Logic depth and power consumption in self-timed circuits: A case-study

Eduardo Boemo
School of Computer Engineering
Universidad Autonoma de Madrid
Ctra. de Colmenar Km.15,
28049 Madrid, España.
http://www.ii.uam.es/~ivan

J. Herrera Camacho
E.U..I. Telecomunicación,
Univ. Politécnica de Madrid
Km.7 de la Ctra. de Valencia
Madrid - 28031 (España)
http://www.upm.es/centros/euitt.html

J. Meneses
E.T.S.I. Telecomunicación,
Univ. Politécnica de Madrid
Ciudad Universitaria
28040 Madrid, España.
http://www.etsit.upm.es

Abstract.- *In this paper the non-linear relationship between power consumption and throughput in two-phase self-timed (ST) pipelines is demonstrated. This effect is explained by connecting two well-known phenomena: first, the adaptation between instantaneous logic depth and data rate, an attribute inherent to two-phase self-timed pipelines; and second, the increment of data path power consumption with the logic depth, due to the influence of spurious activity. Thus, in a ST pipeline, the slope $\partial P/\partial f$ for a given frequency will be the resultant between the glitch power reduction and the datapath, synchronisation, and off-chip power increments. This effect have been validated experimentally by the construction and measurement of a self-timed pipeline multiplier on FPGAs.*

I. Introduction

The relation between logic depth and power consumption is well known [Cha92], [Lem94]. It can be explained by considering that the average dynamic power consumption of a CMOS circuit is:

$$P = \sum_{all\ nodes} c_n \ f_n \ V_{DD}^2 \qquad [1]$$

where V_{DD} is the power supply voltage, c_n the capacitance of each node, and f_n the actual frequency of each node. Thus, via f_n, the dynamic power consumption also depends on the number of glitches, or the spurious transitions of each node until it reaches the right logic state. Although glitches do not affect the final results on synchronous circuits, they can significantly contribute to the growth in the power consumption. Depending on the logic depth, the glitches generated at initial stages may produce an avalanche effect on the circuit activity.

In [She92] it is shown that the glitch power component can be up to 20% of the total chip consumption, a percentage that can reach between 30% and 70% for combinatorial circuits [Cha92]. As a consequence, a high-speed technique like pipelining also become a direct way to reduce power consumption: the intermediate registers, which are the essence of pipelining, works as barrier against the propagation of glitches, decreasing the spurious activity. This use of pipelining have been emphasised in several papers: in [Nol92] it is pointed that a careful pipelining may consume up to one order of magnitude less than the combinatorial version; in [Lei95] the consumption of pipelines with different number of the stages is analysed; in [Mus95] a novel low-power strategy based on a similar effect is developed; and finally, an study of pipeline and wave pipeline power consumption on FPGAs and 1µ Standard Cells is presented in [Boe95, 96].

In this paper, these previous works are extended to asynchronous pipelines; in particular, the Two-Phase Bundle Convention or Micropipelines scheme proposed by Sutherland [Sut89] have been applied. This strategy was prefered to other asynchronous options because its clearness and moderate hardware requirements.

Self-timed systems has been repeatedly mentioned as a low-power design alternative, mainly due to the elimination of the global clock lines as well as their stoppable clock behaviour, inherent to this synchronisation; however, to the best of our knowledge, one of the most peculiar feature of ST circuits, the non-linear relationship between power consumption and throughput, have not been reported in the scientific literature. This phenomenon is a consequence of the Sutherland protocol singularity : the logic depth is not a constant value; it depends on the data rate.

II. Instantaneous logic depth in two-phase ST systems

In a ST pipeline, each stage closes their input latches after receiving the *req* signal from previous stage, and only opens them again when the next stage sends the corresponding *ack*. Nevertheless, if the time separation between consecutive data is bigger than the pipeline period, some stages will receive the *ack* but not the next *req* signal. In this case, these stages will maintain transparent their latches, waiting for the new data. Depending on the

relation between the input data rate and the maximum pipeline frequency, all pipeline stages could be in transparent mode. In this case, the new data will pass through a combinatorial circuit, and the glitches generated on first stages are able to produce the avalanche effect on node activity.

