Statistical Skew Modeling and Clock Period Optimization of Wafer Scale H-tree Clock Distribution Network

Xiaohong Jiang and Susumu Horiguchi
Graduate School of Information Science
Japan Advanced Institute of Science and Technology,
JAIST, Tatsunokuchi, ISHIKAWA 923-1292, JAPAN
(Email: jiang@jaist.ac.jp; hori@jaist.ac.jp)

Abstract **- Available statistical skew models are too conservative in estimating the expected clock skew of a well-balanced H-tree. New closed form expressions are presented for accurately estimating the expected values and the variances of both the clock skew and the largest clock delay of a well-balanced H-tree. Based on the new model, clock period optimizations of wafer scale H-tree clock network are investigated under both conventional clocking mode and pipelined clocking mode. It is found that when the conventional clocking mode is used, clock period optimization of wafer scale H-tree is reduced to the minimization of expected largest clock delay under both area restriction and power restriction. On the other hand, when the pipelined clocking mode is considered, the optimization is reduced to the minimization of expected clock skew under power restriction. The results obtained in this paper are very useful in the optimization design of wafer scale H-tree clock distribution networks.**

Key words **– H-tree, clock skew, clock delay, clock period, process variations.**

1. Introduction

The need for careful design of clock distribution for Wafer Scale Integrated circuits (WSI) has been widely recognized [1]. The advances in monolithic-WSI technology [2] have demonstrated that clock and signal distribution can severely limit WSI system performance because clock skew becomes a very significant problem. Clock skew may arise mainly from unequal clock path lengths to various modules and from process variations that cause clock path delay variations [3,4]. To reduce clock skew, a common way is to use the well-balanced H-tree technique [5,6]. The uncontrollable clock skew of well-balanced H-tree is due to variations in process parameter that affect the interconnect impedance and, in particular, any distributed buffer amplifiers. When estimating the clock skew, either a worst-case or a statistical approach may be utilized. A worst-case approach can usually cause an unnecessarily long clock period. In a statistical approach, on the other hand, the clock parameters may be chosen so that the probability of timing failure is very small, but not zero. This usually results in a shorter clock period. Available literature dealing with statistical clock skew modeling [7,8] approaches the problem from a standpoint that clock paths are assumed to be independent, so an upper bound of expected clock skew is obtained. The model is too conservative when it is used to estimate the expected skew of a well-balanced H-tree clock network because the stronger correlations among paths are neglected. For different level H-trees, the expected clock skews estimated by using the old model are at least two times the actual expected skews as shown in this paper. In the case where the clock frequency is limited by the skew rather than by the minimum time between two successive events propagated through H-tree [5], an unnecessarily long clock period will be caused by the old skew model. To avoid the conservative result of old model, a new model is developed in this paper to accurately estimate the expected values and the variances of both the clock skew and the largest clock delay of a well-balanced H-tree.

Based on the new model, clock period optimization of wafer scale H-tree clock network are investigated when both the intra-wafer process variations and the inter-wafer process variations are considered. We found that when the conventional clocking mode is used, the clock period of a wafer scale well-balanced H-tree is dominated by its largest clock delay, and the optimization of clock period is reduced to the minimization of expected largest clock delay under both area restriction and power restriction. On the other hand, when the pipelined clocking mode is used, the clock period of wafer scale well-balanced H-tree is determined by its clock skew, and the clock period optimization is reduced to the minimization of expected clock skew under only the power restriction.

The paper is organized as follows: The new models for clock skew and largest clock delay are developed and verified in Section 2. The optimization of clock period under the conventional clocking mode is discussed in Section3, Section 4 focus on the optimization of clock period under the pipelined clocking mode, and Section 5 summarizes the contributions of this paper.

2. Modeling the clock skew and the largest clock delay of well-balanced H-tree

For a well-balanced H-tree clock distribution network having M clock paths, let pd_i be the actual propagation delay of *i-th* clock path, then the largest clock delay, ξ, and the smallest clock delay, η, of the network can be defined as

$$\xi = \max\{pd_1, \cdots, pd_M\} \quad (1)$$

$$\eta = \min\{pd_1, \cdots, pd_M\} \quad (2)$$

Thus the clock skew, χ, of the network is given by

$$\chi = \xi - \eta \quad (3)$$

When process variations are considered, the delay of a path is modeled by normal distribution [4,8]. To model the clock skew, χ, random variables ξ and η should be first characterized. The clock skew model developed in this paper is based on the following assumption.

Assumption 1**:** For a well-balanced H-tree clock distribution network in which clock paths depend on each other, both its largest clock delay and its smallest clock delay can be modeled by normal distributions when process variations are considered.

This assumption takes its roots in the available results [9,10]. The assumption makes it easy to analyze the correlation that exist between ξ and η, and most important, the mean values and the variances of both clock skew and the largest clock delay

estimated by using the assumption are very accurate as shown in this paper.

Before developing the model of clock skew and the largest clock delay, the H-tree itself must first be defined. Without loss of generality, the well-balanced H-tree has *N* hierarchical levels, where *N* denotes the tree depth. Level 0 branch corresponds to the root branch, and level *N* branches to the branches that support leaves. Level *i* branch begin with level *i* split point and end with level *i*+1 split point, with level 0 split point corresponds to primary clock input point. The H-tree illustrated in Fig.1 is drawn for *N*=8 (256 paths), which is used to distribute the clock signals to 256 processors implemented by WSI in a 4 -inch wafer.

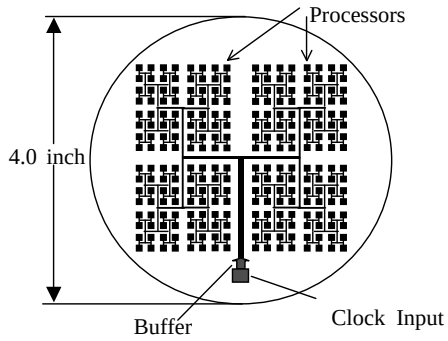

Fig.1 An H-tree clock distribution network for 256 processors in a 4 inch WSI (Clock buffers are not illustrated here).

