Statistical Timing Analysis in Sequential Circuit for On-Chip Global Interconnect Pipelining [*]

Lizheng Zhang
ECE Department
University of Wisconsin
Madison, WI 53706-1691
lizhengz@cae.wisc.edu

Yuhen Hu
ECE Department
University of Wisconsin
Madison, WI 53706-1691
hu@engr.wisc.edu

Charlie, Chungping Chen
EE Department
National Taiwan University
Taipei 106, Taiwan.
cchen@cc.ee.ntu.edu.tw

ABSTRACT

With deep-sub-micron (DSM) technology, statistical timing analysis becomes increasingly crucial to characterize signal transmission over global interconnect wires. In this paper, a novel statistical timing analysis approach has been developed to analyze the behavior of two important pipelined architectures for multiple clock-cycle global interconnect, namely, the flip-flop inserted global wire and the latch inserted global wire. We present analytical formula that is based on parameters obtained using Monte Carlo simulation. These results enable a global interconnect designer to explore design trade-offs between clock frequency and probability of bit-error during data transmission.

Categories and Subject Descriptors

B.8 [**Hardware**]: Performance and Reliability

General Terms

Algorithms, Performance, Design, Reliability

Keywords

Statistical Timing Analysis, Interconnect Pipelining

1. INTRODUCTION

In this paper, we performed statistical timing analysis on two types of pipelined, multiple clock-cycle, global interconnect architectures: a flip-flop inserted global interconnect wire [1–7], and a latch inserted global interconnect wire. Our goal is to analyze the effects of parameter variation and noise on the successful transmission of data in such a circuit. Specifically, we assume that the statistical variation of individual parameters, such as clock skews and jitter,

process parameters, and noise statistics are known. Using these statistics, we are able to derive analytical formula that characterize the distribution of timing delay of these multiple clock-cycle global interconnect wires. These distributions then allow one to estimate the probability of erroneous transmission of a single bit through the wire at a particular clock frequency. By varying this expected operating clock frequency, a bit-error-rate versus clock frequency plot can be drawn for each of these two global interconnect architectures. A plot for both the flip-flop inserted and the latch inserted global interconnect system allows one to compare the different features of these two architectures.

The results presented in this paper differ from these earlier results and are significant in several aspects:

- Existing statistical timing analysis results [10, 12–19] focused on combinational logic. The nonlinear effect of sequential elements such as flip-flops or latches have not been investigated

- An explicit, closed-form expression of the probability density functions (*p.d.f.*s) of data arrival time has been derived for both flip-flop-pipelined global interconnect and latch-pipelined global interconnect.

- These *p.d.f.*s facilitate the estimation of the bit error rate (BER) of each type of global interconnect and therefore mathematically characterize the performance of these communication channels.

- By evaluating the BER against different expected operating frequency, designer is able to explore the design space such as the number of pipelined stages, wire segment lengths, etc. to optimize the global interconnect design.

2. DFF-PIPELINED WIRE TIMING

A typical DFF-based interconnect pipelining stage is shown in Figure 1. The positive-edge-triggered DFF has a size s, the wire length is l, the source driver has a size z, and the receiving driver has a size w.

We denote τ_{setup} to be the set up time of a DFF, τ_{prop} to be the propagation delay from D to Q after the positive clock edge; and τ_{wire} to be the propagation delay from the output of DFF at stage i to the input D of DFF at stage $i + 1$. For the DFF at the $(i + 1)^{th}$ stage to properly latch on a data bit, the propagation delay

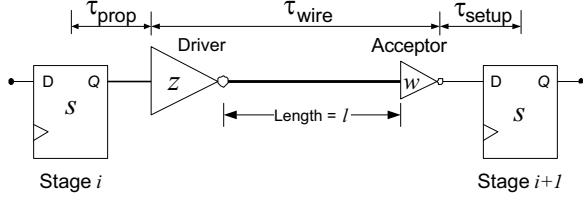$$p_i = \tau_{prop} + \tau_{wire} \tag{1}$$

