



Universidad Nacional del Sur

Proyecto Final de Carrera
Implementación de un Nutt-TDC en FPGA.

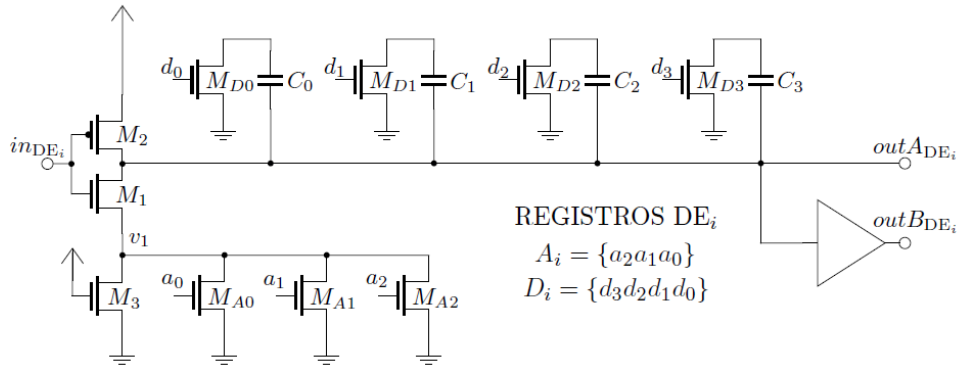
Filsinger Miqueas

Contents

1	Introducción	3
2	Estado del arte	4
2.1	Principios de la arquitectura tipo Nutt	5
2.2	Tests, mediciones y resultados	6
3	Arquitectura	8
3.1	Tapped Delay Line	8
3.2	Decoder	10
3.3	Medición gruesa y Árbitro	12
3.4	Implementación	14
	Referencias	15

Introducción

La performance de un TDC se evalúa principalmente en términos de precisión y linealidad. La precisión es la mínima unidad que el instrumento es capaz de medir, mientras que el concepto de linealidad refiere a la capacidad de mantener la respuesta entrada-salida de forma lineal. Son interesantes además algunos otros aspectos como el rango de medición, tiempo muerto, consumo de energía, y frecuencia de muestreo, sin embargo se omitirán en este trabajo.



Estado del arte

En esta sección se explicarán algunos de los últimos avances propuestos en la bibliografía. Hasta ahora se ha hablado de aquellos aspectos esperables en un TDC, pero es importante tener en cuenta que existen distintos paradigmas en su diseño dependiendo de su destino: ASIC o FPGA. La libertad de colocación de componentes (placement & routing) con la que se cuenta en el diseño de ASICs permite minimizar retardos indeseados que en el diseño de FPGA no tienen remedio, por tal, esta sección será fuertemente basada en [4] ya que propone un estudio, clasificación y análisis de las distintas arquitecturas bajo el paradigma de la construcción en FPGA, algo que no se había hecho hasta el momento.

Machado propone clasificar las distintas arquitecturas según los componentes que utiliza, de esta forma propone la taxonomía que se muestra en la Figura 2.

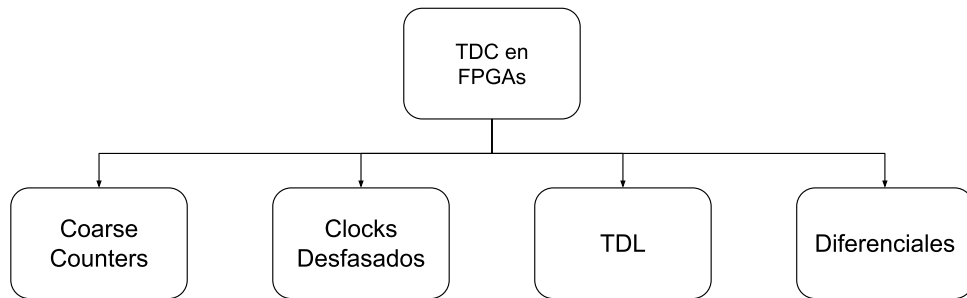


Figure 2: Clasificación de las distintas arquitecturas.

Las arquitecturas de tipo *Coarse Counter* (contador grueso), son aquellas que utilizan un simple contador para estimar el ancho de pulso de una señal, según la cantidad de clocks del sistema que entran en ella; su precisión es entonces el periodo de reloj. Las arquitecturas basadas en *Clocks Desfasados* son todas aquellas que utilizan más de un clock para estimar la medición: algunas de ellas deciden la medición a partir de saber cuál de los clocks puede observar primero la llegada del pulso, mientras que otros utilizan elementos adicionales para la interpolación. Por otro lado está tal vez una de las arquitecturas más utilizadas, siendo la que se aplica en este trabajo, y es aquella que utiliza *Tapped Delay Lines* (TDL: líneas de retardo registradas) para realizar la interpolación y generar la medición fina. Esto es, se ingresa el pulso de entrada en una cadena de elementos que introducen un retardo uniforme; al muestrear el estado de la cadena entonces se infiere cuántos retardos la señal cruzó, con lo que se estima una medición. Dentro de esta clasificación existen distintas formas de utilizar esta forma de interpolación, algunas arquitecturas extienden el rango de medición al combinar también un Coarse Counter, la topología Nutt, otras utilizan varias cadenas de retardo con el fin de mejorar la precisión realizando un promedio de ellas, y otras combinan más de una cadena y más de un clock. En cualquiera de sus formas, se clasifican según su componente principal de interpolación: la cadena de retardos. Por último, las arquitecturas *Diferenciales* se basan en medir la diferencia de retardo entre dos elementos. En estas existen principalmente dos tipos: cadenas de retardos compuestas de dos elementos de retardo distintos, y osciladores en anillo. Por otro lado si buscamos medir el intervalo temporal entre una señal y el clock entonces llamaremos al TDC *síncrono*, si el intervalo puede comenzar y terminar en cualquier momento entonces lo llamaremos *asíncrono*.

2.1 Principios de la arquitectura tipo Nutt

Las arquitecturas TDL pueden ser implementadas de distintas formas, como muestra la Figura 3.

