

Article

Integrated Circuit Conception: A Wire Optimization Technic Reducing Interconnection Delay in Advanced Technology Nodes

Mohammed Darmi ¹, Lekbir Cherif ¹, Jalal Benallal ^{1,*} , Rachid Elgouri ² and Nabil Hmina ¹

¹ Laboratory of Systems Engineering, National School of Applied Sciences, Ibn Tofail University, BP 242, Av. de L'Université, Kénitra 14 000, Morocco; mohammed.darmi@gmail.com (M.D.); lekhir.cherif@gmail.com (L.C.); hmina@univ-ibntofail.ac.ma (N.H.)

² Laboratory of Electrical Engineering & Telecommunication Systems, National School of Applied Sciences, Ibn Tofail University, BP 242, Av. de L'Université, Kénitra 14 000, Morocco; elgouri.rachid@yahoo.fr

* Correspondence: jalal.benallal@gmail.com; Tel.: +212-661-760-453 (J.B.)

Received: 31 July 2017; Accepted: 25 September 2017; Published: 4 October 2017

Abstract: As we increasingly use advanced technology nodes to design integrated circuits (ICs), physical designers and electronic design automation (EDA) providers are facing multiple challenges, firstly, to honor all physical constraints coming with cutting-edge technologies and, secondly, to achieve expected quality of results (QoR). An advanced technology should be able to bring better performances with minimum cost whatever the complexity. A high effort to develop out-of-the-box optimization techniques is more than needed. In this paper, we will introduce a new routing technique, with the objective to optimize timing, by only acting on routing topology, and without impacting the IC Area. In fact, the self-aligned double patterning (SADP) technology offers an important difference on layer resistance between SADP and No-SADP layers; this property will be taken as an advantage to drive the global router to use No-SADP less resistive layers for critical nets. To prove the benefit on real test cases, we will use Mentor Graphics' physical design EDA tool Nitro-SoC™ and several 7 nm technology node designs. The experiments show that worst negative slack (WNS) and total negative slack (TNS) improved up to 13% and 56%, respectively, compared to the baseline flow.

Keywords: technology nodes; optimization; global route; detail route; wire delay; worst negative slack (WNS); total negative slack (TNS); back-end-of-line (BEOL); non-default-rules (NDRs); self-aligned double patterning (SADP)

1. Introduction

The ultimate goal of an integrated circuits (IC) physical design electronic design automation (EDA) tool is to accomplish any physical implementation, with respect to functional design (no timing violations), and to honor all of the rules of physical design. In advanced technology nodes, from 16 nm all the way to 7 nm, the routing step becomes the most complicated task in the physical implementation of ICs. This is mainly due to the increase in design rules that must be respected, from one technological node to another. Practical examples of the new design rules in 7 nm technology node are the self-aligned double patterning (SADP) [1] and aggressive end-of-line (EOL) rules. In a chip where performance is increasingly more than a simple requirement [2], the perspective of delay due to unexpected physical design challenges is a frightening prospect. These new rules create an urgent need for new techniques that help to improve all quality aspects during physical IC implementation from floor-planning, placement, and clocktree synthesis (CTS) until the route [3]. In this paper, we will take an important physical characteristic that comes with 7 nm SADP as an advantage to optimize wire delay by reducing the wire resistance. This comes from the fact that, due to the manufacturing

process, the SADP layers, compared to No-SADP layers, have a high resistance. It is true that the previous tool versions already take the wire resistance into account during the wire promotion in the global route, but in this paper we are presenting an additional wire promotion that is SADP-based. A direct impact will be significant timing optimization without increasing cell area utilization. As a side effect, the consumed power by the wires will be reduced. As proof of concept, we will apply the proposed routing algorithm, using Nitro-SoC™, (2016.2, Mentor Graphics Corporation, Wilsonville, OR, USA, 2016), the Mentor Graphics' physical design EDA tool, and a significant regression of 7 nm designs. The remainder of the paper is organized as follows: first, we present the global trends of delay sensitivity in advanced technology nodes, and we then try to explain the relationship between the wire delay and its resistance; next, we describe our solution as a new routing optimization method; timing is optimized via net delay reduction; finally, we present the experimental results on the selected regression, and a comparison study between baseline and the new method is also included.