2.1 Evaluation of the mean values and the variances of clock skew and the largest clock delay

For a *N* hierarchical levels well-balanced H-tree, let d_i, (i = 0,…,N) be actual delay of the *i-th* branch of a clock path, ξ_i be the largest clock delay and η_i be the smallest clock delay of the sub H-tree starting from *i-th* level split point. Then:

$$\xi_i = \max\left\{\left(d_{(i+1)1}+\xi_{(i+1)1}\right),\left(d_{(i+1)2}+\xi_{(i+1)2}\right)\right\} =$$
$$\left\{\left(d_{(i+1)1}+\xi_{(i+1)1}\right)+\left(d_{(i+1)2}+\xi_{(i+1)2}\right)+\left|\left(d_{(i+1)1}+\xi_{(i+1)1}\right)-\left(d_{(i+1)2}+\xi_{(i+1)2}\right)\right|\right\}/2$$
$$\eta_i = \min\left\{\left(d_{(i+1)1}+\eta_{(i+1)1}\right),\left(d_{(i+1)2}+\eta_{(i+1)2}\right)\right\} =$$
$$\left\{\left(d_{(i+1)1}+\eta_{(i+1)1}\right)+\left(d_{(i+1)2}+\eta_{(i+1)2}\right)-\left|\left(d_{(i+1)1}+\eta_{(i+1)1}\right)-\left(d_{(i+1)2}+\eta_{(i+1)2}\right)\right|\right\}/2$$

Where $d_{(i+1)1}$ and $d_{(i+1)2}$ are independent samples of $d_{(i+1)}$, $\xi_{(i+1)1}$ and $\xi_{(i+1)2}$ are independent samples of $\xi_{(i+1)}$, $\eta_{(i+1)1}$ and $\eta_{(i+1)2}$ are independent samples of $\eta_{(i+1)}$. Based on the Assumption 1, the mean values and the variances of ξ_i and η_i are given by following expressions based on the symmetry of well-balanced H-tree and the properties of normal variables [11].

$$E(\xi_i) = E(d_{i+1}) + E(\xi_{i+1}) + \frac{\sqrt{D(d_{i+1})+D(\xi_{i+1})}}{\sqrt{\pi}} \tag{4}$$

$$E(\eta_i) = E(d_{i+1}) + E(\eta_{i+1}) - \frac{\sqrt{D(d_{i+1})+D(\eta_{i+1})}}{\sqrt{\pi}} \tag{5}$$

$$D(\xi_i) = \frac{\pi-1}{\pi}\cdot\left[D(d_{i+1})+D(\xi_{i+1})\right] \qquad D(\eta_i) = \frac{\pi-1}{\pi}\cdot\left[D(d_{i+1})+D(\eta_{i+1})\right] \tag{6}$$

Where $E(\cdot)$ and $D(\cdot)$ represent the mean value and the variance of a random variable, respectively.

Above process indicate clearly that $E(\xi_i)$, $D(\xi_i)$, $E(\eta_i)$ and $D(\eta_i)$ can be obtained by using $E(\xi_{i+1})$, $D(\xi_{i+1})$, $E(\eta_{i+1})$, $D(\eta_{i+1})$, $E(d_{i+1})$ and $D(d_{i+1})$.Then a recursive approach is obtained to evaluate the mean values and the variances of both the largest clock delay and the smallest clock delay of a well-balanced H-tree.

Applying (4)-(6) to the *N* level well-balanced H-tree recursively, we have:

$$E(\xi) = \sum_{i=0}^{N} E(d_i) + \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D(d_{N-i+k})} \tag{7}$$

$$E(\eta) = \sum_{i=0}^{N} E(d_i) - \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D(d_{N-i+k})} \tag{8}$$

$$D(\xi) = D(\eta) = \sum_{i=0}^{N}\left(\frac{\pi-1}{\pi}\right)^{i}\cdot D(d_i) \tag{9}$$

The results of (7)-(9) and (3) indicate that the expected clock skew and the variance of the skew of a *N* level well-balanced H-tree can be estimated by:

$$E(\chi) = E(\xi) - E(\eta) = \frac{2}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D(d_{N-i+k})} \tag{10}$$

$$D(\chi) = D(\xi)+D(\eta)-2\cdot r\cdot\sqrt{D(\xi)\cdot D(\eta)} = 2\cdot(1-r)\cdot\sum_{i=0}^{N}\left(\frac{\pi-1}{\pi}\right)^{i}\cdot D(d_i) \tag{11}$$

Where r is the correlation coefficient of ξ and η, and r can be recursively evaluated for a network [20]. It can be seen from (10) that the mean value of clock skew of a well-balanced H-tree is determined completely by the variances of branches delay, and the clock skew is accumulated in a complicated way.

2.2 Statistical model for Wafer scale H-tree

The parameters variations in the manufacturing process of CMOS digital circuits cause path delays to deviate from the designed values, and thus cause clock skew in a well-balanced H-tree. These variations are usually classified as intra-die variation, inter-die variation, intra-wafer variation, inter-wafer variation, intra-lot variation, and inter-lot variation, etc. Since we are interested in the wafer-scale H-tree, so only the intra-wafer and the inter-wafer variations are considered in the statistical modeling to avoid arriving at intractably complex models.

In general, inter-wafer and intra-wafer parameters variations can be modeled by normal distributions or uniform distributions [12]. Here the normal distributions are used, but the results obtained can easily be extended to other distributions. Let p^i_{inter} be the value of *i-th* parameter determined only by the inter-wafer normal variations $N(\mu^i_{inter},\delta^i_{inter})$ with mean value, μ^i_{inter}, and inter-wafer standard deviation, δ^i_{inter}, then the actual value of the parameter will be determined by the normal distribution $N(p^i_{inter},\delta^i_{intra})$ with mean value, p^i_{inter}, and intra-wafer standard deviation, δ^i_{intra}. Thus, when both the inter-wafer and intra-wafer variations are considered, the actual value p^i of *i-th* process parameter can be expressed as

$$\begin{aligned} p^i &= p^i_{inter} + \delta^i_{intra}\cdot\tau^i_{intra} \\ &= \mu^i_{inter} + \delta^i_{inter}\cdot\tau_{inter} + \delta^i_{intra}\cdot\tau^i_{intra} \\ \tau_{inter},\tau^i_{intra} &\cong N(0,1) \end{aligned} \tag{12}$$

Where τ_{inter} is the random variables associated with inter-wafer parameter variations and will be same for all process parameters in a wafer. τ^i_{intra} is the random variable associated with the intra-wafer variation of *i-th* parameter. τ^i_{intra} is independent of τ_{inter}, and the correlation between τ^i_{intra} reflects the correlation between parameters.

