Appendix 3

Clock Skew Model

Xiaohong Jiang and Susumu Horiguchi [JIA-01]

1. Introduction

The evolution of VLSI chips toward larger die sizes and faster clock speeds makes clock distribution an increasingly important issue. Clock skew modelling is important in the performance evaluation and prediction of clock distribution networks, because at high speeds clock skew becomes a very significant problem.

Clock skew may arise mainly from unequal clock path lengths to various modules and process variations that cause clock path delay variations. There are worst-case and statistical skew models not suitable for modelling the clock skews of general clock distribution networks in which clock paths are not identical.

The worst-case approach (Kung and Fisher model) can usually cause an unnecessarily long clock period. The statistical models handle the problem from the point of view that all clock paths are assumed to be identical and independent. Kugelmass and Steiglitz model predicts an upper bound of expected clock skew. Zarkesh-Ha, Mule' and Meindl model is too conservative for estimating the clock skew of a well-balanced clock network that has identical but strongly correlated clock paths (for example, a well-balanced H-tree).

In order to provide a more accurate and more general statistical skew model for general clock (in which clock paths can be not identical), these authors propose a new approach to estimate the mean value and variance of clock skew of general clock distribution networks. Based on the new approach, a closed-form model is also obtained for well-balanced H-tree clock distribution networks. The paths delay correlation caused

by the overlapped parts of path lengths is considered in the new approach, so the mean value and the variance of clock is accurately estimated for general clock distribution networks.

2. Clock skew modelling

For a given CDN (clock distribution network), let $t(l_o, l_i)$ denote the signal propagation time on the unique path from the clock source l_o to the sink l_i. The maximal clock delay ξ and the minimal clock delay η of the CDN can be defined as:

$$\xi = \max_i \left\{ t\left(l_0, l_i\right) \right\} \qquad (1)$$

$$\eta = \min_i \left\{ t\left(l_0, l_i\right) \right\} \qquad (2)$$

The clock skew between two sinks l_i and l_j is the delay difference $|t(l_o, l_i) - t(l_o, l_j)|$ and clock skew χ of the CDN is in general defined as the maximum value of $|t(l_o, l_i) - t(l_o, l_j)|$ over all sink pairs l_i and l_j and in the CDN. Thus, χ is given by

$$\chi = \max_{i \neq j} \left| t\left(l_0, l_i\right) - t\left(l_0, l_j\right) \right| = \xi - \eta \qquad (3)$$

Process variations are subject to two sets of factors: systematic factors, like power supply fluctuations, which can be controlled by proper techniques and factors that are random, and therefore uncontrollable by improved techniques. Therefore, the random factors determine the achievable performance of a circuit. We want to model the clock skew of general CDNs when the random factors are considered.

When random process variations are considered, variations of paths delay are modelled by normal distributions. To model the clock skew χ, random variables ξ and η should be first characterized. The model is based on the following two assumptions:

- *Assumption 1:* A CDN can in general be represented by a binary tree. We assume that both the maximal clock delay and the minimal clock delay in each subtree (and

also the whole binary tree) of the CDN can be modelled by normal distributions when process variations are considered. The assumption makes it easy to analyze the correlation that exists between the maximal and the minimal delay in a subtree. This correlation analysis is critical in determining the variance of skew in each subtree (and also the whole binary tree) of the CDN.

- *Assumption 2:* The delay along a clock path is the sum of the uncertain independent delays of the branches along the given path. Correlation between the delay of any two paths is determined only by the overlapped parts of their length.

The clock paths of a CDN usually have some common branches over their length, and these common branches cause correlation among the delays of these paths. The above assumption enables a complete analysis of this kind of correlation.

In addition to the delay correlation described in Assumption 2, the correlation among paths delay may also be caused by the correlated intra-die variations of these parameters involved in that delay (e.g., threshold voltages, resistances, etc.). However, finding the correlation coefficient of these parameters is, in practice, quite uncomfortable and difficult. So authors neglect these kinds of correlation in as indicated in Assumption 2. In general, the intra-die process parameters' correlation will lead to the paths delay in the same chip tending to be positive dependent. In this case, Assumption 2 will guarantee that the expected value of clock skew will still be upper bounded by the corresponding values estimated using our approach.

Compared to the old upper bound of expected skew of a well-balanced CDN where all the clock paths are assumed completely independent, our estimates are enhanced significantly, because the paths delay correlation caused by the common branches of paths length are completely considered. Furthermore, the new approach is applicable to general CDNs, whereas the old models is only applicable to the well-balanced CDNs in which clock paths are identical.

From (3), the mean value and the variance of χ are given by:

$$E(\chi) = E(\xi) - E(\eta) \tag{4}$$

$$D(\chi) = D(\xi) + D(\eta) - 2\rho\sqrt{D(\xi) \cdot D(\eta)} \tag{5}$$

Here, $E(\cdot)$ and $D(\cdot)$ represent the mean value and the variance of a random variable, respectively, and ρ is the correlation coefficient of ξ and η. The parameters $E(\xi)$, $E(\eta)$, $D(\xi)$, $E(\eta)$ and ρ should be accurately estimated for a CDN to allow the accurate modelling of clock skew.