2. Delay Sensitivity in Advanced Technology Nodes

In advanced technology nodes, the delay due to interconnect becomes more important than the delay in the gate. This is due to the increased delay sensitivity to the interconnect parasitic [4]. Figure 1 illustrates the escalating interconnect Resistance-Capacitance (RC) delay with nodes scaling.

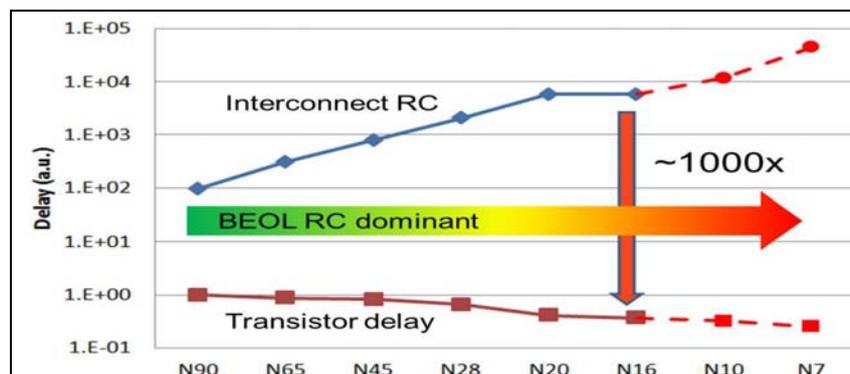


Figure 1. Growing gap between transistor delays and interconnect delay in advanced technology nodes. Reproduced with permission from [5], IEEE, 2013.

With an approximately $10\times$ growing gap for two process nodes, this escalating RC delay represents a significant part of the increase in gate density. This is due to the exponential increase in buffers and drivers counts, and a similar increase in “white” area kept for post-layout buffer insertion [5]. As performance divergence between transistors and interconnects continues to increase, designs have become interconnect-limited [6].

On the other hand, the delay is more sensitive to the wire resistance, more than it is to the wire capacitance [7], as highlighted by Figure 2.

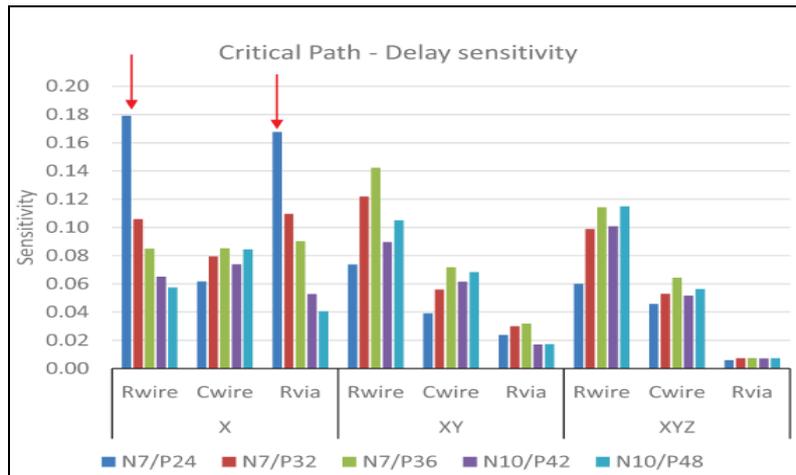


Figure 2. Critical path and delay sensitivity. Reproduced with permission from [7], IEEE, 2016.

Variables in the figure are defined as follows:

X represents local connections, i.e., short connections within cells. In general, the bottom layer is used.

XY represents intermediate connections between cells and cores/modules.

XYZ represents global connections, i.e., long wires.

To address the increasing complexities, physical designers need new routing techniques that can perform a full chip parasitic and timing optimization from placement to the route.

3. Wire Delay Model

The net delay is the time difference, between when a signal is first applied to the net, and when it reaches the other devices connected to that net. The net delay is a direct effect of the finite resistance and capacitance of that net; it is also known as the wire delay.

Wire delay is a function of R_{net} and C_{wire} ($C_{net} + C_{input}$) [8].

For the delay calculation in EDA tools, it is true that the delay is calculated using complex timing models such as Non-Linear Delay Model (NLDM) and Composite Current Source (CCS), but a simple model could be used for a rough estimation of delays like the “Elmore Delay formula” [9] presented in Figure 3. It can be concluded that reducing wire resistance implies wire delay (τ_{DN}) reduction.

