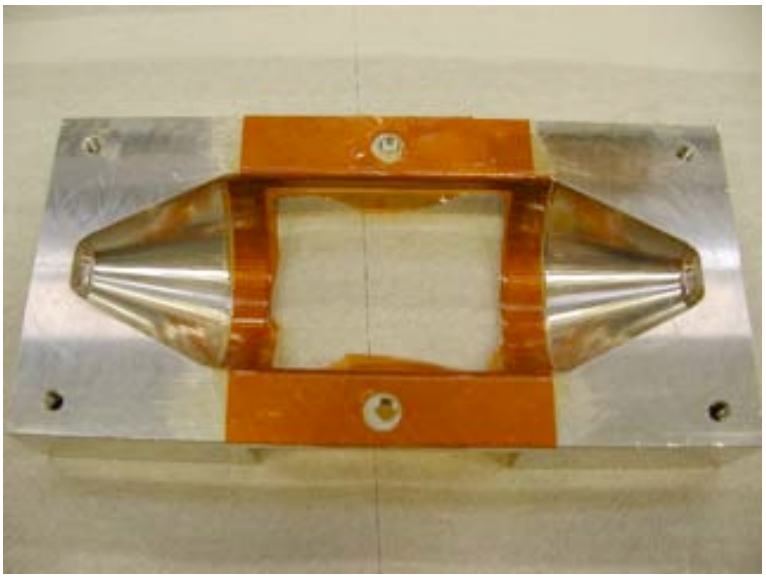

Figure 84 The Kapton - resin - aluminum joint.

1.1.8.2.4 The Cooling Test

Following the good results coming from the structural tests, a cooling system (see Figure 85, Figure 86 and Figure 87) was setup in February 2010.

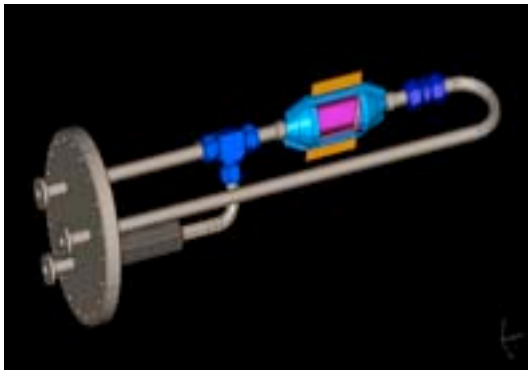

Figure 85 Cooling circuit design.

Figure 86 Cooling circuit implementation.

The system parts are listed below:

- vacuum chamber housing the prototype; pressure inside is 10^{-6} mbar;
- vacuum pumping system : scroll = 5 m^3/h and turbo 70 l/s;
- power and thermocouples feed-through;
- flowmeter;
- nitrogen in/out circuit

The test was limited by the maximum nitrogen flow rate of 7 m^3/h allowed by the pumping system.

In the test the detector was simulated, see Figure 87, by 16 resistors 222 Ω - 2 W with a total power of 32 W. The temperature was measured by 3 K-type thermo couplers mounted on the detector. A data monitor recorded the temperatures and the flow rate during the test.

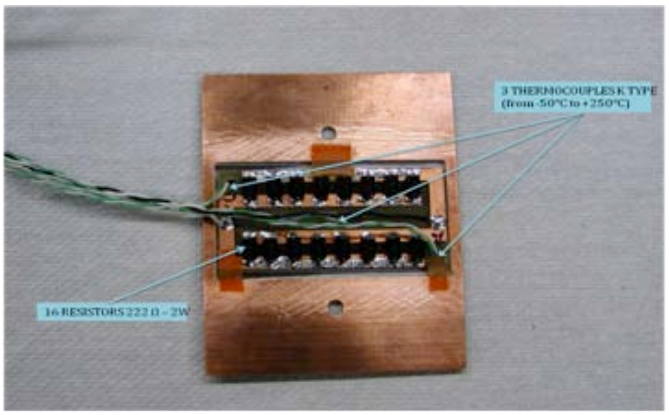

Figure 87 Detector mock up.

1.1.8.2.5 Cooling Results

The cooling results of the test are shown in Table 9, where

- P (W) is the power set on the detector
- Fr (m^3/h) is the nitrogen flow rate
- Tin (°C) is the measured temperature on the detector from the nitrogen inlet
- Tmid (°C) is the measured temperature on the middle of the detector
- Tout (°C) is the measured temperature on the detector from the nitrogen outlet
- Taverage (°C) is the average of the measured temperatures

Table 9 Measured temperature versus power & flow rate.

