Routing Methodology for Minimizing Interconnect Energy Dissipation

Atsushi Sakai, Takashi Yamada, Yoshifumi Matsushita

Materials and Devices Development Center,
SANYO Electric Co., Ltd.
180 Ohmori, Anpachi-cho, Anpachi-gun,
Gifu 503-0195, Japan
Tel: +81-584-64-5218

{sakai, yamada, matusita}@ul.rd.sanyo.co.jp

Hiroto Yasuura

Department of Computer Science and
Communication Engineering, Kyusyu University
6-1Kasuga-koen, Kasuga,
Fukuoka 816-8580, Japan
Tel: +81-92-583-7620

yasuura@c.csce.kyushu-u.ac.jp

ABSTRACT

In this paper, we propose a new physical design technique for sub-quarter micron system-on-a-chip (SoC). By optimizing the routing grid configuration for the automatic place and route methodology, coupling effects such as crosstalk noise and coupled energy dissipation are almost eliminated with only a small cost to the runtime. Experiments are performed on the design of an image processing circuit using a sub-quarter micron CMOS process with multi-layered interconnects. Simply by employing our proposed technique, net switching energy dissipation can be reduced by about 10% maximum without any area penalty. This significant energy reduction greatly accelerates the performance of SoCs.

Categories and Subject Descriptors

B.7.2 [**Integrated Circuits**]: Design Aids – *Placement and routing.*

General Terms

Design.

Keywords

Analysis, crosstalk, energy dissipation, routing, SoC.

1. INTRODUCTION

In the case of SoCs fabricated in sub-quarter micron or nanometer-scale technologies, on-chip interconnection is an important performance limiting factor [2]. With technology scaling, the transistor size, the line width and the line-to-line interval have all been scaled. However, the line height has not yet been scaled, in order to avoid any increase in the wiring resistance that scaling would cause. Therefore, the proportion of gate and vertical line-to-line capacitance with respect to the overall capacitance decreases, while that of the horizontal line-to-line capacitance increases. This large horizontal line-to-line capacitance in adjacent lines, which is the main source of cross-coupling capacitance between signal wires, results in unwanted coupling effects such as an increase in crosstalk noise [4][15] and coupled energy dissipation [7][16].

Crosstalk noise is a voltage change induced by the cross-coupled capacitance in adjacent lines. It causes logical malfunctions because of the signal distortion due to noise [8].

Coupled energy dissipation is caused by the charging and discharging of the coupling capacitance between signal wires. The energy varies with the switching activity of the signals in adjacent lines [16]. Because it is practically difficult to identify the state of switching activity occurring between adjacent lines in a large-scale SoC, energy constraints should be satisfied in the worst-case switching conditions.

To reduce these unwanted coupling effects, various circuitry and routing design techniques have been proposed. When considering circuitry techniques, crosstalk noise is reduced by inserting repeaters, whereby the charge retaining capability of the victim nets is improved [1][9]. In addition, in the case reported in [5], the variations in effective line-to-line capacitance caused by switching activity are mitigated by inserting repeaters (inverters) between adjacent lines at staggered points. Changes made to the circuitry would have a direct effect on energy, and might result in an increase in the chip area. Insertion of repeaters at staggered points is effective in a handcrafted or semi-custom design, however, it is difficult to locate the most suitable insertion points if using the automatic place and route methodology that is common in today's large-scale SoC design. Furthermore, the increase in the number of repeaters is directly linked to an increase in energy dissipation.

Routing techniques have also been proposed as a solution to reducing coupling effects. One technique is to change the line width and spacing without changing the line pitch given by the design rules [3][5]. By using this technique, the line capacitance or the line resistance changes considerably when the line widths are greatly widened and as a result the energy dissipation is unfavorably affected. Another technique is to increase the distance to adjacent lines for chosen nets [11]. By using this method, the degree of congestion in un-chosen nets increases and as a result the overall crosstalk noise occurring on a chip cannot be decreased. Another different technique is to establish a power line and other lines that have no electric-potential

GLSVLSI'03, April 28-29, 2003, Washington, DC, USA.

variations on either side of a noise-sensitive line for shielding [5][6][10]. By using this method, many wiring resources are required for shielding and the design complexity is also increased. Moreover, additional coupling capacitance still remains between the signal and shielding lines.