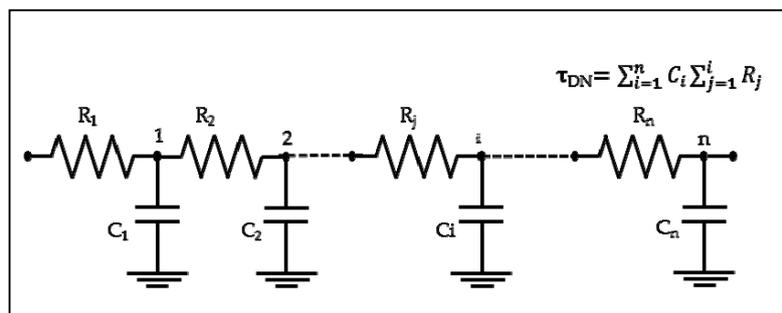


Figure 3. Wire delay model—Elmore Delay formula.

The resistance of a wire is proportional to its length L and inversely proportional to its cross-section A .

The resistance of a rectangular conductor is shown in Figure 4. Where the constant ρ is the resistivity of the material (in W-m) and H is a constant for a given technology [9].

R_s is the sheet resistance of the material, having unit of Ω/sq (Ω -per-square). This expresses that the resistance of a square conductor is independent of its absolute size. To obtain the wire resistance, the sheet resistance is multiplied by its ratio (L/W), as shown in Figure 4.

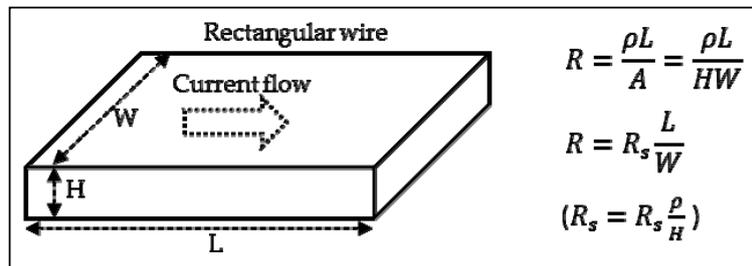


Figure 4. Wire resistance (R) and sheet resistance (R_s).

4. Wire Optimization to Reduce Net Delay

Traditional buffering and upsizing techniques to reduce interconnect delay are no longer as effective due to the area and power impact.

To minimize design costs and to better predict system performance, there is a need to use new techniques to optimize interconnects delay. In addition to what already exists in the EDA tool, we must take maximum advantage of the re-routing transforms during optimization for setup and hold instead of using size and buffer transforms.

In 7 nm technology, the spectrum of metal layer resistivity is large, from the top lesser resistive layers to the bottom SADP layers (M0, M1, M2, and M3).

In addition, wire etching induces great resistance differences based on metal density and spacing, with different values in congested areas versus sparse routing areas, different values for double-spacing around the wire or on only one side, and so on.

Table 1 shows the sheet resistance for each layer of 7 nm node technology used in our case study.

Table 1. Sheet resistance (R_s) in 7 nm node technology.

Layers	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
R_s (Ω/sq)	2.774	2.389	1.279	0.904	0.904	0.904	0.904	0.344	0.344	0.034

Metal 0 and Metal 1 are highly used by standard cells; the router mainly uses layers from Metal 2. From Table 1, changing wire from Metal 2 to Metal 4 reduces its sheet resistance by 54% and by 46% from Metal 3 to Metal 4.

This is an interesting breakpoint, a potential reduction of the total wire resistance by more than 46%.

Our proposed routing optimization technique takes this technology propriety as a solution to reduce the critical net delay by re-routing the delay-critical nets with upper layers.

To limit any side effect that can affect routing quality and timing, the router is driven through non-default-rules (NDRs).

The main guidance is to use layers up to Metal 4 (No SADP layers) to route the critical nets. Nets that are not critical or less critical will be routed with all layers.

The wire optimization of timing critical nets using NDRs enables the optimization and routing engines to balance efficiently between the timing and congestion.

