Niveles de abstracción de un diseño

- Sistema
  - Partición HS-SW, memoria, Dynamic Power Management (DPM), Interfaces
- Algoritmos
  - Complejidad, concurrencia, paralelismo
- Arquitectura
  - Concurrencia, paralelismo, pipelining, tipo de datos, Power Management.
- Diseño del circuito
  - Mapeo, lógica, codificación, rutas, glitches
- Tecnología
  - Proceso de fabricación (V threshold), escalado

Co-funded by the Erasmus+ Programme of the European Union

Eliminación de glitches

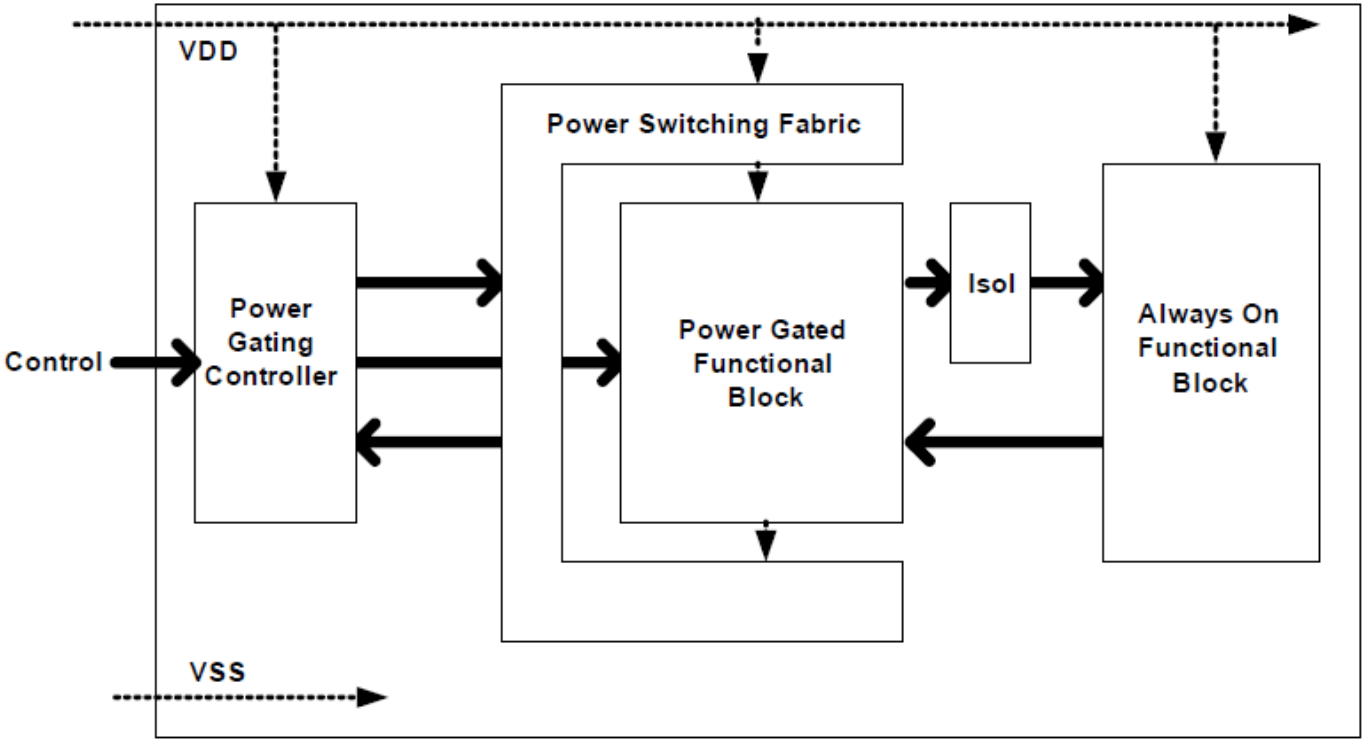
- Gate freezing (F-gate)
- Gate sizing and path balancing
- Multiple threshold transistor

- Eliminar azares
- Codificar en Gray
- Disminuir profundidad de lógica con FF

Deshabilitar parte de un circuito

- Power Gating
- Clock Gating, Clock Enable
- Data Gating, Blocking inputs
- Latches (transparents)

Power Gating



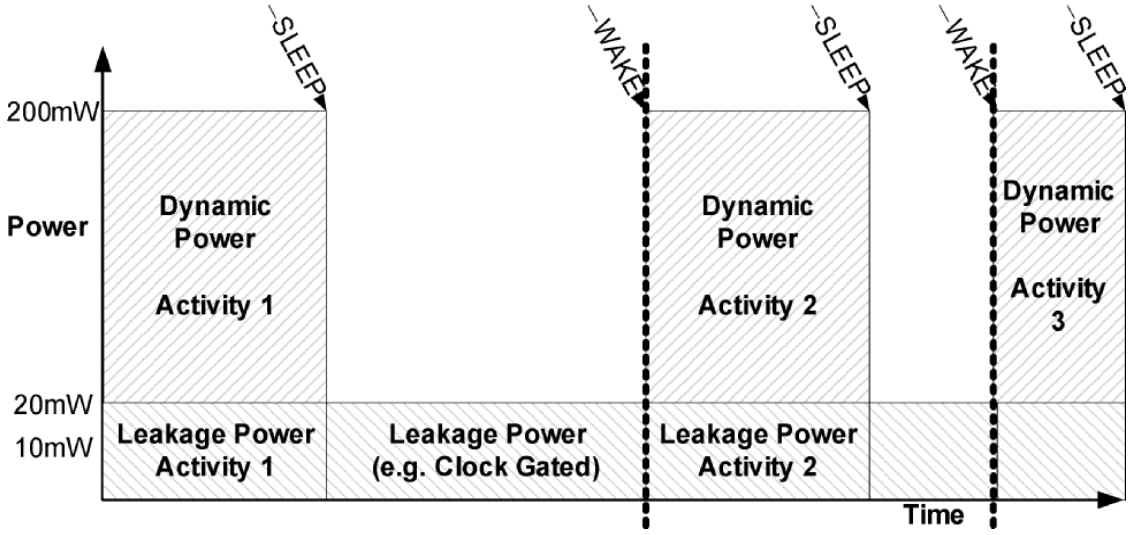
Libro:

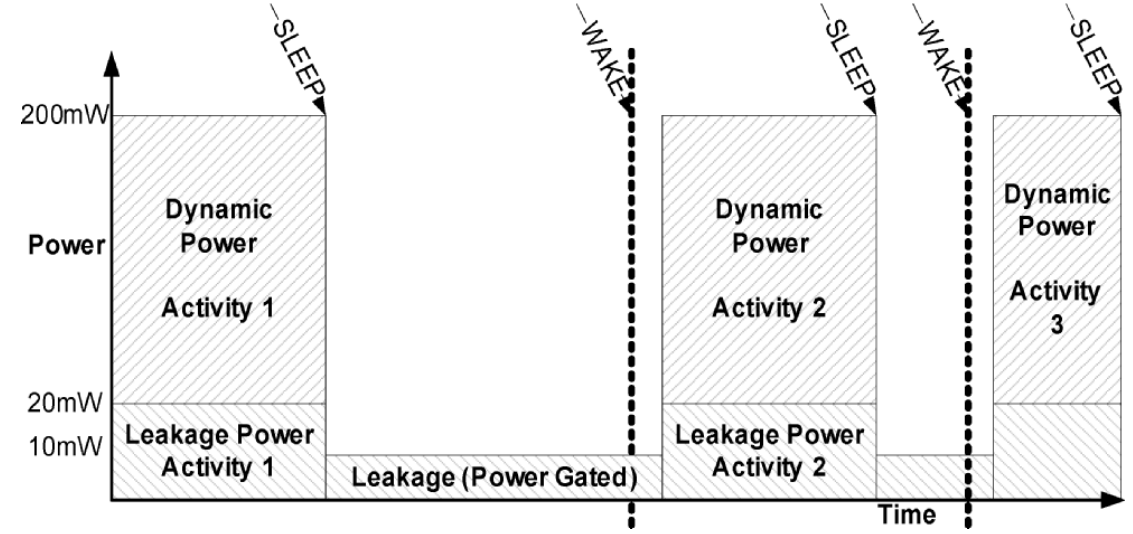
Low Power Methodology Manual: For System-on-Chip Design

Keating, M., Flynn, D., Aitken, R.,Gibbons, A., Shi,

New York, NY: Springer, 2008.

