Yield Prediction of High Performance Pipelined Circuit with Respect to Delay Failures in sub-100nm Technology

Animesh Datta, Saibal Mukhopadhyay, Swarup Bhunia, and Kaushik Roy
School of Electrical and Computer Engineering, Purdue University, IN-47906, **USA**
{adatta, sm, bhunias, kaushik}@ecn.purdue.edu

Abstract

*In **nano-scaled technology, large variations in process parameters produces wide delay spread in high performance circuit. In this paper we develop analytical models for yield prediction with respect to delay variation of pipeline design. We have addressed the converse problem of estimating the design space for individual pipe stages based on a target yield. For an example 4 stage pipelined circuit proposed analytical models are verified to predict yield within 2% of results obtained from Monte-Carlo Hspice simulation.***

1. Introduction

The profitability of high performance system design and manufacturing strongly depends on the fabrication yield, defined as the ratio of operational chips to the total number of fabricated chips [1]. Traditionally high performance circuits are designed with specific target frequencies. However, due to process variations the operating frequency of individual units may vary substantially [4]. In deep submicron technologies, uncertainties in device and interconnect characteristics (e.g. effective gate length (L), oxide thickness (T_{ox}), random placement of dopant atoms), die location dependence, and line edge roughness often lead to heterogeneous and non-monotonic relationships among the process parameters. This produces large spread in gate delays and hence results substantial circuit delay variation (which determines the highest operating clock frequency). Consequently, estimating circuit performance and designing high performance systems with high ***design yield*** (probability that the design will meet certain delay target) under parameter variations have emerged as serious design challenges in sub-100nm technology [4]. In our subsequent discussions we will consider yield of a design as its design yield in terms of delay failures.

The correctness of circuit performance estimation largely depends on the accuracy of the delay model used for circuit design. Static delay models no longer seem accurate in estimating circuit delay in sub-100nm technology nodes, since it does not model the uncertainty due to parameter variations. Statistical delay models, based on probabilistic timing analysis (PTA), on the other hand, have emerged as an efficient alternative. A statistical delay model has the capacity to model the uncertainty in circuit delay evaluation due to parameter fluctuations. There are several other factors that bring uncertainty in the delay analysis such as inaccuracy in parasitic prediction, clock slew, model-to-hardware correlations errors etc. In statistical timing analysis, delays of the circuit elements (gate and interconnect) are random variables usually Gaussian with a particular mean and standard deviation [10]. Hence the delay of a circuit is a complex function **of** its components delay parameters.

In this paper we analyze and model the statistical delay variation of a pipeline. We present detailed mathematical analysis for the effect of statistical distribution of individual stage delays on the overall pipeline yield. Our analysis considers inter-die, within-die variations and correlations among different delay variations. Our proposed delay model is verified to predict design yield within 2% of Monte-Carlo ***Hspice*** simulation results. In particular this paper makes the following contributions:

a) An analytical model to estimate the yield of a pipelined circuit, given the individual stage delay distributions and correlation parameters for inter- and within-die device parameter variations.

b) Estimation of feasible design space with respect to mean (μ) and standard deviation (STD or a) of individual stage delays, given a target yield and a frequency constraint. The feasible design space is specified in terms of upper and lower bounds on mean and standard deviation of the stage delays. Design choices within this space represent trade-off between area/power and probability of achieving target frequency.

Rest of the paper is organized as follows. In section **2** we build analytical models for delay and yield computation for a pipelined circuit. Section 3 explores feasible pipeline design space that meets target yield. In section 4 we present various design trade-offs of pipeline design and effect of unbalancing stage delays on the overall design yield. Finally section 5 concludes the paper.

2. Statistical modeling of pipeline yield and reliability analysis

Delay failures are a major source of defects in the nano-scaled technologies. The estimation of pipeline design yield has *to* consider all sources of delay variation for each stage. In this section we develop mathematical models to capture the effect of variation to achieve a good estimation of yield in the design phase.

2.1. Process Parameter Variation

Process variation is a variation in physical process parameters among the different fabricated circuits. In general, parameter variations can be distinguished into two components [2, 9]:

- ***Die-to-die physical variations (inter-die):*** These are independent of design implementation and cause systematic variations in electrical characteristics within the chip.

Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS'05)
1530-1591/05 $20.00 © 2005 **IEEE**

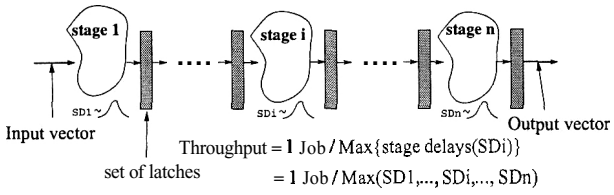

Figure 1: A pipeline circuit

- ***Within-die physical variations (intra-die):*** These can be distinguished into four sub-categories namely; a) Wafer level variations due to non-uniformities b) Die level variations caused by imperfections in lithography c) Random residuals due to random dopant fluctuation. The first two are usually modeled as a correlated systematic component whereas the last category forms the uncorrelated random component in intra-die process variation.

The effect of these parameter variations manifests in the circuit operation as the variation in device currents and voltages relation. Essentially any change of device parameters like L, W, T_{ox} can be modeled as a change in threshold voltage (V_{th}). We lump the effect of all kinds of systematic variation of parameter variations into a set of normal Gaussian random variables (RV) as components of threshold voltage. Considering systematic components of the intra-die parameter variation, the threshold voltage of an NMOS or PMOS device can be modeled as:

$$V_{th} = V_{tho} + \Delta V_{th}, \text{ where } \Delta V_{th} = \Delta V_{th\_inter} + \Delta V_{th\_intra} \qquad (1)$$

where, ΔV_{th_inter} and ΔV_{th_intra} represent threshold voltage shift, due to inter-die and intra-die variations, respectively.

2.2. Estimation of Yield for a Pipelined circuit

The delay of a pipeline stage (SD) consists of the clock-to-Q delay of the flip-flop ($T_{C\text{-}Q}$), propagation delay through the combinational logic (T_{comb}) and the setup time (T_{setup}) [3]. Using the simplified delay models as in [8], the delay of a combinational logic stage (T_{Comb}) with logic depth of L is expressed as:

$$T_{Comb} = \sum_{i=1}^{L} \frac{C_l V_{dd}}{\beta W_i (V_{dd} - V_{th})^{\alpha}} = \sum_{i=1}^{L} \frac{C_l V_{dd}}{\beta W_i (V_{dd} - V_{th})^{\alpha}}, \; 1 < \alpha < 2$$
$$\approx \sum_{i=1}^{L} \frac{C_l V_{dd}}{\beta W_i (V_{dd} - V_{th})^{\alpha}} \left(1 + \frac{\alpha \Delta V_{th}}{V_{dd} - V_{th}}\right) \qquad (2)$$

where C_l is the load capacitance of i^{th} gate, and W_i is its width. We ignored higher order terms of ΔV_{th} since, in sub-100nm technology α is closed to one and $\Delta V_{th} << (V_{dd} - V_{th})$. Authors in [5] show that the other two component of the stage delays (T_{setup}, $T_{C\text{-}Q}$) can also be modeled as Gaussian RV's. Let us assume that $SD_1, \ldots, SD_n$ are normal random variables representing the delay of each pipeline stage (3) (Fig. 1). The frequency of operation of the pipeline will be determined by the delay of the slowest stage. In other words, the overall pipeline delay (T_d) is given by:

$$T_d = \underset{i=1,\ldots,N}{Max}(SD_i) = \underset{i=1,\ldots,N}{Max}(T_{C-Q^i} + T_{Comb^i} + T_{Setup^{i+1}}) \qquad (3)$$

As SD_is are random variables, T_d will also be a random variable. The distribution of T_d determines the probability of meeting the target delay (T_{TARGET}) constraint. We define the probability of achieving the target delay (P_D) as:

$$P_D = P\{T_d \le T_{TARGET}\} = P\{\max_i SD_i \le T_{TARGET}\} \qquad (4)$$

The distribution of T_d depends on mean and standard deviation of the variables SD_is and the correlation among them. The variables SD_is are assumed to have different means and the standard deviations (known as Differently Distributed random variables (DD)). Depending on whether the variables are correlated with each other, we define three different cases:

(a) Independent and Differently Distributed random variables (IDD), which assume correlation between any two variables SD_i and SD_j ***(ρ,)*** to be zero. This corresponds to the delay distributions considering the intra-die variation in each stage delays, neglecting the spatial and electrical correlation among the stages.

(b) Perfectly Dependent and Differently Distributed (PerfDDD), which assumes complete correlation between any two variables SD_i and SD_j (ρ_{ij}= 1 or -1). This case corresponds to the inter-die variation. Since due to inter-die distribution of the parameters, delay of all the stages will be changed in the same direction (i.e. delay of all the stages will increase or decrease) we assume ρ_{ij} =1 in the PerfDDD case.