The dependence between instantaneous logic depth and data rate is illustrated on Figs. 1, 2 and 3. They correspond to a post-layout simulation of an 8-bit Guild array multiplier prototype, pipelined every two cells, designed using 1µ Standard Cells of ES2. Three different situations are shown: 1) an input data rate close to the maximum pipeline frequency ; 2) an intermediate input data rate; and 3) a data rate several times longer than the maximum combinatorial array delay. The chronographs show the signals that control the latches; they allow the number of successive transparent stages between two locked line of latches to be observed.

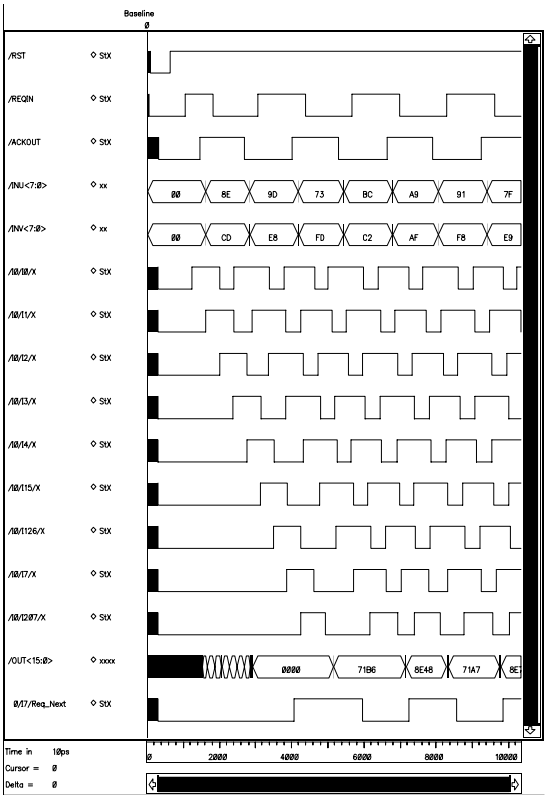

Fig. 1: The signals named with X, control the latches of successive stages (the latches are transparent when "0" and locked when "1"). In this case, the frequency is close to maximum pipeline speed ; thus, the usual behaviour of a non-overlapped two-phase pipeline can be observed.

The consequence of the preceding analysis in terms of power consumption is remarkable. Considering that an increment of frequency produces a logic depth reduction, and assuming that a logic depth reduction also diminishes the spurious activity, the power consumption of ST pipelines will not present a linear dependence with the frequency operation. The slope $\partial P/\partial f$ for a given frequency will be the resultant between the glitch power reduction and the datapath, synchronisation, and off-chip power increments.

3. Experimental Results

In order to demonstrate the thesis of this paper, a model of the Guild ST multiplier was constructed and measured using a XC3090PC84-100 FPGA [Xil95]. The prototype made use of 310 CLBs, an excessive value because the latches were implemented using LUTs. Each asynchronous control block (ACB) is composed of a Müller-C cell, an ordinary XOR gate, and a Toggle cell [Sut89]. Considering that each ACB needs three cascaded CLBs to be implemented, its size limits the practical minimum pipeline logic depth. The partitioning process into CLBs was done manually, meanwhile the placement-routing was performed automatically by using the default tool parameters. The maximum speed (measured) resulted close to