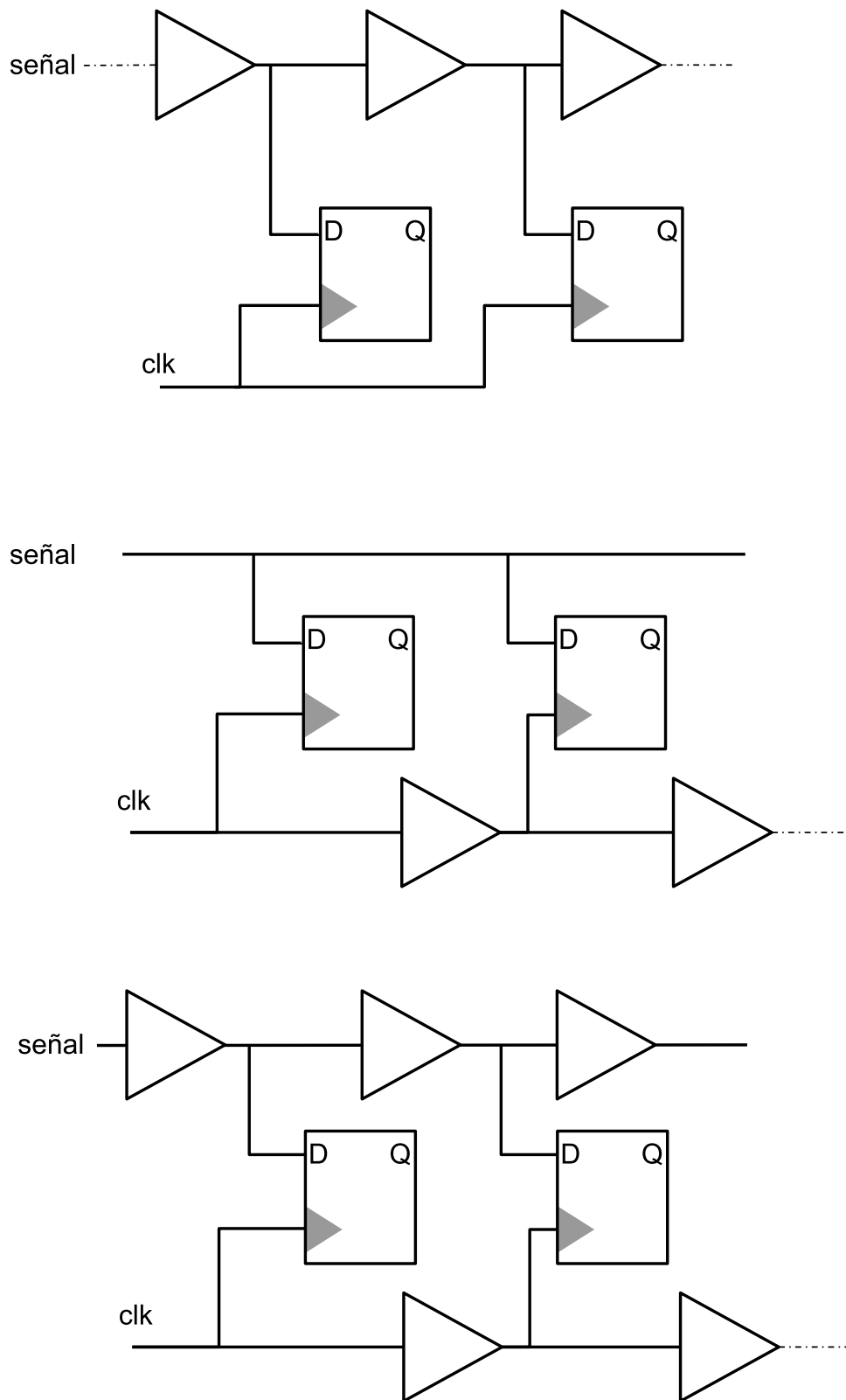


Figure 3: Diferentes formas de aplicar el método Vernier.

Basicamente la señal alimenta una cadena de retardos, la cuál es registrada a cada clock.

Cuando se detecta el pulso de Start, el valor de la cadena se guarda para luego procesarlo, y componer el resultado de parte fina, que idealmente es el espacio entre líneas rojas que muestra la Figura 4. Lo mismo se realiza cuando la señal egresa de la cadena de retardos. Finalmente el resultado se compone de una parte gruesa, y puede calcularse como

$$T_m = (n_{coarse} + 1) T + \tau_{start} - \tau_{stop}$$

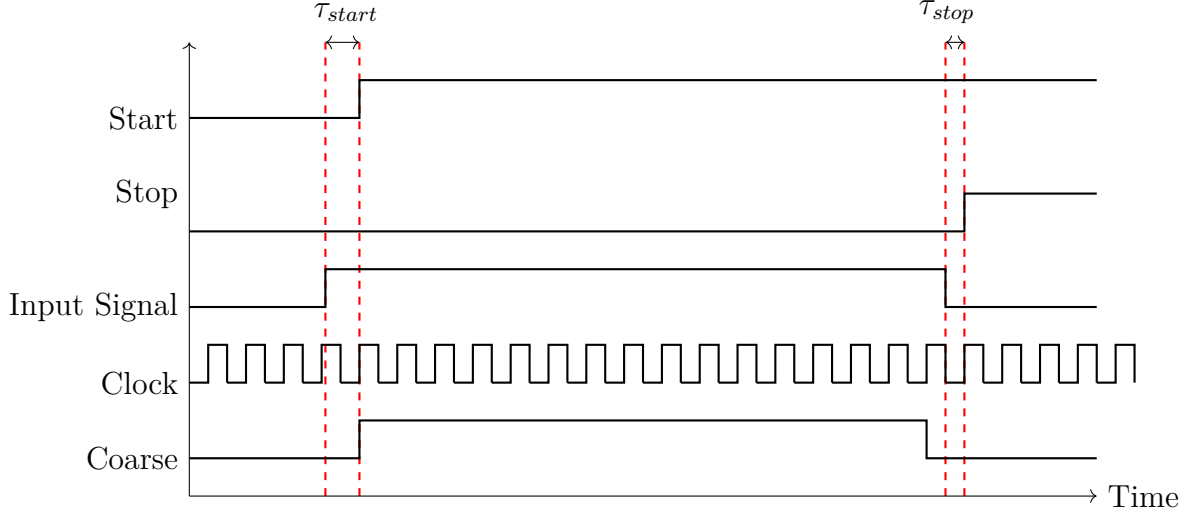


Figure 4: Esquema de medición fina del pulso de entrada.

