Timing en FPGAs y Standard Cells

Parte 1: Fundamentos

Eduardo I. Boemo

Escuela Politécnica Superior
Universidad Autónoma de Madrid
Ctra. de Colmenar Km.15. 28049 Madrid, España.

http://www.ii.uam.es/~ivan
e-mail: eduardo.boemo@uam.es

Introducción

- **Problema**: Las herramientas EDA de FPGA escondes varios conceptos electrónicos al usuario que diseña a alto nivel.
- **Consecuencias:**
  - Aunque se escondan los detalles electrónicos, no significa que se escondan sus consecuencias sobre el circuito final.
  - Las herramientas tienen opciones avanzadas de diseño que requieren conocer algunos temas en profundidad.
- **Estrategia DIE**: Utilizar circuitos integrados tipo *Standard Cells* (no hay datos similares disponibles en Xilinx) para ayudar a entender resultados de una implementación FPGA y mejorar la formación de diseñadores de la UAM.

Capacidad

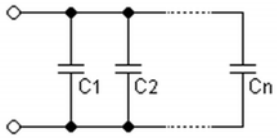
- Dos conductores separados por un aislante pueden almacenar carga.
- Q → Campo eléctrico → Diferencia de potencial V.
- C se define como la relación entre la carga que se almacena y el potencial V.
  - **C = Q/V** [Culombio/Voltio = Faradios].

C = Є A/d (capacitor o condensador plano)

A= área del cap.
d= distancia entre placas

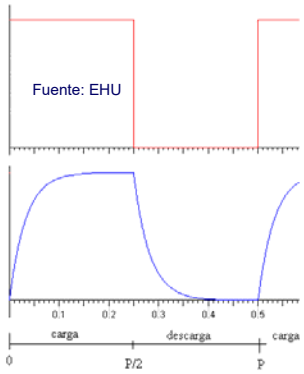
Capacidades en paralelo se suman

- C = Q/V



- Si están en paralelo → V es la misma para todos los
- C_{TOTAL} = C1 + C2 + … + Cn

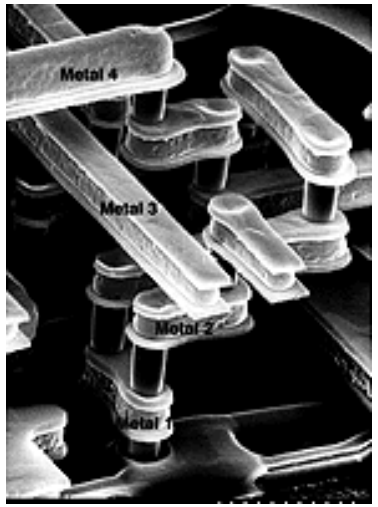
Carga y descarga de un capacitor



- Carga a tensión constante:
  - $V_{cap} = Vcc\ (1 - e^{-t/RC})$
- Carga a corriente constante
  - $V_{cap} = (i/C)\ t$
- Energía almacenada en un C:
  - $E_{cap} = ½\ C\ V^2$

Capacidad

Las pistas de un circuito integrado son capacitores casi "de libro".

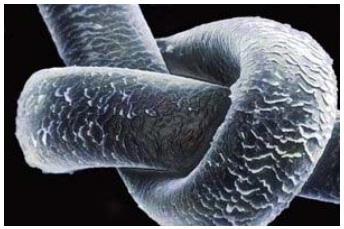


RC como parámetros distribuidos

	Capacitance to ground (aF/µm)	Coupling capacitance (aF/µm)	Resistance/ length (Ω/µm)
Metal 3	18	9	0.2
Metal 2	47	24	0.3
Metal 1	76	36	0.3

- Femto f 10^{-15}
- Atto a 10^{-18}
- Zepto z 10^{-21}
- Yocto y 10^{-24}

µm = millonésima parte de un metro
= milésima del mm.



Diámetro cabello humano:
15 (muy fino) a 170 µm.

