Modeling of thermally induced skew variations in clock distribution network

Modeling of Thermally Induced Skew Variations in Clock Distribution Network

Alessandro Sassone, Wei Liu, Andrea Calimera, Alberto Macii, Enrico Macii and Massimo Poncino

Politecnico di Torino, 10129, Torino, ITALY

***Abstract*—Clock distribution network is sensitive to large thermal gradients on the die as the performance of both clock buffers and interconnects are affected by temperature. A robust clock network design relies on the accurate analysis of clock skew subject to temperature variations. In this work, we address the problem of thermally induced clock skew modeling in nanometer CMOS technologies. The complex thermal behavior of both buffers and interconnects are taken into account. In addition, our characterization of the temperature effect on buffers and interconnects provides valuable insight to designers about the potential impact of thermal variations on clock networks. The use of industrial standard data format in the interface allows our tool to be easily integrated into existing design flows.**

I. Introduction

With higher integration of MOS transistors in VLSI circuits, the power consumed, and more important the power consumed per unit area (i.e., the power density), have increased dramatically resulting in high on-chip temperatures. In addition, due to the extensive use of power management strategies (like power-gating, dynamic voltage-scaling, and adaptive body biasing), integrated circuits are prone to show a non uniform power density distribution across the die. As a main consequence, local hotspots (i.e., localized layout region with higher temperatures compared to the average die temperature) are becoming more common. It has been reported in [1] that thermal gradients larger than 50 °C already exist in high performance ICs.

High operating temperatures combined with large on-chip temperature fluctuations can induce sensible variation on the electrical characteristics of both active (i.e., transistor) and passive (i.e., interconnection) devices. Concerning wires, the resistance of metal increases linearly with temperature leading to significant delay degradation [2]. Concerning devices, in sub-90nm technologies, the propagation delay through a gate (or cell) can vary in complex ways as temperature changes [3]. That is, depending on various parameters, such as cell size, load, supply voltage, and threshold voltage (V_{th}), the delay can either increase (Direct Temperature Dependence) or decrease (Inverted Temperature Dependence).

These complicated thermal dependencies may cause significant device mismatch, and thus make circuit design and analysis extremely challenging. This is especially true when considering critical design elements that are vital to the operation of synchronous circuits, like the clock distribution networks (CDN), also called clock-tree when the network is routed following a tree-like shape. A typical CDN is made up of buffered global interconnects which span the entire circuit layout and that, for this reason, is more sensible to temperature variations on the substrate. The consequence is the possibility of an increase in clock skew which may induce synchronization errors and, in the worst case, circuit failures [4], [5]. Hence, designing robust clock distribution networks for nanometer technologies requires synthesis and analysis tools that account for the impact of spatial and temporal temperature variations.

From a design viewpoint, several recent works propose temperature-aware clock tree optimizations in which the temperature is considered as a direct design variable during the synthesis of a clock-tree. In [6], for instance, the authors modified the traditional Deferred-Merge Embedding (DME) method (traditionally used for minimizing total wirelength and power consumption with zero or bounded skew constraints) to minimize clock skew for both uniform and non-uniform thermal profiles. In [7], the authors optimize clock trees that are subject to time-varying thermal profiles by perturbing an initial clock tree obtained by DME method to minimize worst case skew and wirelength for all possible thermal profiles. In [8], the authors used sequence linear programming methods to minimize clock skew in a buffered clock tree. Finally in [9], [10], [11], the authors studied runtime skew minimization by using dynamically adjustable delay clock buffers, whose drive strength is tuned according to the temperature variations monitored by thermal sensors in different regions of the die.

In this work, we face the other side of the problem, that is, the development of tools for efficient yet accurate analysis of temperature induced clock-skew variation. We propose a clock skew modeling method, which is essential for accurate clock network analysis in high performance ICs where large thermal gradients exist. The proposed method is applied at post clock tree synthesis (CTS) stage and takes advantage of the accurate parasitics extraction from the layout. The method takes into account both the performance degradation in clock buffers, and metal interconnects. We provide two simulation techniques for analyzing the extracted clock tree, one for accuracy and the other for efficiency. The proposed method can be seamlessly integrated with existing back-end physical design flow as plug-ins to provide thermal aware clock skew analysis. Two benchmark circuits implemented in a 65nm technology are used to analyze the impact of within-die thermal gradients on clock skew, considering temperature effect on both the active devices and the interconnect system.