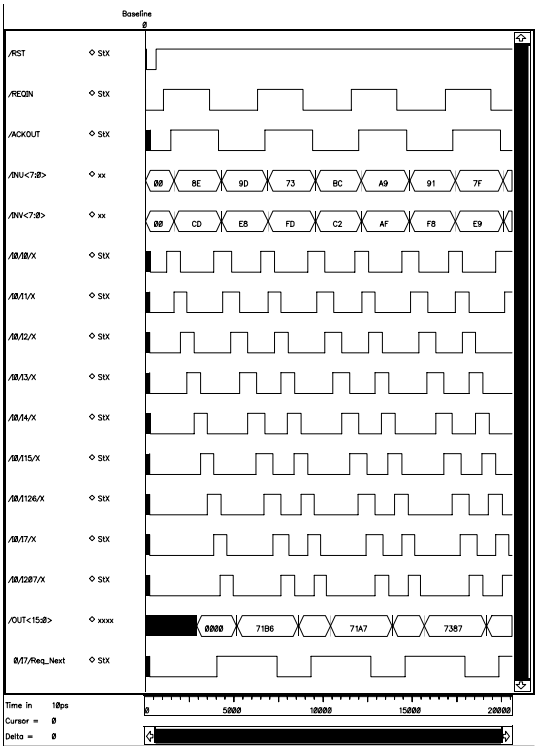

Fig.2 : In this case, the data rate is lower that the pipeline period (note that the time scale has been compressed in relation with the previous figure). Now, several successive latches are simultaneously transparent, producing an increment of the actual logic depth. For this example, the logic depth vary between 5 and 6 pipeline stages.

10 MHz, a modest value for the pipeline grain adopted. For the chip selected, the maximum speed of a ST application was 17 MHz, corresponding to a 16x4 FIFO memory prototype.

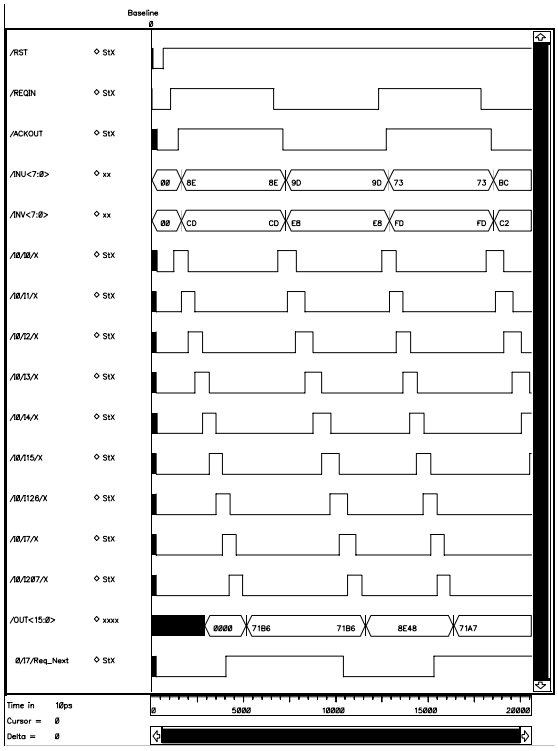

Fig.3 : In this chronograph, the input data period is several times lower than the maximum circuit propagation delay. In consequence, every latches remain transparent waiting for a new data. So, all the pipeline is full combinatorial , and there is no glitch barrier along the circuit.

In order to test the circuit, a conventional pattern generator was utilised. A set of input vectors, each of them followed of an event in the input *req* line, were send in a synchronous way; that is, without taking into account the state of the *ack* input signal. In this case, if the *req* period is longer than the phase between the *ack* and *req* signals, the chip will work without error. Although this method has the same drawback of clocked systems (a security death-time zone must be added), it allowed synchronous instruments to be applied. At the output, to reply every *req* event by generating the corresponding output *ack* where also mandatory. Otherwise, a dead-lock is propagated into the pipeline. This problem was solved by implementing a feedback (whose delay is equal to one pipeline stage) between the pair *req-ack* of the last stage. The same idea was applied during the design phase in order to use a synchronous simulator.

In Figure 4, the average power consumption (measured) versus input data rate is shown. Several different slope regions can be identified. At low-frequency, all the circuit exhibits a combinatorial behaviour ; thus the power consumption increases linearly with the frequency. However, near 0.8 MHz some stages start working in a two-phase fashion, blocking the propagation of glitches ; it causes the later power consumption reduction. The same effect appears near 3.5 MHz, and finally, at 4.2 MHz. From 6.5 MHz on, the circuit starts working at their maximum pipeline degree, and the power consumption again grows linearly with the frequency.