Capacidades de una FPGA: XC3020

Die area:	39,600 mil2
Matrix height (Y):	480 µm
Matrix width (X):	370 µ m
Matrix transistor resistance:	0.5–1k W
Matrix transistor parasitic capacitance:	0.01–0.02 pF
PIP transistor resistance:	0.5–1k ohm
PIP transistor parasitic capacitance:	0.01–0.02 pF
Single-length line (X, Y):	370 um, 480 µ m
Single-length line capacitance:	0.075 a 0.1 pF
Horizontal Longline (8X):	8 cols.=2960 µ m
Horizontal Longline metal capacitance:	0.6 pF

Fuente: Smith

Uno de los pocos datos internos publicados de una FPGA

Timing en FPGAs y Standard Cells

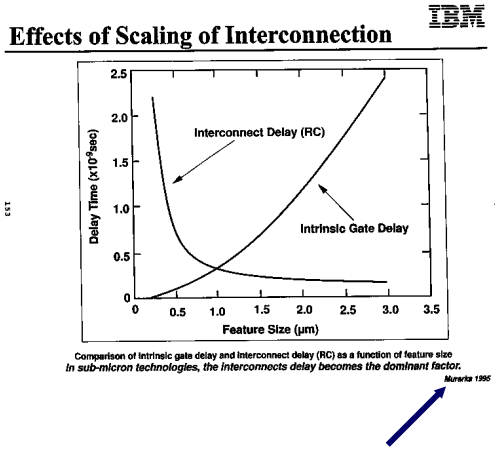
Parte 2: Números concretos

Eduardo I. Boemo

Escuela Politécnica Superior
Universidad Autónoma de Madrid
Ctra. de Colmenar Km.15. 28049 Madrid,
España.

http://www.ii.uam.es/~ivan
e-mail: eduardo.boemo@uam.es

Retardos intrínsecos y extrínsecos



Comparison of intrinsic gate delay and interconnect delay (RC) as a function of feature size
In sub-micron technologies, the interconnects delay becomes the dominant factor.
Murarka 1995

El retardo de una puerta se separa en:

- **Intrínsico**: retardo interno de conmutación de la puerta. Es un valor fijo.
- **Extrínseco**: retardo debido a la carga de la capacidad que debe manejar la salida de la puerta. No es fijo, depende de la carga

1995: hace más de 25 años!

SClib
ATC18

INV0Dx — **Inverting Buffers**

Specification — **Datasheet Version: 1.5.5-1.0.0, Jan 2002** — **Symbol**

INV0D0, INV0D1, INV0D2, INV0D4, INV0D7 and INV0DA are Inverting Buffers with 0.5x, 1x, 2x, 4x, 7x and 10x drive capabilities.

Function Table

Input	Output
I	ZN
L	H
H	L

Cell Parameters

Code	Parameter	inv0d0	inv0d1	inv0d2	inv0d4	inv0d7	inv0da	Unit
X	Length	1.1	1.7	2.2	3.4	5.0	5.6	um
nTran	Transistor Count	2	2	4	8	9	13	trans
Power	AC Power Dissipation	0.006	0.019	0.037	0.073	0.117	0.165	uW/MHz

Input Specifications (Fanin)

Name	Pin Description	inv0d0	inv0d1	inv0d2	inv0d4	inv0d7	inv0da	Unit
i	Data In	0.002	0.005	0.011	0.022	0.035	0.050	pF

Output Specifications (Fanout)

Name	Pin Description	inv0d0	inv0d1	inv0d2	inv0d4	inv0d7	inv0da	Unit
zn	Data Out	0.029	0.137	0.278	0.555	0.946	1.28	pF

Typical Propagation Delays (VDD=1.8V, Temp.=25°C, Input Slope=1ns)

Code	From	To	inv0d0	inv0d1	inv0d2	inv0d4	inv0d7	inv0da	Unit
tpdhl	i	zn	0.063	0.042	0.030	0.030	0.030	0.035	ns
tpdlh	i	zn	0.147	0.111	0.105	0.105	0.101	0.105	ns
dtpdhl	any	zn	6.571	1.716	0.813	0.406	0.261	0.177	ns/pF
dtpdlh	any	zn	11.336	2.570	1.284	0.642	0.377	0.276	ns/pF

Capacidad y Retardo

Neil Weste
Kamran Eshraghian

Fanin, Fanout y Retardos

Fanout (in pF):

- *Is the total capacitance that a signal will have to drive; this includes gate capacitance as well as interconnect capacitance.*
- *The fanout figure is based on output fall and rise time staying within reasonable limits*.

