

**IN-PIXEL TIME DIGITAL
CONVERTER FOR TIME-OF-FLIGHT
PET IMAGING**

IN-PIXEL TIME DIGITAL CONVERTER FOR TIME-OF-FLIGHT PET IMAGING

By

Ebrahim Nemati Hosseinabadi
Bachelor of Science
University of Tehran
June 2010

A THESIS
SUBMITTED TO THE SCHOOL OF GRADUATE
STUDIES IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

McMaster University
Hamilton, Ontario, Canada

© Copyright by Ebrahim Nemati Hosseinabadi, September 2012

MASTER OF APPLIED SCIENCE (2012)
(Electrical and Computer Engineering)

McMaster University
Hamilton, Ontario

TITLE: In-Pixel Time Digital Converter for Time-of-Flight
PET Imaging

AUTHOR: Ebrahim Nemati Hosseinabadi
B. Sc., (Electrical Engineering)
University of Tehran, Tehran, Iran

SUPERVISOR: Prof. M. Jamal Deen

NUMBER OF PAGES: xiii, 121

*To my dearest in the world,
My parents and my sister*

Abstract

In the past decades, great advances in biomedical imaging towards using less invasive and more sensitive imaging modalities have enabled early detection of diseases through timely diagnosis of patients. Positron emission tomography (PET) imaging, as one of the recent imaging technologies, provides imaging from cellular-level metabolic changes in tissues. This gives PET imaging a substantial lead in detecting disease in their very early stages. PET imaging provides high sensitivity and chemical specificity. However, it suffers from low resolution compared to other imaging methods. Time of Flight (ToF) PET imaging, one of the derivations of the PET, improves the imaging by exactly determining the position of the annihilation event using a time digital converter (TDC). By achieving the timing information of the incident anti-parallel photons coming from an event in ToF PET scanner, the TDC helps to determine the exact location of the event. So, it increases the resolution of the PET scanner.

A ToF PET custom-designed TDC has been proposed in this work. The designed TDC offers relatively high resolution and dynamic range (DR) to satisfy some PET imaging specifications. To increase the sensitivity and reduce the noise and latency, in-pixel design of TDC is desired. Therefore, a time digital converter that is specifically designed for ToF PET should follow a strict set of criteria in its design procedure. A three-staged hierarchical TDC was designed and implemented in $0.13\mu\text{m}$ standard CMOS technology to reduce the total number of delay elements for this area limitation issue. Also, a novel half-CLK period interpolation idea was proposed to reduce the total size of the TDC even more. A counter and half-CLK counter construct the coarse stage of the TDC. A delay locked loop (DLL) works as the first fine interpolator, while, the Vernier delay line (VDL) acts as the second fine interpolation stage.

A high resolution of 39ps was achieved with a relatively high DR of $1.28\mu\text{s}$ and the measured DNL and INL of $0.2T_{\text{LSB}}$ and $0.4T_{\text{LSB}}$. Due to all area reduction techniques used, the final designed TDC measures for 0.11 mm^2 , which is much smaller than other similar TDCs with the same resolution and DR. As the amount of delay in the delay elements in the TDC are susceptible to any change in the environmental changes, a delay locking method was used to compensate for process, voltage and temperature (PVT) variations.

Acknowledgements

This section is dedicated to acknowledge the various help and support that I received during past two years of my academic journey in conducting this research. I continuously received a lot of support and encouragement from numerous people around me during this journey for which I am so grateful.

First and foremost, I would like to thank my supervisor, Prof. M. Jamal Deen, for trusting me to work in this area and for believing in me during my research. I would like to thank him for being permanently involved with my research and for teaching me not only the way to do outstanding graduate research, but also the way to improve my academic skills such as teaching, organization and communication. His devotion, professionalism and exemplary academic powers were always a great push for me towards academic excellence.

Next, I would like to thank Dr. Hao Peng and Dr. Nicola Nicolici for being involved in my research during past year and taking time on commenting and providing feedback on this work and for being in my defense committee. I would also like to thanks Dr. Tapas Mondal and Dr. Ravi Selvaganapathy for always being helpful and supportive in my research work and for teaching me numerous great lessons during past two years.

I offer my gratitude to all the members of Microelectronics group for their kind and warm collaboration and guidance whenever I faced trouble and distress. I would specially like to thank Darek Palubiak, Mohammad Reza Dadkhah, Hossein Kassiri and Sumit Majumder for being such helpful friends. Also, special thanks to Dr. Ognian Marinov for his great experience and comments on my work.

Last but not least, I would like to express my immense gratitude to my family. My father, Professor Abdolali Nemati, and my mother, Shamsieh, for devoting their life to their children and for being such great mentors during my entire life. Also, I would like to thank my lovely sister, Marzieh, for being such a close and kind friend to me. Thanks to you all, for always being there for me.

Table of Contents

Abstract	iv
Acknowledgments	v
Table of Contents	vi
List of Figures	ix
List of Tables	xiii
INTRODUCTION	1
1.1. Biomedical Imaging Systems	1
1.1.1. X-ray Computed Tomography	4
1.1.2. Magnetic Resonance Imaging	4
1.1.3. Ultrasound	5
1.2. Positron Emission Tomography	6
1.2.1. Operation Principle	6
1.2.2. Benefits and Challenges	8
1.2.3. Data Acquisition and Image Reconstruction	10
1.2.4. Applications	11
1.3. Time of Flight PET Imaging	11
1.3.1. Operation Principle	12
1.3.2. Time Correlated Single Photon Counting	13
1.3.3. Improved Image Quality in ToF PET	15
1.4. Contribution and Organization of the Thesis	16
PET DETECTORS' COMPONENTS	18
2.1. System Architecture	18
2.2. Scintillation Crystal	19
2.2.1. Operation of Crystal	19
2.2.2. Important Parameters of Crystal	21
2.2.3. Scintillators for ToF PET Imaging	22
2.3. Photodetector	23
2.3.1. Operation Principle	24
2.3.2. Photodetectors for ToF PET Imaging	27
2.3.2.1. Photomultiplier Tubes	27
2.3.2.2. Avalanche Photodiode	29
2.3.2.3. Single Photon Avalanche Diode	30
2.4. Readout Electronics	34
TIME DIGITAL CONVERTERS	36
3.1. TDC Operation Principle	36

3.2. Important Parameters in TDC Design	38
3.2.1. Resolution	39
3.2.2. Dynamic Range	39
3.2.3. Convergence Speed	40
3.2.4. Accuracy	40
3.2.5. Power and Area Consumptions	43
3.3. Basic Ideas in TDC Design	44
3.3.1. Analog Approach	44
3.3.2. Counter-Based Implementation	45
3.3.3. Clock Cycle Interpolation: Tapped Delay Line	46
3.3.4. Delay Locked Loop	48
3.4. Sub-Gate Delay Resolution TDCs	49
3.4.1. Vernier Delay Line	50
3.4.2. Pulse Shrinking Method	51
3.4.3. DLL Array TDC	52
3.4.4. Local Passive Interpolation	54
3.4.5. Time-Amplification Method	55
3.4.6. Hierarchical TDC	56
PROPOSED TDC PIXEL	58
4.1. Introduction	58
4.2. Proposed TDC Architecture	60
4.3. Digital Counter as the Coarse TDC	64
4.4. DLL as the First Interpolation Stage	67
4.5. Half CLK Period Interpolation Idea	69
4.6. VDL as the Second Interpolation Stage	71
4.7. Realization of the Circuit Blocks	76
4.7.1. Voltage-Controlled Delay Cell	77
4.7.2. Phase Detector and Charge Pump	79
4.7.3. Thermometer-to-Binary Encoder	81
4.7.4. Half-CLK-Period Counter	82
4.7.5. Schmitt Trigger and Output Buffer.....	82
TDC SIMULATION RESULTS	85
5.1. Counter as the Coarse TDC	85
5.2. DLL as the first Interpolator	88
5.3. VDL as the Second Interpolator	96
5.4. Operation of Entire TDC	98
TDC MEASUREMENT RESULTS	103
6.1. Layout of the TDC	103

6.2. Test Setup Board	104
6.3. TDC Characterization	107
6.4. TDC Accuracy	109
6.5. Linearity of the TDC	110
SUMMARY AND FUTURE WORK	114
Appendix 1	117
References	121

List of Figures

Figure 1.1 Main building blocks of medical imaging system	1
Figure 1.2 Non-optical Imaging Systems [1]	3
Figure 1.3 Coincidence events detection using detector ring [1]	7
Figure 1.4 Three types of coincidence events which leads [10]	9
Figure 1.5 Annihilation event localizing in ToF PET [12]	12
Figure 1.6 Position Uncertainty of annihilation inside LOR in PET (left) vs. ToF PET (right)	13
Figure 1.7 Time-correlated single photon counting technique [9]	14
Figure 2.1 PET scanner ring (left) and detector structure (right)	19
Figure 2.2 Scintillator for transforming high energy gamma rays into visible light	20
Figure 2.3 Scintillator output light after interaction with incident gamma ray	21
Figure 2.4 Photomultiplier tube (PMT) structure [9]	28
Figure 2.5 Avalanche photodiode structure [22]	30
Figure 2.6 Structure of SPAD designed in our group [25]	31
Figure 2.7 Typical quenching methods in SPADs: left: Passive [27] & right: Active [28] quenching	32
Figure 2.8 I-V characteristic (left) and output voltage shape (right) of SPAD [25]	33
Figure 2.9 SPAD readout electronics including TDC, memory and circuits to convey data to FPGA	35
Figure 3.1 Input-output characteristic of a 2b ideal TDC [47]	38
Figure 3.2 Single-shot standard deviation according to quantization error	41
Figure 3.3 Analog implementation of TDC: architecture (top), timing diagram (bottom)	45
Figure 3.4 Tapped delay line TDC: architecture (top) and timing diagram (down)	47
Figure 3.5 Delay locked loop (DLL) structure	48
Figure 3.6 Vernier delay line TDC technique: structure (top) timing diagram [47] (down)	50
Figure 3.7 Pulse shrinking TDC structure: using two inverters with different sizing as delay stage	52

Figure 3.8 TDC implementation using the DLL array: architecture (left), timing diagram (right) [43]	53
Figure 3.9 Local passive interpolation technique: interpolated signals (left), realization (right) [41]	54
Figure 3.10 Time amplifier implementation (left) and its output-input characteristics (right) [44]	55
Figure 3.11 Hierarchical TDC architecture using a DLL as the coarse TDC	56
Figure 3.12 Timing diagram for a hierarchical TDC	57
Figure 4.1 Simplified architecture of the TDC	62
Figure 4.2 Timing diagram for the 3-stage interpolation TDC	63
Figure 4.3 The architecture of the designed synchronous counter	64
Figure 4.4 Hit signal synchronization circuit (left) and its operation principle (right)	65
Figure 4.5 Block diagram of dual synchronization method	66
Figure 4.6 Operation of dual synchronization circuit in two critical cases	67
Figure 4.7 DLL structure (top) and tapped delay line (down)	68
Figure 4.8 The new proposed DLL structure based on Half CLK cycle interpolation idea	70
Figure 4.9 New timing diagram of proposed TDC (left) added time due to half-CLK counter outputs (right)	71
Figure 4.10 VDL structure used as the 2 nd interpolation stage of the proposed TDC	72
Figure 4.11 Two flip-flop synchronizer structure (Bottom Left) and its application for generating residue for second fine TDC (Top) and its timing diagram (Bottom Right)	74
Figure 4.12 Final architecture of the proposed TDC after applying stated changes	76
Figure 4.13 Some popular voltage-controlled delay cells: (a) n-voltage controlled (b) p-voltage controlled (c) np voltage controlled delay cells and (d) current starved cascades inverters [58]	77
Figure 4.14 Structure of delay cells of the proposed TDC- current starved structure with additional capacitive load control	78
Figure 4.15 Schematic diagram (top) and operation (bottom) of phase detector used in this work	80

Figure 4.16 Designed charge pump schematic (left) and Operation (right)	80
Figure 4.17 Thermometer-to-binary encoder gate level diagram (right) and MUX to make it compatible with half-CLK cycle interpolation idea (left)	81
Figure 4.18 Schmitt-trigger transfer characteristics (left) and it schematic (right) [62]	83
Figure 4.19 Tapered buffer structure [62]	84
Figure 5.1 The schematic of the designed synchronous counter	86
Figure 5.2 Simulation results for 200MHZ CLK frequency	86
Figure 5.3 Counter 1-cycle error due to lack of synchronization	87
Figure 5.4 Counter accuracy after and before synchronization, CLK freq. = 200MHz, Input time = 33.33ns	88
Figure 5.5 Schematic of the designed DLL	89
Figure 5.6 Schematic of the designed tapped delay line	89
Figure 5.7 Schematic of the designed phase detector	90
Figure 5.8 Schematic of the designed charge pump	90
Figure 5.9 Propagation delay of designed delay cell vs. input bias voltage	91
Figure 5.10 Effect of intermediate cap on capability of locking in DLL in 500MHz CLK frequency	92
Figure 5.11 Schematic and layout of the designed delay cell	93
Figure 5.12 Tapped delay line output-input characteristics	94
Figure 5.13 Simulation results for the operation of the phase detector (left) and charge pump (left)	94
Figure 5.14 Operation of phase detector and charge pump together	95
Figure 5.15 Operation of DLL to lock the two signals at beginning and end of the delay chain	95
Figure 5.16 Locking process of the control voltages to feed the start and stop chains' delay elements	96
Figure 5.17 Output-input characteristic of the VDL	97
Figure 5.18 Test setup for obtaining the accuracy of the TDC	98
Figure 5.19 Accuracy test of TDC: number of counts per time slot for input time of 16.67ns ...	99
Figure 5.20 Counts per time bin histogram for DLL and VDL	101

Figure 5.21 DNL and INL report for the TDC found through simulations	102
Figure 6.1 Layout and Photomicrograph of the TDC prototype chip	104
Figure 6.2 Test setup for measuring the performance of TDC	105
Figure 6.3 PCB designed for testing the prototype TDC	107
Figure 6.4 Optronic-TRRC1 as the delay generator for generating stop signals	108
Figure 6.5 TDC output-input characteristics performed in sweep measurement setup	108
Figure 6.6 TDC accuracy test achieved from 1000 measurements with a fixed time input	109
Figure 6.7 Nonlinearity measurements of the tapped delay line	110
Figure 6.8 Nonlinearity measurements for delay locked loop	111
Figure 6.9 Counts distribution for 64 bin of the TDC to measure nonlinearity for 5000 measurements	112
Figure 6.10 Dynamic nonlinearity of the designed TDC for 5000 measurements	112
Figure 6.11 Integral nonlinearity of the designed TDC for 5000 measurements	113

List of Tables

Table 2.1: Common scintillators used in ToF PET scanners	23
Table 7.1: Summary of the measured performance of the prototype TDC	115
Table 7.2: Comparison table for characteristics of the prototype TDC	115

Chapter 1

INTRODUCTION

1.1. Biomedical Imaging Systems

As a general term, imaging means to generate pictorial representation of an object or process [1, 2]. Biomedical imaging is referred to creation of vivid images from the biological body. However generating images from organs or tissues removed from the live organism could be useful; it is not recognized as biomedical imaging. The main purpose of medical imaging is to reveal subtle changes of tissues to detect disease such as cancer [1, 2, 3]. Another application is clinical trial to continuously study the safety and effectiveness of a therapeutic agent. There are several different techniques to create images of organisms from both interior and exterior tissues. Medical imaging systems in spite of their imaging techniques and application, follow the same basic functions to realize imaging: 1) collecting of the incoming photons from the tissue, 2) filtering for unwanted wavelengths, 3) performing photon-to-electron conversion using detectors, 4) readout mechanism for extracting data from detectors, 5) signal processing to obtain image from the raw data coming from detector [1, 3, 4]. Fig. 1.1 shows a general sketch of a medical imaging system.

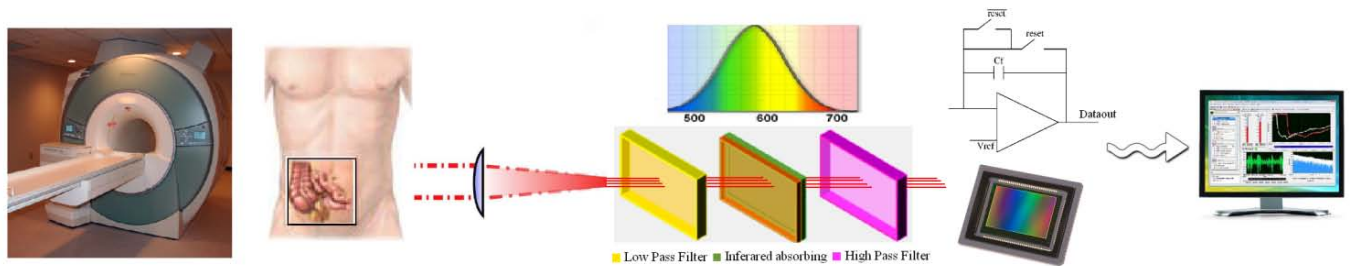


Fig. 1.1: Main building blocks of medical imaging system

To provide appropriate images from the organ, an imaging system must satisfy excellent performance in each of its building blocks. Deviation from a perfect medical imaging system however is inevitable due to the following imperfections [1]:

- Inaccurate depiction due to spatial deformation, nonlinearity, blurring and artifacts.
- Incomplete representation due to mapping an inadequate number of object properties.
- Irreproducibility due to noise and random fluctuations accompanied with the signal.

Imaging techniques can basically be divided into two categories: Optical and non-optical imaging techniques. Optical imaging is based on absorption and scattering of the light. The light that has been emitted from deep inside of tissue reaches the surface after several times of scattering events. Based on the intensity and spatial and spectral resolution of this light, some positional information of the tissue can be obtained [5, 6]. Optical imaging techniques include endoscopy, fluorescence imaging, optical coherence tomography and transillumination imaging. Accessibility, low infrastructure cost as well as superior sensitivity are the main reason for popularity and growth in use of optical imaging techniques. Another important advantage of optical imaging is that the radiation is non-ionizing which makes it possible to use with reasonable dose repeatedly without harm to the patient. However, the major drawbacks of optical imaging have been its lower spatial resolution and limited region of coverage [3, 5, 6].

Non-optical imaging, unlike optical imaging, uses high energy rays instead of visible light which penetrates to the body with relative ease. These high energy photons provide imaging from deep inside of the tissues. In addition, invisible properties of human body such as local values of blood volume, tissue metabolism, perfusion and oxygen utilization can be achieved using non-optical imaging systems [1, 7]. Imaging using invisible radiation initiated with discovery of X-ray by Rontgen in 1895 where visible image of an invisible object (bones of his wife's hand) was made for the first time. X-ray computed tomography (X-ray CT), ultrasound, magnetic resonance imaging (MRI) and its derivatives; single photon emission computed tomography (SPECT) and positron emission tomography (PET) constitute different types of non-optical imaging methods [1, 3, 5]. Each of these imaging modalities employs a specific wavelength out of visible wavelength range. Fig. 1.2 represents distribution in wavelength

domain. In most of these non-optical imaging techniques, the concept is to detect radiation from the target after administrating radioactive materials. The concentration of this work is on this type of imaging which is called nuclear imaging that includes X-ray CT, PET and SPECT.

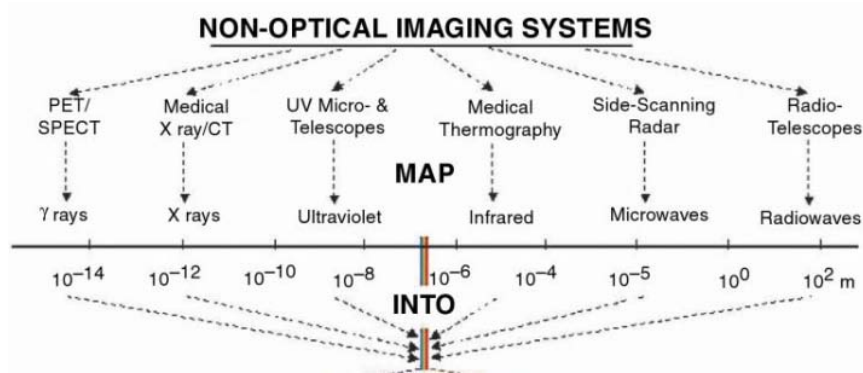


Fig. 1.2: Non-optical Imaging Systems [1]

The main advantage of nuclear imaging is their high sensitivity and spatial resolution, but more importantly capability of measuring radioactivity from deep inside of the tissue. This reduces the dependence of the detected signal on the depth at which the signal is emitted. On the other hand, one of the main disadvantages of nuclear imaging is the involvement of ionizing radiation which makes frequent clinical imaging trials impossible. Also it should be noted that radioactive decay cannot be controlled, so the presence of a non-specific background signal in the image is inevitable [2, 5].

Nuclear imaging techniques can be again divided into ex-vivo and in-vivo categories. The ex-vivo auto-radiographic method employs phosphorous storage plates or real time auto-radiographic systems while in-vivo techniques which is also called “emission tomography (ET)” techniques try to inject radionuclide materials in human body that produce high energy photons such as X-ray and gamma ray. Activity of the point of interest in target can be achieved by measuring the decay rate of radiation from that point [1, 5, 8]. This can be accomplished using different techniques which will be explained later. PET and SPECT are two ET techniques which utilize radionuclide materials to generate gamma rays inside the tissue. Emission tomography is capable of providing information on distribution of glucose metabolism, blood flow and receptor concentration which can be used to locate tumors.

As one type of ET, a complete review on PET imaging will be given as it is the main concentration of this work. But before that, a brief introduction on some of the main non-optical imaging systems will be presented.

1.1.1. X-ray Computed Tomography

X-ray CT, also known as CT scan, is a biomedical imaging technique which employs computer-processed x-rays to provide topographic images by reconstructing image from a large number of x-ray transmission measurements (more than 500,000) [8]. Since its revolutionary invention in early 1970, CT became a standard imaging procedure in a huge number of medical facilities around the world. X-ray CT is capable of differentiating between tissues with 1% of physical density difference due to its inherent high contrast resolution. The high resolution of CT which is about tens of microns makes it possible to sensitively detect very fine changes in bone structure and remodeling in bone density [5, 8]. The main challenge in X-ray CT is the radiation dose injected to the subject that causes biological effects depending on dose and frequency of injection. High degree of ionization which happens due to X-rays can damage directly or indirectly DNA structure which may lead to cancer. Nowadays, the most important subject to investigate in CT is how to obtain high-quality images with minimum radiation dose. Another challenge is to reduce the scattered x-ray flux as CTs are prone to detection of scattered radiation.

1.1.2. Magnetic Resonance Imaging

Magnetic resonance imaging or MRI is a medical imaging technique which utilizes the nuclear magnetic resonance of the atoms inside the body by measuring their interaction with both a large external magnetic field and RF wave to produce highly detailed image of internal structure of the body. The operation principle of the MRI is as follows: the body is first targeted to a high magnetic field. Protons of the target are aligned either in parallel or anti-parallel of the direction of magnetic field. The varying molecular structure and the amount of hydrogen in tissues affect how the proton behaves in the magnetic field. The more a structure has hydrogen

extent, the more it is magnetized in external field. After being targeted to magnetic field, an RF signal is applied because of which some of the spins absorb energy and change their energy state going to higher energy levels. When the RF signal is disabled, some of these spins that had moved to higher energy state give off their energy to their lattice causing the magnetization to regrow along the external magnetic field again. This regrowth takes place with a rate which is given by the tissue relaxation parameters i.e. each pixel of target will have its specific magnetic regrow rate based on its material. So a pictorial description from inside of the body can be achieved by recruiting this regrow information distribution.

One of the key advantages of MRI is its high soft-tissue contrast which makes it possible to detect disease in soft tissues as well as hard tissues such as bones. Another advantage is its safety compared to X-ray CT for patient as it uses magnetic field and RF signal which are non-ionizing. Higher magnetic field leads to higher imaging resolution which is not always possible. Another way to increase the resolution is to increase the acquisition time which implies patient to stay without any motion for a long time for achieving a high quality image (this can be counted as one limitation of MRI). In-vivo images with spatial resolution of about hundreds of micron can be obtained using MRI in several minutes [5, 7]. Low sensitivity of the MRI systems compared to other imaging techniques such as optical and nuclear imaging is another disadvantage of this technique. Also, it should be noted that MRI is not capable of detecting some important element such as oxygen [5, 9]. Generating high magnetic field and RF signal and designing high-sensitivity detectors are other important challenges in MRI system design which makes MRI an expensive medical imaging technique.

1.1.3. Ultrasound

Ultrasound imaging is another harmless imaging technique used for visualizing body structures such as bones, muscles, joints etc. Ultrasound refers to a sound wave with frequency above human hearing range (2-18 MHz is usually used in medical imaging). In ultrasound imaging first, ultrasound wave is generated and introduced to the target. Part of this sound is reflected wherever there is density change in the tissue and these reflected echoes are then gathered using a detector and turned into an image using reconstruction methods. In addition to its

safety for the patient (non-ionizing) this method provides low-cost high-resolution real-time imaging. One of the key disadvantages of this system, however, is its small depth of penetration. Actually there is a trade-off between spatial resolution and depth of penetration. This means, in order to use this method for higher penetration depth, lower frequencies should be used which leads to smaller spatial resolution [5, 9]. Last but not least medical imaging system that will be examined in this work is PET imaging which will be described in the next section.

1.2. Positron Emission Tomography

Positron emission tomography (PET) is another medical imaging technique which was proposed back in 1950s for the first time when it was realized that 3D imaging is possible based on the high energy photons emitted due to annihilation of positrons in a particular radioactive material. In early 1980s, PET imaging became a certain diagnostic tool in medical imaging. During past 30 years considerable advancements in PET imaging happened in instrumentation and technology both in detectors and signal processing aspects. This helped PET imaging become one of the most popular in-vivo molecular imaging modalities nowadays. In this section, we describe PET imaging operation principle, benefits and challenges as well as its applications.

1.2.1. Operation Principle

PET imaging is a tracer approach in which a labeled compound needs to be introduced to the target first. A tracer in PET is a biological molecule that carries a positron emitting isotope with it. This tracer should have the capability to be distributed in the points of interest which has to be distinguished from other points. Many isotopes can be employed as PET agents such as ^{11}C , ^{13}N , ^{15}O and ^{18}F . These are short-live positron emitters which are prepared using different chemical methods. For instance, ^{11}C , cyanide, is formed by first bombarding $\text{N}_2 + 5\% \text{H}_2$ gas

target with 20-MeV protons in the cyclotron target to produce carbon-labeled methane, $^{11}\text{CH}_4$. This product after being combined with ammonia and passed over a platinum wool catalyst at 1000°C generates $^{11}\text{CN}^-$, which is subsequently trapped in NaOH [1, 8]. Generated ^{11}C can be used to label glucose for example. Since glucose is the primary source of energy and tumor tissue consumes more energy than normal tissue, labeled glucose tends to accumulate where tumor stays. This takes couple of minutes. After that, the tracer nuclei decay by emitting positrons which is a positive electron. Positrons travel through the tissue and quickly give up their energy and combine with electrons forming hydrogen-like orbiting pairs called positronium. Positronium is unstable and decays via annihilation and emits two anti-parallel 511-KeV gamma rays with the speed of light. These two gamma rays will be detected then with a detector ring that surrounds the patient. Fig. 1.3 shows this detector ring with a diagram that shows how a coincidence is detected after annihilation takes place.

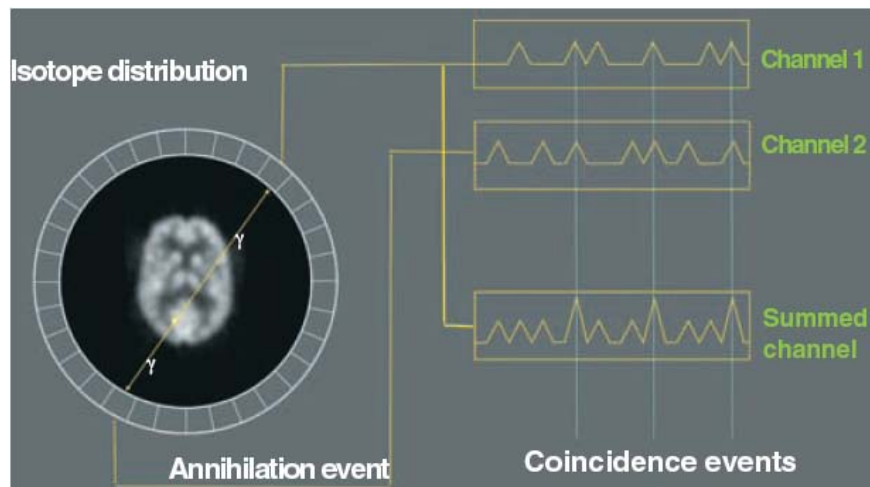


Fig. 1.3: Coincidence events detection using detector ring [1]

As can be seen from Fig. 1.3, when a coincidence event is sensed in a pair of detectors, “Summed Channel” signal shows a rise in its amplitude almost twice the amplitude of each pulse. Simply a comparator can be used to sense when the signal passes a particular threshold to detect coincidences. Whenever an annihilation has been detected by a pair of detectors, this means the annihilation must have happened somewhere along the line connecting these two detectors which is called line of response (LOR). Indeed, exact point of annihilation on the LOR cannot be measured. To form the 3D image of the target a huge number of measurements

should take place from different angles. This is called tomographic acquisition. Acquisition from detector pairs at various angular views followed by appropriate reconstruction algorithms leads to estimation of the tracer distribution within the object with a finite spatial resolution. These reconstruction algorithms are more complex than CT and SPECT as the data set collected using PET is much poorer due to PET's limited spatial resolution.

1.2.2 Benefits and Challenges

A very important advantage of PET scanning over other imaging techniques such as CT and MRI is that PET can provide cellular-level metabolic changes in tissues. This gives PET a significant lead in determining diseases in their early stages as diseases process usually starts with functional changes in the cellular level. Unlike PET, CT and MRI imaging systems have to wait for the disease to change the structure of the tissue or organ to be detected. This feature of PET also can be used to determine if the tumor is benign or malignant before attempting surgery. In addition, neurological illnesses such as epilepsy, Alzheimer's disease, and other dementias can be detected in their early stages using PET.

PET is a non-invasive imaging system. It does involve exposure and ionizing but compared to CT the radiation dose is much less. So this method is much safer than CT. However, usually a combination of CT and PET is employed to give full 3D view of organs. Other great benefits of PET over other imaging systems are its high sensitivity and chemical specificity. One major challenge in PET system is its resolution. There are many factors which degrade the resolution of PET. As PET is sensitive to just coincidence in a pair of detectors, any deviation of the gamma rays more than a certain value leads to error. To illustrate this problem, Fig. 1.4 depicts three conditions of coincidence detection. A true coincidence happens from a single annihilation which is detected by a pair of detectors. This means no scattering has happened before being detected by detectors. A scattered coincidence, however, happens when one or both of the gamma rays generated from a single annihilation are scattered before detection with the detectors. Finally random coincidence happens just due to coincidence of two annihilations at the same time which may lead to detection of two gamma rays by two

detectors as can be seen in Fig. 1.4, Scattered and random coincidences lead to error and thus spatial resolution limitation.

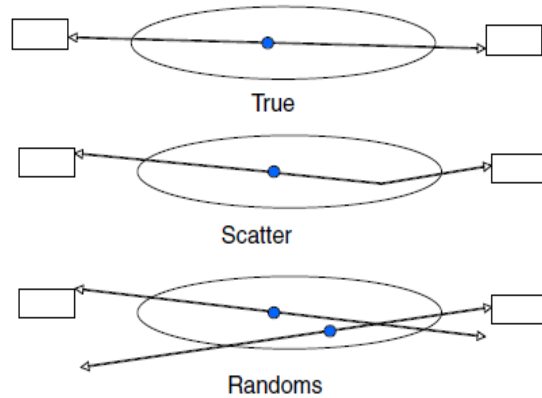


Fig. 1.4: Three types of coincidence events which leads [10]

Another factor which limits the resolution of PET is called parallax error which is a result of different depths of penetration of gamma rays in the crystal before complete absorption [1]. So, one of vital challenges in PET imaging is to determine depth of penetration of gamma rays in the crystals in order to prevent wrong LOR assignment. Also, another limitation on resolution comes from the fact that the reconstructed image in PET is the image of annihilation locations; however, annihilation location is different from positron emission location. This is due to range of positron before annihilation which usually adds uncertainty of less than 1mm [7]. As stated by [8], “Combining values for these factors for the PET-600 tomograph, we can estimate a detector-pair spatial resolution of 2.0 mm and a reconstructed image resolution of 2.6 mm.”

To solve the problem of spatial resolution limitation, the number of coincidences should be increased to reduce the noise in the image. Methods to increase the spatial resolution have been investigated in [1]. Increasing the radiation dose, using efficient scintillators and detectors and more of the energy spectrum and solid angle are some approaches to increase the coincidence event and reduce the noise. Besides, the spatial resolution advancement challenge, another challenge in designing PET systems is to design fast scintillators and detectors (as main components of PET system) especially in cases where high doses of radiation is used. Deadtime of scintillators and detector determines the maximum event rate achievable in PET system.

New scintillators and detectors with fast response time have been proposed during the past few years which pushed PET imaging toward being used for higher event rates.

Another important limitation of PET imaging system is its high cost which keeps this technology from being spread out as fast as it was expected to. The high cost of PET system is mostly due to the high cost of cyclotrons needed to produce short-lived radionuclides. In addition, producing the radiopharmaceuticals after radioisotope preparation should take place on-site. Employing a third-party supplier for radionuclides is an attempt to resolve this limitation; however, this restricts PET system to tracers with high half-life such as fluorine-18.

1.2.3 Data Acquisition and Image Reconstruction

To be detected, high energy gamma rays first should be transformed to low energy visible light by a crystal. The crystal generates a burst of low-energy photons after receiving a gamma ray. These photons are achieved using a photodetector which could be a photomultiplier (PMT) or a photodiode. This photodetector transforms the incoming visible light to appropriate output electrical signal which then can be processed and stored by readout electronics. We will explain the operation of crystal, photodetector and readout electronics as the main components of PET detectors completely in the next chapter.

The next step after data acquisition is to reconstruct the image based on the collected raw data. This raw data which basically is the collection of incident gamma rays detected by each pair of detectors is used to simultaneously solve a set of equations for the activation of each parcel of tissue along each of the LOR. But, before that, usually pre-processing of the data should be done. This includes estimation and subtraction of random and scattered coincidence event and detector dead-time correction. Usually, the PET system employs 15 to 47 transaxial layers for this purpose. The lead shield prevents activities from patient that causes false counts while tungsten septa is used as another filtering layer to reject coincidence occurred due to scattering [8, 10]. After data pre-processing, they will be grouped into projection images called sinograms which then are sorted by the angle of each view. Then, an image reconstruction technique similar to the one used in CT is employed to reconstruct the 3D image from sinograms. Filtered back projection (FBP) and expectation-maximization are two common

techniques that are used for image reconstruction in PET systems in which statistical estimation methods such as likelihood of annihilation event estimation are used [2, 8]. Describing in detail of these techniques is out of the frame of this work.

1.2.4 Applications

As we mentioned before, PET imaging, unlike CT and MRI (which are capable of detecting anatomic structure and changes of tissues), is capable of detecting molecular biology of tissues and organs even before anatomic changes. This is because specific radio-chemical can be employed to target a particular function of a tissue or an organ of body. This makes PET a great imaging technique to detect disease and disorders in their very initial steps. Oncology and neuro-imaging are considered as two main applications of PET imaging. In oncology, a tracer such as F-18 is used which taken up by glucose-using cells. This results in radiolabeling of tissues with high glucose consumption such as most of the cancers. Neuroimaging, as another popular application of PET, employs specific radiotracers to image the activity of brain.

1.3. Time of Flight PET Imaging

The distribution of annihilation events in PET imaging systems is found by reconstruction methods using computed tomography, as mentioned in the previous section. However, it was known, even before employing these methods for PET (in early 1980's), that the exact location of annihilation events can be found accurately by direct measuring the time difference of annihilation photons called Time of Flight (ToF) PET [10, 11]. But, the limitation on the response time of the detectors and scintillators and lack of capability for measurement of time intervals less than nanosecond kept ToF from becoming the dominant method for PET image acquisition. During past decade, however, interest in development of novel ToF PET imaging systems has increased due to vast progress in proposing new fast scintillation materials and photodetectors as well as advancement in resolution of the time interval measurement [11, 12, 13].

1.3.1. Operation principle

In PET, detection of two anti-parallel gamma rays indicates the LOR in which annihilation happened with no more information on its exact location on LOR. As a result, image formation takes place using different measurements from different angles; tomographic images are generated through some reconstruction methods such as filter back projection (FBP) in which counts are distributed uniformly along the LOR. In ToF PET imagers however, the times of arrival of photons are also measured. As can be seen from Fig. 1.5, the time difference between the arrivals of two photons simply can be correlated to the position of the annihilation event with respect to the center of the field of view (Δx):

$$\Delta x = \frac{c}{2} \Delta t \quad c \text{ is the speed of light}$$

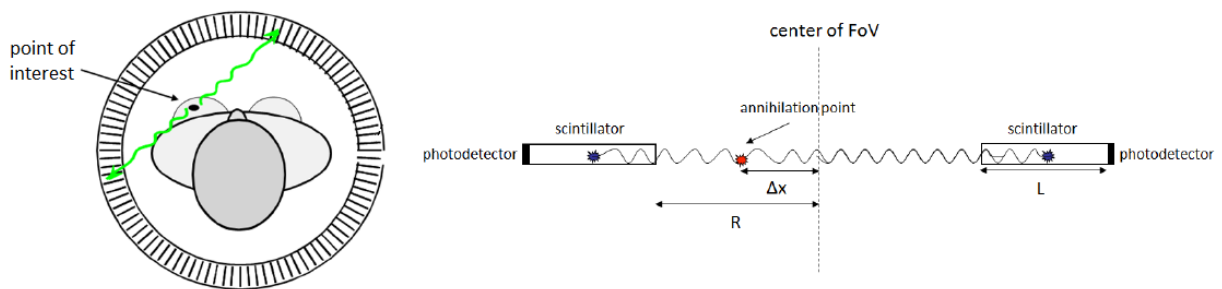


Fig. 1.5: Annihilation event localizing in ToF PET [12]

This correlates the spatial resolution to time resolution of time interval measurement. With perfect ToF measurement (infinite time resolution), the exact point of annihilation for each event can be detected and no reconstruction is needed, however due to imperfections in the time-of-flight measurement, reconstruction algorithms should be employed. But, in the algorithm, instead of distributing the counts uniformly all over the LOR, they are localized within a probability distribution with a full-width half-maximum (FWHM) around the annihilation point (Fig. 1.6). This helps to reduce the statistical noise and increase the quality of the reconstructed image [9, 11, 14].

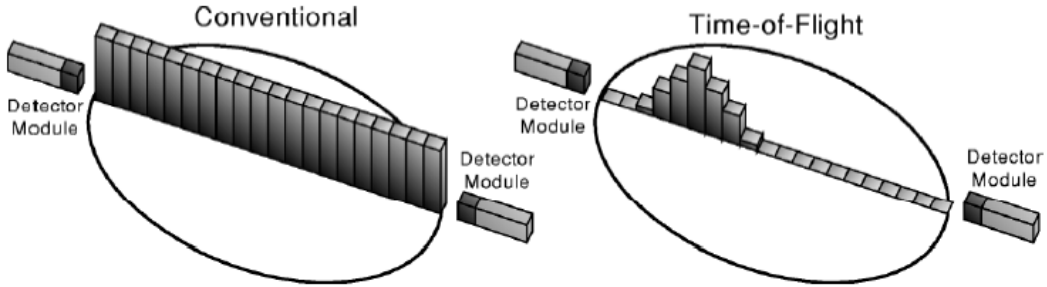


Fig. 1.6: Position uncertainty of annihilation inside LOR in PET (left) vs. ToF PET (right)

As can be seen in Fig. 1.6, due to the limited time resolution, the exact point of the annihilation event cannot be determined and Δt is blurred by a variance $\sigma_{\Delta x}^2$ which leads to blurring in event's position. As a result, timing measurement with high accuracy plays a leading role to provide high-quality image in ToF PET imaging systems.

1.3.2. Time-Correlated Single Photon Counting

The detector's output signal shows the incoming photons where each photon is represented with a pulse randomly distributed in a pulse train. These pulses should be counted and their timing information should be collected to realize ToF PET imaging. Timing information of incoming pulses can be achieved using two techniques: The multichannel-scalar and time correlated single photon counting (TCSPC) systems [6, 9]. In both systems, a discriminator such as a comparator is needed to pass the pulses with amplitude more than a threshold value and remove the unwanted pulses. Then a counter measures the number of pulses. In multichannel-scalar technique, a high speed memory is employed and by fast switching to different memory locations, the counted number of photons coming from a large number of consecutive channels is stored. This technique is the photon-counting equivalent of digital oscilloscope [9]. The counted number of photons in this method can be achieved using gated photon technique and then output of counter is inserted into the fast memory. In gated photon counting, the number of pulses is counted only during a short time window which is defined by the gate generator. The advantage of the multichannel-scalar technique is that it directly delivers the signal by detecting a large number of photons in a sweep. However, it has poor time resolution and count rate.

In very low-level light and high repetition rates, which is the case for many applications in biomedical imaging, the probability of detecting one photon in one signal period is far less than one. So it is not necessary to design for cases with possibility of detection more than one photon in one period. Instead, the decay rate of incoming optical waveform can be estimated by measuring the time of arrival of the photons in the signal period (if there was any) and build up a histogram of the photon arrival times (Fig. 1.7). This is called time-correlated single photon counting (TCSPC) [6, 9, 15].

In ToF PET imaging, the time difference of two incoming pulses from a single annihilation event determines the exact location of the event in LOR with a certain spatial accuracy. Let's call this a pixel. In order to build the histogram of optical waveform for each pixel in the PET chamber, timing information of each photon reaching to each of the detector should be recorded. Then, based on that, any detected coincidence will be assigned to a pixel in the chamber by determining the exact time difference of arrival times of two gamma rays. Then histogram of the photon hits for each pixel is formed to determine the decay constant and thus activity of each pixel inside the target. More information on decay rate detection is related to signal processing part of PET system which is out of the frame of this work.