For a fixed value of random variable τ_{inter} (i.e. when only the intra-wafer variations are considered), the mean values and the variances of clock skew and the largest clock delay of wafer scale H-tree are only determined by intra-wafer parameter variations

and given by (7)-(11), and these values will vary with the variation of τ_{inter} . Thus, when both the inter-wafer and intra-wafer are considered, the mean values and the variances of clock skew and the largest clock delay of wafer scale H-tree can be modeled as:

$$E(\xi)=\int_{-\infty}^{+\infty}\left(\sum_{i=0}^{N}E(d_i)+\frac{1}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D(d_{N-i+k})}\right)\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{\tau_{inter}^2}{2}}d\tau_{inter} \quad (13)$$

$$E(\chi)=\int_{-\infty}^{+\infty}\left(\frac{2}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D(d_{N-i+k})}\right)\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{\tau_{inter}^2}{2}}d\tau_{inter} \quad (14)$$

$$D(\xi)=\int_{-\infty}^{+\infty}\left(\sum_{i=1}^{N}\left(\frac{\pi-1}{\pi}\right)^{i}\cdot D(d_i)\right)\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{\tau_{inter}^2}{2}}d\tau_{inter} \quad (15)$$

$$D(\chi)=\int_{-\infty}^{+\infty}\left(2\cdot(1-r)\cdot\sum_{i=1}^{N}\left(\frac{\pi-1}{\pi}\right)^{i}\cdot D(d_i)\right)\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{\tau_{inter}^2}{2}}d\tau_{inter} \quad (16)$$

2.3 Yield estimates for skew and the largest delay

The clock period of a H-tree network is in general determined by both the clock skew and the largest clock delay of the network. With the estimates of mean values and the standard deviations of both ξ and χ in hand, it is possible for us to estimate the yields of ξ and χ. As indicated in the **Assumption 1,** the yield of ξ can be estimated by normal distribution $N(E(\xi),D(\xi))$ with mean, $E(\xi)$, and variance, $D(\xi)$. On the other hand, clock skew can be modeled by log-normal distribution as verified by extensive simulation results [10]. Then 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 be evaluated as:

$$P(\chi<x)=\int_0^x\frac{\log e}{\sqrt{2\pi}\cdot\delta\cdot t}\exp\left[-\frac{1}{2}\left(\frac{\log t-\mu}{\delta}\right)^2\right]dt \quad (17)$$

Here parameters μ and δ are given by:

$$\mu = \log\left(\frac{[E(\chi)]^2}{\sqrt{D(\chi)+[E(\chi)]^2}}\right) \quad (18)$$

$$\delta = \sqrt{\log e\cdot\log\left(\frac{D(\chi)+[E(\chi)]^2}{[E(\chi)]^2}\right)} \quad (19)$$

So once the mean values and the variances of both ξ and χ are estimated by the model developed in Section 2.2, the yields of ξ and χ can be estimated by the normal and log-normal distributions, respectively.

2.4 Old skew model for H-tree

Under the assumption that all the paths are independent, an upper bound of expected clock skew $E^{upper}(\chi)$ of a well-balanced H-tree is asymptotically given by [7]:

$$E^{upper}(\chi)=\delta_1\cdot\left[\frac{4\ln M-\ln\ln M-\ln 4\pi+2C}{(2\ln M)^{1/2}}+O\left(\frac{1}{\log M}\right)\right] \quad (20)$$

With the variance of clock skew is given by

$$D^{upper}(\chi)=\frac{\pi^2\delta_1^2}{6\ln M}+O\left(\frac{1}{\log^2 M}\right) \quad (21)$$

Where δ_1 is the standard deviation of path delay, C=0.5772… is Euler's constant, and $O(\cdot)$ is the higher order terms. When both the inter-wafer and intra-wafer are considered, the expected skew $E^{upper}(\chi)$ and the variance $D^{upper}(\chi)$ of the old model should also be evaluated as that of (13)-(16) .

2.5 Verification of the new model

To verify the new skew model, extensive simulations and theoretical calculations are conducted that based on an assumed 10.0×10.0 cm^2 WSI as illustrated in Fig.1, along with an 16×16 array of medium grained processing elements (PE) each with 4.0×4.0 mm^2 effective area and 6.0×6.0 mm^2 tile area in 1μm CMOS technology*. The process parameters and estimation of delay variation are based on a predicted 1μm CMOS technology [8]. As that did in [13], the inter-wafer standard deviation of a process parameter is assumed to be 15% of its inter-wafer nominal, and the intra-wafer standard deviation of a process parameter is assumed to be 10% of its intra-wafer nominal. Based on the mean delay value and the delay variance of each branch, both simulation and the theoretical approach can be used to estimate the clock skew and the largest clock delay of a H-tree. In theoretical approach, algorithm presented in Section 2.2 is used to evaluate the parameters $E(\xi)$, $E(\chi)$, $D(\xi)$ and $D(\chi)$ of the H-tree. In the simulation approach, the actual delay of a branch is simulated by a normal random variable. The actual delay of a path is the sum of these actual delays of the branches along the path. Then actual largest clock delay, the smallest clock delay and the clock skew of the H-tree can be simulated by the expressions (1)-(3). Parts of the simulation results and theoretical results are summarized in Fig.2 – Fig.4.

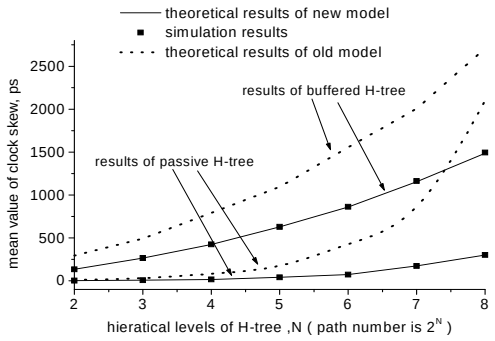

Fig.2 Simulation results and theoretical results in mean values of clock skew.