3. Parameter estimation

A recursive approach for evaluating the parameters $E(\xi)$, $E(\eta)$, $D(\xi)$, $D(\eta)$ and ρ of general CDNs is presented here. Based on this algorithm, closed-form expressions of clock skews and the maximal clock delay of well-balanced H-tree CDNs are also developed.

3.1. Algorithm for general CDNs

A CDN can generally be represented by a binary tree, so a simplified binary tree shown in Figure 1 is taken as an example to illustrate the evaluating process of these parameters. The evaluating process is then applied to general CDNs. All the paths in Figure 1 are partitioned into independent branches s_0, s_1, s_2, ... by the branch split points in the clock tree, where d_i is the actual delay of s_i. The branch split point (i, j) in the clock tree is associated with a set of random variables $(\xi_{ij}, \eta_{ij}, \chi_{ij})$, here ξ_{ij}, η_{ij} and χ_{ij} are the maximal clock delay, the minimal clock delay and the clock skew of the subtree starting from the split point, respectively. Each random variable here is characterized by both its mean value and its variance.
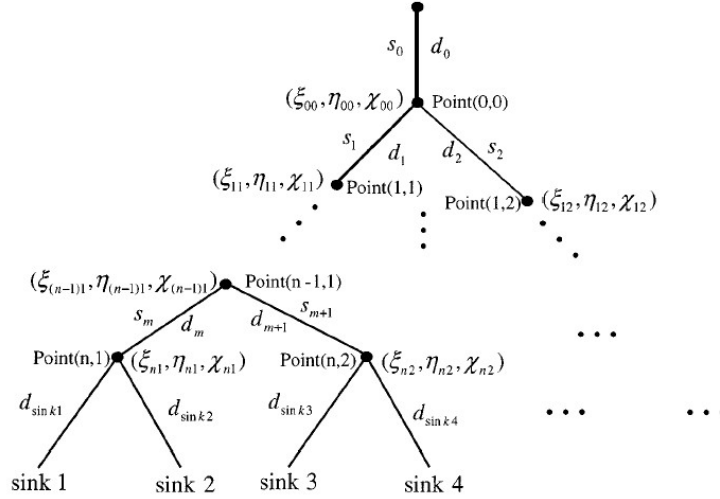
Figure 1: Binary clock tree.

To illustrate that the parameters $E(\xi)$, $E(\eta)$, $D(\xi)$, $D(\eta)$ and ρ of the simplified binary clock tree can be evaluated recursively, we begin with the evaluating process of $(\xi_{00}, \eta_{00}, \chi_{00})$. Let branch s_i also be associated with a set of random variables, (ξ_i, η_i, ρ_i), with ξ_i being the maximal clock delay and η_i being the minimal clock delay of the subtree starting from branch s_i, and being ρ_i the correlation coefficient of ξ_i and η_i. Thus

$$\xi_1 = \xi_{11} + d_1, \quad \xi_2 = \xi_{12} + d_2$$
$$\eta_1 = \eta_{11} + d_1, \quad \eta_2 = \eta_{12} + d_2 \tag{6}$$

$$\rho_1 = \frac{D(\xi_1) + D(\eta_1) - D(\chi_{11})}{2\sqrt{D(\xi_1) \cdot D(\eta_1)}}$$

$$\rho_2 = \frac{D(\xi_2) + D(\eta_2) - D(\chi_{12})}{2\sqrt{D(\xi_2) \cdot D(\eta_2)}} \tag{7}$$

Then we have:

$$\xi_{00} = \max\{\xi_1, \xi_2\} = \frac{\xi_1 + \xi_2 + |\xi_1 - \xi_2|}{2}$$

$$\eta_{00} = \min\{\eta_1, \eta_2\} = \frac{\eta_1 + \eta_2 - |\eta_1 - \eta_2|}{2} \tag{8}$$

$$\chi_{00} = \xi_{00} - \eta_{00} \tag{9}$$

Equations (6)–(9) indicate that the results of branch split point *(0, 0)* are determined by and can be evaluated from both the corresponding results [*(ξ_{11}, η_{11}, χ_{11})* and *(ξ_{12}, η_{12}, χ_{12})*] of the next lower level split points *(1, 1), (1, 2),* and the delay of the branches [s_1 and s_2] connecting the point to those next lower level split points. So once the results of *(ξ_{ij}, η_{ij}, χ_{ij})* are obtained for each lowest-level split point (i.e., the split point from which no further branch split points can be found in the subtree starting from that split point – points *(n, 1), (n, 2), …*), the process above can be used recursively to evaluate the mean values and the variances of clock skew, the maximal clock delay and the minimal clock delay of a general CDN in a bottom-up manner.

