

# Estimación de Capacidad en FPGAs Comerciales

Elías Todorovich<sup>1</sup> y Nelson Acosta

INCA-INTIA, Facultad de Ciencias Exactas

Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA)

Campus universitario, Paraje Arroyo Seco, 7000 TANDIL, Argentina

{ etodorov, nacosta}@exa.unicen.edu.ar

Palabras clave: FPGA, Estimación de capacidad, Consumo, Diseño digital

## Resumen

La estimación de consumo medio en circuitos CMOS es un problema aun no resuelto completamente debido a la dificultad en la estimación de la actividad de conmutación. Cuando se trabaja con FPGAs comerciales, a este problema se añade la falta de información sobre las capacidades físicas de los nodos del circuito, que se necesita también para estimar el consumo de potencia. En este trabajo se presenta una técnica general para estimar la capacidad de los nodos en un diseño implementado en una FPGA comercial basada en la relación entre capacidad, retardo medio y *fan out* del nodo. Los resultados preliminares indican que esta solución para estimar capacidades físicas es aplicable con un error tolerable.

## Introducción

Para calcular el consumo de potencia se debe multiplicar para cada nodo del circuito bajo estudio, la capacidad y la actividad.

$$P = 0.5 V_{dd}^2 f_{clk} \sum_{nodos} C E(sw) \quad (1)$$

$C$  es la capacidad física del nodo,  $V_{dd}$  es el voltaje de alimentación,  $E(wd)$  es el número medio de transiciones por ciclo de reloj, y  $f_{clk}$  es la frecuencia de reloj.

Entonces debe conocerse la capacidad de cada nodo: puertas más pistas. Hay dos situaciones en las que obtener esta información se hace difícil. En primer lugar, en etapas de diseño tempranas los módulos el diseño no tiene correspondencia con una tecnología y no se sabe como será el emplazamiento ni el rutado. Es más: no se conoce como se va a mapear el diseño especificado a alto nivel ni que pistas van a aparecer en la lista de conexiones del diseño físico. Este problema se notará CE-HL<sup>2</sup>. En segundo lugar, puede tenerse un diseño emplazado y rutado, pero los fabricantes de dispositivos no proveen información de las capacidades de las pistas ni los datos que permitan calcularla con exactitud. Esto es típico en diseño con FPGAs comerciales. Este problema se notará CE-PL<sup>3</sup>. La gran mayoría de la literatura encontrada trata el problema CE-HL. Sin embargo hay algún trabajo que hace su aportación al problema CE-HL en FPGAs, incluso para los dispositivos comerciales.

La capacidad de las pistas depende del área de las mismas(ver [Rab96]), o sea largo por ancho de la pista. El ancho es una decisión tecnológica pero el largo depende del emplazamiento y rutado de los

---

<sup>1</sup> Becario CONICET

<sup>2</sup> Capacitance Estimation for a High Level description

<sup>3</sup> Capacitance Estimation for a Physical Level description

bloques lógicos. Para el problema CE-HL se debería predecir el resultado de un futuro emplazamiento y rutado.

### Trabajos Relacionados

Varios trabajos tratan el problema CE-HL. Los basados en la Regla de Rent u otras observaciones empíricas reportan errores mayor al 50%. Por ejemplo [Don79] encontró un límite superior para el largo medio de interconexión en emplazamientos bidimensionales. Sin embargo este límite superior es dos veces el valor obtenido experimentalmente lo cual no da la precisión requerida por la mayoría de las aplicaciones. En [Str96] se continúa este trabajo y se proponen mejoras, superando limitaciones en la aplicación de la regla de Rent tomando en cuenta características de los procesos de emplazamiento. La regla de Rent se enuncia en el Apéndice 1.