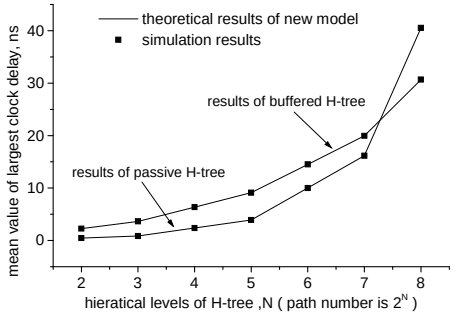

Fig.3 Simulation results and theoretical results in mean values of the largest clock delay.

* A processing element with significant local memory and processing power would of course be much larger. However, we are considering medium grain PE's so that 4mm×4mm is a reasonable size for this case example.

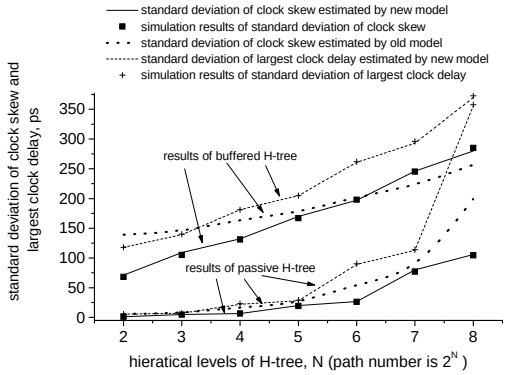

Fig.4 Simulation results and theoretical results in variances of clock skew and the largest clock delay.

To verify furthermore the yield models presented in Section 2.3, the simulation results and the theoretical results in yields of skew and the largest delay of the wafer scale H-tree are summarized in Fig.5 and Fig.6. Here *m* inverters are inserted in each clock path, each inverter is *h* times the minimum inverter of 1μm CMOS technology, and line width of the clock path is *W*. Two combinations of *m*, *h* and *W* are used in the verification. For comparison, we also provide in Fig.5 the simulation results of yield of clock skew when all paths are assumed independent.

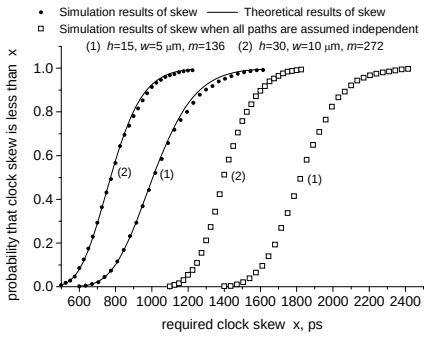

Fig.5 Simulation results and the theoretical results in yield of clock skew of H-tree for two combinations of parameters *m*, *h* and *W*.

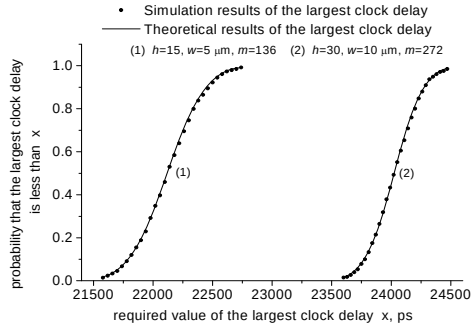

Fig.6 Simulation results and the theoretical results in yield of the largest clock delay of H-tree for two combinations of parameters *m*, *h* and *W*.

The results in Fig.2 show clearly that the old model is too conservative when it is used to estimate the expected skew of a well-balanced H-tree, the expected skew estimated by old model is at least two times the actual expected skew for different level H-tree. The yield results in Fig.5 illustrate further that the assumption of old model that all clock paths are independent is too conservative in estimating the clock skew of H-tree. In the case where the clock frequency is limited by the skew rather than by the minimum time between two successive events propagated through the H-tree [5], an unnecessarily long clock period will be caused by using the old skew model.

On the other hand, the new model developed in this paper can be used to get an accurate estimate of mean values and variances for both the clock skew and the largest clock delay of a well-balanced H-tree as shown in Fig.2-Fig.4, so the too conservative results of old skew model can be avoided by using the new model. The yield results in Fig.5 and Fig.6 indicate that when the mean values and variances of both the clock skew and the largest clock delay are accurately estimated, the yields of clock skew and the largest clock delay of H-tree can be further accurately estimated by log-normal and normal distribution, respectively. Furthermore, the closed expressions (13)-(16) indicate clearly how the clock skew and the largest clock delay are accumulated along the clock paths and with the increase of H-tree size. This enable a suitable H-tree size is selected for a specified clock frequency, and this also enable the optimization to be made to minimize the clock period and thus improve the speed for a fixed size H-tree network. In the following two Sections, we focus on the clock period optimization of the wafer scale H-tree in Fig.1.

3. Clock period optimization of wafer scale H-tree in the conventional clocking mode

By using the conventional clocking method, the clock period, *T*, is required to be greater than the longest clock delay,ξ. Another requirement is the 10% rule of thumb relating the skew,χ, to the clock period [5]. Thus the clock period must be

$$T = \max\left(\xi,\ 10\chi\right) \tag{22}$$

Due to the very symmetrical design of well-balanced H-tree, a lower clock skew is expected, but larger propagation delay from clock input to a processor will be caused because of its relative long clock path. We can reduce the propagation delay by appropriately inserting drivers (inverters) in a clock path. For a H-tree path, drivers can be inserted as shown in Fig.7

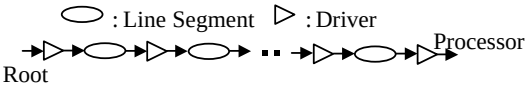

Fig. 7 Insertion of drivers in an H-tree metallic path.