- En Xilinx se entiende por fanout simplemente el número de entradas conectadas a una salida

Tiempo de propagación:

- ***tpdhl*** *propagation delay, high-to-low.*
- ***tpdlh*** *propagation delay, low-to-high.*
- ***dtpdhl*** *differential (load-dependent) propagation delay.*
- ***dtpdlh*** *differential (load-dependent) propagation delay.*

Fuente: *Atmel Corp. 2001*

Retardo Intrínseco y Extrínseco en Xilinx

Table 5-1: **General Slice Timing Parameters**

Parameter	Function	Description
Combinatorial Delays		
T_{ILO} (1)	A/B/C/D inputs to A/B/C/D outputs	Propagation delay from the A/B/C/D LUT inputs of the slice, through the look-up tables (LUTs), to the A/B/C/D outputs of the slice.
T_{ITO}	A/B/C/D inputs through transparent latch to AQ/BQ/CQ/DQ outputs	Propagation delay from the A/B/C/D LUT inputs of the slice, through the LUTs to the AQ/BQ/CQ/DQ outputs of the slice sequential elements (configured as a latch).

Zynq-7000 SoC (Z-7030, Z-7035, Z-7045, and Z-7100): DC and AC Switching Characteristics

How much fast is a FPGA?

CLB Switching Characteristics

Table 53: **CLB Switching Characteristics**

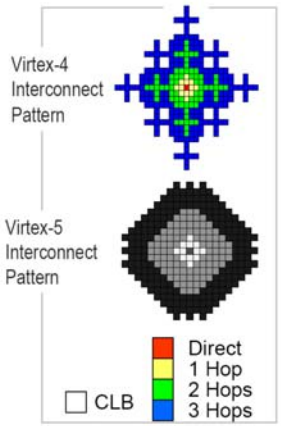
Symbol	Description	Speed Grade				Units
		-3	-2	-1	-1L	
Combinatorial Delays						
T_{ILO}	An – Dn LUT address to A	0.06	0.07	0.07	0.09	ns, Max
	An – Dn LUT address to AMUX/CMUX	0.18	0.20	0.22	0.25	ns, Max
	An – Dn LUT address to BMUX_A	0.28	0.31	0.36	0.40	ns, Max
Sequential Delays						
T_{CKO}	Clock to AQ – DQ outputs	0.29	0.33	0.39	0.44	ns, Max
T_{SHCKO}	Clock to AMUX – DMUX outputs	0.36	0.40	0.47	0.53	ns, Max
Setup and Hold Times of CLB Flip-Flops Before/After Clock CLK						
T_{DICK}/T_{CKDI}	A – D input to CLK on A – D Flip Flops	0.30/ 0.17	0.36/ 0.18	0.43/ 0.20	0.44/ 0.25	ns, Min
$T_{CECK\_CLB}/T_{CKCE\_CLB}$	CE input to CLK on A – D Flip Flops	0.20/ 0.00	0.25/ 0.00	0.32/ 0.00	0.29/ 0.00	ns, Min
T_{SRCK}/T_{CKSR}	SR input to CLK on A – D Flip Flops	0.39/ –0.07	0.44/ –0.07	0.52/ –0.07	0.58/ –0.09	ns, Min
T_{CINCK}/T_{CKCIN}	CIN input to CLK on A – D Flip Flops	0.16/ 0.12	0.19/ 0.14	0.24/ 0.16	0.22/ 0.21	ns, Min

Retardo de interconexión en FPGA?

- Diagonally symmetric interconnect pattern enabling faster routing
  - Fewer hops
  - Faster connections

Path Delay	Virtex-4	Virtex-5	Improvement
One Hop	751 ps	665 ps	13 %
Two Hops	906 ps	723 ps	25 %

Note: Full path delay, flip-flop to flip-flop crossing a LUT; all timing numbers are obtained with advanced speed file for Virtex-5



Fuente: Xilinx Inc.

Retardo de las patas en FPGA?

IOB Pad Input/Output/3-State Switching Characteristics

Table 44 summarizes the values of standard-specific data input delay adjustments, output delays terminating at pads (based on standard) and 3-state delays.

T_{IOPI} is described as the delay from IOB pad through the input buffer to the I-pin of an IOB pad. The delay varies depending on the capability of the SelectIO input buffer.

T_{IOOP} is described as the delay from the O pin to the IOB pad through the output buffer of an IOB pad. The delay varies depending on the capability of the SelectIO output buffer.