ARM+Synopsys Ejemplos

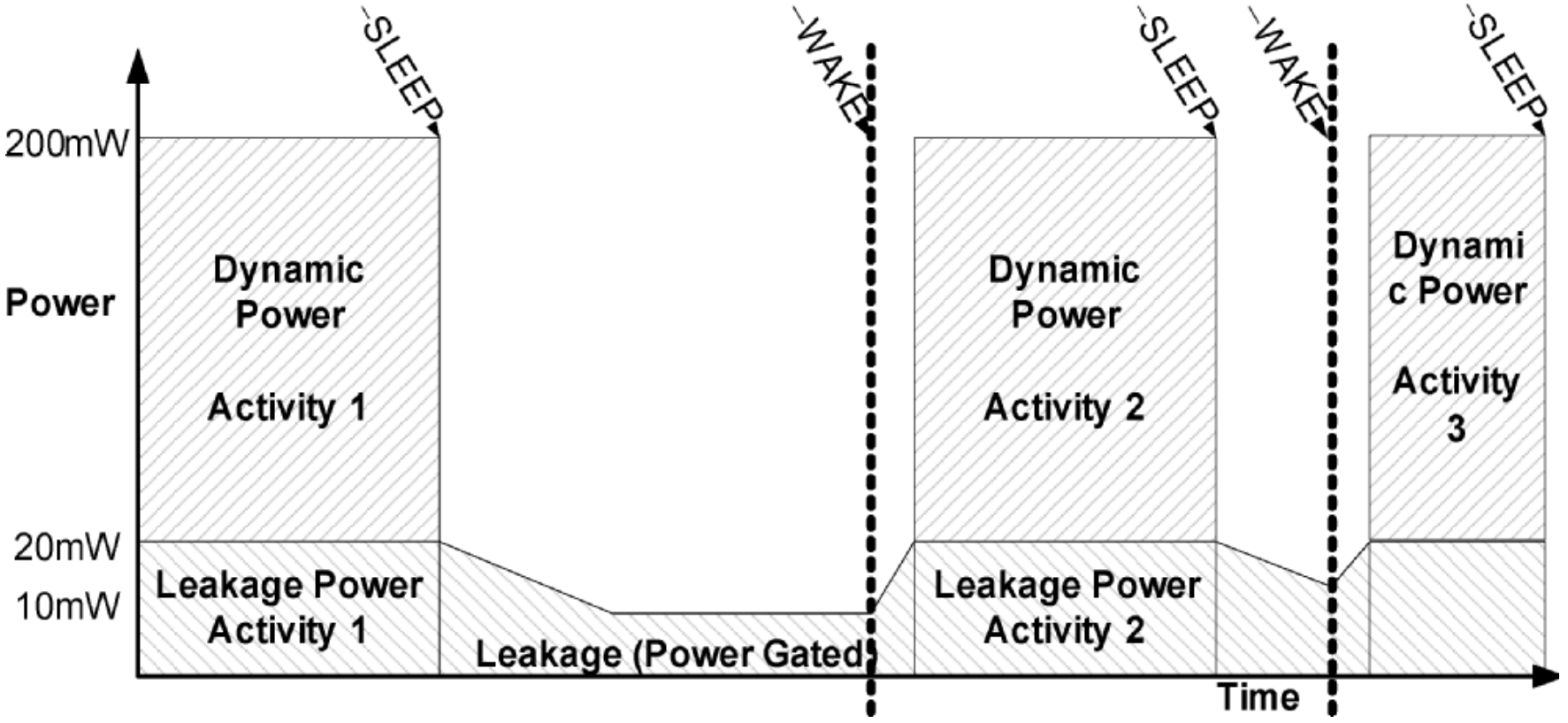


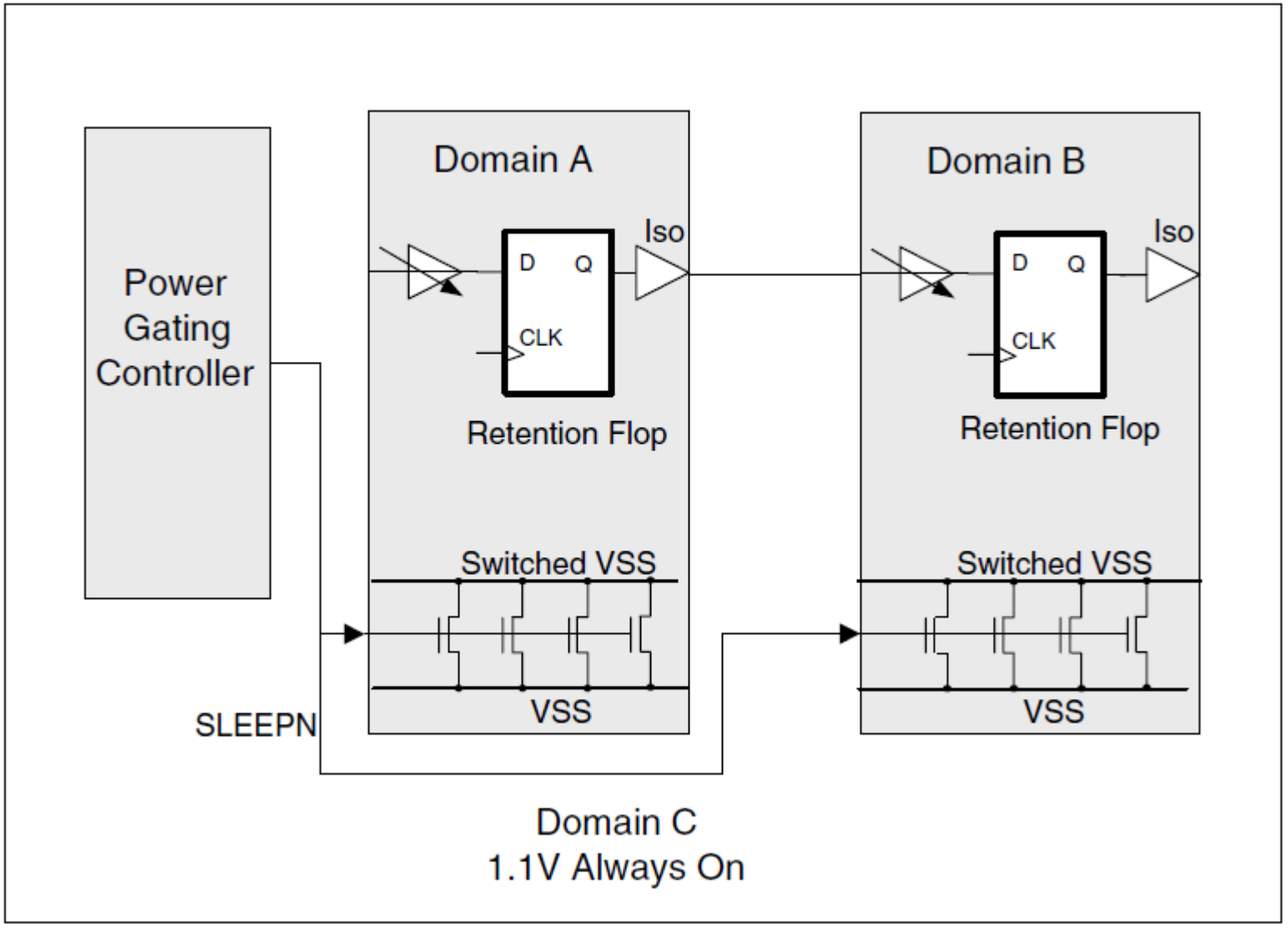


- Activity profile without power gating

- Activity profile with power gating

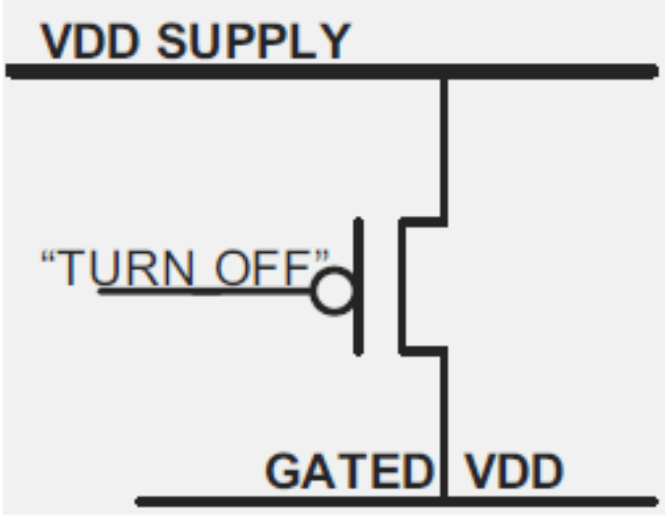
Realistic power gating



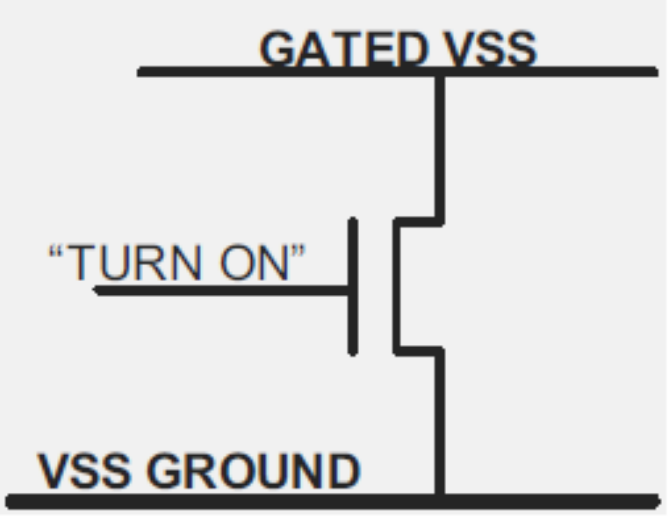

