A novel sub-10 ps resolution TDC for CMOS SPAD array

Vincenzo Sesta, Federica Villa, Enrico Conca, Alberto Tosi
Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano
Milan, Italy
vincenzo.sesta@polimi.it; alberto.tosi@polimi.it

***Abstract*—In this work, we present a novel Time-to-Digital Converter (TDC) for single-chip integration in Single-Photon Avalanche-Diode (SPAD) array and digital Silicon Photomultiplier (SiPM). Such novel detector-timing electronics combination will be suitable for Time-Correlated Single-Photon Counting (TCSPC) applications and direct Time-Of-Flight (TOF) measurements. The proposed TDC is based on a 200 MHz 4-bit counter that guarantees a Full-Scale Range of 80 ns. Two interpolators exploit the sliding scale technique to reduce the Differential Non-Linearity (DNL). Besides the coarse interpolation, the multi-stage interpolators have a novel dual-fine interpolation that guarantees a resolution as good as 7 ps, with a conversion time (< 50 ns) much shorter compared to typical architectures based on Vernier delay lines.**

Keywords—Singlr-Photon Avalanche Diode (SPAD), time interval measurement, Time-to-Digital Converter (TDC), Vernier delay line

I. INTRODUCTION

A Time-to-Digital Converter (TDC) is an electronic circuit capable of converting a time interval in a digital word. A time interval is usually defined by two electrical pulses: the START pulse indicating the starting point of a time interval and the STOP pulse indicating its end. Such components are essential in systems where single-photon arrival times must be measured with a timing resolution in the order of picoseconds. Direct time-of-flight (TOF) LIDARs [1], [2], [8] high-energy physics experiments [3], positron-emission tomography scanners [4] are only few examples of such systems. For most of these applications, the Time-Correlated Single Photon Counting (TCSPC) technique [5] is the best choice for reconstructing the optical waveforms. At its core, besides the light source and the single-photon detector, there is a TDC with very high timing resolution.

In this work, we present a novel TDC suitable for development of array of SPADs in TCSPC applications, but also for direct TOF LIDAR. The converter will be integrated together with a Single-Photon Avalanche-Diode (SPAD), a detector with single photon sensitivity and low timing jitter (few tens of picoseconds). The TDC is currently in fabrication in a 0.35 μm CMOS technology, since SPADs with high performance can be designed in this technology node [6]. The main TDC specifications were set by TCSPC technique: resolution of few picoseconds, differential non-linearity (DNL) lower than few percent of the Least Significant Bit (LSB) and conversion time shorter than 100 ns, to reach high conversion rate.

The proposed TDC has a fully-differential architecture to reduce disturbances and noise. Two stages of interpolation are used for having high resolution, while high linearity is achieved thanks to the "Sliding Scale technique" [7]. A new dual-fine interpolation is used to reduce conversion time without affecting the resolution. Full-scale range is 80 ns and a resolution as fine as 7 ps, with conversion time shorter than 50 ns.

The paper is organized as follows: Section II illustrates the single-photon avalanche diode arrays. Section III describes the TDC architecture. Section IV draws conclusions and perspectives.

II. SINGLE-PHOTON AVALANCHE-DIODE ARRAYS

Various example of CMOS SPAD arrays has been developed and reported in literature. Some of them are either composed by a relatively low number of pixels operated in parallel or by quite a large number of detectors that are in any case multiplexed [2]. The former arrays generally lack in compactness while the latter ones do not exploit the full parallelism and high frame-rate of SPADs. Moreover, those arrays are generally tailored to perform either photon "counting" or "timing" mode.

Many of them have been developed specifically for TCSPC or direct TOF applications [1], [2], [8]. These SPAD arrays are composed by a few number of pixels in fully-parallel architecture in which the TDC and sensing electronic is integrated together with each SPAD. The in-pixel TDC has excellent performance in terms of timing resolution (few tens of picoseconds) and linearity (DNL lower than few percent of LSB) because are defined by the specified applications. However, the excellent precision comes at expense in silicon area, reducing the overall active area, despite the considerable number of pixel.