II. Background on Thermal Effects

A. Temperature Effect on Interconnect Delay

The electrical resistance of metal has a linear relationship with temperature and can be expressed as:

$$R(x) = R_0(1 + \beta \cdot T(x)) \tag{1}$$

where R_0 is the resistance at reference temperature, β is the temperature coefficient ($1/^\circ$C) and $T(x)$ is the temperature profile along the length of the wire. The value of β for copper at room temperature is $3.9E-3$, which means for every $10\,^\circ$C increase in temperature, the resistance would increase by 3.9%.

According to the distributed RC Elmore delay model [12], signal propagation delay through the interconnect of length L can be written as:

$$D = R_d\left(C_L + \int_0^L c_0(x)dx\right) + \int_0^L r_0(x)\cdot\left(\int_x^L c_0(\tau)d\tau + C_L\right)dx \tag{2}$$

where R_d is the driver cell's ON resistance, $c_0(x)$ and $r_0(x)$ are the capacitance and resistance per unit length at location x and C_L is the load capacitance.

Combining Equation (1) and Equation (2), we can obtain a temperature dependent interconnect delay model:

$$D = D_0 + (c_0L + C_L)r_0\beta\int_0^L T(x)dx - c_0r_0\beta\int_0^L x\cdot T(x)dx \tag{3}$$

where

$$D_0 = R_d(c_0L + C_L) + \left(c_0r_0\frac{L^2}{2} + r_0LC_L\right) \tag{4}$$

is the Elmore delay of the interconnect corresponding to the unit length resistance at reference temperature.

According to [2], the temperature within the interconnect for a given substrate temperature can be expressed as:

$$T(x) = T_{sub} + \frac{\theta}{\lambda^2}\left(1 - \frac{\sinh\lambda x + \sinh\lambda(L-x)}{\sinh\lambda L}\right) \tag{5}$$

where θ and λ are constants for a chosen metal layer in a specific technology node and depend on the thermal conductivity of metal and insulator, on their geometries, and on the electrical parameters of the interconnect (current density and resistivity).

The peak temperature rise is equal to θ/λ^2 for interconnects whose lengths are larger than the heat diffusion length. As the clock network is usually routed in global metal layers, the distance from the substrate is large (t_{ins} is larger than local metal layers) and as a result the temperature rise in the clock network can be quite high.

B. Temperature Effect on CMOS Transistor Delay

The output transition time of a CMOS transistor can be obtained from computing the ratio between the total amount of charge transferred and the charging/discharging current. The ON state drain current I_{ds} of a short channel MOSFET is usually expressed using the *alpha-power* model [13]:

$$I_{ds}(T) = \begin{cases} \mu(T)\frac{W}{L_{eff}}P_l(V_{gs} - V_{th}(T))^{\frac{\alpha}{2}}V_{ds} & V_{ds} \le V_{dsat} \\ \\ v_{sat}(T)WP_s(V_{gs} - V_{th}(T))^{\alpha} & V_{ds} \ge V_{dsat} \end{cases} \tag{6}$$

In the expression for I_{ds}, three parameters are strongly dependent on temperature: carrier's mobility μ in the linear region, saturation velocity v_{sat} in the saturation region and threshold voltage V_{th}.

The temperature dependence of the carrier mobility is expressed as,

$$\mu(T) = \mu(T_0)\left(\frac{T_0}{T}\right)^m \tag{7}$$

where T is the junction temperature, T_0 is the nominal temperature (typically at $300\,K$) and m is the temperature coefficient, which is ideally 1.5 but can vary depending on the process. The temperature dependence of the saturation velocity has a more linear relationship with temperature, and the dependence is weaker than that of the mobility,

$$v_{sat}(T) = v_{sat}(T_0) - h(T - T_0) \tag{8}$$

The temperature coefficient h has an extracted value around $150\ ms^{-1}\cdot K^{-1}$. The temperature dependence of the threshold voltage can be expressed as,

$$V_{th}(T) = V_{th}(T_0) - k(T - T_0) \tag{9}$$

where k is the temperature coefficient, whose value is measured to be around $0.8\ mV\cdot K^{-1}$.

It is evident that all three parameters decrease as temperature increases, however, they affect the drain current in different ways. While a lower μ (in the linear region) or v_{sat} (in the saturation region) causes the drain current to decrease, a lower V_{th} (in both linear and saturation regions) causes the drain current to increase. Depending on which parameter dominates, the delay of a transistor will either increase or decrease as temperature increases.