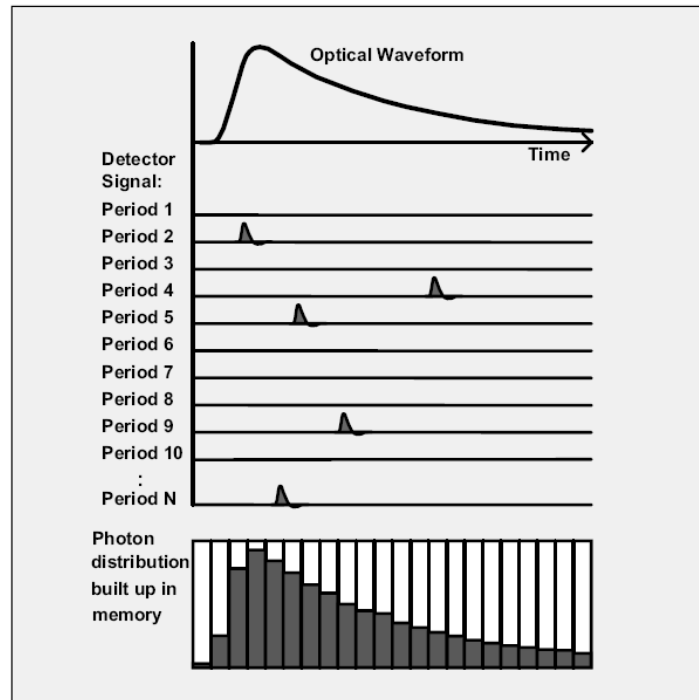


Fig. 1.7: Time-correlated single photon counting technique [9]

As can be seen from Fig. 1.7, using TCSPC, arrival times of the photons in each period have been recorded to form the distribution of the photons along the period. As can be seen, photons tend to arrive at the first half of the period which leads to exponential shape of the decay for the photon's arrival histogram.

1.3.3. Improved Image Quality in ToF PET

In non-ToF PET imaging systems, the FBP algorithm reconstructs the image by incrementing each pixel that lies on the LOR to the number of measurements corresponding to that LOR. So, basically, the coincidence event contributes to all the pixels inside the LOR rather than the pixel from which it was originated. This adds uncertainty or variance to the measured position of the pixel. This variance is much less in ToF PET imaging as incrementing is applied just to some of the pixels close to the annihilation pixel along the LOR. This variance reduction not only is fulfilled for true events but also for scattered and random events. This reduces the noise due to unwanted coincidence as the timing resolution improves [11, 13].

Using ToF PET can significantly reduce the axial blurring and statistical noise (which is one of the main limitations of PET systems) due to the reasons mentioned above. Therefore, it improves the signal-to-noise ratio of the measurements and thus the quality of the achieved image [10, 11, 13]. [16] did a complete comparison between the ToF and non-ToF PET imaging system performances through simulations and measurements. It is indicated in that paper, that ToF PET achieves similar or better image quality compared to non-ToF with only 1/6 the total counts. This means that the sensitivity and image quality gain in ToF PET is 6 times of its equivalent non-ToF PET system. It was also mentioned in that paper, that ToF PET provides faster convergence of reconstruction algorithm with better timing resolution. These reasons clearly justify why ToF PET imaging has received a lot more attention compared to conventional PET during past decade.

The main difference in realizing ToF PET compared to non-ToF is a building block to record the timing information of the incoming photons. This block is called Time digital converter (TDC). The concentration of this work is to design a TDC that satisfies the specifications needed for ToF PET imaging system. Next Chapters of this thesis are going to

explain time digital converter design issues for ToF PET imaging systems and the proposed TDC and its performance.

1.4. Contribution and Organization of the Thesis

This thesis will address a suitable in-pixel timing measurement circuit with a high resolution and wide dynamic range by proposing a novel time digital converter (TDC) for recording the times of arrival of the incident photons in the time-of-flight PET imaging system. The main contribution was to reduce the total size of the TDC and keep the resolution and dynamic range of TDC high enough for ToF PET application. A three-stage TDC was designed to reduce the number of delay elements needed. Besides, a novel half-CLK interpolation method was proposed to reduce the total size of the interpolation stages by half. In addition, a new delay element was designed with a wide delay range and fairly small area which made it possible to operate the TDC under frequency range between 100 to 500MHz. Moreover, PVT variation cancellation was provided using delay locked loop for both delay lines in the TDC to keep it working properly during process and environmental changes.

Chapter 2 will discuss the main components of a ToF PET imaging system, including different scintillation crystals suitable for this application, a review on different photodetectors. Finally, the readout circuit that has been usually employed to collect the incident photons' information will be introduced. A brief introduction on how timing information of the photons is collected by TDC will be provided.

Chapter 3 delivers a complete review on different time digital converter designs. But first, the operation principle of the TDC and important parameters in TDC such as resolution, dynamic range, conversion speed, accuracy and power and area consumption will be discussed. Then, different TDC design techniques including both categories of basic poor-resolution TDCs and sub-gate delay resolution TDCs will be presented.

In Chapter 4, the specifications of TDC needed for ToF PET imaging system will be examined first. Then, the proposed TDC architecture will be introduced. A detailed explanation

on all of the building blocks of the TDC will be provided next. After that, the improvements on the TDC prototype based on the novel ideas such as the new delay element and the half-CLK interpolation will be presented and final TDC architecture will be offered. Finally, the way of realization of some main circuit components of the design will be addressed.

Chapter 5 is devoted to the simulation results for the designed PCB. Results for different stages of the TDC will be presented and a complete simulation result set on the operation of the whole TDC including resolution test, accuracy test and nonlinearity test will be provided.

In chapter 6, the results from the measurement of the designed and implemented TDC will be offered. The measurement setup includes capability to measure the operation of the TDC, its accuracy and nonlinearity measurements and its performance during different bias conditions. Finally a summery on the designed TDC together with the opportunities for future work will be delivered.

Chapter 2

PET DETECTORS' COMPONENTS

2.1. System Architecture

This chapter describes different components of detectors in ToF PET scanner system. After preparation of the target by administrating the desired radioactive material which accumulates in the areas of interest in body, positron annihilation starts and acquisition is performed by PET detectors. PET acquisition is done using a ring of detectors surrounding the target. This ring consists of a huge number of detectors to detect the incident gamma rays in all directions with high spatial resolution. Fig. 2.1 shows PET ring together with the structure of each detector. As can be seen, each detector consists of a scintillation crystal, a photodetector and readout electronics. The scintillation crystal absorbs high-energy gamma rays and generates a burst of visible low-energy photons. These low energy photons then can be detected by photodetectors which usually perform in the range of visible light wavelengths. Detection of each photon in a photodetector is followed by a pulse generated in the output of detector which should be sensed and processed using the readout electronics. This includes circuits blocks such as signal shapers, counters, memories and in case of time-correlated measurement (ToF PET imaging) a Time Digital Converter (TDC). The TDC records the arrival times of incident photons to perform time-of-flight estimation, as explained in Chapter 1.

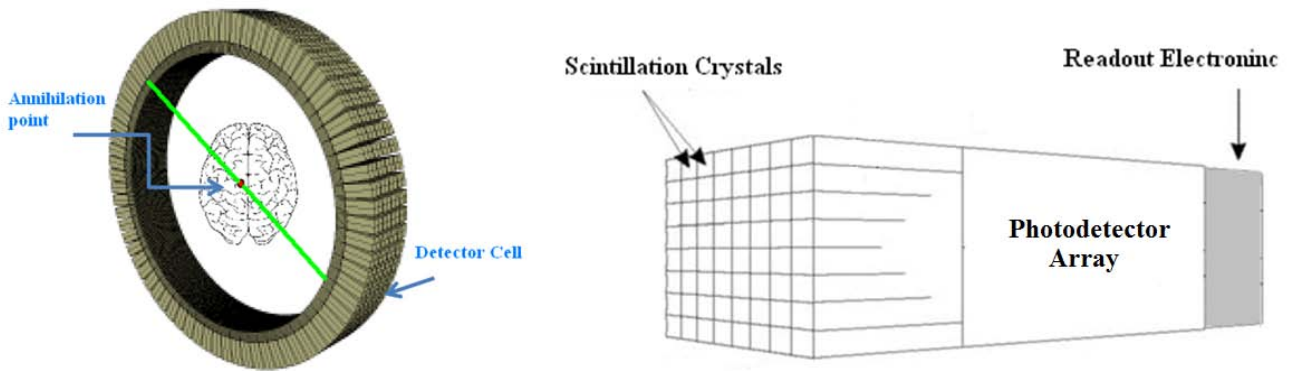


Fig. 2.1: PET scanner ring (left) and detector structure (right)

This chapter consists of description of each of these parts and briefly reviews their recent advancement in the literature. A quick overview on different types of these components and a brief comparison between them will be presented. This includes different types of scintillators and photodetectors and common method for reading out the data from the scanner.

2.2. Scintillation Crystal

As we mentioned in the previous chapter, the idea of detecting an annihilation point based on the timing information of a photon pair has been proposed from the very first moments of introducing PET, but, a key limitation to realizing this idea was the restricted timing response of scintillators. As stated in [13]: “The key advance that enabled modern TOF PET was the development of new scintillator materials which had faster decay time (reducing dead time) and considerably higher light output”. Scintillators of various materials have been used in a very broad range of applications since 1950s. Since then, one of the largest consumers of scintillators has been the medical imaging market. In this section, we will briefly describe scintillator operation and its usage in PET imaging systems.

2.2.1. Operation of Crystal

Photodetectors should transform incoming photons into electrical signal in order to be processed and stored by readout electronics. But usually, photodetectors work in an energy range much lower than the energy of gamma rays emitted from annihilation events. As the energy of a gamma ray (few thousand to million electron-volts) is much greater than the average binding energy of electrons in photodetectors [1], an intermediate substance is used before photodetector in order to convert high energy photons to low energy visible light. This material is called scintillation crystal or scintillator (Fig. 2.2). As can be seen in the figure, gamma ray has been absorbed and low-energy photon has been generated instead by crystal. This visible light is then transferred to photodetector to be detected.

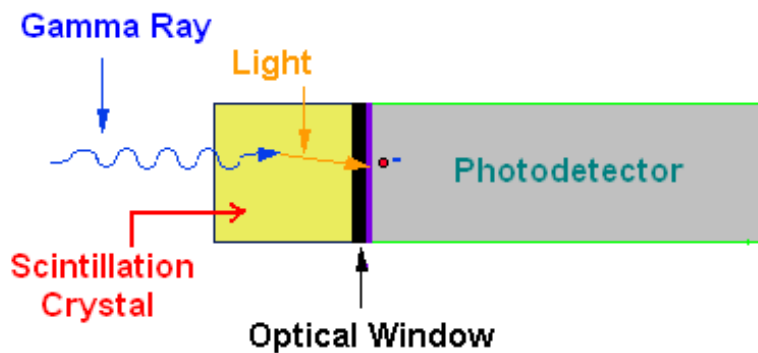


Fig. 2.2: Scintillator for transforming high energy gamma rays into visible light

The energy of gamma rays is transferred to one of the bound electrons of scintillators which results for this electron to be ejected. This will remain a hole in the matter which makes the crystal excited. After a very short period an electron from higher energy level fills this hole to bring the atom to its lowest possible energy level. This releases an X-ray photon which again is absorbed by another photon in crystal before leaving the crystal. This again leaves a hole which is going to be filled by another outer level electron. This procedure continues until the whole energy of incident gamma ray is transformed into visible low-energy photons which can now leave the crystal. This produces a burst of visible light during scintillation process with a typical shape shown in Fig. 2.3.

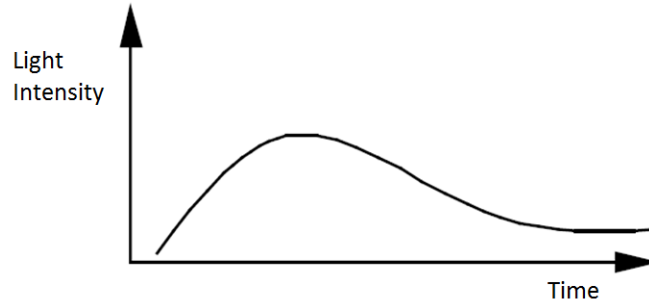


Fig. 2.3: Scintillator output light after interaction with incident gamma ray

This visible light after being inserted to the photodetector, is transformed into electrical signal which is then capable of being collected, processed and stored (details on operation of photodetector and readout circuit will be given in next sections).

2.2.2. Important Parameters of Crystal

There is huge range of scintillators of different types that have been proposed with different appearances and different characteristics. The ideal scintillator, though, is the one with high light output, high energy resolution, fast timing response, high stopping power, good linearity and high density with large photoelectric cross-section [1, 12, 13]. Also it must be easy to grow in large crystals.

The most important factor in the scintillation crystal is its output light which is related to both energy resolution and efficiency of the detector, i.e. it would be easier to detect the incident gamma ray by the detector. The detector efficiency is the ratio of detected particle to the total number of incident photons. A high density in the crystal helps reduce the parallax error as the variability of depth of interaction (DOI) of incoming gamma rays will be reduced. Also it increases the detector sensitivity to the coincident events [13]. In addition, high density reduces the scintillator size and allows for a compact detector [17]. The stopping power is the average energy loss of the particle per unit path length. A fast time response or short decay is another important factor of a good scintillator. Response time is the time needed for the whole output light burst to be generated from the incident ray. This factor plays a significant role, especially in time interval measurement. Other factors in scintillators are the low gamma ray

output, spectral sensitivity matched to that of the photodetector and also the cost of scintillator.

There is no scintillator that has all of these factors all together. So, in order to employ the appropriate one, careful examination of the application requirement should be fulfilled. Now, let's take a brief look into the literature for common scintillators for PET imaging and their characteristics.

2.2.3. Scintillators for ToF PET Imaging

As pointed out by most of the works in the literature on scintillators, high output power, very short decay time and high density are the main requirements for a scintillator to be employed for ToF PET imaging [1, 12, 13]. Two scintillators that were extensively used and studied at the initiation of ToF PET scanners in 1980s were Cesium Fluoride (CsF) and Bismuth Germanate (BGO) [16]. CsF had a fast timing response, however it provided very poor light output compared to BGO. With advancement in technology and introducing new materials with better characteristics, some scintillators with fewer compromises for ToF PET imaging were proposed. Lutetium oxyorthosilicate (LSO) was one of these crystals which provided 4 times higher light output and 7 times shorter decay time compared to BGO [1, 11] which made it an ideal crystal to be used for ToF PET imaging. During past decade BGO and LSO were the most common scintillators used in PET scanners.

Recently a new hexagonal crystal has been introduced which showed great promising improvements with its unique properties [12]. Lanthanum(III) bromide (LaBr_3) shows faster response time and more luminance. The decay time of LaBr_3 is about half of LSO and its output light is about two times. Day by day, better scintillators are introduced, but cost and compatibility are some important issues which determine which scintillator is preferred to be used in industry. Table 2.1 includes some common scintillators which have been studied and employed more for ToF PET imaging. As can be seen in Table 2.1, peak emission wavelength is another important factor in scintillators; for maximum performance, the peak wavelength of scintillator should be matched with maximum absorption wavelength in photodetector. This increases the photon detection probability (PDP) of the scanner.

	NaI (Tl)	BaF ₂	BGO	LSO	GSO	Plastic	LaBr ₃
Density (g/cm ³)	3.67	4.89	7.13	7.40	7.90	1.03	4.89
Effective atomic number (Z)	51	54	74	66	59	12	46
Decay Time (ns)	230	0.8	300	40	65	2-5	15-26
Photon yield/keV	36	1.8	5.4	27	10.8	10	63
Peak Emission Wavelength (nm)	410	220	480	420	430	various	380
Hydroscopic	Yes	No	No	No	No	No	Yes

Table 2.1: Common scintillators used in ToF PET scanners

2.3. Photodetector

Direct detection of gamma rays without transforming them into visible light is possible by employing semiconductor materials with very large band-gaps and atomic number than silicon such as cadmium zinc telluride (CdZnTe) [1]. However, this method is not a commercial method due to the need for unconventional technology and due to the complexity of design of the readout circuit on the new substrate materials. Therefore, scintillators are widely used to transform gamma rays into lower energy-level photons first. Visible light, which is generated by the crystal, should then be detected using the photodetectors. Photomultiplier Tubes (PMT) have been the most popular photodetector during past decades for PET scanners due to all their advantages including their high sensitivity and fast response time. However, during the past years, new compact semiconductor devices have been studied such as avalanche photodiodes (APDs) which provide more options than PMTs in many ways. In this section we will examine different types of photodetectors and then we will concentrate on single photon

avalanche diodes (SPAD) as this has been the photodetector under concentration in our research group and in this work.

2.3.1. Operation Principle

Photodetectors convert low-energy visible photons into an appropriate electric signal. In PET scanners, photodetectors are used to generate appropriate pulses from the light that is coming from the scintillator. There are several characteristics that define photodetectors. Gain, sensitivity, response time, efficiency, dark count rate (DCR) and after-pulsing probability are some main factors which are used to define the performance of a photodetector in ToF PET applications.

- **Gain**

Gain of the photodetector is the number of secondary electrons generated from a single incident photon. As the mechanism of current generation in different detectors is different, their gains usually differ from each other. In addition to the photodetection mechanism, gain also depends on the bias voltage, temperature and environmental factors. It also should be noted that the gain mechanism acts as a random process [9] and a photodetector covers a range of gain rather than a single fixed gain. This means that the number of secondary electrons can be different for different incident photons. However, this should not be of concern as long as gain is high enough to distinguish between the signal and noise [18]. In addition, variation of the gain must be small enough so that the comparator in the detection electronics doesn't miss a true photon.

- **Sensitivity**

Sensitivity of the photodetector plays a very significant role in determining its performance especially in low-level light applications. Sensitivity of a photodetector is defined as the minimum amount of light which is needed to generate an appropriate output signal with a defined signal-to-noise ratio (SNR). For very low-levels of light in some applications such as TCSPC, in which the ultimate goal is to provide single photon detection, a very high sensitivity is

required. This has become possible with advancement in technology in designing extremely-sensitive photodetectors. PMTs and single photon avalanche diodes (SPAD) are some good examples of such high-sensitivity detectors.

- **Response Time**

Response time in photodetectors is the time from incident light hitting the detector to the time that a corresponding signal is generated. During this time, the photodetector cannot detect any other incidents. So for very high rate of incoming photons this time should be very short. On the other hand, for a very low emission rate, this is not a concern. For time-of-flight PET imaging, the timing information of the incoming photons should be recorded in addition to their number. Since, it is very important that photodetector has a very fast timing response to prevent from implying long time offsets to the time interval to be measured. More importantly, this time offset should be fixed as the time detection process is extremely susceptible to the offset variation introduced by the detector. It is now possible to find very fast detectors in industry with response times less than 1ns [9, 17, 18].

- **Timing Jitter**

Timing jitter or jitter is the statistical time variations from absorption of photons by the detector to sensing by the readout circuit i.e. the statistical variation in response time of the detector. Timing jitter is usually given in full-width half-maximum (FWHM) of the distribution of the arrival time of photons. This parameter, like quantum efficiency, is wavelength dependent. This is because mean penetration depth of photon in the detector is strongly wavelength dependent [17]. Timing jitter is an even more important parameter than response time for ToF PET imaging. This is because what is detrimental for time-correlated single photon counting is variation in time offset added by detector rather than the time offset itself.

- **Quantum Efficiency**

Quantum efficiency of the photodetector is the probability to release a photoelectron or charge carrier from an incident photon. This parameter is similar to sensitivity of the photosensitive device which again should be high enough in low-level light applications such as TCSPC where

no photon should be missed. Quantum efficiency is completely a wavelength-dependent parameter. This is because photons with energy range closed to the band-gap of silicon have more chance to be absorbed and generate a photoelectron. The material and type of photodetector are other factors which affect the quantum efficiency.

- **Dark Count Rate**

Beside carriers that are generated due to incident photons, there are some other carriers which might be generated due to some other reasons in photodetector such as thermal noise and tunneling [6, 17]. This adds a noise to the measurement result. The number of counts per second of these unwanted generated carriers is called dark count rate (DCR). In other word, DCR is the number of avalanches occurring in SPAD when there is no light shining on active area. It is very important that the device provide a very small DCR to keep the accuracy of detector high enough.

Thermal carrier creation, which is the main source of dark count generation, happens due to the traps exist between the valance and conductance bands of silicon. Traps are lattice defects that can hold and release carriers. During operation of SPAD, electrons might receive enough energy from thermal energy to move to these traps. These electrons when released from the traps can have enough energy to initiate an avalanche. This type of avalanche creation is called thermal dark count generation. Tunneling is another mechanism of dark count generation. It happens when a high reverse electric field is applied to PN junction. This gives the electrons in valance band chance to penetrate through the band-gap and turn out to be free carriers. This method of carrier generation is called tunneling which happens in very high electrical fields in the order of 10^6 V/cm or higher [17, 19]. This type of noise is strongly dependent on the doping of the junction and excess bias. Reducing the doping and excess bias can help to reduce this noise. SPAD can be constructed nowadays with very small amount of tunneling-assisted dark noise. Thermal dark noise, however, cannot be suppressed that much. Cooling is the main method to decrease thermally-assisted dark count rate [6, 20].

- **After-Pulsing Probability**

Photodetectors show an increased probability of generating pulses shortly after detection of a photon. These unwanted pulses which do not belong to an incident light are called after-pulses. These pulses add noise to the measurement and should be minimized or cancelled to provide accurate photon detection. In PMTs, after-pulses are believed to be the result of ion feedback, luminescence of the dynode material and the glass of the tube [9, 21]. However, in semiconductor-based detectors, this is mostly due to releasing of some carriers which have been trapped during the detection period in the semiconductor [20, 21]. Unlike DCR, after-pulsing is correlated to number of count and different techniques are employed to minimize after-pulsing probability based on the type of the detector. In TCSPC applications, this minimization is necessary as the high-repetition rate in TCSPC causes after-pulses from many signal periods to pile up and generate a signal-dependent background with significant amplitude. Techniques of after-pulse reduction include reducing the bias voltage and output current as well as intentionally increasing the response time of the detector.

2.3.2. Photodetectors for ToF PET Imaging

Vast progress in proposing new crystals with high output light power and fast decay time leaves timing response of photodetectors the bottleneck in time-correlated photon counting [16, 20, 21]. In order to be used for ToF PET imaging, the photodetector must provide a reasonably high gain with a very short response time. High gain decreases the chance for statistical noise and affects the timing accuracy [12] which is a key factor in ToF PET imaging. In addition, the photodetector must provide high efficiency in the wavelength range of light generated by scintillator to completely absorb photons. Low quantum efficiency of the photodetector at the peak emission wavelength of scintillator results in a considerable reduction of absorbed photons.

Basically there are three types of photodetectors commonly used in PET scanners. Photomultiplier tubes (PMT) have been the most popular photodetector in PET imaging systems. However, the need for compact detectors with CMOS technology compatibility for

implementing in-pixel signal processing circuits has been a great push toward silicon-based photodetectors such as APD and SPAD.

2.3.2.1. Photomultiplier Tubes

Photomultipliers (PMT) are considered as the most widely used photodetector for low-level light detection and TCSPC applications [9, 10]. Their super-high gain and sub-nanosecond time resolution helped them spread as the gold standard photodetector for time of flight PET imaging [12, 21].

A PMT is a vacuumed device with a photocathode and anode which multiplies the photoelectron carriers (which are generated by the incident photon hitting the photocathode) using several amplifying stages called dynodes (Fig. 2.4). Photoelectrons generated from incident photons are concentrated and accelerated with a high electric field which helps them hit the first dynode with an extremely high speed. This causes secondary electrons to be generated from the dynode and these electrons are then directed to other cascaded dynodes which cause further multiplication of carriers. Finally electrons generated in last dynode are collected by anode creating a high enough current on the output of PMT.

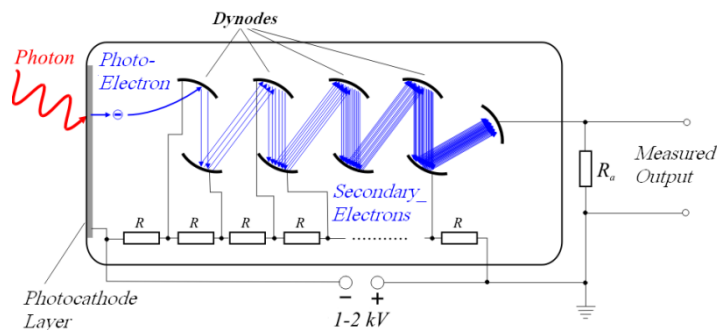


Fig. 2.4: Photomultiplier tube (PMT) structure [9]

PMTs offer a huge gain of 10^6 which implies high sensitivity. Also they are low noise which makes the SNR of the detector considerably high. Big sensitive area and fast time response are some other characteristics of PMTs which makes them ideal detectors for ToF PET imaging [10, 12, 21]. Another significant advantage of PMT compared to other detectors is its stability against a wide range of temperature variation [12].

Early PMTs for PET imaging had one anode and were used in one-scintillator-per-PMT architecture. This was limiting the spatial resolution of PET scanners and the capability to fabricate multi-ring systems as existing PMTs were too large [1]. From early 1980s, designers tried to use multi-channel PMTs for PET systems in which a PMT was shared by many scintillators. This combined the advantages of standard PMTs with the capability to integrate small arrays. In multi-channel PMTs, there are several electron-multiplying channels with separate anode for each, but they share the same photocathode. This helps to increase spatial resolution which was formerly a serious limitation. PMTs with up to 256 channels are now available from Hamamatsu photonics.

Beside its great advantages, PMTs suffer from several serious limitations. Their need for very high bias voltages of about 1KV and their bulky size make them an inappropriate choice for high-level integration applications such as high-channel density PET scanners. Their maximum quantum efficiency is limited to 50% and their high sensitivity to even very weak magnetic fields forces the designers to shield the PMTs even from earth's magnetic field. Another disadvantage of PMT which seriously affects performance of ToF PET scanners is its sensitivity to position and direction of incident light. This variation can sometimes cause a 50% variation in time response and sensitivity of the PMT which can be detrimental for time-correlated photon counting systems [1, 20].

2.3.2.2. Avalanche Photodiodes

A simple idea to implement photon detection is a single PN junction. Incident photon hitting the depletion region of the PN junction can cause electron-hole pair generation. This pair is then conducted out of the depletion region and causes an electric signal. This simple implementation has very high quantum efficiency of more than 90% and provides compactness and high capability of integration. However, they have no internal gain and provide low SNR which limits their operation only to applications with high-level of incident light [1, 21].

Another similar device which solves the limitation of gain in PIN diodes is avalanche photodiode (APD) which exploits the photoelectric effect to convert the light into electrical signal and achieves internal gain of up to 1000 through avalanche multiplication. Electron-hole

pair is generated in the same manner of PIN diode but under a high reverse bias voltage. Carriers are drifted because of this high electric field region in the P-N junction and accelerated inside the diode. These high-energy carriers then generate secondary electron-hole pairs by impact ionization (Fig. 2.5).

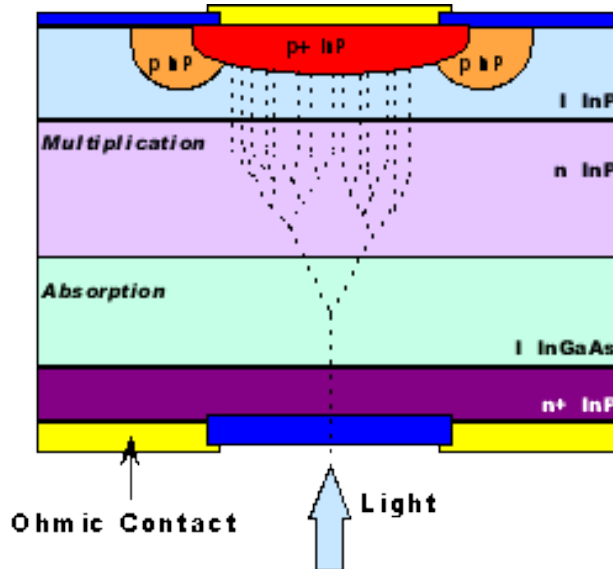


Fig. 2.5: Avalanche photodiode structure [22]

Again, similar to PIN diodes, APD provides compactness which leads to better integration of scanners with higher spatial resolution. High quantum efficiency, low dark current, insensitivity to magnetic field and low bias-voltage operation are some other benefits of APDs [1, 10, 21]. While unlike PIN diode, it provides internal gain, its gain is far smaller than PMT which makes using preamplifier necessary in order to be used for ultra-low light applications. It is good to mention that the time response of avalanche photodiode is poorer than PMTs, which makes it undesired for high sample rate applications such as for time-of-flight PET applications.

2.3.2.3. Single Photon Avalanche Diodes

In PMTs, incident light is focused on the photocathode, from where photoelectrons are extracted to initiate multiplication. This is called “external photo effect” [9] which causes photoelectron to be emitted in all directions including back into the photocathode. This is why quantum efficiency of PMTs cannot go over 50%. Unlike PMTs, photodiodes benefit from “internal photo effect” which makes their quantum efficiency close to 1. APDs provide this high

quantum efficiency while they still lack sufficient gain needed for low-light applications. Geiger-Mode APD (G-APD) as a new device similar to APD has been proposed recently to address these limitations in APDs [23, 24]. As this device can have multiplication gain similar to PMTs, it is also called a silicon photomultiplier (SiPM). Due its high gain, this device has been known to be highly sensitive to even detect a single photon; therefore, it's also called single photon avalanche diode (SPAD). As SPADs are the photodetectors which has been used in our group and in this work for TCSPC, we are going to describe its operation and advantages in detail to help readers completely understand its suitability for ToF PET imaging.

Operation of APD has been explained. The SPAD works very similar to APD but, instead of being biased close to breakdown voltage (as in APD), SPAD is biased well above the break down. This high bias voltage helps free carriers to accelerate more and increases the probability to generate secondary carriers. This is why SPAD offers multiplication gain far above the gain of APD (in the order of 10^7). High electric field and bias voltage demands special design in structure of SPAD. To clarify, any local variation of the field within the SPAD can be damaging to the device. Basically the edges of diode are where these variations in electric field happen and break down in the edges is not surprising if not protected. To avoid this issue, SPADs demand a specific structural as protection to edge breakdown. A lesser doped region surrounding the active area that lowers the electric field in the borders is a common technique that is used to protect the SPADs from edge break down [20, 21]. This lesser doped area is called guard ring. Fig. 2.6 shows internal structure of a SPAD which was designed in our group [25].

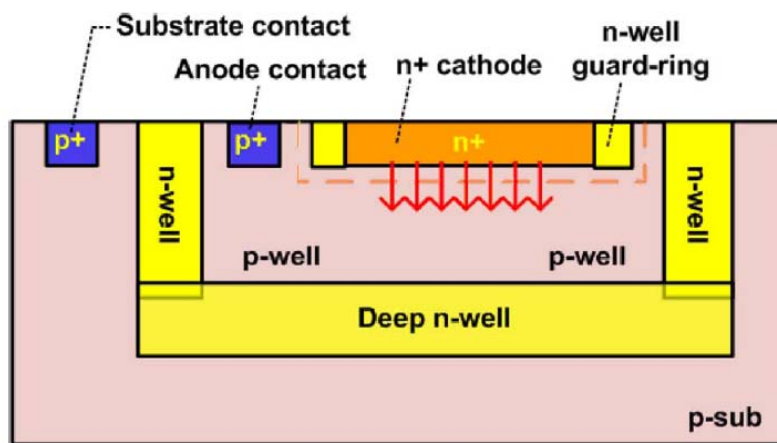


Fig. 2.6: Structure of SPAD designed in our group [25]

As can be seen in Fig. 2.6, when implanting an n-well inside a p-well that is isolated by a deep n-well, the n-well does not go deep enough to touch the deep n-well. This suppresses edge breakdown using an n-well guard ring around active area [25].

Biasing a SPAD in Geiger mode results into a huge carrier generation after incident light enters active region of the diode. This leads to output pulses with amplitudes in the order of millivolts even when SPAD is hit by a single photon. This carrier multiplication if not controlled can destroy the diode. Thus, a quenching circuit which restores the diode to its normal state after detection of each photon is required. There are two quenching methods: passive and active quenching. Fig. 2.7 shows circuits for implementing these two types of quenching. Typically, quenching is achieved by temporarily lowering the bias voltage of SPAD below the breakdown voltage. In passive quenching circuit, resistor (R_L) acts as key element to quench the current. When a photon hits the SPAD, avalanche current will discharge the capacitance in output node. This high current puts a high voltage on R_L which causes the bias voltage of SPAD to decrease below the breakdown voltage and thus output current to be reduced. When bias voltage reaches breakdown voltage, the rate of decrease becomes slower due to long time constant of $R_L C_{out}$. That is why passive quenching is not suitable for fast applications due to its long recovery time [26].

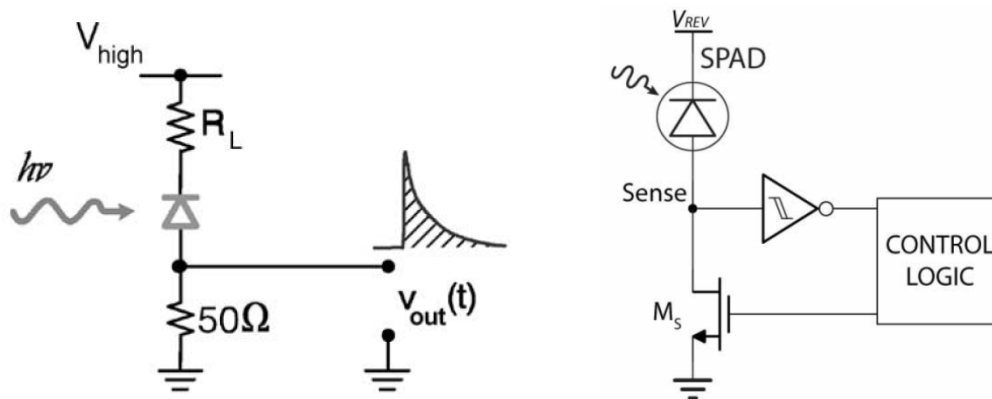


Fig. 2.7: Typical quenching methods in SPADs: left: Passive [27] & right: Active [28] quenching

Another method that does not have the limitation of passive quenching is active quenching (Fig. 2.7-right). As can be seen, the control logic will be used to sense the rising of the avalanche and quickly react back on the SPAD forcing the bias voltage of SPAD going below breakdown voltage [26]. This can be simply fulfilled by changing the current of a source transistor connected to the

SPAD. By using control logic and adjusting the gate voltage of this control transistor, the bias voltage of the SPAD is forced to go below breakdown voltage shortly after onset of avalanche and then return back to below it for next photon detection.

Avalanche event discharges the output capacitor where quenching helps it to recharge back. This forms a pulse on the output node. So, by sensing this very short pulse in the output of the SPAD, a photon arrival incident can be detected. For more clarification on shapes of output pulse of SPAD, fig. 2.8 shows measured pulse shapes for both passive and active quenching circuit that has been designed in our group [25]. Both pulses in Fig. 2.8 show a fall on output voltage during avalanche and reset back to the original value in recovery period. But, the total deadtime of SPAD which equals to sum of quenching and reset periods is much smaller in active quenching circuit (300ps) in comparison to passive quenching circuit (12ns). This promising improvement in speed of SPAD pixels with active quenching circuit, suggests using them for high speed applications such as ToF PET imaging.

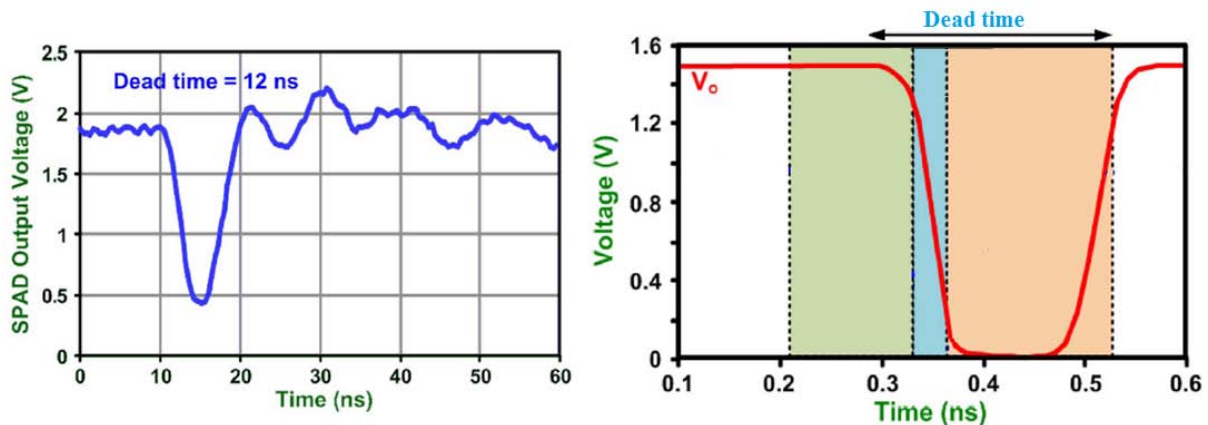


Fig. 2.8: I-V characteristic (left) and output voltage shape (right) of SPAD [25]

By resolving limitations of both PMT and APD, SPAD presents an ideal photodetector for low-level light detection. It provides superb gain up to 10^7 which makes single photon detection possible. In addition, its incredible timing resolution of about hundreds of picoseconds, which is due to thin width of SPAD together with its high amplitude and fast breakdown [21], makes this device an ideal solution for time-correlated photon counting [10]. Moreover, it is insensitive to magnetic field and it offers good linearity over a wide range of incident light wavelength. Its compatibility with standard CMOS technology also is a great advantage which enables

integration of electronics together with the detector. Like APDs high quantum efficiency is an inherent characteristic of SPAD which goes up to 90% [10, 20, 21].

Besides all its advantages over PMT and APD, SPAD suffer from some shortcomings. Thermal noise is a serious issue in SPADs. This is due to energy traps inside the band-gap of the silicon which leads to high dark count rate in SPADs in the order of KHz to MHz [10, 12, 21]. Also, the SPAD structure is more complicated and needs careful design considerations. Another limitation in SPAD is cross-talk which happens due to micro-plasma formation during breakdown of a cell in which electronics are lifted to high band. After quenching when they relax, photon are emitted which can travel to a neighboring cell and cause an avalanche there [21]. [25] shows one of recent SPAD design that has been recently proposed in our research group which is going to be used together with the proposed TDC in this work in a pixel in near future. The proposed SPAD which has been designed in 0.13 μm CMOS technology shows promising sensitivity to perform single photon detection with deadtime of 300ps and dark count rate of 1 KHz in excess bias voltage of 0.5 V.

2.4. Readout Electronics

Pulse that is generated on the output of photodetector usually has an inappropriate shape to be detected with digital circuits. The output of the photodetector first is formed into a standard shape before further signal processing. A pulse shaper is usually employed to transforms SPAD's output pulse into an appropriate pulse which can be used by readout electronics in the next stage. In PET imaging, these shaped pulses are then counted to indicate number of gamma rays coming from each pixel in the PET chamber. So, one vital circuit block in readout circuit of PET scanner is the counter. Based on the output data of the counters in each pixel, the number of annihilation in each LOR can be assessed and then the radioactivity of each pixel can be estimated. In ToF PET however, the arrival time of incident photons should be recorded as well. For this, a time digital converter (TDC) has to be employed to build up the histogram of arrival times of incident photons.

After data generation by the TDC and counter, these data need to be stored and then transferred outside the chip. A high speed SRAM has been designed in our group to store the data and then transfer it to FPGA using serializers and output buffers. Fig. 2.9 depicts the pixel architecture of a SPAD with such readout electronics.

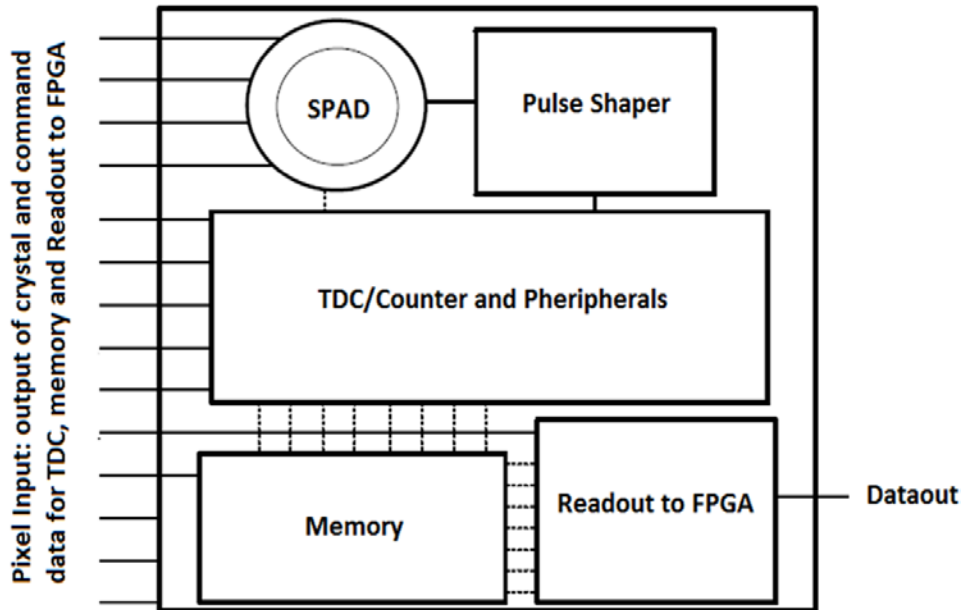


Fig. 2.9: SPAD readout electronics including TDC, memory and circuits to convey data to FPGA

Different types of photodetectors and scintillation crystals were analyzed in this chapter. However, the concentration of this work would be the design of a time digital converter to acquire timing information of incident pulses in a time-of-flight PET imaging system that exploits SPAD as photodetector. Next two chapters of this thesis are specified to describe basics rules in designing TDCs and different TDC structures used for ToF PET scanners. Then we'll illustrate a novel TDC design for in-pixel photon time of arrival measurements.

