A Polynomial Time Optimal Diode Insertion/Routing Algorithm for Fixing Antenna Problem *

Li-Da Huang, Xiaoping Tang, Hua Xiang, D. F. Wong

Department of Computer Sciences
University of Texas at Austin
Austin, TX 78712 U.S.A.
{lida, tang, xiangh, wong@cs.utexas.edu}

I-Min Liu

Silicon Perspective Corporation
3211 Scott Blvd., Suite 100
Santa Clara, CA 95054 U.S.A.
liu@siperspective.com

Abstract

Antenna problem is a phenomenon of plasma induced gate oxide degradation. It directly affects manufacturability of VLSI circuits, especially in deep-submicron technology using high density plasma. Diode insertion is a very effective way to solve this problem. Ideally diodes are inserted directly under the wires that violate antenna rules. But in today's high-density VLSI layouts, there is simply not enough room for "under-the-wire" diode insertion for all wires. Thus it is necessary to insert many diodes at legal "off-wire" locations and extend the antenna-rule violating wires to connect to their respective diodes. Previously only simple heuristic algorithms were available for this diode insertion and routing problem. In this paper, we show that the diode insertion and routing problem for an arbitrary given number of routing layers can be optimally solved in polynomial time. Our algorithm guarantees to find a feasible diode insertion and routing solution whenever one exists. Moreover, we can guarantee to find a feasible solution to minimize a cost function of the form $\alpha \cdot L + \beta \cdot N$ where L is the total length of extension wires and N is the total number of vias on the extension wires. Experimental results show that our algorithm is very efficient.

1 Introduction

Yield and reliability of VLSI circuit have always been an important item on the agenda of IC manufacturers. With the continuous and rapid increase in complexity of VLSI designs and fabrication technologies, manufacturing yield and product reliability is becoming one of the most important issues among the other existing ones, such as small die size, high speed, low-power and so on [1, 2]. Fine feature size of modern IC technologies are typically achieved using plasma-based processes. As the technology enters the deep-submicron era, more stringent process requirements make some advanced high density plasma (HDP) reactors adopted in the production lines to achieve fine-line patterns [5]. However, these plasma-based processes have a tendency to charge conducting components of a fabricated structure. The existing experimental evidence indicates that charging may affect the quality of the thin oxide. This is called "antenna effect" (also called "plasma induced gate oxide damage") [2, 3]. During metallization, long floating interconnects act as temporary capacitors and store charges gained from the energy provided by fabrication steps such as chemical mechanical polishing (CMP). A random discharge of the floating node due to subsequent process steps could permanently damage transistors rendering the IC useless [4, 6]. For example, the exposed polysilicon and metal structures connected to a thin oxide transistor will collect charge from the processing environment (e.g., reactive ion etch) and damage the transistor when the discharging current flows through the thin oxide.

Although the mechanism of the gate oxide damage is not very well understood, the precise experimental relationships between the amount of damage and the antennas have been studied [4, 6, 7]. Since the plasma damage is caused by the electrical charging of devices during plasma processes, the damage increases with an increase in the area of the exposed conductor (antenna) during the plasma process. In order to reduce or prevent damage to the gate oxide from the plasma process, and thus to ensure reliability of VLSIs, a circuit layout rule that considers the antenna effect (antenna rule) is employed. The conventional antenna rule restricts a maximum antenna size or antenna ratio allowed for circuit layout. Recent studies show that the damage considering all plasma-based manufacturing operations increases in proportion to both the area and the perimeter of the antennas [9]. A more accurate model considering the cumulative oxide damage is discussed in [13]. These models provide a good guideline for router or physical layer EDA tools to help in reducing the antenna effect damage and getting higher yield and reliability.