P (W)	Fr (m^3/h)	Tin (°C)	Tmid (°C)	Tout (°C)	Taverage (°C)
0,14	2,2	13,8	14,2	14,1	14,03
0,56	2,2	19,3	20,3	19,7	19,77
1,25	2,2	24,6	26,4	25,2	25,40
1,25	3,7	16,8	18,8	17,7	17,77
1,25	7,2	-17,2	-16,1	-16,6	-16,6
3,48	3,7	27,4	32,4	29,6	29,80
3,48	7,2	-13,8	-10,0	-11,90	-11,9
13,9	3,7	74,3	90,6	79,2	81,40
13,9	7,2	16,3	19,9	17,7	18

The results show a quite good uniformity of the temperature, with the maximum in the middle zone. The 7 m^3/h flow rate limited the power dissipation at 14 W. After the first test all the parts were checked and there was no evidence of damages due to the temperature excursion (from -17 °C to 80 °C).

In a second test in March 2010 the cross section of the cylindrical vessel was reduced of a factor 3 by adding two inner parallel flat walls at a distance of 12 mm. The amount of material was the same: flat walls thickness 10 µm and round walls thickness 40 µm, with a breaking safety factor 2.

The aim was to increase the thermal exchange at a reduced flow rate, improving the laminar flow and to compare these results with those of the first test.

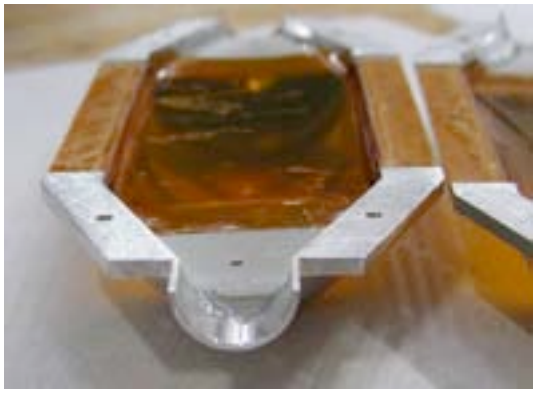

Figure 88 Flat Kapton wall on half of the vessel.

The cooling results of the test are shown in Table 10, where the symbols are the same as described for Table 9.

Table 10 Measured temperature versus power & flow rate

P (W)	Fr (m^3/h)	Tin (°C)	Tmid (°C)	Tout (°C)	Taverage (°C)
3,5	3,3	-	19,2	20,1	19,7
8,9	5,1	-	13,2	14	13,6
20	5,3	-	53,2	55,3	54,2
31,3	7,2	-	22,5	23,5	23,0

Figure 88 shows one flat wall glued on half of the vessel . The pressure between the flat walls and the round walls was balanced by two holes connecting the volumes.

The Tin is not reported because the first thermocoupler was not thermally connected to the detector.

The temperature shows a quite good uniformity and the 7m3/h flow rate allows the operation of the detector at nominal power.

Comparing the results of the two tests the reduced cross section given by the flat walls improve the cooling, increases the flow speed and reduces the needed flow rate.

1.1.9 Electro-Mechanical Integration

Each of the three GTK stations will be installed inside a dedicated vacuum vessel at three different positions along the NA62 experiment beam line. The vessels for GTK1 and GTK2 are identical and are 292 mm x 239 mm x 190 mm in size, whereas GTK3 is 550 mm x 560 mm x 2000 mm as it contains the CHANTI detector planes[5] as well, see Figure 89 shows the vessel design for GTK1 and GTK2.