(c) Partially Dependent and Differently Distributed random variables (PDDD), which considers the exact correlation between the variables SD_i and SD_j i.e. ($-1 < \rho_{ij} < 1$). This is the most general case and it corresponds to the presence of both inter-die and intra-die (with spatial correlation) variation.

2.3. IDD Random Variables

If the variables SD_i, …, SD, are independent of each other then the cumulative distribution function of their maximum, which determine the maximum pipeline frequency, is given by [6,10]:

$$\Pr\{\max_i SD_i \le w\} = \Pr\{(SD_1 \le w) \ldots (SD_n \le w)\}$$
$$= \Pr\{SD_1 \le w\} \ldots \Pr\{SD_n \le w\} = \prod_{i=1} \Pr\{SD_i \le w\} \qquad (5)$$

Using (5) and assuming SD_is are Gaussian variables, P_D is given by:

$$P_D = \Pr\{T_d \le T_{TARGET}\} = \prod_{i=1}^{n} \Phi\left(\frac{T_{TARGET} - \mu_{SDi}}{\sigma_{SDi}}\right) \qquad (6)$$

where, μ_{SDi}, σ_{SDi} are mean and standard deviation of the delay of i^{th} stage and Φ represents the Cumulative Distribution Function (CDF) of a standard normal Gaussian variable with mean = μ_{SDi} and STD = σ_{SDi}. The estimation of the mean and the standard deviation of the overall pipeline design are quite complicated (involved). Using Jensen's inequality [113 the simple lower bound on the mean of T_D is given by:

$$E[T_d] = E[\max SD_i] \ge \max\{E[SD_i]\} \qquad (7)$$

This states that the mean of the overall pipeline delay will be larger than the maximum of the mean delay of all the stages. Estimation of the exact value of the mean (or a stronger upper and lower bound) and the standard deviation is very complex [6, 11]. In this section, we will present two approaches to

Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS'05)
1530-1591/05 $20.00 © 2005 **IEEE**

estimate the mean and the STD of the overall pipeline delay.

Approach-1: Let X_1 and X_2 are two Gaussian random variables with mean μ_1, μ_2 and standard deviation σ_1 and σ_2. The mean and STD of $Y = \max(X_1, X_2)$, μ_y and σ_y can be obtained as [6, 11]:

$$
\begin{aligned}
&m_0 = 1 \\
&m_1 = \mu_1\Phi(\alpha) + \mu_2\Phi(-\alpha) + a\varphi(\alpha) \\
&m_2 = (\mu_1^2 + \sigma_1^2)\Phi(\alpha) + (\mu_2^2 + \sigma_2^2)\Phi(-\alpha) + (\mu_1 + \mu_2)a\varphi(\alpha) \\
&\textit{where} \\
&a = (\mu_1 - \mu_2)/a \;;\; a^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho \\
&\mu_y = m_1 \\
&\sigma_y = \sqrt{m_2 - m_1^2}
\end{aligned}
\tag{8}
$$

where, $\varphi(\alpha) = (2\pi)^{-1/2}exp\ (-\alpha^2/2)$ (standard normal density with mean = 0, STD = *1* at a), ρ is the correlation coefficient between X_1 and X_2 and m_i denotes i^{th} non-central moment of Y [10]. Using (8), mean of the maximum of 'n' IDD variables can be approximated as follows:

$$
\begin{aligned}
&E[\max(X_1, X_2, ..., X_{n-1}, X_n)] \\
&= E[\max(X_1, X_2, ..., \max(X_{n-1}, X_n))] \\
&= E[\max(X_1, X_2, ..., N_{n,n-1})]
\end{aligned}
\tag{9}
$$

where, $N_{n,n-1}$ represents the normal approximation to max (X_n, X_{n-1}). This process is repeated until two variables are left. The final expected value is then computed using (9). **A** similar approach is used to estimate the STD. In our analysis we refer this method by IDD model. This model estimates the design yield within ***2%*** of ***Hspice*** Monte-Carlo analysis (Table 1). However, mean values are more accurately estimated within *±0.5%* of Monte-Carlo analysis (Table 1) using this approach.

Fig. 2(a)-(b) shows the overall delay distribution of a pipeline composed of transmission gate Master-Slave D Flip-Flops and 40 inverter pipeline in 70nm Berkeley predictive technology model (BPTM) [21] with $\sigma_{intra-die}$ = 40 mV obtained from Monte-Carlo simulation in ***Hspice*** and employing proposed IDD model. In the case of IDD variables (Fig. 2) the above approach produces good estimate of mean and standard deviation values. Fig. 2a shows that modeled delay distribution for a pipeline logic depth of 20 along with the Monte-Carlo simulation results. It shows that based on independent stage delay assumption the overall pipeline delay distribution is almost a Gaussian standard normal distribution. Fig. 2b shows this model is also valid for unbalanced pipeline, when logic depths in different stages are different.