Maly et al. [2] proposed a method for detecting antenna condition and calculated both the area and the perimeter of antennas using a general purpose design rule checking (DRC) program. However, the method does not indicate any measure to feedback the antenna information to a layout generation. On the other hand, Wang et al. [3] proposed a channel router which considers the antenna effect. They introduce a layer restriction to a conventional channel router and this limits the maximum length of the wires with antenna problem. But this restriction increases the die size and net length. It is hard to estimate antenna before the layer assignment is completed, so they estimate the antenna to be all the wires from gates to the highest layer wires. This inaccuracy of estimation causes a severe layering restriction in that all the routings from the gates must go through the higher layer pattern near the gates. Shirota et al [8, 9] proposed a router which combines a traditional router with a modification of wires for reducing the antenna effect damage using a rip-up and reroute method. This method reduces the damage with only a little penalty on die size and performance. But this method will introduce vias when modifying the route, and will impact signal integrity especially in high speed designs. Chen et al [13] adopted diode insertion method to fix antenna problem. In [13], only a simple heuristic algorithm was presented for this diode insertion and routing problem. Ideally diodes are inserted directly under the wires that violate antenna rules. But in today's high-density VLSI layouts, there is simply not enough room for "under-the-wire" diode insertion for all wires. Thus it is necessary to insert many diodes at legal "off-wire" locations and extend the antenna-rule violating wires to connect to their respective diodes. Previously only simple heuristic algorithms were available for this diode insertion and routing problem. In this paper, we show that the diode insertion and routing problem can be optimally solved

*This work was partially supported by the National Science Foundation under grant CCR-9912390.

in polynomial time. Our algorithm guarantees to find a feasible diode insertion and routing solution whenever one exists. Moreover, we can guarantee to find a feasible solution to minimize a cost function of the form $\alpha \cdot L + \beta \cdot N$ where L is the total length of extension wires and N is the total number of vias on the extension wires. Experimental results show that our algorithm is very efficient.

Section 2 will review the antenna effect and Section 3 will give the problem formulation. An optimal diode insertion and routing algorithm will be introduced in Section 4. Section 5 presents the experiment results, and concludes the paper.

2 Antenna Effect

Charging occurs when conductor layers not covered by a shielding layer of oxide are directly exposed to plasma. The amount of such charging is proportional to this plasma-exposed area. If the charged conductor layers are connected only to the gate oxide, Fowler-Nordheim (F-N) tunneling current will discharge through the thin oxide and will cause damage to it.

During the discharge, the larger gate area will have less damage. Therefore, a basic antenna rule defines the area ratio between the conductor area and the gate area. For example, a $200\mu m$ long by $1\mu m$ wide polysilicon wire connected to two channel regions of $2\mu m$ by $0.6\mu m$ and $1\mu m$ by $0.6\mu m$ has an antenna ratio of 111. Therefore, the polysilicon rules require that the area of the polysilicon over field oxide divided by the area of the transistor gate (thin oxide area) must be less than Np, which is a limit that depends on the process and on design targets.

On the other hand, if the amount of charging collected by connected conductor layer patterns could be released through a low impedance path, such as previously formed diffusion layer pattern (e.g., source/drain), it will not introduce the gate oxide damage, see Fig. 1. These conductor layers, connected only to the gate oxide, are called antennas.

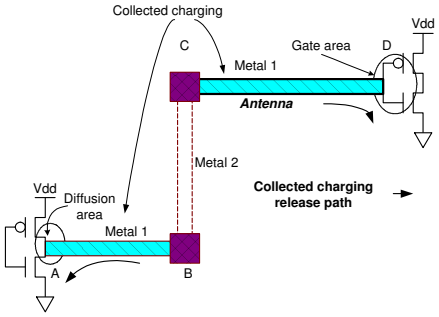

Figure 1: Before metal 2 is constructed, the charging on the wire segment (C, D) discharges through the gate area and causes the antenna problem. The charging on the wire segments (A, B) will discharge through diffusion area, so the device will not be damaged.

A more accurate analysis of the cause of the charging collected during the deep sub micron VLSI manufacturing operations [9] shows that the perimeter length of conductor layer patterns must also be included into calculation. There are three types of plasma-based manufacturing processes:

1. Conductor layer pattern etching processes: The amount of accumulated charge is proportional to the perimeter length of conductor layer patterns. Etching processes divide conductor layer plates into innumerable routing patterns. In the late stage of the processes, the perimeters of the routings are directly exposed to plasma.