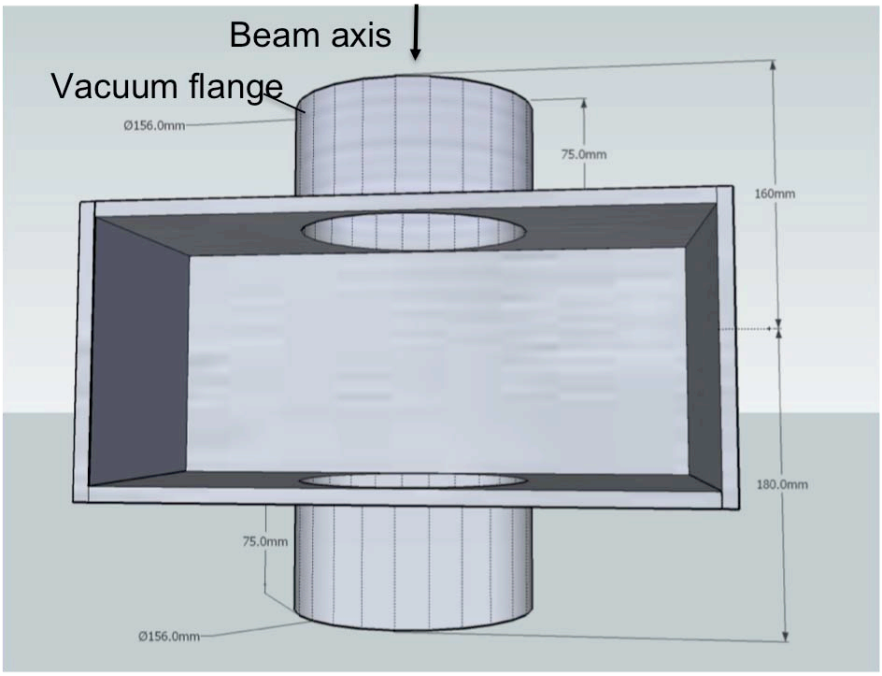

Figure 89 GTK 1 and 2 vessel.

The vessels will be fixed and precisely aligned on the beam axis. The vessel for GTK3 and the CHANTI will be fixed in place whereas GTK1 and 2 vessels are mounted on an xy-table for automated alignment to GTK3 and the beam axis. Connections with the beam pipe are made using bellows allowing the vessels to move.

Due to the radiation damage in the sensor, the modules need to be replaced on a regular basis, see chapter 1.1.3.1. The electro-mechanical integration is based on the principle that one GTK module – the assembly and its services - is an integral and compact component, which can be inserted and removed with a minimum intervention in the vessel. The Gigatracker assembly carrier is a printed circuit board which supports mechanically the assembly and provides all ASIC and sensor power supplies, control and read-out signal connections on differential signals and carries optical

[5] The CHANTI detector layout is described in section 2.4.2.

components outside the vacuum vessel and acts as vacuum feed-through for the electrical lines and the cooling connections. Figure 90 shows a drawing of the GTK assembly carrier for the micro-channel cooling option. On the left hand side the PCB has an opening for the sensor assembly and the micro-channel cooling plate. In the middle a flange with an O-ring is tightly glued around the PCB acting as vacuum feed-through. On the right hand side optical components, already outside the vacuum, are placed.

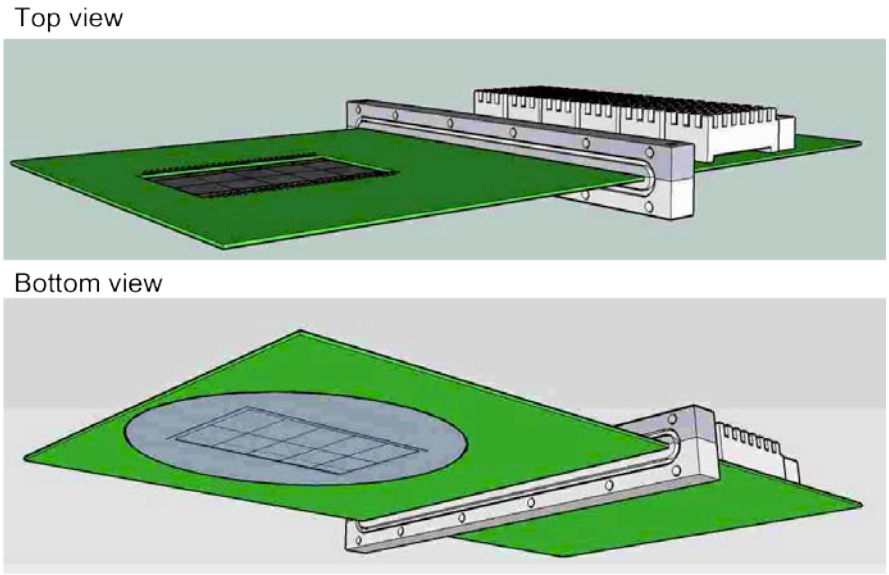

Figure 90 GTK assembly carrier.

Mechanical guides assure precise alignment of the assembly carrier and thus the silicon assembly with the vessel. Figure 91 shows the vacuum feed-though and the positioning pins. Figure 92 shows the module inserted in the vessel.