Hay métodos más precisos para estimar la capacidad media de pistas. En [Lan96] se usa como regla empírica para estimar el largo medio de pistas en bloques compuestos, la siguiente ecuación:

$$L = k \frac{\sqrt{A}}{3} \quad (2)$$

Donde

$A$  es el área que ocupa el módulo en el chip y

$k$  es el cociente de calidad entre un buen emplazamiento y otro aleatorio

Para bloques del camino de datos (*datapath*) el largo medio de pistas ya no depende del área sino del largo del mismo:

$$L = k \frac{L_{datapath}}{3} \quad (3)$$

Los errores reportados en [Lan96] son diferencias con resultados simulados y no medidos y se resumen en la Tabla 1:

Estimación largo medio de pista	% error
bloque compuesto	7
Reloj	22
en el datapath	17

Tabla 1.

Las capacidades se calculan finalmente multiplicando estos largos medios de pista por la capacidad por unidad de largo. Esta última se puede estimar con precisión si se conoce la tecnología destino y las reglas de diseño.

En [Kat95] se hace una estimación de consumo de potencia de varios diseños especificados a nivel comportamental, para sintetizar luego la alternativa de menor consumo que cumpla además con las restricciones de área y velocidad. En este trabajo se describe todo el proceso de síntesis donde se obtiene una especificación a nivel RTL en VHDL a partir de la dada a nivel comportamental, pero el énfasis del artículo está en la estimación de la **capacidad total** conmutada. Por un lado se estima, mediante simulación a nivel de puertas, la capacidad conmutada por vector de entrada para cada módulo de librería. La capacidad de las conexiones se estima partiendo de una cota superior de la cantidad de interconexión. Los valores de capacidad estimada se comparan con los resultados de

simulación de cada diseño sintetizado. Se informa de un error medio de 12 % para un diseño completo. Se dan valores de error para la ruta de datos y la unidad de control pero no se dan valores para la estimación de capacidad de conexiones en forma separada del total del diseño.

Las medidas de largo **total** o **largo medio de pistas** son útiles para los algoritmos de emplazamiento y rutado o para estimación de área pero no son adecuadas para estimación de capacidad para estimación de potencia pues como se dijo más arriba la potencia se calcula multiplicando actividad por capacidad en cada nodo.

Una mejor técnica de estimación de capacidad orientada a la estimación de consumo debería hacerse para el **largo de cada pista** siendo sensible al diseño actual, y predecir las desviaciones con los modelos de largos de pistas distribuidos uniformemente. [Alv97] es la única referencia encontrada donde se estima la capacidad de pistas individualmente. Se presenta una técnica para la estimación de capacidad de pistas que serviría también cuando no se tiene información de emplazamiento y rutado. La técnica requiere como entradas una estimación gruesa del área de los bloques lógicos a ser emplazados, y la lista de conexiones de estos bloques. La heurística utilizada se basa en la vecindad de las células: Usa la topología de conexiones para predecir el posible emplazamiento y estima la distancia de cada celda a las que se conecta de a una. Finalmente usa la distancia rectilínea entre cada par de células (Manhattan), y un árbol de spanning mínimo para determinar el largo de las pistas con más de dos conexiones. Las comparaciones con programas de emplazamiento y rutado comerciales reportan errores inferiores al 12% en el largo medio.

Para el problema CE-PL en el caso de las FPGAs, se usarían las ecuaciones aceptadas para el cálculo de capacidad, al igual que en el nivel físico de un ASIC. Pero esto es posible solamente cuando se tienen todos los datos requeridos por las ecuaciones. Tal es el caso para los investigadores que trabajan con modelos de académicos FPGAs. Sin embargo, cuando se quiere estimar la capacidad de los nodos de un diseño que va a ser implementado en un dispositivo comercial, no se tiene prácticamente ninguna información que permita calcular la capacidad por medio de las ecuaciones mencionadas. Ni siquiera se provee la capacidad calculada por las herramientas de síntesis. El objetivo de este trabajo es estimar la capacidad de los nodos dentro de una FPGA comercial, dados la demora media de las pistas y el *fan out* de cada nodo, que es información que sí proveen las herramientas de síntesis de los fabricantes.