Power
Gating
Controller
Domain A
Iso
D
Q
CLK
Retention Flop
Switched VSS
VSS
Domain B
Iso
D
Q
CLK
Retention Flop
Switched VSS
VSS
SLEEPN
Domain C
1.1V Always On

Power gating transistor structures

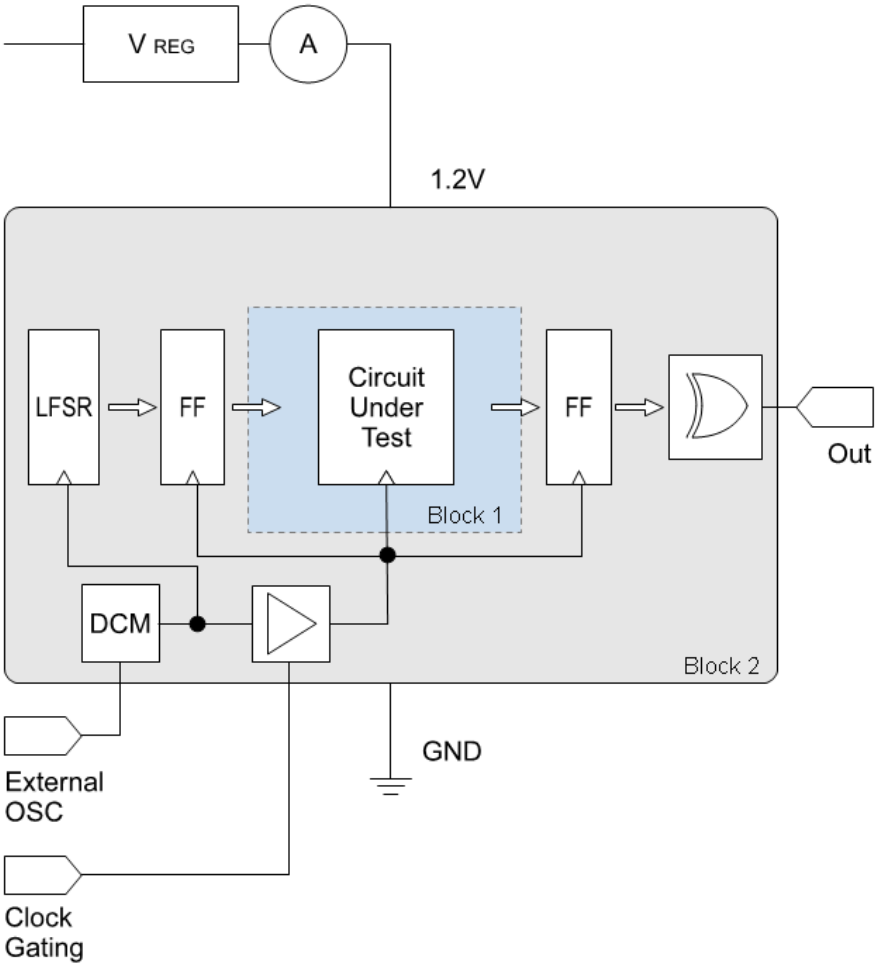
Basic Header-Switch structure:

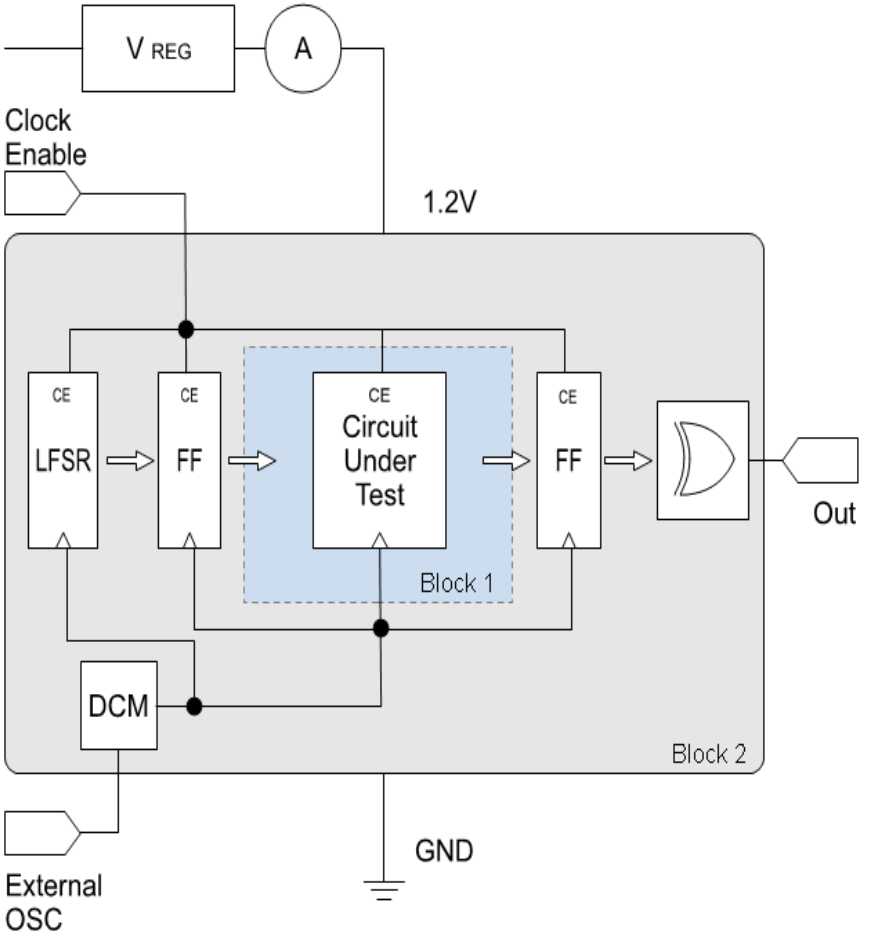


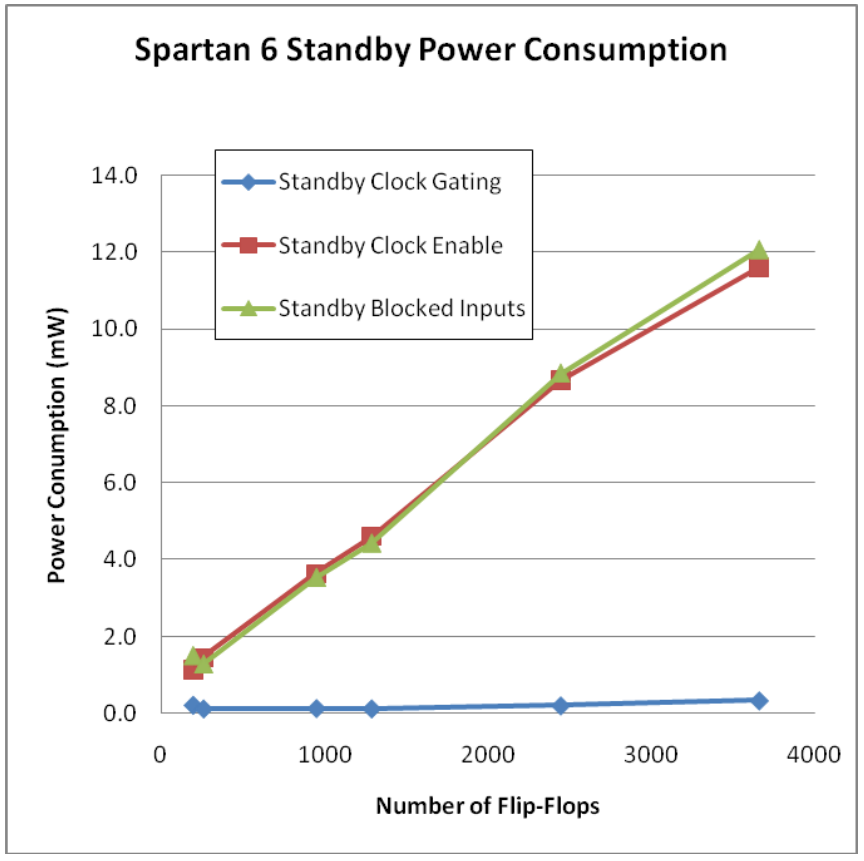
Basic Footer-Switch structure:

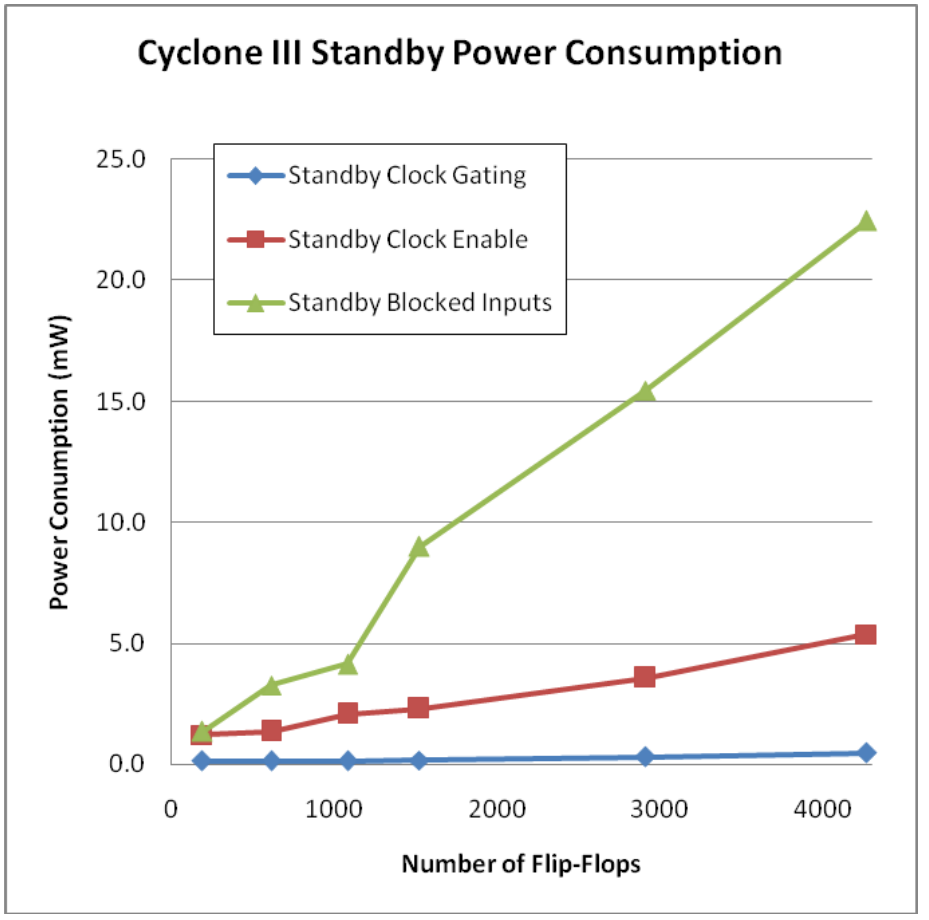


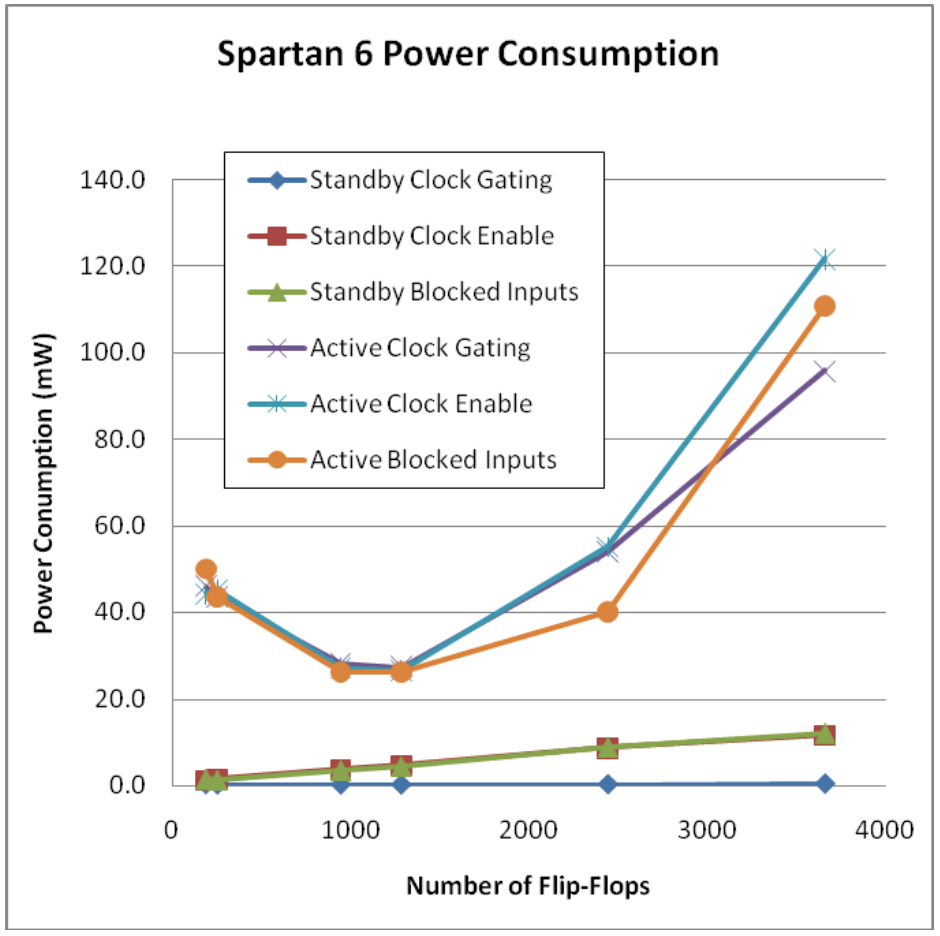
Clock gating

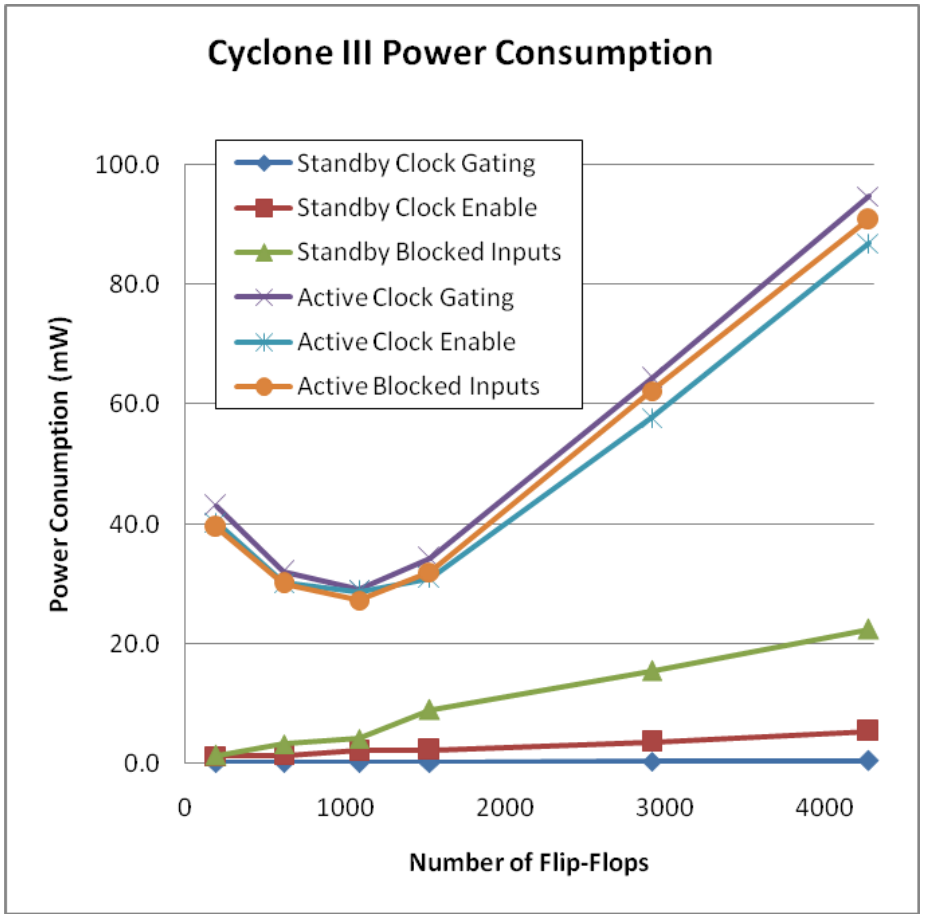





Spartan 6 Standby Power Consumption
Standby Clock Gating
Standby Clock Enable
Standby Blocked Inputs
Power Conumption (mW)
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
0
1000
2000
3000
4000
Number of Flip-Flops


Cyclone III Standby Power Consumption
Standby Clock Gating
Standby Clock Enable
Standby Blocked Inputs
Power Conumption (mW)
25.0
20.0
15.0
10.0
5.0
0.0
0
1000
2000
3000
4000
Number of Flip-Flops





J.P. Oliver, J. Curto, D. Bouvier, M. Ramos, and E. Boemo, "Clock gating and clock enable for FPGA power reduction", in Proc. 2012 VIII Southern Conference on Programmable Logic (SPL), Page(s): 1- 5. IEEE: 2012

Lógica

- Máquinas de estados
  - Asignación de estados, codificación
  - Partición de máquinas de estados
- Conmutatividad
- Profundidad lógica

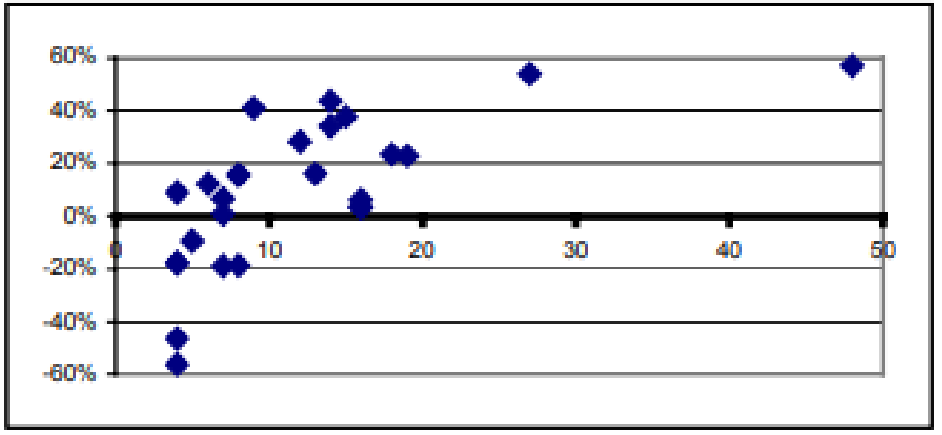
Máquinas de estados

- Asignación de estados (codificación binaria de cada estado)
  - Binario, Gray, Out-oriented, One-hot, Two-hot

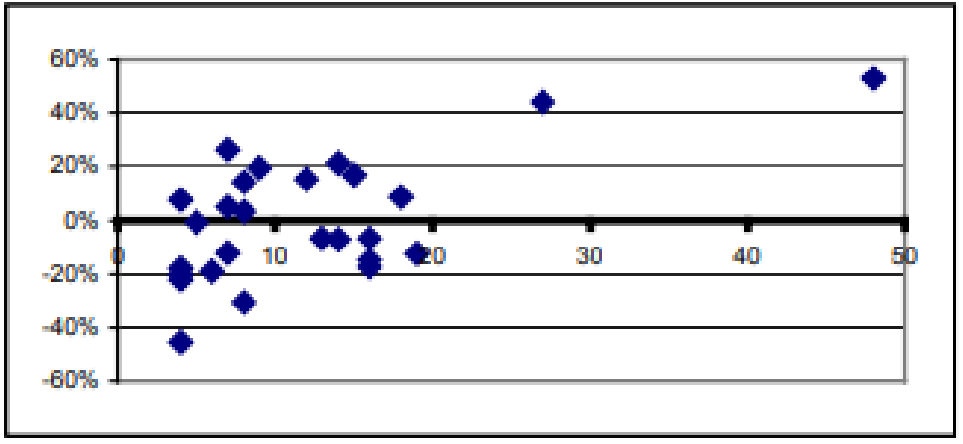
G. Sutter, E. Todorovich, S. Lopez-Buedo, and E. Boemo, "Low-Power FSMs in FPGA: Encoding Alternatives", Lecture Notes in Computer Science, Vol.2451, pp.363-370, Berlin: Springer-Verlag 2002.