C. Thermally-Induced Clock Skew

Clock distribution network is one of the largest on-chip interconnect network and plays a crucial role in the correct operation of synchronous systems. The clock network can be implemented in different topologies, including mesh, tree and hybrid structures.

The mesh structure provides many parallel paths between clock source and clock sinks, thus is more robust against on-chip variations. However, clock mesh is expensive in terms of wirelength, power and routing. The tree structure, on the other hand, is more economical and can usually achieve the shortest wirelength for the implementation of a clock network. However, clock tree is more vulnerable to variations since there is only one path from the clock source to any sink. A hybrid structure, where the clock signal is globally distributed in a mesh and locally routed in a tree, can provide low cost yet robust clock network for high performance designs.

In a clock distribution network, clock edges may arrive at the sinks at different time due to delay unbalance. This variation in arrival time of the clock signal is commonly known as *clock skew*. Clock skew is defined as the maximum difference between the arrival time of the clock signal at the input pin of any two sinks:

$$skew = max(t_i - t_j) \qquad i, j \in S \qquad (10)$$

where S is the set of clock input pins.

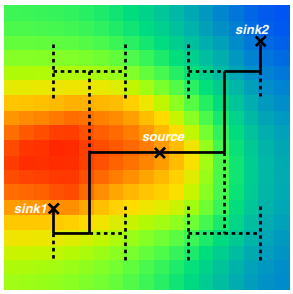


Fig. 1. Paths subject to different thermal profile develop different delays.

In the presence of large thermal gradients, the performance of buffers and interconnects in the clock network can vary which might introduce extra skew. Figure 1 illustrates such an example, where the clock network is implemented in a H-tree and equi-wirelength from the source to sinks is achieved. However, the path from *source* to *sink1* (path1) is subject to a much higher temperature than the path from *source* to *sink2* (path2). Path1, despite being as long as Path2, has a larger delay due to the increase in wire resistance. The temperature induced clock skew would not be discovered in traditional clock network analysis tools.

III. A Framework for Temperature-Aware Clock-Skew Analysis

The main limitation of commercial design tools is that the analysis of the clock skew is done considering a uniform thermal profile, i.e., without considering thermal gradients. As described in the previous sections, however, real-life circuits may show spatial temperature variations that can lead to substantial skew increase. In this section we describe the thermal-aware analysis framework we have implemented to take into consideration the deleterious effects induced by thermal gradients on both the active buffers and the interconnects of the clock-tree.

A relational flow-chart of the proposed framework is shown in Figure 2. The clock skew simulator (dark box in the Figure) is based on SPICE-level simulation and works after the Clock-Tree-Synthesis (CTS). At this stage of the design flow, the HDL description of the design has been already synthesized and placed using standard row-based layout organization, while the clock distribution network is routed and integrated through the circuit.

Having a full placed&routed netlist allows to accurately annotate the parasitics, which are made available using a Standard Parasitic Exchange Format (SPEF). The physical, information together with the power consumption profile of the circuit, are also used to estimate the thermal map of the circuit.

The clock skew analyzer takes as input the aforementioned information, i.e., the circuit netlist, the SPEF file, the thermal map, and performs a temperature-driven delay simulation. To notice that SPICE models provided by the silicon vendors must be given to the simulator. At the end of the simulation, the analyzer generate a final report, also including values of the clock skew, the longest and the shortest timing paths, as well as the maximum and minimum arrival time of all the buffers at each level of the clock tree.

The tool is fully compatible with commercial physical design tools and ready to be integrated with industrial back-end design flows.

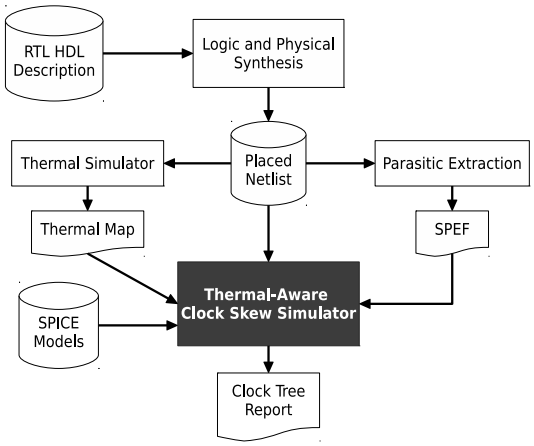


Fig. 2. Post-layout flow for our methodology.