In fact, the results of *(ξ_{n1}, η_{n1}, χ_{n1})* of one lowest-level split point in Figure 1 can be obtained as follows: The mean values and the variances of ξ_{n1} and η_{n1} can be evaluated by using their distribution functions, respectively. The mean value and the variance of χ_{n1} are given by:

$$
\begin{aligned}
E\left(\chi_{n1}\right) &= E\left(\xi_{n1} - \eta_{n1}\right) = E\left(\left|d_{\sin k1} - d_{\sin k2}\right|\right) \\
&= \frac{2\sqrt{D_{\sin k}}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{E_{\sin k}}{\sqrt{D_{\sin k}}}\right)^2\right) + \frac{2\left|E_{\sin k}\right|}{\sqrt{2\pi}} \int_{0}^{\left|E_{\sin k}\sqrt{D_{\sin k}}\right|} \exp\left(-\frac{1}{2}t^2\right) dt
\end{aligned}
\tag{10}
$$

$$
D\left(\chi_{n1}\right) = D\left(\left|d_{\sin k1} - d_{\sin k2}\right|\right) = E_{\sin k}^2 + D_{\sin k} - \left[E\left(\chi_{n1}\right)\right]^2
\tag{11}
$$

where

$$
\begin{aligned}
E_{\sin k} &= E\left(d_{\sin k1}\right) - E\left(d_{\sin k2}\right) \\
D_{\sin k} &= D\left(d_{\sin k1}\right) + D\left(d_{\sin k2}\right)
\end{aligned}
\tag{12}
$$

Based on the results of *(ξ_{00},η_{00},χ_{00})*, the parameters, *E(ξ), E(η), D(ξ), D(η)* and ρ of the whole binary tree are then given by:

$$
\begin{aligned}
E(\xi) &= E\left(\xi_{00}\right) + E\left(d_0\right), \quad E(\eta) = E\left(\eta_{00}\right) + E\left(d_0\right) \\
D(\xi) &= D\left(\xi_{00}\right) + D\left(d_0\right), \quad D(\eta) = D\left(\eta_{00}\right) + D\left(d_0\right)
\end{aligned}
\tag{13}
$$

$$
\rho = \frac{D(\xi) + D(\eta) - D\left(\chi_{00}\right)}{2\sqrt{D(\xi) \cdot D(\eta)}}
\tag{14}
$$

The pseudocode for the parameter estimation algorithm is the following:

- **Parameter estimation for general CDNs**

```
{
Initialization:
      for each L_p ∈ V do
      {
```

$$\xi^b \leftarrow d^b$$
$$\eta^b \leftarrow d^b$$
$$\rho^b \leftarrow 1$$

```
      }
Algorithm:
      while (V not empty) do
      {
            for each L_p ∈ V do
            {
```

$$p_{\xi^{LP}}(x) = p\left(\xi_1^b < x\right) \cdot p\left(\xi_2^b < x\right)$$

$$p_{\eta^{LP}}(x) = 1 - p\left(\eta_1^b > x\right) \cdot p\left(\eta_2^b > x\right)$$

$$E\left(\xi^{LP}\right) = \int_{-\infty}^{+\infty} x \cdot d\left(p_{\xi^{LP}}(x)\right)$$

$$D\left(\xi^{LP}\right) = \int_{-\infty}^{+\infty} \left[x - E\left(\xi^{LP}\right)\right]^2 \cdot d\left(p_{\xi^{LP}}(x)\right)$$

$$E\left(\eta^{LP}\right) = \int_{-\infty}^{+\infty} x \cdot d\left(p_{\eta^{LP}}(x)\right)$$

$$D\left(\eta^{LP}\right) = \int_{-\infty}^{+\infty} \left[x - E\left(\eta^{LP}\right)\right]^2 \cdot d\left(p_{\eta^{LP}}(x)\right)$$

$$Cov\left(\xi^{LP}, \eta^{LP}\right) = Cov\left(\frac{\xi_1^b + \xi_2^b + \left|\xi_1^b - \xi_2^b\right|}{2}, \frac{\eta_1^b + \eta_2^b - \left|\eta_1^b - \eta_2^b\right|}{2}\right)$$

$$E\left(\chi^{LP}\right) = E\left(\xi^{LP}\right) - E\left(\eta^{LP}\right)$$

$$D\left(\chi^{LP}\right) = D\left(\xi^{LP}\right) + D\left(\eta^{LP}\right) - 2 \cdot Cov\left(\xi^{LP}, \eta^{LP}\right)$$

```
                  Remove G^LP = (V^LP, E^LP) from G = (v, E)
```

$$\xi^b \leftarrow \xi^{LP} + d^{LP}$$

$$\eta^b \leftarrow \eta^{LP} + d^{LP}$$

$$E(\xi^b) \leftarrow E(\xi^{LP}) + E(d^{LP})$$

$$E(\eta^b) \leftarrow E(\eta^{LP}) + E(d^{LP})$$

$$D(\xi^b) \leftarrow D(\xi^{LP}) + D(d^{LP})$$

$$D(\eta^b) \leftarrow D(\eta^{LP}) + D(d^{LP})$$

$$\rho^b \leftarrow \frac{D(\xi^b) + D(\eta^b) - D(\chi^{LP})}{2\sqrt{D(\xi^b) \cdot D(\eta^b)}}$$

```
            }
        }
}
```