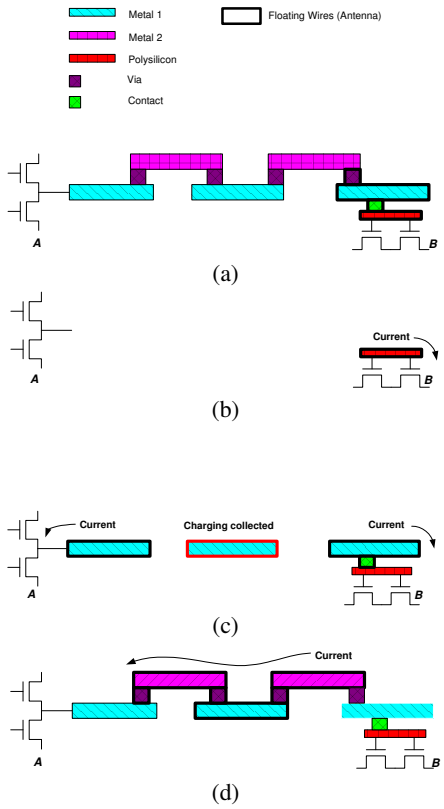

Figure 2: Two metal layers process: (a) An example of connection from an output of device A to an input gate of device B. (b) Late stage of poly layer pattern etching of (a). The collected charge on the poly pattern is discharged through the gate and damages the device B. (c) Late stage of metal 1 layer pattern etching of (a). The charge on the right most metal wire pattern is discharged through the gate and damages device B. The charge on the left most metal wire pattern is discharged through the diffusion area without damaging to device A. The charge in the disconnected metal 1 pattern (i.e., the middle one) is still accumulated. (d) Late stage of metal 2 layer pattern etching of (a). All the charge on metal 2 pattern and the accumulated charge on metal 1 pattern is discharged through diffusion area of device A without damage.

2. Ashing processes: The amount of accumulated charge is proportional to the area of the conductor layer patterns. Ashing processes remove remaining photo resist layer after etching processes of conductor layer. In the late stage of the processes, the area of conductor layer patterns is directly exposed to plasma.

3. Contact etching processes: The amount of accumulated charge is proportional to the area of the total area of the contacts. Contact etching processes dig holes between two conductor layers. In the late stage of the processes, the area of all the contacts on the lower conductor layer pattern is directly exposed to plasma.

As a result, considering all the plasma-based processes, the risk of gate-oxide damage is proportional to the area and perimeter length of antenna routings and inversely proportional to the area and perimeter length of the gate oxide. Fig. 2 shows how the accumulated charging will be released as the higher conductor layer are constructed. Fig. 3 shows the discharging current path for a three metal layers process. The antenna consists of polysili-

con, contact, metal 1, metal 2, and vias. Extension wires can start from metal 1 or metal 2 layer of the violating wire. But if the violating wire consists of only metal 1 layer pattern, the extension wire can not go up to metal 2 layer, and go back to metal 1 later to connect the diode. The discharging path through the diode will be disconnected. Because when the antenna problem occurs, the metal 2 layer patterns are still not constructed.

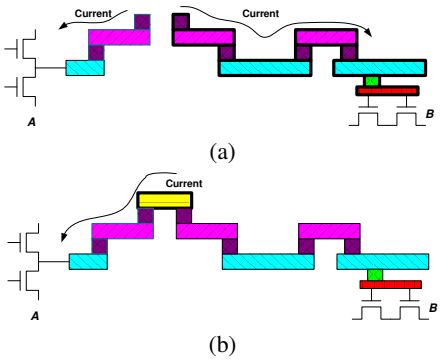

Figure 3: Three metal layers process: (a) The discharging current path at the late stage of metal 2 pattern etching. Highlighted wire segments are antennas. The discharging current from antennas will damage device B (b) At the late stage of metal 3 layer pattern etching, The collected charge on the metal 3 pattern is discharged through the diffusion without damaging to device A.

There are three kinds of solution to reduce antenna effect [13]:

1. Router options: Break signal wires and routes to upper levels by jumper insertion. This reduces the charge amount for each net during manufacturing.