To address these problems, we propose a new physical design technique for large-scale SoCs by optimizing the routing grid configuration for the automatic place and route methodology. The correlation between grid settings and timing has already been investigated and a maximum of 15% improvement in delay has been achieved [13]. In this work, we demonstrate that this technique is also very effective for reducing interconnect energy dissipation by employing post-layout simulation.

2. Techniques to Reduce Coupling Effects

In the automatic place and route methodology for large-scale SoC layouts, the grid-based model and the reserved layer model are generally used [14]. Since many routing resources are needed to avoid short circuits between wires, the routing grid pitch is calculated according to design rules and the smallest value is generally adopted.

Line-to-line coupling capacitance between signal wires is the main cause of coupling effects such as crosstalk noise and coupling power consumption. This capacitance can be separated into two components, which are the horizontal (i.e. intra-layer) line-to-line capacitance (C_H) and the vertical (i.e. inter-layer) line-to-line capacitance (C_V). In deep sub-quarter micron SoCs, C_H is between 2.4-5.6 times larger than C_V [17]. Therefore, reducing C_H by increasing the horizontal line-to-line spacing is considered to be the most effective approach to reduce coupling effects.

The interconnect energy dissipation (E_{wire}) is expressed as follows if the transistor's load capacitance is negligible [15].

$$E_{wire} = \alpha(C_H+C_V)V_{DD}^2 f \quad (1)$$

Here, α is the switching activity ratio, and f is the operating frequency. With scaling, the interconnect portion of the SoC's total energy dissipation will increase [15]. Furthermore, since the major part of C_H is the coupling capacitance between the signal wires and because C_H/C_V increases with scaling, the coupled energy dissipation becomes the major constituent of the total energy dissipation in sub-quarter micron SoCs. From equation (1), it is clear that decreasing C_H directly leads to a reduction of E_{wire}.

To increase the horizontal line-to-line spacing, widening the routing grid pitch on the same layer seems to be the most attractive approach. However, if this approach is employed in isolation, a shortage of routing resources occurs, resulting in an extraordinary increase in runtime for the detailed routing.

The objective of this study is to address the trade-off between the coupling effects reduction and the shortage of routing resources. We propose a new optimization technique that can be used to tune the routing grid configuration, both in the horizontal and vertical directions. We elaborate on the details of this technique in the subsequent sections.

2.1 Optimizing the configuration of the routing grid

An effective way to decrease the coupling effects, and maintain the routing resources is to focus on the layers that have many parallel segments of wire. For this purpose, we propose a new routing technique for the automatic place and route methodology. This technique has two features as follows:

(I) Increasing an individual layer's routing grid pitch in an incremental manner to decrease C_H, thereby reducing the coupling effects.

(II) In addition to (I), choosing appropriate layers for widening the routing grid pitch by considering the trade-off between the reduction of coupling effects and the routing resources.

These concepts are illustrated in Figure 1. In this four metal layer example, we can control the widening of the routing grid pitch of Ma, Mb, Mc and Md independently.

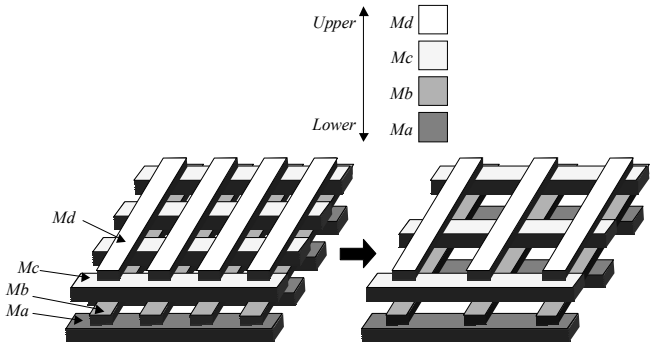

Figure 1. Basic concept of optimization for the routing grid configuration

We re-specify the minimum routing pitch for various layers from their original values, then re-route, re-extract, and perform noise, power, and runtime analysis. These results are compared to the original design to find the optimal wiring pitches by evaluating the cost function (COST).