If *k* inverters which is *h* times the minimum inverter are inserted in the path, the propagation delay, $d_{Propagation}$, of the path is given by [14]:

$$d_{Propagation} = k \cdot T_{segment}$$

$$T_{segment} = 2.3\frac{R_0}{h}\left(\frac{C_{\text{int}}}{k} + hC_0\right) + \frac{R_{\text{int}}}{k}\left(\frac{C_{\text{int}}}{k} + 2.3hC_0\right) \tag{23}$$

Where $T_{segment}$ is the delay per line segment, C_0 and R_0 are the input capacitance and output resistance of the minimum size inverter, C_{int} and R_{int} are the capacitance and resistance of the interconnection line in the path. These parameters are given by:

$$R_0 = \frac{L_T / W_T}{\mu \cdot C_{ox}(V_{DD} - V_T)}, \qquad C_0 = C_{ox} \cdot W_T \cdot L_T$$
$$R_{\text{int}} = \rho \cdot L_{\text{int}} / W_{\text{int}} \cdot t, \qquad C_{\text{int}} = \varepsilon \cdot W_{\text{int}} \cdot L_{\text{int}} / t_{ox} \qquad (24)$$

Where W_T and L_T are the width and length of the transistor, C_{ox} is the gate unit area capacitance, μ is the charge carrier mobility, V_T is the threshold voltage, ρ is the metal resistivity, and ε is the oxide dielectric constant. Here the interconnection line is with width W_{int}, length L_{int} and thickness t on an oxide layer of thickness t_{ox}. By setting derivatives of $d_{Propagation}$ with respect to k and h to zero, optimal values for k and h can be obtained to minimize the propagation delay of the path [14].

Note from (13) and (14) that the mean value of the largest clock delay of a well-balanced H-tree is determined by both the expected delay of single path and the expected skew of the H-tree. Here the expected delay of a path is determined by the expected branches delay in the path. On the other hand, the expected clock skew is determined by the variances of the branches delay in the path. The minimum of mean value and variance of path delay usually do not occur for the same combination of driver number and driver size [8]. Thus minimizing path delay does neither guarantee minimum skew nor minimum largest clock delay, it is just a heuristic design that turn out to provide a good tradeoff between clock skew and the largest clock delay.

The results in Section 2 and above analysis indicate that the clock skew and the largest clock delay of a well-balanced H-tree are completely determined by the mean values and the variances of branches delay. The mean value and the variance of a branch delay are the sum of the mean delay values and the delay variances of the line segments in the branch (here we insert at least one driver in each branch to make the analysis easy), respectively. The mean value of a line segment can be obtained by using (23), and one approach to calculate the delay variance of a line segment due to the variations of process parameters is to express the relation (23) in terms of independent variables. These variables are geometrical dimensions, C_{ox} (unit area gate oxide capacitance), μ (carrier mobility), and V_T (threshold voltage). All these factors can be considered independent [8,15]. Thus, the variance of a line segment can be determined in terms of variances of these independent random variables. For example, the variance, δ_Z^2, of a random variable Z that is a function of some independent random variables, $z = f(x, y, \cdots)$ may be obtained from

$$\delta_z^2 = \left(\frac{\partial f}{\partial x}\right)^2 \cdot \delta_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \delta_y^2 + \cdots \qquad (25)$$

To find the effects of drivers on the clock skew and on the largest clock delay, the new model developed in Section 2 is used to evaluate the wafer scale H-tree clock network shown in Fig.1. Here, process parameters are based on the predicted 1μm CMOS technology [8,16]. The PE area, tile area and the variations of process parameters are same as that of Section 2.5, the width of interconnection line is set as W_{int}=10μm. Fig.8 illustrates the trends of the mean values $E(\xi)$ and $E(10\chi)$ of the H-tree with the variations of driver number and driver size in each path.

The results in Fig.8 indicate clearly that in the conventional clocking mode, the actual clock period of well-balanced H-tree clock distribution is dominated by its largest clock delay rather than by its clock skew. For a driver size, there exist an optimal number of drivers (about 100 in each path) to minimize the mean value of the largest clock delay and thus the clock period of the H-tree. Therefore, in the conventional clocking mode, minimization of clock period of well-balanced H-tree can be implemented by minimizing its largest clock delay. To assure that H-tree can works with high probability around the optimized clock period, the variances of the largest clock delay should also be considered. By finding the minimum clock period for different driver size, and calculating the corresponding variance, we can construct a plot of standard deviation vs. mean value of the largest clock delay (clock period) as shown in Fig.9.

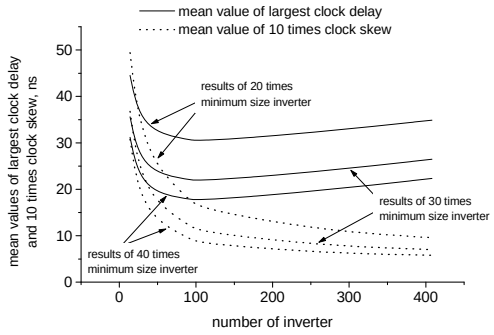

Fig.8 The trends of expected values of clock skew and the largest clock delay with the variations of driver number and driver size in each path.

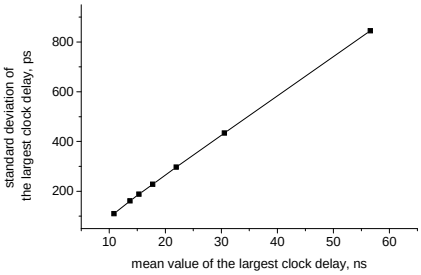

Fig.9 Standard deviation vs. mean value of the largest clock delay.

Fig.9 illustrates a similar relationship as obtained in [17], the standard deviation of the largest clock delay decreases linearly with the decrease of its mean value. Thus, in the conventional clocking mode, the minimization of clock period of wafer scale H-tree can be simply reduced to the minimization of the mean value of the largest clock delay.

It is well known that by increasing the driver size, the minimum clock period can be further reduced as illustrated in the Fig.8. However, beyond a certain point that the period improvement becomes costly in both area and power. By finding the minimum clock period for different driver size, and calculating the corresponding average power dissipation and area of the network, we can construct a plot of power and area vs. the largest clock delay (clock period) as shown in Fig.10. The results in Fig.10 show that in the conventional clocking mode, reducing in clock period carries both extra area and extra power penalty. Furthermore, all of the additional area is in active silicon (transistor sizing) so that yield and reliability are reduced while power dissipation is increased, so both the area requirement and power requirement should be considered in the optimization of clock period.

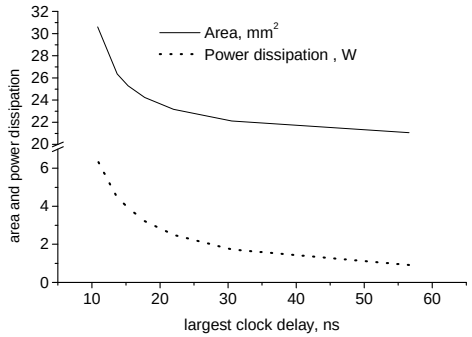

Fig.10 Power and area vs. the expected largest clock delay (notice that the two curves do not follow the same ordinate scale: one is a area and the other is a power).