2. Embedded protection diode: Add protection diodes on every input port for every standard cell. Since these diodes are embedded and fixed, they consume unnecessary area when there is no violation at the connecting wire.

3. Diode inserting after placement and route: Fixing only the wire with the antenna violation which will not waste routing resources. During wafer manufacturing, all the inserted diodes are floating (or ground). Since the input ports are high impedance, the charge on the wire flows through the insert floating/ground diode (composite) instead of flowing into the input port. One diode can be used to protect all input ports that are connected to the same output ports.

Fig. 4 contains three violating wires that need diode insertion. If the violating wire lies above a space that can insert a diode, then the diode is inserted right below that wire. Otherwise, extension wires are necessary to connect the violating wires to a diode insertion legal space, as shown in Fig. 5.

The difference between jumper insertion and protection diode is the impact on the timing characteristics of the wire. Each jumper needs at least two vias. These vias will introduce 10Ω to hundreds Ω as the process moved to $0.18\mu m$ or deeper submicron technology. The via resistance is almost the same order as the wire resistance, so the caused RC delay and the effect on the signal integrity can not be ignored. On the other hand, the diode insertion will introduce diode capacitance. The amount of the capacitance is around $1fF$ for $0.35\mu m$ technology. As compared to the wire capacitance for $0.35\mu m$ technology, which is around $0.1pF$ to several pF depending on the length of the wire, the diode insertion caused delay is negligible. Besides, the via itself is more

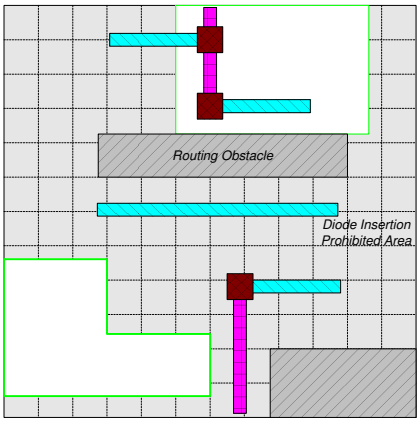

Figure 4: Routing obstacle and diode insertion obstacles are scattered over the routing space.

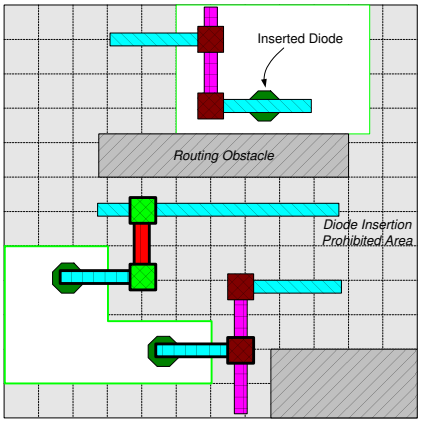

Figure 5: The diode could be inserted under the violating wire. The violating wires in the diode insertion obstacles need wire extension to reach the diodes.

sensitive to process variation than the diode. For example, the distortion in sub-wavelength optical system will make the hole in the via smaller or even close and results into increasing the via resistance or disconnecting the wire. Therefore, it is better to adopt diode insertion than jumper insertion. Furthermore, if there is no space along the violating wire for adding the diode, it is necessary to find a position along the wire to extend the wire to a diode.

3 Problem formulation

Our problem is formulated as finding the best position on the violating wires and extend the wire to the diodes with the minimum cost. Diode location assignment and routing the extension wire are two closely related tasks. Diode assignment alone followed by the wire connection can not achieve the optimal solution, since the interconnection can not be accurately estimated without carrying out actual routing step when doing the diode assignment. Therefore, these two steps must be considered simultaneously.

For example, in Fig. 6, after placement and routing, there are some antenna rule violating wires among routing or diode insertion obstacles. If the diode is inserted by a greedy algorithm in (a, b, c, d) order, then the extension wire for wire c will block wire d to get connected to a diode, as in Fig. 7. Besides, the total length of the extension wire and the number of vias can not be guaranteed to be the minimum. So, simultaneous diode insertion and routing is not a trivial problem. One feasible solution is shown in Fig. 8.

However, there is a better solution with less vias and shorter total length of extension wires as shown in Fig. 9.

