A Multi-Level Transmission Line Network Approach for Multi-Giga Hertz Clock Distribution

Hongyu Chen and Chung-Kuan Cheng

Department of Computer Science and Engineering, University of California, San Diego

Abstract—In high performance systems, process variations and fluctuations of operating environments have significant impact on the clock skew. Recently, hybrid structures of H-tree and mesh [2,15,18,19] were proposed to distribute the clock signal with a balanced H-tree and lock the skew using the shunt effect of the mesh. However, in multi-giga hertz regime, the RC model [15] of the mesh is no longer valid. The inductance effect of the mesh can even make the skew worse. In this paper, we investigate the use of a novel architecture which incorporates multiple level transmission line shunts to distribute global clock signal. We derive the analytical expression of the skew reduction contributed by the shunt of a transmission line with the length of an integral multiple of clock wavelength. Based on the analytical skew expression, we adopt convex programming techniques to optimize the wire widths of the multi-level transmission line network. Simulation results show that the multilevel network achieves below 4ps skew for 10GHz clock.

I. INTRODUCTION

With increasing clock frequency, the clock skew caused by many nondeterministic factors such as process variations, supply voltage fluctuation and temperature gradient consumes a significant portion of clock period. For high performance synchronous circuitry, the design of a robust global clock distribution system which can sustain various parameter variations becomes an increasingly challenging task.

RC shunted networks have been successfully used to reduce the clock skew under process variations. In [11], three wide spine shunts are adopted to reduce the skew between the leaf nodes of a very deep buffer tree. In [15,18,19], a clock mesh driven by a balanced H-tree is used for global clock distribution. In [2,15], multiple level shunts are applied to further reduce the clock skew.

However, when the clock frequency increases to multi-giga hertz range, the inductance effect of the shunt wires becomes significant. At 10GHz clock rate, the time of flight between two corners of a chip is comparable to the clock cycle. The RC model [15] of the shunt effect is no longer valid. The inductance of the shunt can even cause worse skew. We have to view the shunt wires as transmission lines.

On the other hand, a transmission line with a properly tailored length can lock the oscillators together. Galton et al. [7] showed that when the wire length is shorter than one quarter wavelength, the transmission line can synchronize oscillators both in phase and magnitude. Several other researches utilized the synchronization capability of the transmission line in clock distribution by connecting the distributed PLLs together with transmission lines [8,9,20]. In order to compensate the lossy nature of the on-chip transmission lines, O'Mahony et al. [14] use distributed transconductors along the transmission line to generate the standing wave.

In this paper, we present a hybrid structure of H-tree and the transmission line shunts. We simplify the clock to differential signals of sinusoidal waves. The transmission lines are driven at discrete points and bent into spiral pattern

in order to link the drivers of the H-tree. The drivers of the H-tree are shunted level by level. The shunt lengths between the drivers are an integral multiple of wavelength. For an ideal case that the line is lossless, a standing wave can lock the clock drivers to zero skew. For lossy shunts, we derive the skew reduction as a function of the wire width. We optimize the wire widths of the multi-level network based on the analytical skew function.

Comparing with other methods, the proposed global clock distribution architecture enjoys several advantages. First, there is no direct feedback path from the transmission line network to the clock source. The transmission lines are linear network and thus the design and optimization involve no active components. Second, the energy storage capability of the locked standing wave in the transmission line can mitigate the clock jitter. And finally, the power consumption of the whole network is much less because the resonance effect of the transmission line.

Our contributions in this paper include:

1. We derive the analytical expression of the skew reduction effect of a multiple wavelength-long transmission line shunt.
2. We construct multi-level spiral networks for multi-giga hertz global clock distributions.
3. We optimize the wire widths of spirals of each level using convex programming techniques. Simulation results demonstrate that the optimized clock network can achieve below 4ps skew at 10GHz.

The rest of this paper is organized as follows: In Section 2, we formulate the problem. In Section 3, we derive the analytical skew expression for multiple wavelength long transmission line shunted network. In Section 4, we describe the implementation of multiple level transmission line spiral network structure and use convex programming approach to optimize it. We then present the experimental results in Section 5. Finally, we conclude the paper in Section 6.

II. PROBLEM FORMULATION

The goal of this work is to construct a global clock distribution network that has a working frequency higher than 5GHz and is robust against various parameter variations.

In subsection A, we propose to use differential sinusoidal wave to distribute global clock signal. We describe the model of variations in subsection B, and introduce the multi-level network in subsection C. After that, we formulate the optimal spiral sizing problem in subsection D.