2.2 Tests, mediciones y resultados

Particularmente en la arquitectura TDL, la opción predilecta en la bibliografía para calibrar el instrumento es mediante un “code-density test”. Esta prueba consiste en ingresar una señal decorrelacionada del reloj del sistema, de forma que excite todos los elementos de retardo por igual y por lo tanto “golpee” a cada celda por igual, generando un histograma de hits (golpes) en función del número de celda. Si la cadena tiene N elementos, y mide un periodo de reloj T , y además la i -ésima celda captura n_i mediciones, entonces es claro que el tamaño de cada retardo es

$$\tau_i = n_i \frac{T}{N}, \quad (1)$$

luego el retardo medio es

$$\bar{\tau} = \frac{T}{N}. \quad (2)$$

Con esto es posible calcular la DNL (Differential Non-Linearity) como

$$DNL_i = \tau_i - \bar{\tau}, \quad (3)$$

y con esto la INL (Integral Non-Linearity) como

$$INL_i = \sum_{i=0}^{N-1} \tau_i - \bar{\tau}. \quad (4)$$

Conociendo estas propiedades es posible generar una “tabla de calibración”. Una vez la medición final se realiza en el TDC, se corrige mediante una tabla pre-cargada en memoria, entregando un resultado calibrado al instante de medición (on-line calibration [5]). Se puede

contemplar una calibración exclusiva para el flanco *Start* y el flanco *Stop* ([6]), y corregir variaciones en las condiciones PVT ([7]). Es claro que la precisión del TDC no sólo depende de un buen diseño sino que también se ve afectado por propiedades estadísticas, por lo tanto es posible corregir limitaciones de hardware mediante una buena calibración.

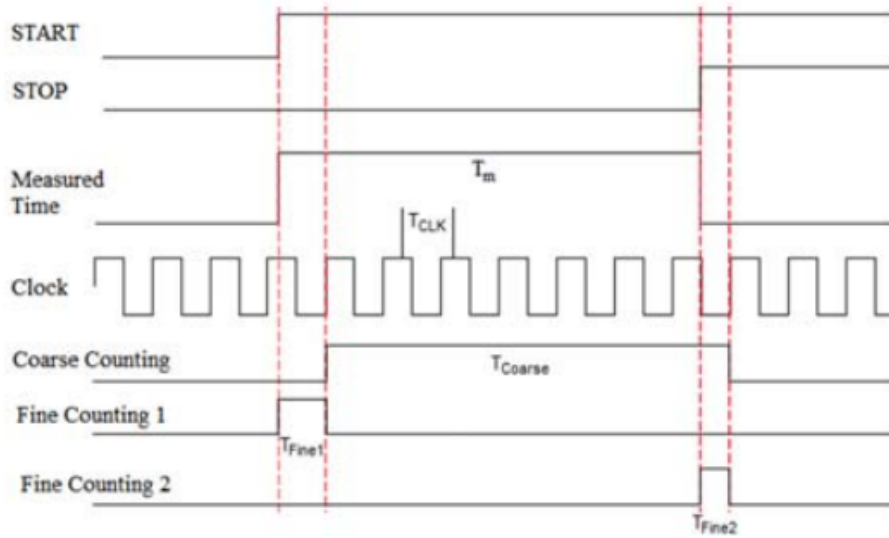


Fig. 1. Functioning principle of a generic TDC.

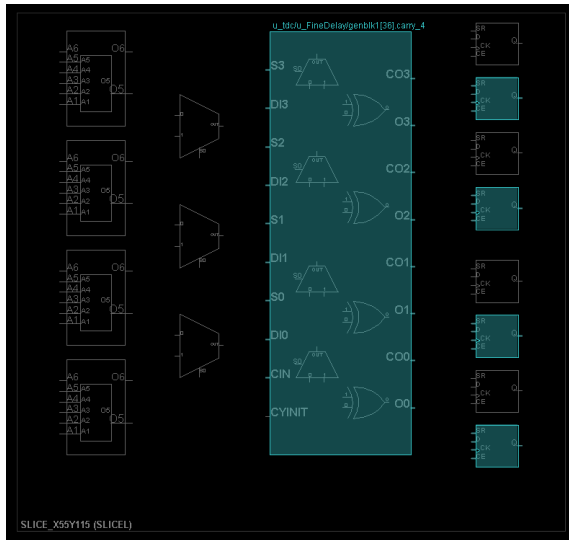
Figure 5: Caption de prueba

Arquitectura

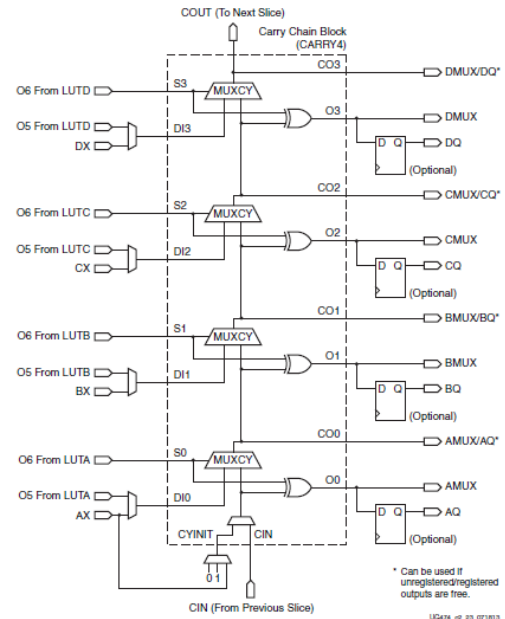
A este tipo de estructuras digitales nos gusta llamarlas *ad-hoc*, ya que desde su concepción hasta su implementación podremos ver que no se puede enmarcar dentro de una estructura típica, sino que su diseño viene estrechamente ligado a su aplicación en particular. A simple vista puede verse que los elementos que se utilizarán, las etapas de procesamiento, y las técnicas implementadas son poco comunes para lo que es el mundo del diseño digital. Se da a continuación una explicación descriptiva de cada bloque fundamental que conforma la arquitectura del *TDC*.