En [Osm00] se implementa una técnica de estimación de consumo y se aplica a FPGAs de Xilinx. La estimación de capacidades es los nodos se basa en la resolución de un sistema de ecuaciones que encuentra la capacidad de cada pista de acuerdo a la distancia Manhattan entre los CLBs que conecta. Sin embargo esta técnica, para ponerse a punto, requiere un número de mediciones físicas de la potencia total consumida del orden  $2N$ , siendo  $N$  la cantidad de CLBs por fila o columna del arreglo de bloques configurables de la FPGA.

### **Cálculo de la Capacidad de una Pista**

En el presente trabajo se estudia la relación entre la capacidad de cada pista y su retardo medio para distintos valores de *fan out*. Para el caso de las pistas simples (Fig. 1), se hace variar el largo y por ende el retardo de las pistas conectando inversores en serie.

La capacidad de cada pista se calcula, en el caso de la fig. 1-a, desarrollando la ecuación (1) para el circuito tal como se materializa en una FPGA, como se muestra en la Fig. 2.

$$P = V_{dd}^2 (C_I f_I + C_N f_N + C_A f_A + C_B f_B + C_O f_O + C_{CK} f_{CK} + C_{EN} f_{EN})$$

Como se explica en [Boe96], la potencia de los pines de salida,  $V_{dd}^2 C_{Ofo}$ , y la potencia de reloj,  $V_{dd}^2 C_{CK} f_{CK}$ , se pueden medir y aislar de la potencia total  $P$ . La habilitación,  $EN$ , se hace una sola vez según se quiera medir la potencia total o la potencia interna del chip. Por lo tanto  $f_{EN}$  es cero y anula este término. Si además se aplica a las entradas la misma señal que para el reloj. Entonces:

$$P - P_O - P_{CK} = V_{dd}^2 f (C_I + C_N + C_A + C_B)$$

$$\frac{P - P_O - P_{CK}}{V_{dd}^2 f} = C_I + C_N + C_A + C_B$$

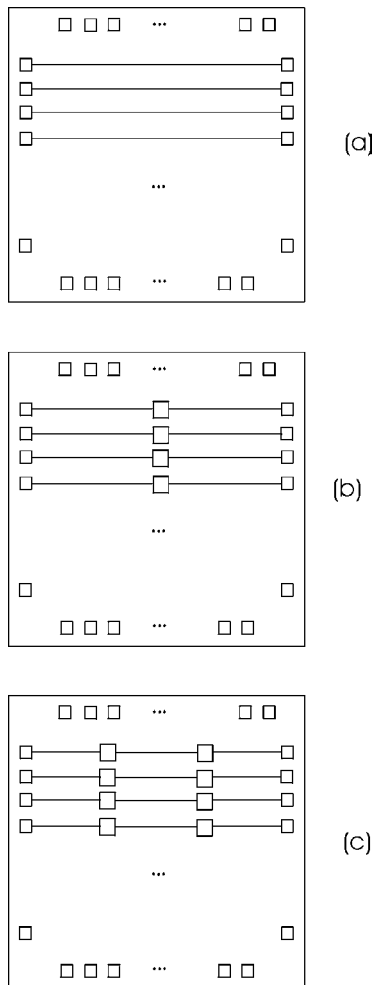


Fig. 1: Pistas con una conexión de entrada y una de salida

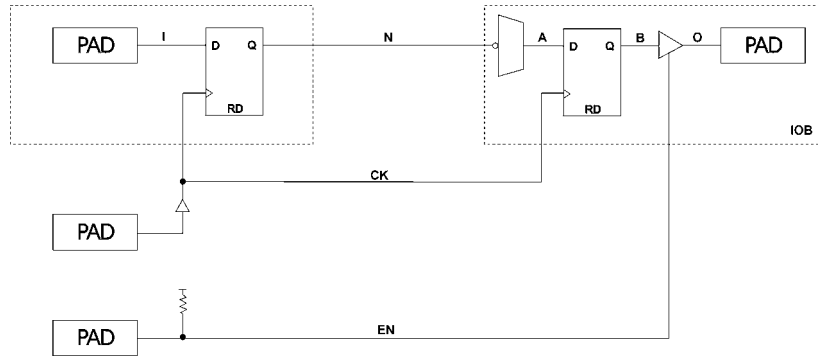


Fig. 2: Materialización de la fig. 1-a en una FPGA.

Se supone que la capacidad del nodo  $N$ , es mucho mayor que la de  $I$ ,  $A$  y  $B$ . De esta manera se propone la siguiente inecuación:

$$C_N < \frac{P - P_O - P_{CK}}{V_{dd}^2 f}$$

Como se ve en la fig. 1-a, este circuito se replica  $n$  veces en la FPGA para de esta manera reducir el error de medición.

$$\frac{P - P_O - P_{CK}}{V_{dd}^2 f} = n(\overline{C_I} + \overline{C_N} + \overline{C_A} + \overline{C_B})$$

Donde aparecen los promedios de las capacidades, entonces:

$$\overline{C_N} < \frac{P - P_O - P_{CK}}{nV_{dd}^2 f} \quad (4)$$

Para tener pistas simples más cortas en la FPGA, como aparece en la fig. 1-b, se puede implementar un circuito con un solo inversor pero en un CLB, y no en un IOB como en la fig. 2. Tal implementación se muestra en la fig. 3.

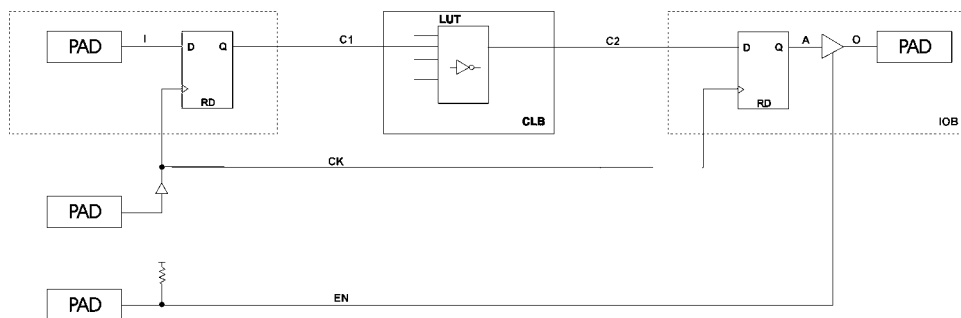


Fig. 3: Materialización de la fig. 1-b en una FPGA.

En el caso de la fig. 3 se tiene:

$$\frac{P - P_O - P_{CK}}{V_{dd}^2 f} = C_I + C_1 + C_2 + C_A$$

Usando las herramientas adecuadas (*floorplaner*), se puede hacer  $C_1 \approx C_2$ . Con el mismo razonamiento que antes se obtiene:

$$2C_1 < \frac{P - P_O - P_{CK}}{V_{dd}^2 f} \text{ y } C_1 < \frac{P - P_O - P_{CK}}{2V_{dd}^2 f}$$

Replicando  $n$  veces el circuito de la fig. 3, se llega a:

$$\bar{C}_1 < \frac{P - P_O - P_{CK}}{2nV_{dd}^2 f} \quad (5)$$

En general, para una cadena de  $s-1$  inversores con  $s$  pistas, replicada  $n$  veces en la FPGA y tratando que los retardos de estas pistas estén ecualizados, se llega a:

$$\bar{C}_1 < \frac{P - P_O - P_{CK}}{snV_{dd}^2 f} \quad (6)$$

### Fan Out 2

Se estudia la capacidad de pistas con *fan out 2*, conectando la salida de un inversor a la entrada de dos registros, como se muestra en la fig. 4.