T_{IOTP} is described as the delay from the T pin to the IOB pad through the output buffer of an IOB pad, when 3-state is disabled. The delay varies depending on the SelectIO capability of the output buffer.

Table 45 summarizes the value of T_{IOTPHZ}. T_{IOTPHZ} is described as the delay from the T pin to the IOB pad through the output buffer of an IOB pad, when 3-state is enabled (i.e., a high impedance state).

Table 44: **IOB Switching Characteristics**

I/O Standard	T_{IOPI} Speed Grade				T_{IOOP} Speed Grade				T_{IOTP} Speed Grade				Units
	-3	-2	-1	-1L	-3	-2	-1	-1L	-3	-2	-1	-1L	
LVDS_25	0.85	0.94	1.09	1.08	1.45	1.54	1.68	1.62	1.45	1.54	1.68	1.62	ns
LVDSEXT_25	0.85	0.94	1.09	1.08	1.53	1.65	1.84	1.73	1.53	1.65	1.84	1.73	ns
HT_25	0.85	0.94	1.09	1.08	1.51	1.62	1.78	1.69	1.51	1.62	1.78	1.69	ns
BLVDS_25	0.85	0.94	1.09	1.08	1.39	1.50	1.67	1.65	1.39	1.50	1.67	1.65	ns
RSDS_25 (point to point)	0.85	0.94	1.09	1.08	1.45	1.54	1.68	1.62	1.45	1.54	1.68	1.62	ns
HSTL_I	0.81	0.91	1.06	1.06	1.45	1.56	1.73	1.71	1.45	1.56	1.73	1.71	ns
HSTL_II	0.81	0.91	1.06	1.06	1.44	1.56	1.74	1.72	1.44	1.56	1.74	1.72	ns

Capacidades I/O en una FPGA? Ejemplo Virtex 6

Table 3: **DC Characteristics Over Recommended Operating Conditions (1,2)**

Symbol	Description	Min	Typ	Max	Units
V_{DRINT}	Data retention V_{CCINT} voltage (below which configuration data might be lost)	0.75	–	–	V
V_{DRI}	Data retention V_{CCAUX} voltage (below which configuration data might be lost)	2.0	–	–	V
I_{REF}	V_{REF} leakage current per pin	–	–	10	µA
I_L	Input or output leakage current per pin (sample-tested)	–	–	10	µA
C_{IN}(3)	Die input capacitance at the pad			8	pF
I_{RPU}	Pad pull-up (when selected) @ V_{IN} = 0V, V_{CCO} = 2.5V	20	–	80	µA
	Pad pull-up (when selected) @ V_{IN} = 0V, V_{CCO} = 1.8V	8	–	40	µA
	Pad pull-up (when selected) @ V_{IN} = 0V, V_{CCO} = 1.5V	5	–	30	µA
	Pad pull-up (when selected) @ V_{IN} = 0V, V_{CCO} = 1.2V	1	–	20	µA
I_{RPD}	Pad pull-down (when selected) @ V_{IN} = 2.5V	3	–	80	µA
I_{BATT}	Battery supply current	–	–	150	nA
n	Temperature diode ideality factor	–	1.0002	–	n
r	Series resistance	–	5	–	Ω

Notes:

1. Typical values are specified at nominal voltage, 25°C.
2. Maximum value specified for worst case process at 25°C.
3. This measurement represents the die capacitance at the pad, not including the package.

Capacidades I/O en una FPGA? Ejemplo Virtex 6

Output Delay Measurements

Output delays are measured using a Tektronix P6245 TDS500/600 probe (< 1 pF) across approximately 4" of FR4 microstrip trace. Standard termination was used for all testing. The propagation delay of the 4" trace is characterized separately and subtracted from the final measurement, and is therefore not included in the generalized test setups shown in Figure 6 and Figure 7.