Figure 1: DFF Pipelined Interconnect Stage

must satisfy a timing constraint

$$0 \le p_i \le T_{CLK} - \tau_{setup} \quad (2)$$

In other words, if a data bit has been already correctly registered by the DFF in i^{th} stage, the probability to have correct data transmission between the ith and $(i+1)^{th}$ stage can be expressed as:

$$q = Pr(0 \le p_i \le T_{CLK} - \tau_{setup}) \quad (3)$$

We model the data arrival time p_i as a random variable with probability density function ($p.d.f.$) $P(p_i)$. Noted that from equation (1), p_i is the sum of two independent random variables, τ_{prop} and τ_{wire} with $p.d.f.$s denoted by $P(\tau_{prop})$ and $P(\tau_{wire})$ respectively. Hence $P(\tau_{stage})$ can be found as the convolution of $P(\tau_{prop})$ and $P(\tau_{wire})$:

$$P(p_i) = P(\tau_{prop}) ** P(\tau_{wire}) \quad (4)$$

where "$**$" is the convolution operator.

We assume that the clock period T_{CLK} and the setup time of the DFF, τ_{setup} are random variables. Hence, $A = T_{CLK} - \tau_{setup}$ will also be a random variable with $p.d.f. P(A) = P(T_{CLK}) ** P(-\tau_{setup})$. Thus,

$$q = \int_{-\infty}^{\infty} \left\{ \int_{0}^{T_{CLK}-\tau_{setup}} P(p_i) dp_i \right\} P(A) dA \quad (5)$$

Since p_i is the sum of timing delays, the probability of the event $p_i < 0$ is zero. Hence above equation can be rewritten as:

$$q = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{T_{CLK}-\tau_{setup}} P(p_i) dp_i \right\} P(A) dA \quad (6)$$

where the lower bound of integration is extended from 0 to $-\infty$. This makes it possible to define another variable $\delta = p_i + \tau_{setup} - T_{CLK}$ with a $p.d.f.$ $P(\delta) = P(\tau_{prop}) ** P(\tau_{wire}) ** P(\tau_{setup}) ** P(-T_{CLK})$ such that

$$q = \int_{-\infty}^{\infty} P(A) dA \int_{-\infty}^{0} P(\delta) d\delta = \int_{-\infty}^{0} P(\delta) d\delta \quad (7)$$

The error probability that a single data bit is transmitted through a N-stage DFF-pipelined global interconnect wire, denoted by BER, is the bit error rate when data is transmitted through this on-chip communication channel. Due to the presence of a DFF, the probability of correct data transmission at each stage is independent of each other. Hence,

$$BER = 1 - q^N \quad (8)$$

Assuming that all timing variables τ_{prop}, τ_{wire}, τ_{setup}, and T_{CLK} have normal distributions, then δ will also have a normal density function with

$$\mu_\delta = \mu_{\tau_{prop}} + \mu_{\tau_{wire}} + \mu_{\tau_{setup}} - \mu_{T_{CLK}} \quad (9)$$

$$\sigma_\delta^2 = \sigma_{\tau_{prop}}^2 + \sigma_{\tau_{wire}}^2 + \sigma_{\tau_{setup}}^2 + \sigma_{T_{CLK}}^2 \quad (10)$$

Hence

$$q = Pr(\delta \le 0) = \frac{1}{2} + erf\left(\frac{\mu_\delta}{\sigma_\delta}\right) \quad (11)$$

where $erf(x) = \frac{1}{\sqrt{2\pi}} \int_0^x exp\left(-\frac{t^2}{2}\right) dt$