A. Differential sinusoidal waves

We adopt differential sinusoidal waves in the global clock distribution. The sinusoidal waveform simplifies the analysis of resonance phenomena of the transmission line. The differential signals provide well controlled current return loop, thus improve the predictability of inductance value.

A practical concern is how to recover a square-wave clock signal to the registers all over the chip. One approach is that

we can use a sinusoidal standing wave to distribute the global clock signal to regions on a chip, and convert it into digital signal at each local region. This approach has been successfully applied to several high-speed clock distribution systems [5][7][14].

We adopt the clock buffer shown in [14] for the conversion. The buffer has two stages. The first stage is a differential amplifier and the second stage uses a cross-coupled inverters and a shunt resistor to convert the sine wave to square one. This buffer can achieve below 1ps amplitude-dependant skew.

B. Model of parameters variations

We utilize a simple linear variation model to represent the systematic spatial variations on wire widths and transistor lengths [12]. For any location (x, y) on the chip, the actual geometrical parameter $d = d_0 + k_x x + k_y y$, where d_0 is the nominal parameter and k_x, k_y are the horizontal, vertical variation coefficient, respectively. We set the maximum variations across the chip to be $\pm 10\%$ of the ideal value [3]. We choose this “pseudo-deterministic” linear variation model because it can be regarded as a “worst case” scenario of the probabilistic variations. However, we can replace this model with more sophisticated variation models in the proposed design framework.

We also take the supply voltage fluctuation into account when we analyze the clock skew. In this paper, we assume that the supply voltages are a set of independent random variables within $\pm 10\%$ of the nominal Vdd value.

C. Hybrid H-tree and shunt network

We tailor the natural frequency of the shunt wire to reduce the skew between clock terminals. Fig. 1 shows a hierarchical transmission line spirals network for global clock distribution. Each spiral consists of a pair of multiple wavelength long coplanar differential transmission line. Clock drivers are evenly distributed on every spiral and the separation between two neighboring drivers is one wavelength. An H-tree distributes the sinusoidal clock signals from the center of the chip to all the clock drivers. The signal arriving time of all the clock buffers on the same spiral is designed to be equal.

Fig. 1 Multi-level Transmission Line Spirals Network

The proposed symmetrical global distribution network distributes low skew clock signals to the lowest level spiral. For each local region, a local distribution tree or mesh is needed to send clock signals from the clock buffers on the lowest level spiral to innumerable clocking elements.

D. Problem Statement

For the same amount of routing area, assigning them to the spirals at different level may have different impact on the clock skew. In this paper, we are interested in the optimal way to distribute the routing resources to the spirals at different levels such that the minimum skew is achieved on the bottom level spiral with given routing area budget.

We formulate this problem as the following optimum hierarchical transmission line spirals sizing problem.

Optimum Hierarchical Transmission Line Spirals Sizing Problem:

Given: model of parameters variations.

Input: H-tree and n-level spiral network

Constraint: total routing area

Object Function: minimize the skew on the bottom level spiral

Output: The optimum wire width w_i of spirals at level i , for $i = 1$ to n , such that the clock skew is minimized

III. SKEW REDUCTION EFFECT OF TRANSMISSION LINE SHUNTS

A. Two sources case

Fig. 2 Two sinusoidal sources with phase shift shunted by a wavelength long transmission line

We use a simplified circuit model shown in Fig. 2 to study the skew reduction mechanism of a one wavelength long transmission line shunt. In Fig. 2, two clock drivers with driving resistance R_s and input phase shift (skew) Φ are connected by an RLGC transmission line of exactly one wave length long. The output at two separated terminals, V_1 and V_2 are synchronized by the transmission line. Fig. 3 shows the simulated wave forms. The input skew between input voltages V_{s1} and V_{s2} is 30 degree, the resulted skew between output voltages V_1 and V_2 is only 0.7 degree.

Fig. 3 Simulated waveform of two sources shunted by one wavelength long transmission line

Assume that input skew is small and $R < \omega L$, by superposition of all possible traveling and standing waves in the transmission line we obtain following skew expression.

$$\Delta \phi = \frac{1 - e^{-\frac{\pi R}{\omega L}}}{1 + e^{-\frac{\pi R}{\omega L}}} \phi \quad (1)$$

From the skew equation (1), we observe that when resistance R approaches zero, the transmission line becomes lossless. As a result, $\Delta \Phi$, the phase shift between voltages V_1 and V_2 , also approaches zero. Two clock drivers get fully synchronized. When R approaches infinity, nodes 1 and 2

are open. There is no shunt effect. The phase shift between nodes 1 and 2 remains the input phase shift, ϕ .