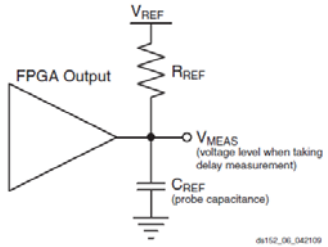


Figure 6: **Single Ended Test Setup**

Table 44: **IOB Switching Characteristics *(Cont'd)***

I/O Standard	T_{IOPI} Speed Grade			
	-3	-2	-1	-1L
LVCMOS18, Fast, 2 mA	0.55	0.61	0.71	0.73
LVCMOS18, Fast, 4 mA	0.55	0.61	0.71	0.73
LVCMOS18, Fast, 6 mA	0.55	0.61	0.71	0.73
LVCMOS18, Fast, 8 mA	0.55	0.61	0.71	0.73
LVCMOS18, Fast, 12 mA	0.55	0.61	0.71	0.73
LVCMOS18, Fast, 16 mA	0.55	0.61	0.71	0.73

- Qué sucederá con los tiempos de un IOB si Ud. carga una pata con más de 1 pF ? (ej. Un pin de una FPGA con 8 pF)

Miscellaneous: Cómo presenta Xilinx sus datos?

Switching Characteristics

All values represented in this data sheet are based on the advanced speed specification (version 1.05). Switching characteristics are specified on a per-speed-grade basis and can be designated as Advance, Preliminary, or Production. Each designation is defined as follows:

Advance

These specifications are based on simulations only and are typically available soon after device design specifications are frozen. Although speed grades with this designation are considered relatively stable and conservative, some under-reporting might still occur.

Preliminary

These specifications are based on complete ES (engineering sample) silicon characterization. Devices and speed grades with this designation are intended to give a better indication of the expected performance of production silicon. The probability of under-reporting delays is greatly reduced as compared to Advance data.

Production

These specifications are released once enough production silicon of a particular device family member has been characterized to provide full correlation between specifications and devices over numerous production lots. There is no under-reporting of delays, and customers receive formal notification of any subsequent changes. Typically, the slowest speed grades transition to Production before faster speed grades.

Timing en FPGAs y Standard Cells

Parte 3: ¿Cómo bajar T?

Eduardo I. Boemo

Escuela Politécnica Superior
Universidad Autónoma de Madrid
Ctra. de Colmenar Km.15. 28049 Madrid,
España.

http://www.ii.uam.es/~ivan
e-mail: eduardo.boemo@uam.es

Bajar retardo de pista (Capacidad) en Diseño FPGA a Alto Nivel

- Para “visualizar” la capacidad total que maneja una determinada puerta, hay que sumar:
  - *Fanout* que soporta cada salida.
  - Capacidad distribuida de cada cable.
- Minimizar retardo requiere:
  - Reducir fan-out (en el significado, cantidad de salidas que maneja una entrada) C↓
  - Reducir *wire length* C↓
  - Aumentar *driving* i ↑
  - Bajar Temp ó subir Vcc

Como minimizar capacidad diseñando a alto nivel?

Retardo o Capacidad alta: implica pistas largas o mucho *fanout*.

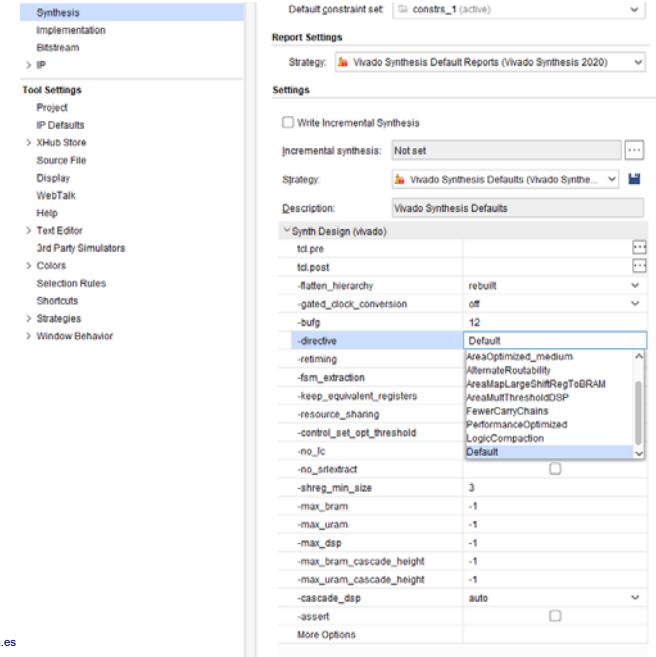
Minimizar capacidad → Minimizar longitud de pista

- Atributos: *time_spec.*
- Minimizar tamaño del circuito
  - → reducir nº de bits
  - Placement manual