In this work, we design a novel TDC suitable in SPAD array employed in these applications that require best-in-class performance in terms of timing resolution (LSB) and linearity. The next section describes in detail the TDC architecture.

III. TDC ARCHITECTURE

The TDC was designed to achieve picoseconds resolution and tens of nanoseconds full-scale range. Such ranges could be achieved through interpolation methods based for instance on Pulse-Shrinking delay line, Tapped delay line, or standard and

cyclic Vernier delay line. In all those methods, linearity is limited by components mismatches, but it can be greatly improved by employing the "Sliding Scale technique", which requires to separately measure the time interval between a reference clock and both asynchronous START and STOP signals. In this way, even if the same START-STOP time interval is converted, different portions of START and STOP interpolator ranges are exploited, thus interpolators' deterministic nonlinearities are converted in stochastic jitter. Therefore, the linearity of the converter is considerably improved. However, the improvement of linearity is paid with a more complex structure (two interpolators instead of one) and a higher quantization error due to the difference between START and STOP interpolation measurements. In fact, the quantization error is:

$$\sigma_q = \sqrt{\frac{LSB_{START}^2}{12} + \frac{LSB_{STOP}^2}{12}} = \frac{LSB}{\sqrt{6}} \quad (1)$$

where LSB_{START} is the resolution of START interpolator, LSB_{STOP} is the resolution of STOP interpolator, which have the same value and represent the overall resolution of the TDC (LSB). The single-shot precision is also degraded by the integral non-linearity (INL) of the interpolators. The single-shot precision of a TDC based on counter and two interpolators can be written as:

$$\sigma_{single-shot} = \sqrt{\sigma_q^2 + \sigma_{INL_{START}}^2 + \sigma_{INL_{STOP}}^2 + \sigma_{ck}^2 + \sigma_{jitter}^2} \quad (2)$$

where σ_{INL_START} and σ_{INL_STOP} are the standard deviations of the respective INLs of the START and STOP interpolator, σ_{ck} is the reference clock jitter and σ_{jitter} is the additional jitter introduced within the TDC.

The proposed TDC architecture, whose block diagram is shown in Fig. 1, makes use of a 4-bit counter with a 200 MHz reference clock (clock period of 5 ns), thus obtaining a full-scale range (FSR) of 80 ns. Two interpolators are used to measure the delay between the START or STOP signals and the first rising edge of the reference clock to make use of the aforementioned "Sliding Scale technique". In a multi-channel implementation of the proposed TDC, multiple STOP and only one START channel suffice if the START event is common to all pixels. To reduce the area occupation and obtain both high resolution and short conversion time, both interpolators make use of two interpolation stages as shown in Fig. 2. The coarse stage uses 16 separate phases of the reference clock to detect the arrival time of the input in respect to them. The second stage uses a modified single-stage Vernier delay line to obtain a finer time resolution.

Fig. 3 illustrates the principle of operation of the proposed TDC. The counter output is proportional to $T_{counter}$, given by the number of clock periods between START and STOP. The START and STOP interpolators measure the time interval between the respective input signal and the reference clock rising edge. The first coarse interpolation measures the delay (T_{c_START} and T_{c_STOP}) between the first clock phase after the rising edge of the input signal and the successive reference clock rising edge. The time interval between the input and the