3.1 Tapped Delay Line

El bloque principal de este diseño, encargado de realizar la *medición fina*, es una cadena de retardos registrados. Para implementar estos retardos existen distintas pero acotadas opciones en una implementación en FPGA. En sus inicios la solución predilecta era utilizar multiplexores en cadena [8], pero para aprovechar las ventajas propias de la estructura de una FPGA se evolucionó a utilizar *Carry4s*, posiblemente utilizado por primera vez en [9]. Estos elementos forman parte de la unidad mínima de trabajo en una FPGA, y su combinación con LUTs, Flip-Flops, y Multiplexores forman lo que se conoce como un *SLICE*¹. Los *Carry4s* en su concepción fueron destinados para implementar sumadores, y tienen la ventaja de estar optimizados para tener el mínimo retardo entre compuerta y compuerta. Además cada salida cuenta con un Flip-Flop disponible inmediatamente luego, como se muestra en la Figura 6.



(a) SLICE L de una Artix-7 en Vivado 2023.2



(b) Carry 4

Figure 6: Elementos disponibles dentro de una SLICE.

Cada Carry4 cuenta con 4 salidas que pasan por una XOR, (O0-O3) y 4 salidas que pasan por cada MUXCY (CO0-CO3), además cuenta con un puerto *CIN* con el propósito de colocarlas en cadena y una entrada auxiliar *CYINIT* desde donde ingresará el pulso externo.

Como la implementación aquí presentada es asíncrona precisamos capturar la señal en su flanco

¹La Artix-7 cuenta con un total de 33650 slices.

de entrada y de salida de la cadena de retardos, por lo que luego de la colocación de la primer columna de Flip Flops (*First Flip Flops*), como se ve en 6, se utilizan otras dos columnas de Flip Flops llamados *Start Flip Flops* y *Stop Flip Flops*. Estos se encargan de registrar la primer columna solamente cuando un flanco de subida o de bajada es detectado, a través de su *Clock Enable*. Así se genera la estructura que se ve a continuación en la Figura 7.

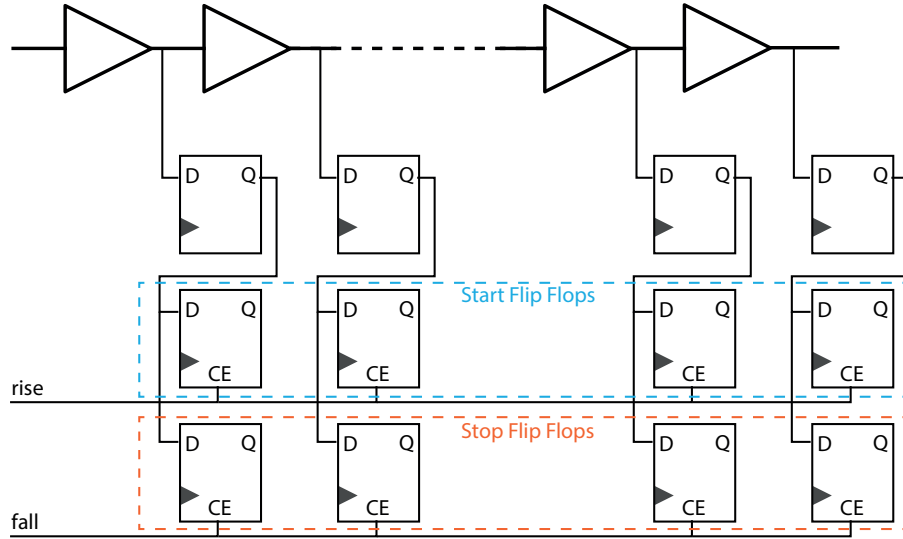


Figure 7: Esquema de la cadena del módulo de interpolación fina.

De esta estructura surgen dos aspectos importantes:

- 1) Es sabido que la transferencia de información analoga a digital puede producir metastabilidades si no se realiza correctamente. La configuración de doble flip flop para registrar los elementos de retardo soluciona este problema introduciendo lo que se conoce como “sincronizador”.
Esto es, si el tiempo de establecimiento de la señal externa es menor al tiempo de establecimiento requerido por un sólo flip flop, este puede obtener un valor desconocido a la salida. Sin embargo introduciendo un segundo flip flop en serie, este no sufre de este problema, ya que para el próximo clock la señal es estable, y esta indeterminación se resuelve decantando hacia un estado. Además la técnica de doble registro reduce los problemas de burbujas en el resultado [10].
- 2) Una propiedad intrínseca del diseño tipo Nutt es que el tiempo del pulso a medir atraviese la cadena en un tiempo no menor al de un clock del sistema. Esta es una condición de diseño y se aprovecha en múltiples bloques, pero en particular aquí se puede ver que el registro de *Start* y *Stop* dependen de un flanco de subida o bajada para capturar la cadena retardos, el cuál tiene una duración de un clock.

Por otro lado, es importante prestar especial atención al routeo y colocación de cada elemento en el bitstream final, pues es preponderante en la calidad de los resultados del TDC. Si los retardos introducidos por el routeo son uniformes para cada elemento de retardo entonces podemos asegurar que la calibración es capaz de corregir en gran parte estos errores. Según [10] es importante colocar la cadena de retardos en las Slices que quedan a continuación de los Clock Buffers, y que los flip flops sean colocados también cerca de la cadena. Esto fue probado empíricamente y se obtuvieron mejores resultados cuando los extremos de la cadena de retardos quedaban simétricos respecto de los clock buffers.

Experimentalmente se observó que cuando el extremo inferior de la cadena es lejana a los buffers,

se pueden observar huecos en el histograma resultante, mientras que al colocarlo simétrico, este espacio se reparte al inicio y al final de la cadena, como se ilustra en la Figura 8.