In parallel to the EDA tool software, there is a reference flow script that is “Tool Command Language” TCL-based to run each part of the Place & Route design flow step by step. It is made by a set of organized TCL scripts that cover different stages from placement to post-route. It includes all

needed command and settings to implement a large variety of designs from low to high complexity. The used Nitro-SoC tool begins with floorplanning and placement, and then handles CTS, and routing. Signal integrity (SI) analysis and multi-corner/multi-mode (MCM) analysis can be performed at any stage during the design flow. The Double Via Insertion, Double Patterning fixing, and Design For Manufacturing (DFM) optimizations are performed at the end of the flow [3,10].

Our added solution is a set of TCL scripts that include Nitro-SoC basic commands. This solution is added exactly after the Post Clock Tree Synthesis (Post-CTS) within reference flow.

Figure 5 shows where we are adding the wire optimization inside the initial flow (baseline flow) and the section below gives more details for each step:

1. Create NDR rules that drive the global route to use Metal 4 and up.
2. Identify the best net targets for optimization, which is the most important step because a good optimization starts from good targets that result in a high delay reduction without any side regression. This is subdivided into four steps:
 - a. Start from a list of all nets that violate setup timing.
 - b. Select only nets with a delay up to a predefined threshold value. It is a procedure that helps in finding nets with a critical delay greater than the delay threshold. These nets are the right targets for our optimization.
 - c. Apply costing to remove nets that violate hold timing. A filtering is performed in order to keep only the nets that do not cause a violated hold timing path.
 - d. Remove nets that already have high No-SADP-layer usage. There will otherwise be no room for additional optimization.

Now we have a good list of target nets to re-route.

3. Apply NDR to a sorted list of final target nets.
4. Run the normal global route followed by a detail route.

Before executing a new global route, it is necessary to delete the old net's global route. Then, perform a new one on the list of target nets.

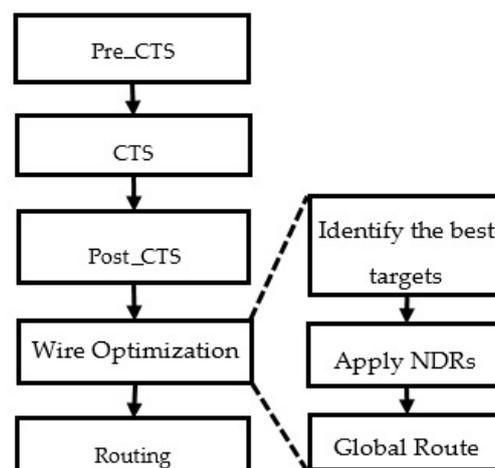


Figure 5. Baseline design flow vs. featured flow.

5. Application and Results

The first experiment was done on three critical nets belonging to the same timing path. Starting from a post-CTS database, we selected three critical nets with a delay higher than a predefined value (10 ps for our use case). We then applied our routing-optimization technique.

As shown in Figure 6, the net resistance is significantly reduced. This result was expected since the router is now constructed to use Metal 4 and up. As a consequence, the net delay is globally reduced by 50%. Figure 7 illustrates this result.

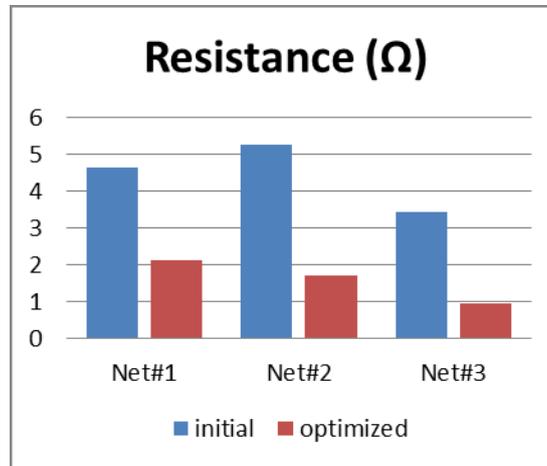


Figure 6. Wire resistance reduction.