Chapter 3

TIME DIGITAL CONVERTERS

3.1. TDC Operation Principle

Accurate measurement of time intervals between two or more physical events is frequently needed in many applications in science and industry to provide information on how close events are from each other. These applications include laser ranging [29, 30, 31], 3-D imaging, time-of-flight mass spectrometry [32], fluorescence life time imaging microscopy (FLIM) and in time-of-flight positron emission tomography (ToF PET) which is the target of this work. Precise time measurement can also be employed in equipment testing in high speed integrated circuits [33, 34, 35] and digital storage oscilloscopes and logic analyzers [36, 37] as well as on-chip testing structures [38, 39]. Applications of accurate timing measurement have been extended recently to clock and data recovery, jitter measurements and all-digital frequency synthesis where digital PLLs and delay locked loops (DLLs) help reduce bogus locking states which happens in RF applications [40, 41, 42].

In time of flight PET imaging, the difference between the times of arrival of two gamma rays has to be calculated for each detected event to give the estimation on the exact point of annihilation in the LOR. The difference between these two incoming pulses is measured using a Time Digital Converter (TDC). TDC quantizes this time difference into a digital word in the same way ADC does. However, in the ADC the variable to be quantized is voltage and current whereas in the TDC it is the time difference between two rising or falling edges [43, 44]. The TDC sometimes is called time counter (TC), time interval counter (TIC), time digitizer (TD), time interval digitizer (TID), time interval meter (TIM). However, the term that is the most common in ToF PET imaging is TDC [39].

TDC has been employed in the field of particle and high-energy physics for more than 20 years for accurately measuring time intervals. Looking back to the history of TDCs, the first TDC implementations were on printed circuit boards (PCB) with emitter coupled logic (ECL) components having fairly large size and high power consumption [29, 45, 46]. New applications and specifications require TDCs to be improved in many ways. Prevalence of TDC in industry and its performance improvement has occurred with a fairly constant speed during past decades. But, in past few years, TDCs have become suddenly popular in mainstream microelectronics which raises the question of why. We know that digital signal processing becomes more popular due to the high noise margin and robustness against noise, coupling and variations but the question is why digitizing timing information through TDC have been under a lot of attention during past couple of years. Well, the answer lies in scaling difficulties of mixed-signal circuits in the deep-submicron regime. While voltage levels decrease continuously in new technologies, the intrinsic gain of a single MOS transistor ($g_m \cdot r_{ds}$) is reduced while noise does not scale, so the signal-to-noise ratio (SNR) degrades. There are some methods to increase the gain such as cascode, but with the reduced supply voltages of scaled technologies, there is often not enough voltage headroom to use cascodes.

While technology scaling is against voltage-domain signal processing, it has been advantageous for temporal behaviour of transistors as it leads to reduction of gate delay and improvement of switching speed for digital circuits. So, the implementation of signal processing circuits in time-domain would immediately take advantage of technology scaling. This is where the role of time digital converter becomes significant [41, 47]. Today, there is an upward trend to investigate time-domain signal processing circuit design which will open a new window in mainstream microelectronic design.

While TDC design has been advanced due to the applications mentioned above, in this work, TDC will be used to digitize the time difference between the signals coming from photodetector in a PET scanner, as explained in the previous chapter. As we indicated, the first TDCs for ToF detection were implemented on PCBs and FPGAs separated from the photodetector. However, the increase in time resolution which occurred due to the demand for higher image quality in imaging systems requires integration of TDCs close to the

photodetector. This is because discrete TDC components can easily become the bottleneck in the performance of the imager due to their high dead-time [48]. That is why on-chip or even in-pixel TDC design has been under majority of attention for ToF PET imaging systems during past decade. In addition, in-pixel integration eases circuit integration and reduces the statistical noise which increases SNR of the measurements [49]. This, however, introduces many challenges in designing TDC such as area limitation, readout speed, jitter and nonlinearity, all of which will be explained in the next section of this chapter.

3.2. Important Parameters in TDC Design

There are many parameters identifying the performance of TDC. One of the key characteristics of TDC is its input-output behavior from which many parameters such as resolution, dynamic range and nonlinearity can be extracted. Fig. 3.1 shows the ideal input-output characteristic for a 2bit TDC which maps the continuous time input (x-axis) to the discrete output word (y-axis).

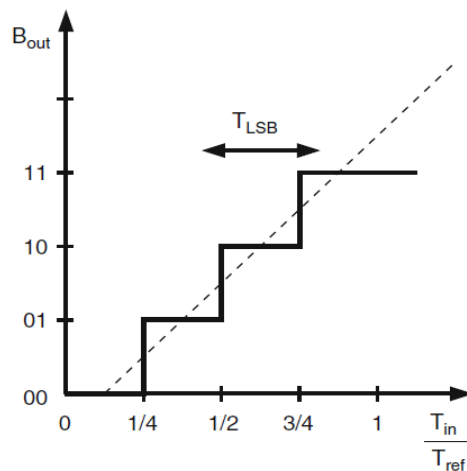


Fig. 3.1: Input-output characteristic of a 2b ideal TDC [47]

T_{ref} is the reference time of the TDC which is the maximum time that can be measured with TDC, similar to V_{ref} in ADCs. On the other hand, T_{LSB} is the minimum time that can be measured which is often called “resolution” of TDC, again similar to the resolution in ADCs. There are some important parameters that have been used more often to describe the performance of

TDCs in literature. These parameters include resolution, dynamic range (DR), conversion speed, accuracy and power and area consumption. The ultimate goal in designing TDCs is to have wide dynamic range with high resolution while having maximum conversion speed and accuracy and minimum power and area consumption. In this section, these important parameters will be described to give the readers a better capability to quantify the performance of TDCs.

3.2.1. Resolution

Resolution of the TDC can be defined both in the unit of time or bits [43, 50]. TDC quantizes the continuous time which means it maps time to discrete output values. This means there is a range of time interval that is mapped to the same output code (Fig. 3.1). Resolution is the width of this time interval or the minimum unit of time that can be measured and it is often called T_{LSB} . Resolution is also defined as the number of bits of the TDC (N), similar to the definition of resolution in ADC. It shows the interpolation level of TDC which is given by:

$$2^N = \frac{\text{Maximum time can be measured}}{\text{Minimum time can be measured}}$$

The resolution depends upon many parameters such as design architecture, circuit characteristics and noise performance. Designers always prefer to have a high resolution or small T_{LSB} as is possible. But, as there is always trade-off between resolution and other parameters of TDC such as dynamic range, area and power. Therefore, usually designers try to design for just enough resolution based on the application specifications.

3.2.2. Dynamic Range

Dynamic range of the TDC, which is usually given in second, is the maximum length of the time window that can be measured with the TDC. This parameter, which is also called the range of the TDC, is sometimes defined in dB. It also can be defined as follows: Input voltage range between the amplitude where signal-to-noise ratio (SNR) is zero dB and full scale [43, 50]. In Fig. 3.1, T_{Ref} indicates the dynamic range of the 2 bit TDC. Like the resolution, TDC should be designed with high enough dynamic range and not more than that based on the specifications

of the targeted application. This has to be noted especially in high-resolution TDCs where having very wide dynamic range leads to bulky, high power systems.

3.2.3. Conversion Speed

In applications where the frequency of the events is high, there is a good chance to miss some of the events happening when TDC is busy measuring the arrival time of the previous event. This is usual when the time needed for the TDC to prepare the digital output code is longer than average period of incident events. This leads to an incomplete spectrum which results in a low-quality final image. Conversion speed, latency, dead time or throughput is the time needed for the TDC from detection of the event to preparing the output word corresponding to that event. It has also been defined as the longest delay path in the TDC [48]. Ref. [51] defined the dead time as the shortest time interval between the end of a measurement and the start of the next one. Conversion speed would be the inverse of the dead time or throughput which corresponds to the maximum event rate that is possible to detect with the TDC.

In many TDC designs which employs delay elements such as Vernier delay line (VDL), pulse shrinking or time amplification methods (as we will see in next sections), the throughput is large which sometimes makes them impossible to use for high-event-rate applications. It is important for TDCs designed for high-repetition rate, real-time applications, to have a fast conversion speed in order to detect all the events.

3.2.4. Accuracy

There are several sources of error in measuring time in TDCs. These sources of error that add the uncertainty to measurements results of TDC, include quantization error, timing jitter and non-linearities. Accuracy, precision or in some cases uncertainty is defined as the average sum of these errors and usually defined as a multiple of the least significant bit. As can be seen from Fig. 3.1, the input code can be calculated based on the output values and T_{LSB} .

$$T_{in} = B_{out}T_{LSB} + \varepsilon \quad 0 \leq \varepsilon < T_{LSB}$$

Here, ε is the quantization error which approaches zero by increasing the interpolation level or number of bits of the TDC. As the occurrence of the events is uncorrelated to the edge of the reference or CLK signal, then quantization error is considered as a random variable. In the ideal case ε behaves like an equally-distributed random signal which contributes only to the noise floor of the measurements [47] and is uncorrelated with the signal. An equally distributed quantization error has an average of:

$$\langle \varepsilon \rangle = \frac{1}{T_{LSB}} \int_0^{T_{LSB}} \varepsilon \cdot d\varepsilon = T_{LSB}/2$$

and a quantization noise power of $T_{LSB}^2/3$. The standard deviation of the quantization error in measurement results can be approximated from Hewlett-Packard equation: $\sigma = \sqrt{\varepsilon(T_{LSB} - \varepsilon)}$

Fig. 3.2 shows the single-shot precision value ε/T_{LSB} vs. the time fraction. As can be seen from this figure, the uncertainty peaks at the middle of the LSB interval in which quantization error has its largest value.

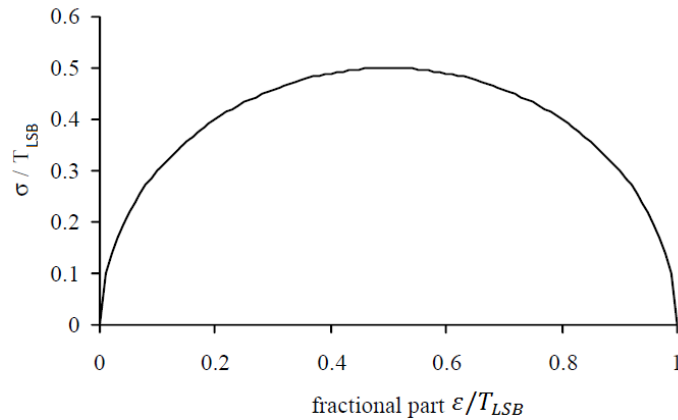


Fig.3.2: Single-shot standard deviation according to quantization error

Nonlinearity of the TDC, as defined in the literature, is any deviation of TDC characteristics from its ideal expected shape. With this definition, any deviation of the step positions and width from their ideal value in TDC input-output diagram leads to nonlinearity [47]. Similar to ADCs, nonlinearity in TDC is usually expressed in two ways: differential nonlinearity (DNL) and integrated nonlinearity (INL). Differential nonlinearity is defined as the deviation of each bin or step size from its nominal actual value (T_{LSB}) or from the average bin width of all bins [43, 50].

INL is defined as the total deviation from the ideal characteristics for each bin for all of the measurements. DNL and INL can be easily measured by introducing a huge number of random start-stop signals and finding the statistical output word-density distribution of TDC. In this method, if C_i is the total count in bin i and C_n is the average number of counts and N is the total number of bins, INL and DNL can be found using following equations:

$$DNL_i = \frac{C_i - C_n}{C_n} \quad INL_i = \sum_{j=1}^i \frac{C_j - C_n}{C_n} \quad \text{where: } C_n = \frac{\sum_{i=1}^N C_i}{N}$$

Nonlinearities basically originate from variation of elements inside the TDC. In most of the designs, there are delay chains to generate delayed version of input signals. Variations of these elements from each other lead to nonlinearities in performance of TDC. These variations can occur due to process, voltage and temperature variations, as well as inappropriate design of the TDC which leads to having different loads at different nodes of the design. DNL expresses the variation of each element compared to others and INL expresses the accumulated error in cascade elements. Nonlinearities can be minimized by properly designing TDC elements and by employing Delay Locked Loops (DLLs). We will explain more about PVT variation control in the next chapter of this thesis.

Time jitter of a signal is the unwanted deviation of the timing information of the signals from its desired real value. Jitter can happen on rise time, fall time and period of the signal. This issue is a serious problem in designing circuits, especially digital and synchronous circuits which are designed to be sensitive to signal edges. In synchronous circuits, as all the circuit blocks are working with one clock signal, any deviation in the input clock leads to a detrimental deviation on the performance and the output of the system. This issue is even more serious in high-speed digital circuits like TDCs with high resolution. In mixed signal circuits such as time digital converters, time jitter introduces random variations in the sampled signal which generate multiplicative de-correlation noise from the coherent signal that is called noise power [52]. The net effect of this is to reduce the total effective number of bits (ENOB).

In the real world, all periodic signals or time intervals have jitter caused by their intrinsic physical way of generation, so it cannot be completely removed from the system. However, it can be controlled in such a way so not to be destructive to the performance and the output of

the system. For example, as the time jitter is heavily dependent on the interval length to be measured [50, 52], it is important to keep the time interval measured by the TDC small enough as this reduces the effective jitter. Another method is to divide the whole time interval into many parts and measure each part by different TDCs to keep the time jitter within acceptable bounds. DLL can also be used to regulate delays and rising edges of the intermediate signals to minimize the uncertainty caused by jitter.

Now let us look at what precise TDC means. Ref. [51] did a great experimental review on accuracy of TDCs and defined a precise time measurement system as a system with a standard uncertainty (S) less than 1ns. It confirms that in best instruments using advanced modern technologies, S is in range of 3 to 10 ps where in TDCs which exploit interpolation it's about 20 ps. The typical uncertainties edges have become even smaller nowadays for sure.

3.2.5. Power and Area Consumption

Power and area consumption may not seem to be so important at first glance, but when it comes to designing an array of detectors in imaging systems such as FLIM or PET, terms like power and area budget become significant issues. In many applications of time interval measurements, it is important to provide high resolution together with wide dynamic range to satisfy the specifications of the design. In most TDC implementation methods, this leads to enormous amount of circuit blocks or delay elements which makes the whole design large. This is inappropriate for in-pixel designs in which detector should be fabricated inside the pixel together with all readout electronics. This also leads to large power consumption which becomes a serious issue in large arrays. A strict in-pixel power budget must be kept in order to minimize IR drops and di/dt effects across the array. However, power consumption is not a serious issue in some applications such as PET imaging (as it is going to be used in a plugged-in, in-hospital application eventually), but underestimating this might cause huge heat generation which implies need for a complex cooling system to be embedded. The area limitation usually is a much more serious issue, which often prevents the TDC to have both wide DR and high resolution.

Due to this trade-off between area, power, resolution and DR of the TDC in imaging systems with arrays of detectors such as PET scanners, a careful compromise should take place in the design of the time digital converters specifically for the targeted application. Therefore, first, the application should be analyzed carefully and an estimation of specifications should be made. Then, the architecture, the whole design, number of elements and all other features of TDC must be calculated. In the next chapter, we will discuss the TDC specifications needed for ToF PET imaging systems based of which the whole TDC design has been performed.

3.3. Basic Ideas in TDC design

There are many implementation methods for designing TDC, each of which has certain advantages and disadvantages. In the following two sub-sections, we will introduce some of the most popular methods and describe their features to give the reader the ability to compare different implementation methods and to understand when each of them can be chosen based on the application and design specifications. In this section, some basic TDC design ideas (usually low-resolution implementations) will be presented and the next section will be focused on high-resolution advanced TDC implementation methods.

3.3.1. Analog Approach

A traditional approach to implement time digital converter is to first transform the input time to voltage or current by a time to analog converter (TAC) and then use an ADC to convert it to digital. As can be seen in Fig. 3.3, in the simplest case, start and stop signals (the signals which define beginning and end of the time interval to be measured) are driven to a pulse shaper to form a pulse with duration equal to the time difference between them. In the next step, an integrator or a sample and hold circuit will charge a capacitor whenever this signal is high. Then, the voltage on capacitor which indicates the time difference is converted to digital word using an ADC. In this method, the resolution and dynamic range of the TDC is heavily

dependent on the resolution of ADC which is limited by analog constraints [47]. This design is simple and easy to implement but, it should be noticed that analog designs are not suitable for technology scaling, they are susceptible to noise and dissipate large static power [43, 47]. Also, all parts of the design such as the time interval logic, charge pump and sample and hold circuit should be linear to meet the linearity of the TDC which is hard to attain. For example, basically charge pumps are made with current sources which are not linear circuits due to their finite output impedance.

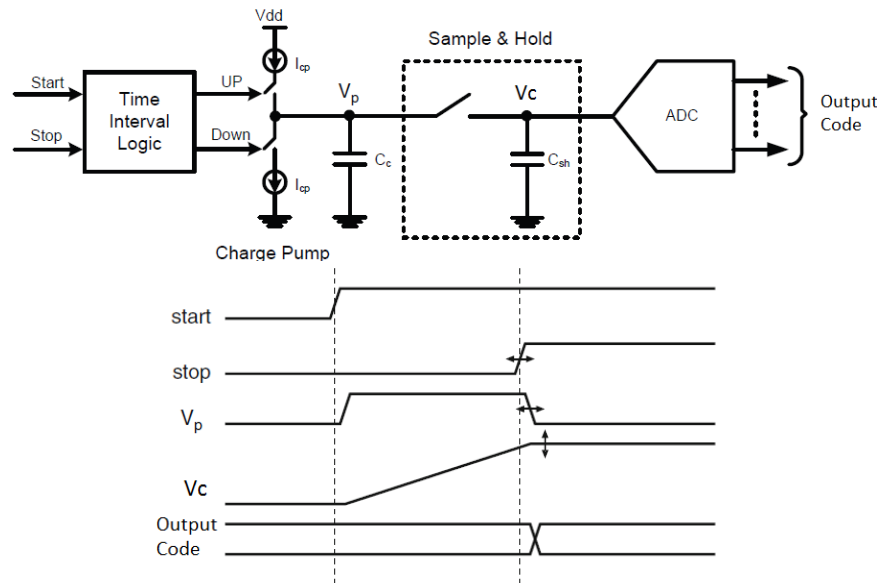


Fig. 3.3: Analog implementation of TDC: architecture (top), timing diagram (bottom)

All in all, although this method is the first thing comes to mind for implementing a TDC. Although this method is simple, it has many constraints due to analog implementation and usually is avoided by designers.

3.3.2. Counter Based Implementation

When the design of the TDC needs to go beyond just a simple timer and be used for many applications as a building block, its suitability in scaling with CMOS technology trend plays an important role. That is why designers tended to switch to digital approaches for implementing TDCs. Also, time-domain signal processing (as one of the key reasons to start new generation of

TDCs) was emphasized as the way to stay away from voltage-domain, so, using analog approach which involves voltage-domain signal processing is in contradiction to the preliminary goal of time digital converter design as mentioned earlier.

The simplest digital approach for implementing a TDC is to employ a counter that counts the number of CLK cycles in between start and stop signals. In this method the start signal enables the counter and the stop signal disables it. The resolution in this design is the CLK period and the stability of the CLK sets the accuracy. This method, however is simple and easy to fabricate and it provides a wide dynamic range and fast conversion speed, it cannot provide high resolution. Resolution can be increased by increasing CLK frequency but targeting to resolutions less than 1ns implies CLK frequencies higher than 1 GHz. However, generating these high-frequency signals not only makes the whole design complex, but also imposes high power consumption and accuracy degradation as CLK jitter becomes significant. Using a lower frequency CLK signal is applicable, but this leads to limited resolution and high quantization error. Quantization error which occurs because the start and stop signals are asynchronous with the reference CLK signal can be very large in low CLK frequencies.

3.3.3. Clock Cycle Interpolation: Tapped Delay Line

The counter method can be employed in applications where a wide dynamic range is required, but not a very high resolution. In order to design for higher resolutions without increasing the CLK signal frequency, interpolation of the CLK signal may be used by employing a delay line in the path of the start signal.

As can be seen from Fig. 3.4, interpolation of the reference signal is fulfilled by introducing a delay chain to the start signal to generate its delayed versions. These delayed versions of the start signal must then be sampled on the arrival of the stop signal using a chain of the DFFs to perform time interval measurement. The basic concept is that, at the arrival of the stop signal, all delay stages which have been already passed by start signal, will have “High” output and all stages which were not, will have “Low” output values (Thermometer code). As a result, the more time difference between start and stop signal, the more “Highs” will be detected in the output of the DFFs.

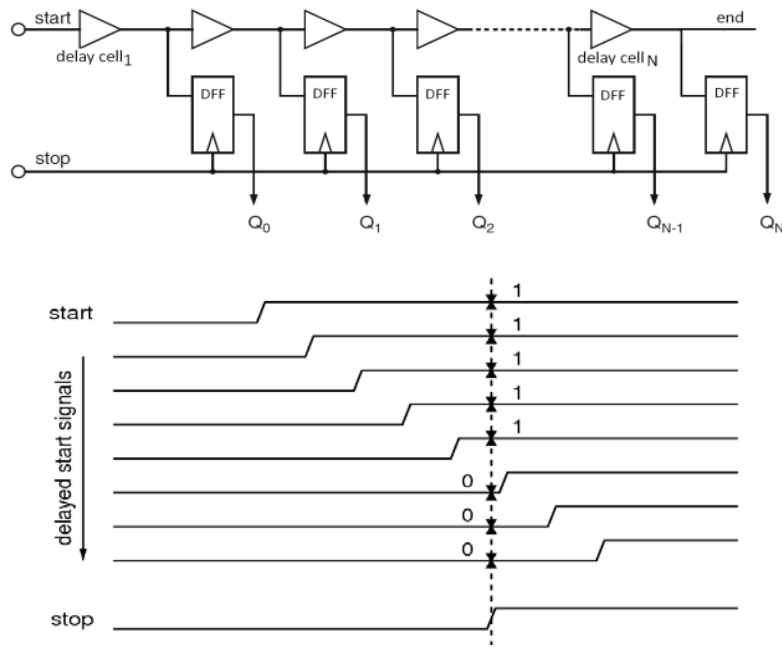


Fig. 3.4: Tapped delay line TDC: architecture (top) and timing diagram (down)

The main difference between this method and counter method is that the resolution here is not determined by the CLK frequency. Indeed, the LSB in this approach is indicated by propagation delay of the delay cells. This means, that the resolution is limited by the minimum delay of the delay elements. So, to increase the resolution, simple fast gates with extremely small propagation delay, such as buffers and inverters, are required. The dynamic range is defined by the total number of delay elements in this approach which can be increased at the expense of power and area consumption. To avoid large, power hungry designs and to increase the range at the same time, a tapped delay line can be used in a loop with a counter, but this is at the expense of adding nonlinearity and complexity to the system.

A tapped delay line TDC provides high conversion speed similar to the counter approach i.e. the data would be ready just after arrival of stop signal. However, this design extremely suffers from nonlinearity as the propagation delay of delay elements is so sensitive to process, voltage and temperature variation and there is a good chance that delay amounts of cells inside the chain differ a lot from each other which leads to a heavy degradation of accuracy. Therefore, in order to keep the input-output characteristics of the TDC close to its ideal shape, a control mechanism is necessary to regulate delay amount of delay cells during different

conditions to keep the measurement result robust. This is done by employing delay locked loops (DLL).

3.3.4. Delay Locked Loop

The tapped delay line TDC performance is very sensitive to process, voltage and temperature (PVT) variations as these variation which can have significant effect on the gate delay of the delay element. To control the amount of propagation delay due to these variations the delay locked loop (DLL) structure can be employed (Fig. 3.5). The operation of DLL is similar to PLL i.e. DLL tries to synchronize the start signal with its delayed version at the end of the delay line. For this purpose, a periodic start signal will be used to feed the delay line and the signal at the beginning and end of the delay line will be injected to a phase detector. Relative to the phase difference between these two signals, the charge pump generates a voltage at its output and this voltage will be fed back to the delay elements to either increase or decrease their propagation delay. This goes on until the rise or fall edge of the end signal locks to the start signal. This keeps the total delay amount of delay line fixed (equal to an integer number of the start signal's period) during the measurements.

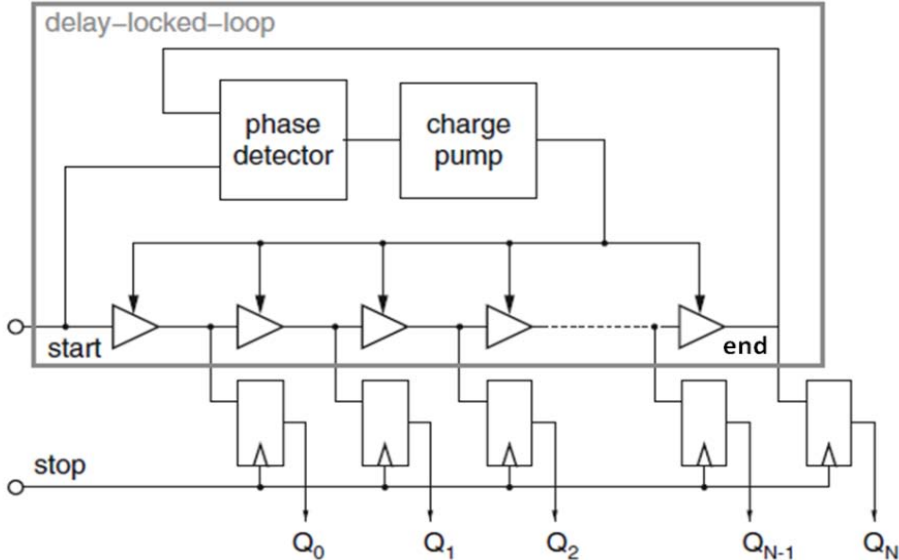


Fig. 3.5: Delay locked loop (DLL) structure

The DLL design provides good resolution and by adding a counter, it offers a wide dynamic range. However the counter adds nonlinearity as the load seen in the last node in the delay chain is different with the load in the intermediate nodes. False locking is another issue which happens when the phase detector tries to synchronize the delayed version of the start signal with a wrong edge of the start. To avoid false locking, a phase-frequency detector can be used instead.

The DLL together with a loop counter provides sufficient performance for many of the applications and it is a robust method as it regulates the TDC parameters using the locked loop. However, for applications that require very high resolution, this method cannot be used like other previous mentioned methods. This is because the minimum time interval that can be measured is limited to the gate delay of the delay cells. The fastest gates such as inverters are not able to provide delays less than 100ps. So, to achieve resolutions better than that, a new category of methods should be used. These methods are called “sub-gate delay resolution approaches”.

3.4. Sub-Gate Delay Resolution TDCs

The only common feature between all of the methods that were mentioned in last section was their low resolution. In DLL, the resolution is limited to the minimum gate-delay which is limited by technology. This implies using a completely different approach in applications where a least significant bit less than 100ps is required. These methods are called sub-gate delay resolution approaches as they try to reach resolution beyond the minimum gate delay of digital gates. As we will explain in next chapter, ToF PET requires a resolution better than 100ps. Consequently, based on the ultimate goal of this project, a TDC with resolution beyond sub-gate delay should be designed and implemented. So, some of most popular sub-gate delay TDC implementation methods will be introduced briefly in this chapter. In most of these approaches, the main concept is based on parallel scaled delay elements to realize high-resolution TDCs.

3.4.1. Vernier Delay Line

One of the most popular high-resolution TDC implementation techniques is Vernier delay line (VDL) technique which was first employed for high resolution time measurement in 2000 in [53]. In this method, as can be seen in Fig. 3.6, the structure is similar to the tapped delay line except that the delay elements are used in the stop line as well as in the start line. Delay elements in the start line have delay (t_1) slightly bigger than the delay of delay elements in the stop line (t_2). These delay can be realized by a voltage controlled delay element which is regulated by a separate delay locked loop for both start and stop lines. Due to the slight difference between delay cells of start and stop delay chains, the delay between the start and stop signal will be decreases at each stage until these two signals catch together and that is when output code will be available to sample.

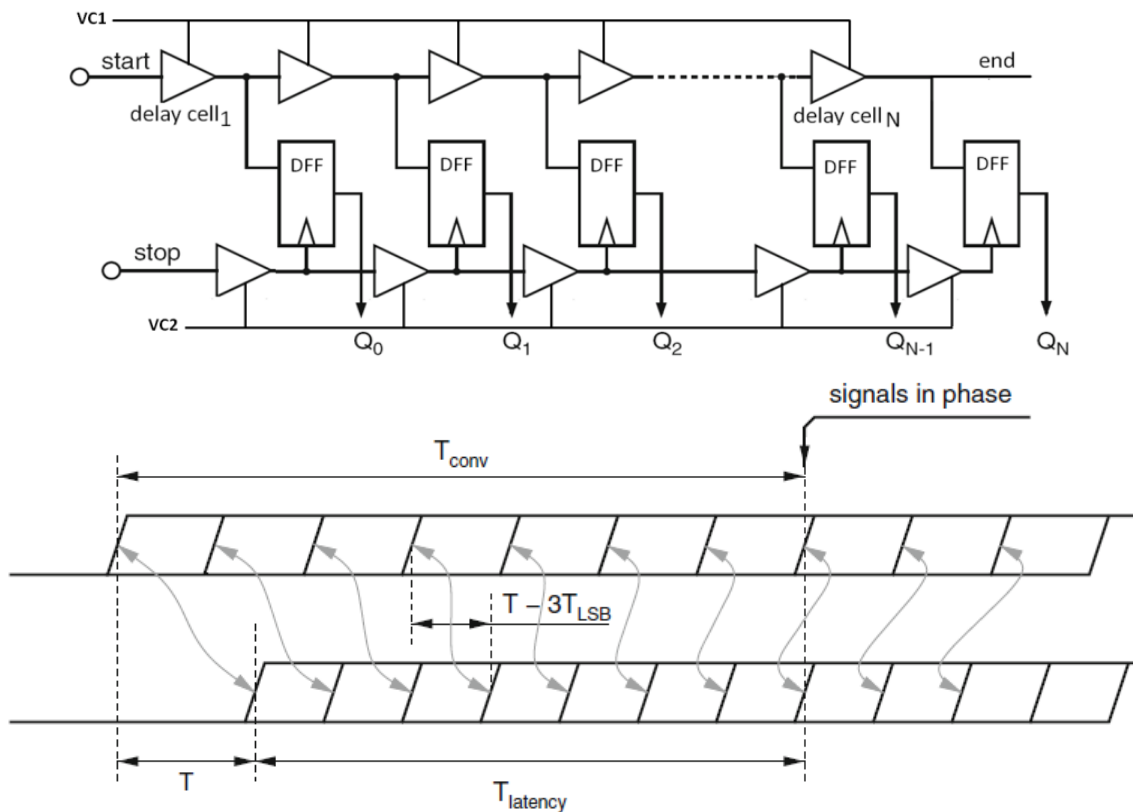


Fig. 3.6: Vernier delay line TDC technique: structure (top) timing diagram [47] (down)

The VDL technique provides a theoretical resolution of zero ($t_1 - t_2$) at the expense of area and power, but reaching resolutions beyond picoseconds is improbable due to noise limits. The dynamic range of the VDL is indicated by the number of delay cells. In order to use a VDL for a wide time measurement range, a huge number of cells should be employed which adds area, power and cost. The accuracy of the VDL is relatively good as it is an all-digital method and the locked loop can be also used to regulate delay cells. The conversion speed is a little bit worse compared to DLL though, as it takes time for the stop signal from appearance at the beginning of the chain to the time it catches up with start signal when data would be ready. Another problem in VDL is that the latency depends on the length of time interval, so, this should be considered when designing the readout topology.

3.4.2. Pulse Shrinking Method

Pulse shrinking is another sub-gate delay resolution method which is used for designing high-resolution TDCs [54]. The idea is to pass the width-coded pulse (width of the pulse indicates duration of time interval) through a delay chain. By passing each of delay stages, the pulse width is decrease by a certain amount which is determined by the delay elements. The pulse goes on until it vanishes, so the number of cells the pulse passes when it is still high determines the width of the pulse. The pulse width is decreased through passing a delay line with intentional asymmetry in each cell. A very basic implementation of a pulse shrinking delay stage can be two in-series inverters with different rise and fall times (Fig. 3.7). Assuming risetimes tr_1 and tr_2 and falltimes tf_1 and tf_2 of the inverters, the reduced pulse width in each stage which indicated least significant bit of the TDC is given by:

$$T_{LSB} = (tr_2 - tr_1) - (td_2 - td_1)$$

Again the least significant bit can be decreased as much as necessary by adjusting the delay amounts to keep the T_{LSB} close to zero until noise becomes significant. Similar to VDL in this method, the number of delay elements defines the dynamic range of TDC. So having a very high resolution and high dynamic range requires a large number of delay stages which increases the area and power consumption.

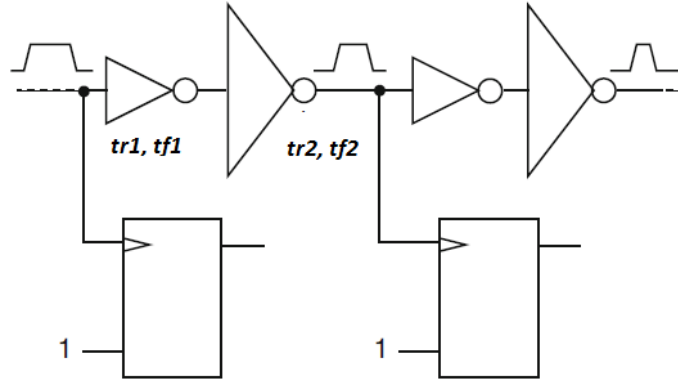


Fig. 3.7: Pulse shrinking TDC structure: using two inverters with different sizing as delay stage

The latency in this method is linearly dependent on the time interval to be measured, similar to VDL and it is usually long compared to other methods because, TDC should wait until the width-coded pulse vanishes to give up its output word. The accuracy of the pulse shrinking method is poor as the delay elements are very sensitive to PVT variations. Also, unlike the tapped delay line and VDL, it is hard to regulate their delay in a DLL structure. Another serious issue in the pulse shrinking method is the DFF output uncertainty in their very last stages. As a minimum pulse width is needed in the input of the DFFs in order to be detected as 1, the least significant bits of the TDC in this method becomes inaccurate when it comes to very high resolutions.

3.4.3. DLL Array TDC

In DLL array TDC which was introduced by [55], the basic idea is to delay the start or reference signal using delay elements as in the DLL, but the difference is that, the signal is going to be delayed in many delay chains instead of only one. Fig. 3.8 shows the architecture of this design. As can be seen from this figure the array is made up from several delay lines. The signal in the input of each of row delay lines has been delayed first with a column delay line which has delay cells with delay different than the ones in the row lines. So, least significant bit of TDC can be calculated as:

$$T_{LSB} = T_{Ref} \left(\frac{1}{m} - \frac{1}{n} \right)$$

where m and n are the number of delay cells with delays t_m and t_n , as in Fig. 3.8. A time diagram for $m=28$ and $n=35$ can be seen in the Figure which leads to time interpolation of reference cycle by 140 times. Again, the resolution can go close to zero by adjusting m and n in this method. However, it is really hard to attain T_{LSB} less than few picoseconds due to the jitter of the signals and delay elements. This method, however, is an effective and accurate way to increase the resolution, but it needs a large array of delay cells to be used for wide-range applications which leads to a bulky power hungry TDC. An array made with a single DLL and RC delay line can act similar to the DLL array in which multiple sampling signals are generated using RC delay cells [56].

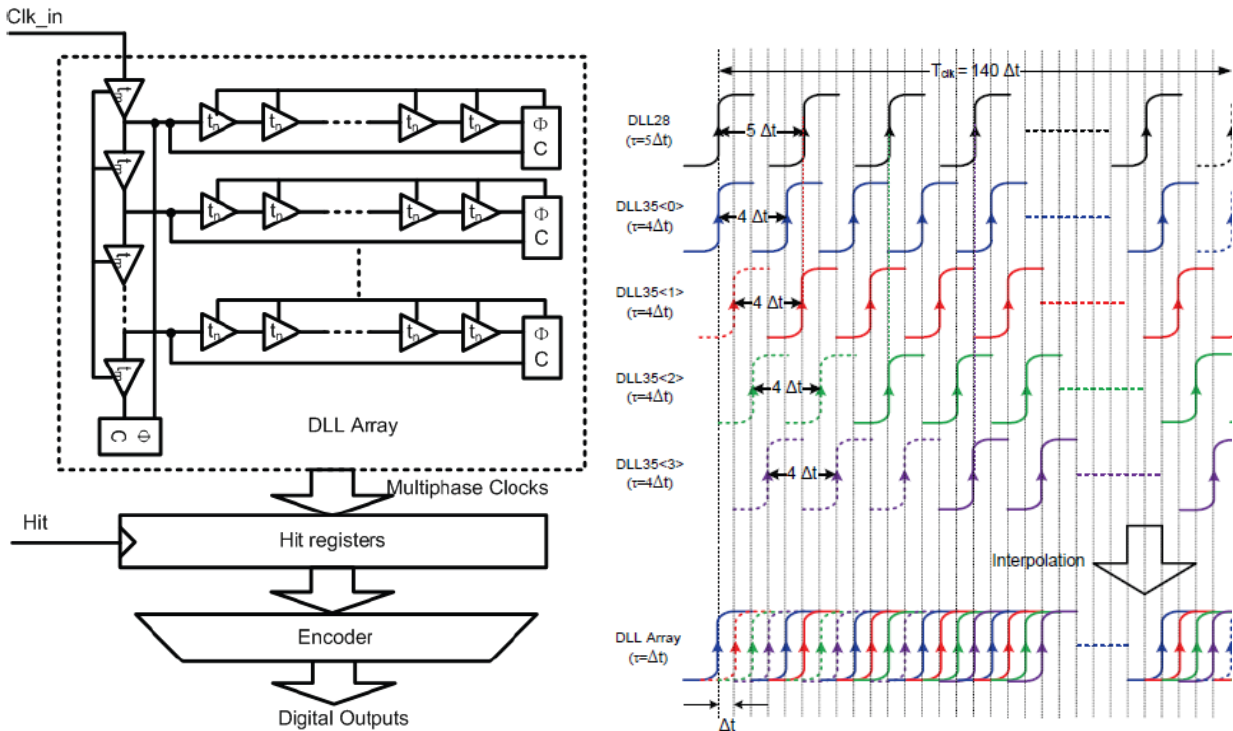


Fig. 3.8: TDC implementation using the DLL array: architecture (left), timing diagram (right) [43]

In the RC delay line method, the resolution can be improved by increasing the number of delay elements and RC delay cells. Also, it might lead to less area consumption than the DLL array if the resistors and capacitors are integrated with the transistors. However, the accuracy is poor due to varying resistor and capacitor values and the fact that they are susceptible to PVT variations. In addition in this method, in order to achieve high resolution, it is necessary to implement very small R and C values inside the array which makes the effect of parasitic

capacitors and resistors significant. Since the model for parasitics in MOS technology is not varying accurate, there will not be accurate delay estimation for each cell which leads to degradation in the TDC's precision [43].

3.4.4. Local Passive Interpolation

The local passive interpolation technique which was introduced in [41] uses the idea similar to voltage interpolation in interpolating flash TDCs. Reaching sub-gate delay resolution is fulfilled by subdividing the coarse time interval given by the delay line. In this method, intermediate signals in between the two adjacent delay cells in delay line are made with mixing them as can be seen in Fig.3.9. Here, V_A and V_B are two signals on two adjacent delay cells and the middle signal is the interpolated one. These interpolated signals can be easily generated using a resistor chain in between each two adjacent nodes in the delay line (Fig. 3.9).

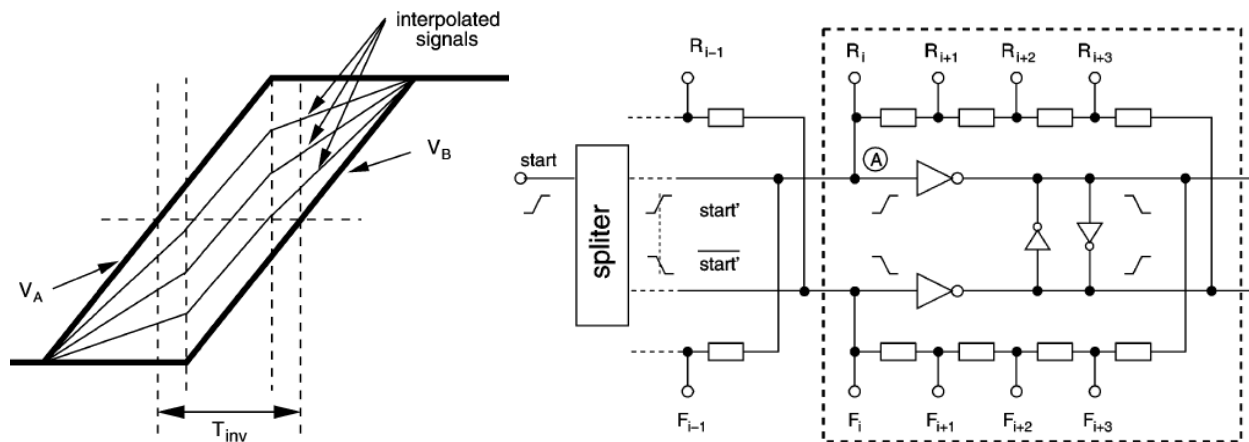


Fig. 3.9: Local passive Interpolation technique: interpolated signals (left), realization (right) [41]

In this method resolution is gate-delay divided by number of passive elements in each stage. The resolution in this method can be increased by reducing gate delay (technology scaling) and also by increasing the number of passive components. Moreover, unlike VDL and the pulse shrinking methods, this approach does not suffer from poor conversion speed and the measurement time is not increasing with higher resolutions. Also, using passive interpolation makes the design a little bit more robust to local variation as the passive interpolation translates any variation to a subdivided variation of the intermediate signals [41].

A very important issue in the local passive interpolation technique is the fabrication of passive elements in MOS which takes a lot of space. In addition, this technique may suffer from resistor mismatch which leads to poor linearity. Moreover, this technique employs power hungry resistors for interpolation which makes this method practically unsuitable for low power applications.

3.4.5. Time-Amplification Method

Providing both high resolution and wide dynamic range usually requires a large number of digital circuit blocks as delay or interpolation elements or interpolators. These interpolation elements usually result in large area and increased power consumption. Inspired by resolution improvement technique by amplifying the difference between the input and the closest coarse level in coarse-fine ADCs, the idea of time amplification to increase the resolution of TDC without requiring large number of delay cells was introduced in [44]. Implementation of this idea is done by exploiting the variable delay of an SR latch subject to nearly coincident input edges. Fig. 3.10 shows the complete structure of the proposed time amplifier and its output-input characteristics. The time amplification method helps to solve the problem of increased area and power for high-resolution high-DR applications. Using this method, extremely short time intervals can be amplified and then measured with a coarse TDC with low resolution.