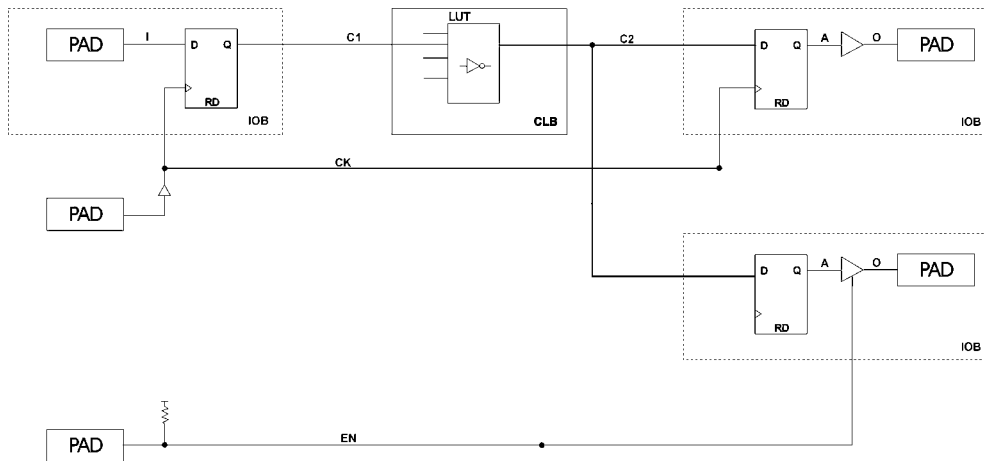


Fig. 4: Estudio de pistas con *Fan Out 2* en una FPGA.

En este caso, considerando que  $C_I$  y  $C_A$  son mucho menores que  $C_1$  y  $C_2$ , se tiene:

$$C_1 + C_2 < \frac{P - P_O - P_{CK}}{V_{dd}^2 f}$$

Si se tiene, como se va a mostrar luego, para pistas con *fan out 1*, una relación lineal entre capacidad y retardo, se puede hacer:

$$C_2 < \frac{P - P_O - P_{CK}}{V_{dd}^2 f} - c_1 t_1$$

donde  $c_1$  es la capacidad por unidad de retardo para nodos con *fan out* uno. Si se replica este esquema  $n$  veces, se tendrá:

$$\overline{C}_2 < \frac{1}{n} \left( \frac{P - P_o - P_{CK}}{V_{dd}^2 f} - c_1 \overline{t}_1 \right) \quad (7)$$

Esta fórmula vale para calcular esta cota superior de la capacidad para nodos con cualquier *fan out* mayor que uno.

### Resultados Experimentales

Los experimentos se hicieron usando el mismo dispositivo (*XC4010PC84-4C* de Xilinx) sobre una placa de construcción propia, y el ambiente de desarrollo Foundation 3.1.

Se calculó la capacidad en nodos simples (*fan out* 1) mediante la Ec.6, y con *fan out* 2, 4, 8 y 16 mediante la Ec. 7. Para cada experimento se hizo la máxima cantidad de replicas del circuito elemental de prueba que permitía el dispositivo seleccionado. De esta manera, la implementación que se descarga en el chip tiene una cantidad de los circuitos propuestos que permite reducir lo máximo posible el error de medición que se produce al medir la corriente que circula por la FPGA, como lo sugiere la Fig. 1.

Para nodos simples el análisis de regresión lineal, asumiendo que la recta pasa por cero presenta los resultados de la figura 5. Con la pendiente así obtenida, se puede estimar la capacidad por unidad de retardo en nodos con *fan out* mayor. Los resultados para *fan out* 2, 4, 8 y 16 se muestran en las figuras de 6 a 9.