In the pseudocode, a CDN is represented by graph $G = (V, E)$ with vertex (split point) set V and edge (branch) set E. The lowest level split point L_P of the graph is associated with random variables $(\xi^{LP}, \eta^{LP}, \chi^{LP})$, as defined above. (ξ^b, η^b, χ^b), are the random variables associated with the branches starting from a lowest level splitting point, with $(\xi_1^b, \eta_1^b, \chi_1^b)$ and $(\xi_2^b, \eta_2^b, \chi_2^b)$ representing the random variables associated with the two branches starting from L_P, respectively. For a CDN, the initial values of ξ^b and η^b are just the actual delay d^b of the branches that support sinks. d^{Lp} is the actual delay of the branch L_P connecting to its parent splitting point, and $G^{Lp} = (V^{LP}, E^{LP})$ is the subgraph starting from L_P.

Since the algorithm carries out the same amount of computation for each split point, the following conclusion can be obtained:

Theorem 1: *"The parameter estimation algorithm given above computes a network $G=(V, E)$ in $O(|V|)$ time".*

The theorem indicates that the parameter estimation algorithm is computationally effective in estimating the parameters, $E(\xi), E(\eta), D(\xi), D(\eta)$ and ρ of a general CDN.

3.2. Clock skew estimation for H-tree CDNs

The H-tree technique is widely used to reduce the clock skew. Due to the very symmetric structure of *H*-tree CDNs, it is possible for us to get a closed form model for clock skew and the maximal clock delay of *H*-tree CDNs.

Before developing the models, the *H*-tree itself must first be defined. Without loss of generality, a well-balanced *H*-tree has *n* hierarchical levels, where *n* denotes the tree depth. The level *0* branch corresponds to the root branch, and level *n* branches to the branches that support sinks. A level *i* branch begins with a level split *i* point and ends with level *i+1* split point. The *H*-tree illustrated in Figure 2 is drawn for *n=6*, which is used to distribute the clock signals to *64* processors.
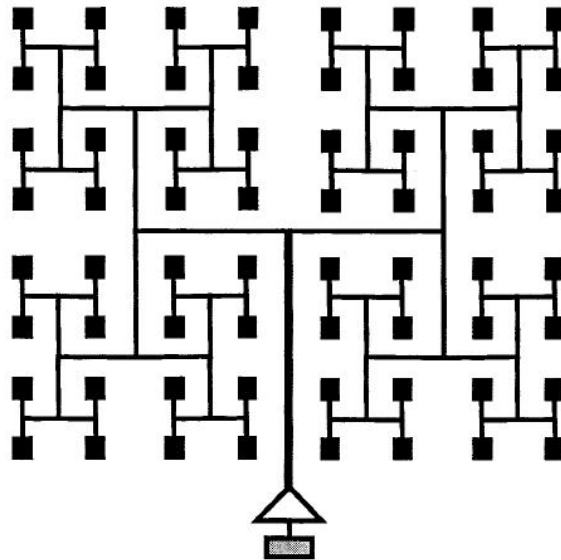
Figure 2: A well-balanced *H*-tree clock distribution network for 64 processors

For a *n* level well-balanced *H*-tree, let d_i, $i=0, ..., n$ be the actual delay of branch *i* of a clock path. The mean values and the variances of the maximal clock delay ξ and the minimal clock delay, η, of the *H*-tree are then given by following equations:

$$E(\xi) = \sum_{i=0}^{n} E(d_i) + \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} \sqrt{\sum_{k=1}^{i} \left(\frac{\pi-1}{\pi}\right)^{k-1} \cdot D(d_{n-i+k})} \qquad (15)$$

$$E(\eta) = \sum_{i=0}^{n} E(d_i) - \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} \sqrt{\sum_{k=1}^{i} \left(\frac{\pi-1}{\pi}\right)^{k-1} \cdot D(d_{n-i+k})} \qquad (16)$$

$$D(\xi) = D(\eta) = \sum_{i=0}^{n} \left(\frac{\pi-1}{\pi}\right)^{i} \cdot D(d_i) \qquad (17)$$

Results (15)–(17) and (3) indicate that the expected clock skew $E(\chi)$ and skew variance $D(\chi)$ of the n level well-balanced H-tree are given by:

$$E(\chi) = E(\xi) - E(\eta) = \frac{2}{\sqrt{\pi}} \sum_{i=1}^{n} \sqrt{\sum_{k=1}^{i} \left(\frac{\pi-1}{\pi}\right)^{k-1} \cdot D(d_{n-i+k})} \qquad (18)$$

$$\begin{aligned} D(\chi) &= D(\xi) + D(\eta) - 2 \cdot \rho \cdot \sqrt{D(\xi) \cdot D(\eta)} = \\ &= 2 \cdot (1-\rho) \cdot \sum_{i=0}^{n} \left(\frac{\pi-1}{\pi}\right)^{i} \cdot D(d_i) \end{aligned} \qquad (19)$$

where ρ is the correlation coefficient of ξ and η, and ρ can be recursively evaluated for a network as discussed in Section 3.1. The closed-form expressions (15)–(19) indicate clearly how the clock skew is accumulated along the clock paths and with the increase of H-tree size. This enables a suitable H-tree size to be selected for a specified clock frequency, and also enables minimization of the clock period, improving the speed for a fixed size H-tree network.