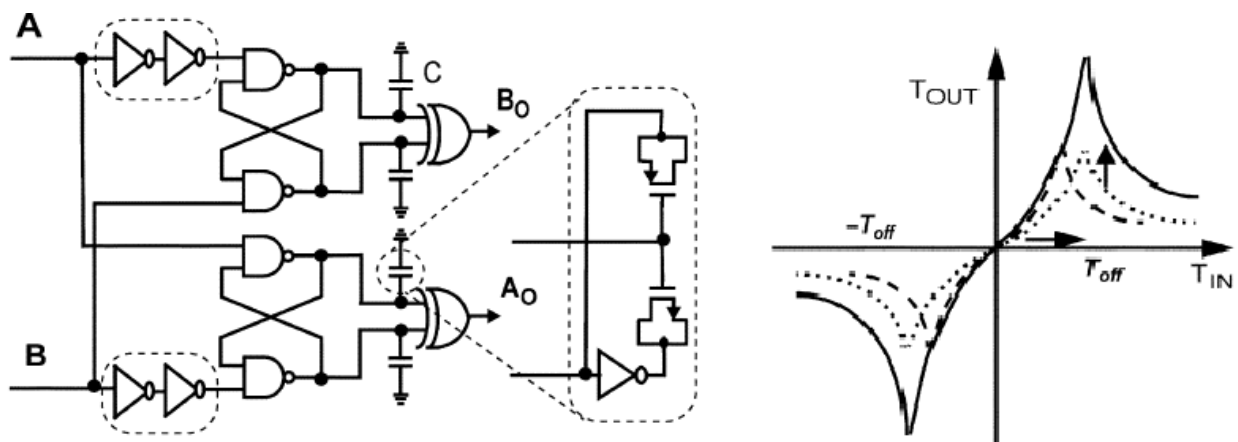


Fig. 3.10: Time amplifier implementation (left) and its output-input characteristics (right) [44]

As can be seen in Fig. 3.10, the problem with this implementation is the nonlinearity in its characteristics and there is just a small linear range for amplification. Indeed, the total linear range is very small (less than 50ps), so it must be used in a coarse-fine measurement system in which the major part of the time interval is measured with a coarse TDC and the residual is amplified and measured again with the same coarse TDC. It should be said that with all its problems, the idea of time amplification is new and it points to new research direction in TDC design.

3.4.6. Hierarchical TDC

The last architecture to be introduced in this chapter is hierarchical or hybrid TDC. In this method, a combination of TDC implementation techniques mentioned in this chapter is used in a coarse-fine architecture. There are variations in the realizations of this method. For example, a counter as coarse TDC together with a DLL as the fine interpolator is one of the most popular hybrid architectures. The DLL-VDL, counter-VLD, counter-time amplifier are some other popular hybrid TDC implementations. Fig. 3.11 shows a hierarchical TDC with DLL performing as the coarse TDC. In this method, by employing different architectures, the TDC can be designed for a specific application i.e. a hierarchical TDC is suitable for an application-specified design. But more importantly, its great benefit is the large area saving achieved using this architecture. This area reduction is obtained by dividing the entire time interval into several parts, as can be seen in Fig. 3.12.

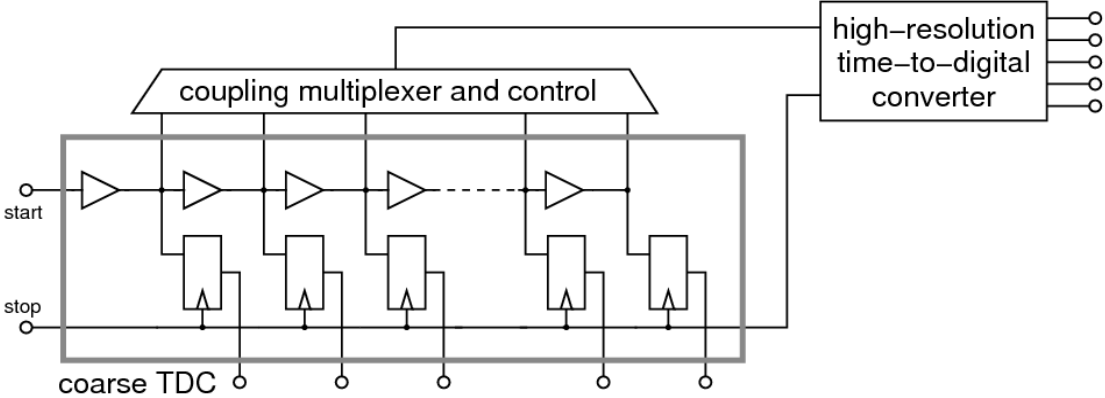


Fig. 3.11: Hierarchical TDC architecture using a DLL as the coarse TDC

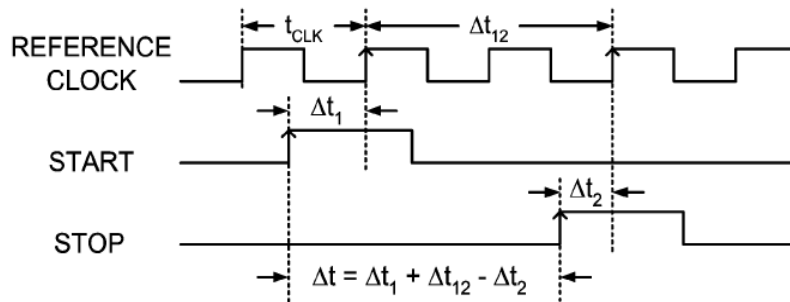


Fig. 3.12: Timing diagram for a hierarchical TDC

The coarse TDC measures the major part (Δt_{12}) and therefore very small portions of the interval have to be measured with a high-resolution TDC (Δt_1 and Δt_2). This means much less delay elements for the VDL and hence, high resolution together with wide dynamic range that is achievable by this technique. This can be very beneficial in in-pixel TDC implementation that is required in ToF PET scanners in which the ultimate goal is to integrate the photodetector, quenching circuits and other digital blocks together with TDC inside the pixel. Good conversion speed and accuracy can be obtained by carefully designing the TDC. The only issue is the nonlinearity and complexity added by the interface between the coarse and fine TDC, so these should be carefully considered in designing Hierarchical TDC. We will explain hierarchical TDC in detailed and illustrate why this method is of the most suitable for our ToF PET application in the next chapter.

Chapter 4

PROPOSED TDC PIXEL

4.1. Introduction

Design procedures typically start with a set of design specifications. Since the goal of this work is to implement an in-pixel TDC that can measure time difference between the arrival times (called time-of-flight) of photon from the target in a PET imaging system, then it is vital to know what the expectations are on the biomedical side of the project. The time interval measurement specifications in ToF PET imaging indicates the parameters needed for a TDC to be designed. Based on literature and on feedbacks from our collaborators in the Molecular Imaging Instrumentation (MII) group in the Medical Physics department, the design specifications are as follows.

Improving time resolution is always appreciated as it reduces statistical noise and advances the image quality of the PET scanner. From expectations in the biomedical side of the project, the spatial resolution should be around a centimeter or less. This spatial resolution is correlated to a timing resolution less than 100 ps ($\Delta t = \Delta s/c$). By assuming the maximum diameter of the PET detector chamber to be around 1 meter, then the dynamic range of the TDC can be found by dividing this value by the speed of light given a dynamic range of 10 to 20 ns. To use the TDC for other testing purposes such as dark count measurement of the SPADs, the dynamic range must be wider. As a result, to satisfy the specifications, a dynamic range of 200ns was targeted in this work. The emission rate from the target depends on the dose of radionuclide that has been injected into the patient. But usually it ranges from 2 to 5 MSingle/s. This implies maximum allowable total deadtime of 200-500ns for scintillator, detector and TDC.

Based on the deadtimes of photodetector and scintillator which are in the range of nanoseconds, the TDC's maximum deadtime is found to be 200ns.

The total size of the TDC can be found from the cell size in PET scanner. This can be found from the scintillator cell size which is between $1 \times 1 \text{mm}^2$ to $3 \times 3 \text{mm}^2$. Considering the total image sensor size of $2 \times 2 \text{mm}^2$ and assuming to have at least 16 by 16 arrays of detectors in each cell, the size of each pixel which contains an SPAD, driving and quenching circuit, some signal processing circuits and a TDC can be estimated to be about $100 \times 100 \mu\text{m}^2$, resulting in a very serious issue of size limitation. As mentioned in the last chapter, to achieve high resolution and also wide dynamic range as in the VDL method, a large number of delay cells are needed and that leads to an area-intensive design. In order to solve this problem, a hierarchical structure can be used in which a major portion of the time interval is measured with a coarse TDC, and the remainder, which is a small portion of time, is measured with fine TDC. Let's consider our case which requires resolution of say 100ps and dynamic range of 200ns. By employing VDL technique, 2000 delay stages are needed. But if we divide the 200ns time interval into two parts, each measured with separate TDCs, then the total number of stages would be a lot less. Again considering VDL approach, by having a VDL TDC with resolution of 100ps with 40 delay stages and another coarse TDC with resolution of 4ns with 50 delay stages as the two building block of a whole hierarchical TDC, then a dynamic range of 200ns is obtained. The total number of delay stages with the method would be $40 + 50 = 90$ which is much less than 2000. Due to this great area saving benefit, we will use the hierarchical TDC implementation in our design. The hierarchical structure and its operation principle have already been explained in the previous chapter.

As mentioned before, the conversion speed in hybrid TDC depends on the conversion speed of its building blocks. By employing TDC structures with low latency such as a counter and DLL, a flash TDC can be achieved even with the hybrid structure. In fact, a hierarchical TDC has a better conversion speed than other circuit implementations. This is because just a very small portion of the time interval needs to be measured with a high-resolution TDC which usually have more deadtime compared to a coarse TDC. Therefore, we implement a fast hierarchical TDC with high resolution and wide measurement range.

Other design specifications are as follows. Employing digital design approaches is a key in our design to completely profit from CMOS technology scaling. As a result, any analog circuit such as the TAC is avoided. In order to compensate for process, temperature and voltage variations which gravely degrade the linearity and therefore the performance of TDC, PVT variation control block is going to be used to regulate the delay cells and maximize the linearity. Power consumption is not a very serious issue in our design as TDC will be used in a PET detector chamber, not a portable device that demands energy saving. Still, careful consideration of power of the system is necessary as high power consumption adds to the requirements for a complex cooling device for the PET chamber.

Based on all these design issues and specifications, a TDC is proposed. The whole architecture of the design and all design procedures are be described in this chapter, together with methods and techniques that have been used.

4.2. Proposed TDC Architecture

Here, we explain the designed architecture of the TDC and we illustrate the reasons for choosing this architecture. Due to all advantages of hierarchical method for our application (designing an in-chip TDC with fairly high resolution and wide dynamic range for ToF PET imaging as mentioned in above), our design is based on this method. In most of the hybrid designs in literature, a reference clock signal with a fixed period has been used and interpolation has been done in this period to increase the robustness of the design. Following this fact, we chose a stable clock signal coming from the FPGA for further interpolation and a digital counter for counting the reference clock cycles in between the two hit signals (start and stop) as a coarse TDC. The advantage of this coarse TDC is that the digital counter provides a simple robust CLK cycle measurement which easily can be extended to increase the dynamic range of the TDC.

The FPGA that is available for our project (Cyclone II) provides reference signals up to 200 MHz generated with internal PLLs. To increase the resolution of TDC, this maximum CLK

frequency was targeted in this work. The CLK frequency of 200MHz is equivalent to 5ns period, that is, the time in which interpolation should take place. The Minimum number of elements or interpolation levels required is quickly found from this CLK period and targeted resolution. As a resolution less than 100ps is one of the key benchmarks of this project, the minimum interpolation level should be 50. This can be fulfilled by employing 50 delay stages in such a way that the total delay of the delay line equals the reference CLK cycle. As a resolution below 100ps should be achieved in this work, sub-gate delay resolution techniques should be employed. This means the delay line should be formed, for example, in a VDL or pulse shrinking structure. But as mentioned in previous chapter, these methods are bulky; especially when a wide dynamic range is desired (5ns is equivalent to 50 stages). Therefore, we decided to execute two interpolation stages by employing two different delay stages. Instead of having 50 elements, two interpolation stages each with 8 elements has been designed (two fine TDCs). This not only reduce the total number of elements from 50 to 16, but it also decrease the total power consumption and cross-talk between the elements due to reduction of the switching nodes. In addition, the number of delay stages with sub-gate delay will be decreased.

The dynamic range of the first interpolation stage should be equal to the reference CLK period which is 5ns. So, by having 8 interpolation levels, the resolution of the first fine TDC would be 625ps. Since this is more than the minimum gate delay of digital circuits, then the first stage can be implemented using one of the low-resolution methods mentioned in the previous chapter. A DLL has been used for this purpose with 8 delay stages with adjustable delay to provide flash operation and good temperature and process stability. For realization of the second interpolation stage, low-resolution techniques are not applicable anymore. The dynamic range of the second fine TDC should be equal to the resolution of the first fine TDC which is 625ps. So the expected resolution would be around 80ps for 8 interpolation levels for the 2nd fine TDC. This implies utilizing a sub-gate delay TDC technique. The VDL technique has been chosen to implement second stage due to its advantages mentioned in previous chapter. The VDL stage is a robust method if delay cell regulation is performed. Also, it does not have the pulse shrinking method's last stages problem and non-linearity of the time amplifier. It seems to be a good candidate for our requirements for the 2nd interpolation stage.

So the whole TDC consists of 3 stages: Counter as coarse TDC, DLL as first fine TDC and VDL as the second fine TDC. An overall sketch of the architecture of the designed TDC consisting of these three TDCs as can be seen in Fig. 4.1. The digital counter measures the number of clock cycles in between the hit signals.

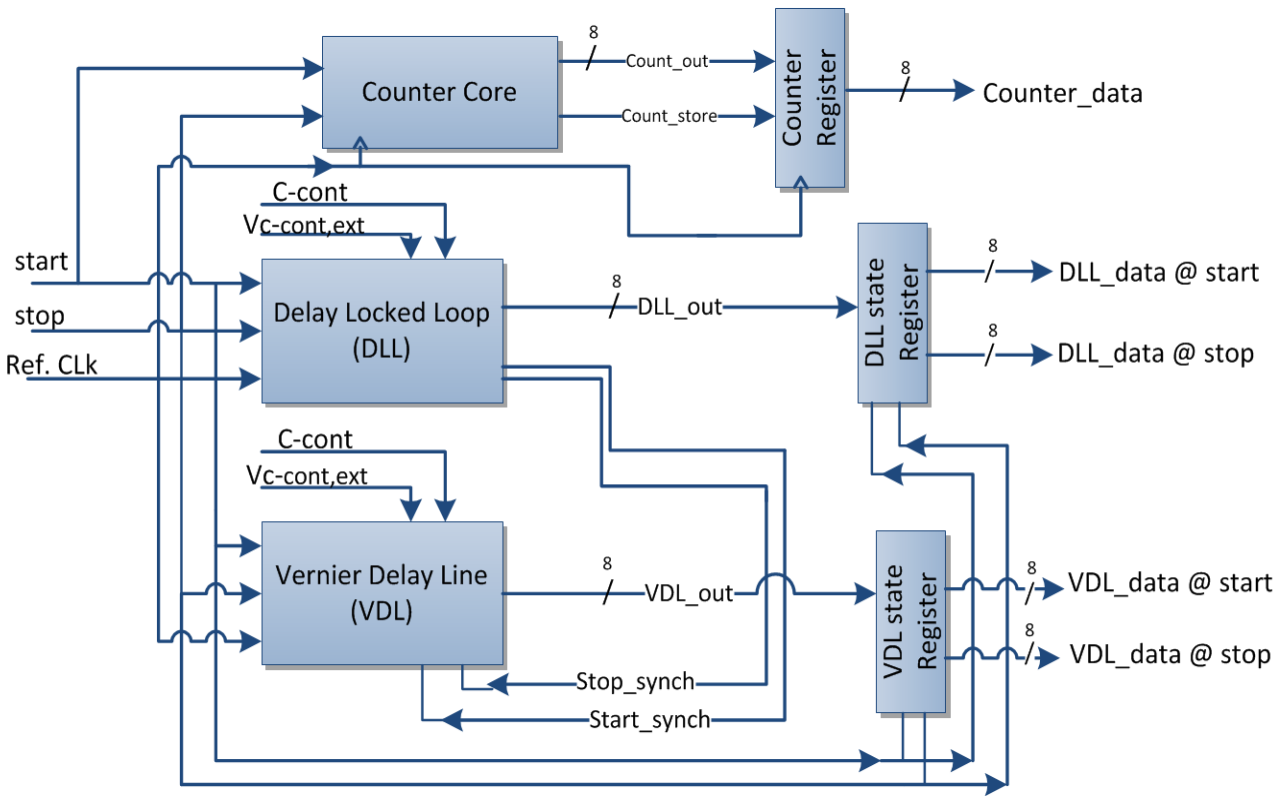
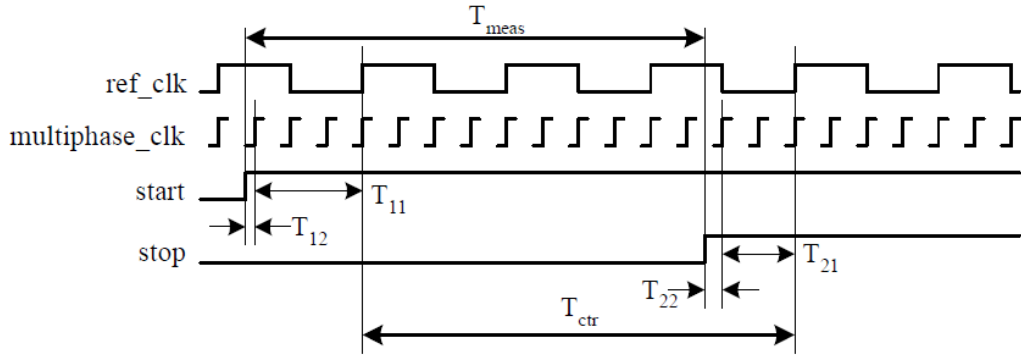


Fig. 4.1: Simplified architecture of the TDC

The time residue in between the start and stop signal and the next rise time of the CLK signal is measured with two fine TDCs in the routine that can be seen in Fig. 4.2. As shown, the time that is measured with the counter is the time difference between the first detected CLK's risetime after the edges of the start and stop signals. This leaves only the time interval between the start and next risetime of CLK signal to be measured with the fine TDCs (same for the stop signal). This helps to reduce the number of elements required for interpolation.



$$T_{\text{meas}} = T_{\text{ctr}} + (T_{11} + T_{12}) - (T_{21} + T_{22})$$

Fig. 4.2: Timing diagram for the 3-stage interpolation TDC

The first fine TDC is implemented with a DLL structure with 8 delay cells that generate multiphase CLK signals (delayed versions of the CLK signal). As a result, the resolution is improved in the fine TDC by a factor of 8. In fact, in the improved version of DLL that was developed in this work, this improvement factor is 16. This has been performed using a novel technique of “half-CLK period interpolation” which will be explained fully in section 4 of this chapter. The DLL measures the time difference between the first multiphase CLK signal after hitting the start signal and the next rise time of the Ref. Signal (same for the stop signal). The remainder time is measured with the 2nd fine TDC which employs a VDL structure with 8 delay stages. This remainder is the time between start (stop) signal and the next risetime of the multiphase CLK signal.

So by dividing the whole time interval into 3 parts, a large area saving can be achieved by keeping both the resolution and dynamic range high. This is due to benefits of the hierarchical structure in which the dynamic range of each stage is equals to the resolution of the next coarser stage. The remaining issue is the interface between the stages i.e. how to manage the connections between the stages to make the whole TDC work as one TDC and generate remainder (the difference between the asynchronous signals and synchronous CLK signals) to be fed to the fine TDCs. The time difference between the multiphase CLK and the reference signal is measured by knowing the state of DLL when the hit signal is reached. This will be illustrated in Section 4.4. The remaining part which is the time difference between the

hit signals and the next multiphase CLK is going to be extracted by the multiphase CLK synchronization technique which will be illustrated in section 4.6.

The operation of all the building blocks of the TDC will be explained in detail together with the synchronization issues and design challenges in the following sections. Section 4.3 concentrates on the operation of the counter and the synchronization issue with the hit signals. Section 4.4 is on the DLL structure that has been used as the first interpolation stage, while Section 4.5 introduces the novel “half CLK period interpolation” idea. Section 4.6 describes the VDL stage designed as the second interpolation stage and its operation. The circuit realization of main building blocks of the TDC such as the voltage-controlled buffer, phase detector and charge pump, will be described in section 4.7.

4.3. Digital Counter as the Coarse TDC

The Synchronous counter is a one of the key building blocks of the proposed TDC which should be enabled with the start signal and sampled by the stop signal in order to find the number of CLK periods in between hit signals. We are using a common synchronous 8-bit counter shown in Fig. 4.3. Figure shows only realization of the first two bits of the counter for the case of simplicity.

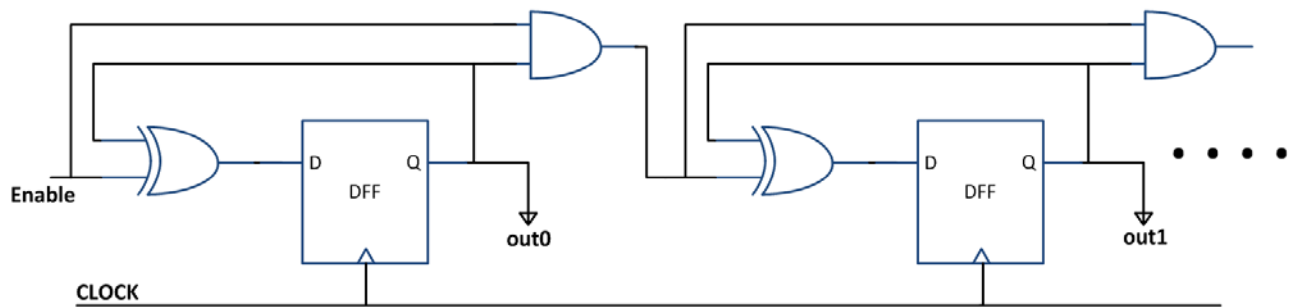


Fig. 4.3: The architecture of the designed synchronous counter

The counter implemented in this work has 8 stages (performing as an 8 bit counter) to provide a wide dynamic range to satisfy application demands. However, implementation of the

synchronous counter might look simple but careful design consideration is needed to keep the counting process accurate. This is because minimum deviation of counter's output from its real value is a CLK period which is at least 2ns in our design. This is huge for a time interval measurement system with resolution less than 100 ps. A serious issue is the sampling process; the counter cannot be sampled asynchronously by the hit signal because the sampling moment could take place right at the transition of the counter when the data is not yet ready. That is why both start and stop signals should first be introduced to the circuit in Figure 4.4. The first DFF is to synchronize the hit signals to the edge of the reference CLK. The second DFF is to give enough time to the first flip-flop to resolve its state. In other words, the unavoidable violation of the timing margin of the first flip-flop increases its delay, so the second flip flop gives the first one enough time (one CLK cycle) to stabilize its state. Another flip flop can be added before these two to expand the hit signals if they are just short pulses. This is done by connecting the hit signal to the DFF's CLK input and VDD to its data input. Fig. 4.4 shows the operation of the synchronization circuit.

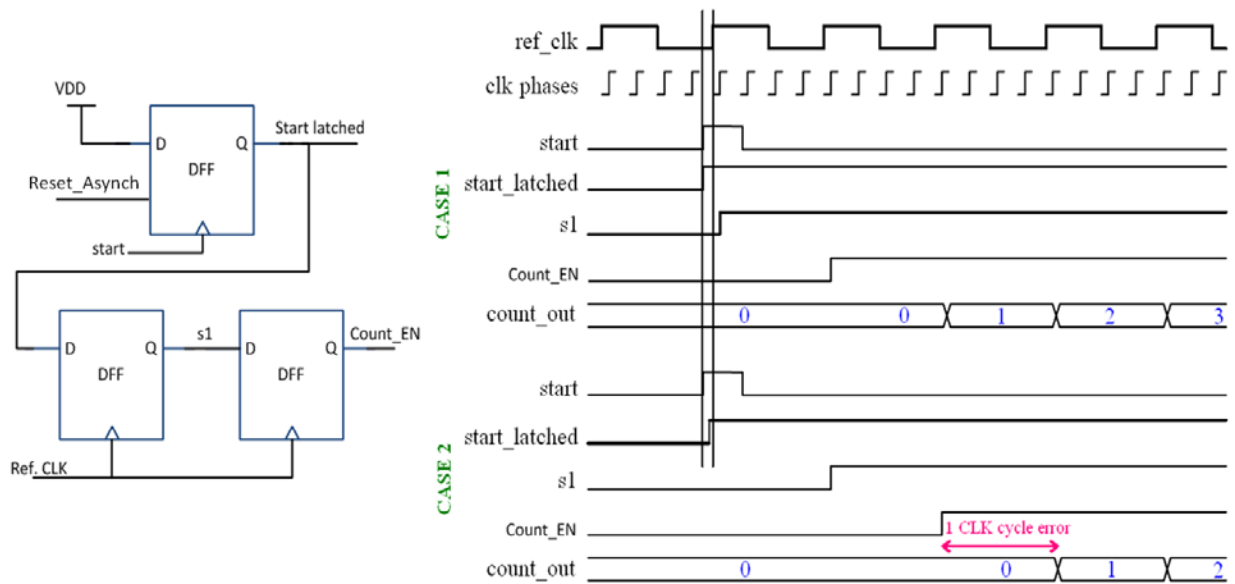


Fig. 4.4: Hit signal synchronization circuit (left) and its operation principle (right)

As can be seen, the data of the counter will be sampled about three CLK cycles after the arrival time of the hit signals. So, in this way, sampling of the counter will be synchronized so that enough time is given to the counter to process its data in such a way that it does not deteriorate its accuracy. Still, there is a problem related to the setup time of DFF which is indicated in Fig. 4.4-right. Two cases have been depicted for the synchronization circuit in Fig. 4.4. In the first case, the hit signal comes right before rising edge of the reference signal but with timing distance more than the setup time of DFF. However, in second case, this difference is less than the setup time of DFF. As can be seen, the synchronization for case number two occurs one CLK cycle later in comparison to the case number one. This is because, in the second case, DFF has not enough time to setup its output. This problem, however, is rare but adds an uncertainty of one CLK cycle to the measurement results which is way more than resolution of TDC. To solve this problem, dual synchronization to both rise and fall times of the reference signal has been fulfilled with the technique that has been introduced in [57].

Fig. 4.5 shows the block diagram of the dual synchronizer. As shown, dual synchronization has been fulfilled by determining if it is safe to synchronize to either rise or fall edge of the CLK signal based on the position of the hit signal. Indeed, the decision is made based on the incidence of the hit signal either in the high or low state of the half-cycle of the reference signal. The state of the next delayed version of the reference signal when the hit signal occurs help to determine that. This solves 1-CLK-Cycle uncertainty problem (Fig. 4.6).

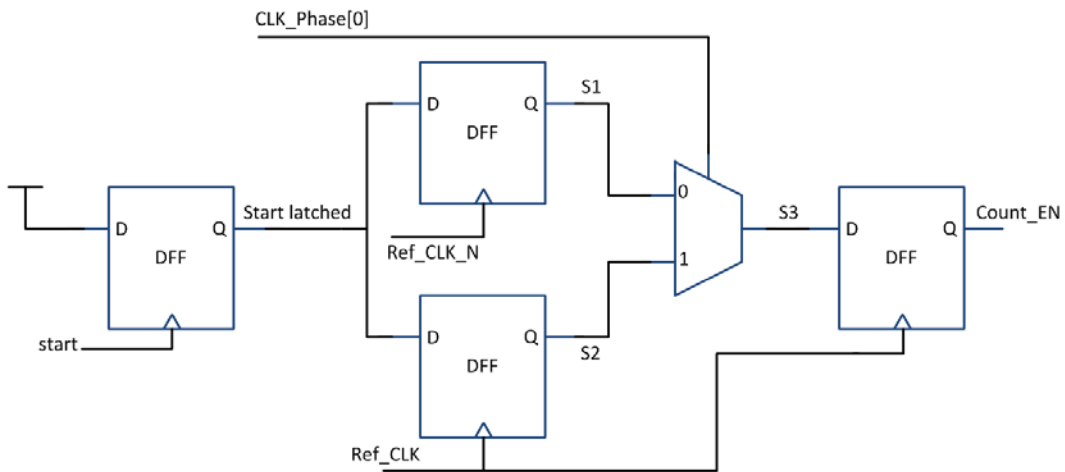


Fig. 4.5: Block diagram of dual synchronization method

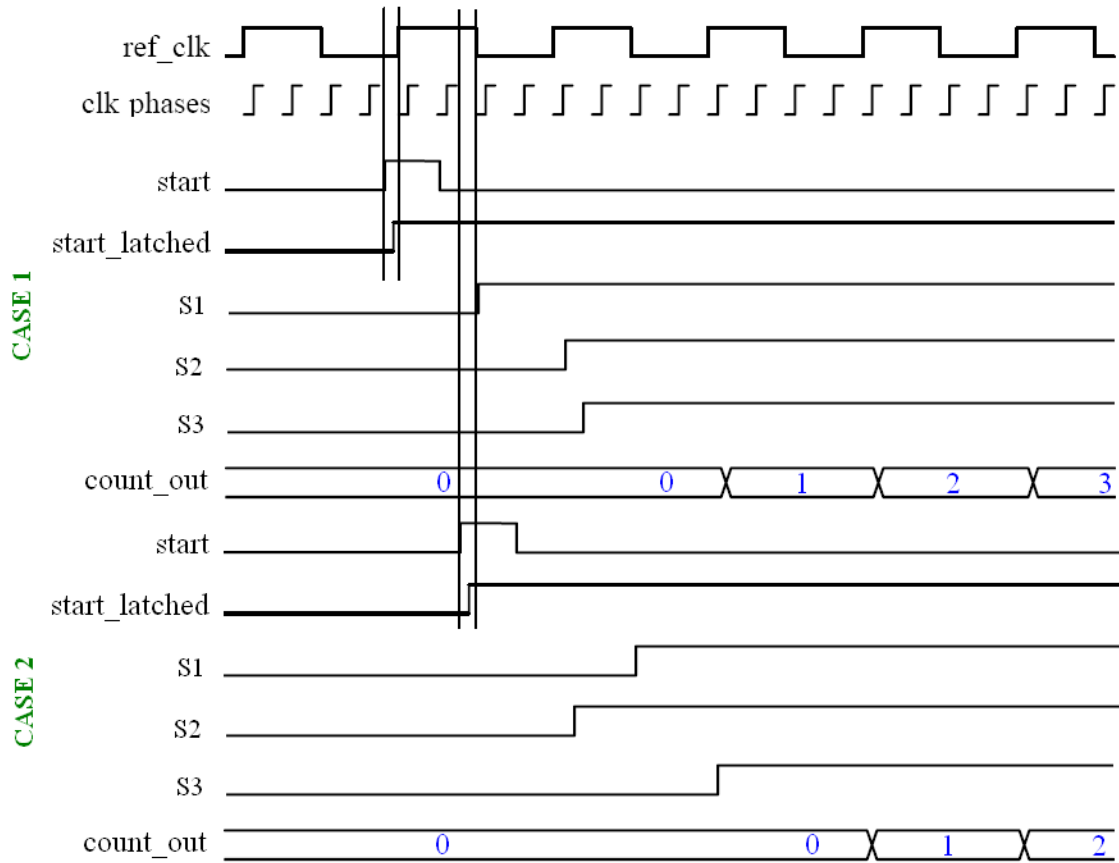


Fig. 4.6: Operation of dual synchronization circuit in two critical cases

4.4. DLL as the First Interpolation Stage

The delay line looped structure provides a flash time measurement with a resolution that can go up to the minimum gate delay and it takes advantage of reference signal locking to regulate the delay of the delay cells when PVT variations occur. Figure 4.7 shows the structure of the DLL implemented in this work. It consisted of a tapped delay line with 8 delay cells (to provide 8 level of interpolation) as the main component. The phase detector detects the phase difference between the input reference signal and its delayed version at the end of the delay chain called “end” signal. The phase difference of two signals is forced to be fixed to an integer multiple of the reference CLK period independent of PVT variation. Based on the phase difference

deviation, the charge pump generates the necessary voltage to feed the buffers to adjust their delay so that the total delay of the line goes back to the desired value. For example, if the total delay of the line is larger than the desired value, that means that, the end signal comes a little bit later than is expected. Based on this difference, the charge pump generates appropriate voltage (V_{c1}) to feed the delay cells (depends on their structure) to increase their speed and compensate for the late end signal. Thus, the phase difference goes back to its looked-for value. The external control of V_{c1} is also applicable by employing a control circuit in cases needed. This is basically consists of just a switch to pass the external voltage (V_{c1_ext}) when V_{c1_cont} signal is high. When V_{c1_cont} is low, the DLL determines the appropriate V_{c1} to be fed to delay line.

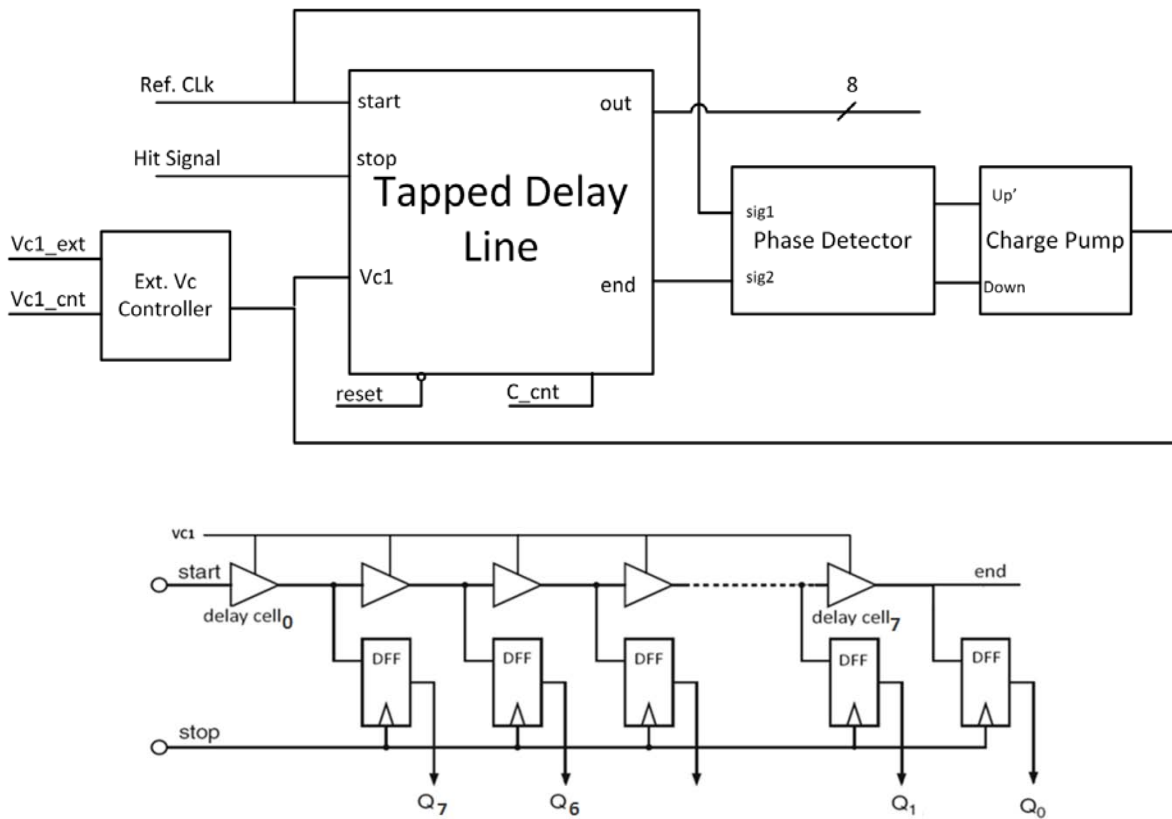


Fig. 4.7: DLL structure (top) and tapped delay line (down)

As can be seen from Fig. 4.2, the time interval that needs to be measured with the first fine TDC or DLL is T_{1A} (time between next phase of the multiphase CLK after the start arrival and the next risetime of the ref. signal) and T_{2A} (time between the next phase of the multiphase CLK after

the stop arrival and the next rise time of the ref. signal). This time is given by the state of the multiphase signal i.e. the states of the flip flops in the tapped delay line. For example, if the hit signal happens at the middle of the ref. cycle, the output of the FFs would be 11110000 in thermometer code. The state of the first interpolation TDC can be collected by the FPGA synchronously which means data can be collected on the edges of the ref. signal. This should be satisfied by storing the data until at least the next risetime of the reference signal. Using DLL structure and this way of reading its output, T_{1A} and T_{2A} time intervals can be measured. It needs to be mentioned that measurement should take place after passing the DLL locking period which is a short time at the beginning of measurements to automatically generate the control voltage (V_{c1}).

4.5. Half CLK Period Interpolation Idea

As shown in Fig. 4.7, the DLL structure locks its two inputs to have a phase difference equal to an integer number of CLK periods. So, the minimum delay that can be achieved by a DLL with a certain number of elements (N) would be T/N (T is the period of the Ref. signal). To increase the resolution of the DLL with a fixed amount of cells and dynamic range, a new idea of “half CLK cycle interpolation” has been developed. In this method, instead of locking the input reference signal and the end signal to have a CLK cycle phase difference, they will have phase difference of half a CLK cycle. This idea was realized by locking the end signal to the inverse of the input reference signal by using a voltage controlled inverter shown in Fig. 4.8. This simple novel idea theoretically reduces the area consumed by the DLL by half. However, it adds the uncertainty of the delay of an inverter to the measurement which is much bigger than the resolution of the TDC. This uncertainty happens because total the delay of the delay line will be locked to a half CLK cycle plus delay of an inverter. So, this offset must be cancelled in a way to keep the accuracy of the TDC high. To do this, a voltage controlled buffer with different sizes was designed in such a way that its delay equals to the delay of the inverter when the control voltage of both are kept the same (coming from output of CP). This adds nonlinearity as the last

delay element of the delay line (cell₇) sees different loading than other cells. This was solved by adding a dummy buffer after delay cell₇ and also after the inverter to increase the linearity, while keeping the symmetry. So based on the idea of the half CLK interpolation the number of elements needed (including delay elements, DFF and output buffers) will be decreased by half in the DLL while keeping the same resolution and dynamic range. This area reduction is a great advantage for in-pixel TDC implementation.

There are two other issues left for consideration to make this idea work. First, using this implementation, initially dynamic range of DLL seems to be half the CLK cycle which is half the resolution of the counter. So a very simple half CLK counter is added to the TDC to compensate for that. Indeed, if the difference between the hit signal and the next positive edge of CLK is more than 1 half CLK cycle, then the output of this counter becomes one, otherwise it becomes zero. We will explain more about half-CLK counter operation and its realization in last section of this chapter.

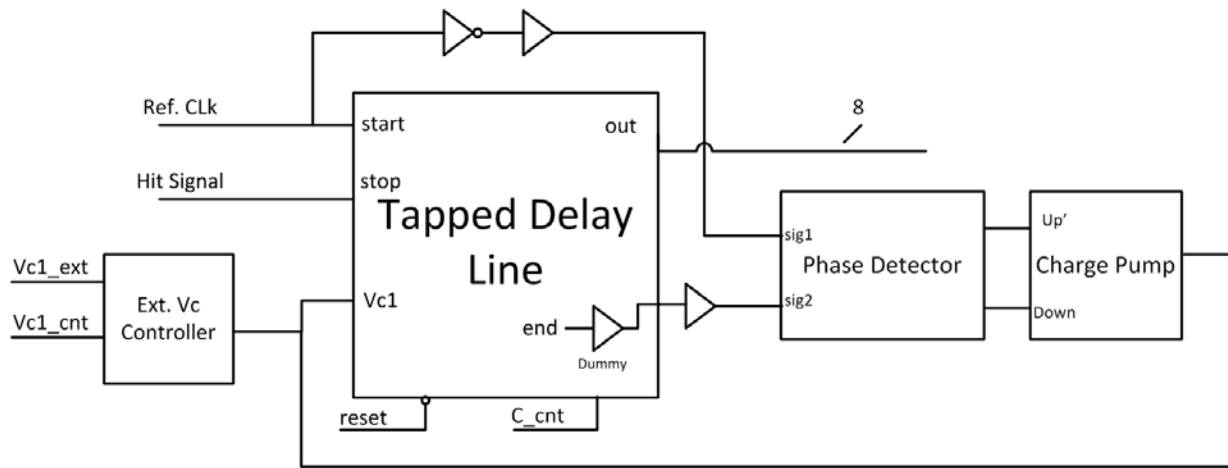


Fig. 4.8: The new proposed DLL structure based on Half CLK cycle interpolation idea

Another problem is that, the output code of the DLL would be different based on the position of hit signal in the reference signal. This means that if the hit signal approaches when the ref. signal is high, then the output will follow the thermometer code. However, if it takes place when ref. signal is low, output of DLL follows the inverse of the thermometer code. So, this needs to be taken into account when reading the data from TDC. A simple multiplexer has been employed as the interface between the output of the DLL and the thermometer-to-binary

encoder to solve this problem. We will describe the operation of this MUX in detail later in this chapter. Fig. 4.9 shows the timing diagram of the proposed TDC including the half CLK interpolation idea and when the half CLK should be added to the measurement result based on the outputs of the counters.