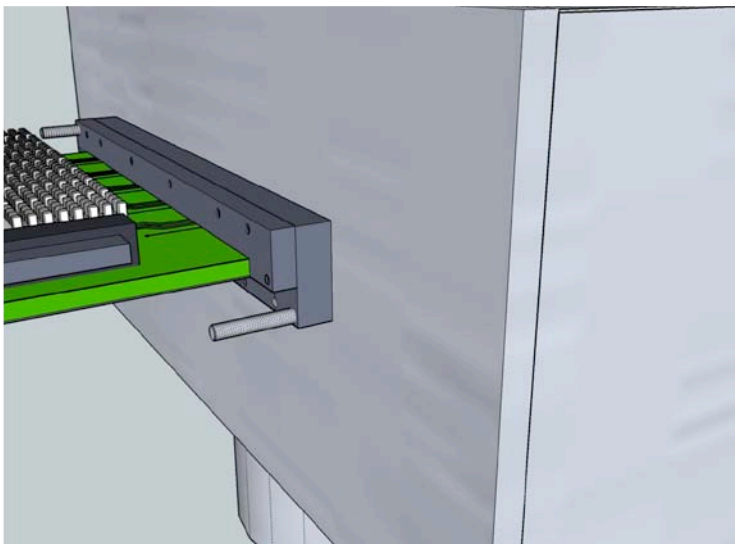

Figure 91 Vaccum feed-through and positioning pins.

Wire bonds provide the electrical connections between the ASIC wire bond pads, the sensor bias connections and the GTK assembly carrier. Each ASIC requires three different power supplies 1.2 V digital, 1.2 V analogue and 2.5 V I/O. In order to minimize the electrical interferences, each ASIC will be powered individually with separate power supplies. Measurements will show whether the analogue and digital supply of an individual chip can be combined on the assembly carrier. For the optical components and clock distribution buffers the card is connected to 3.3 V. This means a total of up to 31 power supply connections for each GTK interface board are required. In order to avoid the need of a precise pitch adapter for the wire bond connection to the printed circuit board, a staggered wire bonding scheme is proposed where the pitch in the PCB does not need to match the ASIC pitch of 73 µm. Figure 93 illustrates the principle where the power connections and the signal connections go to different rows on the PCB.

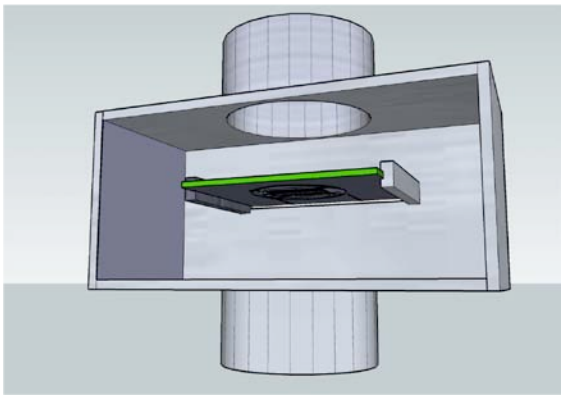

Figure 92 GTK module inserted in the vessel.

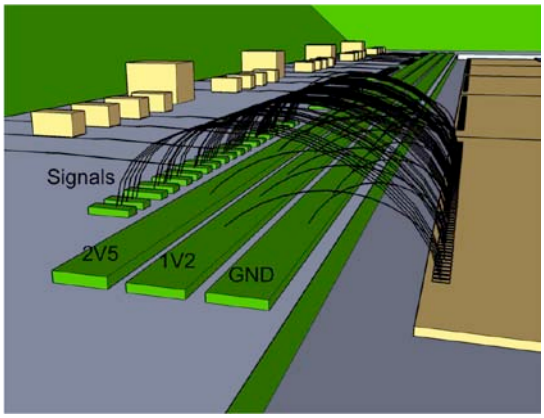

Figure 93 Staggered wire bonding scheme avoids pitch adapter.

Figure 94 shows a cross section of the module. The FR4 acts as rigidifier as well as carrier for the printed circuit. The figure also shows the micro-cooling connection (Nano Port). This design connects the electrical connection on the top side, whereas the cooling is connected on the bottom side avoiding mechanical interference.

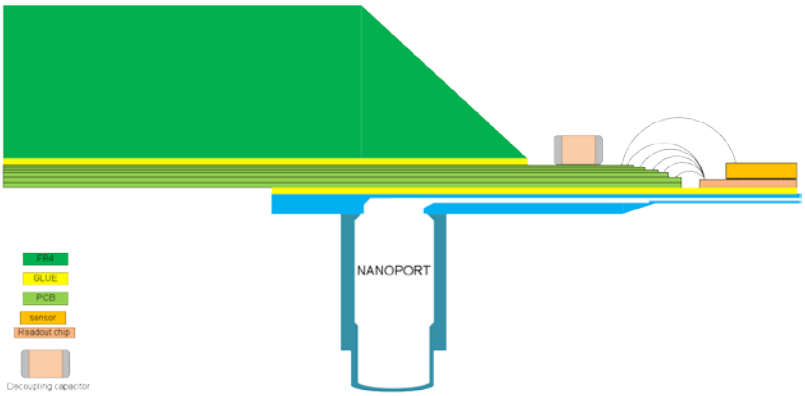

Figure 94 Cross section of GTK module.