B. Multiple sources case

Fig. 4 Multiple sources shunted by a transmission line

We use a random model for multiple sources case. Fig. 4 shows the model of multiple sources shunted by a transmission line. We derive the skew expression under following assumptions: i) the transmission line is infinitely long and the clock buffers are placed evenly on the line with separation of one wavelength; ii) the input phase of each voltage source to be a random number uniformly distributed in $[0, \Phi]$. Because it is an infinitely long line, we can assume there are two nodes a, b having exact phase 0 and Φ , respectively. We compute the expected phase of these two points, and take the difference of the expectations as the skew.

We assume the driving resistance is much larger than the characteristic impedance of the transmission line and the input skew is small. Using similar technique in the derivation of equation (1), we obtain following skew equation.

$$\Delta \phi = \frac{1 - e^{-\frac{3\pi R}{\omega L}}}{1 + e^{-\frac{3\pi R}{\omega L}}} \phi \quad (2)$$

IV. OPTIMAL SIZING OF MULTI-LEVEL SPIRAL NETWORKS

We use a pair of coplanar copper transmission lines to construct the spiral shunt. We follow the technology of [14]. The two parallel differential wires have height 240nm, and the same width w . The separation between them is 2 μ m, and the wires are 3.5 μ m above a ground plane. Typical value of w ranges from 0.5 to 40 μ m.

We use fast field solver to get the frequency dependant resistance, R , and inductance, L . We use linear regression to get the relation between resistance/inductance ratio, R/L , and wire width, w . The $R/L \sim 1/w$ relation displays excellent linearity.

We rewrite the skew function of each level spiral as

$$\Delta \phi = \frac{1 - c_i e^{-\frac{k_i}{w_i}}}{1 + c_i e^{-\frac{k_i}{w_i}}} \phi \quad (4)$$

Where, w_i is the width of the i th level spiral and c_i , k_i are constants for level i spiral. We write the optimal spiral sizing problem as following mathematical programming:

$$\text{Min:} \quad \Delta \phi = (((\phi_1 \frac{1 - c_1 e^{-\frac{k_1}{w_1}}}{1 + c_1 e^{-\frac{k_1}{w_1}}}) + \phi_2) \frac{1 - c_2 e^{-\frac{k_2}{w_2}}}{1 + c_2 e^{-\frac{k_2}{w_2}}} + \phi_3) \dots + \phi_n) \frac{1 - c_n e^{-\frac{k_n}{w_n}}}{1 + c_n e^{-\frac{k_n}{w_n}}}$$

$$\text{s.t.} \quad \sum_{i=1}^n l_i w_i = A$$

In the programming (5), Φ_i is the skew of signal propagation from level $i-1$ to level i spiral. l_i and w_i are length and width of the spiral of level i . The object is to minimize skew under the maximum routing area constraint A .

We proved following lemma.

Lemma: $f(w) = \frac{1 - ce^{-k/w}}{1 + ce^{-k/w}}$ is a convex function on

$w \in [\frac{k}{2}, \infty)$, where, k is a positive constant.

The above lemma suggests that, when the wire of the transmission line is wide enough, the skew~wire-width relation is convex. In order to make the programming convex, we impose a set of minimal wire width constraints to each level spiral. In our experiments, the minimal wire widths of each level mesh are 0.6 μ m, 1.3 μ m, 1.3 μ m. With the minimal wire width constraints for each level spiral, we get following convex program.

$$\begin{aligned} \text{Min:} \quad & \Delta \phi = (((\phi_1 \frac{1 - c_1 e^{-\frac{k_1}{w_1}}}{1 + c_1 e^{-\frac{k_1}{w_1}}}) + \phi_2) \frac{1 - c_2 e^{-\frac{k_2}{w_2}}}{1 + c_2 e^{-\frac{k_2}{w_2}}} + \phi_3) \dots + \phi_n) \frac{1 - c_n e^{-\frac{k_n}{w_n}}}{1 + c_n e^{-\frac{k_n}{w_n}}} \\ \text{s.t.:} \quad & \sum_{i=1}^n l_i w_i = A \\ & w_i > m_i, \forall i \in \{1, 2, \dots, n\} \end{aligned}$$

Due to the convex property of the program (6), we have the following theorem.

Theorem: The local optimum of the programming (6) is the global optimum.

According to the above theorem, many numerical methods such as gradient descendant and line search methods can be adopted to solve this class of programming. In our experiments, we solve them using the optimization package of MATLAB.