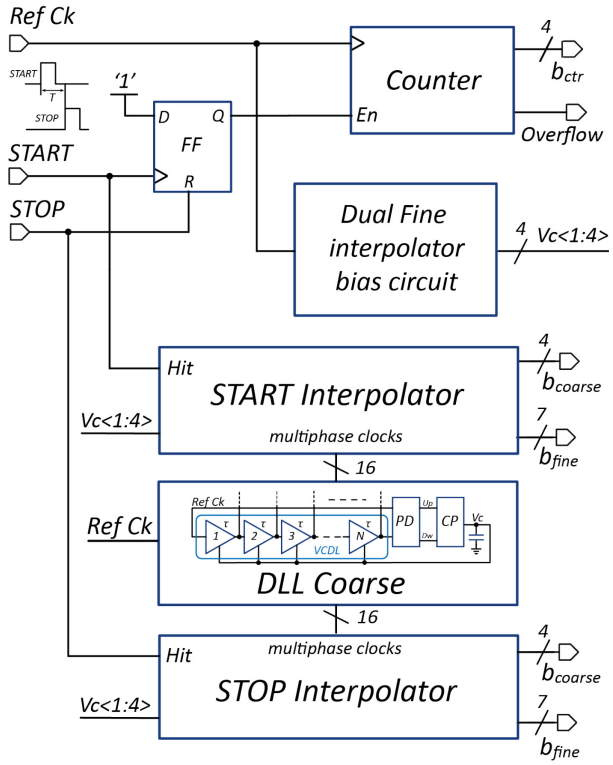

Fig. 1: Simplified TDC block diagram. DLL Coarse is used to generate multiphase clocks, and the dual fine interpolator bias circuit is used to properly bias the fine interpolators. A 4-bit coarse counter and two 11-bit START and STOP interpolators are used to measure time-intervals with long dynamic range and high precision. The overflow bit can be used to extend the FSR.

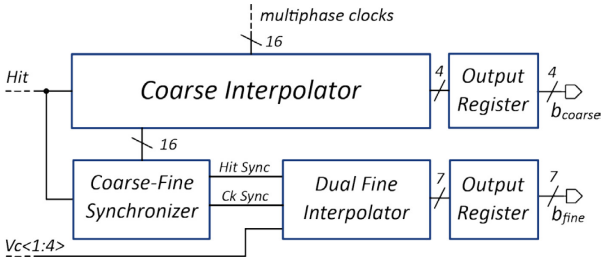

Fig. 2: START/STOP interpolator block diagram. The coarse interpolator measures the delay between the first clock phase after the hit signal and the reference clock rising edge. The synchronizer provides the input signal for the dual fine interpolator, reaching the high resolution of about 7 ps.

successive clock phase (T_{f_START} and T_{f_STOP}) is measured by the dual-fine interpolator as the difference in time between HIT SYNC and CK SYNC pulses, through a coarse-fine synchronizer. The overall time measurement T is obtained by summing the times measured by the counter and the START interpolator minus the delay measured by the STOP interpolator:

$$T = T_{counter} + (T_{c\_START} + T_{f\_START}) - (T_{c\_STOP} + T_{f\_STOP}) \quad (3)$$

A Coarse Delay-Locked Loop (DLL) is used to generate a multiphase clock, while a dual fine interpolator bias circuit, composed by four DLLs, is used to properly bias the dual fine interpolators to set the correct propagation delays and define the resolution of converter. The coarse and dual fine interpolator are described in detail in the following.

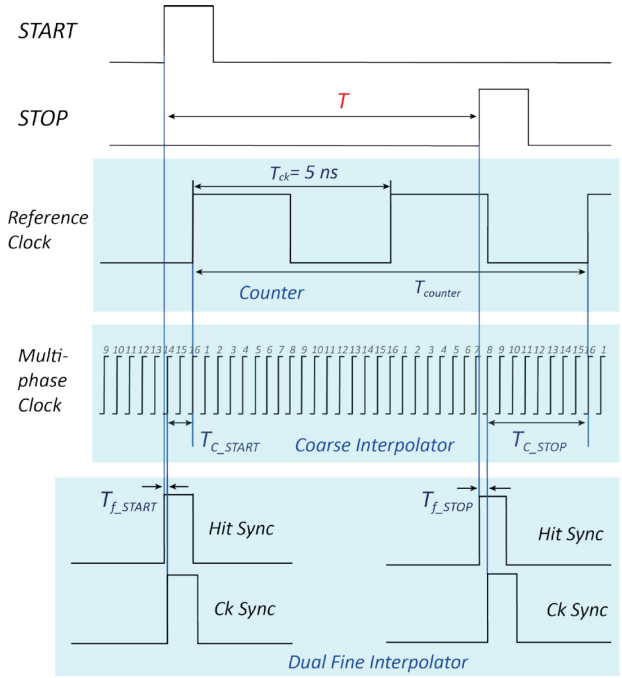

Fig. 3: TDC operation principle. A counter counts the cycles of a 200 MHz clock, between START and STOP events, while START and STOP interpolators resolve the START and STOP occurrence, respectively, within one clock period. Both interpolators use two stages of interpolation: coarse interpolator and dual fine interpolator.