$$COST = CN^{\alpha}EN^{\beta}RT^{\gamma} \quad (2)$$

Here, CN, EN and RT are the number of crosstalk-violated net, the interconnect energy dissipation and the runtime at the routing stage, respectively. And α, β and γ are weighting factors, which are heavily weighted if the parameter is important, but is otherwise zero if the parameter is negligible. The correlation between these parameters can be checked for various routing grid settings to find the optimal one that gives the minimum COST value. Some "interconnect rich" designs are possibly difficult to route when the routing grid pitch is a minimum. In this case, we should change the placement to ease the congestion by re-tuning the cell utilization ratio.

To save on the overall design period, our technique should be applied at the global routing stage instead of at the detailed routing stage for all settings. In this case, RT can be substituted for the number of violations (or other parameters that can express routing congestion).

2.2 Energy Calculation of Interconnect

After the routing has been completed, a gate-level energy calculation is performed to estimate the energy dissipation of the targeted design. To increase the accuracy of the analysis, actual layout information is used. At first, the parasitic components of the wires including the parasitic coupling capacitance are extracted by using a 3-D field solver. Furthermore, the transition times of all input signals are extracted.

The energy calculation is performed with the crosstalk noise taken into consideration. We change the coupling capacitance into the effective capacitance. The effective switching factor is fixed at 1 for all conditions. Coupling from all adjacent lines is summed to estimate the effective capacitance of each net.

Cell libraries are characterized by highly accurate SPICE simulation, taking the operating mode into consideration. For example, RAMs are characterized separately in read and write mode. The overall energy dissipation is the summation of three components, (i) the dynamic cell internal energy, (ii) the net switching energy, and (iii) the cell leakage energy [12]. When estimating component (i) or component (ii) by using a static-based energy calculation, the switching activity ratio must be considered. Since this ratio is difficult to estimate, we use 0.2 for all nets. This figure has been frequently used with reasonable accuracy [15].

In our proposed technique, component (ii) is the most significant parameter for dramatically varying the topology of the wire. Therefore, component (ii) is considered during the routing grid optimization process.

3. Experimental Results

The methods proposed in Section 2 are used to layout an image processing circuit of the 100,000-instance scale, which is part of a SoC fabricated in both 0.13µm CMOS with six copper layers and 0.18µm CMOS with five aluminum layers. Experiments were performed using these circuits to see if the net switching energy could be reduced without sacrificing too much runtime. All EDA tools employed in this work were commercial ones widely used for large-scale SoC layout.

3.1 Results of the 0.13µm Design

The routing grid pitch was increased in five steps from a value specified according to the minimum design rule to a value twice as large. For this design, we defined two cases of layer selection for widening the routing grid pitch. For the first of these we chose the third to sixth layers (case 1) and for the other we chose the fourth to sixth layers (case 2). Since the pin pitch inside cells cannot be modified, the pitch of the first and the second layers was kept to the minimum value so as not to heavily degrade the routing efficiency.

In case 1, as is shown in Figure 2, the number of crosstalk-violated nets (assuming the peak exceeded 30% of V_{DD}) decreased to about one-eighth if we made the routing grid pitch twice as large as the minimum value. In case 2, the number of crosstalk-violated nets decreased to about one-third.

The improvements in energy dissipation achieved in cases 1 and 2 are shown in Figure 3. By increasing the routing grid pitch, the energy also reduced until the pitch reached 1.5. However, widening beyond 1.5, the energy remained at the same value due to the shortage of routing resources. As a result, the maximum reduction in energy was 9% in case 1 and 4% in case 2.

The runtimes of cases 1 and 2, when the routing grid pitch was a maximum, were increased by a factor of eight and two, respectively, compared to when the pitch was at a minimum. Due to the shortage of routing resources, it was impossible to complete the detailed routing when the pitch was widened from the second to sixth layers.

To find the optimum settings for the configuration of the routing grid, the cost function was employed. The cost of this design was evaluated based on equation (2). Here, to evaluate the trade-off between energy improvement and runtime, we set $\alpha=0$, $\beta=4$, and $\gamma=0.2$. The results of the cost calculation are displayed in Figure 4. As is apparent from this figure, a pitch of 1.5 for case 1 was the best solution for achieving the maximum reduction in energy dissipation (9%) with little runtime penalty (less than a fourfold increase) for this design.