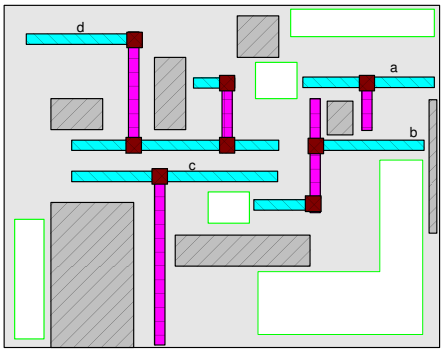

Figure 6: Four antenna rule violating wires.

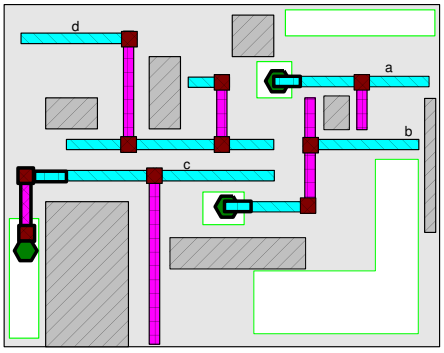

Figure 7: The result of diode insertion by greedy algorithm in *(a, b, c, d)* order. Wire *d* could be blocked because of previous extension wires.

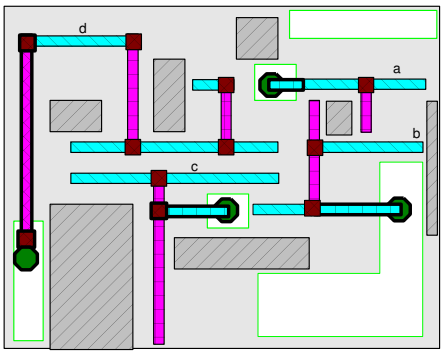

Figure 8: A feasible diode insertion and routing solution.

The diode connection with extension wire is modeled by a grid graph $G = (V, E)$. It contains antenna rule violating wires, diode insertion locations, and routing/diode insertion obstacle information. Each wire needs to be connected to a diode located in a legal space. The grid nodes that are adjacent to each other are connected by an edge representing the wire segments. Each grid in the diode insertion region represents a diode insertion location. The routing obstacles represents the existed wiring and the hard blocks that does not allow over-block routing. The diode insertion obstacles are the occupied diffusion area where diode insertion is prohibited. The problem is stated as follows:

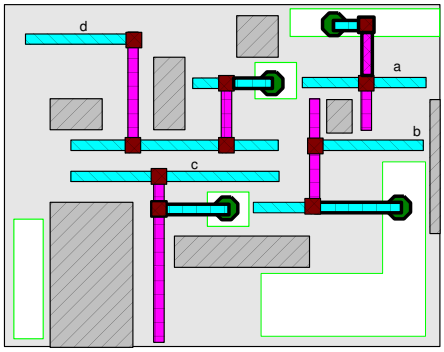

Figure 9: A better solution with less vias and shorter extension wire length.

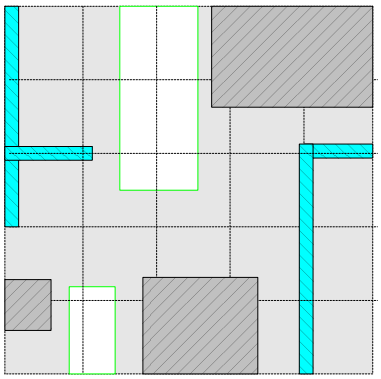

Figure 10: A diode insertion and routing problem in 2-metal process. In this problem, there are two violating wires in metal 1.

Problem 1 DIR **(Diode Insertion and Routing)** *Given a routing grid graph $G(V, E)$, a set of violating wires W, a set of diode insertion positions D, and a set of routing blocks B, find a set of paths P connecting each wire in W to a diode position in D, such that the cost is minimum. The cost function is defined as $\alpha \cdot L + \beta \cdot N$, where L is the total length of extension wires and N is the total number of vias on the extension wires. Each path in P can only be constructed from higher conductor layer to lower conductor layer.*

Each wire in W represents the violating wire that needs the extension wire to reach the diode. Each grid on a wire represents the possible position to extend the wire. D is the set of the diode legal positions. Each element in D occupies one grid on G, and can only be connected to one violating wire. Set B represents the obstacles on the routing space that no wire can across them. We restrict the extension wire being only constructed from a higher conductor layer to a lower conductor layer to avoid disconneting the discharging path when the extension wire go up to higher conductor layers.