V. EXPERIMENTAL RESULTS

In our experiments, we set the chip size to be 2cm by 2cm, and use a three level spiral to distribute clock signal. The clock frequency is 10.336GHz. And the wave length is exactly 1cm. Each of the spirals has 4, 9, and 17 clock drivers respectively. We synthesize a balanced H-tree[4] to distribute clock signal from the center of the chip to the clock drivers. The designed arriving time of all drivers on the same level spiral is equal. With given process variations model, we obtain the worst skew of the signal propagation from one level to the next level based on SPICE simulation. We use these skews as the values of Φ_i in the convex programming. We normalize the routing area to the area of bottom level spiral with 1 μ m wire width.

Table 1. Optimized wire width of each level spiral

Total Area	W1 (um)	W2 (um)	W3 (um)	Skew M (ps)	Skew S (ps)	Impr.(%)
0	0	0	0	23.15	23.15	0%
0.5	1.7	0	0	17.796	20.50	13%
1	1.9308	1.0501	0	12.838	14.764	13%
3	2.5751	1.3104	1.3294	8.6087	8.7309	15%
5	2.9043	3.7559	2.3295	6.2015	6.3169	16%
10	3.1919	4.5029	6.8651	4.2755	5.2131	18%
15	3.6722	6.1303	10.891	2.4917	3.5182	29%
20	4.0704	7.5001	15.072	1.7070	2.6501	37%
25	4.4040	8.6979	19.359	1.2804	2.1243	40%

Table 1 lists the optimized wire width of each level spiral for different total routing area. W1, W2, and W3 are optimal wire widths of level 1, level 2 and level 3 spirals, respectively. For the comparison reason, we also simulate the skew on a single-level spiral network, which only uses bottom level spiral to shunt all the leaf nodes of the H-tree. We let the single level spiral network has same total routing area of the multi-level spirals network. Column 5 and 6 are the skews of multi-level spirals and single level spiral. Column 7 shows the skew improvement of multi-level

spirals over single level spiral. When total routing area is small, the optimal configurations prefer to allocate routing resources to the higher level mesh. With gradually increasing of the routing area, more resources are allocated to the bottom level mesh. Comparing with the single-level spiral, optimized multi-level spiral can reduce the skew by 40%.

We also compare the power consumption of optimized multilevel spiral network and that of single level spiral. In Table 2, the first row are the total routing areas of the multi-level spirals; the second row and the third row list the power consumption of the multilevel spiral and single level spiral with given amount of total routing area. The simulated results show that multilevel spiral can reduce the power consumption by 81%.

Table 2 Power Consumption Comparisons

Area	3	4	5	7	10	15	20	25
PM(mw)	0.4	0.5	0.7	0.9	1.0	1.4	1.5	1.6
PS(mw)	0.83	1.5	2.1	2.64	3.04	4.7	7.2	8.3
reduce(%)	48	67	67	66	67	70	79	81

We test the robustness of our optimized spirals network against supply voltage fluctuations. We perturb the supply voltage of every clock drivers independently by a random number within 10% of its nominal value. We perform 5 experiments on each network, and record the worst case skew and average case skew in the Table 3. We compare the skew of optimized multilevel spiral and single level spiral networks. The last column of Table 3 lists the improvement of the average case skew. Multilevel spiral network improves the skew by up to 55%.

Table 3 Skew in the presence of voltage variations

Area	Skew-S		Skew-M		
	Ave.	Worst	Ave.	Worst	Impr(%)
0	28.4	36.5	28.4	36.5	0%
3	9.75	12.33	8.75	9.07	11%
5	7.32	9.06	6.55	6.91	12%
10	6.31	8.05	4.41	5.41	30%
15	5.03	7.33	2.81	4.93	44%
25	3.83	4.61	1.72	3.06	55%

When the clock frequency deviates from its nominal value or the electrical length of transmission lines varies from integral multiple of the wavelength, the resonance phenomena of the transmission line shunts diminishes. As a result, the synchronization capabilities of transmission line shunts degrade accordingly. We explore the frequency response properties of the proposed multilevel clock network. We set the wire width of the lowest level transmission line to be 5 μ m wide and the clock rate to be 10.33GHz. The -3db bandwidth of the output voltages is 0.42GHz. At 10.33GHz, the minimal skew of 1.38 degrees is achieved. In the frequency range of 10.2GHz to 10.5GHz, the skew lies between 2.5 degrees and 1.38 degrees.