Circuits	FSM characterist.				Area Bin		Area OH		Area Out-O		Area T-H		Delay (ns)				Power mW/MHz			
	inputs	outpts	rules	states	CLBs	FF	CLBs	FF	CLBs	FF	CLBs	FF	Bin	OH	Out-O	T-H	Bin	OH	Out-O	T-H
bbara	4	2	42	7	11	3	8	7	10	3	15	4	30.0	25.6	29.4	31.2	1.39	1.38	1.87	1.46
bbsse	7	7	208	13	36	4	26	13	27	4	36	5	43.1	36.2	34.6	40.1	4.02	3.37	3.14	3.43
bbtas	2	2	24	6	4	3	4	6	3	3	4	3	16.8	12.7	16.7	15.5	1.08	0.95	0.77	0.97
beecoun	3	4	20	4	7	2	10	4	7	2	12	4	21.1	18.6	16.4	28.9	1.33	1.62	1.33	2.36
cse	7	7	91	16	52	4	42	16	48	4	53	5	54.9	39.1	47.4	47.9	3.73	3.50	2.99	3.83
dk14	3	5	56	7	27	3	26	7	25	3	27	4	34.1	32.5	31.7	37.8	4.15	3.88	4.08	3.92
dk15	3	5	32	4	18	2	20	4	20	2	20	4	29.2	28.2	25.6	32.8	3.32	3.02	3.28	3.85
dk16	2	3	108	27	59	5	31	27	50	5	57	7	52.1	35.0	43.3	44.0	8.09	3.73	6.67	6.64
dk17	2	3	32	8	12	3	10	8	13	3	14	4	24.2	27.8	27.3	24.5	2.30	1.94	2.27	2.28
dk27	1	2	14	7	3	3	4	7	3	3	4	4	12.6	20.2	18.6	18.8	0.88	1.08	0.95	1.36
dk512	1	3	30	15	14	4	10	14	9	4	16	5	20.8	20.4	26.0	23.9	2.46	1.54	1.85	2.48
ex2	2	2	56	14	21	4	17	11	12	4	22	5	31.0	21.3	24.4	27.4	3.60	2.03	1.88	3.23
ex3	2	2	20	5	6	3	8	5	7	3	7	3	19.2	18.1	16.7	13.7	1.38	1.52	1.51	1.44
ex4	6	9	21	14	22	4	15	14	19	4	18	5	31.2	29.4	27.0	27.2	2.51	1.66	2.10	2.11
ex5	2	2	16	4	1	2	5	4	4	2	7	4	8.8	20.1	17.7	25.8	0.55	1.26	0.98	1.39
ex6	5	8	34	8	34	3	28	8	29	3	35	4	40.0	31.4	33.6	47.6	4.25	3.59	3.71	4.86
ex7	2	2	16	4	2	2	5	4	2	2	7	4	10.2	14.5	9.5	18.3	0.62	1.16	0.64	1.49
keyb	7	2	170	19	57	5	42	19	50	5	53	6	58.1	41.7	54.9	62.3	6.55	5.05	4.43	6.02
kirkman	12	6	370	16	45	4	43	16	45	4	57	5	38.3	36.2	38.9	36.6	4.14	4.00	3.73	5.21
lion9	2	1	16	4	2	2	2	4	2	2	5	4	8.8	15.1	8.8	25.5	0.44	0.54	0.43	1.04
mark1	5	16	180	12	19	4	15	12	17	4	17	5	30.2	24.6	24.1	30.5	2.50	1.79	2.11	2.41
opus	5	6	29	9	23	4	15	9	20	4	18	4	31.1	33.0	27.8	28.1	2.95	1.74	2.16	2.45
planet	7	19	115	48	113	6	65	48	106	6	99	10	60.6	41.3	54.3	61.1	14.4	6.23	13.2	11.7
prep3	8	8	29	8	13	3	14	8	12	3	18	4	33.3	26.9	26.5	30.9	1.66	2.04	1.42	1.99
prep4	8	8	78	16	39	4	37	16	35	4	41	5	45.9	31.4	41.5	37.7	5.47	5.29	4.37	4.92

Power Saving



One-Hot versus Binary



One-Hot versus Out-oriented

Resultados asignación de estados

- Hasta 57% reducción de consumo
- Hasta 8 estados → codificación binaria
- Más de 15 estados → One-hot

Partición de máquinas de estado

- Dividir la máquina original en máquinas chicas
- Deshabilitar las máquinas que no operan
- Al ser chicas, menos lógica combinatoria, se puede aumentar la velocidad de reloj
- Existen herramientas para hacer esto a nivel de VHDL o Verilog
- Ahorros de hasta 42% pero hay veces que empeora!

Sample	Original FSM		Partitioned One Hot Encoded				Partitioned Binary Encoded				Power Improvement
	OH	Bin	Arch1	Arch2	No Blk	Blk and	Arch1	Arch2	No Blk	Blk and	
Bbsse	3,90	4,70	3,80	3,95	4,04	4,34	3,55	3,76	4,23	3,91	9,0%
Cse	3,85	4,10	3,24	3,46	4,29	5,30	3,00	2,88	3,83	3,59	25,3%
Dk16	3,88	10,00	5,80	5,76	5,81	6,34	7,50	7,01	9,09	9,96	-32,8%
Dk512	1,84	2,80	2,46	2,79	2,44	2,14	2,24	2,51	2,16	1,94	-5,2%
Ex1	7,09	8,56	6,73	6,53	8,11	8,16	6,53	6,11	7,90	7,79	13,8%
Ex2	2,51	4,10	3,40	3,09	2,69	3,26	3,09	2,88	3,58	3,46	-6,5%
Keyb	5,50	7,06	4,73	4,31	7,88	7,69	3,66	4,65	5,25	6,81	33,4%
Kirkman	4,50	4,61	4,90	4,66	4,49	4,50	4,80	4,49	4,83	4,80	0,3%
Mark1	2,70	3,30	3,01	3,01	3,31	3,09	2,66	2,78	2,63	2,88	2,8%
Planet	8,04	16,80	9,18	9,29	10,23	10,01	10,88	11,81	15,18	16,99	-12,4%
Prep4	4,66	5,71	5,44	5,38	6,86	7,55	5,11	4,66	6,86	6,44	0,0%
S386	4,23	4,84	4,08	4,45	4,98	4,98	4,21	4,21	5,55	4,59	3,6%
S820	7,84	9,28	5,81	5,44	8,43	7,98	4,51	4,65	8,83	7,30	42,4%
S832c	7,01	10,21	5,08	5,00	7,64	6,60	4,73	5,04	7,55	6,75	32,6%

Table 2. Power consumption expressed in mW / MHz.

G. Sutter, E. Todorovich, S. Lopez- Buedo, and E. Boemo, "FSM Decomposition for Low Power in FPGA", Lecture Notes in Computer Science, Vol.2438, pp.350-359. Berlin: Springer-Verlag 2002.

Propiedad Conmutativa

- Circuitos digitales aritméticos (p.ej. multiplicador)
  - Cumplen propiedad conmutativa en el resultado
  - Pero consumen distinto

    Power (AxB) ≠ Power (BxA)

Table 5. Dynamic Power consumption of A×B and B×A in mW/MHz. 16 multiplier set.

16 bits Circuits	MaxTog		
	P (A×B)	P(B×A)	PR
Core16	7,57	5,43	28,2%
Exp16	6,42	6,98	-8,1%
Leo16	7,69	6,01	21,8%
Syn16	5,82	7,63	-23,7%
Xst16	7,21	6,06	15,9%
	AvgTog		
Core16	2,45	2,20	10,5%
Exp16	2,41	2,53	-4,5%
Leo16	2,53	2,26	10,7%
Syn16	2,18	2,37	-8,2%
Xst16	2,40	2,30	4,0%

Table 6. Dynamic Power consumption of A×B and B×A in mW/MHz. 32-bits multiplier set.

32 bits Circuits	MaxTog		
	P(A×B)	P(B×A)	PR
Core32	34,12	27,77	16,7 %
Exp32	23,81	29,39	-6,5 %
Leo32	31,40	27,87	9,3 %
Syn32	32,31	35,12	-16,4 %
Xst32	32,29	29,45	9,3 %
	AvgTog		
Core32	11,92	9,94	18,6 %
Exp32	9,56	10,22	-19,0 %
Leo32	11,70	10,62	11,3 %
Syn32	10,04	12,00	-8,0 %
Xst32	11,71	10,62	8,8 %

E. Boemo, G. Sutter, “AxB is different of BxA in terms of power consumption: Some examples on FPGAs”. In 2007 III Southern Conference on Programmable Logic, IEEE Press, 2007.