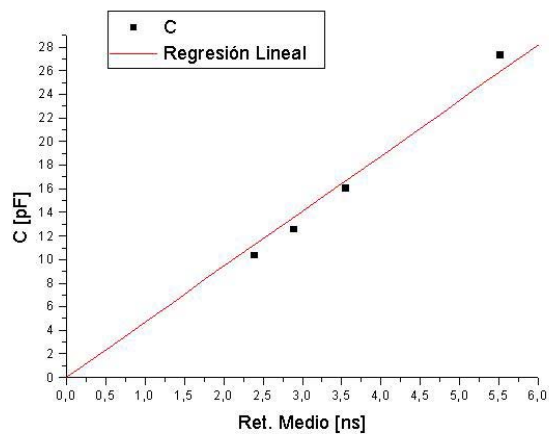


Fig. 5: Relación entre retardo medio y la capacidad del nodo, para nodos con *fan out* 1.

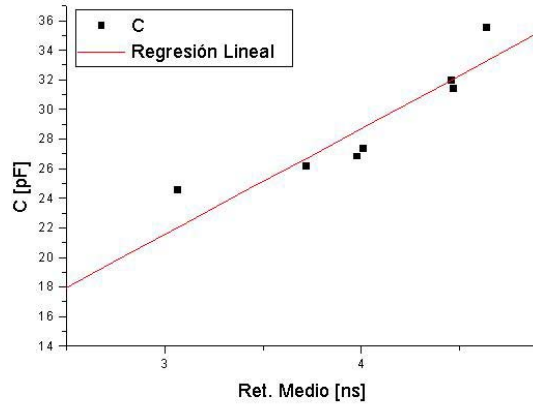


Fig. 6: Relación entre retardo medio y la capacidad del nodo, para nodos con *fan out* 2.

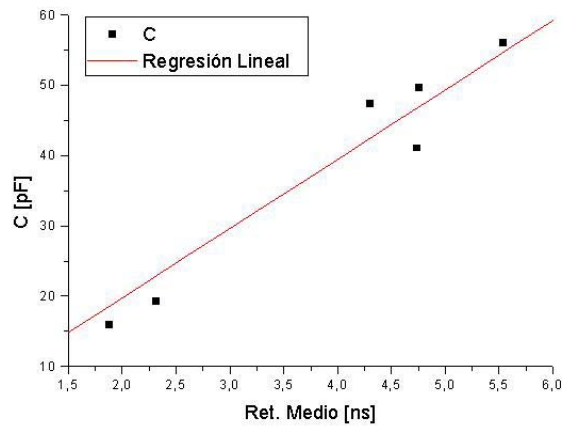


Fig. 7: Relación entre retardo medio y la capacidad del nodo, para nodos con *fan out* 4.

Los valores de las pendientes obtenidas para nodos con diferente fan out, junto con los principales indicadores estadísticos del respectivo análisis de regresión lineal, se muestran en la Tabla 2. Con estos valores para los coeficientes de determinación se puede explicar más del 90 % de la variación de capacidad dado el retardo medio de cada nodo.

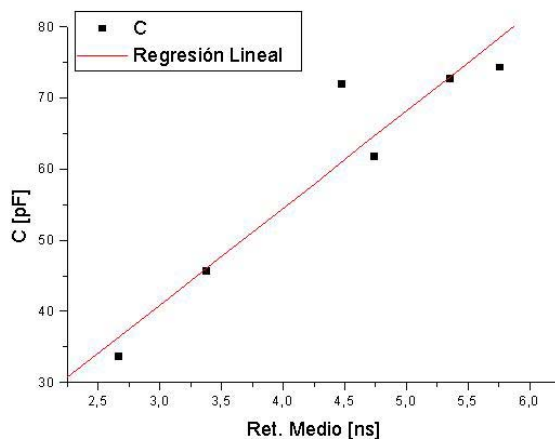


Fig. 8: Relación entre retardo medio y la capacidad del nodo, para nodos con *fan out* 8.