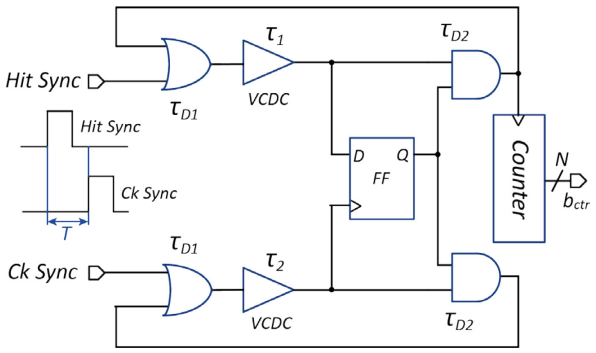

Fig. 4: Single-stage Vernier delay loop based on two loops with only one time-resolving element each, with consequent improvement of conversion linearity.

A. Coarse Interpolator

The coarse interpolator implements a DLL-based multiphase clock interpolation. A DLL uses a voltage-controlled delay line (VCDL), composed of 16 voltage-controlled delay cells (VCDC), through which the 200 MHz reference clock propagates. The total delay of delay line is locked to the reference clock thanks to a phase detector (PD) and a charge pump (CP) which sets the control voltage accordingly. The outputs of the 16 delay cells are tapped off to obtain a multiphase clock where every phase is ideally delayed from the previous one by a constant value equal to $T_{ck} / 16 = 312.5\ ps$ stabilized against PVT (process parameters, supply voltage and temperature) variations. An arbiter circuit is used to sample the state of the multiphase clock in correspondence of the HIT SYNC signal's rising edge; this thermometric value is then converted into a binary value and stored in a 4-bit output register.

B. Dual Fine Interpolator

To achieve a target resolution better than the intrinsic propagation delay of logic cells, a possible implementation is a single-stage Vernier delay line, whose resolution is defined by the difference between propagation delays of two cells, as shown in Fig. 4. This method is based on two delay loops in which the HIT SYNC and CK SYNC pulses cycle through. The propagation delay of the cell in the CK SYNC loop is fixed by the DLL to be shorter than the propagation delay of the cell in the HIT SYNC loop. The cycling is terminated as soon as HIT SYNC arrives later than CK SYNC. The result of conversion is given by the number of cycles needed to reach the termination condition. Furthermore, as the loops are composed of a single cell, the statistical mismatches between the two loops does not result in non-linearity, but just in gain error in the interpolator, which can be easily taken into account in post-processing.
The conversion time of this approach is high and it is given by:

$$T_{conv} = 2^{N} \cdot \tau_{loop1} \quad (4)$$

where τ_{loop1} is minimum total delay loop and N is the number of bits of the counter.

In this work, to meet the resolution given by the Vernier delay line while keeping a short conversion time, we propose a new operation principle and implement a dual-fine interpolation. Fig. 5 illustrates the signal timing of dual fine interpolation. The first fine interpolation is identical to the classic Vernier delay line: HIT SYNC and CK SYNC pulses propagate in two delay loops and gradually, HIT SYNC reaches CK SYNC with low resolution (LSB_1). The first interpolation finishes when HIT SYNC pulse overtakes CK SYNC pulse and the result is stored in a binary code. Instead, the second interpolation measures the residual time between CK SYNC and HIT SYNC that is left when HIT SYNC overcomes CK SYNC. In this case, CK SYNC pulse gradually reaches HIT SYNC pulse with high resolution (LSB_2). Therefore, the second fine interpolation finishes when CK SYNC overtakes HIT SYNC and the result of interpolation is again stored in a binary code. In this approach, the conversion time is given by:

$$T_{conv} = 2^{N_1} \cdot \tau_{loop1} + 2^{N_2} \cdot \tau_{loop2} \quad (5)$$

where τ_{loop1} and N_1 are respectively the total loop delay and number of bits of the counter in the first fine interpolation, while τ_{loop2} and N_2 are referred to the second fine interpolation. In the proposed architecture, we used a 4-bit counter for first fine interpolation and a 3-bit counter for second fine interpolation. Thus, with $\tau_{loop1} = \tau_{loop2} \approx 3$ ns, the conversion time is equal to 48 ns, while in a classic Vernier delay line, with the same number of bits ($N=N_1+N_2$) and same total loop delay, the conversion time would be significantly higher, about 384 ns. Therefore, the conversion time with dual fine interpolation is strongly reduced, maintaining the same

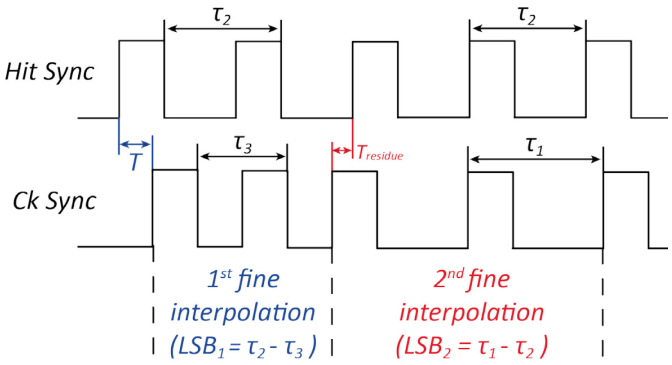

Fig. 5: Signal timing in the dual fine interpolation. In first interpolation, HIT SYNC gradually reaches CK SYNC with low resolution (LSB_1) while the second interpolation resolves the residue of time ($T_{residue}$) with high resolution (LSB_2), reducing conversion time compared to classic Vernier delay line with same resolution.

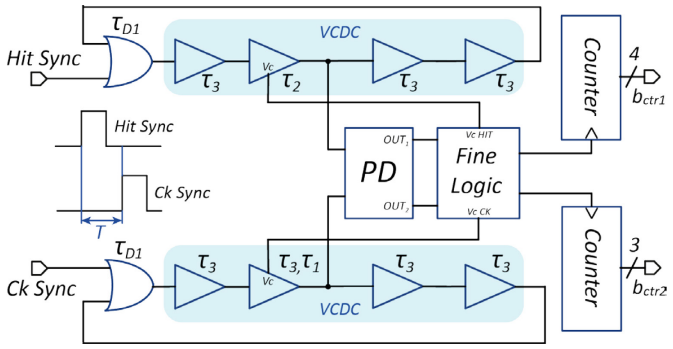

Fig. 6: Block schematic of dual fine interpolator. The phase detector (PD) detects when HIT SYNC overtakes CK SYNC or vice versa. The fine logic block modifies the propagation delay of the cells to change the resolution in first and second fine interpolation.

resolution and number of bits. In Fig. 6 shows the block schematic of the dual fine interpolator. To implement the dual fine interpolation, a phase detector (PD) is needed to verify when HIT SYNC overtakes CK SYNC, and a proper logic is required in order to change the propagation delay of cells (resolution) in the single-stage Vernier delay line.

IV. CONCLUSION

We presented a novel TDC architecture capable of reaching high linearity and high resolution with short conversion time suitable for TCSPC application and TOF measurements. The full-scale range of 80 ns is obtained by using a 4-bit coarse counter in combination with a 200 MHz clock and the resolution of 7 ps is reached thanks to START and STOP interpolators. The new dual fine interpolation allows to achieve 7 ps resolution with shorter conversion compared to classic Vernier delay line architecture. The layout of the overall chip designed in 0.35 µm CMOS technology is shown in Fig. 7.