In the following section, we will develop an algorithm to solve the DIR problem, and present that it is solvable in polynomial time.

4 Algorithm ($DIRMCF$) for Diode Insertion and Routing by Min-Cost Flow

To simplify the presentation, we first present diode insertion in a two conductor layers problem. In this case, only the wires at the lower conductor layer will have the antenna problem. Based on the grid graph, we construct a network graph first, and then

apply a minimum cost network flow algorithm to get the solution. Each path in the solution represents a connection between a violating wire and an inserted diode. The extension wire routing is embedded along the path.

Given $G(V, E)$, W, and D, the network graph $G_F(V_F, E_F)$, edge capacity U_F, and edge cost C_F are constructed as follows:

1. $V_F = \{s, s_{w_i}, t\} \cup V$, where s, t are the source node and sink node respectively, and s_{w_i} is the node added for each wire $w_i \in W$.
2. $E_F = E \cup \{(s, s_{w_i})\} \cup (s_{w_i}, v_{w_{i,j}}) \cup \{(u_k, t)\}$, where $v_{w_{i,j}}$ is the grid point on the violating wire w_i, and u_k is the grid point that a diode can be inserted.
3. Edge capacity: Assign each edge (u, v) in E_F as one. ($i.e.$,, $U_F(u, v) = 1, \forall u, v \in E_F$.)
4. Edge cost: $C_F(s, v) = 0, C_F(u, t) = 0$ for other edge $e \in E$, $C_F(e)$ is assigned as the weight of wire or via.

Fig. 11 is a simple example illustrating the constructed network graph based on a diode insertion and routing problem shown in Fig. 10. In this case, there are two violating wires, and five legal diode insertion locations on the grid graph. Since the extension wire can not cross the routing obstacles, all the edges and nodes inside the routing obstacles could be removed.

Once the solution of the min-cost network flow of G_F is obtained, a flow f in the G_F assigns extension wires and diodes for the violating wires on grid graph G. In general, f consists of a set of disjoint paths in G_F from s to t, each corresponds to a path on the grid graph G connecting a violating wire to a diode insertion location. In Fig. 12, there are two paths in f on the resultant G_F. Each path connects the violating wire to one diode insertion location as shown in Fig. 13. The extension wire routing and diode assignment are considered simultaneously and achieve the minimum total length of the extension wires when the connections are completed. We have the following theorem.

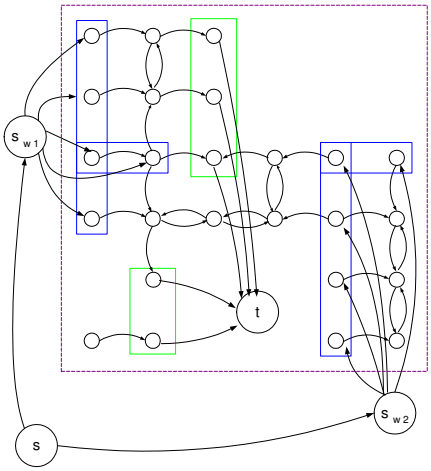

Figure 11: The network graph for Fig. 10.

Theorem 1 *The min-cost flow f in G_F assigns extension wires and diode insertions in G, and $|f| \leq \min(|W|, |D|)$. If $|f| = |W|$, all the violating wires have been connected to inserted diodes with minimum total length. If $|f| < |W|$, not all the violating wire groups can be connected to the diodes.*

The algorithm $DIRMCF$ is summarized as follows:

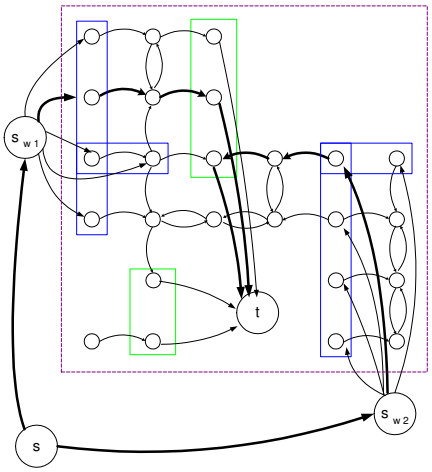

Figure 12: A Min-Cost Flow solution for the network in Fig. 11.