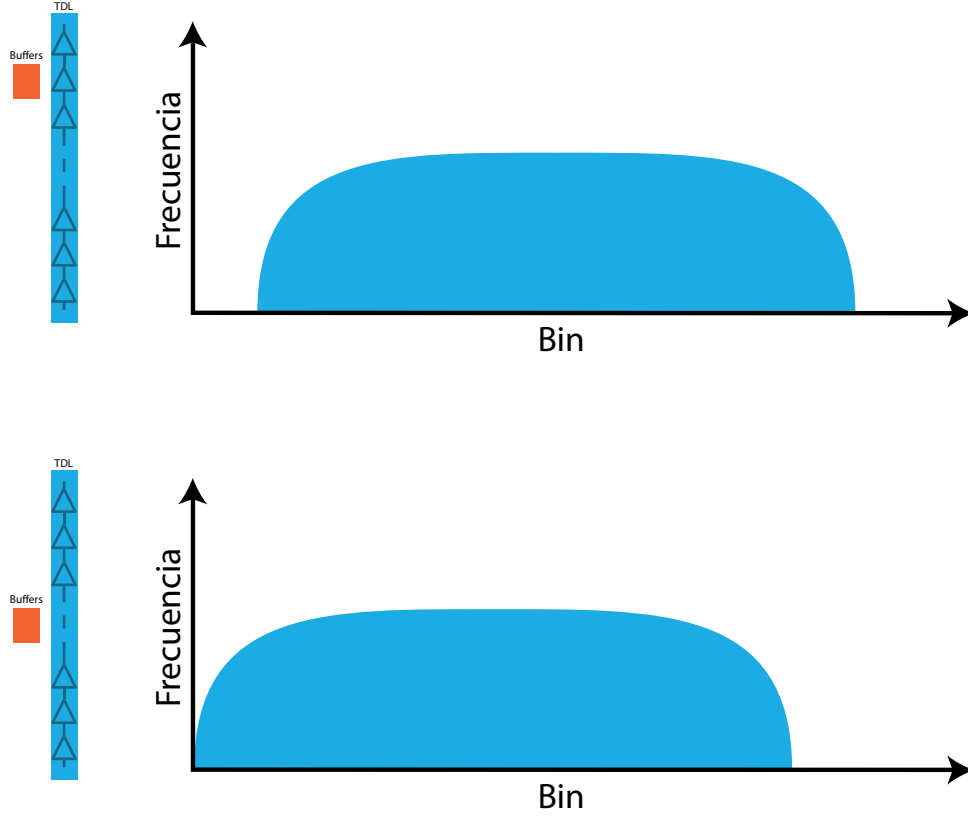


Figure 8: Ejemplo de histograma conseguidos al excitar el TDC con una frecuencia pseudo-aleatoria modificando la posición relativa de la cadena de retardos respecto a los buffers de clock.

Este fenómeno tiene una explicación, y es que si la cadena de buffers mide temporalmente un clock, entonces la posición de los buffers de clock cambia la sincronía de los Flip-Flops si no se realiza con cuidado. Los Flip-Flops extremos son los más susceptibles a tener un mayor desfase que el resto, debido al delay de routeo. Por lo tanto al poner los buffers en una posición equidistante logramos compensar ese delay a ambos extremos, mientras que en el ejemplo de la Figura 8 existe un retardo mayor para el *Carry₄* de entrada.

3.2 Decoder

La propagación de un pulso por la línea de retardos naturalmente genera un código unario sobre los registros. Esto es, a medida que el flanco descendente progresa sobre la línea de retardos deja a su paso un estado alto en cada Flip-Flop de la cadena, mientras que cuando egresa un flanco ascendente deja ceros a su paso, tal como muestra la Figura 9.

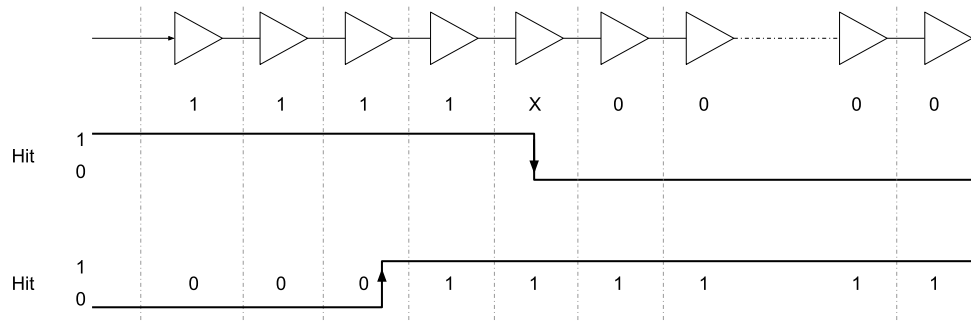


Figure 9: Registro de un pulso digital en el momento que ingresa a la línea, y en el momento que egresa de ella.

La misión principal del decoder es traducir estos valores leídos de los registros *Start* y *Stop* a datos procesables. Aquí se presenta un gran inconveniente presente en todas las soluciones propuestas por bibliografía, pues debido a las variabilidades propias de la arquitectura de una FPGA y de los órdenes de precisión que se pretenden es muy frecuente encontrar burbujas en los registros resultantes, imposibilitando la utilización de decodificadores unarios convencionales. Esto significa que en los registros *Start* o *Stop* se pueden encontrar registros con errores del tipo “11..111001000...”, donde no es posible distinguir una transición de estados clara. Según la bibliografía ([4]) esto se debe principalmente al efecto de *clock skew* cada vez más presente en las últimas tecnologías, gracias al decrecimiento abrupto de tamaño de fabricación. Es posible que la diferencia temporal del flanco de clock (el clock skew) entre dos Flip Flops que registran elementos de retardo adyacentes sea mayor al tiempo de retardo de la celda a registrar, y por lo tanto se produce una burbuja en el registro. Esto se ejemplifica en la Figura 10. Si el Flip-Flop número uno captura la cadena en el momento t_1 y el Flip-Flop número dos lo hace un momento después en t_2 , es posible entonces que el primero decante su estado en un valor 0, mientras que el segundo decante su estado en un valor 1. La burbuja ocurre entonces si $\Delta = t_2 - t_1 \geq \tau$ siendo τ el tiempo de una celda de retardo, y Δ el clock skew.