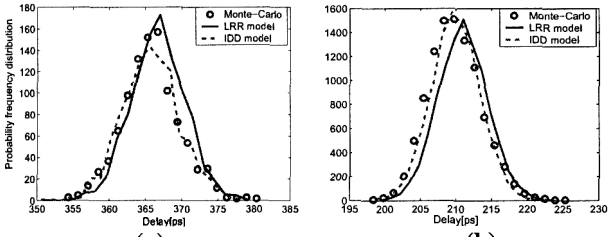

Figure 2: Modeling and Monte-Carlo simulation results for a 5-stage 40-long inverter chain pipeline considering stage delays as IDD variables (a) For 2-stage pipeline of logic depth 20. (b) Logic depth of the inverter chain varies in each stage (avg. logic depth 8).

Approach-2: The mean of a non-negative random variable Y can be obtained as [6]:

$$
E[Y] = \int_0^m \Pr\{Y > x\}\,dx = \int_0^\infty (1 - \Pr\{Y \le x\})\,dx \tag{10}
$$

Using (10) we can evaluate the mean of the $Y = \max(X_1, \ldots, X_n)$ = Pr[$\{X_1 > x\}$ + ... + $\{X_n > x\}$] as:

$$
\begin{aligned}
E[\max_i X_i] &= \int_0^\infty \left(1 - \prod_{i=1}^n \Pr\{X_i \le x\}\right)dx \\
&= \int_0^\infty \left(1 - \prod_{i=1}^n \Phi_i\left(\frac{x - \mu_i}{\sigma_i}\right)\right)dx
\end{aligned}
\tag{11}
$$

The above integration can be computed numerically. Using a similar approach the standard deviation can be estimated as follows:

$$
\begin{aligned}
E\left[\left(\max_i X_i\right)^2\right] &= \int_0^\infty x\left(1 - \prod_{i=1}^n \Pr\{X_i \le x\}\right)dx \\
&= \int_0^\infty x\left(1 - \prod_{i=1}^n \Phi_i\left(\frac{x - \mu_i}{\sigma_i}\right)\right)dx
\end{aligned}
\tag{12}
$$

Since the function 1- $\Phi(x)$ rapidly approaches to zero, the error introduced by the truncation of the above integral is very small. **A** more general result to estimate a tighter upper bound on mean can be given by [17]:

$$
E\left[\max_i X_i\right] \le c + \sum_{i=1,..,N} \int_c^\infty \Pr\{X_i > x\}dx \tag{13}
$$

Where, c is a real number. The smallest upper bound corresponded to the value of c, which satisfies:

$$
\sum_{i=1}^{n} \Pr\{X_i > c\} = 1 \tag{14}
$$

This can be numerically solved to determine c. In case of Gaussian variables the above integral (13) simplifies to (see appendix **A** for a derivation of (15) from (13)):

$$
E\left[\max_i X_i\right] \le c + \sum_{i=1,..,N} (\sigma_{Xi}^2\varphi(c) + (m_i - c)(1 - \Phi_i(c))) \tag{15}
$$

Hence from c, m_i and σ_{X_i} the upper bound of μ_T can be estimated. We refer this as LRR (by Lai, Robinson and Ross [11]) model in our subsequent discussions. Since in the LRR

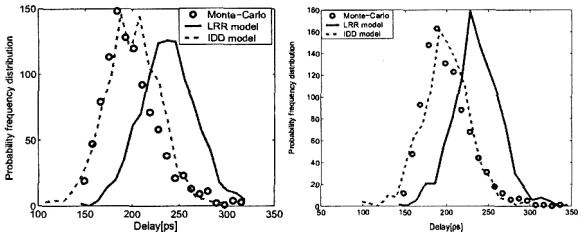

Figure 3: Modeling and Monte-Carlo simulation results for a 5-stage pipeline (each stage consists of 8 inverters) considering stage delays as IDD variables (a) for inter-die distributions and (b) for both inter-die and intra-die distributions.

Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS'05)
1530-1591/05 $20.00 © 2005 **IEEE**

model we use this upper bound as the modeled mean value, this model always gives a conservative estimate of the delay and yield of the design (Fig. 2). The yield estimated from these models matches very closely with the Monte-Carlo analysis (Table 1). When the logic depth varies in different stages LRR model starts to deviate from the Monte-Carlo simulation results (Fig. 3) because the assumption of identically distributed independent variables is no more valid for the stage delay variables. In this case stage with largest logic depth determines the overall pipeline delay distribution.