3.3. Yield of clock skew model

The clock period of a CDN is in general determined by the clock skew of the network. With the estimates of mean values and variances χ in hand, it is possible for us to estimate its yield. Here, the yield of a random variable means the probability that the variable is less than a specified value. For general CDNs, ξ and η are positively correlated (i.e., $\rho>0$) normal variables, and clock skew χ can be modelled by log-normal distribution as verified by available extensive simulation results. The clock skew yield, i.e., the probability that the actual skew of the network, χ, is less than a skew specification x $(P(\chi<x))$, can then be evaluated as:

$$P(\chi < x) = \int_0^x \frac{\log e}{\sqrt{2\pi} \cdot \delta_1 \cdot t} \exp\left[-\frac{1}{2}\left(\frac{\log t - \mu_1}{\delta_1} \right)^2 dt \right] \quad (20)$$

where parameters μ_1 and δ_1 are given by:

$$\mu_1 = \log\left(\frac{[E(\chi)]^2}{\sqrt{D(\chi) + [E(\chi)]^2}} \right) \quad (21)$$

$$\delta_1 = \sqrt{\log e \cdot \log\left(\frac{D(\chi) + [E(\chi)]^2}{[E(\chi)]^2} \right)} \quad (22)$$

Once the algorithm developed in Section 3 estimates the mean value of χ of a CDN, the yield of χ can be approximated by a lognormal distribution.

4. Clock skew calculation in function of its components

4.1. Calculation

The delay of a branch may then be obtained by averaging the rise and fall times. Using Sakurai's model for interconnection delays (90 % of time delay), the delay of a stage composed of a wire interconnecting two buffers (inverters) is:

$$T_{Delay} = 1.02 R_{int} C_{int} + 2.30 \left(R_0 C_0 + R_0 C_{int} + R_{int} C_0 \right) \quad (23)$$

Here R_0 is the output resistance of the driving transistor of minimum size inverter, C_0 is the input capacitance of the driven minimum size inverter, C_{int} and R_{int} are the capacitance and resistance of the interconnection line in the branch. Their expressions are given by:

$$R_0 = \frac{1}{K \cdot (V_{DD} - V_T)}, \quad K = \frac{\mu \cdot C_{ox} \cdot W}{L_{eff}}, \quad C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$

$$C_0 = C_{ox} \cdot W \cdot L_{eff}$$

$$R_{int} = \frac{\rho}{W_{int} \cdot t_{int}} \cdot L_{int} \tag{24}$$

$$C_{int} = \varepsilon_{ILD} \cdot \frac{W_{int}}{T_{ILD}} \cdot L_{int}$$

where:

- W and L_{eff}: width and effective length of the transistor.
- C_{ox}: gate unit area capacitance.
- t_{ox}: gate oxide thickness.
- μ: charge carrier mobility.
- V_T: threshold voltage.
- ρ: metal resistivity.
- ε_{ox}: oxide dielectric constant.
- ε_{ILD}: interlevel dielectric constant.
- W_{int}, L_{int} and t_{int}: width, length and thickness of the interconnection line.
- T_{ILD}: Interlevel dielectric thickness.

In the C_{int} expression, we don't take into account the contribution of fringing fields.

The process parameters and their standard deviations used here are based on the *0.25 µm* CMOS technology predicted by the International Technology Roadmap for Semiconductors (ITRS) and the MOSIS parametric test results of a typical *0.25 µm* technology. The mean values and intra-die standard deviations (SD) of these process parameters are presented in Table 1.

Parameters	Mean	SD
L_{eff} *(µm)*	0.25	0.0075
V_{DD} *(V)*	2.5	0
V_{TN} *(V)*	0.51	0.02
V_{TP} *(V)*	-0.51	0.02

t_{ox} (A°)	50	0.65
μ_n (cm²/V·s)	391	7.82
μ_p (cm²/V·s)	122	2.44
t_{int} (μm)	0.1	0.0013
T_{ILD} (μm)	0.1	0.0013

Table 1: *0.25 μm* Process parameters (mean and standard deviation).