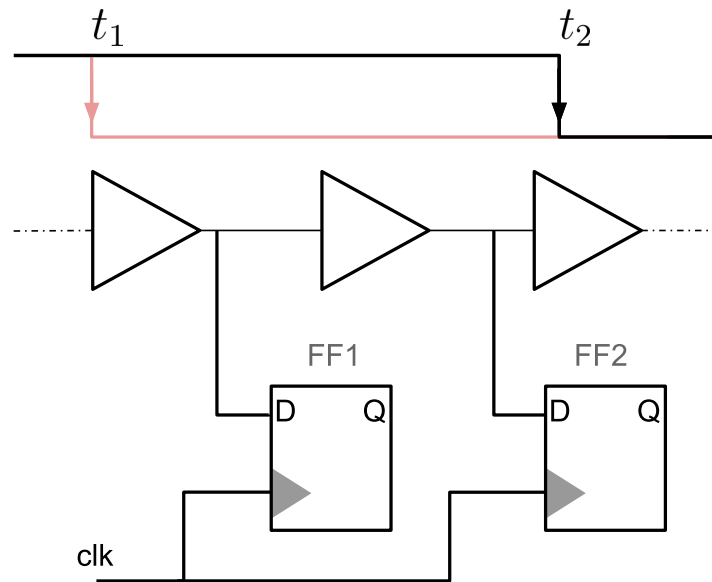


Figure 10: Retardo virtual de propagación que se puede observar cuando existe *clock skew* entre dos flip-flops aledaños.

Para implementar un decodificador se han propuesto distintas soluciones, en esta sección resumiremos las ideas más usadas:

- 1) **Contador de unos (ones counter):** se propuso como una solución en el diseño de ADCs, y luego en [11] se utiliza por primera vez para el diseño de un TDC. La implementación consiste básicamente en contar la cantidad de estados altos en el registro para definir la posición aproximada del flanco. Si se puede asegurar que la cantidad de burbujas es despreciable entonces la respuesta del decodificador se acerca a la real. Esto es beneficioso porque se prescinde del ordenamiento de un registro, es decir la entrada “111100” se decodificaría en la misma respuesta que la entrada “111010”, por lo tanto tiene la habilidad natural de corregir errores de burbuja.
- 2) **Buscador de flancos:** Consiste en realizar un circuito combinacional que busque secuencialmente (look-for) transiciones de estados del tipo “0-1” o “1-0” en el registro de entrada y devuelva la posición de este dentro la cadena, e.g para un registro “0001111111” el decoder debe devolver tres, ya que es la posición donde se detecta la transición. Para eliminar el problema de burbujas se recurre a aumentar el tamaño de la secuencia buscada, es decir si se tiene un registro con burbujas del tipo “0001001111” podemos buscar la secuencia “0001” de forma que únicamente la primer transición es detectada y la burbuja se ignora ([12]). Rápidamente se hace evidente que a medida que el tamaño del filtro incrementa es capaz de corregir burbujas de mayor tamaño, a costa de descartar una mayor cantidad de muestras al principio y final de la cadena.
- 3) **Realineamiento de taps:** Una vez detectados los lugares del registro donde se producen burbujas, se intercambia la posición de salida de estos flip-flops con el de uno cercano ([5]), luego se utiliza un decoder del tipo “buscador de flancos”. Para su implementación primero se debe hacer un “code-density test” que nos permita conocer cuáles son los bins de tamaño cero (burbujas), para luego realizar el reordenamiento de taps. Esta implementación depende fuertemente de las condiciones PVT¹ en las que se utiliza la FPGA.

3.3 Medición gruesa y Árbitro

Una de las ventajas de la implementación tipo Nutt es que nos permite mantener un buen rango de medición sin perder precisión. Esto se logra separando la implementación en dos tareas, la medición gruesa y la medición fina, para esto se utiliza un contador grueso (coarse counter) con el que se mide la cantidad de periodos de clocks que tarda la señal en atravesar la cadena. Aunque esta propuesta parece sencilla precisa de una sincronización previo al procesamiento del resultado, ya que si la señal externa entra en sincronía con el clock del sistema es posible que el contador cuente o no un periodo de reloj menos, perdiendo una precisión de un periodo completo. En [10] se propone la utilización de más de un contador grueso como árbitros para decidir el resultado final del cálculo, el desarrollo es el siguiente:

Se implementaron tres contadores gruesos, cada uno alimentado por un clock derivado del clock del sistema y ligeramente desfasado, como se muestra en la Figura 11.

¹Proceso, voltaje y temperatura.

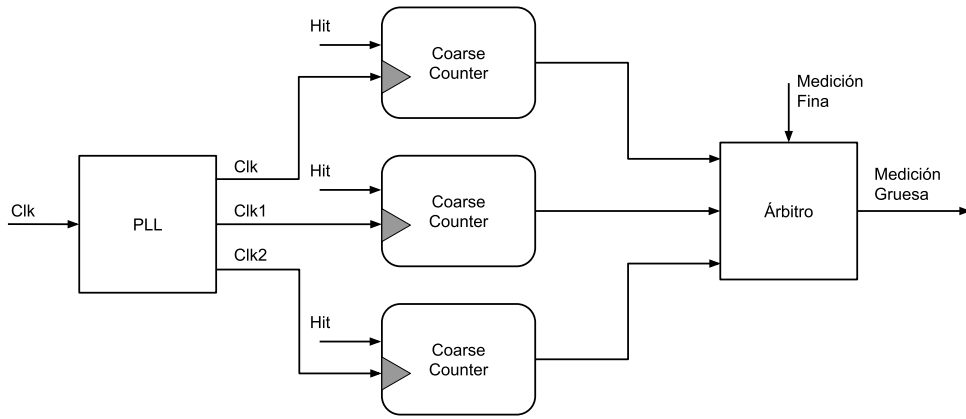


Figure 11: Árbitro implementado.

Estos tres clocks inducen cuatro casos a diferenciar, estos se plantearan para el flanco de entrada pero puede hacerse de la misma forma para el flanco de salida.