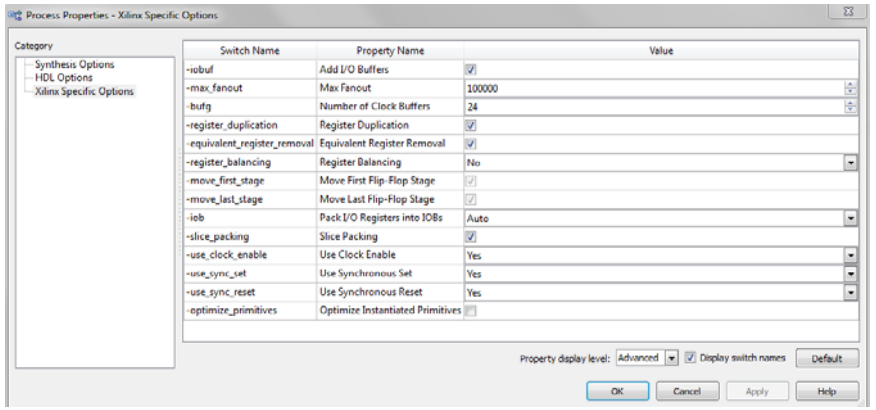
Minimizar capacidad → limitar *fanout*

- Replicar HW
- Atributo *max_fanou*t (Xilinx)
- “Tocar” las opciones por defecto de la síntesis”

Opciones síntesis (Vivado 2020.1)

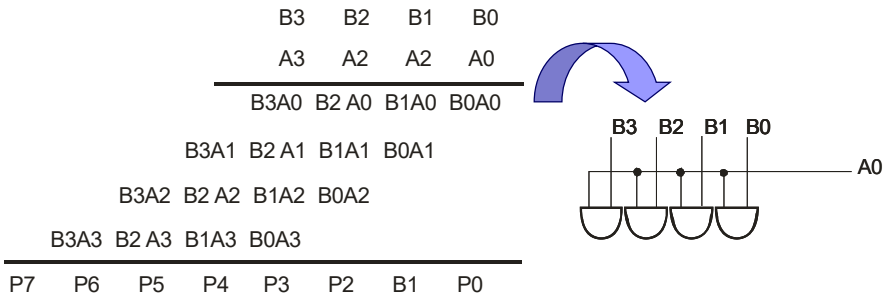


Limitación de *Fanout* en síntesis (ISE 14.7)



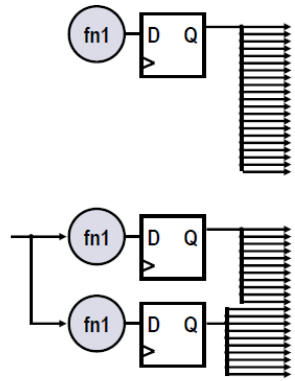
¿Cómo limitar el *fanout* de un circuito?

- Las señales **globales *(broadcasted)*** son muy comunes en los circuitos digitales. ¿Cómo evitarlas o limitarlas?



¿Cómo limitar el *fanout* de un circuito? Recomendación Xilinx

- High-fanout nets can be slow and hard to route
- Duplicating flip-flops can fix both problems
  - Reduced fanout shortens net delays
  - Each flip-flop can fanout to a different physical region of the chip to reduce routing congestion
- Design trade-offs
  - Gain routability and performance
  - Increase design area
  - Increase fanout of other nets



Limitación de *Fanout*: VHDL y XCF.

VHDL

Before using MAX_FANOUT, declare it with the following syntax:

```
attribute MAX_FANOUT: string;
```

After MAX_FANOUT has been declared, specify the VHDL constraint as follows:

```
attribute MAX_FANOUT of {signal_name|entity_name}: {signal|entity} is
"integer";
```

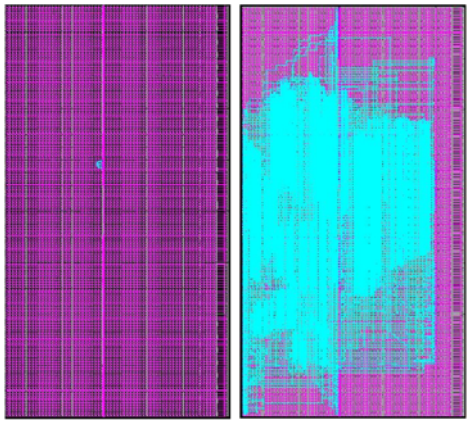
XCF

```
MODEL "entity_name" max_fanout=integer;

BEGIN MODEL "entity_name"

      NET "signal_name" max_fanout=integer;

END;
```