2.4. Perfectly DDD and Partially DDD Random Variables

If the delays of the different stages (represented by random variables X_i, for i =1, 2, ..., n) are correlated, the estimation of P_D, the mean and the standard deviation of the delay becomes complex. Let us consider that $X_1, ..., X_n$ are Normal random variables with the mean $m_1, ..., m_2$. Lets us also consider that the covariance matrix is Σ such that, $\Sigma_{(i,i)} = \sigma_{Xi}$ and $\Sigma_{(i,j)} = \rho_{ij}$. The probability P_D is given by:

$$\Pr\{\max X_i \le w\} = \Pr\{(X_1 \le w)...(X_n \le w)\} \quad (16)$$
$$= \int_0^w ... \int_0^w \varphi(X_1,...,X_n)\,dx_1...dx_n$$

where, $\varphi(X_1,...,X_n)$ is the joint normal PDF of all the variables which is given by:

$$\varphi(X_1,...,X_n) = \frac{\exp\{-\frac{1}{2}(\overline{X} - m_{\overline{X}})\Sigma^{-1}(\overline{X} - m_{\overline{X}})'\}}{\sqrt{(2\pi)^n |\Sigma|}} \quad (17)$$

where $\overline{X}$ and $m_{\overline{X}}$ are the random vectors representing the stage delays and their means respectively, $|\Sigma|$ is the determinant of covariance matrix Σ, and superscript *'t'* denotes transpose operation. Above integral (16) can be evaluated numerically to obtain P_D. Mean and standard deviation of the overall distribution can **be** obtained by considering two variables at a time as in (9). To understand this process considering the correlation, let us consider three normal random variable (x_1, x_2, x_3). Then, we have:

$$W = \max[x_1, x_2, x_3] = \max[\max[x_1, x_2], x_3] = \max[z, x_3] \quad (18)$$

The mean and the STD of z (μ_z, σ_z) can be calculated using the procedure shown in (8). We next assume z to be also normal random variable and repeat the procedure to calculate the distribution of W. However, to proceed, we evaluate the correlation between z and x_3 (say $\rho(z, x_3)$):

$$\rho[z, x_3] = [\sigma_1 \rho_{31} \Phi(\alpha) + \sigma_2 \rho_{31} \Phi(-\alpha)] / \sigma_z \quad (19)$$

where, ρ_{31} and ρ_{32} are the correlation of x_3 with x_1 and x_2, respectively. Progressing in this way, it can be extended to n variables, where n is the number of pipeline stages.

It can be noted that equations for DDD random variables represent a generic scenario, of which independent (case a) and perfectly dependent (case b) are special cases with correlation coefficient ρ between two successive stages 0 and 1 respectively. Fig. 3 shows the overall delay distribution of the inverter chain pipeline of two proposed models along with Monte-Carlo simulation results considering correlation between the stages. We observe that the IDD model approach gives very good estimate of yield (within **1.5%** from table 1), where as LRR model substantially overestimate the mean value of delay. This comes form the fact that upper bound obtained from equation (15) is a tight one only for IDD variable but not for correlated variables. Due to correlation determining value of c from (14) does not provide a stringent bound on mean delay. So we propose to use IDD model in all scenario to get the pipeline delay and yield estimates. However, LRR model gives a good estimate for IDD random variables and provides an upper bound on mean for PDDD and PerfDDD cases.

3. Estimation of Stage Delay Distribution for a Target Yield

Given a target delay, mean and standard deviation of the delay of individual stages need to be optimally designed to maximize P_D. In this section we have presented a methodology to estimate the design space for individual stages in a pipeline circuit to meet a required P_D. To simplify our discussion let us first assume that n stage delays are uncorrelated (i.e. only random intra-die variation is considered). Hence, given a required P_D, we have:

$$\prod_{i=1}^{n} \Phi\left(\frac{T_{TARGET} - \mu_{SDi}}{\sigma_{SDi}}\right) = P_D \quad (20)$$

Since the minimum value of $\Phi(x)$ is 1, to ensure the required value of P_D, each stage at least has to satisfy

$$\mu_{SDi} + \sigma_{SDi} \cdot \Phi^{-1}(P_D) \le T_{TARGET} \quad (21)$$

This will result in a worst-case bound for the feasible region in the μ-σ design space of the pipeline stages (Fig. **4,** worst-case bound line). However, the pipeline stages are often designed for equal delays which will corresponds to the design

