

Low Power High Speed Memory Architecture Using Multi Bit Flip Flops

G.Barathi¹, S.Ramesh Kumar² Dr.R.Ganesan³

¹ PGScholar, Sethu Institute of technology, Madurai.

² Assistant Professor, Sethu Institute of technology, Madurai.

³ Professor/HOD,

Abstract – In modern VLSI designs, power consumed by clocking has taken a major part of the whole design especially for those designs using deeply scaled CMOS technologies. Hence in this paper we propose an algorithm for reducing the power consumption by replacing some flip-flops with fewer multi-bit flip-flops without affecting the performance of the original circuit. The flip-flop replacement will change the location of some flip-flops, leading to violation of timing and placement capacity constraints. To avoid this problem, several techniques are proposed. Utilizing the properties of manhattan distance and coordinate transformation, first we identify those flip-flops that can be merged and their legal regions. Next, a combination table is built to enumerate all possible combinations. Finally, the flip-flops are merged in hierarchical manner. According to the experimental results, our algorithm significantly reduces clock power by 20-30% and besides power reduction minimizing the total wirelength is also considered.

KEY WORDS – Clock power, multi-bit flip-flop manhattan distance, merging.

I. INTRODUCTION

A clock system and a logic part consumes dominant part of the total chip power. The clock system itself consumes 20–45% of the chip power. In this clock system power, 90% is consumed by the flip-flops themselves and the last branches of the clock distribution network which directly drives the flip-flops [1]. This is due to the high switching activity.

$$P_{clk} = C_{clk} V_{dd}^2 f_{clk} \quad (1)$$

where P_{clk} is clock power, f_{clk} is the clock frequency, V_{dd} is the supply voltage, and C_{clk} is the switching capacitance including the gate capacitance of flip-flops (sequential elements) controlled by the clock signal, the interconnect capacitance of the clock network, and the capacitance associated with the buffers/inverters used in the clock network. Several methodologies [2], [3] have been proposed to reduce the power consumption of clocking. Given a design that the locations of the cells have been determined, the power consumed by clocking can be reduced further by replacing several flip-flops with multi-bit flip-flops. During clock tree synthesis, less number of flip-flops means less number of clock sinks. Thus, the

resulting clock network would have smaller power consumption and uses less routing resource.

Applying MBFFs may have the following advantages:

- 1) smaller design area due to shared clock drivers and clock gating cells;
- 2) less delay and power of the clock network due to fewer clock sinks and smaller capacitive load on the clock net;
- 3) controllable clock skew because of common clock and enable signals for a group of flip-flops and reduced depth of a clock tree;
- 4) improved routing resource utilization especially when considering design for testability. The required routing resource for a scan chain is greatly reduced because of fewer cells in a scan chain.

Fig. 1 shows an example of merging two 1-bit flip-flops into one 2-bit flip-flop. If we replace the two 1-bit flip-flops as shown in Fig. 1(a) by the 2-bit flip-flop as shown in Fig. 1(b), the total power consumption can be reduced because the two 1-bit flip-flops can share the same clock buffer. However, the locations of some flip-flops would be changed after this replacement, and thus the wirelengths of nets connecting pins to a flip-flop are also changed. To avoid violating the timing constraints, we restrict that the wirelengths of nets connecting pins to a flip-flop cannot be longer than specified values after this process. Besides, to guarantee that a new flipflop can be placed within the desired region, we also need to consider the area capacity of the region.

A. Related work

Chen *et al.* [4] and Hou *et al.* [5] leverage on register banking at logic synthesis and at early physical synthesis, respectively. However, the subsequent timing and routing cost of the clustered result may somewhat deviate from what is expected at such early stages.

On the other hand, Yan and Chen [7], Chang *et al.* [8], and Wang *et al.* [9] postponed this task to postplacement to further consider the timing and even routing issues. Yan and Chen [7] analyzed the timing-safe region for each flipflop and then constructed an intersection graph to record the pairwise overlapping of these regions. They reduced MBFF clustering to minimum clique partitioning and solved it by iteratively merging flip-flops with fewest compatible flip-flops. However, they assumed the available bit numbers of the given MBFF library are contiguous and unlimited.

Considering a discrete and finite MBFF library, Chang *et al.* [6] proposed the problem of using multi-bit flip-flops to reduce power consumption in the post-placement stage. They use the graph-based approach to deal with this

problem. In a graph, each node represents a flip-flop. If two flip-flops can be replaced by a new flip-flop without violating timing and capacity constraints, they build an edge between the corresponding nodes. After the graph is built, the problem of replacement of flip-flops can be solved by finding an m -clique in the graph. The flip-flops corresponding to the nodes in an m -clique can be replaced by an m -bit flipflop. They use the branch-and-bound and backtracking algorithm [8] to find all m -cliques in a graph. Because one node (flip-flop) may belong to

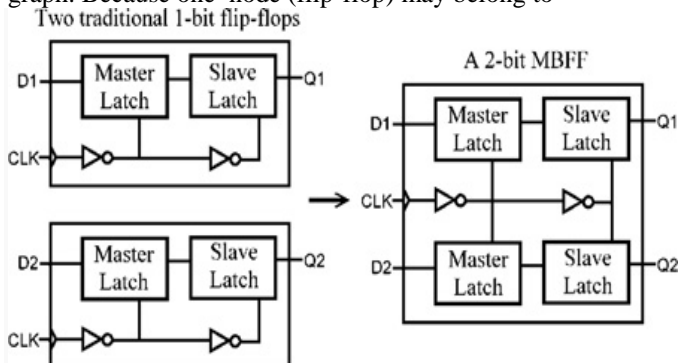


Fig. 1. Replacing two traditional FFs by a 2-bit MBFF

several m -cliques (m -bit flip-flop), they use greedy heuristic algorithm to find the maximum independent set of cliques, which every node only belongs to one clique, while finding m -cliques groups. However, if some nodes correspond to k -bit flip-flops that $k > 1$, the bit width summation of flip-flops corresponding to nodes in an m -clique, j , may not equal m . If the type of a j -bit flip-flop is not supported by the library, it may be time-wasting in finding impossible combinations of flip-flops.

