Abstract

To achieve minimum signal propagation delay, the non-uniform wire width routing architecture has been widely used in modern VLSI design. The non-uniform routing architecture exploits the wire width flexibilities to trade area for performance. However, many additional design rules, which confine the routing flexibilities, are introduced in nanoscale circuit designs. With the increasing difficulties of fabricating nanoscale circuits, the conventional non-uniform routing architecture becomes clumsy. We propose an uniform dual-rail routing architecture to cope with these new challenges. The proposed architecture exploits the anti-Miller effect between two adjacent wires with the same signal source. Hence, the coupling capacitance between these two wires is reduced. The simulation results demonstrate that our proposed architecture provides a signal propagation channel with similar propagation delay, less crosstalk noise, and less power consumption to the conventional non-uniform routing architecture with moderate routing area overheads. In terms of the properties and the scalabilities, we argue that the uniform dual-rail routing architecture is a wire sizing alternative without incurring layout irregularity and stacked vias overheads.

I. INTRODUCTION

With the advances of VLSI design technology, feature sizes of both logic gates and routing wires keep shrinking. The shrinking feature size provides larger chip capacity as well as higher chip complexity for circuit designers to work out a more powerful design than ever. However, down-size scaling brings enormous fabricating difficulties below 130 nm technology node [1]. One of these major difficulties is the photo lithography process. The optical interference and diffraction happen due to the sub-wavelength mask. These optical phenomena distort the patterns and at worst destroy the patterns. Several emerging techniques, such as the optical proximity correction (OPC), the phase shift mask (PSM), and off-axis illumination (OAI), are proposed to cope with these challenges [2]. However, these techniques also brings new constraints for circuit design. Hence, there are many additional design rules introduced to ensure the circuit manufacturability in deep submicron (DSM) era [1]. To satisfy those complex design rules and hence to improve circuit manufacturability, simple and regular layout favors chip implementation.

Besides the growing design for manufacturability (DFM) issues, the uneven shrinking factors among the wire thickness, the wire width, and the gate length will result in serious performance degradation. The shrinking factor of the wire thickness is less than those of both the wire width and the gate length [3]. With this shrinking trend, the aspect ratio of metal wire as well as the coupling capacitance become higher. As a result, the coupling capacitance between adjacent wires contributes significant part of interconnect capacitance. The undesired coupling capacitance incurs circuit delay, crosstalk noise, power consumption, and timing uncertainty [4].

To achieve the optimal signal propagation delay, the non-uniform wire width routing architecture is proposed under the Elmore delay model [5]. This intuitive architecture, which is similar to the urban water supply system, uses thick wire at the source and thin wire at the sink to achieve the optimal signal propagation delay [6]. Although the non-uniform wire width routing architecture gives optimal timing result, using the off-grid routing frequently will result in higher routing complexity and larger layout database. At worst, the layout irregularity incurred by non-uniform routing will cause more routing resource wasted due to the introducing of forbidden routing region design rules [7].

For the purposes of both the best signal propagation delay and the DFM issues consideration, we propose an uniform dual-rail routing architecture with high layout regularity. By exploiting negligible coupling capacitance between two adjacent wires with the same signal source, our routing architecture provides a signal propagation channel with similar propagation delay, less crosstalk noise, and less power consumption to the conventional non-uniform routing architecture. Due to its high layout regularity, the uniform routing architecture can be easily adopted by existing routing frameworks. To the best of our knowledge, there is no previous work which exploits the anti-Miller effect between adjacent wires to trade moderate routing area for both circuit performance and layout regularity in DSM era.

The rest of this paper is organized as follows. The preliminary background knowledge is given in Section II. Section III presents our observations and the proposed routing architecture. The experimental results are drawn in Section IV. Section V concludes this work.

II. PRELIMINARY

In this section, we discuss important background knowledge, the non-uniform routing architecture for optimal signal propagation delay and the capacitive crosstalk effect between adjacent wires.

A. Non-uniform Routing Architecture for Optimal Delay

Signal propagation delay dominates circuit performance in recent years. To improve circuit performance, the top priority is to optimize interconnects between logic gates. Signal propagation delay is due to the wire resistance and
the parasitic capacitance on a wire. The parasitic $RC$ of a wire segment can be approximated by a $\pi$-model. We can use the Elmore delay model to compute the $RC$ delay of a wire segment. Several important researches focus on minimizing the signal propagation delay for a given routing topology. Chen and Wong proposed an optimal wire sizing formula for non-uniform routing [5].

In chip manufacturing process, it is impossible to optimally fabricate a wire with continuous decreasing wire width. The only feasible solution is to map the optimal wire width function into a few wire segments with discrete wire widths. After the wire width of each wire segment is determined, we perform the routing layer assignment. The thick wire can be arranged either in the upper layer or in the same layer of the thin wire. However, either solution has some undesired drawback. Putting the thick wire in the upper layer of the thin wire incurs extra stacked vias especially for the interleaved horizontal/vertical layers. Putting the thick wire and the thin wire in the same layer results in layout irregularity and hence incurs the DFM issues. Therefore, the non-uniform routing architecture is not capable to preserve a regular layout in each metal layer without incurring stacked vias overheads.