El caso base es suponer que el pulso llega antes del flanco de subida de todos los clocks, como muestra el Caso I de la Figura 12, o también llegar entre cada par de flancos como muestran los Casos II, III y IV. Para el primer caso, la línea de retardos captura la señal al inicio de la cadena pues el clock ocurre cerca de la llegada del pulso, sin embargo para el resto de casos donde el flanco de clock ya ocurrió es necesario esperar un periodo al próximo flanco de clock, en consecuencia se captura el flanco de entrada al borde de saturar la cadena de retardos, como muestra la Figura 13.

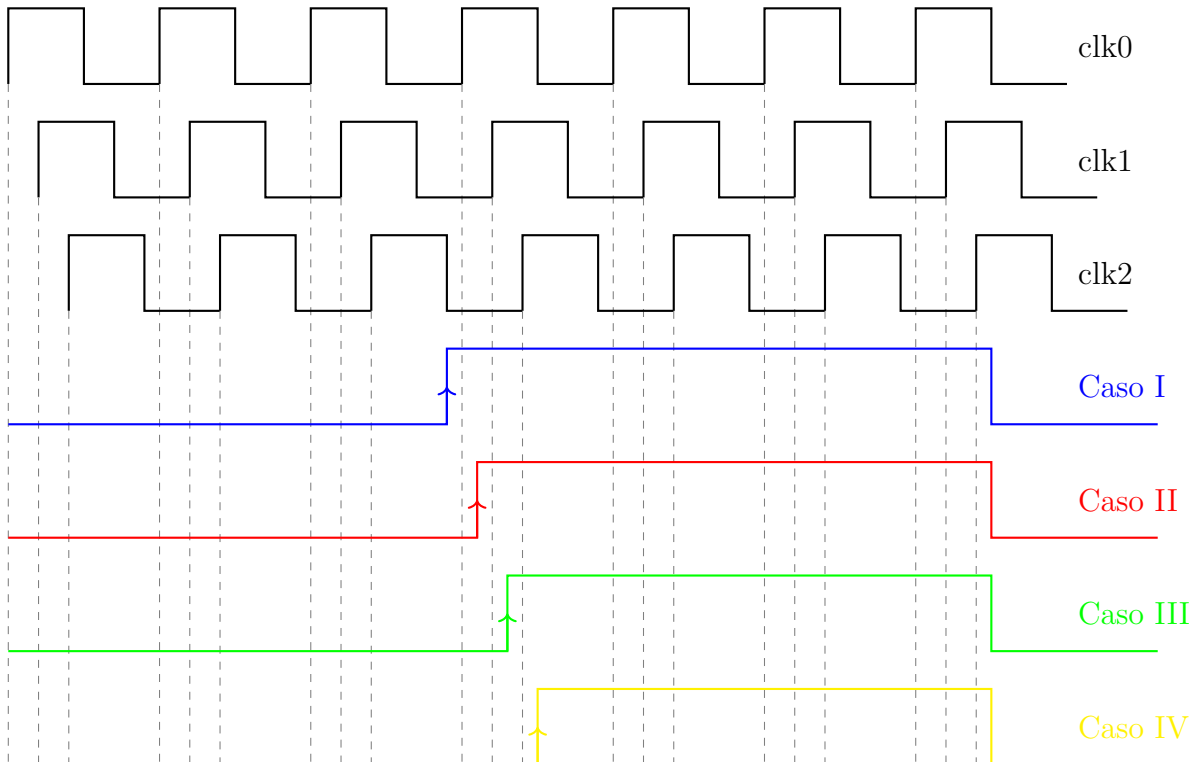


Figure 12: Todos los casos posibles de arbitraje para el flanco Start

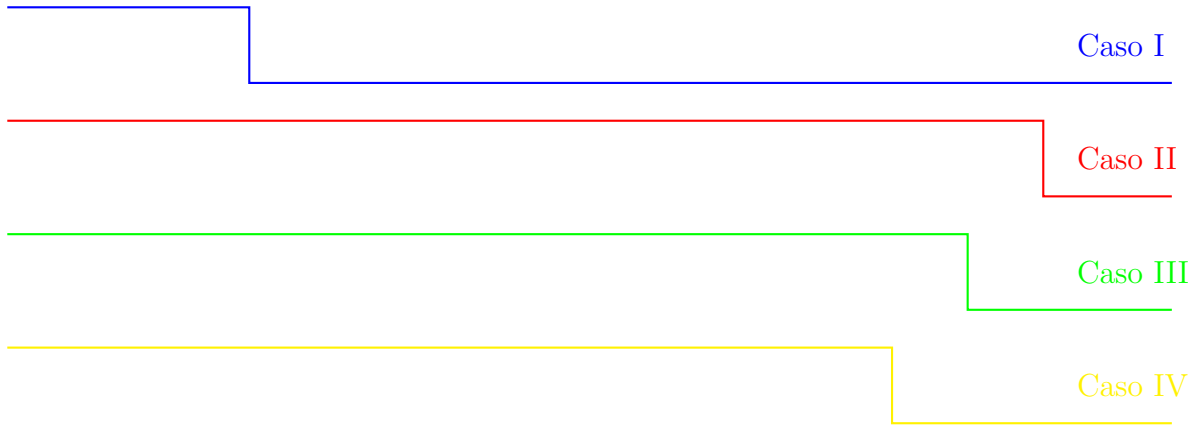


Figure 13: Registro *Start* de la cadena de retardos para cada caso de arbitraje.

Si analizamos cada caso con su coarse counter entonces obtendríamos los resultados que se presentan en la Tabla 1. Rápidamente nos podemos dar cuenta que si todos los Coarse Counters tienen la misma respuesta entonces el resultado de cualquiera de ellos es correcto, pero si no son iguales y el pulso es detectado en el extremo de la cadena, sabemos con certeza que la cuenta correcta es de tres periodos, pues el periodo incompleto está codificado en la respuesta fina. Es entonces posible conociendo las cuatro posibles condiciones de cuentas gruesas generar una respuesta apoyada en cuatro referencias: tres contadores y la respuesta fina.

Coarse counter según:	clk0	clk1	clk2
Caso I	4	4	4
Caso II	3	4	4
Caso III	3	3	4
Caso IV	3	3	3

Table 1: Cuenta gruesa del ejemplo presentado en la Figura 12, según cada clock generado por el PLL.

3.4 Implementación

Finalmente se implementó la arquitectura presentada en [10] que se repite en la Figura 14.