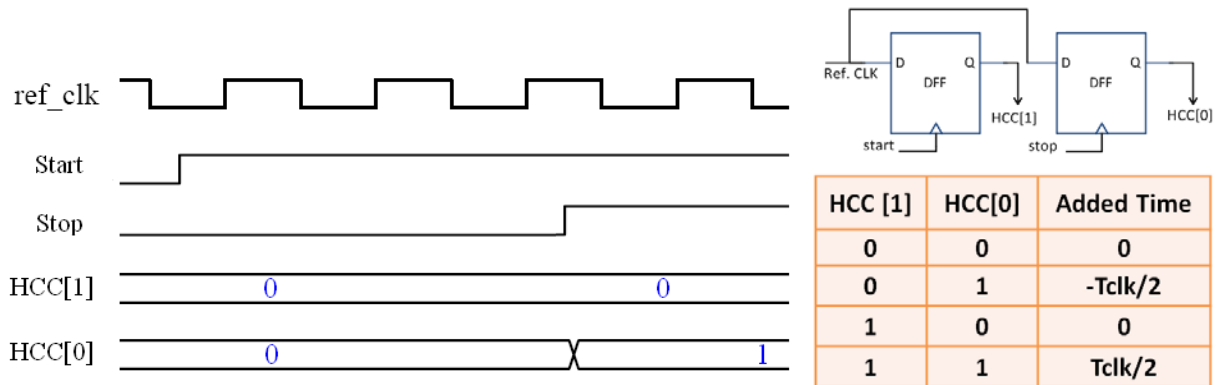


Fig. 4.9: New timing diagram of proposed TDC (left) added time due to half-CLK counter outputs (right)

So based on the idea of half CLK interpolation, the number of elements needed (including delay elements, DFF and output buffers) will be decreased by half in DLL while keeping the same resolution and dynamic range. The nonlinearity was canceled by adding dummy buffers and the dynamic range was kept the same by adding a half CLK counter. The area saving in this method can be significant in DLLs with high levels of interpolation. Therefore, this architecture can be replaced with the conventional DLL structure to save area (while keeping the performance of DLL the same) especially in applications that requires in-pixel TDC implementation which involves very tight area budget in the design.

4.6. VDL as the Second Interpolation Stage

While the counter provides the number of CLK cycles between start and stop, the DLL finds the time between the 1st interpolation signals and the positive edge of ref. signal (T_{1A} and T_{2A}). The time intervals that are left to be measured are T_{2A} and T_{2B} as in Fig. 4.2. These two short time intervals are have to be measured with a high resolution TDC that satisfies the high resolution

demand of ToF PET imaging. For the reasons mentioned at the beginning of this chapter, the VDL technique has been chosen to achieve this high resolution while keeping the precision in an acceptable range by regulating the delay cells of the VDL during PVT variations. Fig. 4.10 shows the architecture of VDL structure in detail. As shown in the figure, the start and stop signals are delayed with different delay amounts and the output of the VDL is sampled when these two signals catch each other. The designed VDL structure has 8 delay cells in each of the start and stop delay lines plus a dummy cell at the end of the each line. The last cell is just for matching the output load of the 8th cell to other cells (to avoid any nonlinearity) and does not participate in the interpolation. As illustrated in the previous chapter, the resolution of VDL is the difference between the delay amounts of the delay cells in the start and stop chain. To generate this difference, different delay cells with slightly different delay amounts should be designed as the first step. One way to do this is to build up two different DLLs with input CLK frequency of one slightly higher than the other. This requires two DLLs which occupy a lot of area. Also, generating two high frequency CLK signals with a slight difference in their frequencies is complicated and usually not applicable with FPGAs.

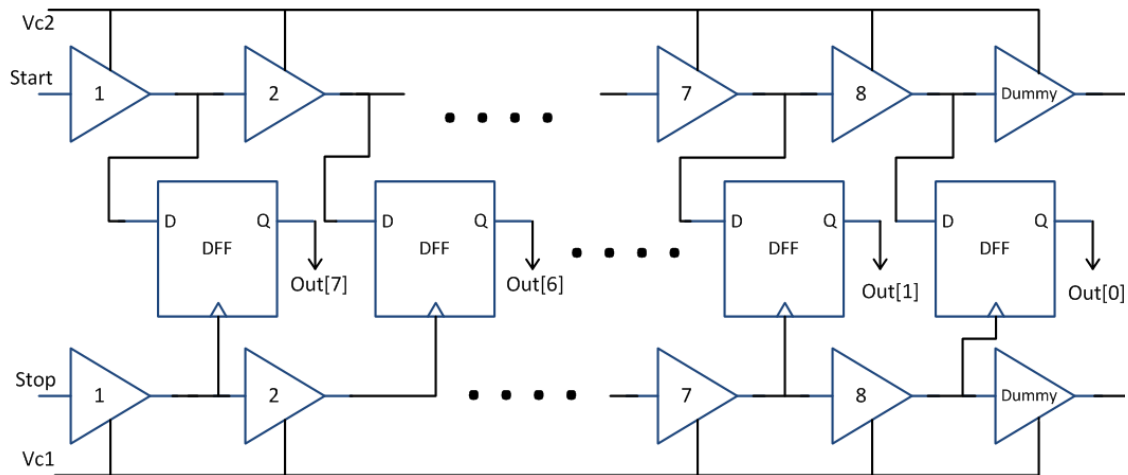


Fig. 4.10: VDL structure used as the 2nd interpolation stage of the proposed TDC

In this work, we used another technique to generate this slight difference in delay amounts of the delay cell. Using this technique, requirement for two DLLs and two CLK signals was eliminated. Only one reference signal has been used. The delay cells in the stop chain are

controlled with the same voltage that has been generated in the first interpolation TDC (V_{c1} - this means the delay of each is $T_{CLK}/8$). The control voltage of the delay cells in the start chain, however, is generated using an extra DLL with the same delay cell structure compared to the DLL used for the stop signal, but with one delay element less (7 delay cells in the chain). So, the delay amount for each delay cell in stop chain is equal to $T_{CLK}/2/7$. Therefore, theoretically, the resolution of the 2nd fine TDC equals to $(T_{CLK}/2/7 - T_{CLK}/2/8)$. For the CLK frequency of 200 MHz which is achievable with FPGAs, the resolution of VDL would theoretically be 44.5ps with a dynamic range of $44.5ps \times 8 = 357ps$, which is a bit more than the resolution of the 1st interpolation stage. So, this should be kept in mind in the data processing after reading out the measurement results.

Now let us look at the input signals which are going to feed the VDL (start and stop). As can be seen from Fig. 4.2, the start signal of the VDL has to be asynchronous hit signals (start or stop input of the whole chip), while the stop input of VDL must be fed with the next delayed version of the reference signal. Fig. 4.11(left) shows how to synchronize an asynchronous hit signals with a synchronous multiphase CLK signal using a two flip-flop synchronizer. The first flip-flop is used to stretch the hit signal with short pulse duration. The second and third DFFs, however, are used to sample the input hit signal synchronously, as illustrated in the figure. The two-flip-flops structure eliminates the nonlinearity due to the waiting time needed for the first DFF. As we mentioned, this waiting time is needed for the first DFF to resolve its state. Indeed, the 2nd flip flop gives the first one a short time to recover from its probable meta-stable state. Fig. 4.11(left) shows how two start and stop inputs of VDL can be achieved by generating a residue signal. This figure depicts the implementation of the time residue finder where the time residue is the time between the hit signal and next interpolated signals from the DLL. But as hit signal is an asynchronous signal that can happen anytime before any of the multiphase CLK signals, the first next interpolation signal can be any of the interpolated signals in the DLL (delayed versions of Ref. signal). So, parallel implementation of the time residue finder has been fulfilled to take into account the probability of being the next synchronous signal for all of the interpolation signals. This has happened by implementing in parallel, 8 two flip-flop synchronizer and then combining them with an 8-input OR gate as shown in Fig. 4.11(right).

This structure helps hit the signal to be synchronized effectively with a GHz level signal by just employing a MHz level CLK signal which eliminates the need for high-frequency reference signal generation and eases the complexity of the design by avoiding RF IC design considerations.

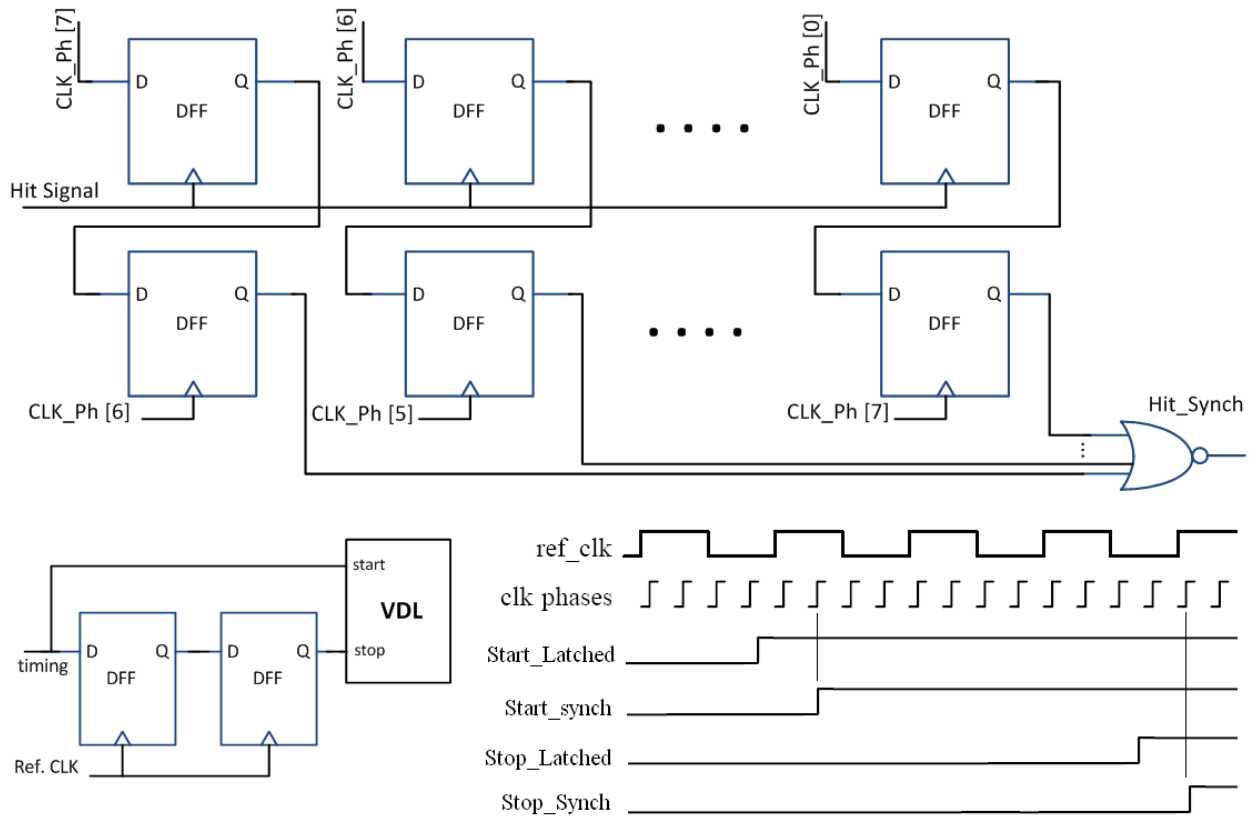


Fig. 4.11: Two flip-flop synchronizer structure (Bottom Left) and its application for generating residue for second fine TDC (Top) and its timing diagram (Bottom Right)

One drawback of the two flip-flop synchronizer architecture is the time offset of one CLK cycle (for settling the first DFF output). This, however, gives the first DFF enough time to resolve its state, increases the time residual range and therefore demands more dynamic range for the VDL. To decrease this time offset, the input port of the 2nd DFF of each of the synchronizers is fed with the next phase of the CLK signal. Based on this idea, a smaller time offset is added to the measurement interval (only a fraction of the CLK cycle is added) while the first flip-flops will have $T_{CLK}/8$ to recover their state which seems to be enough for the fast flip-flop implemented and the low CLK frequency used. We test the metastability of the first DFF by running a sweep on its input signal and it was confirmed that $T_{CLK}/8$ is enough for the first DFF to generate its

output. Another problem is the propagation delay of the OR gate which can delay the synchronous signal. This delay can be more than the dynamic range of the VDL which adds a significant uncertainty to the measured results.

To compensate for these sources of error, some delay compensation techniques can be employed. For compensating for the $T_{CLK}/8$ delay offset in the synchronizer, one delay element in the first DLL with the same structure and same control voltage (V_{c1}) has been used in the propagation path of the asynchronous hit signals;. This adds an intentional delay of exactly $T_{CLK}/8$ to the hit signals. Also, employing dummy logic gates having the same amount of delay with the gates in between the propagation path of the asynchronous and synchronous signals is a beneficial technique which compensates for their delay offset. These techniques help to increase the precision of the fine TDCs and leads to improving linearity in the whole TDC.

The whole architecture of the proposed TDC based on these considerations and improvements is shown in Fig. 4.12. Again, it consists of three main parts: a Counter as the coarse TDC, a DLL as first interpolation stage and a VDL as the second interpolation stage. Synchronizers have been used both for the counter and the interpolators. A half CLK counter together with some other changes is used for the half-CLK period interpolation idea which was illustrated in Section 5 of this Chapter. A second DLL structure was used to generate the delay amount needed for the delay cell in the stop chain in VDL, as was previously described. Two delay elements are used for generating the appropriate delay for offset compensation in the VDL. A Schmitt-trigger circuit has been used for some critical slow input signals such as start, stop, CLK and reset signals to generate signals with fast transition. In addition, output buffers have been employed to buffer the output codes and send them to the large load of the pads of the chip. The thermometer-to-binary encoder is another part of the system which generates binary-encoded outputs from the thermometer coded inputs to decrease the length of output word due to limitation of the number of the pads. Detailed information on operation principle and circuit realization of the Schmitt-Trigger, buffers, thermometer-to-binary encoder and other circuit blocks of the proposed TDC is given in the following section.

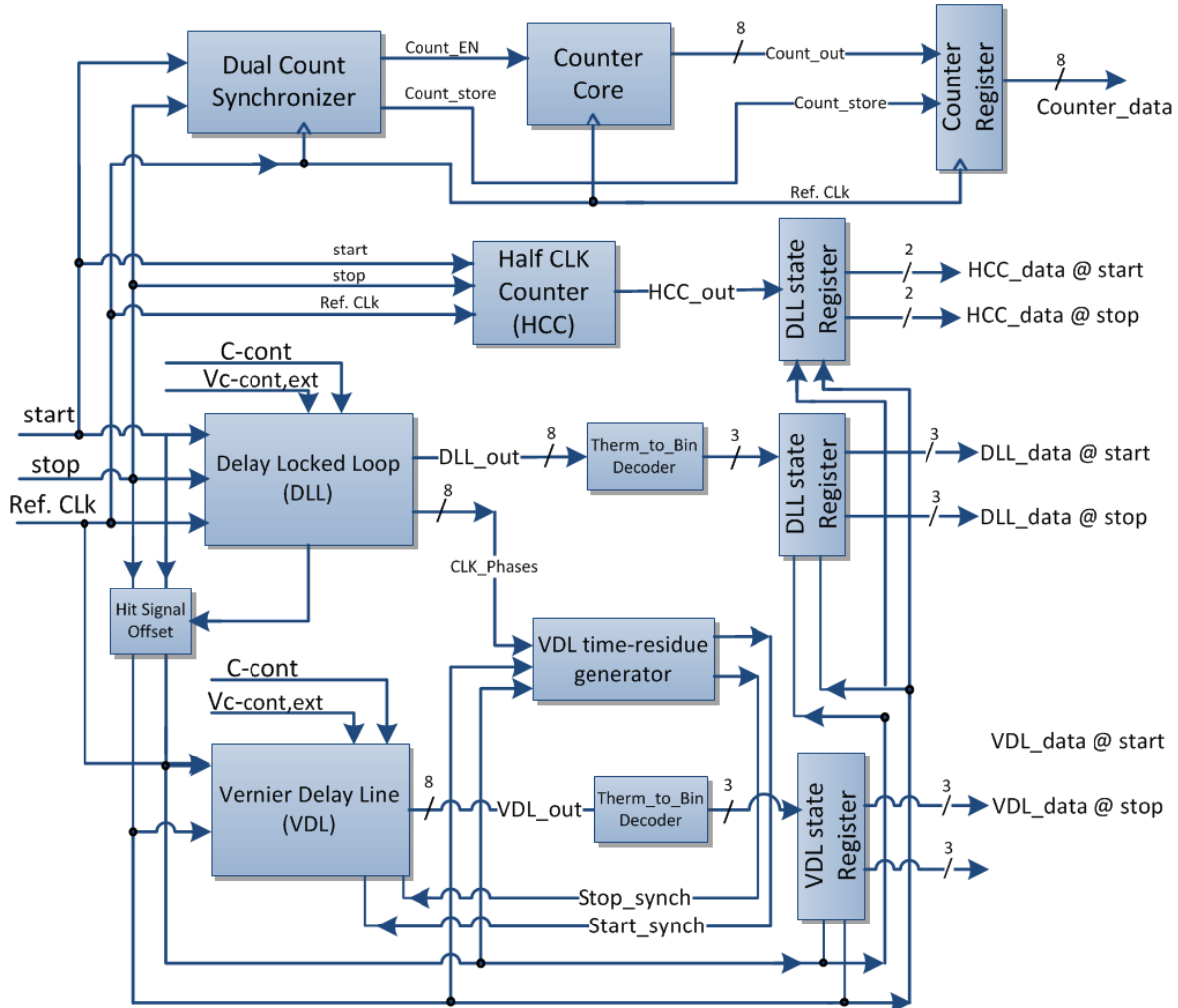


Fig. 4.12: Final architecture of the proposed TDC after applying stated changes

4.7. Realization of the Circuit Blocks

In this section, detailed information on realization of some of the circuit blocks of the proposed TDC will be given. These circuits include voltage-controlled delay element, phase detector and charge pump which are some of the main building blocks of DLL and VDL as well as Thermometer-to-Binary encoder, half-CLK counter, Schmitt-Trigger and output buffers. Explanation on why they have been employed and details on the circuit design will be given.

4.7.1. Voltage-Controlled Delay Cell

The delay elements for our TDC should provide an appropriate range of delay in the order of 300 to 500ps to be used for the designed architecture with 8 interpolation levels in half CLK period where the CLK frequency is 200MHz. In order to be used for different CLK frequencies such as 100MHz and 50 MHz, the delay cells have been designed in a way to provide a wide range of delays. There are many delay cell structures in the literature which can be used in this work. The simplest design consists of two in-series inverters offering a fixed amount of delay. In order for the delay elements to be suitable for the DLL structure, they must provide a range of delay rather than a fixed delay. In addition, this delay should be a function of the voltage of a control node inside the cell. Ref. [58] did a good comparison and analysis on different delay elements with variable and fixed delay amounts.

Fig. 4.13 shows some popular voltage-controlled delay element structures. The N-voltage (p-voltage) controlled delay cell consists of a cascade inverter pair with additional pull-down (pull-up) transistor with control voltage V_N (V_P) to control the current passing through the inverter. In this way, their propagation delay can be controlled (Fig. 4.13-a,b). the PN-voltage controlled delay cell (Fig. 4.13-c) employs control transistors in both pull-up and pull-down networks which add one extra level of control to the delay element at the expense of the reduction in the output voltage range. In all of these structures, the signal integrity is poorer than a simple two-inverter delay cell due to the additional resistance and cap introduced by the control transistors. Also, they consume more power. But, they offer wider range of propagation delay which is essential as a key building block of the DLL and VDL circuits in our design.

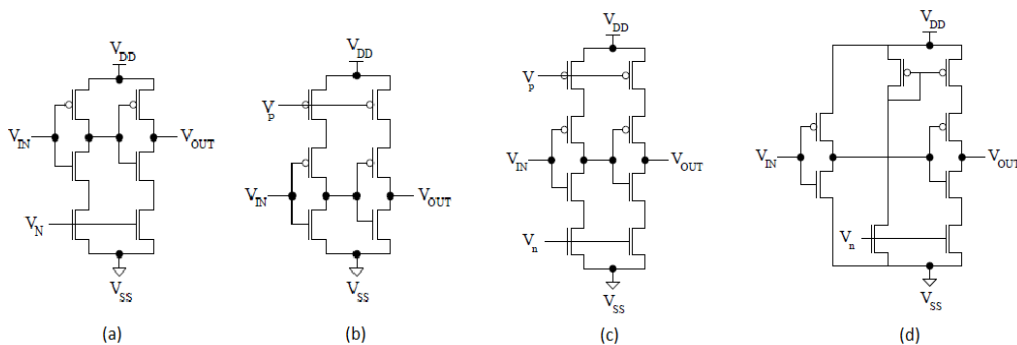


Fig. 4.13: Some popular voltage-controlled delay cells: (a) n-voltage controlled (b) p-voltage controlled (c) np voltage controlled delay cells and (d) current starved cascades inverters [58]

The current-starved delay cell is similar to the cascade inverters, but it employs two extra transistors to control the maximum current passing through the second inverter. The delay range that can be achieved with this delay element is considerably larger than with other structures due to the control transistor that limits the amount of output injected current [58]. However the signal integrity is poor compared to cascaded inverters.

Other delay element structures in the literature have been implemented by applying some changes on these basic structures to increase their delay range or frequency bandwidth. Ref. [59] employed current-starved delay cells in a different structure to improve their noise and frequency response. Ref. [60] proposed a bit more complex delay cell by employing N voltage-controlled current paths and using a transistor network to digitally control the delay range. Delay cell for this project was chosen to have the appropriate delay range with a low number of transistors to reduce the area consumption and with acceptable signal integrity to generate sharp undistorted delayed signals. So, the basic current-starved structure has been used in this work with two additional transistors to increase the delay range of the cell. The delay range expansion was done by adding a capacitance to the intermediate node of the delay cell to shift the delay range (Fig. 4.14).

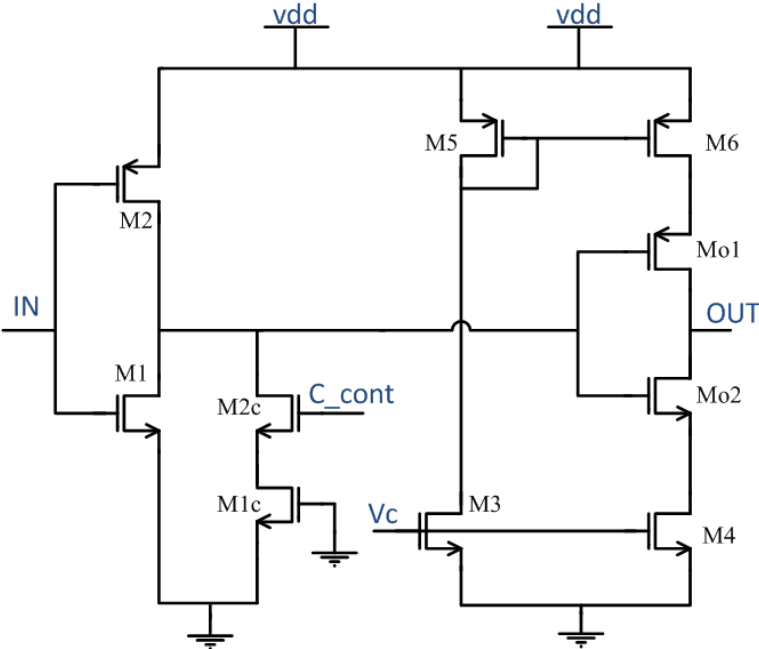


Fig. 4.14: Structure of delay cells of the proposed TDC- current starved structure with additional capacitive load control

This delay cell takes advantage of two degrees of controllability. Voltage V_C controls the current passing through the inverters and therefore their delay is similar to the basic current-starved structure. In order for this delay cell to be used for different CLK frequencies, the range of the propagation delay of the cells should be expanded. For this, two MOS transistors have been used to adjust the capacitance in the intermediate node. When the V_{C-Cap} is high, the capacitive load is maximum. This means that the delay range of delay cell is shifted up (maximum delay of 1.5ns is applicable to be used for CLK frequencies as low as 50 MHz). However, when V_{C-Cap} is low, the M1c and M2c transistors' capacitance C_{DS} are in series with each other, considerably reducing the load seen in intermediate node and making it possible to use the delay cell in high frequencies up to 500 MHz. The proposed delay provides a wide range of propagation delay (between 250ps to 1.5 ns) in simulations which makes its usage possible for a frequency range between 50 to 500 MHz in the proposed TDC.

4.7.2. Phase Detector and Charge Pump

Process, voltage and temperature variations lead to deviation of the performance of delay cells from their ideal performance. These deviations which occur in propagation delay, risetime, falltime and signal integrity cause nonlinearity in the TDC operation. To compensate for these deviations, the DLL structure has been employed to regulate delay cells during PVT variations. Two significant building blocks of the DLL are the phase detector and the charge pump. The phase detector needed for our project has to provide a low phase offset and fast operation with small number of transistors. For this, a dynamic phase detector proposed in [61] was employed. This phase detector provides high speed and low phase offset at high frequencies due to symmetry in the designed circuit (Fig. 4.15-a). As can be seen from the circuit, only 12 switches have been used to generate \overline{up} and \overline{down} signals which are the input signals of the charge pump. Operation of phase detector is shown in Fig. 4.15-b. This figure illustrates how the width of the up and down pulses is proportional to the phase difference between two input signals. When the DCLK signal is ahead of the reference, the “up” signal becomes high. On the other hand, the “down” signal becomes high when reference signal is ahead of DCLK.

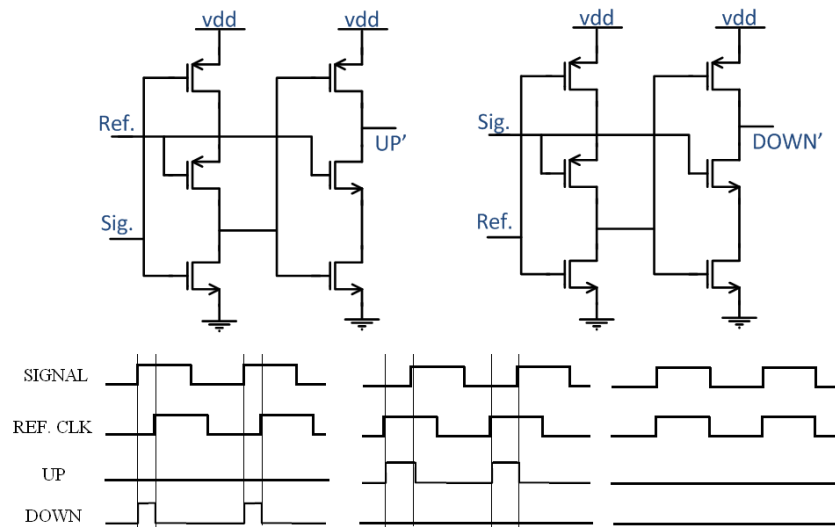


Fig. 4.15: Schematic diagram (top) and operation (bottom) of phase detector used in this work

These up and down signals are then used to feed the charge pump in the DLLs so that an appropriate output control voltage can be generated to regulate delay cells. Figure 4.16 shows the schematic diagram and the operation of the charge pump used in this work.

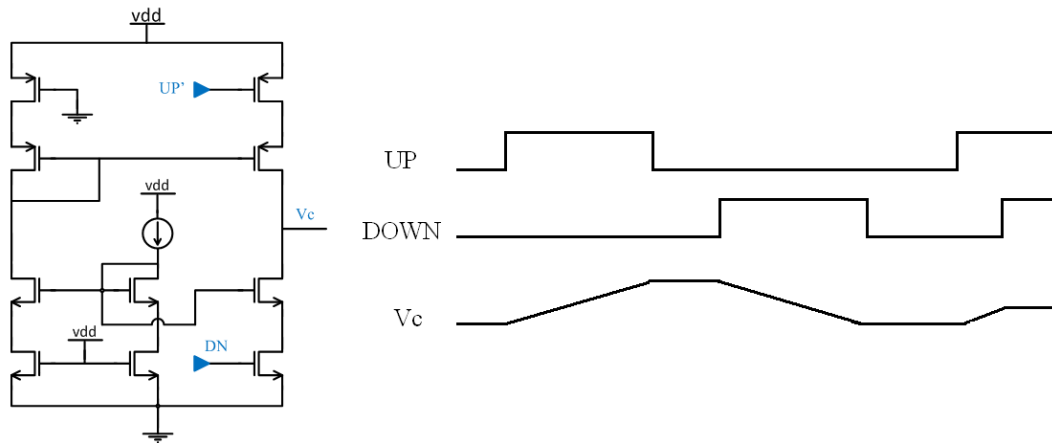


Fig. 4.16: Designed charge pump schematic (left) and operation (right)

As can be seen from this figure, when *up* signal is high, charge is injected to the output capacitor from VDD, while the capacitor is discharged when the *down* signal is high. Therefore, the voltage on the output capacitor is determined by the width of the up and down pulses. In order to make a fixed current independent to the input signals, the mirror architecture has been used. The current that is generated with this current source is then mirrored to the output chain. The amount of this current determines the amount of charge injected to the output capacitor and thus the duration of locking period.

4.7.3. Thermometer-to-Binary Encoder

As can be seen from the TDC architecture in Fig. 4.13, there are 8 outputs for both the DLL and VDL, which together with the output signals coming from the counter require a large number of pads to read them from the chip. As the interpolation occurs in half the CLK cycle, the outputs of the DLL and VDL follow the thermometer or inverse-thermometer code, as illustrated in Section 5 of this chapter. So, in order to reduce the number of pads as well as output buffers, the output data has been encoded to binary using thermometer-to-binary encoder (TBE). Fig. 4.18(right) shows the schematic diagram of this circuit. To reduce the number of TBE used for data encoding and to reduce circuitry, the outputs of the DLL are coded both in thermometer and inverse-thermometer code based on the state of reference CLK signal. So, a mux circuit was designed. It passes the DLL output when reference signal is high, and the inverse of DLL output when reference is low (Fig. 4.17).

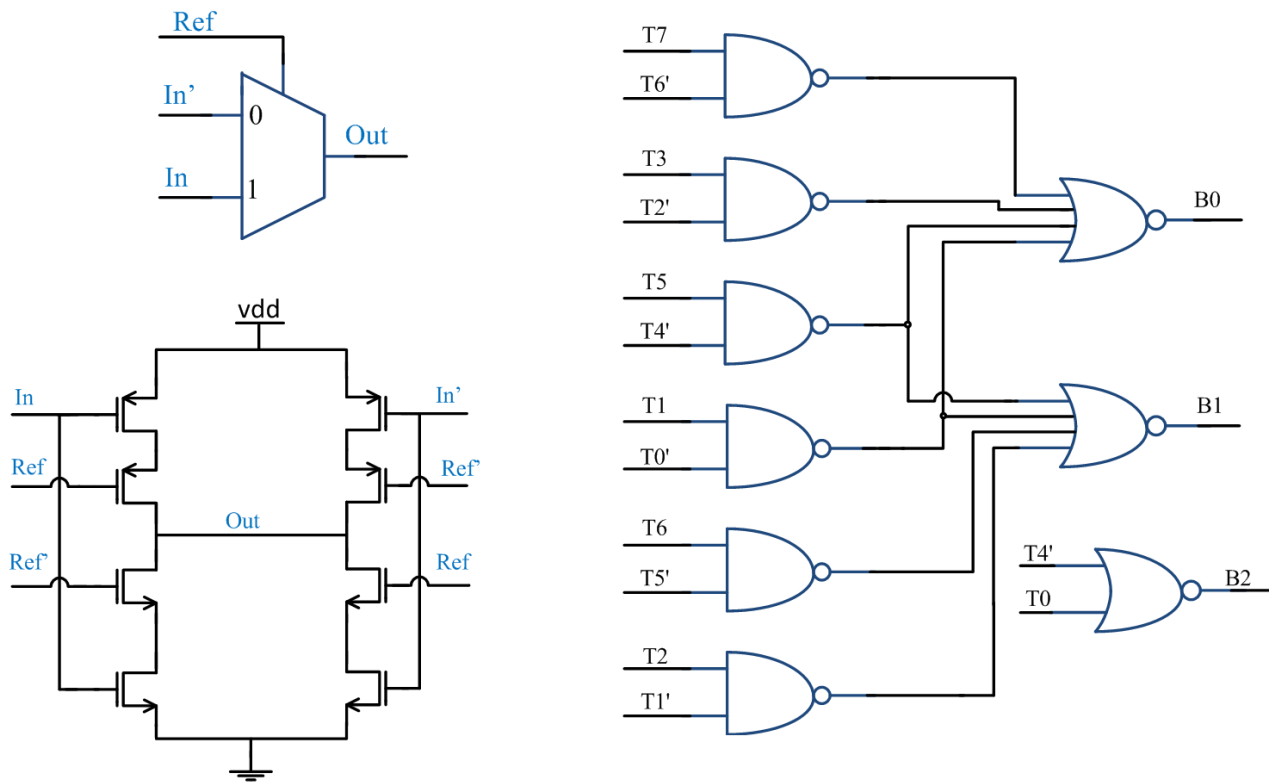


Fig. 4.17: Thermometer-to-binary encoder gate level diagram (right) and MUX to make it compatible with half-CLK cycle interpolation idea (left)

Using this MUX, the thermometer-to-binary encoding becomes compatible with the half-CLK cycle interpolation idea. Using MUX adds delay, but as it occurs after time measurement and as this delay is smaller than the reference CLK, it does not change the performance of TDC.

4.7.4. Half-CLK-Period Counter

As we mentioned in section 5, to realize the half CLK cycle interpolation idea, a half-CLK counter should be implemented to count for the residual half CLK cycle. Indeed the total number of half_CLKs between the start and stop edges. Counting the number of half_CLKs can be achieved by increasing the resolution of the main counter. This can be fulfilled either by doubling the input reference CLK signal's frequency, or by making the counter sensitive to both the rise and fall edges of the CLK signal. But these methods work at the expense of doubling the size of counter for the same dynamic range. This is in contradictory with our purpose to minimize the size of the system. As the half-CLK counter is just performing after the hit signal arrival time, and before enabling the main counter. It should just indicate if this time interval is bigger than a half-CLK cycle or not, so the design can be done in a much simpler way. The design has been done based on the following idea which previously has been shown in Fig. 4.9. If the reference CLK is high when the hit signal arrives, there would be a half CLK cycle before next positive edge of CLK signal. On the other hand, there will not be any if reference signal is low when hit signal arrives. So the half-CLK counter would be just a DFF with reference signal connected to its data input and the hit signal connected to its trigger input. This circuit has been used for both start and stop signal to detect if there is a half-CLK cycle between them and next positive edge of the CLK signal. Each of half-CLK counters provide a 1 bit output from which time needs to be added to the time measured with counter and interpolation stages can be found. This added time depends on the state of two half-CLK counters' outputs and can be seen in table in Fig. 4.9.

4.7.5. Schmitt Trigger and Output Buffer

In order to appropriately interpolate the reference CLK signal with the delay elements, each of the delay elements should have a sharp signal going from 0 to VDD or from VDD to 0 as fast as

possible. This helps DFFs in the tapped delay line or VDL to appropriately decide if their input is 0 or 1 when the trigger signal comes. In order to have sharp edges in the output of the delay elements, the input signals to the TDC such as start, stop and reference signal should be sharp enough. But, these input signals are coming from signal generators or the FPGA which usually have high rise and fall times. One circuit that is commonly used to sharpen the signals coming from external sources is the Schmitt-Trigger. Unlike conventional inverter or buffer, a Schmitt-Trigger has different transfer curves in high-to-low and low-to-high transitions. This is called the Hysteresis effect. This helps to have different upper and lower switching voltages. By adjusting these two voltages close to VDD and ground, we can provide sharpened edges from the input signals with high rise and fall times. Fig. 4.18 shows the transfer characteristics of a common Schmitt-Trigger and the Schmitt-Trigger transistor-level schematics that have been used in this work [62].

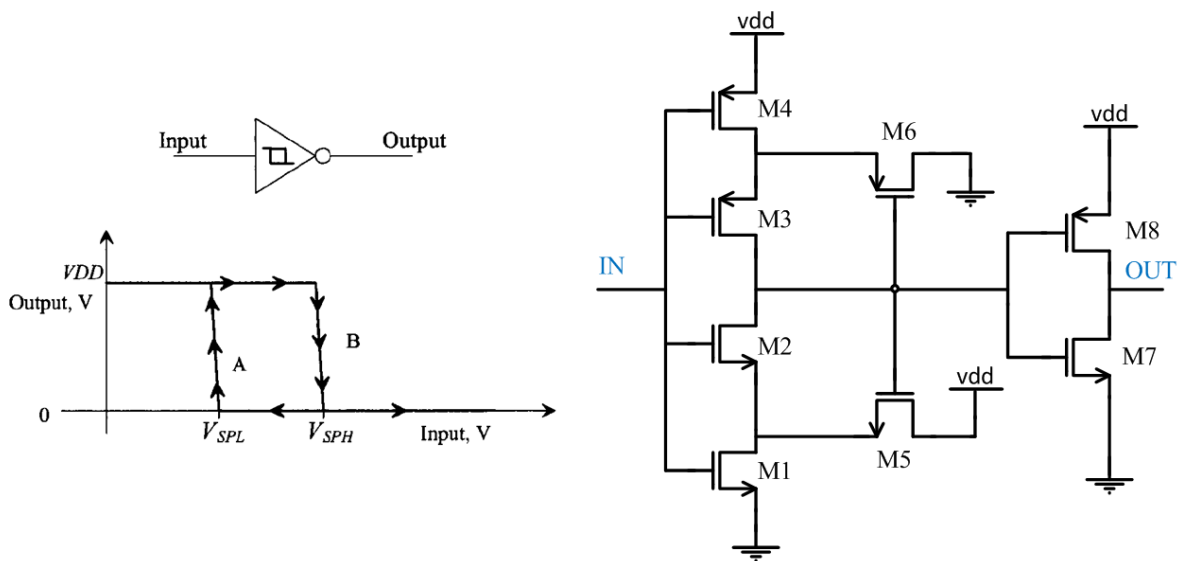


Fig. 4.18: Schmitt-trigger transfer characteristics (left) and its schematic (right) [62]

As can be seen from the schematic of Schmitt-Trigger, if the output is low, then M6 is on and M3 is off. This means only p-channel portion is of concern. On the other hand, if the output is high, M3 is on and M6 is off, which means we are concerned with the n-channel portion of the circuit. This helps to have different upper and lower switching point voltages. Detailed explanation on the circuit operation of Schmitt-Trigger can be found in [62].

Another important circuit block in our TDC which indeed is common to most of the analog/mixed signal integrated circuits (IC) is the output buffer. Usually the IC is connected with the external devices (which could be the FPGA) with the pads. Pads together with the input load of the next circuit connected to IC impose a large capacitive load. This load implies using a buffer with high driving capability after the designed circuit and before the pads of the IC. A simple buffer with very large transistor sizes is usable for this purpose. But, as in-pixel design should be always concerned with area limitation, a different solution has been used in our work. The circuit that has been used for driving the high load of the pads is called tapered buffer and is shown in Fig. 4.19.

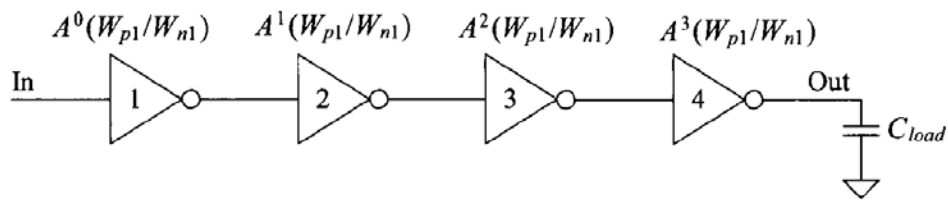


Fig. 4.19: Tapered buffer structure [62]

This buffer not only provides high output drive capability but also occupies less area compared to the conventional two-inverter buffers. It only requires appropriate calculating of the sizing for each of the stages based on the number of them. More information on this calculation can be found in [62].

An overall review on the designed TDC in this chapter was presented in this chapter together with the detailed information on the building blocks of the design. In the next chapter we will present simulation results of each of these building blocks and the entire TDC.

Chapter 5

TDC SIMULATION RESULTS

The main concentration of this chapter is the simulation results from the designed TDC. Each stage has been tested separately to confirm their performance. First, the performance of the counter as the coarse TDC will be presented. After that, simulation results from DLL and VDL as the interpolation stages or fine TDCs will be provided. During this, we will examine behavior of some of the key building blocks of the TDC such as delay element, phase detector and charge pump. After test results from each of the stages of the TDC, simulation results for the entire TDC which includes assessment on its accuracy, nonlinearity and resolution will be offered.

5.1. Counter as the Coarse TDC

The first stage of the TDC which measures the number of CLK cycles between the start and stop signal is the counter. The counter architecture was discussed in the previous chapter. It consists of 8 stages which provide a wide measurement range between 512ns to 2.56 μ s for CLK frequencies from 100MHz to 500MHz. The whole schematic of the counter can be seen in Fig. 5.1. The simulation result for that is shown in Fig. 5.2. The counter was tested for CLK frequencies of 100MHz, 200MHz and 500MHz and the simulation results were satisfactory for all these input signals.