Therefore, in the conventional clocking mode, the clock period optimization of wafer scale well-balanced H-tree can be formulated as:

$$\text{Minimize} \int_{-\infty}^{+\infty}\left(\sum_{i=0}^{N} E(d_i)+\frac{1}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D\left(d_{N-i+k}\right)}\right)\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{\tau_{inter}^2}{2}}d\tau_{inter}$$

$$\text{subject to} \quad \begin{matrix} Area < A_{spec} \\ Power < P_{spec} \end{matrix} \qquad (26)$$

In practice, one can perform this optimization by using the genetic algorithms [21]. For example, when the area requirement is 25mm^2 and power dissipation requirement is 4 W, we found the minimum mean value of the largest clock delay, $E(\xi)$, is 17.8 ns, and the corresponding standard deviation $\sqrt{D(\xi)}$ given by (15) is 228 ps. This occurs when number of the drivers is 100 in each path and the size of the drivers is 40 times the minimum size inverter. To assure a very high (>99%) probability of system success, we chose a confidence level of $3\sqrt{D(\xi)}$ based on the normal distribution as discussed in Section 2.3. Then the H-tree can works with very high (>99%) probability at frequency larger than $1/\left(E(\xi)+3\sqrt{D(\xi)}\right)\approx 54\text{MHz}$.

4. Clock period optimization of wafer scale H-tree in the pipelined clocking mode

The results in Section 3 show that in the conventional clocking mode, the clock period of a well-balanced H-tree network is seriously limited by its largest clock delay that is usually larger for H-tree clock network. The limitation can be released by using the pipeline techniques [5][18]. Given the fact that Fig.7 path structure of a well-balanced H-tree is pipelined, to insert the same kind of event (transition to zero or one), one just has to wait until the previous occurrence of the event propagates to the third stage. Again, another requirement is the 10% rule of thumb relating the skew to the clock period. Thus, the clock period will be

$$T = \max\left(2\times T_{segment},\ 10\times\chi\right) \qquad (27)$$

Fig.11 illustrates the trends of the mean values $E(T_{segment})$ and $E(\chi)$ of the H-tree with the variations of driver number and driver size in each path.

The results in Fig.11 indicate clearly that when drivers are used to divide clock paths into line segments, the mean delay of a line segment is considered smaller than the expected clock skew of the H-tree. Thus, in the pipelined clocking mode, the clock period of wafer scale well-balanced H-tree is dominated by its clock skew rather than by the delay of a line segment. For a driver size, there exist an optimal number of drivers to minimize the mean value of the clock skew and thus the clock period of the H-tree. Therefore, in the pipelined clocking mode, minimization of clock period of well-balanced H-tree can be implemented by minimizing its clock skew.

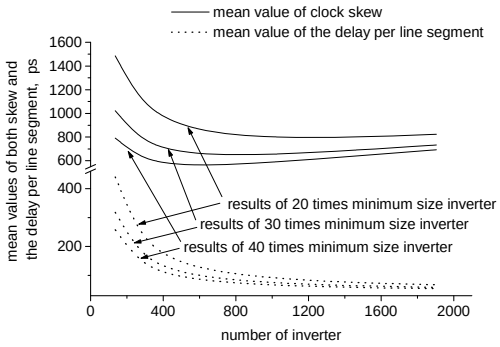

Fig.11 The trends of expected values of clock skew and line segment delay with the variations of driver number and driver size in each path.

For a driver size, it can be seen from Fig.8 and Fig.11 that the driver number required to minimize the clock period in the pipelined clocking mode is larger than that of conventional clocking mode. In the pipelined clocking mode, the clock period is dominated by clock skew, and the mean value of clock skew is governed by the variances of branches delay in a path. So the results in Fig.8 and Fig.11 indicate that the variation in capacitance due to the variations in line dimensions (width and thickness) is the main cause in the path delay variation and relative inverter delay variation is minor. The above discussion results in a same conclusion as that obtained in [8]. Drivers are effective in making the line delay linear with the line length, but they are even more effective in reducing the standard deviation of the line delay.

To assure that H-tree can works with high probability around the optimized clock period, the variances of the clock skew should also be considered. By finding the minimum clock period for different driver size, and calculating the corresponding variance, we can construct a plot of standard deviation vs. mean value of clock skew (clock period) as shown in Fig.12.

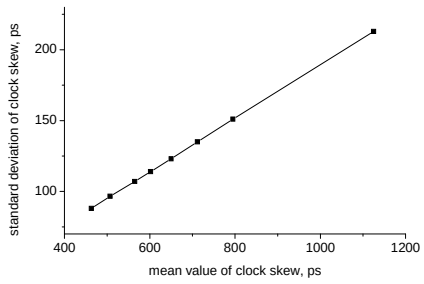

Fig.12 Standard deviation vs. mean value of clock skew.

Again, Fig.12 illustrates a similar linear relationship as that of Fig.9. The standard deviation of clock skew also decreases linearly with the decrease of its mean value. Thus, in the pipelining clocking mode, the minimization of clock period can be simply reduced to the minimization of the mean value of clock skew.

It is shown in Fig.11 that by increasing the driver size, the minimum clock period can be further reduced. By finding the minimum clock period for different driver size, and calculating the corresponding average power dissipation and area of the network, we can construct a plot of power and area vs. the clock skew (clock period) as shown in Fig.13.

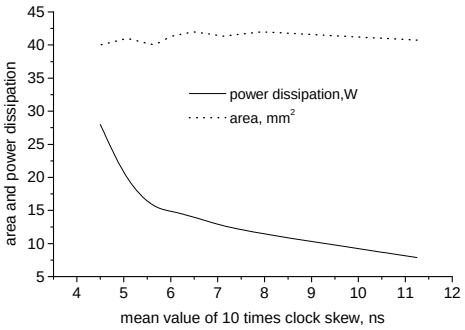

Fig.13 Power and area vs. the expected clock skew (notice that the two curves do not follow the same ordinate scale: one is a area and the other is a power).