#Stage x logic depth	Target delay (ps)	Monte-Carlo simulation			IDD model			LRR model	
		μ (ps)	σ (ps)	Yield (%)	μ (ps)	σ (ps)	Yield (%)	μ (ps)	Yield (%)
8 x 5	160	154.80	2.82	96.30	154.10	2.68	98.62	156.03	93.12
5 x 8	200	197.65	3.27	77.93	197.91	2.72	77.72	198.89	75.00
5 x *	215	209.82	3.67	91.66	209.95	3.42	93.00	210.92	88.35
2 x 20	370	365.67	3.93	87.8	365.75	3.83	86.69	366.70	80.57
5x8 inter	240	199.5	29.21	90.4	199.69	28.9	91.83	238.27	68.77
5x8 inter +intra	240	200.43	28.62	90.8	199.7	28.0	92.38	235.62	56.05

Table 1: Pipeline design delay distribution and yield values from IDD and LRR models and Monte-Carlo simulations (* in 3rd row represents variable logic depths)

Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS'05)
1530-1591/05 $20.00 © 2005 **IEEE**

for equal μ_{SDi} and σ_{SDi} for all stages (i.e. $\mu_{SD1} = \ldots = \mu_{SDn} = \mu_{SD}$ and $\sigma_{SD1} = \ldots = \sigma_{SDn} = \sigma_{SD}$). In this case, the μ-σ design space can be bounded by:

$$\mu_{SD} + \sigma_{SD}.\Phi^{-1}(P_D^{1/n}) \leq T_{TARGET} \quad (22)$$

Hence, design for equal stage delay, results in a tighter constraint on the feasible values of μ and σ (Fig. 4 shows 2 such equality bounds for $n = n1, n2$ with $n1 < n2$). If the stage delays are correlated with each other the estimation of the feasible region becomes more involved [6, 11]. However, since the cumulative distribution of P_D is jointly Gaussian (section 2.1), we can easily derive the following inequality:

$$\mu_{T_d} + \sigma_{T_d}\Phi^{-1}(P_D) \leq T_{TARGET} \quad (23)$$

Using Jensen's inequality on equation (23), we get:

$$\mu_{SD} \leq \mu_{T_d} \leq T_{TARGET} - \sigma_{T_d}\Phi^{-1}(P_D) \quad (24)$$

where μ_{SD} is the maximum of the individual stage delay. On top of that, even with correlated variables, individual stage delays have to satisfy the bound in (21). Due to the complex nature of the CDF with correlated stage delays, it is difficult to get a bound on the distribution as in (24). However, from empirical analysis, we have observed that the maximum standard deviation of all stages matches closely with σ_{T_d} obtained for a given yield.

The minimum bound on μ_{SD} and σ_{SD} can be obtained from the minimum allowable depth and process specification (Fig. 4, minimum μ, σ bound lines). Monte-Carlo analysis of a stage delay with an inverter chain of length equivalent to minimum allowable logic depth (without violating setup and hold times of the flip-flops) can determine the minimum mean and standard deviation values for a stage delay distribution.

4. Observations of pipelined circuit yield

Using the analytical models presented in section 2 and 3, in this section we have analyzed the effect of varying logic depth and unbalancing of the stage delays on the pipeline yield.

4.1. Perfectly Balanced vs. Unbalanced Pipeline Design

Incorporating imbalance among the pipeline stage delays can have positive impact on yield under statistical delay model. This is because a balanced pipeline has more number of critical paths and increasing number of critical paths adversely affects the yield [4]. Transistor sizing and logic re-structuring are methods that can be effectively applied to incorporate imbalance in an otherwise balanced pipeline thereby reducing the number of critical paths in the pipeline. We performed two experiments using the proposed models to observe the impact of imbalance among pipeline stages. In the first case, we introduced discrepancy in the logic depth of three stages (Fig. 5, case B), and in the second case, we used transistor sizing to insert a delay variances among stages (Fig. 5, case C). It can be observed that an unbalanced pipeline can result better yield than a balanced one for the same target delay as shown in Table 2. Consider a balanced Alu-Decoder design with 3 stages each of logic depth 4, when we re-size the transistors to incorporate some imbalance among the stage delays the effective number of critical paths decreases and hence the yield improves about **5%** (in Table 3, row 4, 4-4-4 sized configuration represent unbalanced stage delay design). Interestingly our proposed models able to correctly predict this yield improvements in terms of mean and STD of delay parameters. Using the proposed model we obtain smaller mean and STD values for an unbalanced pipeline design (row 4 **vs.** row 3 in Table 2). However, this is not predictable from the static delay analysis. We should note that, for an unreasonably high imbalance, we get diminishing return with respect to yield when pipeline performance is governed by the μ of the slowest stage. We observe this behavior in table 2-3 (mean and standard deviation of circuit delay for '4-5-3' configuration is more than nominal size '4-4-4' configuration). So at the same target delay 4-5-3 configuration produces a poor yield.