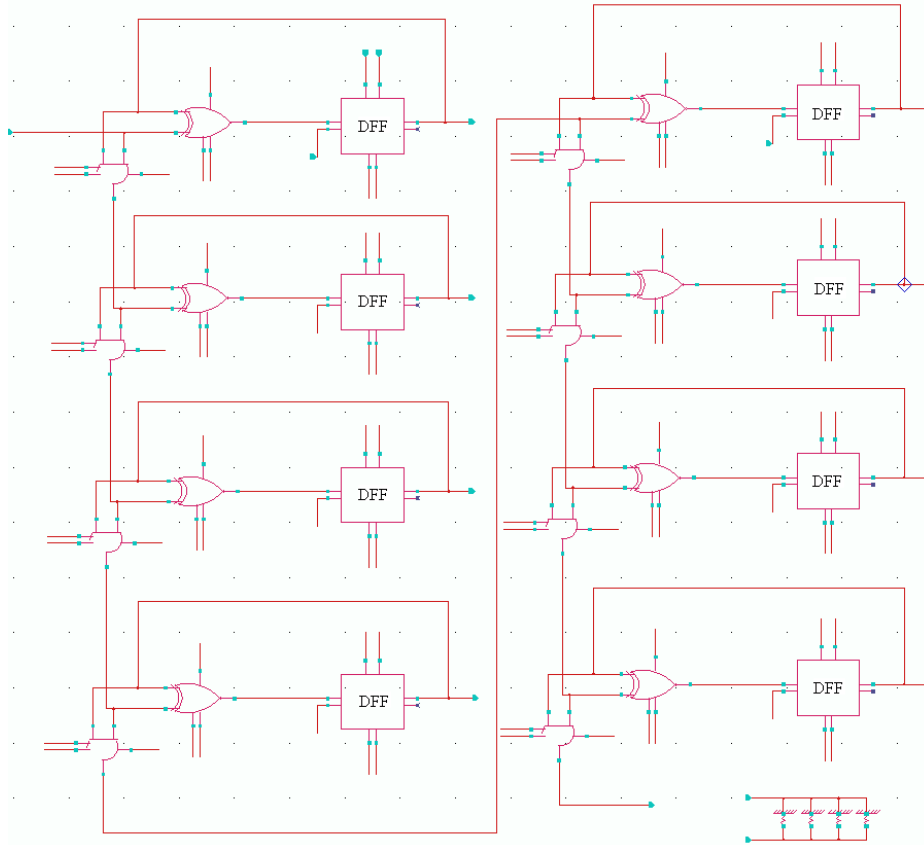


Fig. 5.1: The schematic of the designed synchronous counter

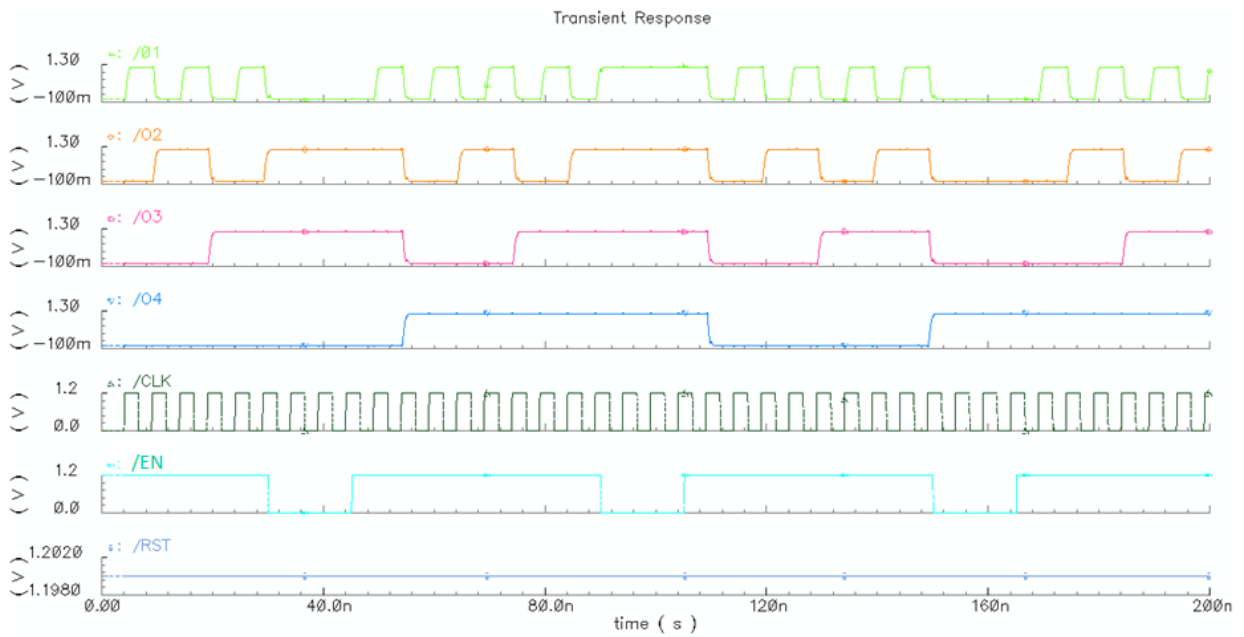


Fig. 5.2 Simulation results for 200MHZ CLK frequency

As can be seen from the Fig. 5.2, counting process is working properly. Now as we know the hit signals are asynchronous with the ref. CLK signal. Because of that, as we explained in Chapter 4, a CLK cycle might get skipped if the incident hit signal is very close to the rising edge of the ref. signal. Fig. 5.3 shows this situation and the way a CLK cycle may be skipped. This surely degrades the accuracy of TDC.

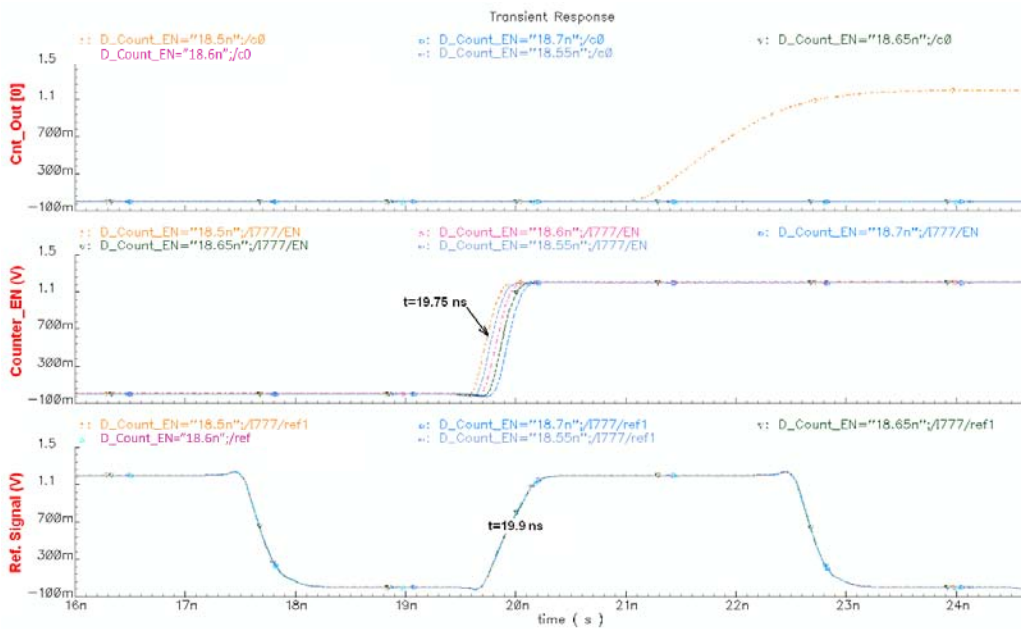


Fig. 5.3: Counter 1-cycle error due to lack of synchronization

As can be seen in Fig. 5.3, when the enable signal of the counter gets very close to the edge of reference CLK (19.9ns), the output of the counter does not increase. As shown in the figure, for the edges of the enable signal more than 19.75, the out of the bit 0 of the counter becomes 0 instead of 1. This means if the hit signal edge is closer than 150ps to the reference CLK, then one CLK cycle will be skipped which implies one CLK period error in the measurement results. So, as we mentioned in chapter 4, a dual synchronization circuit has been employed to prevent this error from happening. Fig. 5.4 shows the accuracy performance of the counter before and after using dual synchronization method. We used a periodic enable signal with frequency of 30MHz to enable the counter and a stop signal which was synchronized to the next edge of the enable signal. So, the input time of the counter would be 33.33ns and the counter should have the ideal value of 6. We run the test for 50 times for different initial delay of enable signal and the results are shown in Fig. 5.4.

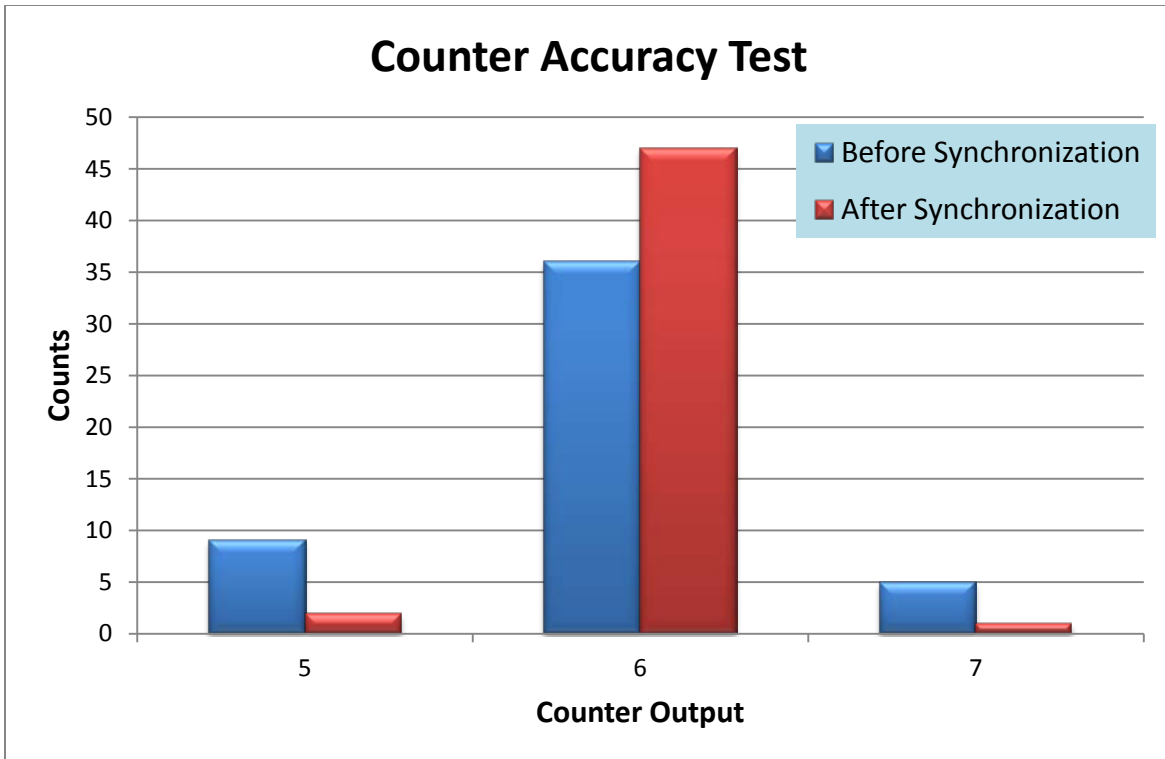


Fig. 5.4: Counter accuracy after and before synchronization, CLK freq. = 200MHz, Input time = 33.33ns

As shown in Fig. 5.4 the one-CLK-cycle error was avoided in the counter after employing dual-synchronization circuit. Deep analysis of this synchronization circuit has been offer in Chapter 4.

5.2. DLL as the First Interpolator

The first interpolation stage is the delay locked loop (DLL) in the proposed TDC. The advantage for using the DLL is that, in addition to providing a good resolution, it has good accuracy and linearity due to the delay locking process which compensates for PVT variations. The important issue in designing DLL is the stability. It is important that the DLL could lock properly to the fixed appropriate delay amount that has been designed for. To design a stable DLL, many parameters should be taken into account. First, delay elements should be designed in a way to produce stable and appropriate delay range. In addition, phase detector must work linearly which means the width of the output signal should be linearly dependent on the difference in the

phase of the input signals. Finally, charge pump (CP) must be designed carefully. The amount of current for both up and down channels must be the same. Also, current source in the charge pump must be designed appropriately. High current can cause instability, while small current slows the charging and response time of the CP. Now let us look at the designed circuit and their simulation results in DLL. Fig. 5.5 shows the schematic of the designed DLL circuit and Figures 5.6, 5.7 and 5.8 depicts its main building blocks: Tapped delay line, phase detector and charge pump.

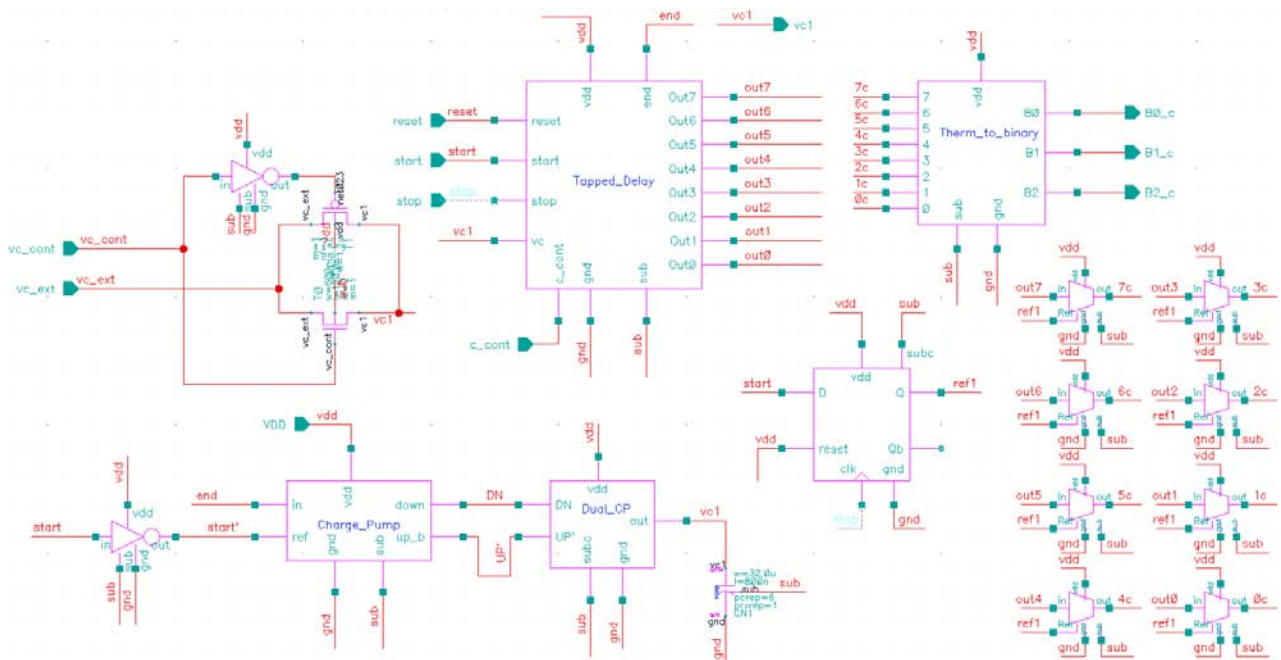


Fig. 5.5: Schematic of the designed DLL

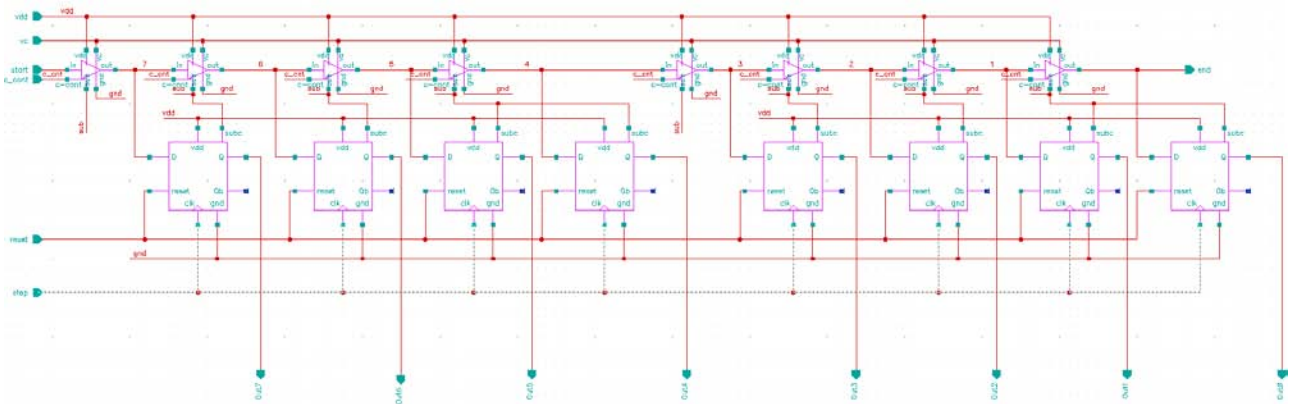


Fig. 5.6: Schematic of the designed tapped delay line

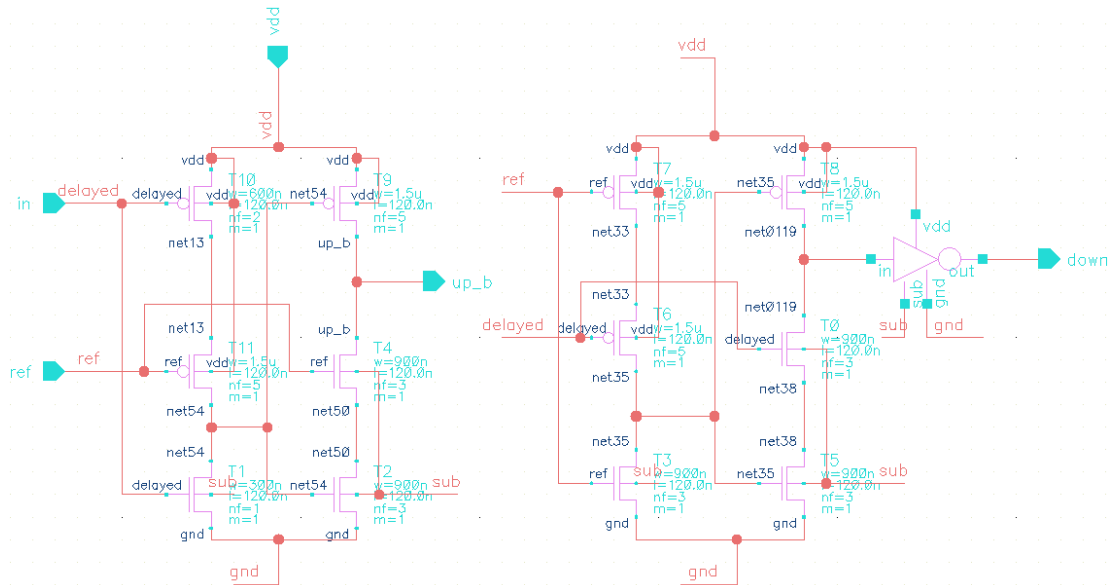


Fig. 5.7: Schematic of the designed phase detector

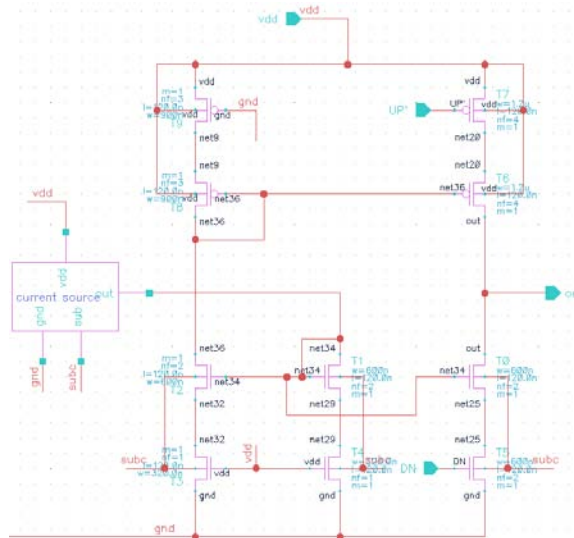


Fig. 5.8: Schematic of the designed charge pump

The first step in assessing the DLL in simulation is to see if the tapped delay line is working properly. As the main element of the tapped delay line, first, the delay element has been simulated and examined. As we mentioned in previous chapter, a current starved cascaded structure has been used in this work with two additional transistors to widen its delay range. The schematic of the delay cell has been shown in previous chapter in Fig. 4.14. Figure 5.9 shows the propagation delay of the delay cell versus applied bias voltage for both cases when C-cont is 1 and 0.

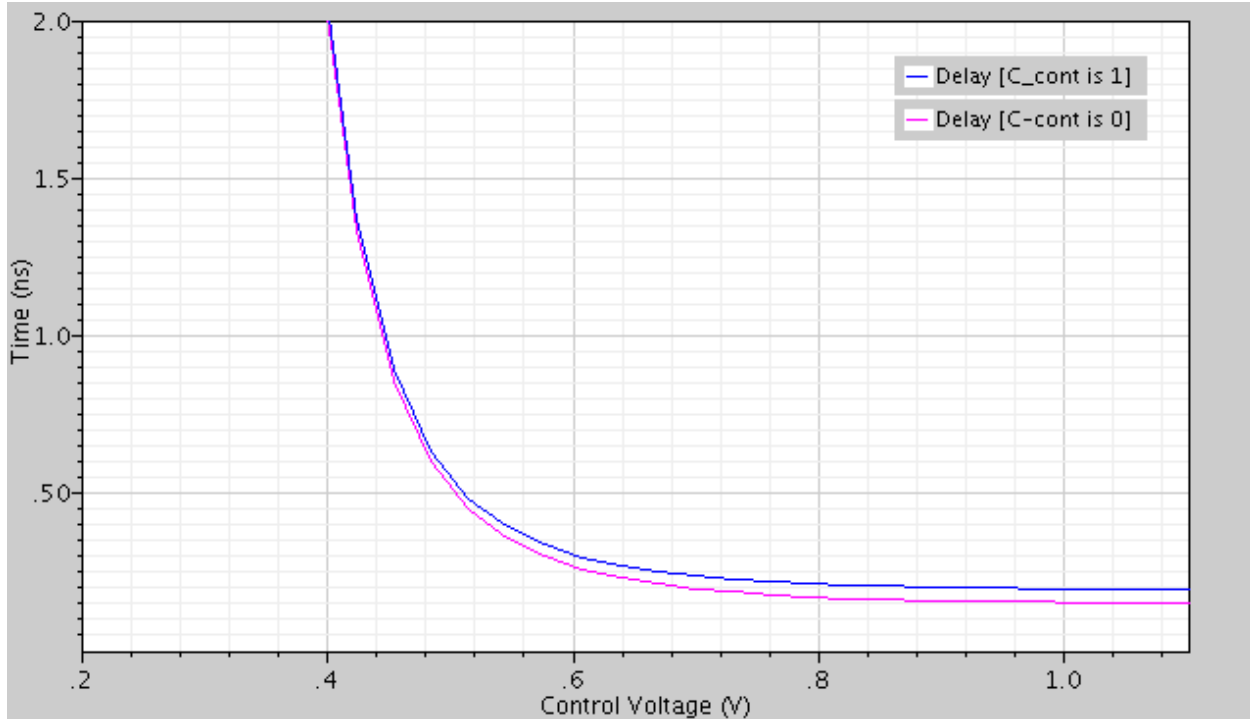


Fig. 5.9: Propagation delay of designed delay cell vs. input bias voltage

As can be seen by Fig. 5.9, when C-cont is 1, the M2c transistor becomes on and capacitors of M1c will be added to the intermediate node which is large because M1c has been intentionally sized large. But when C-cont is 0 capacitors of M1c and M2c will be in series which brings intermediate cap load to a smaller value. This is because M2c was intentionally sized very smaller than M1c. That is why when C-cont is 1 the propagation delay of the delay cells is slightly bigger than those when C-cont is 0. This small delay difference when is seen in a delay line, becomes a big deal in adjusting the performance of the DLL and VDL. In fact, this addition in the intermediate cap of the delay element helps widening the delay range of the delay cells, so they could be used for low-frequency CLK signals as well. For instance, DLL was not able to lock in Ref. CLK frequency of 100MHz when C_cont was 0. However, when C_cont was 1 locking was fulfilled. On the other hand, using 500MHz CLK frequency was not possible when additional intermediate cap was added (C-cont=1). However, locking was possible in high frequencies such as 500MHz when C-cont was 0. Figure 5.10 shows the problem of locking in 500MHz and how it was solved by shifting the delay of delay cells down by changing C-cont to 0.

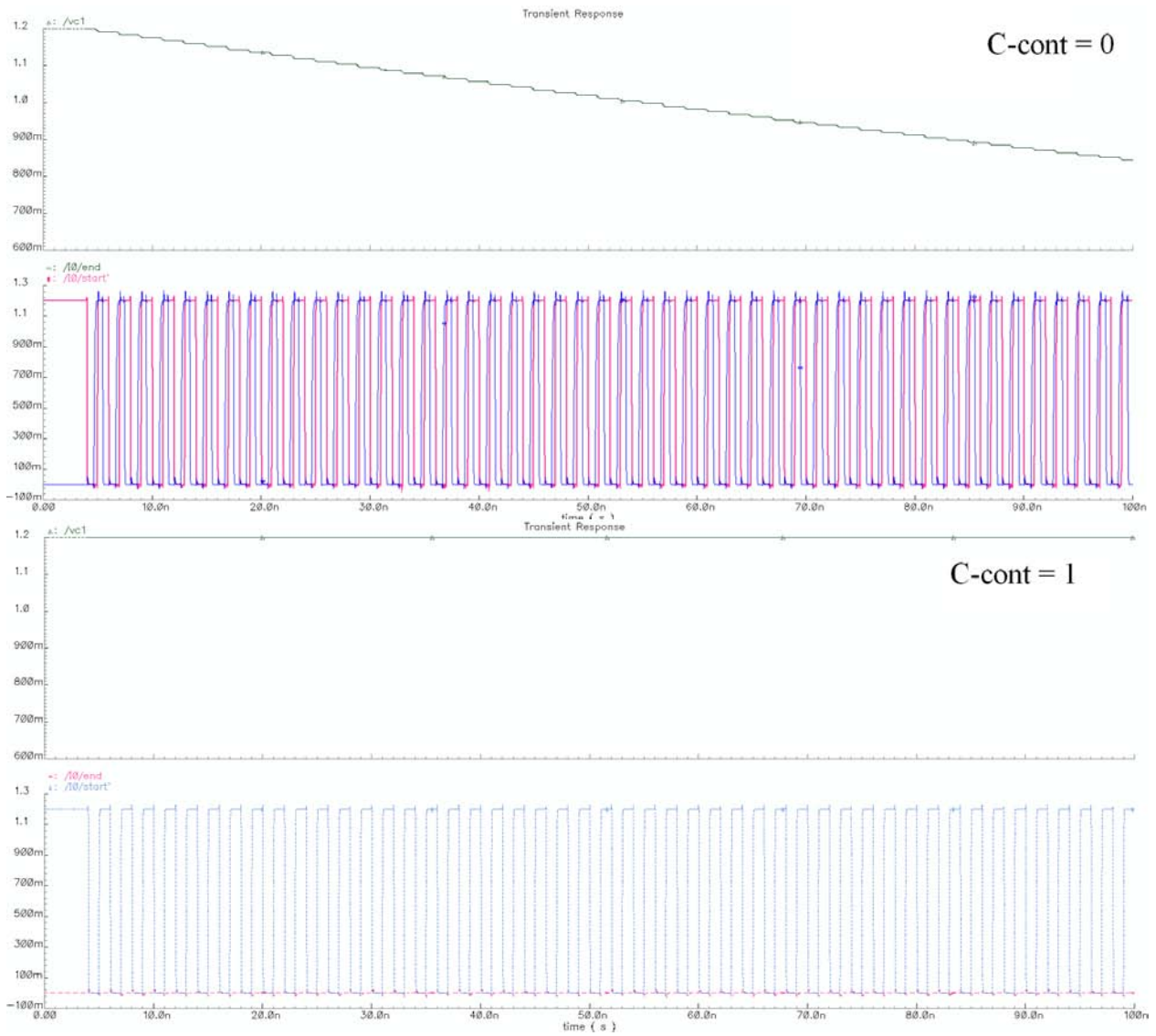


Fig. 5.10: Effect of intermediate cap on capability of locking in DLL in 500MHz CLK frequency

Clearly, figure 5.10 shows that DLL fails to lock start and end signal when C-cont is 1 (bottom diagram). On the other hand locking was successful and auto-adjustment of Vc was performed when C-cont was 0 (top diagram).

Another simulation that has been conducted for the delay elements is their temperature behavior. Figure 5.11 shows the delay cell structure and the layout of the circuit. As can be seen, using some fairly sized transistors in the intermediate node of the buffer we could extend the propagation delay of the delay cell and use it for wider range of frequencies. The delay elements were designed in such a way to minimize for the size. The total size of the delay cell

designed for the TDC is $9\mu\text{m}\times 9\mu\text{m}$. Using only 10 transistors and trying to avoid using large capacitors (for widening the delay range of the delay cells) were some methods to reduce the total size of the designed delay cell.

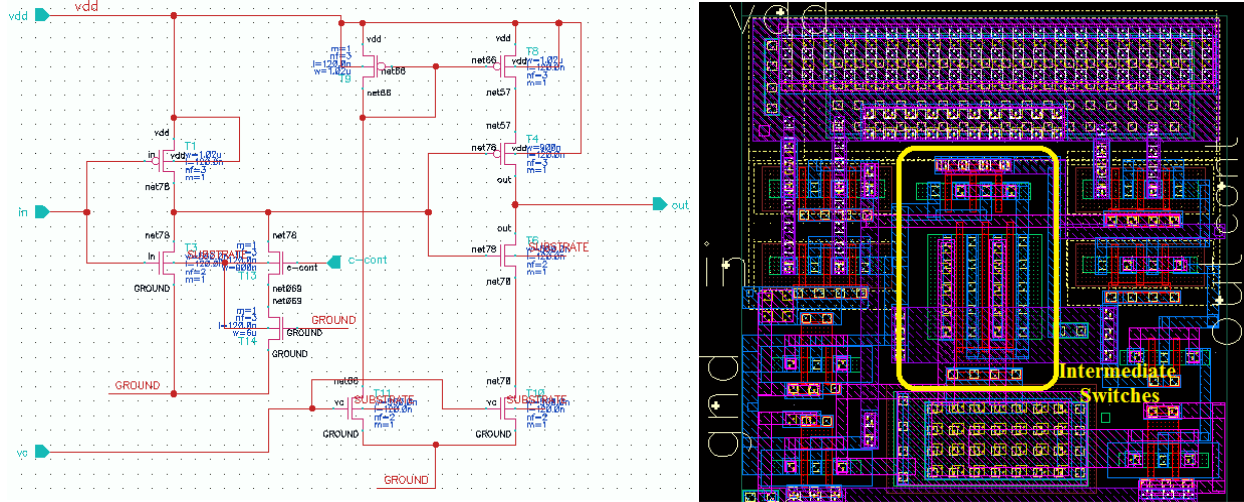


Fig. 5.11: Schematic and layout of the designed delay cell

After testing the behavior of delay elements as one of the main building blocks of the TDC, tapped delay line should be simulated. To test the tapped delay line, appropriate start and stop signal should be used based on the simulation type. For that, a set of start and stop signals was fed to the tapped delay line with different phase differences. Fig. 5.9 shows output of tapped delay line versus phase difference between the start and stop signal. As can be seen the width of each of the steps are almost the same which means that tapped delay line is working properly and linearly. This was simulated with CLK frequency of 200MHz. This means the half-CLK duration is 2.5ns and each of the delay elements should have $2.5/8 = 312.5\text{ps}$. As shown in Fig. 5.12, characteristic is almost completely linear.

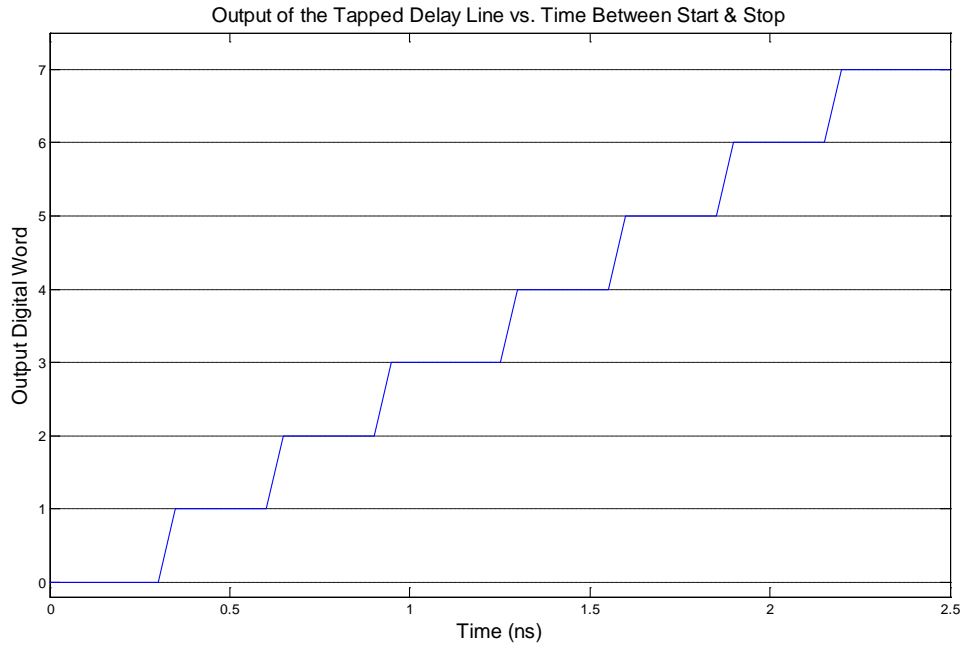


Fig. 5.12: Tapped delay line output-input characteristics

Fig. 5.13 shows the simulation results for the phase detector and the charge pump. The left figure depicts the operation of the phase detector and the states of “up” and “down” signals based on the relative position of the lead and lag signals. The right figure shows how the current of up and down signal paths based on generated output voltage.

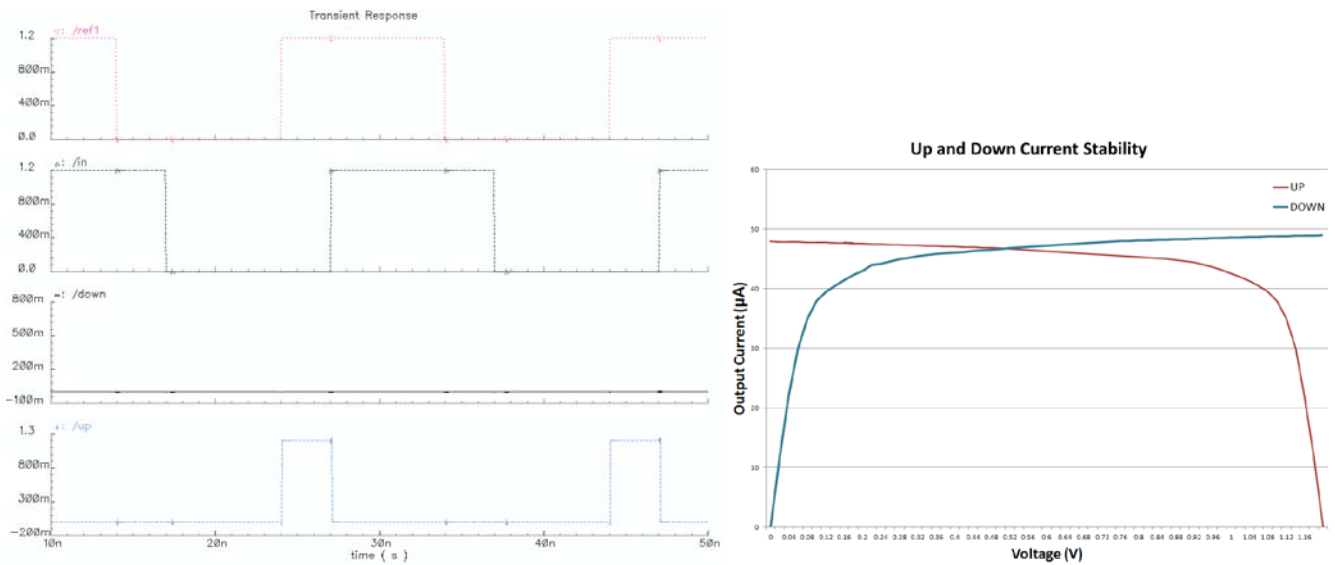


Fig. 5.13: Simulation results for the operation of the phase detector (left) and charge pump (left)

As can be seen from Fig. 5.13, UP and DOWN currents are almost the same between 0.2 to 1V. This means charge pump performs appropriately in this range, which is fairly a wide range for controlling the delay elements. Fig. 5.14 shows how phase detector and charge pump work together. It can be seen in this figure that, when UP signal is 1 output node is charging up but when DOWN signal is one output is discharged with the same amount. When both of them are 1 or 0 the output voltage keeps its value.

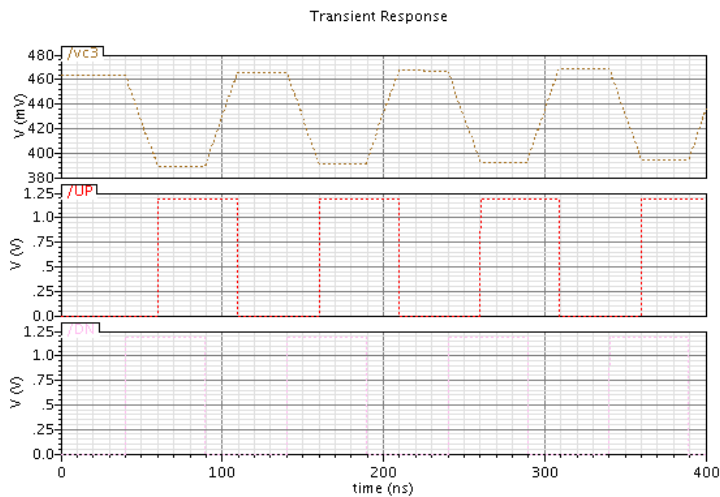


Fig. 5.14: Operation of phase detector and charge pump together

Now let us look at the way DLL locks start signal and the end signal (the signal at the end of the delay chain) by generate appropriate control voltage (V_c) at the output of charge pump (Fig. 5.15).

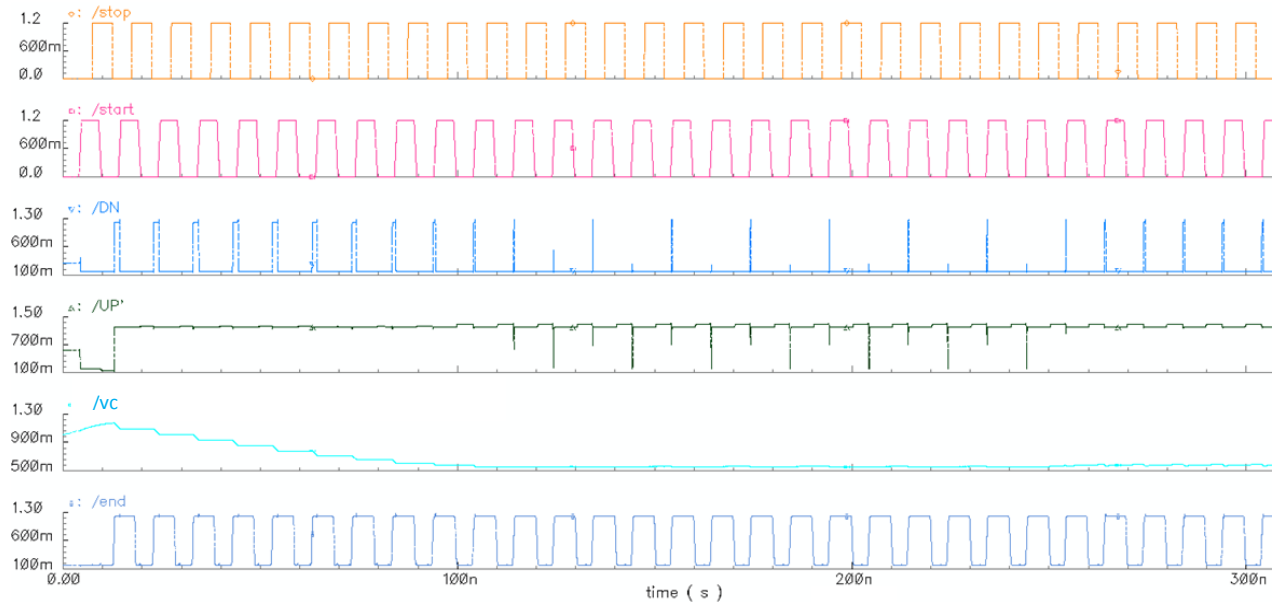


Fig. 5.15: Operation of DLL to lock the two signals at beginning and end of the delay chain
 As can be seen from Fig. 5.15, at time zero, the start and the end signal have a slightly different phase which leads to generating DOWN signal which decreases Vc. This reduction in the output voltage of CP forces the delay elements to slow down a little bit. And this increase in the delay of the delay elements gradually helps the end signal to catch up with the start signal.

5.3. VDL as the Second Interpolator

Vernier delay line (VDL) as the third and finest stage of the TDC plays an important role in determining the resolution and accuracy of the TDC. Details on the operation of VDL have been offered in chapter 4. A second delay locked loop has been used to regulate the 2nd row of the delay cells in VDL during unwanted PVT variations. Actually, the delay elements in the stop chain of VDL are controlled with the same control voltage that has been generated from the first DLL (Vc1). The start chain's delay elements are fed with the control voltage that is generated with the 2nd DLL (Vc2). Fig. 5.16 shows how DLLs generate Vc1 and Vc2 for the VDL.

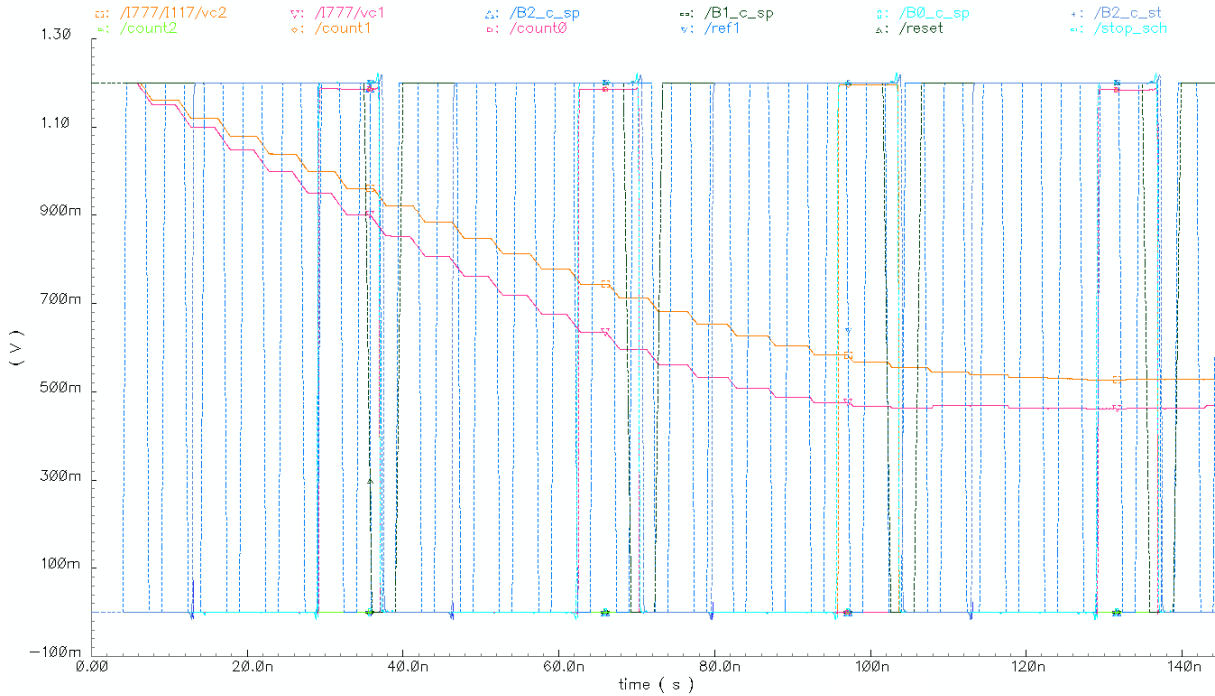


Fig. 5.16: Locking process of the control voltages to feed the start and stop chains' delay elements As can be seen in Fig. 5.16, V_{c2} is slightly bigger than V_{c1} which implies delay elements in the start channel to have slightly larger delay amount. But this difference is regulated with the delay locked loop and the way of implementing the structure of VDL (as was explained in chapter 4) so that desired resolution and dynamic range can be obtained for the VDL stage. Similar to simulation of DLL in previous section, we ran a sweep test to determine the output-input characteristic of the TDC and find its resolution and linearity (Fig. 5.17).