The results in Fig.13 show that reducing in clock period in the pipelined clocking mode only carries extra power penalty. Compared to the results in Fig.8, the power dissipation in the pipelined clocking mode is larger than that of the conventional clocking mode. This is due to the increase in both the number of drivers be used and the clock frequency obtained. However, this is the price to pay for achieving high frequencies, because the problem of high-power dissipation at high-speed is intrinsic in CMOS. As illustrated in Fig.11 that when larger drivers are used to reduce the clock skew, the fewer number of the drivers are required to minimize clock skew, then a negligible area variation is caused as shown in Fig.13. Thus, in the pipelined clocking mode, the clock period optimization of wafer scale well-balanced H-tree can be formulated as:

$$\text{Minimize} \quad \int_{-\infty}^{+\infty}\left(\frac{2}{\sqrt{\pi}}\sum_{i=1}^{N}\sqrt{\sum_{k=1}^{i}\left(\frac{\pi-1}{\pi}\right)^{k-1}\cdot D\left(d_{N-i+k}\right)}\right)\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{\tau_{inter}^2}{2}}\,d\tau_{\text{inter}}$$

$$\text{subject to} \quad Power < P_{spec} \tag{28}$$

For example, when the power dissipation requirement is 20 W, we found the minimum mean value of clock skew, $E(\chi)$, is 0.507 ns, and the corresponding standard deviation $\sqrt{D(\chi)}$ is 0.0966 ns, and this occurs when number of the drivers is 408 in each path and the size of the drivers is 50 times the minimum size inverter. As discussed in Section 2.3, the clock skew of well-balanced H-tree clock is accurately modeled by log-normal distribution. E.g., the simulation results of clock skew of above arrangement are summarized Fig.14, and the *Kolmogorov-Test* for goodness of fit [19] indicates that the common logarithm of the clock skew is accurately modeled by normal distribution.

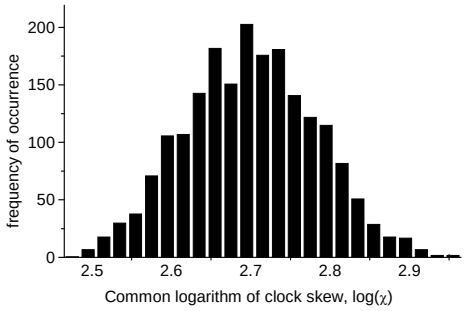

Fig.14 Distribution of clock skew of a well-balanced H-tree.

Then the yield of clock period, i.e. the probability that the actual clock period is less than a specified period value x ($P(T < x)$), can be given as

$$\begin{aligned} P(T < x) &= P(10\chi < x) = P(\chi < 0.1\cdot x) \\ &= \int_0^{0.1\cdot x}\frac{\log e}{\sqrt{2\pi}\cdot\delta\cdot t}\exp\left[-\frac{1}{2}\left(\frac{\log t-\mu}{\delta}\right)^2\right]dt \end{aligned} \tag{29}$$

Here parameters μ and δ are given by expressions (18) and (19), respectively.

To assure a very high (>99%) probability of system success, we chose the specified value of clock period is 7.6 ns such that

$$P(T < 7.6) = P(\chi < 0.76) > 99\%$$

Then the H-tree can works with very high (>99%) probability at frequency larger than 1/(7.6 ns) ≅ 131MHz. The results indicate that speed of Well-balanced H-tree can greatly improved by using the pipelined clocking mode, but the improvement in speed carries extra power dissipation. Such trade-off is crucial to the success of WSI systems.

5. Conclusions

H-tree technique is widely used for clock distribution. Due to the unavoidable random process variations, robust design of wafer-scale H-tree clock distribution network is very important. The robust design of H-tree relays on reliable statistical skew model. Available statistical skew is too conservative and not suitable for the purpose. A new statistical model was developed in this paper to accurately estimate the expected values and the variances of both clock skew and the largest clock delay of well-balanced H-tree. The new model indicates clearly how the clock skew is accumulated along the clock path and how the clock skew related to the largest clock delay, this enables the optimization design (in the sense of clock period minimization) and robust design (in the sense of high probability of system success) of H-tree clock distribution networks to be made when the process variations are considered. Based on the new model, the clock period optimization of wafer scale H-tree is investigated under two clocking modes. We found that when the conventional clocking mode is used, the optimization of clock period is reduced to the minimization of the mean value of the largest clock delay under both area restriction and power restriction. On the other hand, when the pipelined clocking mode is used, the optimization of clock period is reduced to the minimization of the mean value of clock skew under only power restriction. The optimization process also guarantees the robustness of the design

in the sense that standard deviation of obtained clock period decrease linearly with the decrease of the mean value of the clock period. Furthermore, the results in this paper indicate that the variation in capacitance due to the variations in line dimensions (width and thickness) is the main cause of the path delay variation and relative inverter delay variation is minor, inverters are effective in making the line delay linear with the line length, but they are even more effective in reducing the standard deviation of the line delay, and the optimization of clock period of wafer scale well-balanced H-tree can be implemented by appropriately inserting inverters in clock paths.

Acknowledgments

This work was supported in part by Grant-In-Aid for Scientific Research in JSPS (Japan Society for the Promotion of Science).

References

[1]N.Nigam and D.C.Keezer, "A comparative study of clock distribution approaches for WSI," Proc. IEEE Int'l Conf. Wafer Scale Integration, pp.243-251, Jan. 1993.

[2]N.G.Sheridan, C.M.Habiger and R.M.Lea, " WSI clock and signal distribution: a novel approach, " Proc. IEEE Int'l Conf. Wafer Scale Integration, pp.235-242, Jan. 1993.

[3] M.D'ABREU , "Understanding of the fabrication process - key to design and test of mixed signal ICs". Proc. European Test Workshop, 1998

[4] M.Eisele, J.Berthold, D.Schmitt-landsiedeld and R.Mahnkopf, "The impact of intra-die device parameter variations on path delays and on the design for yield of low voltage digital circuits," IEEE Trans.VLSI Syst., vol.5, no.4, pp. 360-368,1997.

[5] M.Nekili, G.Bois and Y.Savaria, " Pipelined H-trees for high-speed clocking of large integrated systems in presence of process variations," IEEE Trans. VLSI Syst., vol.5, no.2, pp.161-174, 1997.