A. Clock-Skew Simulator

The clock-skew simulator is in charge of collecting the physical and thermal information, and run a temperature-driven clock simulation. A key step during this process is the generation of a SPICE-level netlist of the clock-tree. Such a netlist, besides including the topological structure of the clock-tree, i.e., how and which buffers are interconnected each others, has to report detailed physical information about the parasitics associated to the metal interconnects, as well as the operating temperature of each and every element.

We implemented customized TCL scripts to be used for parsing the input files and automatically generate a SPICE-compliant netlist. The clock tree is modeled as a distributed RC networks (whose values are read from the SPEF) spaced up by the buffers. For the buffers we used spice models provided by the silicon vendor. The sinks of the clock tree (i.e., the flip-flops belonging to the logic circuits) are modeled as load capacitances. Also in this case, we used the datasheet provided by the silicon vendor. The root of the tree is driven by an ideal supply voltage. The locations of clock buffers and RC elements in the netlist are analyzed against the thermal map to set their temperature values. Temperature of a buffer is simply the value

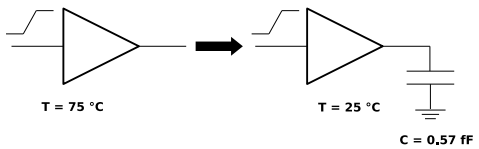
at the same location in the thermal map, while temperature of an RC element is the value computed using Equation (5). This allows to take into consideration the self-heating effect of wires.

We provide two different simulation strategies for analyzing the extracted clock tree. The main difference is the underlying tool used as simulation engine. The first technique aims at accuracy, and indeed relies on Synopsys' HSPICE, a highly accurate industrial SPICE simulator. Main limitations of HSPICE are the large simulation time, as well as the CPU and RAM usage. This may limit the maximum size of the clock-tree that can be simulated. The second technique aims at efficiency, and instead relies on Synopsys' HSIM, a fast-SPICE simulator which degrades the quality of the simulation (ignoring second order effects) to achieve substantial performance and capacity gains.

B. Accounting for Temperature During SPICE Simulations

In order to analyze the effects that thermal gradients induce on the clock skew, it is important to use simulation engines which allow different element of the netlist to work at a different temperature. While HSPICE supports such feature, HSIM does not; that is, while with HSPICE it is possible to assign to each and every element a specific temperature, HSIM allows to simulate circuit with different, but uniform temperatures.

To overcome this drawback, we implemented a kind of *capacitive-based thermal emulator*. The latter consists of an extra capacitor connected to the output pin of each buffer. The main effect of such additional load is to increase the propagation delay of the buffer. The amount of delay degradation is linearly proportional to the value of the capacitance. We empirically extracted a look-up-table (LUT) which provides us with the right capacitive load to be used for emulating a given shift on the operating temperature[1].



$Tran\,(ps)$	$Temp\,(^\circ C)$	$Cap\,(fF)$
20	50	0.30
20	75	0.57
20	100	0.81
20	125	1.03

Fig. 3. Modeling of thermal behavior in clock buffers for *HSIM*.

Figure 3 shows an example of the proposed technique. The delay of a buffer at 75 °C is calculated by simulating the buffer coupled with a capacitor at 25 °C. The LUT lists the capacitance values when considering a transition time of 20 ps. As temperature increases, a larger capacitance value is needed to match the increasing thermally induced delay.

[1]This technique is effective whenever the buffers show a direct temperature dependence, that is the case for our library

C. Thermal Simulator

Clock skew analysis relies on the accurate estimation of temperature profile within the chip. The heat generated in the transistor junctions during signal transition is mainly transferred to the ambient environment through conduction. In general, heat conduction can be modeled using Fourier's law. In our work, we use the Finite Difference Method (FDM) to solve the steady-state heat diffusion problem. Our thermal simulation method is described in detail in [14] and we briefly summarized it below.

Using the FDM method, a chip is meshed into a 3D grid of thermal cells. Thermal cells in the mesh are modeled as a subcircuit composed of resistors, capacitors and current sources based on the analogy between heat diffusion in the thermal domain and current flow in the electrical domain. The thermal mesh can therefore be converted to an equivalent RC circuit to be solved using circuit analysis techniques.