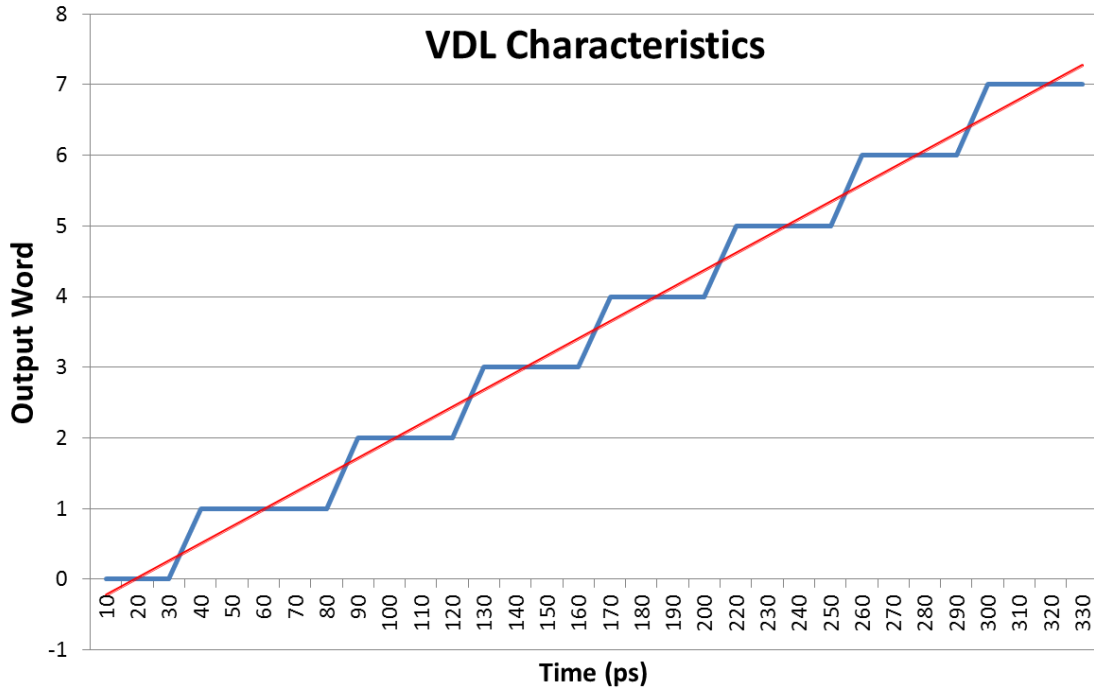


Fig. 5.17: Output-input characteristic of the VDL

As can be seen there from this figure, the characteristic is almost linear there is a small nonlinearity at the beginning and at the end of the time period. This is due to different load seen at the output nodes of the delay elements in the beginning and end of the delay chain compared to intermediate delay cell. The resolution can be found from the width of the steps which is about 40ps and the dynamic range as shown is around 320ps for the VDL which is what we expected from the calculations that were given in chapter 4.

5.4. Operation of Entire TDC

The ultimate step in testing the design is to test the entire TDC. There are different methods for determining the performance of the TDC. A sweep in the input of the TDC and check how the output is changing is one of the methods which lead to output-input characteristics of the TDC. It provides an overview on the performance and linearity of the TDC. But, it need a lot of

simulation runs for sweeping all over the dynamic range and finally the shape of the achieved diagram would be something similar to output-input characteristics of the stages from which TDC was constructed.

Another technique to test the accuracy of the TDC is to do a very large number of measurements with the same input and look how the output of the TDC changes. The input here is the time difference between the start and stop signal and the output is the digital word consists of the digital outputs of the counter, half-CLK counter (HCC), DLL and VDL. Drawing the histogram of the output code and measuring the variance of it gives a very clear idea about the accuracy of the TDC. Fig. 5.18 shows the test designed setup for this purpose.

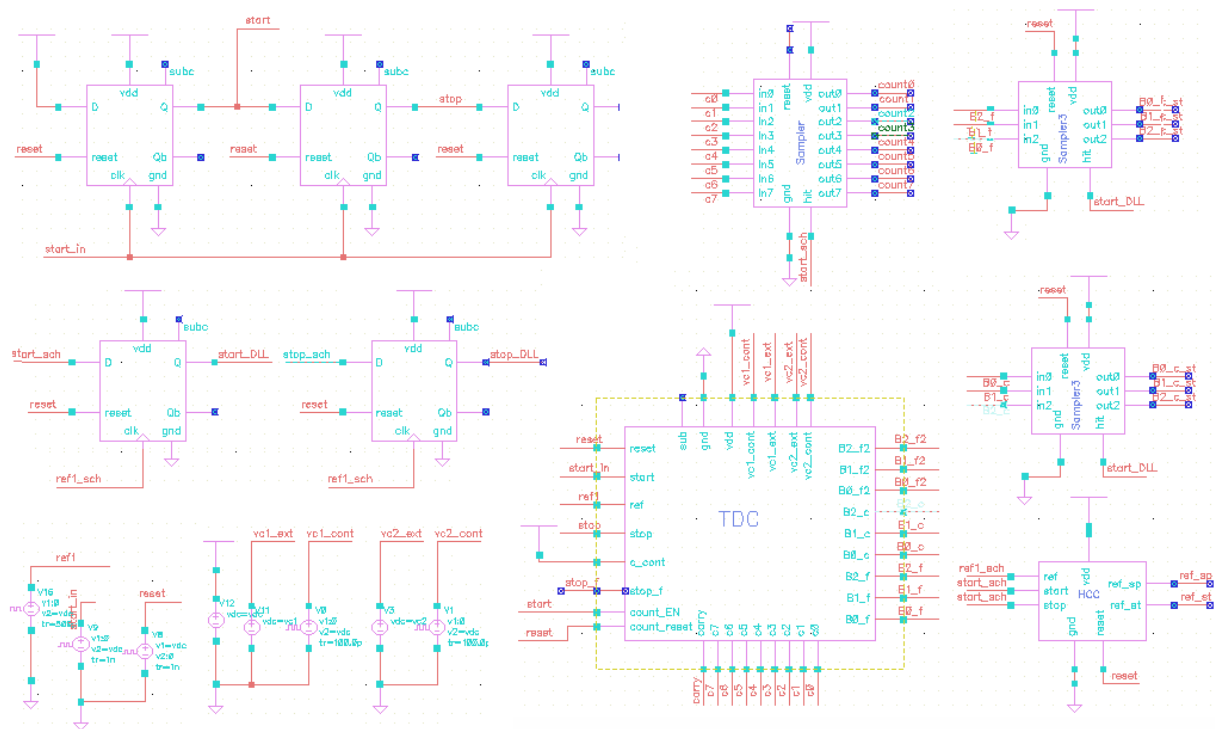


Fig. 5.18: Test setup for obtaining the accuracy of the TDC

As can be seen from the figure, the TDC is in the middle. There are some registers on the right side to store the outputs from counter, HCC, DLL and VDL when the command signals come. On the left side, there are some DFFs to generate those commands signals based on the start and reference CLK. Also, one three were used (left-top corner) to generate the hit signals that should feed the TDC. A latched-start signal is generated from the very short start pulse using the first DFF and the next two DFFs are used to generate a stop and a read signal synchronous with the start signal but with a fixed delay amount phase difference. As can be seen using this

idea, the phase difference between the start and stop signal is always kept fixed and equal to the period the start signal. Then, a huge number of measurements on the TDC can be performed without changing its input to assess its accuracy. Fig. 5.19 shows the histogram of the number of shots vs. time slots of the TDC for 100 numbers of measurements. The start signal's period is 16.667ns which is equal to the time difference between the hit signals.

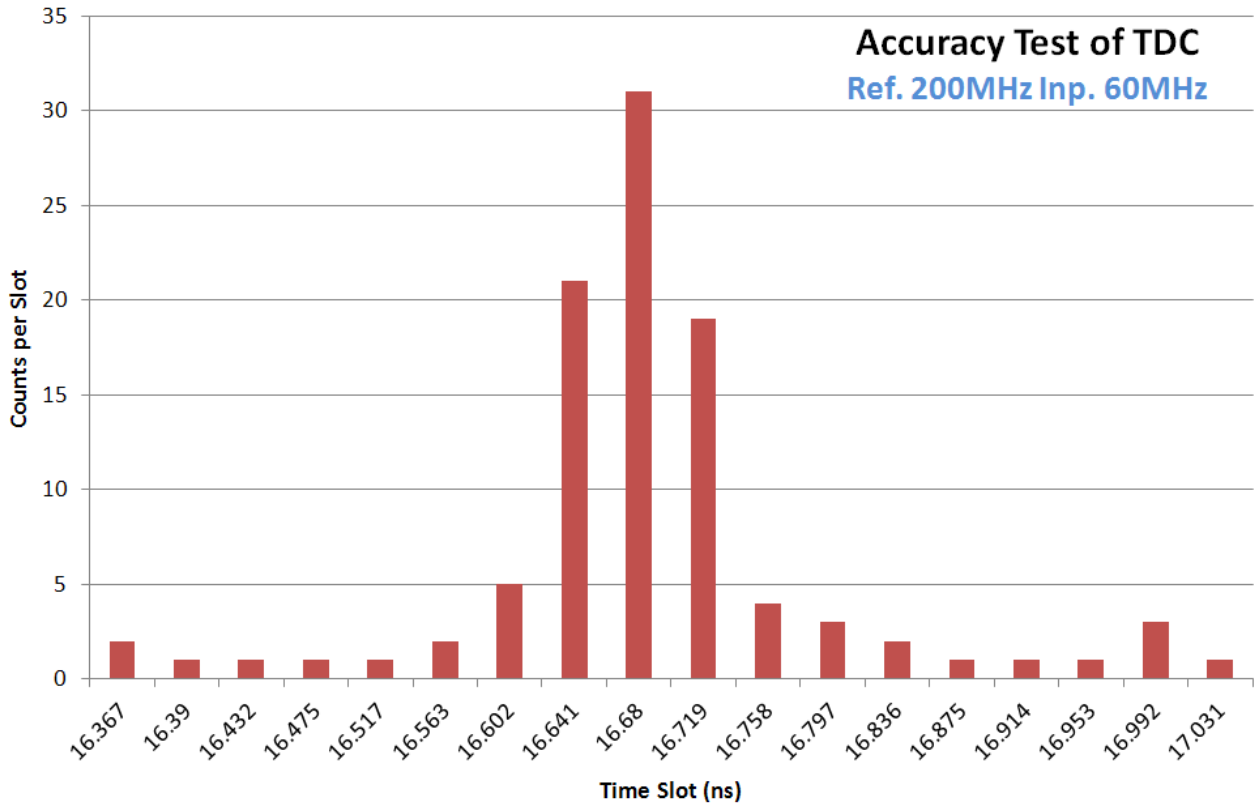


Fig. 5.19: Accuracy test of TDC: number of counts per time slot for input time of 16.67ns

From this Figure, many parameters can be extracted. First of all, this figure shows that TDC is working properly to determine the timing information of the input signals. In fact, the average value of the histogram which is 16.68ns is very close to the expected value of 16.67ns. In addition, this is the result only for 100 measurements. By increasing the number of measurements, the statistical errors can be reduced.

The most important parameter in TDC is the resolution or T_{LSB} of the system. This figure shows that the timing information can accurately be achieved with the designed TDC within

± 40 ps. Indeed the T_{LSB} , as we expected is 40ps. Other factors that can be extracted from this figure are the variance and precision of the histogram. Variance for this histogram for 100 measurements is 12.2ps which shows that TDC offers good accuracy in simulations. However, this value might get worse for the results for the same setup in measurements. As can be seen the number of counts suddenly increases 16.992ns and 16.367. These sudden changes are due to wrong DLL output that caused a big difference in the measured time compared to the expected value (312ps shift). But the number of counts in these time slots is a lot less than number of counts for 16.68ns time slot, as was expected.

The next test setup is to find the simulation results for the nonlinearity of the TDC. Theoretically all bins of the TDC should have the same width. However, in real word, widths of the bins are different from each other. This happens due to variations between the delay cell and the load seen by each node in the delay lines. As we explained in chapter 3, DNL and INL are two factors describing the nonlinearity of the TDC. To measure them, a large number of measurements should take place with random time inputs. For this, a two input signals with two different frequencies has been used for start and stop signals. This leads to different input time for each of the measurements and by doing this experiment for a very long time we were able to find the histogram of the counts per each bin in the TDC. As the TDC has two interpolation stages each with 8 bins, there are totally 64 bins within a CLK cycle. Figure 5.20 shows the distribution of the counts per each of the bins in the CLK cycle. We recorded outputs of the VDL and DLL for 1000 measurements to form this histogram. Dynamic nonlinearity and integral nonlinearity can be found from this histogram as going to be mentioned.

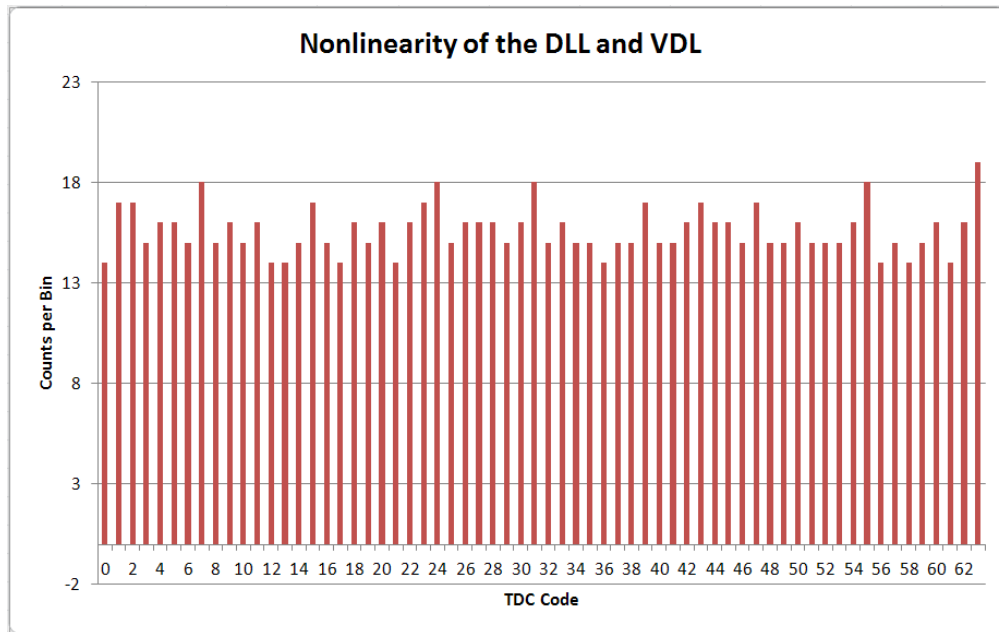
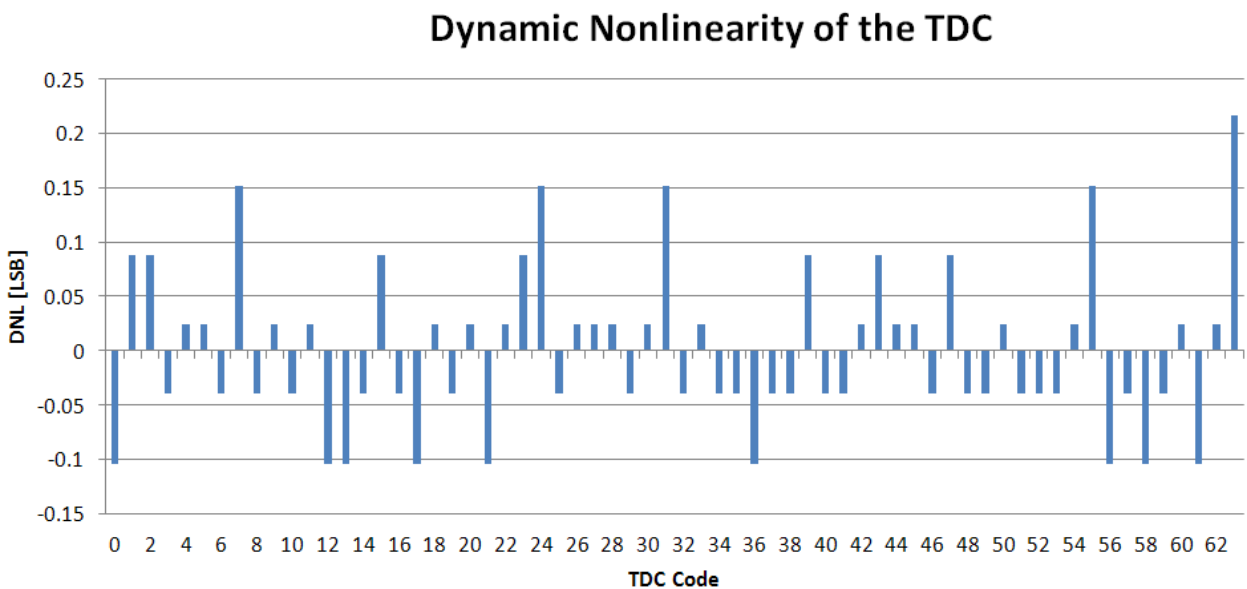


Fig. 5.20: Counts per time bin histogram for DLL and VDL

As can be seen from fig. 5.20, the histogram peaks a little on at the end of VDL at the last bin of the VDL's cycle. Also it peaks at the last time bin where the last node of DLL is. This could be due to different capacitive load at the end of the delay lines, for last delay cell in DLL and VDL. Based on this histogram, the DNL and INL of the TDC have been found and can be seen in Fig. 5.21.



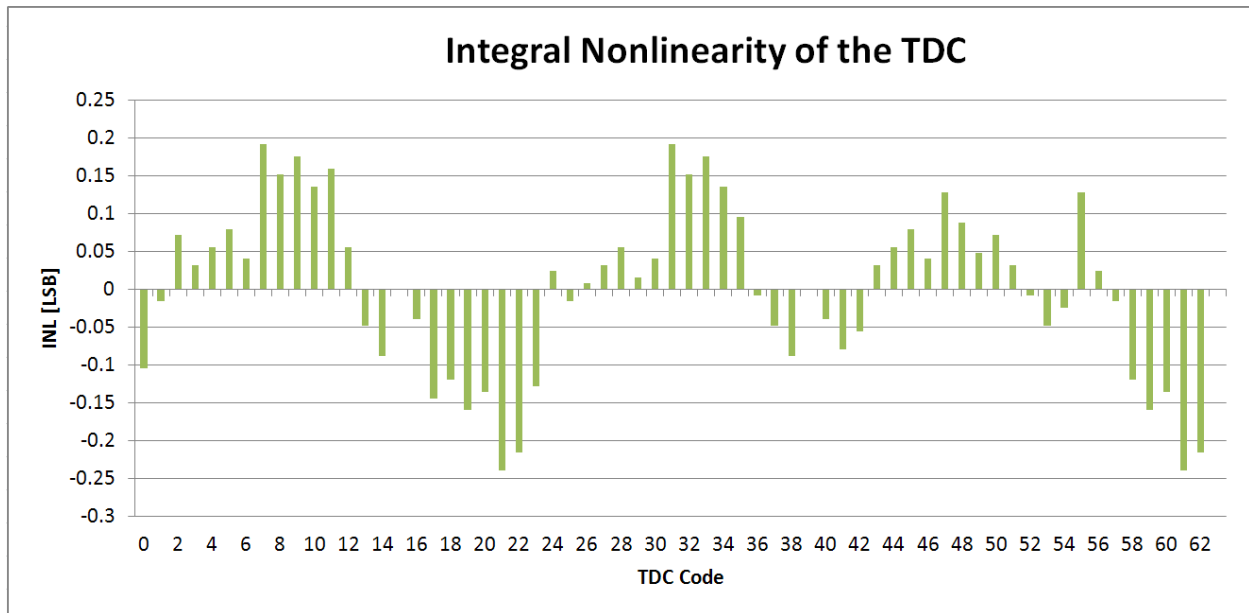


Fig. 5.21: DNL and INL report for the TDC found through simulations

As can be seen from Fig. 5.21, the TDC showed $0.15T_{\text{LSB}}$ of dynamic nonlinearity and about $0.25T_{\text{LSB}}$ of integral nonlinearity through simulations.

Chapter 6

TDC MEASUREMENT RESULTS

6.1. Layout of the TDC

The prototype TDC was fabricated in 0.13 μm standard CMOS technology process sponsored by CMC microsystems. The photomicrograph of the chip is shown in Fig. 6.1. The total die area is 2mm \times 2mm including the routes, pads and an additional test structures. The area used by the entire TDC measures for 0.11mm² from which about 40% is occupied by the routings. Unlike the TDC circuit, these routings were not optimized as the matter of area consumption. This means even more area optimization can be done in future works. DLL, VLD and counter occupy 0.00824, 0.01495 and 0.0126 mm² respectively which totally measures for 0.04mm². A total number of 36 pads were used to provide power supply and signals to the TDC and transfer the data out from the counter, HCC, DLL and VDL to the FPGA.

Several issues were considered in designing the layout of the TDC. We tried to lay out all the delay cells of each of the stages together in a compact area to minimize the statistical mismatch to which they correspond. Another important point was matching the signal paths of the start and stop signals to avoid delay mismatch between the hit signals.

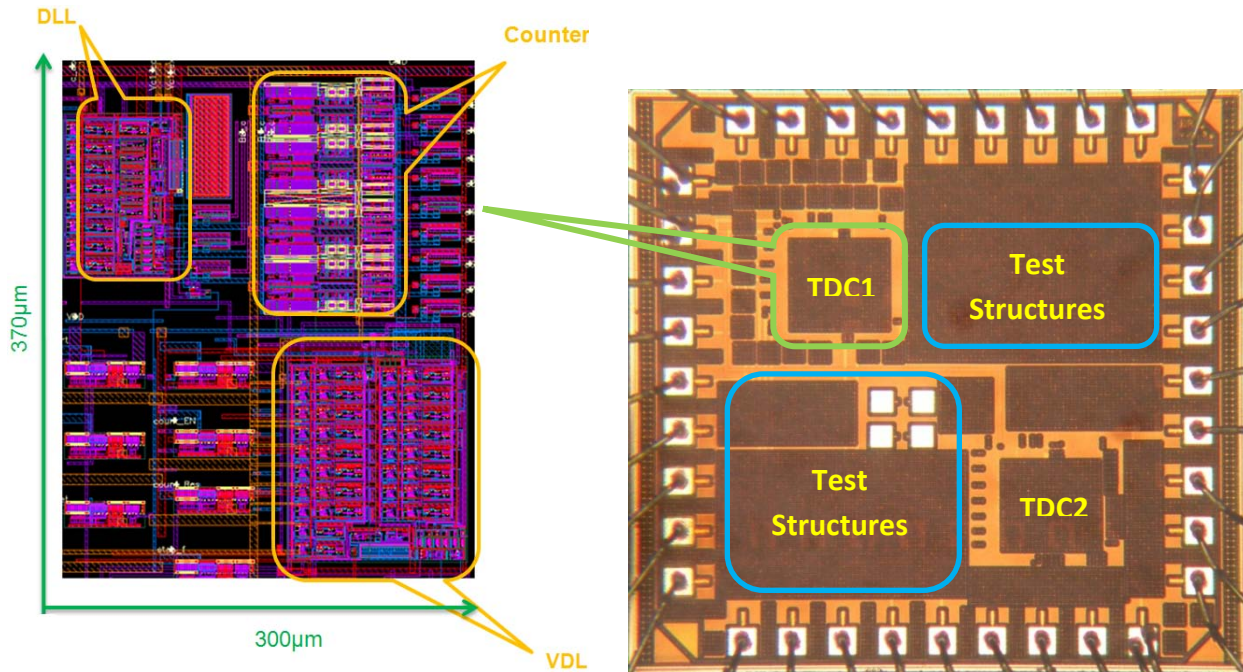


Fig. 6.1: Layout and Photomicrograph of the TDC prototype chip

6.2. Test Setup Board

The measurements were conducted using a printed circuit board (PCB) as the base for the test setups. TDC chip was mounted directly on the board to minimize the parasitic capacitances and inductance of the package. A measurement setup similar to what was shown in previous chapter for simulations was employed to design the PCB.

To measure the accuracy of the system, a constant time input should be injected to the TDC. So, a start and stop signal with a fixed phase difference were used as the input signals of the TDC. In order to do that, we employed the same idea that we explained in chapter 5. For that, stop signal was generated from that using an external delay generator. This passive delay generator was used for sweep test as well.

Another test that was conducted for the TDC was the sweep test in which we used a fixed start signal and its delayed versions and use them as a set of stop signals. For that, we used a passive delay unit called Optronic-TRRC1. We will describe the operation of this device later.

Figure 6.2 shows the test setup board that has been designed to fulfill the measurement tests of the TDC. To power up the PCB and TDC, four LM317 regulators were used to generate 5, 2, 1.65 and 1.5 voltages to supply the TDC and the components on the PCB. These regulators provide stabilized voltages in the range of 1.2 to 35 volts with current drive capability up to 1.5A. There are two modes of injecting input signals to the board. Signals can be provided from either the external signal synthesizers or the FPGA. Three switches have been employed prior to the input buffers to select if the start, stop and reference CLK signals are coming from the FPGA or the external signal generators. OPA3695 opamp was used to buffer and level-shift the input signals. This type of opamp provided by Texas Instrument offers ultra-high slew rate with extremely small propagation delay which helps providing fast-transition signals with sufficient drive current and appropriate amplitude. These signals are fed to the TDC chip then.

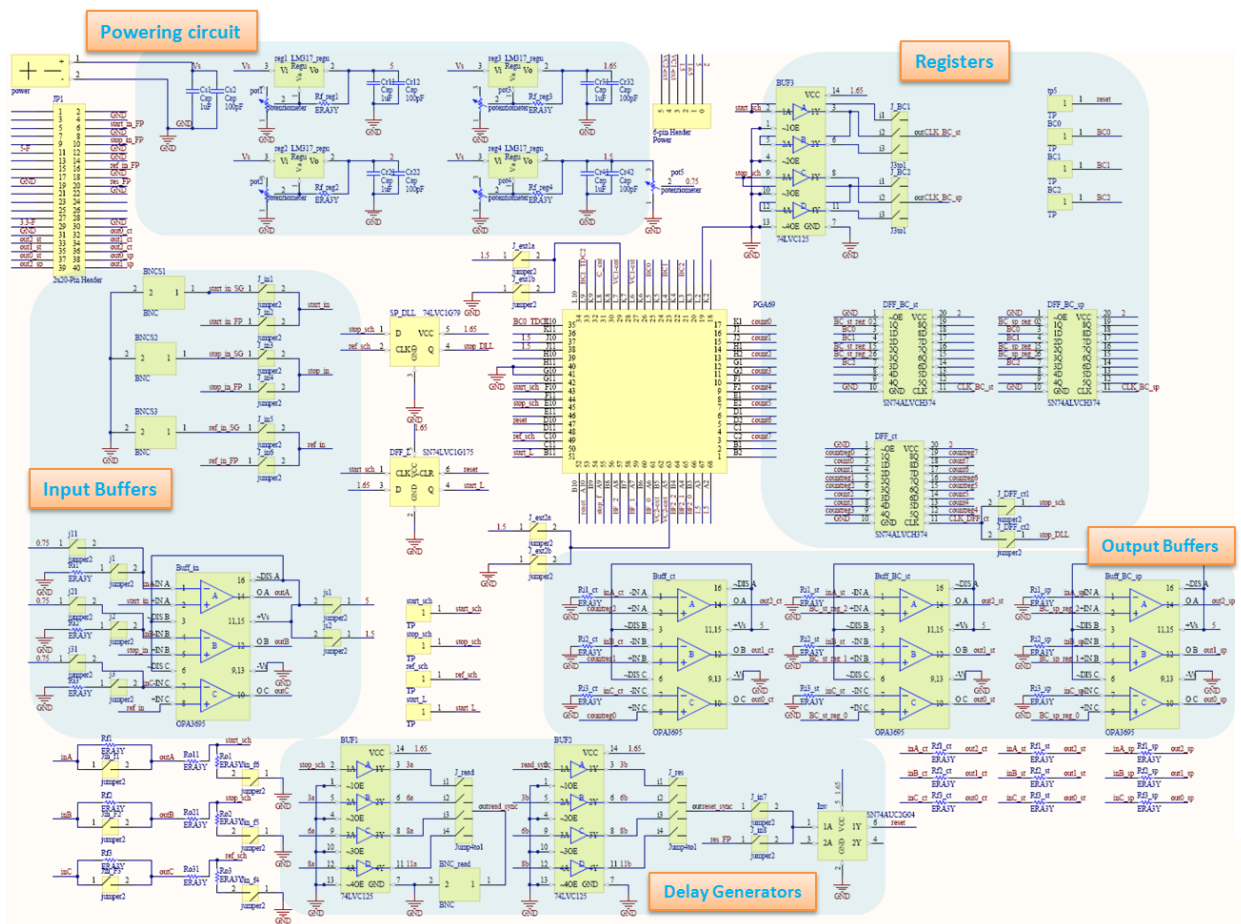


Fig. 6.2: Test setup for measuring the performance of TDC

Delayed versions of the stop signal are generated on the board as read and reset signals. Generating some command signals such as read and reset from the stop signal helps to enable synchronized reading the data and resetting the TDC right after each measurement. A string of buffers with fixed delays was used to generate these signals. The number of buffers in the signal path can be adjusted with jumpers to add or remove a buffer from the chain.

The next step after generating appropriate input signal including start, stop, reference, read and reset, they will be injected to the TDC chip so that measurements could be performed. After each measurement event, TDC output which includes the output data from the counter, HCC, DLL and VDL will be latched using DFFs. These DFFs should be fast in order to minimize the latency of the measurements. Since the signals coming from TDC have amplitude of 1.5V, they should be amplified in order to satisfy the minimum amplitude requirement of the FPGA. The FPGA we are using accepts input signals with minimum 3.3V amplitude. OPA3695 was used for this purpose. This opamp performs level-shifting on the signals and drive them to the FPGA with sufficient drive current.

The final fabricated PCB is shown in Fig. 6.3. The total size of the PCB is 15cm×15.7cm. The input reference signal has the frequency of 200MHz and is generated from the outside of the PCB from an external signal synthesizer (Anritsu MG3694A). Therefore, start and stop signals that are coming from either the FPGA or external pulse generators are asynchronous with respect to the reference CLK signal. The output signals are sent to FPGA, however, Tektronix TDS3054 oscilloscope was used in parallel to record and show the output signals whenever needed. The total jitter for the measured time combining the rms jitter of the typical start and stop signals and reference CLK signal (coming from external pulse generators - Anritsu MG3694A) was measured to be ~5 ps.

Anritsu MG3694A provides frequency range up to 40 GHz. It has excellent phase noise performance and offers high output power. The oscilloscope that has been used (Tektronix TDS3054) is capable of recording signals with the frequencies up to 500MHz and it provides high sample rate of 5 GS/s.

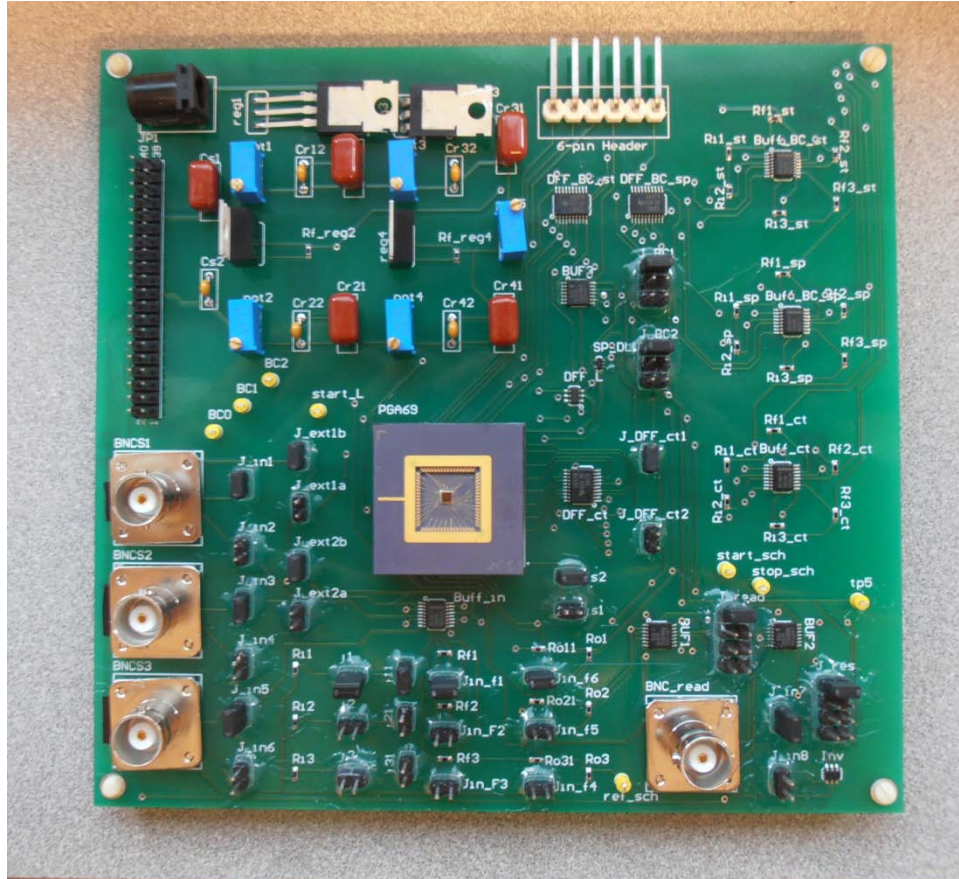


Fig. 6.3: PCB designed for testing the prototype TDC

6.3. TDC Characterization

In order to characterize the TDC, a sweep on the time input of the TDC should be made. Therefore, a start signal together with a set of stop signals was generated with different phase delays with the start signal. Phase delay generation was done using a passive delay unit, called Optronics-TRRC1. Optronics-TRRC1 is capable of applying a fixed and accurate delay to its input. As can be seen from the photograph of TRRC1 shown in fig. 6.4, the delay amount is controlled with some switches on the device. As shown in this figure, these switches are capable of adding delays between $\frac{1}{32}$ to 4 times of the input signal's period. Using a start signal with the frequency of 50 MHz and changing the switches' states, a sweep in the input signal which is the phase difference between the start and stop was fulfilled. Using these start and stop signals,

the TDC was measured over 0 to 40ns with a time step of 625ps. Fig. 6.5 shows the output-input characteristic of the TDC. The sweep test was repeated a couple of times and measured output for each point was averaged to reduce statistical variations.



Fig. 6.4: Optronics-TRRC1 as the delay generator for generating stop signals

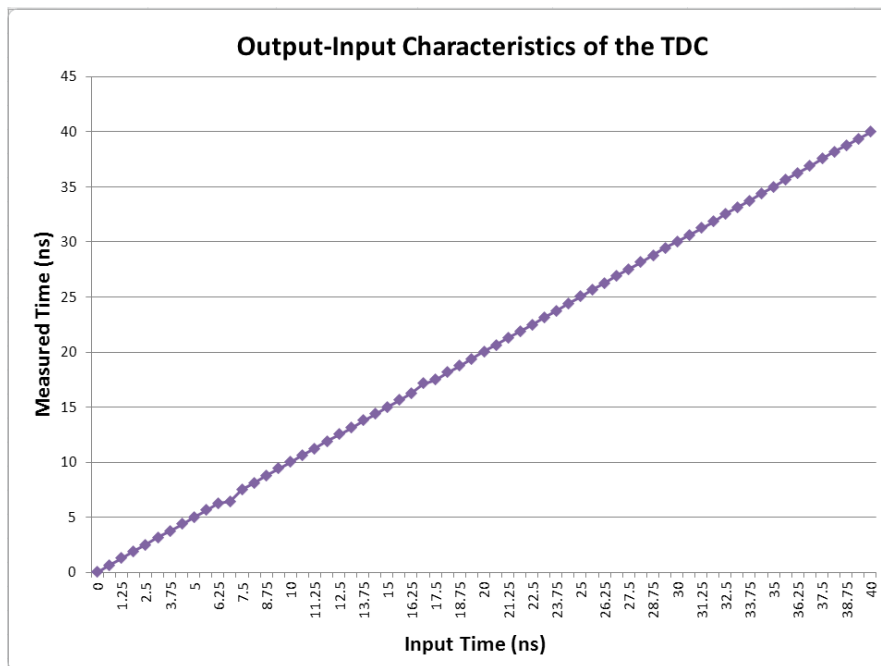


Fig. 6.5: TDC output-input characteristics performed in sweep measurement setup

Figure 6.5 shows that TDC operates with good linearity over the measured portion of dynamic range. However, there are some points with relatively large deviation from their expected values. In order to provide numerical values for linearity performance of the TDC, DNL and INL of the TDC was measured. We will explain about them later in this chapter.

6.4. TDC Accuracy

Figure 6.5 just shows the measurement results for a limited set of points each were achieved after averaging the results for couple of measurements to minimize the statistical variations. So, this figure does not incorporate to the resolution and accuracy of the TDC well enough. In order to measure those parameters an accuracy test similar to the accuracy test introduced in chapter 5 has been employed. Optronics-TRRC1 delay unit was used to generate a fixed phase difference between the start and stop signals over the whole measurement period and the output data was collected. To perform a trustable accuracy test, output data form 1000 measurements were collected. Fig. 6.6 shows the results for the accuracy test of the TDC. To compare with the simulation results, they were added to the figure as well.

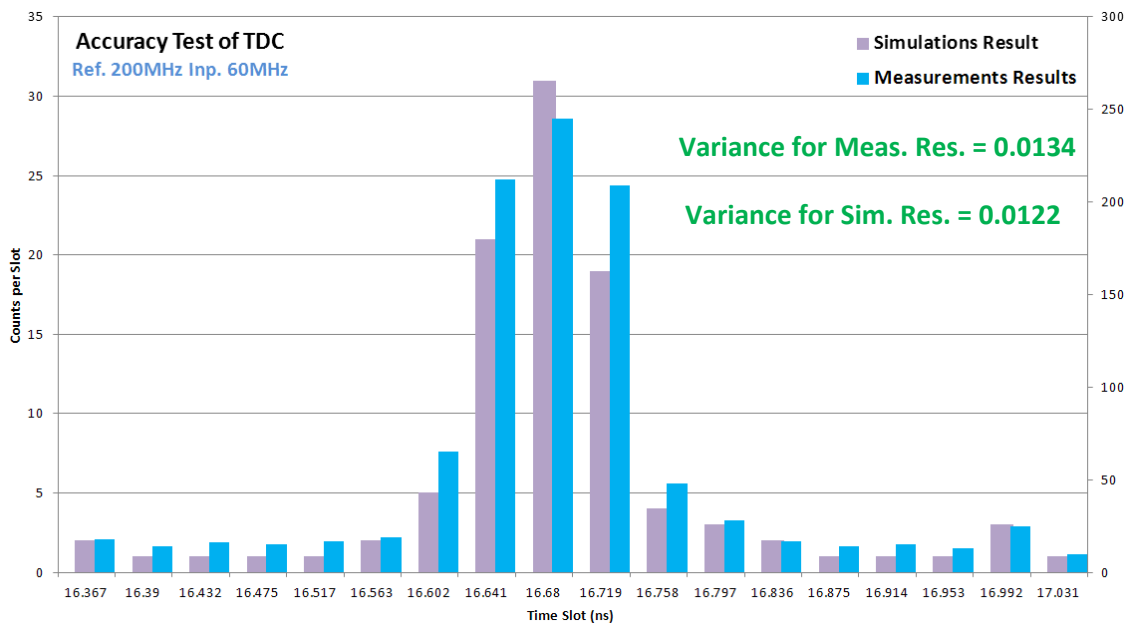


Fig. 6.6: TDC accuracy test achieved from 1000 measurements with a fixed time input

As can be seen from this figure, the measurement results show more variance compared to the simulation results, as we expected. This is because of the added parasitic capacitors and resistors in the intermediate nodes of the delay chain that increases the nonlinearity of the system. The variance of the measurement results was calculated to be 0.0134 which is slightly bigger than the variance calculated for the simulation results (0.0122). This difference was

expected due to non-idealities in measurements situation compared to the simulations. Capacitive loads of each of the nodes are very sensitive to the parasitics and mismatch between the nodes. Due to these reasons, uniformity of the bins of the TDC should be more degraded in fabricated TDC compared to the simulated one. But, still the results present good accuracy in the performance of the TDC. As can be seen in fig. 6.6, the resolution which is the distance between each of the bars is set to be 39ps.

6.5. Linearity of the TDC

To measure the nonlinearity of the TDC, a very large number of measurements should be taken. These measurements should be randomly distributed in the time input axis (phase difference between start and stop signals). In order to provide the TDC this set of start and stop signals, two external signals with different frequencies of 20MHz and 30MHz were used. Data from the TDC for over 1000 number of measurements were collected. Figure 6.7 shows the results for the number of counts per each bin of the tapped delay line which was designed with eight delay cells and no delay locked loop. This gives a sense on the nonlinearity of the coarse TDC when no delay regulator is employed.