Here, V_{DD} is not considered as a random variable since the power supply in a system is globally controlled. Furthermore, the standard deviation of the width of a transistor is 0.02 *μm* for n-MOS and 0.05 *μm* for p-MOS as estimated from MOSIS. The n-MOS transistor and p-MOS transistor in the minimum inverter of the technology are assumed to have the gate width/length of *0.37 μm /0.25 μm* and *1.1 μm /0.25 μm*, respectively.

As indicated in Assumption 2, the intra-die parameters correlations are neglected in this model. Thus, R_0 in (24) will be independent from C_0. One approach to calculating the delay variance of a branch due to the variations of process parameters is to first express the relations (24) in terms of independent variables. The delay variance of the branch can then be determined in terms of variances of these independent random variables. For example, the variance, σ_z^2, of a random variable z that is a function of independent random variables, $z=f(x,y,...)$, may be obtained from:

$$\sigma_z^2 = \left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + ... \qquad (25)$$

We are going to consider the following variables to calculate the variance of T_{delay}: $V_T, \mu, t_{ox}, L_{eff}, W, T_{ILD}, W_{int}, t_{int}$.

Therefore, the variance of the delay in a branch is the following:

$$\sigma_{T_{Delay}}^2 = \left(\frac{\partial T_{Delay}}{\partial V_T}\right)^2 \sigma_{V_T}^2 + \left(\frac{\partial T_{Delay}}{\partial \mu}\right)^2 \sigma_\mu^2 + \left(\frac{\partial T_{Delay}}{\partial t_{ox}}\right)^2 \sigma_{t_{ox}}^2 + \left(\frac{\partial T_{Delay}}{\partial L_{eff}}\right)^2 \sigma_{L_{eff}}^2 + \left(\frac{\partial T_{Delay}}{\partial W}\right)^2 \sigma_W^2$$
$$+ \left(\frac{\partial T_{Delay}}{\partial T_{ILD}}\right)^2 \sigma_{T_{ILD}}^2 + \left(\frac{\partial T_{Delay}}{\partial W_{int}}\right)^2 \sigma_{W_{int}}^2 + \left(\frac{\partial T_{Delay}}{\partial t_{int}}\right)^2 \sigma_{t_{int}}^2 \qquad (26)$$

where:

$$\frac{\partial T_{Delay}}{\partial V_T} = \frac{\partial T_{Delay}}{\partial R_0}\frac{\partial R_0}{\partial V_T} = 2.30\left(C_0 + C_{int}\right)\frac{R_0}{V_{DD} - V_T}$$

$$\frac{\partial T_{Delay}}{\partial V_T} = \frac{\partial T_{Delay}}{\partial R_0}\frac{\partial R_0}{\partial \mu} = 2.30\left(C_0 + C_{int}\right)\frac{R_0}{\mu}$$

$$\frac{\partial T_{Delay}}{\partial t_{ox}} = \frac{\partial T_{Delay}}{\partial R_0}\frac{\partial R_0}{\partial t_{ox}} + \frac{\partial T_{Delay}}{\partial C_0}\frac{\partial C_0}{\partial t_{ox}} = 2.30\left(C_0 + C_{int}\right)\frac{R_0}{t_{ox}} + 2.30\left(R_0 + R_{int}\right)\frac{C_0}{t_{ox}} \quad (27)$$

$$\frac{\partial T_{Delay}}{\partial L_{eff}} = \frac{\partial T_{Delay}}{\partial R_0}\frac{\partial R_0}{\partial L_{eff}} + \frac{\partial T_{Delay}}{\partial C_0}\frac{\partial C_0}{\partial L_{eff}} = 2.30\left(C_0 + C_{int}\right)\frac{R_0}{L_{eff}} + 2.30\left(R_0 + R_{int}\right)\frac{C_0}{L_{eff}}$$

$$\frac{\partial T_{Delay}}{\partial W} = \frac{\partial T_{Delay}}{\partial R_0}\frac{\partial R_0}{\partial W} + \frac{\partial T_{Delay}}{\partial C_0}\frac{\partial C_0}{\partial W} = 2.30\left(C_0 + C_{int}\right)\frac{R_0}{W} + 2.30\left(R_0 + R_{int}\right)\frac{C_0}{W}$$

$$\frac{\partial T_{Delay}}{\partial T_{ILD}} = \frac{\partial T_{Delay}}{\partial C_{int}}\frac{\partial C_{int}}{\partial T_{ILD}} = \left(1.02R_{int} + 2.30R_0\right)\frac{C_{int}}{T_{ILD}}$$

$$\frac{\partial T_{Delay}}{\partial W_{int}} = \frac{\partial T_{Delay}}{\partial R_{int}}\frac{\partial R_{int}}{\partial W_{int}} + \frac{\partial T_{Delay}}{\partial C_{int}}\frac{\partial C_{int}}{\partial W_{int}} = \left(1.02C_{int} + 2.30C_0\right)\frac{R_{int}}{W_{int}} + \left(1.02R_{int} + 2.30R_0\right)\frac{C_{int}}{W_{int}}$$

$$\frac{\partial T_{Delay}}{\partial t_{int}} = \frac{\partial T_{Delay}}{\partial R_{int}}\frac{\partial R_{int}}{\partial t_{int}} = \left(1.02C_{int} + 2.30C_0\right)\frac{R_{int}}{t_{int}}$$

Also we have to consider that every branch has a different length. It doubles it length each two levels.