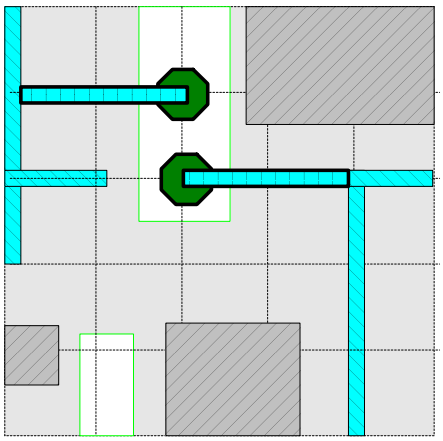

Figure 13: The extension wires and inserted diodes for the problem in Fig. 10. based on the min-cost flow solution in Fig. 12.

Algorithm $DIRMCF(G, W, D, B)$

1. Remove the nodes and edges in G corresponding to B.
2. Construct the network graph $G_F(V_F, E_F)$.
3. Assign the edge capacity and cost (i.e., U_F, and C_F).
4. Apply min-cost flow algorithm on G_F.
5. Generate the layout of the extension wires, and allocate the inserted diodes according to the flow from step 4.

There are several polynomial-time algorithm to find the min-cost flow in a given network graph (the step 4)[14]. All the other steps can be done in $O(E)$. We conclude that the timing complexity of $DIRMCF$ is $O(VE \log(V^2/E) \log(V))$. Since the cost from s_{w_i} to the boundary grids is always less then the cost to the interior grids, the diodes are assigned only on the boundary grids of the diode insertion region. Therefore, we can leave the grids and edges on the boundary of diode insertion region of G, and remove all other inside grids and edges.

Our algorithm can be easily extended to solve the problem with an arbitrary given number of routing layers. We present the

extension to the 3-layer case and the extension to general multi-layer case is obvious. For three conductor layers process, the antennas consists of two lower conductor layers and vias. The extension wires can only start from higher conductor layer of the antenna through vias to lower conductor layer and connect the diodes. So when constructing the network graph from a two conductor layers routing grid graph G, each via on node u in G will introduce an edge in G_F from the node u_F^h in G_F corresponding to u at the higher conductor layer to the node u_F^l in G_F corresponding to u at the lower conductor layer on the grid graph. Fig. 14 is part of the network graph constructed from a grid graph with two conductor layers (e.g., Matel 1 and Metal 2). The source node s, the violating wire s_{w_i}, and sink node t are not shown in Fig. 14. Notice that the direction of via edges can only be from the higher layer to the lower layer. By assigning the cost β on these via edges of the network graph, and the cost α on the conductor edges, the total cost could be represented as $\alpha \cdot L + \beta \cdot N$, where L and N are the total length and the total number of vias of the paths.

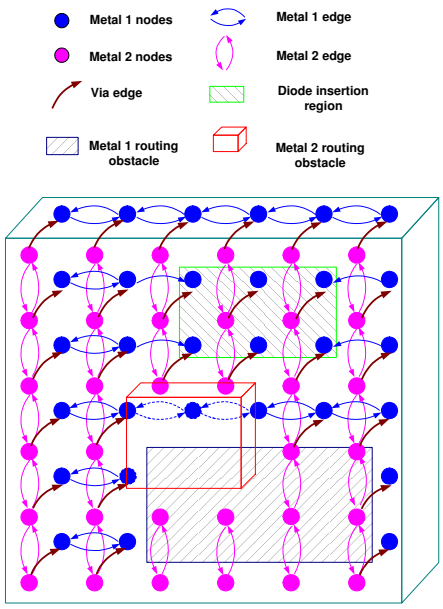

Figure 14: Network for 3-metal process. Each via edge represents a possible via on the grid graph. The direction of the via edge in network graph can only be from a higher layer to a lower layer.