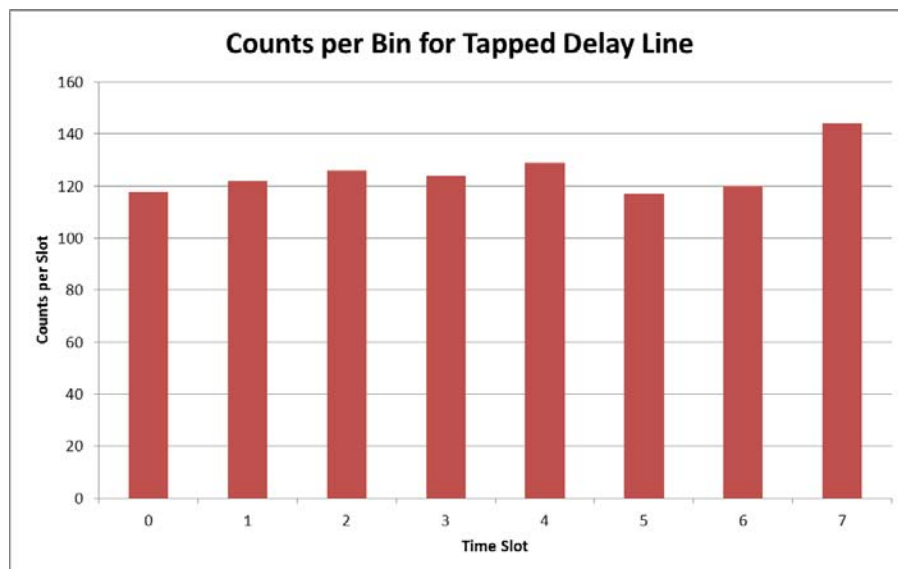


Fig. 6.7: Nonlinearity measurements of the tapped delay line

At the next step, we ran the same measurement test setup for the DLL. The nonlinearity of the DLL for its 8 bins is shown in Fig. 6.8. Again, 1000 measurements were collected to generate the diagram. As can be seen from this figure, the linearity has been improved by locking the delay of the whole delay chain to the period of a fixed reference CLK signal using DLL method. This method helps reducing the process, voltage and temperature variations. Therefore, delay mismatch between the internal nodes of the delay line will be reduced.

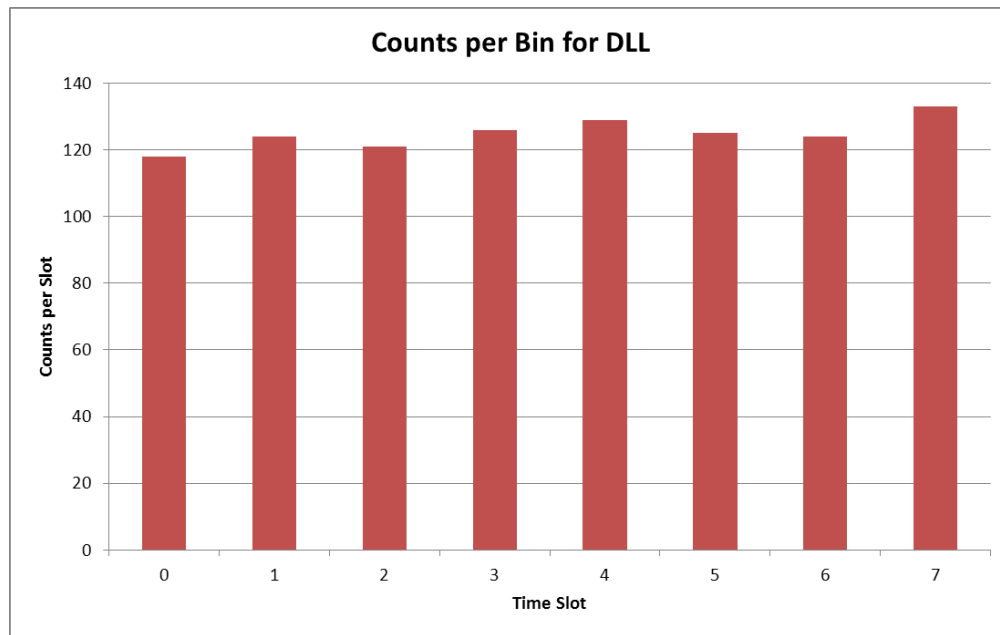


Fig. 6.8: Nonlinearity measurements for delay locked loop

Measuring the nonlinearity of the DLL as a part of the entire hierarchical TDC, presents only nonlinearity behaviour for the coarser part of the TDC. In the next step, counts' distribution for all of the bins of the TDC consisting of the bins of DLL and VDL will be provided. DLL has 8 bins and in each of its bins, 8 bins for VDL exist. Therefore, there are 64 levels of interpolations in half of a CLK period and therefore 64 bins for DLL and VDL. The same experiment for measuring the nonlinearity behaviour of interpolation bins simulation-wise in chapter 5. Figure 6.9 shows the counts distribution for the 64 bins of the TDC which was collected from a set of 5000 measurements.

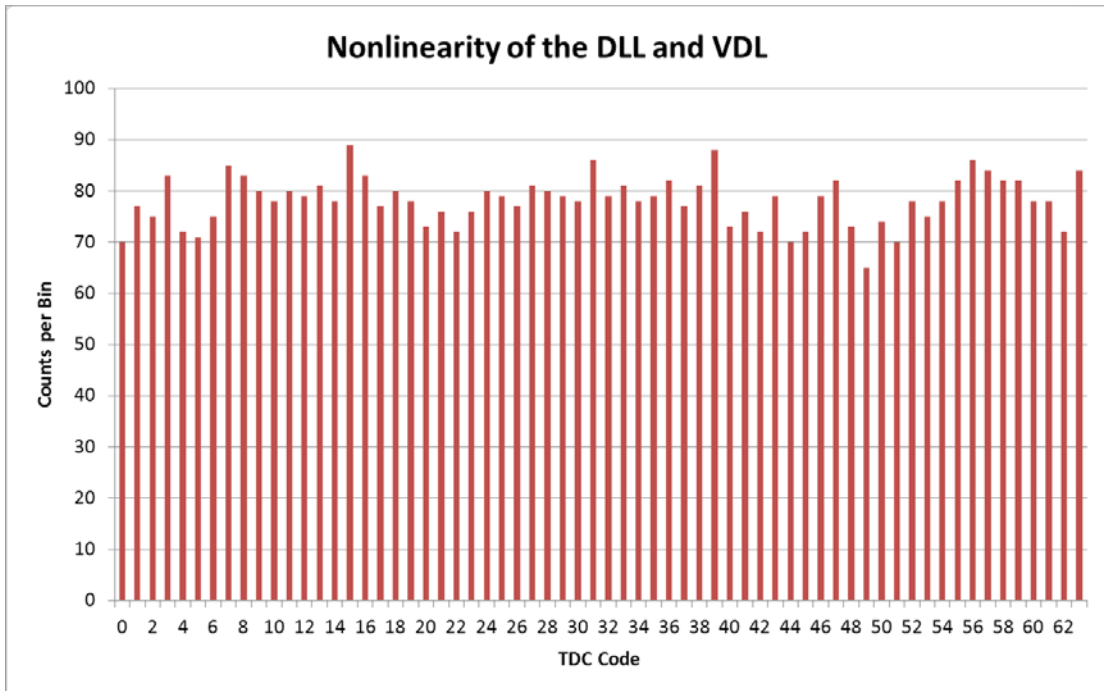


Fig. 6.9: Counts distribution for 64 bin of the TDC to measure nonlinearity for 5000 measurements

Dynamic and integral nonlinearity of the TDC was calculated from the collected data in figure 6.9. The DNL and INL calculation was done by the formula that was introduced in chapter 3. Fig. 6.10 and 6.11 shows the measured nonlinearity performance of the designed time digital converter.

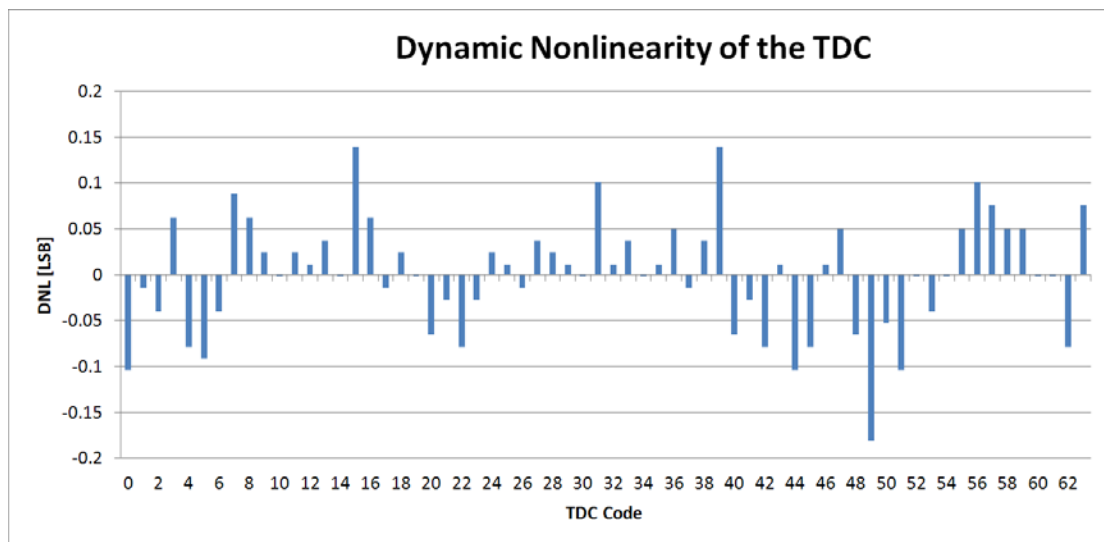


Fig. 6.10: Dynamic nonlinearity of the designed TDC for 5000 measurements

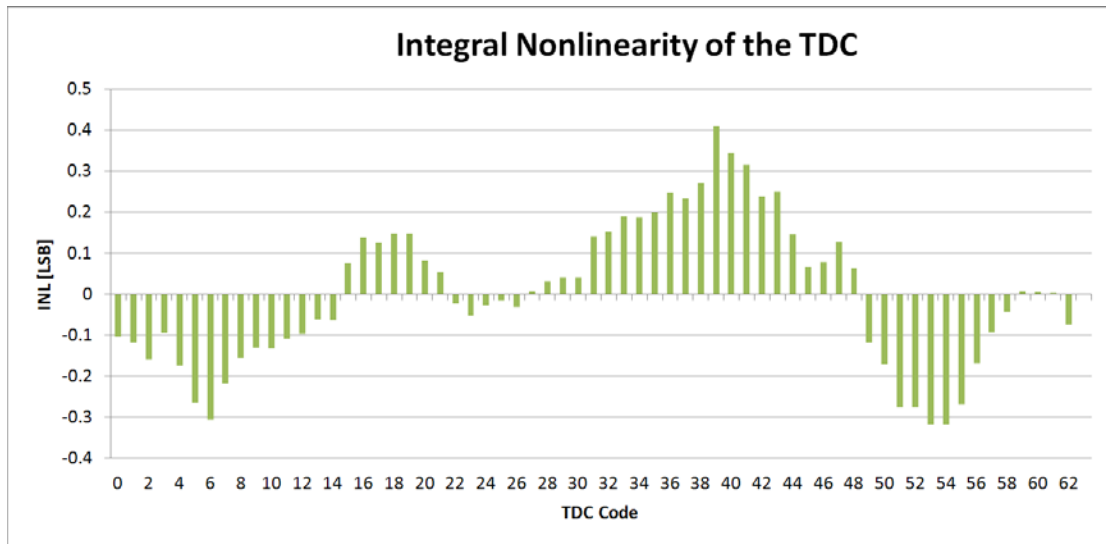


Fig. 6.11: Integral nonlinearity of the designed TDC for 5000 measurements

As can be seen from these figures, prototype TDC has the maximum DNL of $0.2T_{LSB}$ and maximum INL of about $0.4T_{LSB}$. This clearly shows that, the designed TDC offers good linearity compared to recent similar works.

Chapter 7

Summary and Future Work

Summary

A complete review on all of the performance characteristics of the TDC prototype is shown in Table 7.1. The resolution of the prototype TDC is measured to be 39ps which was delivered using 3 stages of interpolation. A dynamic range of 1.28 μ s was obtained through an 8bit counter used as the coarse TDC. Nonlinearity was measured to be maximum of 0.2LSB for the DNL and 0.4LSB for the INL of the TDC. Finally, the area occupied with the entire TDC and its surrounded routings accounts for 0.11mm² which was obtained due to hierarchical architecture of the TDC. Three stages helped avoid using large number of delay cells. Half CLK interpolation idea also helped to minimize the area occupied with TDC and makes it useful for in-pixel time measurement applications. A comparison table, Table 7.2, compare the characteristics of the prototype TDC with some of the recent TDC design works in the literature that were close to our design method and application.

Parameter	Value
Technology	0.13 μ m standard CMOS
Power Supply voltage for TDC	1.5 V
Power Supply voltage for peripheral circuits on PCB	1.5, 1.65, 2 and 5 V
TDC Size	0.11 mm ²
Ref. CLK frequency range	100-500 MHz (the values in the table are for 200MHz CLK)
LSB resolution	39 ps
Dynamic Range	1.28 μ s for 200MHz CLK signal
Dynamic Nonlinearity (DNL)	0.2 TLSB
Integral Nonlinearity (INL)	0.4 TLSB
PVT variation regulation	was enabled through DLLs
Power consumption	39 mW

Table 7.1: Summary of the measured performance of the prototype TDC

Authors	LSB	Dynamic Range	Word Length	DNL	INL	Area	PVT Calibration	Technology	Power	In-pixel Design	Published
A. S. Yousif et al	31 ps	2 ns	6	0.625 LSB	0.725 LSB	0.49 mm ²	Phase Locked Loop	0.13μ CMOS	1 mW	✓	IEEE TNS 2007
E. Charbon et al	55 ps	55 ns	10	0.08 LSB	1.89 LSB	NA	Phase Locked Loop	0.13μ CMOS	550 mW	✓	ISSCC 2011
R. C. Jaeger et al	8 ps	32.8 ns	12	-	-	0.26 mm ²	Just Delay Line Symmetry	0.13μ CMOS	7.5 mW	✓	JSSC 2010
J. P. Jansson et al	12.2 ps	204 μs	15	-	0.66 LSB	7.5 mm ²	-	0.35μ CMOS	40 mW	-	JSSC 2006
P. Chen et al	50 ps	250 ns	12	-	1.1 LSB	0.225 mm ²	Just Delay Line Symmetry	0.35μ CMOS	-	✓	IEEE TNS 2006
A. Mäntyniemi et al	1.2 ps	327μs	28	-	2.67 LSB	4.32 mm ²	Delay Locked Loop	0.35μ CMOS	33 mW	-	JSSC 2009
B. K. Swann et al	97ps	80ns	8	0.2 LSB	0.3 LSB	2.88 mm ²	-	0.5μ CMOS	175 mW	-	JSSC 2004
This Design	39 ps	1.28 μs	14	0.2LSB	0.4LSB	0.11 mm²	Dual Delay Locked Loop	0.13μ CMOS	39 mW	✓	

Table 7.2: Comparison table for characteristics of the prototype TDC

The purpose of designing this TDC was to use it for PET imaging system. In order to provide a compact accurate PET scanner, readout and signal processing circuits should be implemented in-pixel, beside the photodetector. So, one of the main concentration of this work was to minimize the occupied area with the TDC so it could be fitted inside the pixel. This usually conducted by limiting the resolution or dynamic range of the TDC. However, in this work we tried to keep the resolution and DR high enough for PET application and reduce the size as well. That is why 3 stage TDC and half-CLK interpolation idea were proposed in this work.

Future Work

For future works even more minimizing on the size of the TDC can be done. Further optimization on the area can be performed. By minimizing the routes and connections and removing the spaces between the DLL, counter and VDL even a smaller TDC can be built. Also as a novel idea, sharing one DLL for a couple of neighbor TDCs could be investigated. This helps to eliminate using one DLL per pixel. However, this may increase the nonlinearity as the

propagation delay of delay cells of a TDC might be controlled with the DLL of the neighbor TDC. For applications which require less dynamic range, reducing the number of the bits of the counter or even eliminating it can greatly help reducing the total size of the TDC.

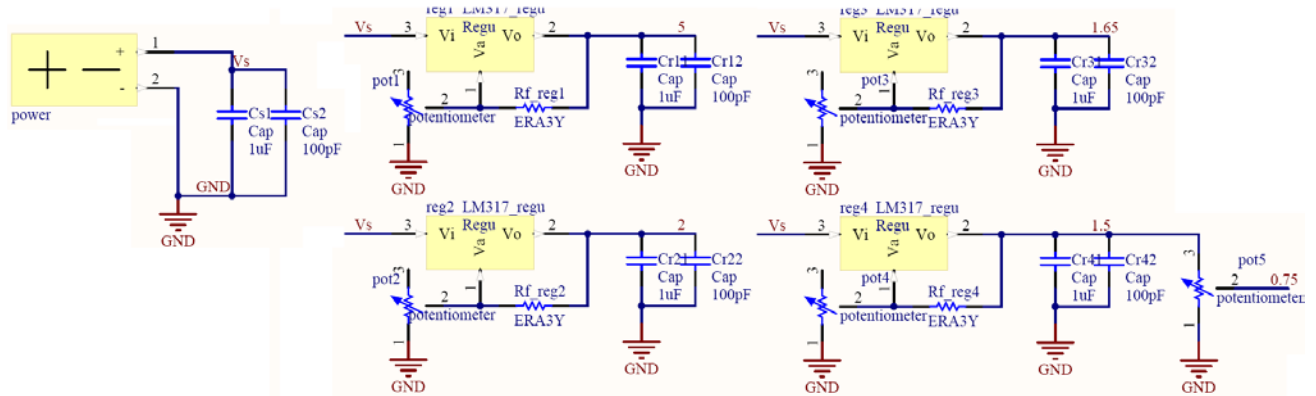
To increase the linearity of the system, special attention can be paid to each of the nodes of the delay line to match them carefully. This helps to have better matching between the loads seen by each of the nodes and increase the linearity of the TDC. However this might cost more area consumption. As some delay control blocks may need to be added to control the delay of each of the cells separately. An integral look-up table can be added to the TDC based on the measurement results of the designed prototype. This helps to assign specific delay for each of the cells. The trade-off is again increase in the total size of the TDC.

Resolution is another parameter that can be improved, but this comes with increased number of delay cells which leads to area enlargement. It seems that further resolution improvement is not necessary because the time resolution is limited by the jitter of the photodetector. However, with future optimization on SPADs and by employing better and faster photodetectors, increasing the resolution of TDC can be done.

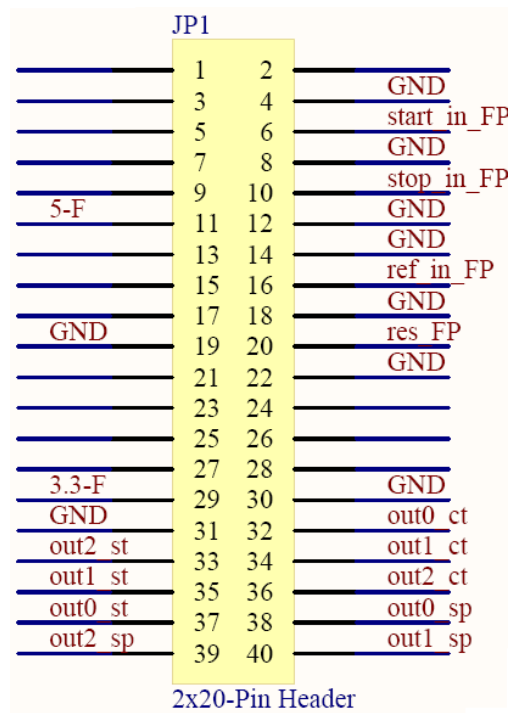
Finally, it should be noted that this TDC can also be used for other time-correlated photon counting application such as fluorescence lifetime imaging as they also require in-pixel time measurement. In order to do that, specifications of the TDC should be slightly adjusted for these applications by adding or removing the number of bits in counter to adjust the dynamic range. Resolution of the designed TDC seems to be high enough for these applications. Also, some applications may require strict power budget for the TDC which should be considered as well. Moving toward using more advance technologies such as 90nm and 65nm leads to using smaller VDDs and decreasing the total power of the system.

Appendix 1.

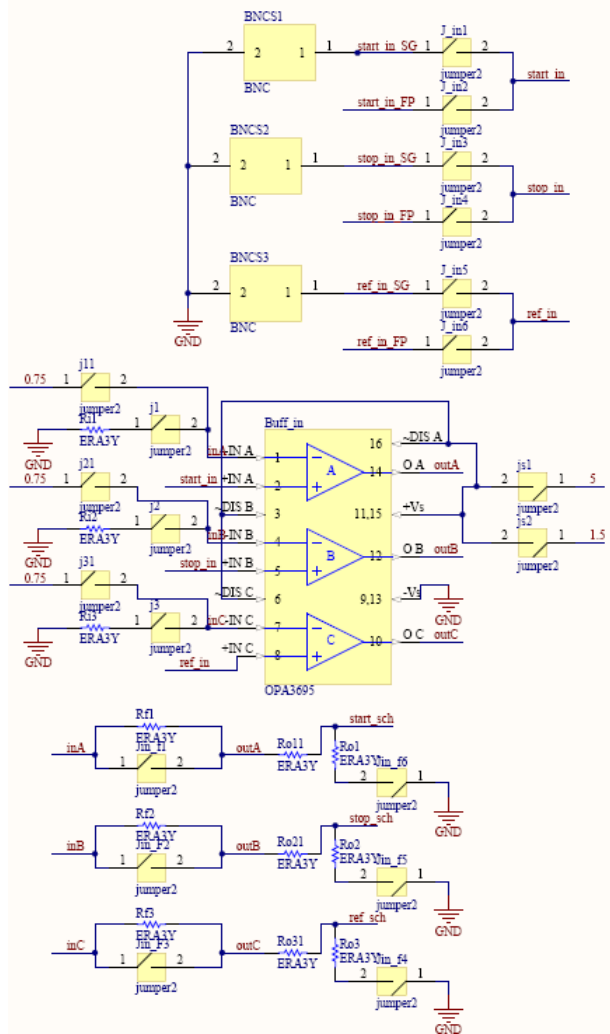
Power Supply Circuit:



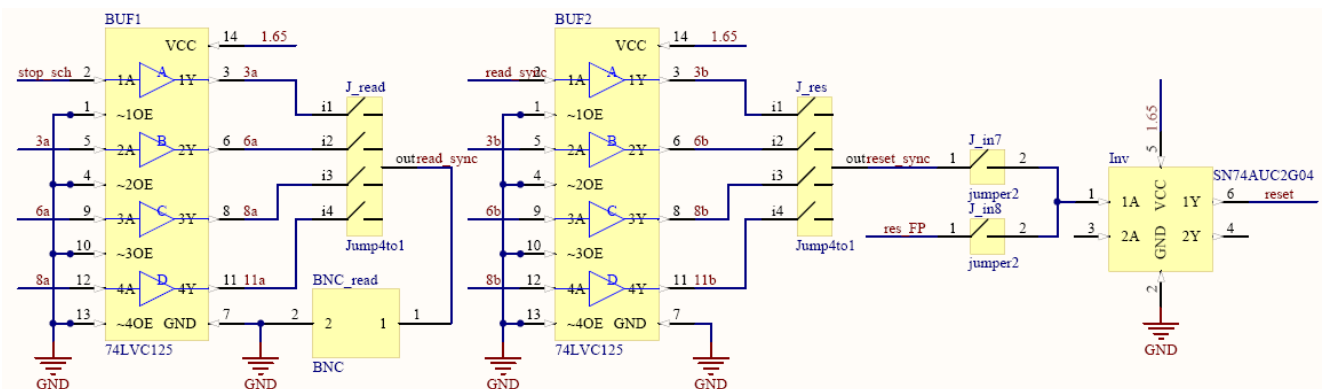
FPGA Connector:



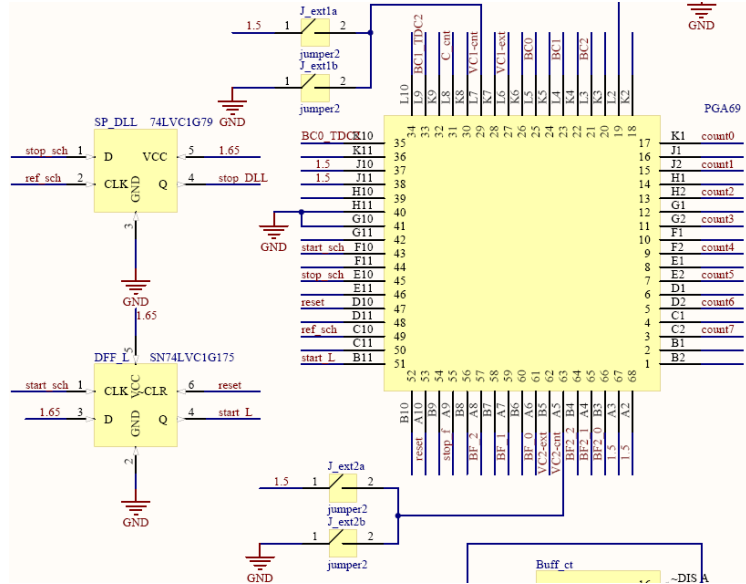
Input Signal Preparation:



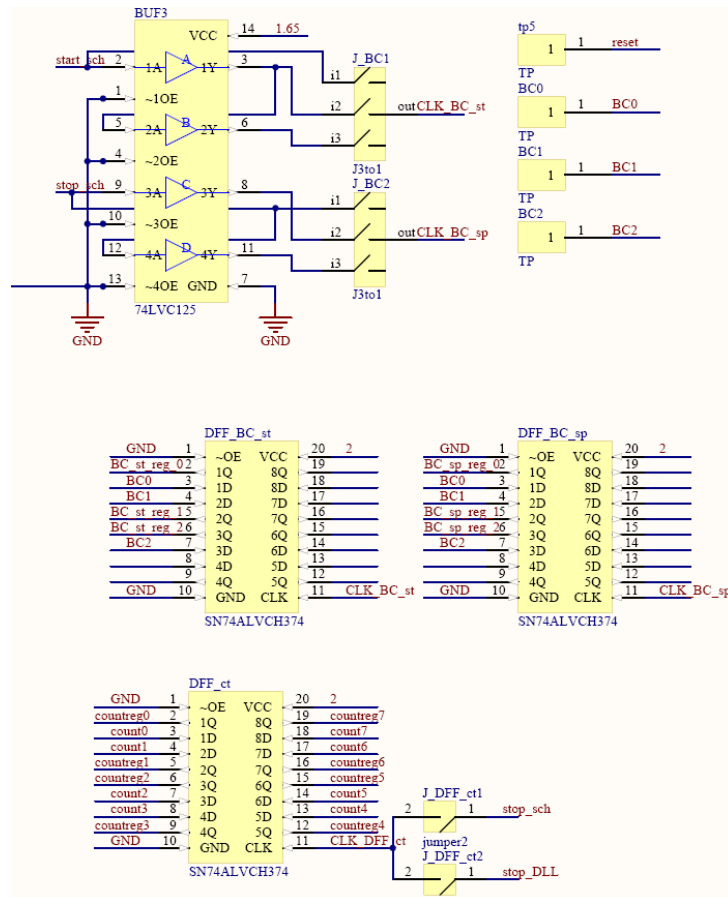
Read and Reset Signal Generators:



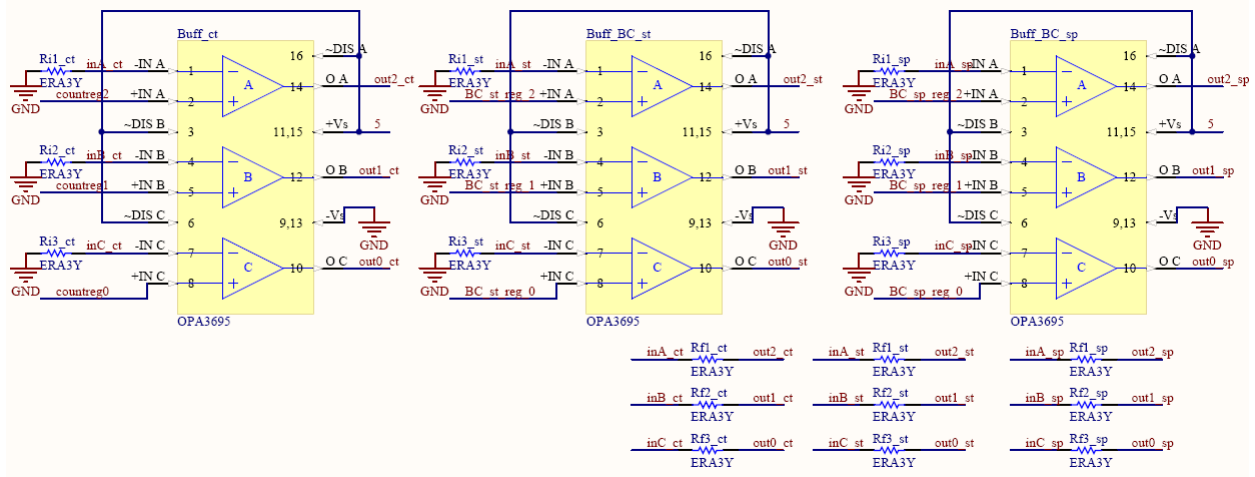
Chip stand:



Output Registers:



Output Buffers:



References

- [1] M. N. Wernick, J. N. Aarsvold (Eds.), "Emission Tomography: The Fundamentals of PET and SPECT," Elsevier, 2004.
- [2] GB. Saha, "Basics of PET Imaging: Physics, Chemistry, and Regulations," Springer, 2005.
- [3] A. Dhawan, B. D'Alessandro, and X. Fu, "Optical imaging modalities for biomedical applications," *IEEE Rev. Biomed. Eng.*, vol. 3, no. 1, pp. 69–92, Dec. 2010.
- [4] M. Bigas, E. Cabruja, J. Forest, and J. Salvi, "Review of CMOS image sensors," *Microelectron. J.*, vol. 37, pp. 433–451, 2006.
- [5] S. R. Cherry, "In vivo molecular and genomic imaging: New challenges for imaging physics," *Phys. Med. Biol.*, vol. 46, pp. R13–R48, 2004.
- [6] P. Seitz, A.J.P. Theuwissen, Eds. "Single-photon imaging", Springer, Berlin Heidelberg, 2011.
- [7] J. Tian, J. Bai, X. Yan, S. Bao, Y. Li, W. Liang, and X. Yang, "Multimodality molecular imaging," *IEEE Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 48–57, Sep. 2008.
- [8] K. M. Mudry, R. Plonsey, J. D. Bronzino, Biomedical imaging, CRC Press, 2003.
- [9] W. Becker, "Advanced Time-Correlated Single Photon Counting Techniques", Springer, 2005.
- [10] T. K. Lewellen, "Recent developments in PET detector technology," *Phys. Med. Biol.*, vol. 53, 2008, pp.287–317.
- [11] W. W. Moses, "Time of flight in PET revisited," *IEEE Trans. Nucl. Sci.*, vol. 50, no. 5, pp. 1325–1330, Oct. 2003.
- [12] V. C. Spanoudaki, and C. S. Levin, "Photo-detectors for Time of Flight Positron Emission Tomography (ToF-PET) ", *Sensors*, vol. 10, pp. 10484-10505, 2010.
- [13] W. W. Moses, "Recent advances and future advances in time-of-flight PET", *Nucl. Instrum. Methods Phys. Res.*, vol. 580, pp. 919-92, 2007.
- [14] T. K. Lewellen, "Time-of-flight PET", *Sem. Nucl. Med.*, vol. 28, pp.268, 1998.

- [15] A. Nassalski; M. Moszynski; A. Syntfeld-Kazuch, et al., "Multi Pixel Photon Counters (MPPC) as an Alternative to APD in PET Applications," *IEEE Transactions on Nuclear Science*, vol. 57, pp. 1008–1014, 2010.
- [16] S. Surti, J. S. Karp, L.M. Popescu, M. E. Daube-Witherspoon, and M. Werner, "Investigation of Time-of-Flight Benefit for Fully 3-D PET", *IEEE Trans. Med. Imag.*, Vol. 25, no. 5, pp. 529-538, May 2006.
- [17] J. Karp, S. Surti, M. E. Daube-Witherspoon, and G. Muehllehner, "Benefits of time-of-flight in PET: Experimental and clinical results," *J. Nucl. Med.*, vol. 49, no. 3, pp. 462–470, Mar. 2008.
- [18] C. S. Levin and E. J. Hoffman, "Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution," *Phys. Med. Biol.*, vol. 44, pp. 781–799, 1999.
- [19] M. El-Desouki, M. Jamal Deen, Q. Fang, L. Liu, F. Tse and D. Armstrong, "CMOS image sensors for high speed applications," *Sensors*, vol. 9, pp. 430-444, 2009.
- [20] D. V. O'Connor, D. Phillips, "Time-correlated single photon counting," *Academic Pr.*, 1984.
- [21] D. Renker, "New trends on photodetectors," *Nucl. Instrum. Meth.*, vol. 571, pp. 1–6, 2007.
- [22] http://www.olson-technology.com/mr_fiber/glossary-a.htm
- [23] D. L. Snyder, "Some noise comparisons of data-collection arrays for emission tomography-systems having time-of-flight measurements," *IEEE Trans. Nucl. Sci.*, vol. NS-29, pp. 1029–1033, Feb. 1982.
- [24] W. W. Moses, "Potential uses for improved coincidence timing accuracy in PET," *J. Nucl. Med.*, vol. 43, p. 229P, May 2002.
- [25] D. Palubiak, M. M. El-Desouki, and Q. Fang, "High-Speed, Single- Photon Avalanche-Photodiode Imager for Biomedical Applications," *IEEE Sensors Journal*, Vol. 11, pp. 2401 – 2412, Oct. 2011.
- [26] S. Cova, M. Ghioni, A. Lotito, I. Rech, and F. Zappa, "Evolution and prospects for single-photon avalanche diodes and quenching circuits," *J. Mod. Opt.*, vol. 51, no. 9/10, pp. 1267–1288, 2004.

- [27] F. Zappa, S. Tisa, S. Cova, P. Maccagnani, D. B. Calia, R. Saletti, R. Roncella, G. Bonanno, and M. Belluso, "Single-photon avalanche diode arrays for fast transients and adaptive optics," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 1, pp. 365–374, Feb. 2006.
- [28] S. Tisa, F. Guerrieri, F. Zappa, "Variable-Load Quenching Circuit for Single-Photon Avalanche Diodes", *Optics Express*, Vol. 16 pp.2232-2244 (2008).
- [29] J. Kostamovaara and R. Myllylä, "Time-to-digital converter with an analog interpolation circuit," *Rev. Sci. Instrum.*, vol. 57, pp. 2880–2885, 1986.
- [30] R. Ahola, "A Pulsed Time-of-Flight Laser Rangefinder for Fast, Shortrange, High Resolution Applications", *Acta Univ. Oulu*, C 38, pp. 29-32, 39-45, 1987.
- [31] S. N. Vainshtein, V. Rossin, A. Kilpela, J. Kostamovaara, R. Myllyla, K. Määttä, "Internal Q switching in semiconductor lasers: High intensity pulses in the picosecond range and spectral peculiarities," *IEEE J. Quant. Electr.*, vol 31, 1015-1021, 1995.
- [32] H. Brockhaus and A. Glasmachers, "Single particle detector system for high resolution time measurements," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 4, pp. 707–711, Aug. 1992.
- [33] T. Otsuji, "A picosecond-accuracy, 700-MHz range, Si-bipolar time interval counter LSI," *IEEE J. Solid-State Circuits*, vol. 28, pp. 941–947, Sept. 1993.
- [34] V. P. Ladygin, P. K. Manyakov, N. M. Piskunov, "Time-of-flight trigger with digital selection of events", *NIMPRS A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Volume 357, Issues 2–3, Pages 386-390, 1995.
- [35] A. P. Heinson *et al.*, "Measurement of the branching ratio for the rare decay $K_L0 \rightarrow \mu + \mu^-$ ", *Phys. Rev.*, D 51, 985, 1995.
- [36] J. B. Rettig and L. Dobos, "Picosecond time interval measurements," *IEEE Trans. Instrum. Meas.*, vol. 44, pp. 284–287, Apr. 1995.
- [37] K. Park and J. Park, "Time-to-digital converter of very high pulse stretching ratio for digital storage oscilloscopes," *Rev. Sci. Instrum.*, vol. 70, no. 2, pp. 1568–1574, Feb. 1999.
- [38] A. Mantyniemi, T. Rahkonen, and J. Kostamovaara, "A CMOS time-to-digital converter (TDC) based on a cyclic time domain successive approximation interpolation method," *J. Solid-State Circuits*, vol. 44, no. 11, pp. 3067–3078, Nov. 2009.
- [39] G. W. Roberts, M. Ali-Bakhshian, "A brief introduction to time-to digital and digital-to-time converters," *IEEE Trans. Circuits Sys. II*, vol. 57, no. 3, pp. 153-157, 2010.

- [40] B. Markovic, A. Tosi, F. Zappa, S. Tisa, "Smart-pixel with SPAD detector and Time-to-Digital Converter for Time-Correlated Single Photon Counting," *2010 IEEE Photonics Society Annual Meeting*, pp.181-182, Nov. 2010.
- [41] S. Henzler, S. Koeppe, D. Lorenz, W. Kamp, R. Kuenemund, and D. Schmitt-Landsiedel, "A local passive time interpolation concept for variation-tolerant high-resolution time-to-digital conversion," *IEEE J. Solid-State Circuits*, vol. 43, no. 7, pp. 1666–1676, Jul. 2008.
- [42] R. B. Staszewski, S. K. Vemulapalli, P. Vallur, J. L. Wallberg, and P. T. Balsara, "1.3 V 20 ps time-to-digital converter for frequency synthesis in 90-nm CMOS," *IEEE Trans. Circuits Syst. II, Brief Papers*, vol. 53, no. 3, pp. 220–224, Mar. 2006.
- [43] <http://cdn.intechweb.org/pdfs/12913.pdf>
- [44] M. Lee and A. Abidi, "A 9b, 1.25 ps resolution coarse-fine time-to digital converter in 90 nm CMOS that amplifies a time residue," *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 769–777, Apr. 2008.
- [45] K. Matta and J. Kostamovaara, "A high-precision time-to-digital converter for pulsed time-of-flight laser radar applications," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 521–536, 1998.
- [46] R. Rankinen, K. Maatta, and J. Kostamovaara, "Time-to-digital conversion with 10 ps single shot resolution," *Proceedings 6th Mediterranean Electrotechnical Conference*, vol. 1, pp. 319 – 322, 1991.
- [47] S. Henzler, "Time-to-Digital Converters," Springer, 2010.
- [48] A. S. Yousif and J. W. Haslett, "A fine resolution TDC architecture for next generation PET imaging," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 10, pp. 1574–1582, Oct. 2007.
- [49] D. Stoppa, F. Borghetti, J. Richardson, R. Walker, L. Grant, R. Henderson, M. Gersbach, and E. Charbon, "A 32 32-pixel array with in-pixel photon counting and arrival time measurement in the analog domain," in *Proc. ESSCIRC*, pp. 204–207, 2009.
- [50] G. Hungerford and D. J. S. Birch, "Single-photon timing detectors for fluorescence lifetime spectroscopy," *Meas. Sci. Technol.*, vol. 7, pp. 121–135, 1996.
- [51] J. Kalisz, "Review of methods for time interval measurements with picosecond resolution," *Metrologia*, vol. 41, pp. 17–32, 2004.
- [52] V. S. Reinhardt, "A Review of Time Jitter and Digital Systems," in *the Proc. IEEE IFCS*, pp. 38-45, 2005.

- [53] P. Dudek, S. Szczepanski, and J. Hatfield, "A high-resolution CMOS time-to-digital converter utilizing a Vernier delay line," *IEEE J. Solid-State Circuits*, vol. 35, pp. 240–247, Feb. 2000.
- [54] P. Chen, S. I. Liu, and J. Wu, "A CMOS pulse-shrinking delay element for time interval measurement," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 47, no. 9, pp. 954–8, Sep. 2000.
- [55] J. Christiansen, "An integrated high resolution CMOS timing generator based on an array of delay locked loops," *IEEE J. Solid-State Circuits*, vol. 31, no. 7, pp. 952–957, 1996.
- [56] M. Mota and J. Christiansen, "A high-resolution time interpolator based on a delay locked loop and an RC delay line," *IEEE J. Solid-State Circuits*, vol. 34, no. 10, pp. 1360–1366, Oct. 1999.
- [57] A. Mantyniemi, T. Rahkonen, and J. Kostamovaara, "A 9-channel integrated time-to-digital converter with sub-nanosecond resolution," in *Proc. IEEE MWSCAS*, vol. 1, pp. 189–192, Aug. 1997.
- [58] N. R. Mahapatra, A. Tareen and S. V. Garimella, "Comparison and analysis of delay elements," *IEEE MWSCAS*, vol.2, pp. 473-6, 2002.
- [59] P. Chen, C. C. Chen, J. C. Zheng, and Y. S. Shen, "A PVT insensitive Vernier-based time-to-digital converter with extended input range and high accuracy," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 4, pp. 294–302, Apr. 2007.
- [60] J. P. Jansson, A. Mantyniemi, and J. Kostamovaara, "A CMOS time-to-digital converter with better than 10 ps single-shot precision," *IEEE J. Solid-State Circuits*, vol. 41, no. 6, pp. 1286–1296, Jun. 2006.
- [61] Y. Moon, J. Choi, K. Lee, D. Jeong, and M. Kim, "An all-analog multiphase delay-locked loop using a replica delay line for wide-range operation and low jitter," *IEEE J. Solid-State Circuits*, vol. 35, pp. 377–384, Mar. 2000.
- [62] R. J. Baker, "CMOS Circuit Design, Layout and Simulation, Third Edition," Wiley-IEEE, 2010.