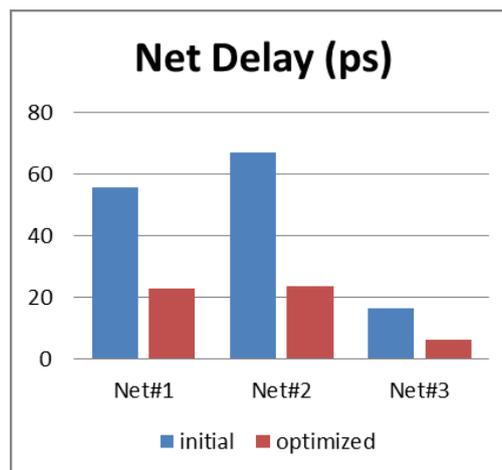


Figure 7. Wire delay reduction.

Figures 8–10 show the metal layers distribution for Net#1, Net#2, and Net#3. It can be seen that, in the optimized routing, the used layers are up to Metal 4.

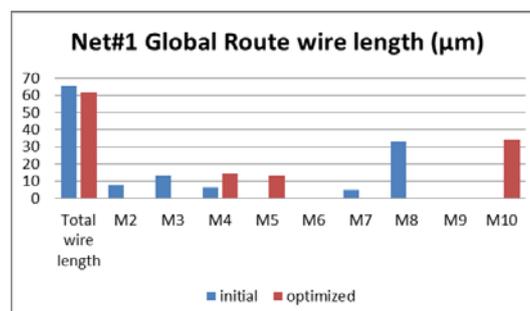


Figure 8. Net#1 global route wire length distribution.

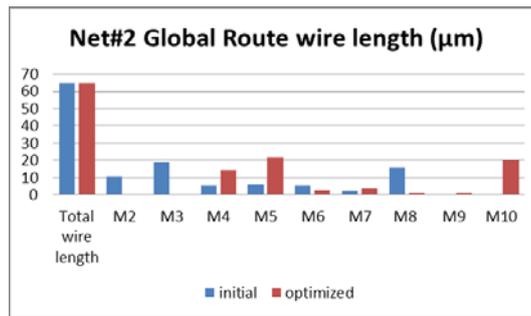


Figure 9. Net#2 global route wire length distribution.

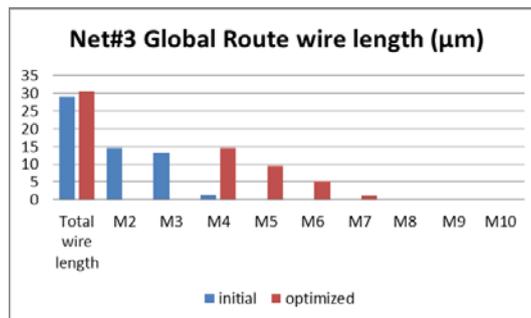


Figure 10. Net#3 global route wire length distribution.

Figures 11–13 show the corresponding wires of the same net before and after detail net re-routing.

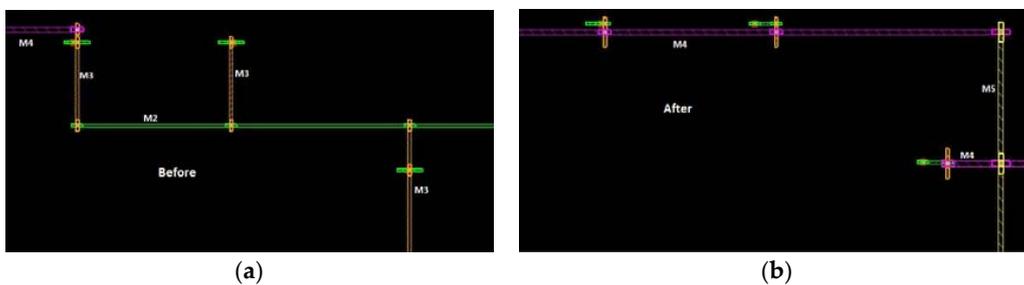


Figure 11. Net#1: detail wire used layers before (a) and after (b) optimization.