3. LATCH-PIPELINED WIRE TIMING

A typical latch-pipelined interconnect stage is shown in Figure 2. Again, we assume the positive level-triggered latch has a size s, the wire length is l, the sending driver size is z and the receiving driver size is w.
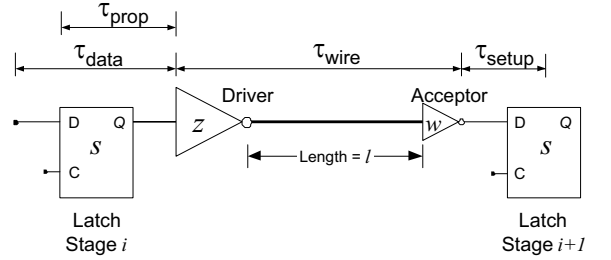

Figure 2: Latch Pipelined Interconnect Stage

There are three delays associated with a latch: (i) τ_{data}, the delay between data input terminal D and output terminal Q when clock is high. (ii) τ_{prop}, the delay between positive clock edge and Q. (iii) τ_{setup}, the time data signal must held steady before the falling edge of the clock signal so that the latch may latch on to the correct data value. We also denote by τ_{wire} the delay incurred over the wire segments within a pipelined stage.
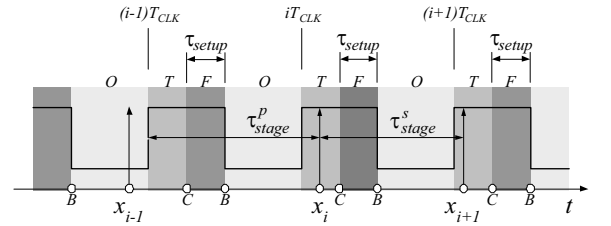

Figure 3: Latch Timing

In figure 3, the timing constrains of a latch-piplined stage are illustrated. The square wave in solid line represents the ideal clock waveform with a period equal to T_{CLK}. The positive clock edge is designated as the origin of each clock period that extends over $\{t | (i-0.5)T_{CLK} \le t \le (i+0.5)T_{CLK}\}$. We assume 50% duty cycle so that the width of the clock pulse is $0.5T_{CLK}$. We use point B to mark the falling edge of clock pulse, and point C to mark the point $0.5 \cdot T_{CLK} - \tau_{setup}$ within each clock cycle.

We denote x_i as the arrival time of a single bit data to stage i and $p_i = x_i - i \cdot T_{CLK}$. If the data arrives within τ_{setup} time units prior to the falling clock edge B, then the D latch may not have sufficient time to register this new data, and the data transmission will fail. Hence the interval $F = \{t | C = 0.5T_{CLK} - \tau_{setup} < t < 0.5T_{CLK}\}$ is denoted as the *faulty region* within each clock period.

As illustrated in figure 3, x_{i-1} falls inside an *opaque region* $O = \{t | -0.5 \cdot T_{CLK} \le t \le 0\}$ prior to the rising clock edge.

During this interval, the clock signal is low, and the latch is off. Thus, the latch appears to be *opaque* to the data signal at its input D. The data will remain there until the rising clock edge at the end of this interval. Then it will propagate through the latch and wire in pipelined stage i and arrive the output of the acceptor at time x_i. Thus $x_i = (i-1)T_{CLK} + \tau_{prop} + \tau_{wire}$ Or equivalently, $p_i = -T_{CLK} + \tau_{prop} + \tau_{wire}$. Note that p_i is independent of p_{i-1}.

In figure 3, x_i arrives after the rising clock edge ($t = 0$) and falls within a *transparent region* $T = \{t | 0 \leq p_i \leq 0.5T_{CLK} - \tau_{setup}\}$. Since the latch is on, it appears *transparent* to the input signal at D. The incoming data immediately starts passing through the latch, subsequent buffers, and wire. Hence $x_{i+1} = x_i + \tau_{data} + \tau_{wire}$, and $p_{i+1} = p_i + \tau_{data} + \tau_{wire} - T_{CLK}$. To summarize,

$$p_{i+1} = \begin{cases} \tau_{prop} + \tau_{wire} - T_{CLK} & if\, p_i \in O; \\ p_i + \tau_{data} + \tau_{wire} - T_{CLK} & if\, p_i \in T. \end{cases} \quad (12)$$

If $p_i \in F$, then the data transmission is considered as a failure and there will be no p_{i+1}.