Since meshing the chip at the tiny size of metal wires would result in an excessive number of thermal cells, we only obtain temperatures on the device layer and temperatures on metal layers are computed using Equation (5). We use SPICE to solve the equivalent RC circuit to obtain the nodal voltage within each thermal cell, which, according to the thermal-electrical analogy, is in fact the temperature in the center of the thermal cell. Using the layout and power consumption information of the standard cells at the post-placement stage, our thermal simulator can produce a highly accurate thermal map. The obtained thermal profile is then used by the clock skew analyzer to estimate performance degradation in the clock buffers and metal interconnects.

IV. Experimental Results

In this section, we first provide a full characterization of the performance degradation of the clock buffers and the interconnects due to temperature variations. Second, we present the results obtained using our tool when applied on two benchmark circuits mapped onto an industrial 65nm technology provided by STMicroelectronics.

A. Characterization of Delay in Buffers

As described in Section II-B, the thermal behavior of CMOS transistors can be quite complex. It is, therefore, very important to know the trend of thermally induced delay variation in clock buffers in order to understand the potential impact of temperature effect on clock tree.

Synopsys HSPICE has been used to characterize the thermal profile of the buffers. Different fanout loads have been considered, while the temperature has been progressively swapped from $25°C$ to $125°C$. The collected data have been also used to extract the temperature-to-capacitance LUT described in Section III-B.

Figure 4 shows the results for a medium sized buffer, the *BFX4*. As one can see, the buffer shows a direct temperature dependence, i.e., the delay increases with temperature (on average 5.4% from room temperature to $125°C$), and for any

fanouts. Similar numbers have been observed for all the other cells.

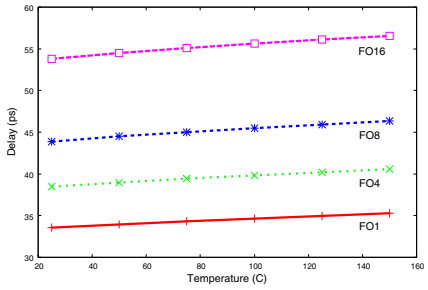


Fig. 4. Temperature induced delay increase in clock buffers.

B. Characterization of Delay in Interconnects

The thermal behavior of metal wires is much simpler than transistors, as the resistance increasing linearly as temperature rises.

In order to optimize root-to-sinks signal propagation, standard clock tree synthesis algorithms avoid long wire segments. This helps to minimize the transition times and the raise/ fall edges of the the clock signal, as well as to avoid antenna rules violation. In our technology, for instance, we experimented a maximum wirelength of few hundreds of micrometers ($<$500μm). Using such short wires can guarantee small flight times across the network that, as consequence, shows a path delay dominated by clock buffers.

Figure 5 shows a comparison between thermally-induced delay degradation on wires of different lengths and a buffer, the $BFX4$. The delay of the buffer is close to the delay of long wires ($>$500 μm), but quite far from shorten wires, that are those used in our clock-tree. However, as technology scales to finer geometries, we expect the crossing point will move towards shorter wires since the delay of cells and global interconnects scale in the opposite direction.

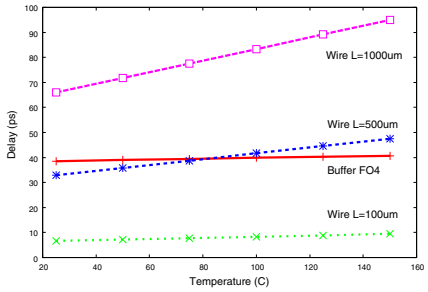


Fig. 5. Temperature induced delay increase in interconnects.

C. Clock Skew in Benchmark Circuits

Two are the circuits used as benchmarks for testing the implemented clock-skew analyzer: a quad-core floating point unit, called $FPU4$ hereafter, and a configurable synthetic benchmark circuit $SYNTH$.

The FPU4 is composed of four identical double precision FPUs that can be power-managed (i.e., activated or turned-off) depending on the actual workload. The SYNTH is a two dimensional grid of "micro-heater" blocks where the underlying power density can be precisely controlled to create arbitrary thermal maps. Each micro-heater consists of a variable number of different sized parallel inverter chains. The synthesized clock-trees consist of 143 buffers with 3472 sinks for FPU4, and 184 buffers with 4800 sinks for SYNTH.