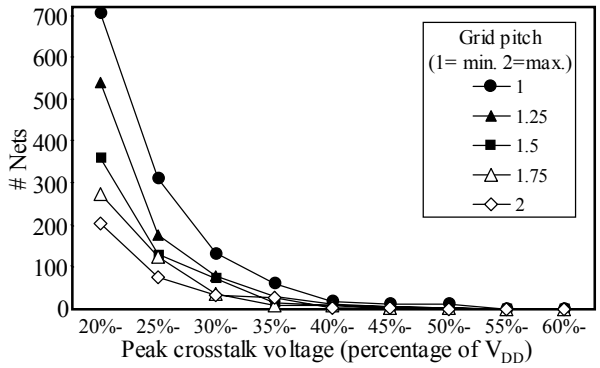

Figure 2. Peak crosstalk distribution vs. routing grid pitch in case 1 for the 0.13µm SoC

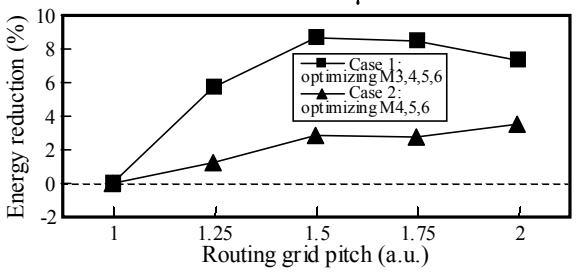

Figure 3. Energy reduction in cases 1 & 2 for the 0.13µm SoC

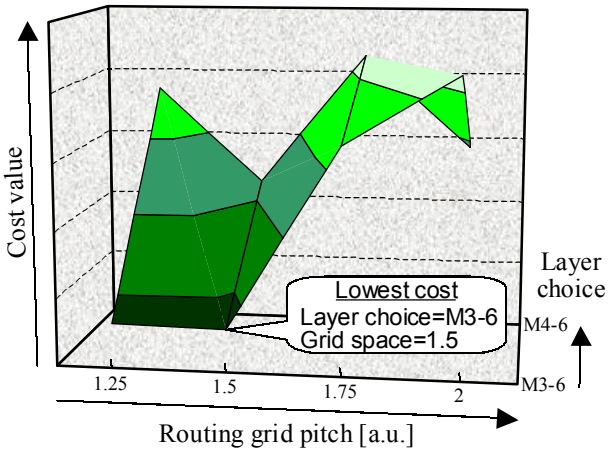

Figure 4. Cost evaluation for the 0.13µm SoC

3.2 Results of the 0.18μm Design

In a similar way to the 0.13μm design, the routing grid configuration was optimized to minimize crosstalk noise and net switching energy dissipation. By widening the routing grid pitch of the third to fifth layers, the number of crosstalk-violated nets was decreased by about one-ninth if we made the routing grid pitch twice as large as the minimum value. In addition, the 8% energy reduction was achieved at a pitch of 1.5 with little runtime penalty (less than 50% increase).

3.3 Summary of Energy Reduction

The energy reductions that we achieved in terms of layer choice for widening when employing the proposed technique are summarized in Figure 5. The technique can achieve dramatic reductions in energy dissipation once the optimum routing grid settings are obtained.

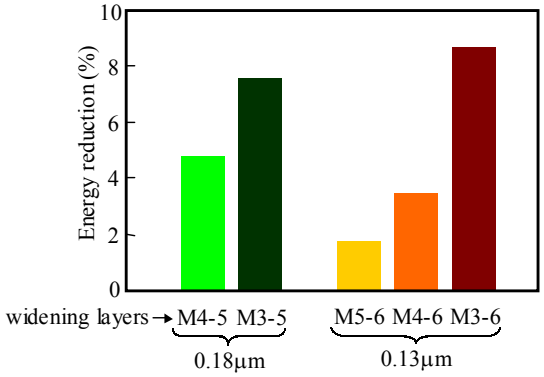

Figure 5. Summary of energy reduction for the 0.13μm/0.18μm SoC

4. Conclusions

In this study, we have demonstrated that coupling effects such as crosstalk noise and coupling energy dissipation can be reduced by simply optimizing the configuration of the routing grid without process improvements or any area penalty. Experimental results showed that our proposed technique proved to be effective in reducing coupling effects without an unreasonable increase in runtime. From the results of the energy calculation, the energy was reduced by about 10% maximum.