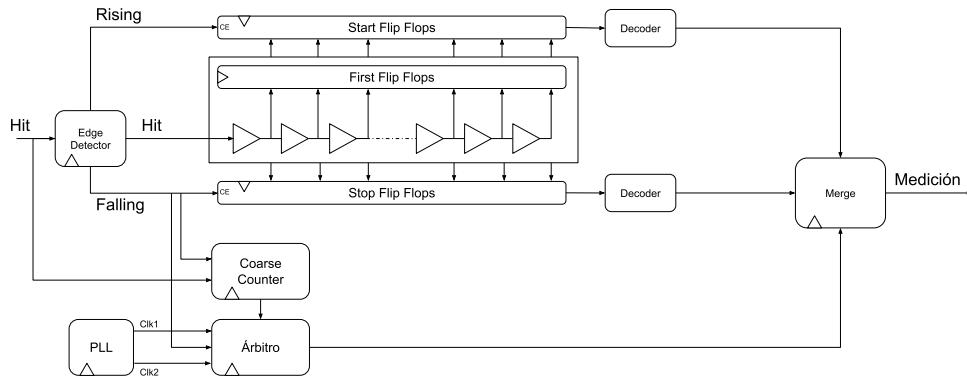


Figure 14: Arquitectura del TDC. Se ignora el routeo del clock.

Referencias

- [1] Samuele Altruda et al. “PicoTDC: a flexible 64 channel TDC with picosecond resolution”. In: *Journal of Instrumentation* 18.07 (July 2023), P07012. ISSN: 1748-0221. DOI: 10.1088/1748-0221/18/07/p07012. URL: <http://dx.doi.org/10.1088/1748-0221/18/07/P07012>.
- [2] K. Maatta and J. Kostamovaara. “A high-precision time-to-digital converter for pulsed time-of-flight laser radar applications”. In: *IEEE Transactions on Instrumentation and Measurement* 47.2 (Apr. 1998), pp. 521–536. ISSN: 0018-9456. DOI: 10.1109/19.744201. URL: <http://dx.doi.org/10.1109/19.744201>.
- [3] Juan Ignacio Morales. “Diseño de sistemas microelectrónicos basados en alta resolución temporal”. PhD thesis. Bahía Blanca: Universidad del Sur, 2020.
- [4] Rui Machado, Jorge Cabral, and Filipe Serra Alves. “Recent Developments and Challenges in FPGA-Based Time-to-Digital Converters”. In: *IEEE Transactions on Instrumentation and Measurement* 68.11 (2019), pp. 4205–4221. DOI: 10.1109/TIM.2019.2938436.
- [5] Chong Liu and Yonggang Wang. “A 128-Channel, 710 M Samples/Second, and Less Than 10 ps RMS Resolution Time-to-Digital Converter Implemented in a Kintex-7 FPGA”. In: *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 773–783. ISSN: 1558-1578. DOI: 10.1109/tns.2015.2421319. URL: <http://dx.doi.org/10.1109/TNS.2015.2421319>.
- [6] Wassim Khaddour, Wilfried Uhring, Foudil Dadouche, Norbert Dumas, and Morgan Madec. “Calibration Methods for Time-to-Digital Converters”. In: *Sensors* 23.5 (Mar. 2023), p. 2791. ISSN: 1424-8220. DOI: 10.3390/s23052791. URL: <http://dx.doi.org/10.3390/s23052791>.
- [7] X. Qin et al. “A 1.15-ps Bin Size and 3.5-ps Single-Shot Precision Time-to-Digital Converter With On-Board Offset Correction in an FPGA”. In: *IEEE Transactions on Nuclear Science* 64.12 (2017), pp. 2951–2957. DOI: 10.1109/TNS.2017.2768082.
- [8] J. Kalisz, R. Szplet, J. Pasierbinski, and A. Poniecki. “Field-programmable-gate-array-based time-to-digital converter with 200-ps resolution”. In: *IEEE Transactions on Instrumentation and Measurement* 46.1 (Feb. 1997). Conference Name: IEEE Transactions on Instrumentation and Measurement, pp. 51–55. ISSN: 1557-9662. DOI: 10.1109/19.552156. URL: <https://ieeexplore.ieee.org/document/552156/citations#citations> (visited on 07/16/2024).
- [9] Claudio Favi and Edoardo Charbon. “A 17ps time-to-digital converter implemented in 65nm FPGA technology”. In: *Proceedings of the ACM/SIGDA international symposium on Field programmable gate arrays*. FPGA '09. New York, NY, USA: Association for Computing Machinery, Feb. 2009, pp. 113–120. ISBN: 978-1-60558-410-2. DOI: 10.1145/1508128.1508145. URL: <https://doi.org/10.1145/1508128.1508145> (visited on 07/15/2024).
- [10] Rui Machado, Luis A. Rocha, and Jorge Cabral. “A novel synchronizer for a 17.9ps Nutt Time-to-Digital Converter implemented on FPGA”. In: *2018 AEIT International Annual Conference*. Oct. 2018, pp. 1–6. DOI: 10.23919/AEIT.2018.8577365. URL: <https://ieeexplore.ieee.org/document/8577365> (visited on 07/16/2024).

- [11] Yonggang Wang, Jie Kuang, Chong Liu, and Qiang Cao. “A 3.9-ps RMS Precision Time-to-Digital Converter Using Ones-Counter Encoding Scheme in a Kintex-7 FPGA”. In: *IEEE Transactions on Nuclear Science* 64.10 (Oct. 2017). Conference Name: IEEE Transactions on Nuclear Science, pp. 2713–2718. ISSN: 1558-1578. DOI: 10.1109/TNS.2017.2746626. URL: <https://ieeexplore.ieee.org/document/8022888> (visited on 07/25/2024).
- [12] Jinyuan Wu. “Several Key Issues on Implementing Delay Line Based TDCs Using FPGAs”. In: *IEEE Transactions on Nuclear Science* 57.3 (June 2010), pp. 1543–1548. ISSN: 1558-1578. DOI: 10.1109/tns.2010.2045901. URL: <http://dx.doi.org/10.1109/TNS.2010.2045901>.