Experimental results are reported in Table I and Table II for the two type of simulation engines we used, HSPICE and HSIM respectively. For each benchmark, different thermal maps have been analyzed (individual rows in the two tables): *Uniform* consists of a flat temperature distribution of 125 ($^\circ$C) across the layout (typical case of standard physical design tools); *Real* is the temperature distribution obtained from our thermal simulator considering a typical workload and the ambient temperature at 25 ($^\circ$C); the remaining thermal maps emulate hot surrounding components and have been obtained by forcing a high temperature (i.e., 150 $^\circ$C) at the four edge of the circuit layout, North ($N150$), East ($E150$), South ($S150$), West $W150$.

For both the two tables, Column $\Delta Temp$ shows the maximum temperature gradient inside the circuit layout; Columns *Longest Path* and *Shortest Path* show the path delay across the longest and shortest paths of the tree (the difference between the twos represents the clock-skew); Column *Global Skew* reports the maximum skew on the clock-tree also showing the variation w.r.t. the *Uniform* case; Columns *Sim Time* and *Mem Usage* show the execution time and memory usage taken by the simulation framework. In the rightmost column of Table II we also reported the accuracy error of HSIM w.r.t. HSPICE.

The clock skew obtained when considering real thermal maps, i.e., the *Real* case, is always larger than the *Uniform* case: 12.2% and 10.2% larger in the worst case for FPU4 and SYNTH respectively. This highlights once again how standard corner-based tools understimate clock-skew in the presence of thermal gradients. Moreover, one should observe how sourrounding elements may play a foundamental role in determining the actual clock-skew of a circuit. Large skew variations w.r.t. the uniform case (23.2% and 26.8% for FPU and SYNTH respectively in the worst-case) denote in fact a strong dependence from sources of heat other than internal active transistors. Needless to say, extracting such analysis with standard tools is hard to be obtained, if not impossible.

Concerning the two simulation engines, we empirically observed that HSIM can run in a smaller time and allocate less RAM while guaranteeing reasonable small errors ($<$2.5% in most of the cases). It is worth emphasizing that the use of HSIM is allowed only thanks the capacitive-based thermal emulator we introduced in SectionIII-B. Ongoing experiments are also showing that when using larger benchmarks, the memory requirement of HSPICE rocket to the point that simulation will abort. On the contrary HSIM can still complete.

V. Conclusions

In the presence of large thermal gradients, the delay degradation in clock buffers and interconnects may cause additional

TABLE I
EXPERIMENTAL RESULTS (HSPICE TECHNIQUE)

Benchmark	*Thermal Map*	$\Delta Temp$ (°C)	*Longest Path (ps)*	*Shortest Path (ps)*	*Global Skew (ps)*	*Sim Time (s)*	*Mem Usage (MB)*
FPU4	Uniform	0.0	311.31	271.88	39.4	25	89.6
	Real	60.0	320.50	276.29	44.2 (12.2%)	27	89.6
	N150	93.0	326.11	278.20	47.9 (21.5%)	26	89.6
	E150	79.3	326.28	279.90	46.4 (17.7%)	27	89.6
	S150	50.1	324.47	275.90	48.6 (23.2%)	26	89.6
	W150	57.2	325.25	277.03	48.2 (22.3%)	27	89.6
SYNTH	Uniform	0.0	318.39	239.38	79.0	34	115.5
	Real	9.8	329.93	242.89	87.0 (10.2%)	35	115.5
	N150	18.9	343.10	242.96	100.2 (26.8%)	34	115.5
	E150	33.3	342.63	248.27	94.4 (19.4%)	37	115.5
	S150	37.1	341.56	248.33	93.2 (18.0%)	35	115.5
	W150	33.7	341.20	246.73	94.5 (19.6%)	36	115.5

TABLE II
EXPERIMENTAL RESULTS (HSIM TECHNIQUE)