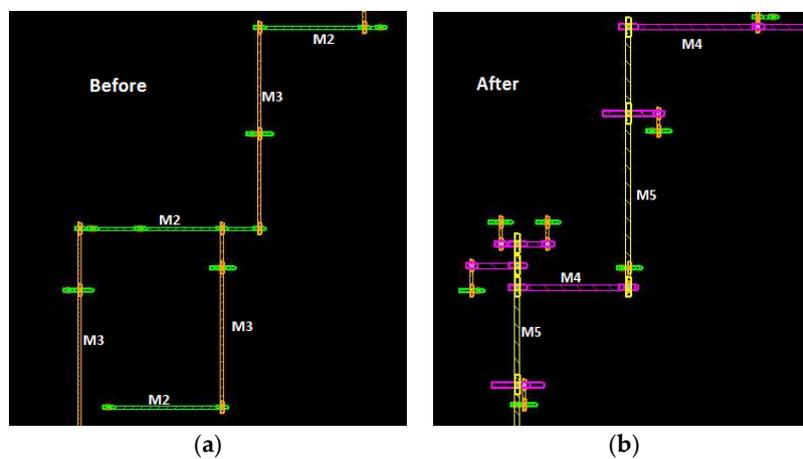


Figure 12. Net#2: detail wire used layers before (a) and after (b) optimization.

Figure 14 shows the expected benefit of the delay reduction; the worst negative slack (WNS) is optimized from -157.4 ps to -56.4 ps.

This first experiment on three critical nets demonstrates a WNS improvement up to $\times 2.8$, with simple net re-routing. This presents an opportunity for physical designers to significantly reduce timing violations using routing optimization.

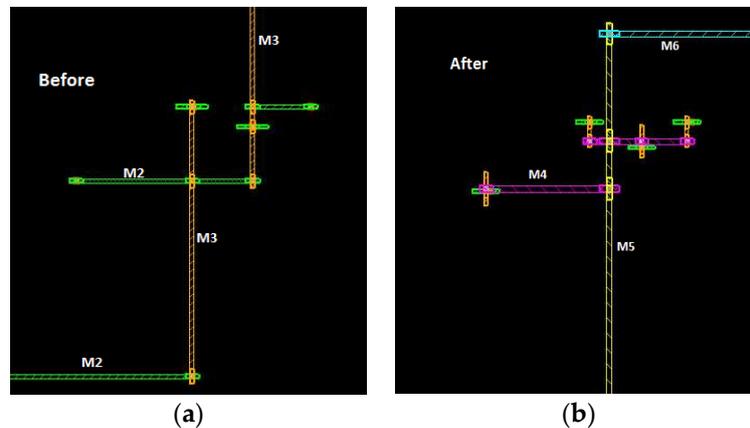


Figure 13. Net#3: detail wire used layers before (a) and after (b) optimization.

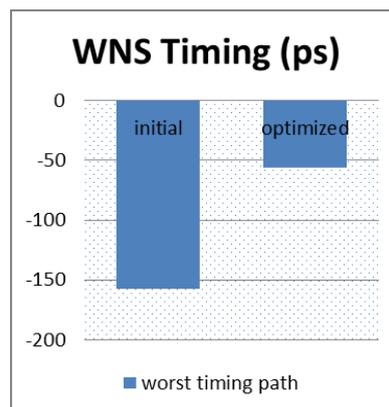


Figure 14. Worst negative slack (WNS) optimization.

The complexity of resistance and capacitance variations makes it nearly impossible for the human mind to determine which combination of layers and via structures to use for a given net in order to obtain the best possible timing and routability.

However, at the same time, we can easily drive a place and route tool such as Nitro-SoC to take advantage of lesser resistivity of wires and vias. This can be achieved through NDRs.

In a real design, the task is more complicated because there is a need to take into account many constraints such as timing hold and setup, max transition, and route congestion. The best way to prove the benefits of a new route optimization methodology is to apply it on large design types. Thus, to go farther than just applying the wire optimization on a few nets of one design, we applied this new route optimization on several types of designs. This will be the aim of the second experiment.

All test cases designs are made with a full place and route flow starting from the floorplan, with the following conditions:

- N7-SADP technology node with stack layers [M0 . . . M12] is used;
- the starting databases are the same for the baseline and feature flow;

- the power grid is achieved using the M11 and M12 top layers, going through others layers to standard cells power ports.

Average area occupancy for each layer for both power and ground nets is distributed as presented in Table 2.

Table 2. Power Grid distribution of area-occupancy per layer.

Layers	M0	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
Occupancy (%)	25	4.4	8.4	2.4	15.8	11.2	9.6	11.2	9.6	15.6	13.8	21	29.6

- clock-tree synthesis is compiled using layers [M4 . . . M10] as preferred layers, with appropriate non-default spacing/width rules; this helps to provide better latency and skew.