Each of the 10 ASICs sends 6Gb/s data on differential lines to the optical components on the assembly carrier. The TDCs inside the ASICs require a low clock jitter signal. The electrical layout of the PCB is driven by these two constraints and requires a study and simulation of signal integrity and

special attention in the routing. Prototype GTK assembly carriers are going to be produced in order to optimize this layout.

Due to high variation of temperature in the cooled operation mode and in the un-cooled stand-by mode, the silicon components are not glued to the GTK assembly carrier, but are placed in a precisely machined opening on the GTK assembly carrier. Carbon fiber clips keep the silicon component in position (see Figure 95) and avoid mechanical stress due to different thermal expansion coefficients of silicon and FR4.

The design of the assembly carrier is carried out so that it is compatible with both cooling options. An integration design has been conducted for the gas cooling option. Figure 96 shows the design.

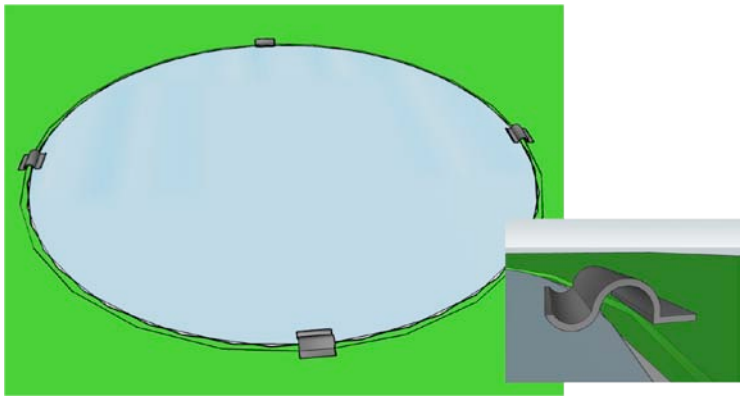

Figure 95 Carbon fiber clips.

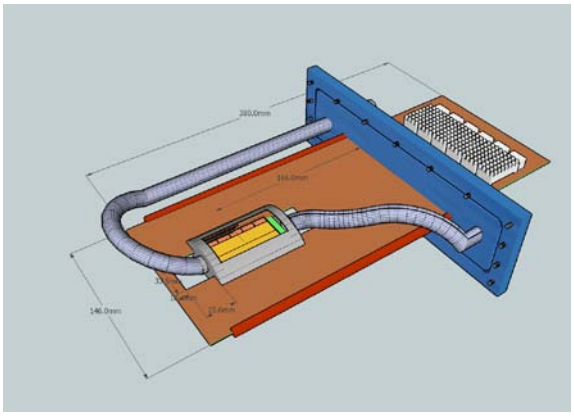

Figure 96 Gas cooling integration.

1.1.10 GTK Production, Test & Assembly

In the context of the GTK project several tasks are challenging and need dedicated research for production and assembly optimisation. In short these tasks are:

- Design and test of the GTK ASIC
- Assembly bump bonding and thinning.
- Design, optimisation, production and test of the cooling system
- Integration of GTK module – assembly of sensor-ASIC assembly with cooling elements.

- Integration of GTK module – assembly of silicon assembly with GTK assembly carrier.