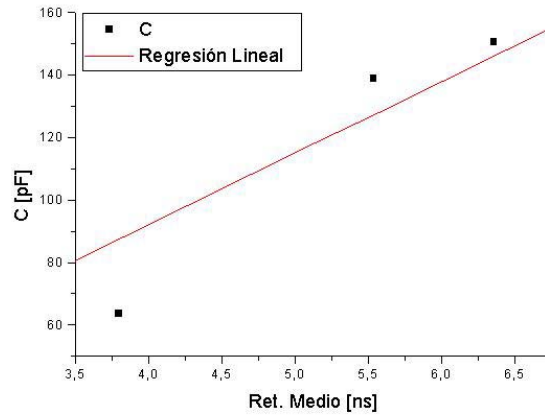


Fig. 9: Relación entre retardo medio y la capacidad del nodo, para nodos con *fan out* 16.

Fan Out	Pendiente [pF/ns]	Error	Coef. de Det. R
1	4,70	0,15517	0,99911
2	7,18	0,15689	0,90597
4	9,87	0,40935	0,97479
8	13,63	0,49914	0,94442
16	23,00	2,04289	0,98117

Tabla 2: Resumen de los resultados del análisis de regresión lineal para cada valor de *fan out*.

Si se grafica la pendiente en función del *fan out*, se obtiene la figura 10. En este caso se obtiene un mejor ajuste asumiendo una relación polinomial de grado 2 (Ec. 8).

$$P = A + B1.F + B2.F^2 \quad (\text{Ec. 8})$$

Los resultados del análisis de regresión se resumen en la Tabla 3.

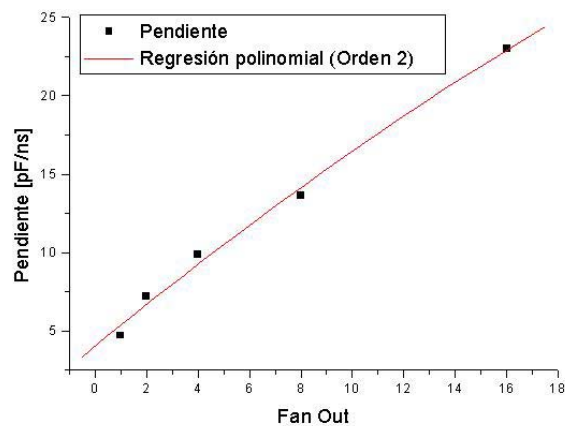


Fig. 10: Relación entre *fan out* y su correspondiente capacidad por unidad de retardo.

Parámetro	Valor	Error
<b>A</b>	4,03155	0,85833
<b>B1</b>	1,34865	0,29871
<b>B2</b>	-0,01052	0,01687
	<b>R-Cuadrado</b>	
	0,99339	

Tabla 3: Resumen de los resultados del análisis de regresión para las pendientes en función del *fan out*.

De los análisis anteriores se propone la siguiente fórmula para estimar la capacidad en los nodos de la FPGA estudiada, expresada en pF, dados su retardo medio,  $d$  en ns, y su *fan out*,  $F$ .

$$C = (4,04 + 1,35F - 0,01F^2) \cdot d \quad (\text{Ec. 9})$$

Si se usa esta fórmula en las mediciones realizadas se obtiene la distribución en el error relativo que se muestra en la Fig. 11. El error se mantiene por debajo del 25 %, a excepción de un caso que es un nodo con *fan out* 16, que tiene un residuo particularmente alto, como se ve en la Fig. 9.