5 Experimental results and Conclusion

The $DIRMCF$ is implemented in C++ programming language, and tested on Sun Sparc Ultra10 (440MHZ, 256 MB DRAM). Since $DIRMCF$ is an exact algorithm to solve DIR problem, we demonstrate the efficiency by different settings of the test cases.

The grid size of our test cases are ranging from 20x20 to 150x150. Table 1 lists the results of $DIRMCF$. The performance of $DIRMCF$ is very well. All violating wires in the test cases are connected to diode insertion locations. In test case $T5$, there are 120 violating wires, 8 routing obstacles and 8 diode insertion obstacles. The diode insertion regions are all irregular shapes. Most test cases are solved within 10 seconds. $T5$ is solved within 40 seconds because of large grid size and many violating wires.

In conclusion, a polynomial time optimal diode insertion and routing algorithm is proposed to fix the antenna problem resulting from deep sub-micron plasma based technology. Our algorithm can construct the extension wires and diode insertion simultaneously and minimize the cost, which is a linear combination of via cost and wire length cost. By adjusting the weights of the combination, we have different diode insertion schemes. Our algorithm guarantees to obtain optimal solutions, and the simulation result shows it is very efficient.

name	wire	grid size	runtime(s)
$T1$	4	20x20	0.03
$T2$	5	25x20	0.06
$T3$	15	50x50	0.41
$T4$	60	100x100	6.62
$T5$	120	150x150	39.32

Table 1: Experimental results of $DIRMCF$ algorithm

References

[1] H. T. Heineken, J. Khare, W. Maly, P.K. Nag, C. Ouyang, et.al., CAD at the design-manufacturing interface, *Proc. ACM/IEEE Design Automation Conf.*, pp. 321-326, 1997.

[2] W. Maly, et.al., Detection of an antenna effect in VLSI designs, Research report No. CMUCAD-96-17, May. 1996.

[3] K. P. Wang, M. Marek-Sadowska, W. Maly, Layout design for yield and reliability,*5th ACM SIGDA Physical Design Workshop*, 1996.

[4] H. Shin, C. -C. King, C. Hu, Thin Oxide Damage by plasma etching and ashing process, *Proc. Of IRPS*, 1992.

[5] S. Krishnan, et. al., Assessment of charge-induced damage to ultra-thin gate MOSFETs, *ITEM 97*, pp. 445-448

[6] H. Watanabe, et.al., A wafer level monitoring method for plasma-charging damage using antenna PMOSFET test structure, *IEEE Trans. Semiconductor manufacturing* vol. 10, no. 2, may. 1997.

[7] R. Rakkhit, et.al., Process induced oxide damage and its implications to device reliability of submicron transistors, *Proc. Of IRPS*, 1993.

[8] H. Shirota, et.al., A new rip-up and reroute algorithm for very large scale gate arrays, *CICC '96*, 1996.

[9] H. Shirota, et.al., A new router for reducing "Antenna Effect" in ASIC design. *IEEE Custom Integrated Circuits Conference*, pp. 601-604, 1998.

[10] M. Terai, et.al., A new efficient routing method for channelless sea-of-gates arrays, *CICC '94*, no. 28.3, 1994.

[11] Takahashi, et.al., Adaptive cut line selection in min-cut placement for large scale sea-of-gates arrays, *ICCAD*, pp. 428-431, 1994.

[12] Ko Noguchi, Akira Matsukoto, Noriaki Oda, A model for evaluating cumulative oxide damage from multiple plasma processes. *IEEE 38th Annual International Reliability Physics Symposium, 2000*

[13] Peter H. Chen, Sunil Malkani, Chun-Mou Peng, James Lin, Fixing antenna problem by dynamic diode dropping and jumper insertion, *Proc. Quality Electronic Design, ISQED 2000*, pp 275-282.

[14] Ravindra K. Ahuja, Thomas L. Magnanti, James B. Orlin, Network Flows *Prentice Hall. 1993*