Table 3 shows characteristics of the used designs and Table 4 shows the ratio of sequential to combinational logic.

Table 3. Characteristics of used designs.

Design	Design Characteristics
Design#1	Normal: max freq: 2.5 GHz, physical area: 30065 μm^2 , Number of instances: 229,396, Number of macros: 0, Number of nets: 233,846, lib type: H240
Design#2	Normal: max freq: 2.5 GHz, physical area: 38998 μm^2 , Number of instances: 215,835, Number of macros: 0, Number of nets: 220,023, lib type: H240
Design#3	Normal: max freq: 2.5 GHz, physical area: 32632 μm^2 , Number of instances: 215,302, Number of macros: 0, Number of nets: 220,058, lib type: H240
Design#4	Normal: max freq: 2.5 GHz, physical area: 32632 μm^2 , Number of instances: 215,302, Number of macros: 0, Number of nets: 220,058, lib type: H300
Design#5	Complex: max freq: 1.66 GHz, physical area: 0.53 mm^2 , Number of instances: 1,521,984, Number of macros: 80, Number of nets: 1,583,731, lib type: H240
Design#6	Complex: max freq: 1.66 GHz, physical area: 3.9 mm^2 , Number of instances: 23,533, Number of macros: 86, Number of nets: 648,689, lib type: H240

Table 4. Ratio of sequential vs. combinational logic.

Design	#Sequential	#Combinational	Ratio (%)
Design#1	11,424	219,443	5%
Design#2	11,425	205,622	6%
Design#3	11,426	205,078	6%
Design#4	11,444	216,265	5%
Design#5	56,880	1,326,688	4%
Design#6	95,022	547,014	17%

The wire optimization is performed at the end of the post-CTS step to improve the timing during the global route; the quality of results (QoR) are measured at the end of the detail route. Table 5 shows how much the WNS and TNS (total negative slack) are reduced, when adding the new wire optimization flow, taking advantage of layer resistance to the normal flow.

In all test cases, the re-route of critical nets with No-SADP layers shows a good reduction of both the WNS and the TNS.

Table 5. Quality of results (QoR) summary.

Design	Featured flow, Baseline flow, %(Baseline-Featured)/Baseline			
	Post-CTS (clocktree synthesis)		Route	
	WNS (worst negative slack, ps)	TNS (total negative slack, ns)	WNS (ps)	TNS (ns)
Design#1	-192.3, -224.1, 14%	-222.5, -346.3, 36%	-215, -228.9, 6%	-283.1, -286.6, 1%
Design#2	-136.8, -181.5, 25%	-535.4, -652, 18%	-131.9, -151.9, 13%	-384.9, -417.4, 8%
Design#3	-16, -64.7, 75%	-1.15, -22.2, 95%	-51, -48.7, 5%	-5.6, -3.6, 56%
Design#4	-33, -73.4, 55%	-6.6, -53.1, 88%	-64, -65.1, 2%	-17.3, -28.2, 39%
Design#5	-38.7, -51.4, 25%	-2.92, -9.61, 70%	-171.6, -193.4, 11%	-55.76, -58.02, 4%
Design#6	-93, -165, 44%	-0.96, -6.39, 85%	-397, -439, 10%	-50.2, -55.8, 10%

We added the new feature post-CTS and we ran the same routing flow both in baseline and featured flows, which means that the cell density should remain the same before and after the route, as we perform net re-routing without any additional optimization, i.e., without added buffers, inverters, or cell sizing. Table 6 shows the cell density before and after the route, and Table 7 shows the buffer–inverter cell count in the initial database and in the routed database for each test case.

Table 6. Before route and after route cell density in featured flow and in baseline flow.

Design	Featured Flow (%)	Baseline Flow (%)
Design#1	63.07	63.07
Design#2	63.94	63.94
Design#3	63.23	63.23
Design#4	66.66	66.66
Design#5	63.35	63.35
Design#6 (Top)	17.55	17.55

Table 7. Buffer–inverter cell count in the initial database and in the routed database.

Design	Initial db	After Route	Add Buffers-Inverters
Design#1	43,648	43,262	−386
Design#2	40,471	37,979	−2492
Design#3	39,941	40,668	727
Design#4	44,522	46,622	2100
Design#5	256,860	284,369	27,509
Design#6	136,504	254,041	117,537

The objective of this article is to demonstrate that we can improve timing by simple global-route net re-routing. We show here that the post-route optimization can absolutely be simplified, especially due to less TNS. The intent at the end is to have this feature native in the global route.