Pipelining

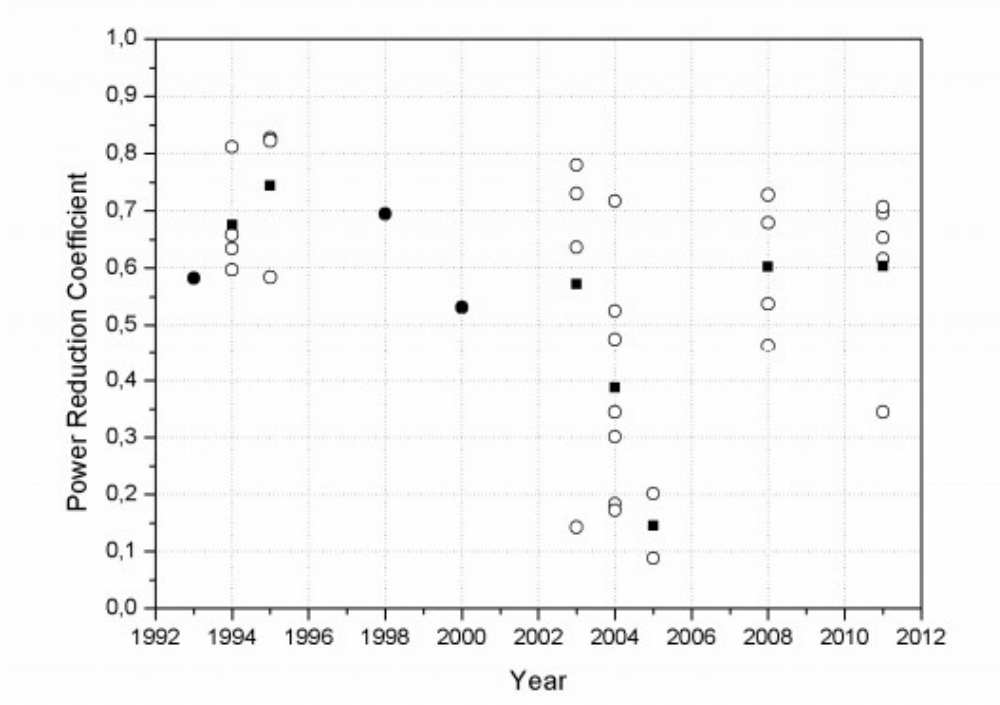


Figure 1. Summary of power reduction by pipeline in 35 published results.

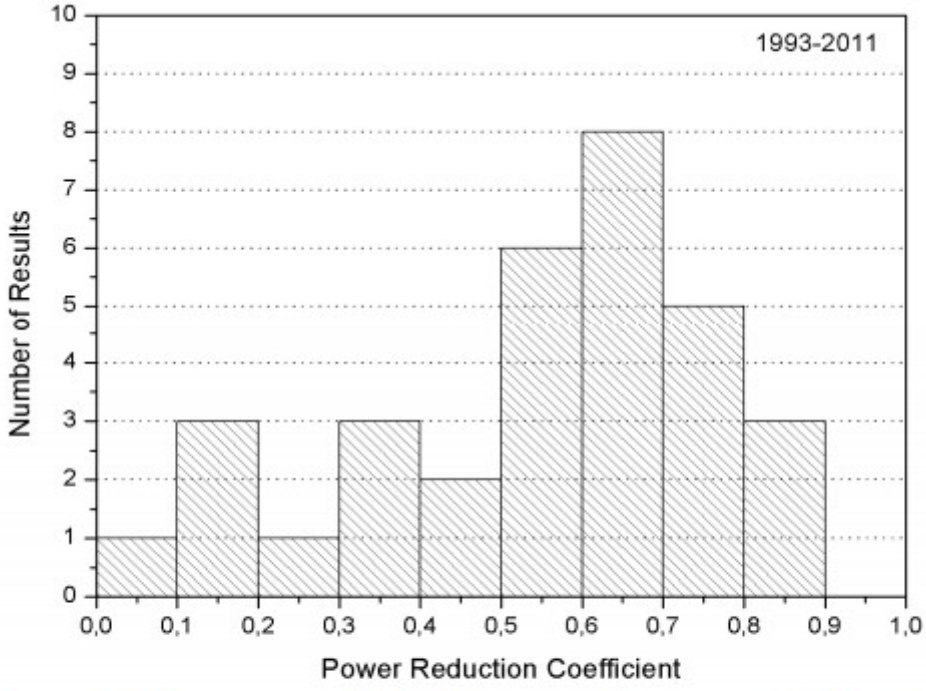


Figure 2. Histogram of the best results reported about the effect of pipeline

(All papers: simulation and measurement results)

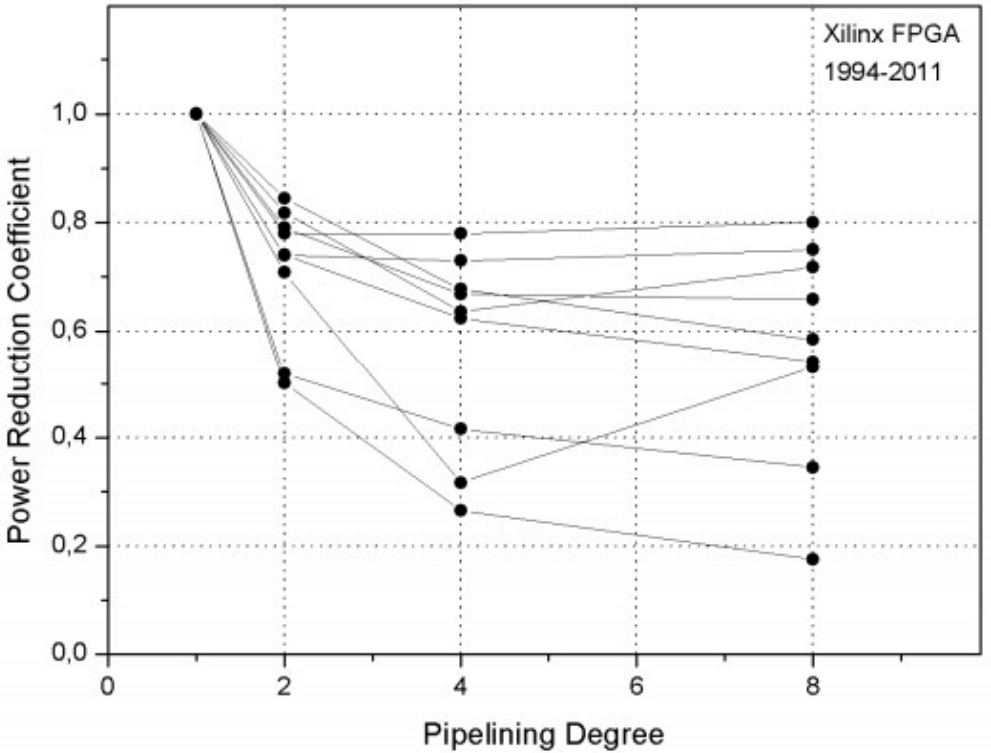


Figure. 4. Power reduction factor vs. different levels of moderated pipelining degree (Xilinx FPGAs)

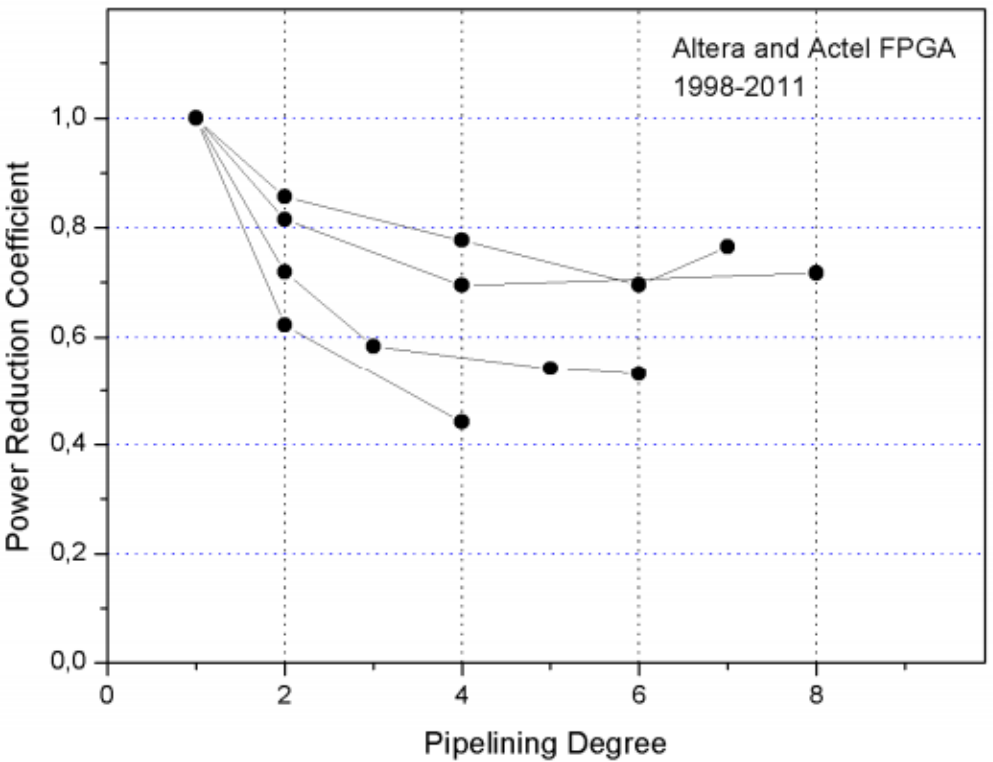


Fig. 5. Power reduction factor versus different levels of moderated pipelining degree (Altera and Actel FPGAs)