B. Capacitive Crosstalk Effect

There are three major wire capacitance categories, the coupling capacitance, the area capacitance, and the fringe capacitance. The coupling capacitance lies in between two adjacent wires; while both the area capacitance and the fringe capacitance lie in between a wire and its upper and lower adjacent wires. For simplicity, we use ground capacitance to represent both the area capacitance and the fringe capacitance. The capacitance is derived as follows [8]:

$$\text{Capacitance} \propto \frac{\text{coupling length}}{\text{separation distance}^\alpha},$$

where $1 < \alpha < 2$.

The crosstalk effect varies according to the signal transition directions. For any two adjacent wires, the propagation delay is reduced when these signals transit to the same direction; otherwise the propagation delay is increased. To model the impact of capacitive crosstalk effect, the concept of effective wire capacitance is widely used

$$C_{eff} = C_{\text{ground}} + \beta \cdot C_{\text{coupling}}$$

where $\beta$ is the switching factor. Kahng, Muddu, and Sarto analyzed the switching factor in modern VLSI routing [9]. The switching factor is around 0 (signals transit to the same direction) to 3 (signals transit to different directions).

III. UNIFORM ROUTING ARCHITECTURE

In this section, we argue the observations from the aforementioned issues for high performance routing and hence propose the uniform dual-rail routing architecture.

A. Observations on High Performance Routing

Due to the manufacturing difficulties incurred by photolithography process, one wire orientation is allowed in a metal layer. A metal layer with pure horizontal/vertical routing wires will be enforced in advanced technology [10]. With the number of design rules explodes, a simple and regular layout is required to meet these complex design rules. On the other hand, the wire width is not a constant from source to sink in order to obtain the optimal signal propagation delay. Therefore, we need to incur extra stacked vias to connect wire segments with different widths in different interleaved horizontal/vertical layers for better signal propagation delay.

Since the signal propagation delay is due to the parasitic resistance and capacitance of a wire, reducing the $RC$ is the key to optimize circuit performance. The most effective way to reduce the wire resistance is to use a thick wire on the upper metal layers. However, thick wires also increase the burden of the ground capacitance and the coupling capacitance [3]. Therefore, using the thick wire in the upper layer may not be the best choice for performance improvement.

B. The Uniform Dual-rail Routing Architecture

We propose an uniform dual-rail routing architecture, which merges two adjacent wires into one signal propagation channel. Due to the anti-Miller effect, the coupling capacitance between these two merged wires can be greatly reduced. Hence, the wire resistance of this new routing channel is reduced by two without increasing the coupling capacitance. By using this dual-rail routing architecture, circuit layout becomes regular and hence increases the manufacturability. The schematics of both the conventional and our routing architectures are drawn in Figure 1. In Figure 1(a), the conventional non-uniform wire width routing architecture propagates signals among several metal layers with extra stacked vias overheads. In Figure 1(b), the uniform dual-rail routing architecture can be fabricated in fewer metal layers without incurring stacked vias.

The $\text{SPICE}$ netlists of the conventional and our routing architectures are drawn in Figure 2(a) and Figure 2(b), respectively. In Figure 2, $R$, $C_c$, $C_g$, $L$, and $M$, represent the resistance, coupling capacitance, ground capacitance, self inductance, and mutual inductance of the wire, respectively. The suffixes, $w$ and $n$, indicate the thick (wide) wire and the thin (narrow) wire, respectively. The middle wires in Figure 2 represent netlists of different routing architectures; while the upper and lower wires are common wires. For simplicity, we use $\text{victim}$ and $\text{aggressor}$ to indicate the middle wire and the upper/lower wire, respectively.

IV. SIMULATION RESULTS

In this section, we conduct a series of $\text{SPICE}$ simulations for more understanding about the characteristics of dual-rail and non-uniform routing architectures in terms of the properties and the scalabilities.

Our simulations target on 65 $\text{nm}$ process technology. The wire segment length is $25 \mu m$ for each distributed $\pi$-model. The $\text{source}$ is a 16X-inverter and the $\text{sink}$ is a 2X-inverter. The
technology parameters are adopted from public documents, the predictive technology model (PTM) [11] and the international technology roadmap for semiconductors 2006 (ITRS) [3]. Table I lists the parameters used in our simulations. Note that the dielectric constant \( k \) is suggested to be 2.2 in PTM, we use the updated value (2.8) from ITRS in our simulations. However, the simulation results of \( k = 2.2 \) comply with those of \( k = 2.8 \). Besides, the wire width of the non-uniform routing architecture is set to 0.14 according to the minimum design rules of the upper metal layers with greater wire thickness. Notice that the cross-section of thick wire is \( 0.049 \times 0.049 = 1.225 \) times larger than that of the dual-rail wire, which means the wire resistance of dual-rail is 22.5% greater than that of the thick wire in our simulations.