Once the mean delay value and the delay variance of each branch are evaluated for a network, the theoretical approach developed can be used to estimate the mean value, the variance and the yield of the clock skew of the network.

4.2. Verification

To verify the new approach, transistor level Monte Carlo simulations are also conducted. In the simulation, each basic parameter in (24) is simulated by a normal random variable. To agree with the conditions used in the theoretical approach, the correlation between K and V_T, between μ and C_{ox} is neglected as discussed above. The actual delay of a branch is then evaluated from the random values of these basic parameters using (24). The actual delay of a path is the sum of the actual delays of the

branches along that path. The actual maximal clock delay, minimal clock delay and clock skew of the network are then determined by (1)–(3). For a specified value, the yield of a parameter (clock skew) is estimated by the ratio of number of simulations in which the parameter is less than the specified value to the total number of simulations.

The first network considered is well known as the *H*-tree approach shown in Fig. 2 (for brevity, inverters are not illustrated in the following networks). Due to the very symmetrical design of *H*-tree clock networks, all clock paths within the *H*-tree are identical, and the old statistical model can be used to get an upper bound of its expected clock skew when all the paths are assumed to be independent. According to the old model, an upper bound of expected clock skew of a well-balanced *H*-tree is asymptotically given by:

$$E^{upper}\left(\chi\right) = \sigma \cdot \left[\frac{4\ln N - \ln\ln N - \ln 4\pi + 2C}{\left(2\ln N\right)^{1/2}} + O\left(\frac{1}{\log N}\right) \right] \quad (28)$$

with the variance of clock skew being given by:

$$D^{upper}\left(\chi\right) = \frac{\pi^2 \sigma^2}{6\ln N} + O\left(\frac{1}{\log^2 N}\right) \quad (29)$$

where:

- σ: standard deviation of path delay.
- $C \approx 0.5772$: Euler's constant.
- N: number of paths (number of processing elements).
- $O(\cdot)$: higher order term.

In a n-level *H*-tree, there are a total of $2^{n+1}-1$ branches, and it can be used to distribute clock signals to 2^n elements. For two combinations of parameters h and W_{int}, in a wide range, and when the numbers of processors are 4, 8, 16, 32, and 64, the theoretical results (obtained using both the old and new models) and simulation results of clock skews of corresponding *H*-trees are summarized in Figure 3.
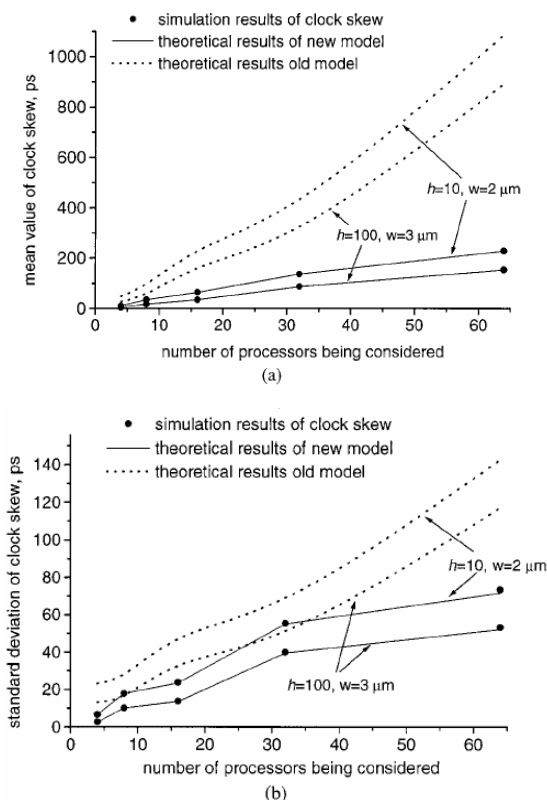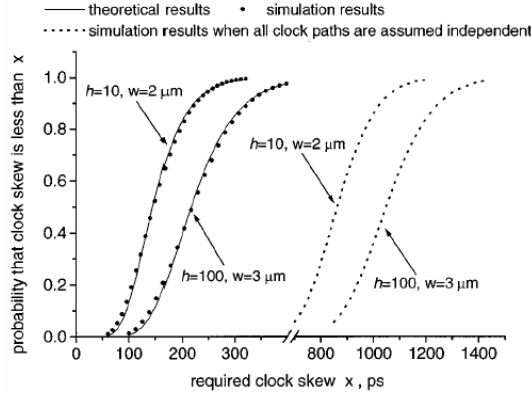
Figure 3: Simulation results and theoretical results of clock skew of *H*-tree networks when different numbers of processors are considered. (a) Mean value of clock skew. (b) Standard deviation of clock skew.