5. REFERENCES

[1] Alpert, C.J., Devgan, A., and Quay, S.T. Buffer Insertion for Noise and Delay Optimization. *Proceedings of DAC*, pp.362-367, 1998.

[2] Cong, J., He, L., Khoo, K.-Y., Koh, C.-K., and Pan, Z. Interconnect Design for Deep Submicron ICs. *Proceedings of ICCAD*, pp.478-485, 1997.

[3] Cong, J., He, L., Koh, C.-K., and Pan, Z. Global Interconnect Sizing and Spacing with Consideration of Coupling Capacitance. *Proceedings of ICCAD*, pp.628-633, 1997.

[4] Gal, L. On-Chip Cross Talk – the New Signal Integrity Challenge. *Proceedings of CICC*, pp.251-254, 1995.

[5] Kahng, A.B., Muddu, S., Sarto, E., and Sharma, R. Interconnect Tuning Strategies for High-Performance ICs. *Proceedings of DATE*, pp.471-478, 1998.

[6] Khatri, S.P., Mehrotra, A., Brayton, R.K., and Sangiovanni-Vincentelli, A. A Novel VLSI Layout Fablic for Deep Sub-Micron Applications. *Proceedings of DAC*, pp.491-496, 1999.

[7] Kim, K.-W., Baek, K.-H., Shanbhag, N., Liu, C.L., and Kang, S.-M. Coupling-Driven Signal Encoding Scheme for Low-Power Interface Design. *Proceedings of ICCAD*, pp.318-321, 2000.

[8] Larsson, P. and Svensson, C. Noise in Digital Dynamic CMOS Circuits. *IEEE Journal of Solid-State Circuits*, vol. 29, pp.655-662, 1994.

[9] Li, D., Pua, A., Srivastava, P., and Ko, U. A Repeater Optimization Methodology for Deep Sub-Micron, High-Performance Processors. *Proceedings of ICCD*, pp.726-731, 1997.

[10] Ma, James D.Z., and He, L. Formulae and Applications of Interconnect Estimation Considering Shield Insertion and Net Ordering. *Proceedings of ICCAD*, pp.327-332, 2001.

[11] Nicoletta, C., Alvarez, J., Barkin, E., Chao, C.-C., Johnson, B.R., Lassandro, F.M., Patel, P., Reid, D., Sánchez, H., Siegel, J., Snyder, M., Sullivan, S., Taylor, S.A., and Vo, M. A 450-MHz RISC Microprocessor with Enhanced Instruction Set and Copper Interconnect. *IEEE Journal of Solid-State Circuits*, pp.1478-1491, 1999.

[12] Pedram, M. Power Minimization in IC Design: Principles and Applications. *ACM Trans. Design Automation Electron. Syst.*, vol.1, no.1, pp.3-56, 1996.

[13] Sakai, A., Yamada, T., Matsushita, Y., and Yasuura, H. Reduction of Crosstalk Noise by Optimizing 3-D Configuration of the Routing Grid. *Proceedings of ASP-DAC*, pp.49-52, 2003.

[14] Sherwani, N. Algorithm for VLSI Physical Design Automation Second Edition. *Kluwer Academic Publishers*, 1995.

[15] Sylvester, D., and Keutzer, K. Getting to the Bottom of Deep Submicron. *Proceedings of ICCAD*, pp.203-211, 1998.

[16] Taylor, C.N., Dey, S., and Zhao, Y. Modeling and Minimization of Interconnect Energy Dissipation in Nanometer Technologies. *Proceedings of DAC*, pp.754-757, 2001.

[17] Yim, J.-S., and Kyung, C.-M. Reducing Cross-Coupling among Interconnect Wires in Deep-Submicron Datapath Design. *Proceedings of DAC*, pp.485-490, 1999.