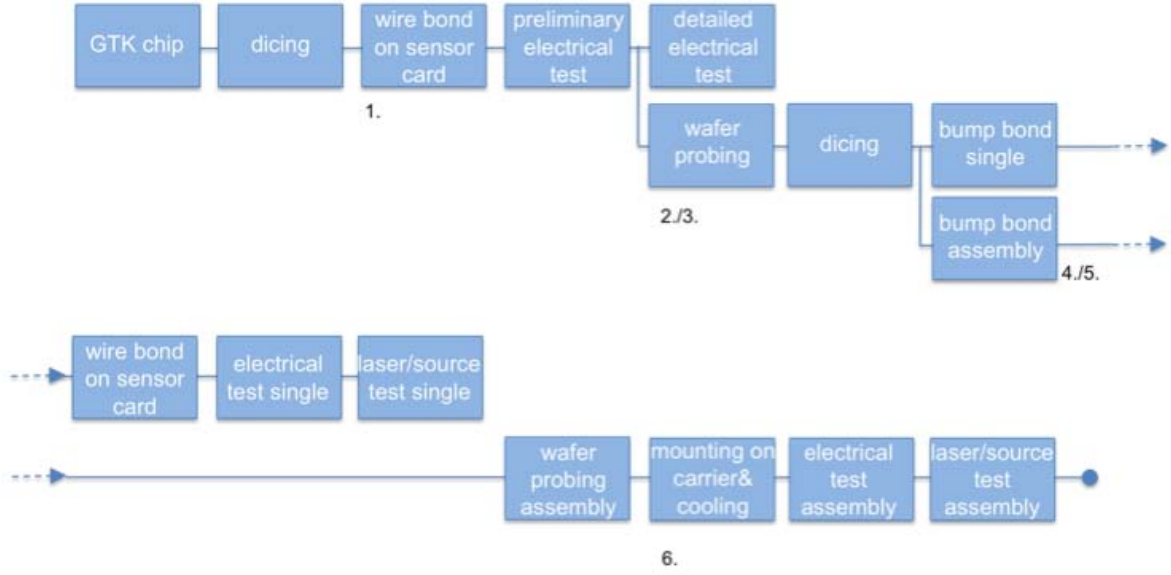

Figure 97 GTK test& production flow chart.

In order to streamline the R&D phase and the production phase in terms of completion time it is mandatory that the tasks listed be tackled in parallel. Consequently for the optimisation process on the integration the final GTK assembly (sensor/ASIC) is not yet available and mechanical and thermal equivalent dummy structures are produced to elaborate a production process. Integrated heating elements mirroring precisely the dissipation on the final ASIC on the dummy structures allows the cooling qualification. The same dummy structures are also used to emulate the electrical behaviour of the ASIC by connecting high speed drivers to it. Thus they also will be used to electrically qualify the GTK assembly carrier PCB. Once the GTK assembly is available the production process is optimized and allows a stream-lined integration.

Figure 97 shows the qualification chain flow chart once the GTK ASIC has been produced. The flow diagram shows several required parallel tasks and assumes the following points (as noted in the figure):

1. Sensor card, where the single chip sensor/ASIC assembly is glued and wire bonded is available when ASIC is ready.
2. Wafer prober has been setup in advance including mechanics, electrical test systems and firmware/software.
3. Wafer probing is run in parallel without the need of the expert performing the detailed tests.
4. Bump bonding & thinning understood in advance using dummy assemblies.
5. Bump bonding of single & assembly is done in parallel.
6. Mechanics assembly & integration with GTK assembly carrier and cooling verified and optimized with dummy assemblies before.

1.1.11 DAQ - Off Detector Readout

1.1.11.1 Introduction

The aim of the off detector readout system is to provide the interface to the on detector electronics, the DAQ system, the trigger system and the NA62/GTK detector control system (DCS). Figure 3 on page 5 shows a full GTK block diagram. The following sections give an introduction to the system and its interface. Section **Error! Reference source not found.** on page **Error! Bookmark not defined.** in the Trigger and DAQ chapter give a detailed description of the readout.

1.1.11.1.1 Detector Interface:

In particular the system is connected to the 3 GTK stations via optical fibers. For each of the 10 chips in a station 2 – 4 multi gigabit serial data links totalling a data rate of 6 Gb/s and chip and one slow status data link are foreseen. In the direction to the detector one clock link and one slow serial configuration link per chip is foreseen. The GTK off detector read-out system reads the raw data stream from the high speed serial links, stores the entire data flow until the L0 trigger decision and selects all data in a time window of 75 ns around the event indicated by the L0 decision which arrives with a latency of 1 ms. In an upgrade phase of the trigger system the latency might go up to several ms. The design of the GTK off-detector read-out will take this possibility already into account.

Serial status information is accepted by the off detector read-out for each chip individually. It continuously evaluates and forwards them to the GTK DCS. Serial status information contain temperature values, acknowledge of commands sent on the serial link and read-back of configurations. Furthermore both the data stream and the status data are monitored for its data consistency in the off detector read-out FPGA processor and consequently by the GTK DCS.

The low jitter clock (< 20 ps jitter) is sent to each chip individually on equal length fibers. The serial configurations are sent to each chip individually allowing to address each chip independently. The serial data contains information such as global and local pixel thresholds, programmable read-out parameters and programmable bias circuitry values.

In total the number of links connected to the off-detector read-out is: 60 or 120 gigabit serial links for a total data rate of more then 120 Gb/s, 30 status links, 30 clock links and 30 configuration links.

1.1.11.1.2 Trigger Interface:

The GTK off detector electronics expects a constant latency trigger signal. It is planned to use the TTC protocol to transmit trigger information. As the GTK off-detector system is placed in the control room and thus not subject to radiation the TTCrx chip can be replaced by an FPGA if it turns out to be more practical. In case of a non-constant latency L0 trigger signal a time tag synchronized with the GTK time domain must be attached to the trigger signal allowing extraction of the corresponding time slice from the GTK read-out memory. Two signals are foreseen to communicate from the GTK read-out to the NA62 trigger system. The throttle signal indicates to the NA62 trigger system that the present trigger rate is too high and buffers are going to overflow. The error signal indicates that

data loss already has occurred and the GTK read-out system attempts to recover process integrity and/or needs a reset.