E. Boemo, J.P. Oliver, and G. Caffarena, "Tracking the Pipelining-Power Rule along the FPGA Technical Literature", Proc. ACM 2013 FPGA World, Stockholm, Sweden. September, 2013

Registrar I/O

- Poner FF fundamentalmente en las patas de salida
- Glitches con capacidades más altas que en el circuito interno
- Experimentos en Xilinx muestran que quitar el FF del pin de salida puede aumentar en consumo en un 20% al 45%

Algoritmo

- Cambios en el algoritmo puede influenciar mucho el consumo
- Ejemplo: Multiplicadores modulares
  - Operación muy usada en criptografía necesaria para calcular exponenciales
  - Dados $x < m$, $y < m$, $m < 2^n$ calcular $z = x.y \bmod m$

Algoritmo

- Se analizan 3 versiones del algoritmo
  - Multiply and Reduce
  - Shift and Add
  - Montgomery multiplier
- Versiones combinatorias y secuenciales

Resultados

Area-Time-Power of the combinational multipliers

	M_r	s_a	mont.
Energy (nJoules)	96,0	186,4	92,7
Area (CLBs)	85	157	102
Time (ns)	186	201	167

Area-Time-Power of the sequential multipliers

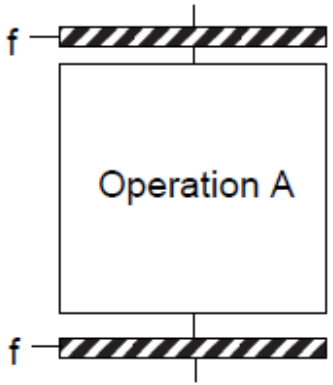
	m_r	s_a	mont.
Dynamic Energy (nJoules)	71,5	52,4	38,6
Synchronization Energy (nJoules)	46,8	26,2	27,2
Combinational Energy (nJoules)	24,7	26,2	11,1
Area (CLBs)	57	33	34
Flip - Flops	67	37	31
Total Time (ns)	320	465	249

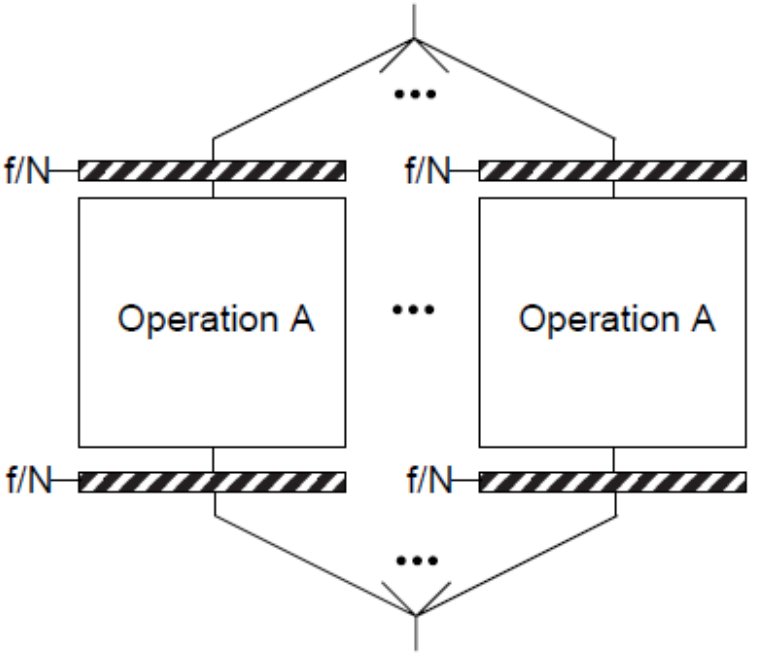
G. Sutter, J. Deschamps, and E. Boemo, "Area-Time-Power of Modular Multipliers implemented in FPGA", Jornadas de Computación Reconfigurable y Aplicaciones, JCRA 2004, Barcelona, España, 13-15 de Septiembre de 2004.

Arquitectura y sistema

- Paralelism
- Pipelining (segmentación)
- Time-Multiplexed Architectures
- Power Management
- Partitioning SW-HW
- Data representation
- Memory optimization

Reduction in clocking frequency





(a) Reference:

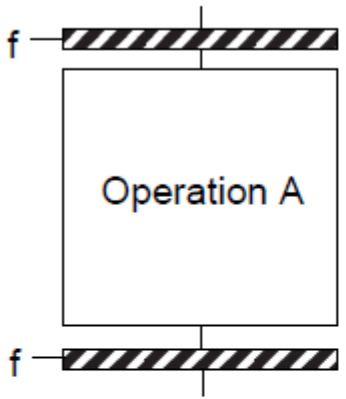
Capacitance = C
Frequency = f
Voltage = V
Power **= P ∝ CV²f**

(b) Parallel:

Capacitance ~ NC
Frequency = f/N
Voltage ~ V/N (neglecting V_t)
Power **~ P/N²**

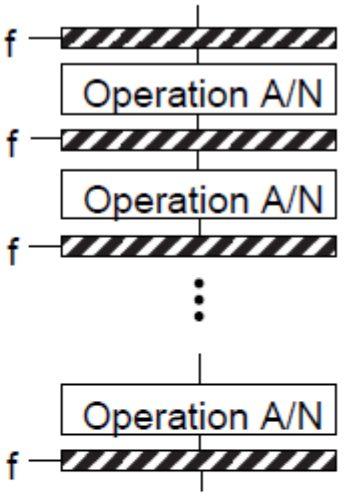
Landman, Paul Eric. "Low-power architectural design methodologies." PhD diss., University of California, Berkeley, 1994.

Voltage scaling and pipelining for low power



(a) Reference:

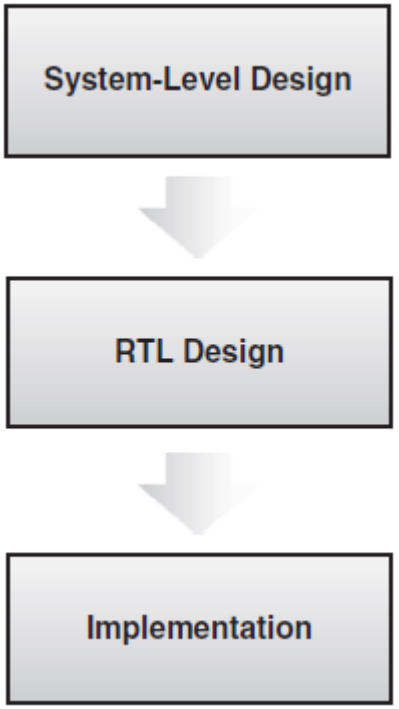
Capacitance = C
Frequency = f
Voltage = V
Power = P ∝ CV^2f



(b) Pipelined:

Capacitance ~ C
Frequency = f
Voltage ~ V/N
Power ~ P/N^2

Generalized Low-Power Design Flow
Design Phase Low-Power Design Activities



- Explore architectures and algorithms for power efficiency
- Map functions to s/w and/or h/w blocks for power efficiency
- Choose voltages and frequencies
- Evaluate power consumption for different operational modes
- Generate budgets for power, performance, area

- Generate RTL to match system-level model
- Select IP blocks
- Analyze and optimize power at module level and chip level
- Analyze power implications of test features
- Check power against budget for various modes

- Synthesize RTL to gates using power optimizations
- Floorplan, place, and route design
- Optimize dynamic and leakage power
- Verify power budgets and power delivery

¿Cómo seguir?

- Modelado y estimación
  - Modelos
  - Herramientas
- Medidas
  - Técnicas
  - Equipos
- Medidas vs Estimaciones