Note that when $p_i \in T$, p_{i+1} is dependent on p_i. This is not the case of a DFF pipelined global interconnect where the delay within each pipelined stage is independent of other pipelined stages.

In order to calculate the probability of successful transmission of a single data bit through a N-stage D latch-pipelined global interconnect, one needs to compute the *p.d.f.* of the data bit arrival time $P(p_i)$ for each pipelined stage. This task is much more complicated for a latch-pipelined global wire than a DFF-pipelined global wire, because p_i may depend on p_{i-1} if $p_{i-1} \in T$, or p_i may not be defined if $p_{i-1} \in F$. Therefore, we will derive a formula for a conditional *p.d.f.* $P(p_{i+1}|p_i \in O \cup T)$ instead. Since $\{p_i \in O\}$ and $\{p_i \in T\}$ are disjoint events, one has

$$P(p_{i+1}|p_i \in O \cup T) \quad (13)$$

$$= \frac{Pr(p_i \in O)}{Pr(p_i \in O) + Pr(p_i \in T)}P(p_{i+1}|p_i \in O)$$

$$+ \frac{Pr(p_i \in T)}{Pr(p_i \in O) + Pr(p_i \in T)}P(p_{i+1}|p_i \in T)$$

where $P(p_{i+1}|p_i \in O) = P(\tau_{prop}) ** P(\tau_{wire}) ** P(-T_{CLK})$; $P(p_{i+1}|p_i \in T) = P(p_i|p_i \in T) ** P(\tau_{data}) ** P(\tau_{wire}) ** P(-T_{CLK})$ and

$$P(p_i|p_i \in T) = \begin{cases} 0 & (p_i \notin T) \\ \dfrac{P(p_i|p_{i-1} \in O \cup T)}{Pr(p_i \in T)} & (p_i \in T) \end{cases} \quad (14)$$

Given the initial *p.d.f.*, $P(p_0)$ as the input of the first stage, then, recursively, $P(p_{i+1}|p_i \in O \cup T)$ can be computed once $P(p_i|p_{i-1} \in O \cup T))$ is computed.

For $i > 1$, given that the first $i - 1$ stages have successfully transmitted a one-bit data, the probability that at the ith stage the data bit is successfully delivered to the $(i + 1)^{th}$ state is:

$$q_i = Pr(p_i \in O \cup T | p_{i-1} \in O \cup T) \quad (15)$$

$$= \iint_{-\infty}^{+\infty} \left[\int_B^C P(p_i|p_{i-1} \in O \cup T)dp_i \right] P(B)P(C)dBdC$$

where random variables $B = -0.5T_{CLK}$, $C = 0.5T_{CLK} - \tau_{setup}$ and their *p.d.f.*s are $P(B) = P(-0.5T_{CLK})$ and $P(C) = P(0.5T_{CLK}) ** P(-\tau_{setup})$ respectively. For $i = 1$,

$$q_1 = \iint_{-\infty}^{+\infty} P(B)P(C)dB\, dC \quad (16)$$

To deliver a data bit through N latch-pipelined stages without error requires that it is delivered correctly at each of the N stages. Hence, the overall bit error rate(BER) can be evaluated as:

$$BER = 1 - \underbrace{q_1 \cdot q_2 \cdots q_N}_{N} = 1 - \prod_{i=1}^{N} q_i \quad (17)$$