[6] H.B.Bakoglu, J.T.Walker and J.D.Neindl, "A symmetric clock distribution tree and optimized high-speed interconnections for reduced clock skew in ULSI&WSI circuits," Proc. IEEE Int. Conf. Computer Design: VLSI in Computers, ICCD'86.

[7] S.D.Kugelmass, and K.Steiglitz, "An upper bound of expected clock skew in synchronous systems," IEEE Trans. Comput., vol.39, no.11, pp.1475-1477, 1990.

[8] M.Afghahi, and C.Svensson, " Performance of synchronous and asynchronous schemes for VLSI systems," IEEE Trans. Comput, vol.41,no.7, pp.858-872, 1992.

[9] T.Gneiting, and I.P.Jalowiecki, "Influence of process parameter variations on the signal distribution behavior of wafer scale integration devices," IEEE Trans. Components, Packaging, and Manufacturing Technology – Part B, vol.18, no.3, pp.424-430, 1995.

[10] X.H.Jiang and S.Horiguchi "Distribution analysis of clock skew and clock delays for general clock distribution networks," JAIST Research Report (ISSN.0918-7553) IS-RR-2000-014, pp.1-20, June. 2000.

[11] Jim Pitman, "Probability," pp.454-461, New York,1993.

[12] HIT-Kit: Statistical Circuit Simulation, Austria Mikro System International AG, 1998.

[13] D.C.Keezer and V.K.Jain, "Design and evaluation of Wafer Scale clock distribution," Proc. IEEE Int'l Conf. Wafer Scale Integration, pp.168-175, 1992.

[14] H.B.Bakoglu and J.D. Meindl, "Optimal interconnection circuits for VLSI," IEEE Trans. Electron Devices, Vol.ED-32, no.5, pp.903-909, 1985.

[15] R.Lakshmikumar, A.Hadaway, and M.A.Copeland, " Characterization and modeling of mismatch in MOS transistors for precision analog design," IEEE J. Solid-State Circuits, Vol.SC-21, pp.1057-1066, Dec.1986.

[16] C.Svensson and M.Afghahi,, " On RC line delay and scaling in VLSI systems," Electronics Letters, vol.24, no.9,pp.562-563,1988.

[17] X.H. JIANG and G.A.ALLAN, "Efficient delay yield estimate of digital circuits," Electronics Letters, vol.35, no.24, pp.2109-2110, 1999.

[18] A.L.Fisher and H.T.Kung, "Synchronizing large VLSI processor arrays," IEEE Trans. Comput.,vol.C-34,1985.

[19] 'Probability and Statistics,' Academic Press, Peking, China, 1985.

[20]X.H.Jiang and S.Horiguchi "A new statistical skew model used for clock period optimization of H-tree clock distribution networks," JAIST Research Report (ISSN.0918- 7553) IS-RR-2000-012, pp.1-26, May.2000.

[21] Michael D.Vose, The Simple Genetic Algorithm: Foundations and Theory, MIT Press, 1999.

Figure Captions:

Fig.1 An H-tree clock distribution network for 256 processors in a 4 inch WSI (Clock buffers are not illustrated here).

Fig.2 Simulation results and theoretical results in mean values of clock skew.

Fig.3 Simulation results and theoretical results in mean values of the largest clock delay.

Fig.4 Simulation results and theoretical results in variances of clock skew and the largest clock delay.

Fig.5 Simulation results and the theoretical results in yield of clock skew of H-tree for two combinations of parameters *m*, *h* and *W*.

Fig.6 Simulation results and the theoretical results in yield of the largest clock delay of H-tree for two combinations of parameters *m*, *h* and *W*.

Fig. 7 Insertion of drivers in an H-tree metallic path.

Fig.8 The trends of expected values of clock skew and the largest clock delay with the variations of driver number and driver size in each path.

Fig.9 Standard deviation vs. mean value of the largest clock delay.

Fig.10 Power and area vs. the expected largest clock delay (notice that the two curves do not follow the same ordinate scale: one is a area and the other is a power).

Fig.11 The trends of expected values of clock skew and line segment delay with the variations of driver number and driver size in each path.

Fig.12 Standard deviation vs. mean value of clock skew.

Fig.13 Power and area vs. the expected clock skew (notice that the two curves do not follow the same ordinate scale: one is a area and the other is a power).

Fig.14 Distribution of clock skew of a well-balanced H-tree.

Biographies:

Xiaohong Jiang received BS and MS degrees in applied mathematics in 1989 and 1992,respectively, and the PhD degree in Solid-State Electronics and Microelectronics in 1999, all from Xidian University, Xi'an, China. He is currently a JSPS (Japan Society for the Promotion of Science) postdoctoral research fellow at the Japan Advanced Institute of Science and Technology (JAIST). Dr. Jiang was a research associate in the Department of Electronics and Electrical Engineering, the University of Edinburgh from Mar. 1999 - Oct. 1999. Dr. Jiang's research interests include IC yield modeling, timing analysis of digital circuits, clock distribution and fault-tolerant technologies for VLSI and WSI, performance analysis and modeling of optical interconnection networks. He has published more than 20 technical papers in these areas.

Susumu Horiguchi graduated from Department of Communication Engineering, Tohoku University in 1976, and received the MS. and Dr. degrees both from the same university in 1978 and 1981, respectively. Currently and since 1992, he has been a full Professor of the Graduate School of Information Science at the Japan Advanced Institute of Science and Technology (JAIST). He was a faculty of Department of Information Science at Tohoku University from 1981 to 1992. He was a visiting scientist of IBM Thomas J. Watson Research Center from 1986 to 1987 and a visiting professor of The Center for Advanced Studies at the University of Southwestern Louisiana in 1994. He has been conducting his research group as the chair of Multi-Media Integral System Laboratory at JAIST. He has been involved in organizing many international workshops, symposia and conferences sponsored by IEEE, ACM, and IEICE. His research interest has been mainly concerned with parallel computer architectures, VLSI/WSI architectures, interconnection networks, parallel computing algorithm, massively parallel processing, and Multi-Media Integral System. Dr. Horiguchi is a senior member of IEEE Computer Society and a board member of The Information and System Society of IEICE.