Benchmark	*Thermal Map*	$\Delta Temp$ (°C)	*Longest Path (ps)*	*Shortest Path (ps)*	*Global Skew (ps)*	*Sim Time (s)*	*Mem Usage (MB)*	*Error (%)*
FPU4	Uniform	0.0	306.79	267.05	39.75	24	56.25	0.8
	Real	60.0	312.54	267.35	45.19 (13.7%)	26	56.25	2.2
	N150	93.0	316.02	268.19	47.82 (20.3%)	26	56.25	-0.2
	E150	79.3	319.04	268.15	50.89 (28.0%)	26	56.25	9.7
	S150	50.1	315.48	268.07	47.42 (19.3%)	27	57.37	-2.4
	W150	57.2	314.84	267.56	47.28 (18.9%)	26	56.25	-1.9
SYNTH	Uniform	0.0	313.34	235.35	77.99	36	62.19	-1.3
	Real	9.8	322.17	235.71	86.46 (10.9%)	37	61	-0.7
	N150	18.9	336.55	236.42	100.13 (26.8%)	38	68.75	-0.0
	E150	33.3	331.26	235.97	95.29 (19.4%)	38	68.62	1.0
	S150	37.1	330.24	235.96	94.28 (18.0%)	38	68.75	1.1
	W150	33.7	332.90	236.38	96.52 (19.6%)	38	68.75	2.2

path unbalance in the clock distribution network and exhibits as an increase in clock skew. We showed that the thermally induced delay degradation is more significant in interconnects than buffers although path delay is still dominated by buffers in the 65 nm process. The exact changes in clock skew caused by temperature variation can only be obtained by performing a detailed modeling of the temperature dependent delay analysis of the clock network. The thermal aware clock skew analysis tool proposed in our work, provides accurate yet efficient simulation based skew analysis that can take into account the temperature effect.

REFERENCES

[1] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," in *Proc. of Design Automation Conference*, June 2003, pp. 338 – 342.

[2] A. Ajami, K. Banerjee, and M. Pedram, "Modeling and analysis of nonuniform substrate temperature effects on global ULSI interconnects," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 6, pp. 849–861, June 2005.

[3] R. Kumar and V. Kursun, "Reversed temperature-dependent propagation delay characteristics in nanometer cmos circuits," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 53, no. 10, pp. 1078 –1082, oct. 2006.

[4] S. Bota, M. Rosales, J. Rosello, A. Keshavarzi, and J. Segura, "Within die thermal gradient impact on clock-skew: a new type of delay-fault mechanism," in *Proc. of International Test Conference*, oct. 2004, pp. 1276 – 1283.

[5] S. Bota, J. Rossello, C. de Benito, A. Keshavarzi, and J. Segura, "Impact of thermal gradients on clock skew and testing," *IEEE Design Test of Computers*, vol. 23, no. 5, pp. 414 –424, may 2006.

[6] M. Cho, S. Ahmedtt, and D. Pan, "Taco: temperature aware clock-tree optimization," in *Proc. of International Conference on Computer-Aided Design*, Nov. 2005, pp. 582–587.

[7] H. Yu, Y. Hu, C. Liu, and L. He, "Minimal skew clock embedding considering time variant temperature gradient," in *Proc. of International Symposium on Physical Design*, 2007, pp. 173–180.

[8] K. Athikulwongse, X. Zhao, and S. K. Lim, "Buffered clock tree sizing for skew minimization under power and thermal budgets," in *Proc. of 15th Asia and South Pacific Design Automation Conference (ASP-DAC)*, jan. 2010, pp. 474 –479.

[9] A. Chakraborty, K. Duraisami, A. Sathanur, P. Sithambaram, L. Benini, A. Macii, E. Macii, and M. Poncino, "Dynamic Thermal Clock Skew Compensation Using Tunable Delay Buffers," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 6, pp. 639–649, June 2008.

[10] T. Ragheb, A. Ricketts, M. Mondal, S. Kirolos, G. Links, V. Narayanan, and Y. Massoud, "Design of thermally robust clock trees using dynamically adaptive clock buffers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 2, pp. 374 –383, Feb. 2009.

[11] J. Long, J. C. Ku, S. Memik, and Y. Ismail, "Sacta: A self-adjusting clock tree architecture for adapting to thermal-induced delay variation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 9, pp. 1323 –1336, Sept. 2010.

[12] W. C. Elmore, "The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers," *Journal of Applied Physics*, vol. 19, no. 1, pp. 55–63, Jan. 1948.

[13] T. Sakurai and A. Newton, "Alpha-power law mosfet model and its applications to cmos inverter delay and other formulas," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 2, pp. 584 –594, apr 1990.

[14] W. Liu, A. Calimera, A. Nannarelli, E. Macii, and M. Poncino, "On-chip Thermal Modeling Based on SPICE Simulation," *Proc. of 19th International Workshop on Power And Timing Modeling, Optimization and Simulation (PATMOS 2009)*, pp. 66–75, Sept. 2009.