Ver además:

- **EQUIVALENT_REGISTER_REMOVAL** is a synthesis constraint. It enables or disables flip-flop optimization related only to the flip-flops described on RTL level. (Instantiated flip-flops are not removed)

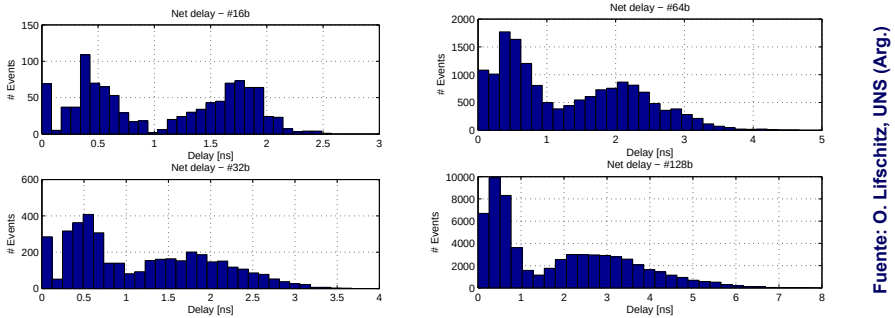
Área (Tamaño) y Retardos: Un caso de estudio en Xilinx



Fuente: O. Lifschitz, UNS (Arg.)

- Familia de multiplicadores de 2x2 a 128x128
- V5 – ISE 12.1
- PPR automático.
- Basado en LUTs.
- Retardo = f(area)
  - Pistas globales
  - Distancia

Área (Tamaño) y Retardos: Un caso de estudio en Xilinx



Fuente: O. Lifschitz, UNS (Arg.)

Concepto de "Deration": Ejemplo Actel

- Tj y Vcc modifican los retardos nominales del *chip*.
- Vcc además influye cuadráticamente en la potencia.
- Tj depende de los datos de entrada (actividad).
- Tj depende de la estrategia de *power management.*

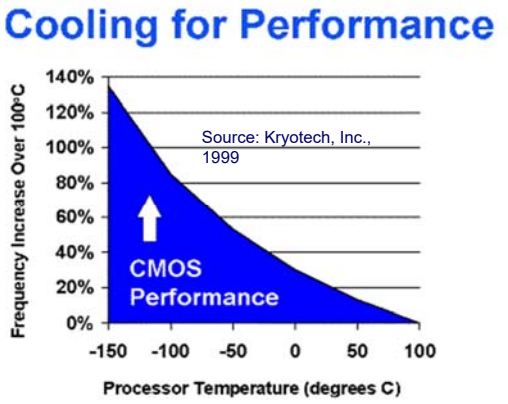
Temperature and Voltage Derating Factors

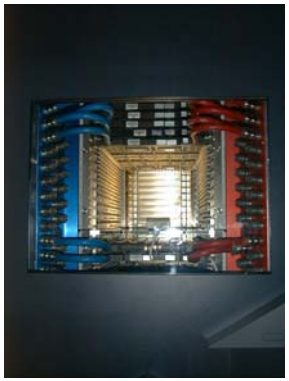
(Normalized to Worst-Case Commercial, T_J = 70°C, V_{CCA} = 2.3V)

	Junction Temperature (T_J)						
V_{CCA}	–55	–40	0	25	70	85	125
2.3	0.75	.079	0.88	0.89	1.00	1.04	1.16
2.5	0.70	0.74	0.82	0.83	0.93	0.97	1.08
2.7	0.66	0.69	0.79	0.79	0.88	0.92	1.02

Merriam-Webster's Collegiate Dictionary: *De.rate vt (1947): to lower the rated capability of (as electrical or mechanical apparatus) because of deterioration or inadequacy.*

ps y Temperatura





Freon cooling at Cray 2
EPFL, © EIB

Reduce temperature to increase speed is well-known. Unfortunatelly, the reverse is also true: Speed decrease 0,3 % per ºC of increment on T