### Table I

**Technology Parameters**

<table>
<thead>
<tr>
<th></th>
<th>Thin wire</th>
<th>Thick wire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wire width (( \mu m ))</td>
<td>0.1</td>
<td>0.14</td>
</tr>
<tr>
<td>Wire spacing (( \mu m ))</td>
<td>0.1</td>
<td>0.14</td>
</tr>
<tr>
<td>Wire thickness (( \mu m ))</td>
<td>0.2</td>
<td>0.35</td>
</tr>
<tr>
<td>Wire cross-section (( \mu m^2 ))</td>
<td>0.02</td>
<td>0.049</td>
</tr>
<tr>
<td>Dielectric constant (( k ))</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Wire resistance (( \Omega ))</td>
<td>27.499</td>
<td>11.224</td>
</tr>
<tr>
<td>Self inductance (( nH ))</td>
<td>0.028</td>
<td>0.025</td>
</tr>
<tr>
<td>Mutual inductance (( nH ))</td>
<td>0.022</td>
<td>0.02</td>
</tr>
<tr>
<td>Ground capacitance (( fF ))</td>
<td>0.617/875</td>
<td>0.8552525</td>
</tr>
<tr>
<td>Coupling capacitance (( fF ))</td>
<td>1.7081</td>
<td>1.834275</td>
</tr>
</tbody>
</table>

A. Property Simulations

At first, we focus on examining the basic properties of different routing architectures in terms of timing, noise, and power. The total wire length is set to 1000\( \mu m \). We incrementally insert one-tens total wire length (i.e. 100\( \mu m \)) of the dual-rail/thick wire from source and then record the propagation delay, the crosstalk noise, and the power consumption. The simulation results are drawn from Figure 3 to Figure 5, respectively.

Figure 3 shows the worst case signal propagation delay. The worst case condition occurs when the victim signal transits to one direction and both the aggressor signals transit to the other direction. The results of crosstalk noise is drawn in Figure 4. The crosstalk noise is the maximum voltage fluctuation of a stable victim when both the aggressor signals transit to the same direction. There are two types of crosstalk noise, \( 0 \rightarrow 1 \) and \( 1 \rightarrow 0 \). The solid lines indicate the \( 1 \rightarrow 0 \) condition, while the dotted lines indicate the \( 0 \rightarrow 1 \) condition. Figure 5 records the power consumption. The peak power consumptions in both the worst propagation delay condition and the best propagation delay condition are recorded. The solid lines indicate the worst propagation delay condition; while the dotted lines indicate the best propagation delay condition.

From Figure 3 to Figure 5, we can see that the dual-rail
routing architecture outperforms the non-uniform routing architecture especially in power consumption. The drawback of uniform dual-rail routing architecture is the extra routing area overheads as compared with the conventional non-uniform routing architecture. Two wire pitches are required for dual-rail wire segments; while only one wire pitch is required for thick wire segments. The maximum area overheads is $1 - \frac{0.281}{0.293} \approx 30\%$ when the whole wire is dual-rail.

B. Scalability Simulations

To ensure the scalability of the dual-rail routing architecture, we next conduct the simulations for different total wire length from 500\(\mu\m\) to 2500\(\mu\m\). The simulation process is similar to the afore-mentioned property simulations. For each of different total wire length condition, we incrementally insert one-tens total wire length of the dual-rail/thick wire from source and then perform the worst case timing simulation. After the best signal propagation delay is found, we record the crosstalk noise and the power consumption accordingly. The simulation results are summarized from Figure 6 to Figure 8.

Figure 6 shows the best signal propagation delay of different total wire length in the worst case condition. Figure 7 and Figure 8 are the results of crosstalk noise and power consumption of different total wire length according to the best dual-rail/thick wire length found in the best signal propagation delay, respectively.

From Figure 6 to Figure 8, we can see that the uniform routing architecture outperforms the non-uniform routing architecture in terms of the crosstalk noise and the power consumption. For the total wire length below 2000\(\mu\m\), these two routing architectures obtain similar signal propagation delay.

V. CONCLUSION

We have proposed an uniform dual-rail routing architecture to cope with the new challenges in DSM era. The proposed architecture exploits the anti-Miller effect between two adjacent wires with the same signal source. Hence, the coupling capacitance between these two wires is reduced. We conduct a series of simulations to examine the characteristics of both the dual-rail and the non-uniform routing architectures in terms of the properties and the scalabilities. The simulation results demonstrate that the dual-rail routing architecture provides a signal propagation channel with similar propagation delay, less crosstalk noise, and less power consumption to the conventional non-uniform routing architecture with moderate routing area overheads. We argue that the uniform dual-rail routing architecture is a wire sizing alternative without incurring layout irregularity and stacked vias overheads.

REFERENCES