The chip contains START and STOP interpolators, DLL coarse and additional bias circuitry. The STOP pulse is an external signal provided to the chip, whereas the START pulse

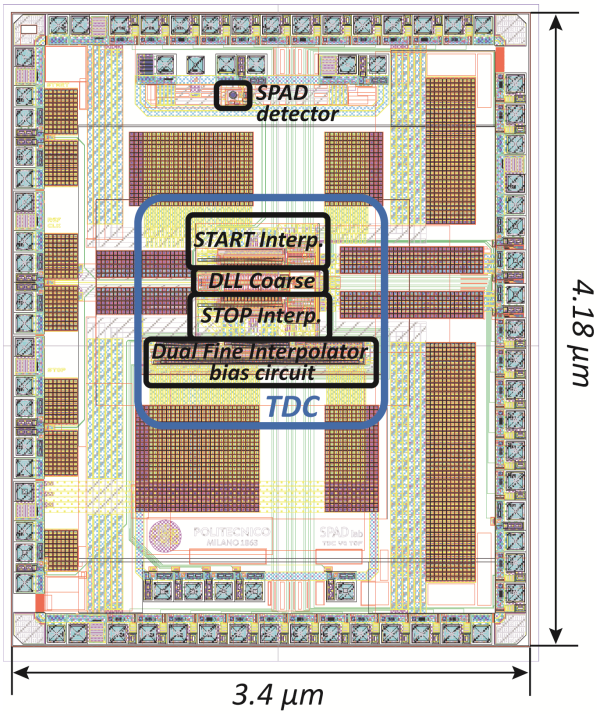

Fig. 7: Prototype chip layout containing TDC with global electronics and a SPAD detector with sensing circuit. The occupation area of single interpolator is about 0.31 mm^2.

can be provided either from outside the chip or from a SPAD photodetector with 30 µm active area diameter integrated with the TDC. The prototype chip is being manufactured and will be characterized implement the converter in a SPAD array.

REFERENCES

[1] M. Perenzoni, D. Perenzoni and D. Stoppa, "A 64 × 64-pixels digital silicon photomultiplier direct tof sensor with 100-MPhotons/s/pixel background rejection and imaging/altimeter mode with 0.14% precision up to 6 km forspacecraft navigation and landing," IEEE J. Solid-State Circuits, vol. 52, no. 1, pp. 151–160, Jan. 2017..

[2] D. Portaluppi, E.Conca, F.Villa, "32 × 32 CMOS SPAD Imager for Gated Imaging, Photon Timing, and Photon Coincidence" IEEE J. Sel. Topics Quantum Electron., vol. 24, no. 2, pp. Mar./Apr. 2018.

[3] Wu Gao, Deyan Gao, Tingcun Wei, Hu-guo, C. YannHu, "A high-resolution multi-channel time-to-digital converter(TDC) for high-energy physics and biomedical imaging applications[C]", IEEE conference on industrial electronics and applications, pp. 1133-1139, 2009.

[4] A. S. Yousif and J.W. Haslett, "A fine resolution TDC architecture for next generation PET imaging," IEEE Trans. Nucl. Sci., vol. 54, no. 5, pp. 1574–1582, Oct. 2007.

[5] V. O'Connor and D. Phillips, "Time-correlated single photon counting," Academic Press, London, 1984.

[6] D. Bronzi, F. Villa, S. Tisa, A. Tosi, F. Zappa, "SPAD Figures of Merit for Photon-Counting, Photon-Timing, and Imaging Applications: A Review," *IEEE Sensors Journal*, vol. 6, no. 1, pp.3-12, Jan. 2016.

[7] B. Markovic, S. Tisa, F. A. Villa, A. Tosi, and F. Zappa, "A high-linearity, 17 ps precision time-to-digital converter based on a single-stage vernier delay loop fine interpolation," IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 60, no. 3, pp. 557–569, Mar. 2013.

[8] F. Villa *et al.*, "CMOS Imager with 1024 SPADs and TDCs for singlephoton timing and 3-D time-of-flight," *IEEE J. Sel. Topics Quantum Electron.*, vol. 20, no. 6, pp. 364–373, Nov. 2014.