1.1.11.1.3 Detector Control Interface:

The GTK off-detector read-out is connected to the GTK PC based DCS system via a dedicated bus system. The GTK DCS initiates via the GTK off-detector read-out the downloading and verification of configuration as well as online verification of parameters and data flow.

1.1.11.1.4 Safety Interlock Interface:

The GTK off-detector read-out system is directly connected to the safety interlock without passing through a PC. In case of a detection of a temperature deviation outside of the programmable limits the read-out interacts with the power supplies via the safety interlock feature.

1.1.11.1.5 DAQ Interface:

The GTK off-detector read-out system is connected to the DAQ system with Gigabit Ethernet links pushing data upon a L0 trigger decision. Two Gigabit Ethernet links per GTK read-out chip are foreseen resulting in 60 Gigabit Ethernet links for the GTK.

1.1.12 GTK System Components, Auxiliaries

Figure 3 shows the full GTK block diagram.

The three GTK stations are located within few meters from each other and the radiation effects harming commercial electronics is negligible already a few 10 cm from the beam area. Consequently low voltage power supply modules for the read-out chip, the HV modules for the sensors, the safety interlock system and the environmental monitoring will be located in a service area next to the GTK some 10 m distant from the stations. Control busses to the units allow remote configuration and monitoring of voltage, currents and environmental parameters from the control room. The magnetic field in the GTK service area is expected to be negligible.

1.1.12.1 Low Voltage Supply

Each individual read-out chip is supplied individually. The read-out chip requires 2.5V (< 1A) and 1.2V (< 3A). Thus for all three stations 30 2.5V and 30 1.2 V LV power channels are required. One channel is connected to one read-out chip with two remote sense lines connected directly on the module allowing to controlling each read-out chip individually. The common ground connection of one station is established directly in the station. Consequently the power supply modules provide floating outputs. Each channel can be set and monitored with a resolution of 10 mV. The current limit and the current monitoring have a resolution of 10 mA with a sampling period of at least 500 ms. The modules can be disabled by the safety interlock system. In case of over voltage or over current the modules switch off automatically using a hardwired internal interlock system. The modules have a hardware adjustable voltage and current limit, which cannot be overridden by software.

1.1.12.2 High Voltage Supply

Each of the sensors in the three stations is individually supplied with bias voltage. The voltage can be adjusted from 0 V to 700 V with a resolution of 1 V. Current limits are set with a resolution of 100 nA with a maximum of 500 µA per sensor station. The actual sensor leakage current is monitored continuously and included in the GTK interlock system. The leakage current provides a measure of the accumulated radiation damage and the temperature of the sensor. Thus monitoring the current on a long term basis provides feedback on the radiation induced performance degradation and detects a potential need for retuning of the analogue front-end. Monitoring the current on short term provides temperature information, which is fed to the safety interlock system in addition to measurements of on-chip temperature sensors. The current monitoring resolution is around 100 µA with a sampling period of about 500 ms. The modules can be disabled by the safety interlock system. The modules have a hardware adjustable voltage and current limit, which cannot be overridden by software.

1.1.12.3 Safety Interlock

Operation of the stations in a controlled environment is very important. As the modules have little mass they need to be prevented from thermal run-away in case of a cooling failure. This need is accentuated as the modules are operated in vacuum at a temperature of 5 degree Celsius or below and there is no cooling effect via convection. In case the thermal gas cooling option (section 1.1.8.1) is chosen during power supply failure the interlock must act on the cooling as a thermal run-away to cold temperatures can destroy the module as well. Each read-out ASIC contains thermal sensors which are accessible via 2 pads and in addition are read-out via the status link of the ASIC. The direct hard-wired reading of the thermal sensor is also operative when the ASIC is not powered. Humidity sensors are placed close to the stations.

The core of GTK interlock system is a fully hardware based system with a reaction of 1 s and it acts on the power supplies of one chip individually. For each read-out chip it takes the hardwired temperature sensor and the humidity sensor into account on a real time basis. Via the status bus of each chip and the read-out electronics, the temperature values are fed into the read-out processor, which also have one hardwired output for each read-out chip. Each of these wires is also taken into account by the interlock system. This configuration allows a safe and from software process independent operation, as direct hard wired connections to the sensors and hardware allow a fast interlock action.