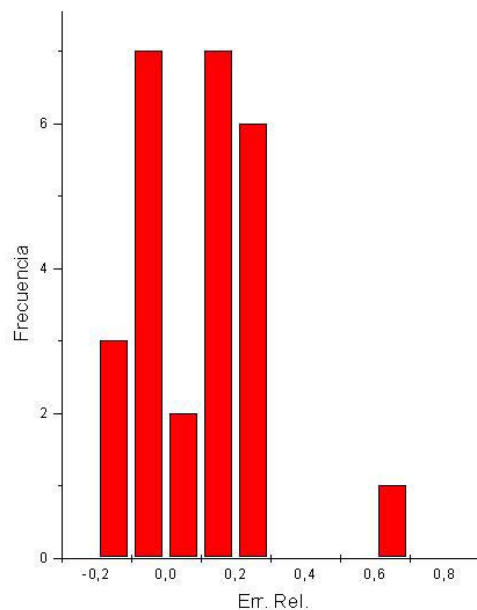


Fig. 11: Distribución del error relativo en las mediciones respecto de los valores estimados mediante la Ec. 8.

## Conclusiones

En el contexto actual de diseño para FPGAs comerciales, donde los fabricantes no dan información sobre la capacidad física de los nodos del circuito, se hace difícil estimar el consumo de potencia.

En este trabajo se propone una fórmula para estimar la capacidad de los nodos en función de su retardo medio y su *fan out*. Los coeficientes de la fórmula propuesta deben obtenerse para cada FPGA mediante mediciones físicas de circuitos especialmente pensados para aislar las magnitudes de interés, y un análisis de regresión.

Los resultados preliminares indican que el error es tolerable como para usar esos valores de capacidad en estimación de consumo de potencia.

## Referencias

[Alv97] Atila Alvandpour and Christer Svensson , "A Wire Capacitance Estimation Technique for Power Consuming Interconnections at high levels of abstraction" PATMOS'97, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, September 8-10, 1997.

[Boe96] E. Boemo, "Contribution to the Design of Fine-grain Pipelined VLSI Arrays", Ph.D. Thesis, ETSI Telecomunicación, Universidad Politécnica de Madrid, 1996.

[Kat95] Srinivas Katkoori, Nand Kumar and Ranga Vemuri, "High Level Profiling Based Low Power Synthesis Technique", International Conference on Computer Design (ICCD) 1995.

[Lan96b] Landman, P.E., Rabaey, J.M., "Activity-Sensitive Architectural Power Analysis", IEEE Transactions on CAD, Vol. 15, No. 6, June 1996, pp. 571-587.

[Don79] W. E. Donath, "Placement and Average Interconnection Lengths of Computer Logic", IEEE Trans. on Circuits and Systems CAS26 (4) (1979), pp. 272-277.

[Osm00] T. Osmulski; J.T. Muehring; B. Veale; J. M. West; H. Li; S. Vanichayobon; S-H Ko; J.K. Antonio; S.K. Dhall "A Probabilistic Power Prediction Tool for the Xilinx 4000-Series FPGA", Procs. of the 5th International Workshop on Embedded/Distributed HPC Systems and Applications (EHPC 2000), Cancun, Mexico, May 2000, pp. 776-783.

[Rab96] J. Rabaey, "Digital Integrated Circuits: A Design Perspective", Prentice Hall, 1996.

[Str96] Dirk Stroobandt, "Improving Donath's technique for estimating the average interconnection length in computer logic", ELIS Technical Report DG 91-01, June 1996.

Gregory B. Sorkin. "Asymptotically perfect trivial global routing: A stochastic analysis", IEEE Transactions on Computer-Aided Design, CAD-6(5):820-827, September 1987.

## Apéndice 1

La Regla de Rent es una ley empírica que relaciona los puertos de un módulo con la lógica que contiene mediante:

$$t_k \approx AK^p$$

Donde:

$t_k$  es el número medio de terminales en el módulo bajo estudio,

$K$  es el tamaño del módulo medido en puertos u otra unidad de ese tipo,

$A$  es el número medio de terminales de las unidades que componen el módulo bajo estudio (por ej. El número medio de puertos de las puertas),

$p$  es una constante entre 0 y 1 vinculada con la topología.