VI. CONCLUSIONS

When the clock frequency thrust into multi-giga hertz regime, transmission line shunts demonstrate the unique potential of achieving low skew low jitter global clock distribution under parameter variations. We propose a new architecture of global clock distribution that incorporates multiple level transmission line spirals.

We derive the analytical expression of the skew reduction

contributed by multiple wavelength long transmission line shunts. The skew formulas display very good fidelity to the spice simulation. Based on the theoretical analysis of the transmission line shunt behaviors, we use a convex programming technique to optimize the wire width of each level spiral in the multiple level spirals clock network. The experimental results show that: 1. Comparing with the traditional RC buffer tree, the optimized multilevel transmission line spiral network can reduce the clock skew by 84%; 2. With 10% random supply voltage fluctuation, our proposed multilevel transmission line spiral network provides a less than 2ps global clock distribution at 10GHz; 3. the power dissipation overhead of the transmission line shunts is only 1.6mW.

VII. ACKNOWLEDGEMENT

The authors would like to thank Prof. Peter Asbeck for helpful discussions and anonymous reviewers for their constructive suggestions. They also would like to thank the support from NSF grant MIP-9987678, California MICRO program, and Cal(IT)2.

REFERENCES

- [1] F.E. Anderson, et al, The Core Clock System on the Next Generation Itanium™ Microprocessor, ISSCC 2002 Session 8.5.
- [2] N. Bindal, T. Kelly, N. Velastegul, and K. L. Wong, "Scalable sub-10ps Skew Global Clock Distribution for a 90nm Multi-GHz IA Microprocessor," ISSCC 2003, Session 19.7
- [3] D. Boning and S. Nassif, Models of Process Variations in Device and Interconnect, in Design of High Performance Microprocessor Circuits, Editors: A. Chandrakasan, W. Bowhill, F. Fox, IEEE Press, 2000
- [4] C.K. Cheng, et al, Interconnect Analysis and Synthesis, 2000, Wiley Interscience.
- [5] V.L. Chi, "Salphasic distribution of clock signals for synchronous systems," IEEE Trans. Comput., vol. 43, pp.597-602, May 1994
- [6] M.P. Desai, R. Cvijetic, J. Jensen, "Sizing of clock distribution networks for high performance CPU chips," DAC 1996 pp..389-394
- [7] I. Galton, D.A. Towne, J.J. Rosenberg, and H.T. Jensen, "Clock Distribution Using Coupled Oscillators," ISCAS pp.
- [8] V. Gutnik and A.P. Chandrakasan, "Active GHz clock network using distributed PLLs," IEEE J. Solid-State Circuits, vol. 35, no. 11, pp. 1553-1560, Nov. 2001
- [9] L. Hall, M. Clements, W. Liu, and G. Bilbro, "Clock distribution using cooperative ring oscillators," in Proc. 17th Conf. Advanced Research in VLSI, Sept. 1997, pp. 15-16
- [10] M. Kamon, M. J. Tsuk, and J. K. White, FastHenry: A multipole-accelerated 3-d inductance extraction program. IEEE Trans. on Microwave Theory and Techniques, 42(9):1750-8, September 1994.
- [11] N.A. Kurd, et al, A Multigigahertz Clocking Scheme for the Pentium® 4 Microprocessor, IEEE Journal of Solid-State Circuits, Vol. 36, No. 11, Nov. 2001 pp. 1647-53.
- [12] V. Mehrotra, Modeling the Effects of Systematic Process Variation on Circuit Performance, Ph.D. Thesis, MIT, May, 2001
- [13] M. Mori *et al.*, "A Multiple Level Network Approach for Clock Skew Minimization with Process Variations," UCSD Technical Report, CS2003-0756
- [14] F. O'Mahony *et al.*, "Design of a 10GHz Clock Distribution Network Using Coupled Standing-Wave Oscillators," DAC 2003, pp. 682-687, June 2003
- [15] M. Orshansky, L. Milor, P. Chen, K. Keutzer and C. Hu, Impact of Spatial Intrachip Gate Length Variability on the Performance of High-Speed Digital Circuit, IEEE trans. on CAD, p.544-553, vol. 21, No. 5, May 2002
- [16] P.J. Restle, et al, A Clock Distribution Network for Microprocessors, IEEE Journal of Solid-State Circuits, Vol. 36, No. 5, May 2001 pp. 792-99.
- [17] P.J. Restle, et al, The Clock Distribution of the Power4 Microprocessor, ISSCC 2002, Session 8.4.

- [18] M. Saint-Laurent, M. Swaminathan and J.D. Meindl, "On the micro-architectural impact of clock distribution using multiple PLLs," ICCD 2001, pp.214-220