Although the interlock system operates independently from any computer it is connected to a computer based detector control system (DCS) which logs the environmental parameters, such as temperature, from the hardwired sensors, temperature measurement from the read-out stream, humidity values, low and high voltage power supply voltage and current and cooling system parameters.

1.1.12.4 Detector Control System

The detector control system is a computer based system which:

- sets the low voltage voltages and current limits on each individual chip;
- sets the high voltage, current limits and initiates the voltage ramp up/down;
- configures the read-out ASIC front-end configuration (global discriminator thresholds, individual discriminator thresholds, enables/disables individual pixels, set trim DACs);
- configures programmable parts in the read-out section of the ASIC;
- configures programmable parts in the off-detector read-out electronics;
- monitors the status of the detector elements (environmental parameters, data consistency) regularly and reads back configuration registers;
- configures and sends an input to the interlock system;
- logs permanently the system parameters; and
- issues warnings and alerts.

Bibliography

1. **Gigatracker Working Group.** The GIGATRACKER: Addendum to the NA62 proposal. *Internal Note NA62-07-08.* Nov. 2007.

2. **Fiorini, M., et al.** Test of Silicon Sensors for a High Rate Pixel Detector for the NA62 experiment. *Internal Note NA62-08-01 .* 2008.

3. **Anelli et al.** Optimum Segmentation and Thickness of Silicon Pixel Detectors for Signal to Noise Ratio and Timing Resolution. *Nuclear Science Symposium Conference Record IEEE, Vol 2.* 2006, pp. 671-680.

4. *Single Event Effects in 130 nm, CMOS technologies.* **Faccio, F., Cervelli, G. and Marchioro, A.** 2005, LNL annual report 2005, ISSN 1828-8545, 70 . LNL Annual Report, ISSN 1828-8545, 70.

5. *An ultra fast 100ps, 100micron 3D-pixel imager.* **A. Kluge et al.** San Jose, USA : SPIE, 2009. Proceedings of the 2008 SPIE conference. Volume 7249, 724909.

6. *Pixel Read-Out Architectures for the NA62 GigaTracker,.* **P. Jarron et al.** Naxos (Greece) : Heliotopos Conferences, 2008. TWEPP08.

7. *A High-Resolution Time Interpolator Basedon a Delay Locked Loop.* **Christiansen, J. and Mota, M.** 10, Oct. 1999, IEEE Journal of Solid-State Circuits, Vol. 34, pp. 1360-1366.

8. *A multi-channel 24.4 ps bin size Time-to-Digital Converter for HEP applications.* **Mester, C.** Naxos, Greece : Heliotopos Conferences, 2008. TWEPP08. TWEPP 2008-09-26.

9. **Ho, R.** Efficient on-chip global interconnects. *Symp. VLSI Circuits.* 2003.

10. **Bashirullah, R.** A hybrid current/voltage mode on-chip signaling scheme with adaptative bandwidth capability. *IEEE Trans. VLSI.* 2004.

11. *Silicon fusion bonding for fabrication of sensors, actuators and microstructures.* **Barth, P.** 1-3, 1990, Sens. Act. A,, Vol. 23, pp. 919-926.

12. **J. Thome et al.** CMOSAIC: 3D Stacked Architectures with Interlayer Cooling. [Online] http://www.nano-tera.ch/projects/67.php.

13. *Two-Phase Cooling of 3D Chip Stacks.* **Y. Temiz et al.,.** Lausanne : EPFL, 2010. Proceedings. of the EPFL-CMI MicroNanoFabrication Annual Review Meeting,.

14. *The fabrication of all-silicon micro gas chromatography columns using gold diffusion eutectic bonding.* **A. Radadia et al.,.** 015002, 2010, J. Micromech. Microeng., Vol. 20.

15. *Low material budget microfabricated cooling devices for particle detectors and front-end electronics.* **A. Mapelli et al.** 2010, to be published in Nucl. Phys. B,.

16. *Field Assisted Glass-Metal Sealing.* **Wallis, G. and Pomerantz, D.** 10, 1969, J. Appl. Phys, Vol. 40, pp. 3946-3949.

17. **S. Ilie et al.** *Chemical and radiolytical characterization of perfluorocarbon fluids used as coolant for LHC experiments.* Geneva : s.n. CERN - EDMS doc N°842110.

18. **Bruus, H.** *Theoretical Microfluidics.* USA : Oxford University Press, 2008.

19. *Effect of Manifold Design on Flow Distribution in Parallel Micro-Channels.* **Webb, R.L.** Maui, Hawai : ASME, 2003. International Electronic Packaging Technical Conference and Exhibition.