4. EXPERIMENTS AND RESULTS

We performed Monte Carlo simulation based on $0.18\mu m$ technology parameters to estimate the *p.d.f.*s of timing parameters τ_{setup}, τ_{data}, τ_{prop}, and τ_{wire}. The nominal design parameters used are: $s = 16\mu m$, $w = 24\mu m$, $z = 40\mu m$. The latches are designed such that $\tau_{setup} = \tau_{data}$. The random wire delay has a mean value μ_w that is a function of the length of the wire segment l. The standard deviation of the wire delay, σ_w is assumed to be a constant. The resulting timing parameters' *p.d.f.* are described in table 1 below:

Table 1: Distribution Parameters for Delay Elements

	μ		σ	
	DFF	Latch	DFF	Latch
τ_{setup}	94ps	119ps	9ps	13ps
τ_{prop}	120ps	116ps	13ps	13ps
τ_{wire}	μ_w		9ps	
T_{CLK}	μ_c		17ps	

Given the BER estimate, the expected throughput rate will be lower, and can be represented as:

$$Thput = \frac{1 - BER}{\mu_c} \quad (18)$$

If error correcting code (ECC) is inserted, the attainable throughput rate will be further compromised.

In Figure 4, the STA analysis results of the BER and throughput of both a DFF based global interconnect wire and a latch based global interconnect wire are plotted. This pipelined global interconnect wire has 8 pipelining stages with the legnth of each wire segment equal to 1.4mm($\mu_w = 110ps$).

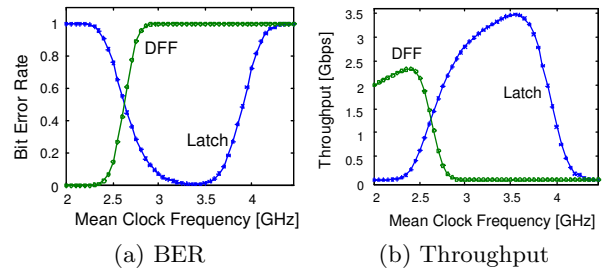(a) BER   (b) Throughput

Figure 4: BER and Throughput v.s. Mean Clock Frequency

Shown on the left of Figure 4(a), the BER of a DFF pipelined global interconnect grow swiftly after the clock frequency exceeds a threshold. The data indicates that the BER is maintained below 1% if the clock frequency is kept below 2.34 GHz.

Interestingly, the latch-based pipeplined global interconnect has a BER that is high when the clock frequency is either too high or too low. When clock frequency becomes higher, BER will also increase is expected. However, when the clock freuqency is too low, the transparency duration T increases; while the propagation delay remain unchanged. As such, a data bit may pass through more than one pipelined stage, causing timing error. From the data, we observe that within a range of clock frequencies $3.34GHz \leq 1/\mu_c \leq 3.44GHz$, the BER is kept below 1%.

To the right side of Figure 4(b), the expected throughput rate are calculated using equation 18. It is seen that the the expected throughput rate of the DFF pipelined interconnect wire reaches a maximum value of 2.35 Gbps; while the latch-pipelined interconnect wire reaches a maximum value of 3.38 Gbps; representing a 44% improvement in throughput. Clearly from a statistical timing analysis point of view, the latch-based interconnect wire has noticeable throughput advantage.

As a comparison, using worst-case timing analysis on the DFF pipelined global interconnect, it is estimated that the maximum operating clock frequency will be 2.2 GHz to ensure the corresponding BER will be less than 1%. Compared to the $2.35GHz$ estimate, it is about 7% off the mark.

5. REFERENCES

[1] H. Shah, P. Shin, B. Bell, M. Aldredge, N. Sopory, and J. Davis, "Repeater insertion and wire sizing optimization for throughput-centric vlsi global interconnects," *IEEE/ACM International Conference on Computer Aided Design(ICCAD 2002)*, pp. 280 –284, 2002.

[2] K. Banerjee and A. Mehrotra, "A power-optimal repeater insertion methodology for global interconnects in nanometer designs," *IEEE Transactions on Electron Devices*, vol. 49, no. 11, pp. 2001 –2007, Nov 2002.