For the case “Design#6”, we have completed the full PNR flow by performing post-route optimization to see how much this new feature can help on closing the timing. In Table 8, the final post-route timing becomes close to be met.

Table 8. Post-route timing closure on Design#6.

Design	Featured flow, Baseline flow, %(Baseline-Featured)/Baseline		
	Post Route		
	WNS (ps)	TNS (ns)	#Violated Endpts
Design#6	-17.19, -22.53, 24%	-0.315, -1.9, 83%	64, 263, 76%

6. Conclusions

In advanced technology nodes, timing and routability convergence becomes a difficult task that complicates a physical designer's work. As the goal of a place and route tool is to natively support all new rules and new process features in these technology nodes and allow for an automatic design closure, new optimization techniques should be developed in all implementation flows. One of these techniques that leads to an important WNS and TNS improvement is the wire optimization. Such a feature reduces net delay by a simple critical net re-routing. This has been proven in a first-use case study; the net delay was reduced by 50%, and the worst negative slack (WNS) by 64%. Moreover, this optimization was performed and validated on six designs with different characteristics and complexities; the WNS and TNS were improved up to 13% and 56%, respectively. None of these transformations impact area utilization.

At the end, interconnect RC delay is an open research area that still requires, among other things, wire length reduction, and wire and via optimization through their resistances and capacitances. All of this work must be done for each new tech node by supporting all new features that are natively inside an EDA tool, so that allow implementation engineers can focus more on the design itself instead of the performance closure.

Author Contributions: M.D. and L.C. conceived and designed the experiments; L.C. performed the experiments with support from J.B.; M.D. and L.C. analyzed the data; J.B. contributed reagents/materials/analysis tools; M.D. and L.C. wrote the paper. R.E. and N.H. supervised the project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Huynh-Bao, T. Statistical Timing Analysis Considering Device and Interconnect Variability for BEOL Requirements in the 5-nm Node and Beyond. Available online: <http://ieeexplore.ieee.org/document/7827016/>. (accessed on 19 January 2017).
2. International Technology Roadmap for Semiconductors. 2013. Available online: <http://www.itrs2.net>. (accessed on 31 July 2017).
3. *Nitro-SoC™ and Olympus-SoC™ User's Manual*, Software Version 2016; November 2016.
4. Prasad, D.; Pan, C.; Naeemi, A. Impact of interconnect variability on circuit performance in advanced technology nodes. In Proceedings of the 17th International Symposium on Quality Electronic Design (ISQED), Santa Clara, CA, USA, 15–16 March 2016; pp. 398–404.
5. Yeap, G. Smart mobile SoCs driving the semiconductor industry: Technology trend, challenges and opportunities. In Proceedings of the Electron Devices Meeting (IEDM), 2013 IEEE International, Washington, DC, USA, 9–11 December 2013; pp. 1–3.
6. Or-Bach, Z. Monolithic 3D IC. 2014. Available online: http://www.eetimes.com/author.asp?doc_id=1322783. (accessed on 31 July 2017).
7. Tōkei, Z. End of Cu roadmap and beyond Cu. In Proceedings of the 2016 IEEE International Interconnect Technology Conference/Advanced Metallization Conference (IITC/AMC), San Jose, CA, USA, 23–26 May 2016; pp. 1–58.
8. Sadrusham, N.J. Net Delay or Interconnect Delay or Wire Delay or Extrinsic Delay or Flight Time. 2008. Available online: <http://asic-soc.blogspot.com/2008/10/netdelay.html>. (accessed on 31 July 2017).
9. Jan, M. Rabaey, Anantha Chandrakasan and Borivoje Nikolic. 2013. Available online: <http://ic.sjtu.edu.cn/ic/dic/wpcontent/uploads/sites/10/2013/04/Rabaey-Digital-Integrated-Circuits-A-Design-Perspective.pdf>. (accessed on 31 July 2017).
10. *Nitro-SoC™ and Olympus-SoC™ Advanced Design Flows Guide*, Software Version 2016.2.R1; November 2016.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).