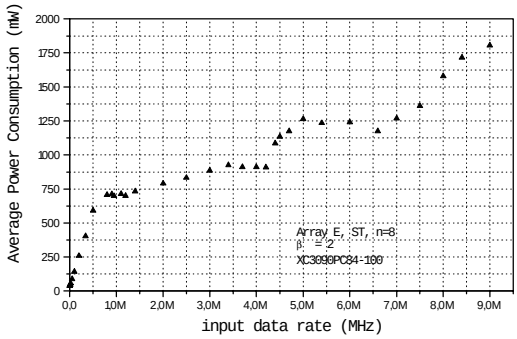

Fig.4: Average power consumption measured versus input data rate.

4. Conclusions

In this paper, the non-linear relationship between power consumption and input data rate on two-phase self-timed pipelines has been explained. A prototype has been constructed and measured in order to validate the arguments. The experiment shows that the power consumption on a self-timed pipeline may even decrease when the operation frequency increases; this occurs in those cases where the reduction in the spurious power (produced by the smaller logic depth) is larger than the increment in other power components.

The results are similar to those previously detected in a set of synchronous arrays pipelined with different granularities [Boe95]; nevertheless, the ST synchronisation has the singularity of presenting different slopes in the same circuit.

An important limitation of this study is that the magnitude of the effect is strongly dependent of the technology and topology used: both factors can magnify or mask the weight of spurious activity. For example, the effect is highlighted in FPGAs, but could result insignificant of Standard Cell, a technology in which synchronisation and off-chip power are higher than datapath power [Boe96].

Acknowledgements

This work has been supported by the CICYT of Spain under contract TIC95-0791. The authors wish to thanks Martina Eckert, Catrin Dramsh, and Sergio López Buedo for their thorough revision of this paper.

References

[Boe95] E. Boemo, G. González de Rivera, S. López and J. Meneses, "On the Usefulness of Pipelining and Wave Pipelining as Low-Power Design Technique", *Proc. PATMOS´95, Fifth Int. Workshop,* pp.252-263, Oldenburg, October 1995.

[Boe96]] E. Boemo, S. López and J. Meneses, "Logic Depth and Power Consumption : A Case-Study on 1u CMOS Standard Cells", Internal Report, Dept of E. Engineering, UPM-Madrid, 1996.

[Cha92] A. Chandrakasan, S. Sheng y R. Brodersen, "Low-Power CMOS Digital Design", *IEEE Journal of Solid-State Circuits*, Vol.27, Nº4, pp.473-484. April 1992.

[Gui69] H Guild, "Fully Iterative Fast Array for Binary Multiplication and Addition", *Electronic Letters*, pp.263, Vol.5, No.12, June 1969.

[Lem94] Z. Lemnios y K. Gabriel, "Low-Power Electronic", *IEEE Design & Test of Computers*, pp.8-13. Winter 1994.

[Liu94] D. Liu y C. Svensson, "Power Consumption Estimation in CMOS VLSI Circuits". *IEEE Journal of Solid-State Circuits*, Vol.29, N°6, pp. 663-670, Jun.1994.

[Nol92] T.G. Noll, "Pushing the Performance Limits due to Power Dissipation of future ASICs", *Int. Symposium on Circuits and Systems*, pp.1652-1655. IEEE Press, 1992.

[She92] A. Shen, A. Gosh, S. Devadas y K. Keutzer, "On average Power Dissipation and Random Pattern Testability of CMOS Combinatorial Logic Networks", *Proc. ICCAD-92 Conf*, pp.402-407. IEEE Press, 1992.

[Sut89] I. Sutherland, "Micropipelines", *Communication of the ACM*, vol.22, nº6, pp.720-734. Jun. 1989.

[Xil95] *The Programmable Logic Array Data Book*, Third Edition, Xilinx Inc. 1995.