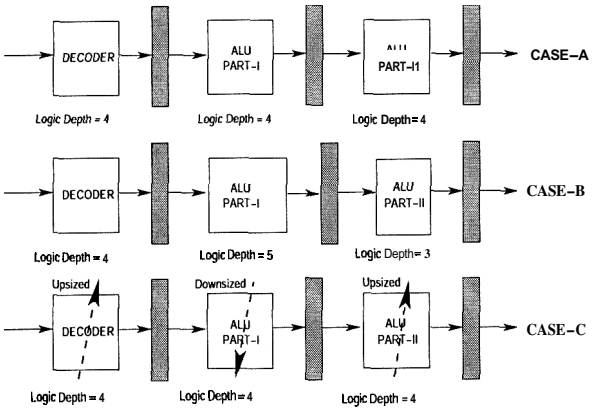

Figure 5: ALU-DECODER pipeline configurations

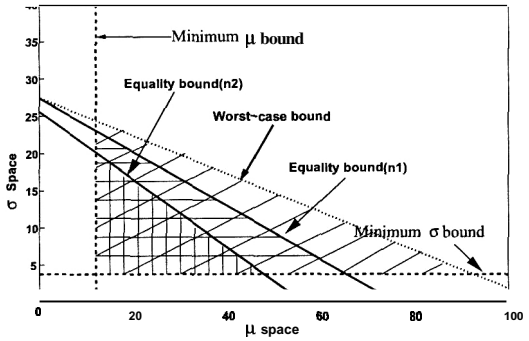

Figure 4: Design space bound for the mean and the standard deviation of stage delay

4.2. Scope of Design Optimization for a Target Yield

In section 3 we derived that a target yield corresponds to a range of viable (μ, σ) values for pipe stage delays. We have

Design configuration	Monte-Carlo (ps)		IDD model (ps)		LRR model (ps)
	μ	σ	μ	σ	μ
2-2-2-2-2-2	134.11	3.21	138.12	3.2	138.24
4-4-4	171.66	3.82	171.64	3.82	171.67
4-4-4 sized	171.08	3.66	171.08	3.66	171.63
4-5-3	188.71	4.34	189.48	3.81	190.4

Table 2: Overall design delay distribution statistics from the proposed models and Monte-Carlo simulation result

Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS'05)
1530-1591/05 $20.00 © 2005 **IEEE**

#Stage x Logic depth	Target delay (ps)	Yield (%)		
		Monte-Carlo	IDD model	LRR model
2-2-2-2-2-2	140	73.4	72.17	70.89
4-4-4	175	81.8	81.03	80.87
4-4-4 sized	175	86.5	85.81	85.75
453	175	91.7	92.62	88.63

Table 3: Overall design yield statistics from the proposed models and Monte-Carlo simulation result

also noted that a pipe stage can be designed for an appropriate (μ, σ), with i) trading-off pipeline depth with number of pipe stages, or ii) reducing the number of critical paths, by incorporating imbalance among pipeline stages, by transistor sizing or by some logic restructuring methods. Ensuring that (μ, σ) of each pipe stage will fall into the viable space of stage delay distribution, we can improve the yield of the pipeline. It gives us an interesting aspect of the design optimization under yield constraint. For a target yield, we can optimize a pipelined circuit for various design objectives. The objective function can be area or power, which are closely related parameters or the probability of achieving a better performance than the targeted one. This scenario is shown in Table 2 (configuration 4-4-4 vs. 4-4-4 sized), which shows that, by resizing we can alter the overall delay distribution of the pipeline, so that it has higher probability of achieving better performance beyond target, while maintaining good yield. The proposed models can be used as important guiding tools in designing pipelined circuit under parameter fluctuations.

5. Conclusions

We have presented analytical models to predict the yield of a pipeline design to achieve a target delay under parameter fluctuations. The models are verified for both the cases of independent and correlated stage delays and their prediction matches within 2% of *Hspice* Monte-Carlo simulation results. The proposed models can be used to capture the effect of both die-to-die and within-die variations in predicting pipeline design delay distribution and estimating design trade-off for yield improvement. A method of estimating bounds on the delay distribution parameters of individual pipe stages to satisfy a target yield is also proposed. The proposed models can be effectively used by the designers to explore design choices with respect to different design objectives, while ensuring target yield under statistical delay distribution.