[3] S. Srinivasaraghavan and W. Burleson, "Interconnect effort -a unification of repeater insertion and logical effort," *Proceedings of IEEE Computer Society Annual Symposium on VLSI*, pp. 55 –61, Feb 2003.

[4] C. Chu and D. F. Wong, "Closed form solutions to simultaneous buffer insertion/sizing and wire sizing," *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 6, no. 3, pp. 343–371, July 2001.

[5] L. Scheffer, "Methodologies and tools for pipelined on-chip interconnect," *IEEE International Conference on Computer Design(ICCD)*, 2002.

[6] P. Cocchini, "Concurrent flip-flop and repeater insertion for high performance integrated circuits," *IEEE/ACM International Conference on Computer Aided Design(ICCAD)*, pp. 268 –273, 2002.

[7] R. Lu, G. Zhong, C. K. Koh, and K. Y. Chao, "Flip-flop and repeater insertion for early interconnect planning," *Proceedings of Design, Automation and Test in Europe Conference and Exhibition, 2002*, pp. 690 –695, March 2002.

[8] I. Lin, J. A. Ludwig, and K. Eng, "Analyzing cycle stealing on synchronous circuits with level sensitive latches," *29th ACM/IEEE Design Automation Conference*, 1992.

[9] B. Taskin and I. S. Kourtev, "Performance optimization of single-phase level sensitive circuits using time borrowing and non-zero clock skew," *TAU 2002*, Dec 2002.

[10] J. J. Liou, K. T. Cheng, S. Kundu, and A. Krstic, "Fast statistical timing analysis by probabilistic event propagation," *Design Automation Conference, Proceedings*, pp. 661–666, June 2001.

[11] J. J. Liou, C. L. Wang, A. Krstic, and K. T. Cheng, "Experience in critical path selection for deep sub-micron delay test and timing validation," *Design Automation Conference, 2003. Proceedings of the ASP-DAC 2003. Asia and South Pacific*, pp. 751 –756, Jan 2003.

[12] A. Agarwal, V. Zolotov, and D. Blaauw, "Statistical timing analysis using bounds and selective enumeration," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 9, pp. 1243 –1260, Sept 2003.

[13] B. Choi and D. Walker, "Timing analysis of combinational circuits including capacitive coupling and statistical process variation," *VLSI Test Symposium, 2000. Proceedings. 18th IEEE*, pp. 49–54, May 2000.

[14] H. Jyu, S. Malik, S. Devadas, and K. Keutzer, "Statistical timing analysis of combinational logic circuits," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 1, no. 2, pp. 126 –137, June 1993.

[15] R.-B. Lin and M.-C. Wu, "A new statistical approach to timing analysis of vlsi circuits," *VLSI Design, 1998. Proceedings., 1998 Eleventh International Conference on*, pp. 507 –513, Jan 1998.

[16] R. Brawhear, N. Menezes, C. Oh, L. Pillage, and M. Mercer, "Predicting circuit performance using circuit-level statistical timing analysis," *European Design and Test Conference, 1994. EDAC, The European Conference on Design Automation. ETC European Test Conference. EUROASIC, The European Event in ASIC Design, Proceedings.*, pp. 332 –337, Mar 1994.

[17] S. Bhardwaj, S. B. Vrudhula, and D. Blaauw, "τau: Timing analysis under uncertainty," *ICCAD'03*, pp. 615–620, Nov 2003.

[18] H. Chang and S. S. Sapatnekar, "Statistical timing analysis considering spatial correlations using a single pert-like traversal," *ICCAD'03*, pp. 621–625, Nov 2003.

[19] A. Devgan and C. Kashyap, "Block-based static timing analysis with uncertainty," *ICCAD'03*, pp. 607–614, Nov 2003.

[20] N. Kurd, J. Barkatullah, R. Dizon, T. Fletcher, and P. Madland, "A multigigahertz clocking scheme for the pentium(r) 4 microprocessor," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 11, pp. 1647 –1653, Nov. 2001.