The results in Figure 3 show that the new model is accurate in estimating the mean values and standard deviations of the clock skew of *H*-tree CDNs, where the delay correlation determined by the overlapped parts of path lengths has been considered as indicated in Assumption 2. Compared to the old estimates of expected clock skew where all the paths in the *H*-tree are assumed completely independent, the new estimates based on Assumption 2 are a significant enhancement. For different sized *H*-trees, the expected clock skews estimated using the old model (26) are at least 2.6 times the expected clock skews estimated using our approach. This is shown in Fig. 3(a). In cases where clock frequency is limited by skew rather than by the minimum time between two successive events propagated through the *H*-tree, an unnecessarily long clock period will result from using the old skew model.

Based on the above estimated results of the mean values and the variances of χ, the yields of χ and can be further estimated as discussed in Section4. For the six-level *H*-tree shown in Figure 2, the theoretical yield results and the simulation yield results of χ is summarized in Figure 4. For comparison, we also present in Figure 3 the simulation results of yield of clock skew when all paths are assumed independent.

Figure 4: Simulation yield results and theoretical yield results of clock skew of an *H*-tree network for two combinations of parameters *h* and *W*.

The results in Figure 4 indicates that the old model's independent assumption leads to very conservative estimates of clock skew yields of *H*-tree CDNs that have identical but strongly correlated clock paths. The results in Figure 5 also show that when the mean values and the variances of χ of the *H*-tree CDNs are accurately estimated by our approach, the yields of their χ is further approximated by log-normal distribution.

Due to the assumption that all clock paths are identical, the old statistical skew model could not be used to model the clock skews of general clock networks with non-identical paths. The new model developed, however, can be used to accurately estimate the mean values and the variances of clock skew of these general CDNs. For a well-balanced CDN (e.g., *H*-tree clock network), the expected clock skews estimated by old model are very conservative because correlation among paths delay are completely neglected. On the other hand, the expected clock skews estimated by the new model are enhanced significantly. For the well-balanced *H*-tree network shown in Figure 2, the results in Figure 3 show that when parameters *h=100* and $W_{int}=3\ \mu m$, the expected clock skew estimated by old model is about *892.5 ps*, a value *5.8* times larger than the actual value. For a traditional clocking mode and when the *10%* rule of thumb relating the

skew to the clock period is used, the actual clock period should be dominated by the maximal clock delay, and the mean value of clock period will be *6.96 ns*. However, when the old skew model is used in the skew estimate, the clock period should be determined by the clock skew rather than the maximal clock delay, and the mean value of clock period will be *8.925 ns*. The old model will thus mislead efforts to reduce the clock period of the network. For a pipelined clocking mode, the clock periods of well-balanced *H*-tree networks will be dominated by the clock skew rather than by the minimum time between two successive events propagated through the *H*-tree, an unnecessarily long clock period will result from using the old skew model.

5. Conclusions

- The old model (upper bound model) is too conservative in estimating the expected skew of a well-balanced CDN. Also, in not general enough to model the clock skew of a non-balanced CDN.

- A closed form model (statistical) of clock skew is presented for well-balanced H-tree CDNs.

- The path delay correlation determined by the overlapped parts of path lengths is completely considered in this approach. Therefore, the mean value and variance of clock skew is accurately estimated for general CDNs.

- Considering process variations in designing a clock distribution network, mean values and variances of delays for all branches should carefully be estimated, then the approach presented here will be useful in evaluating and predicting the network's performance of clock skew.

- Two main assumptions:

  - *Assumption 1:* A CDN can generally be represented by a binary tree. We assume that both the maximal clock delay and the minimal clock delay in

each subtree (and also the whole binary tree) of the CDN can be modelled by normal distributions when process variations are considered. The assumption makes it easy to analyze the correlation that exists between the maximal and the minimal delay in a subtree. This correlation analysis is critical in determining the variance of skew in each subtree (and also the whole binary tree) of the CDN.

- *Assumption 2:* The delay along a clock path is the sum of the uncertain independent delays of the branches along the given path. Correlation between the delays of any two paths is determined only by the overlapped parts of their length.

- To apply the model in a H-tree CDN, it's necessary to know:

  - Interconnection resistance: R_{int}
  - Interconnection capacitance: C_{int}
  - On-resistance of the driving transistor: R_0
  - Input capacitance of the driving inverter: C_0
  - Power supply voltage: V_{DD}
  - Threshold voltage: V_T

  - Threshold voltage deviation (in %): σ_{VT}
  - Charge carrier mobility deviation (in %): σ_{μ}
  - Gate oxide thickness deviation (in %): σ_{tox}
  - Transistor width deviation (in %): σ_W
  - Effective channel length deviation (in %): σ_{Leff}
  - Wire width deviation (in %): σ_{wint}
  - Wire thickness deviation (in %): σ_{tint}
  - ILD thickness deviation (in %): σ_{TILD}

  - Lowest level branch length: L_{int}
  - H-tree levels: n