Appendix A

Here we present a derivation of the 2nd term of (15) from that of (14). Each term of the sum of (14) can be written as:

$$\int_c^\infty \Pr\{X_i > x\}dx = \int_0^\infty \Pr\{X_i > x\}dx - \int_0^c \Pr\{X_i > x\}dx \qquad (A_1)$$

Where X is a nonnegative random variable from (11) we get,

$$E[X] = \int \Pr\{X > x\}dx = \int t f(t)\,dt \qquad (A_2)$$

Where $f_i(t)$ is the PDF of the random variable X_i, now from (A_1) and **(A,)** we've

$$\int_c^\infty \Pr\{X_i > x\}dx = \int_0^\infty t f_i(t)\,dt - \int_0^c \Pr\{X_i > x\}dx$$

$$= \int_c^\infty t f_i(t)\,dt + \int_0^c t f_i(t)\,dt - \int_0^c \Pr\{X_i > x\}dx$$

$$= \int t f_i(t)\,dt - I, \text{where}$$

$$I = \int_0^c t f_i(t)dt - \int_0^c \Pr\{X_i > x\}dx$$

$$= c\int_0^c f_i(t)dt - \int_0^c \left(\int_0^x f_i(t)dt\right)dx - \int_0^c \Pr\{X_i > x\}dx$$

$$= c(1 - \Pr\{X_i \geq c\}) - \int_0^c \Pr\{X_i \leq x\}dx - \int_0^c (1 - \Pr\{X_i \leq x\})dx$$

$$= c\Pr\{X_i > c\}, \text{Hence}$$

$$\int \Pr\{X_i > x\}dx = \int t f_i(t)\,dt - c\Pr\{X_i > c\} \qquad (A_3)$$

Where integration I is evaluated applying integrating by parts method on the first integral. Assuming $f_i(t)$ as a standard normal distribution $f(t) \approx N(m, \sigma)$, we can simplify (A_3) as :

$$\int_c^\infty t f_i(t)\,dt = \int_c^\infty t \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(t-m)^2}{2\sigma^2}} dt$$

$$= \int (t - m)\frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(t-m)^2}{2\sigma^2}} dt + m\int f_i(t)\,dt$$

$$= \frac{1}{2\sqrt{2\pi\sigma^2}} \int_{(c-m)^2}^\infty e^{\frac{-u}{2\sigma^2}} du + m\Pr\{X_i > c\}$$

$$= \frac{\sigma^2}{\sqrt{2\pi\sigma^2}} e^{\frac{-(c-m)^2}{\sigma^2}} + m\Pr\{X_i > c\}$$

$$= \sigma^2 f_i(c) + m\Pr\{X_i > c\} \qquad (A_4)$$

Hence from (A_3) and (A_4), we get the final result of (15).

References

[1] I. Koren et al., "Defect Tolerant VLSI Circuits: Techniques and Yield Analysis," ***Proceedings of the** IEEE,* Vol. 86, Sept. 1998, pp. 1817-1836.

[2] K. **A.** Bowman et al., "Impact of Die-to-Die and Within-Die Parameter Fluctuations on the Maximum Clock Frequency Distribution for Gigascale Integration", *JSSC*, 2002, pp. 183-190.

[3] J. L. Hennessy et al., "Computer Architecture: **A** Quantitative Approach", ***Morgan Kaufmann,*** 3-rd edition, May 2002.

[4] **S.** Borkar et al., "Parameter Variations and Impact on Circuits and Microarchitecture", *DAC* 2003, pp. 338-342.

[5] H. Mahmoodi et al., "Estimation of Delay Variations Due to Random-dopant Fluctuations in Nano-Scaled CMOS circuits", CICC 2004, pp. 17-20.

[6] A. M. Ross, "Useful Bounds on the Expected Maximum of Correlated Normal Variables", **Aug.** 2003.

[7] BPTM: http://www-device.eecs.berkeley.edu/~ptm/

[8] M. C.-T. Chao et al., "Static Statistical Timing Analysis for Latch-Based Pipelined Designs", ICCAD 2004, pp. 468-472.

[9] S. G. Duvall, "Statistical circuit modeling and optimization", *IWSM,* June 2000, pp. 56-63.

[10] **A.** Papoulis, et al., "Probability, Random Variables and Stochastic Processes", ***McGraw-Hill,*** Dec 2001.

[11] C. E. Clark, "The Greatest of a Finite Set of Random Variables", ***Operations Research 9(2),*** Mar-Apr, 1961, pp. 145-162.

Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS'05)
1530-1591/05 $20.00 © 2005 **IEEE**