II. OUR ALGORITHM

Our design flow can be roughly divided into three stages. Please see Fig. 5 for our flow. In the beginning, we have to identify a **legal placement region** for each flip-flop f_i . First, the feasible placement region of a flip-flop associated with different pins are found based on the timing constraints defined on the pins. Then, the legal placement region of the flip-flop f_i can be obtained by the

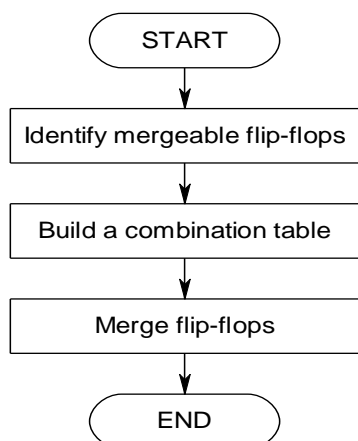
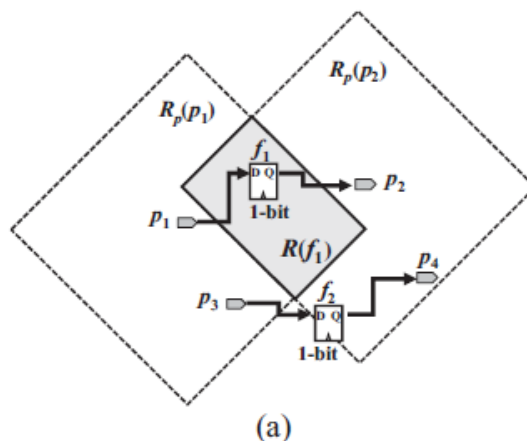


Fig. 2. Flow chart of our algorithm

overlapped area of these regions. However, because these regions are in the diamond shape, it is not easy to identify the overlapped area. Therefore, the overlapped area can be identified more easily if we can transform the coordinate system of cells to get rectangular regions. In the second stage, we would like to build a **combination table**, which defines all possible combinations of flip-flops in order to get a new multi-bit flip-flop provided by the library. The flip-flops can be merged with the help of the table. After the legal placement regions of flip-flops are found and the combination table is built, we can use them to **merge flip-flops**. To speed up our program, we will divide a chip into several bins and merge flip-flops in a local bin. However, the flip-flops in different bins may be mergeable. Thus, we have to combine several bins into a larger bin and repeat this step until no flip-flop can be merged anymore.

A. Transformation of placement space

The replacement of some flip-flops with multi-bit flip-flops would change the routing length of the nets that connect to a flip-flop, it inevitably changes timing of some paths. To avoid that timing is affected after the replacement, the Manhattan distance between pin p_i and flip-flop f_j cannot be longer than the given constraint $S(p_i)$ defined on the pin p_i [i.e., $M(p_i, f_j) \leq S(p_i)$]. Since there may exist several pins connecting to f_i , the legal placement region of f_i are the overlapping area of several regions. As shown in Fig. 3(a), there are two pins p_1 and p_2 connecting to a flip-flop f_1 , and the feasible placement regions for the two pins are enclosed by dotted lines, which are denoted by $R_p(p_1)$ and $R_p(p_2)$, respectively. Thus, the legal placement region $R(f_1)$ for f_1 is the overlapping part of these regions. In Fig. 3(b), $R(f_1)$ and $R(f_2)$ represent the legal placement regions of f_1 and f_2 . Because $R(f_1)$ and $R(f_2)$ overlap, we can replace f_1 and f_2 by a new flip-flop f_3 without violating the timing constraint, as shown in Fig. 3(c).



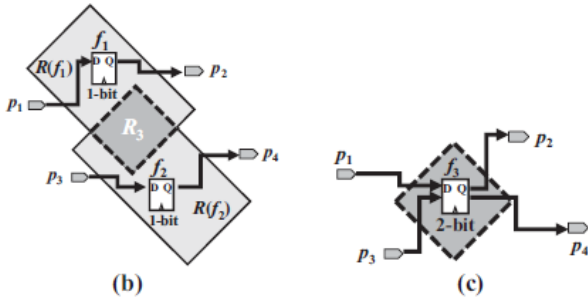


Fig. 3. (a) Feasible regions $R(p_1)$ and $R(p_2)$ for pins p_1 and p_2 which are enclosed by dotted lines, and the legal region $R(f_1)$ for f_1 which is enclosed by solid lines. (b) Legal placement regions $R(f_1)$ and $R(f_2)$ for f_1 and f_2 , and the feasible area R_3 which is the overlap region of $R(f_1)$ and $R(f_2)$. (c) New flip-flop f_3 that can be used to replace f_1 and f_2 without violating timing constraints for all pins p_1, p_2, p_3 , and p_4 .

However, it is not easy to identify and record feasible placement regions if their shapes are diamond. Moreover, four coordinates are required to record an overlapping region [see Fig. 4(a)]. Thus, if we can rotate each segment 45° , the shapes of all regions would become rectangular, which makes identification of overlapping regions become very simple. For example, the legal placement region, enclosed by dotted lines in Fig. 4(a), can be identified more easily if we change its original coordinate system [see Fig. 4(b)]. In such condition, we only need two coordinates, which are the left-bottom corner and right-top corner of a rectangle, as shown in Fig. 4(b), to record the overlapped area instead of using four coordinates.

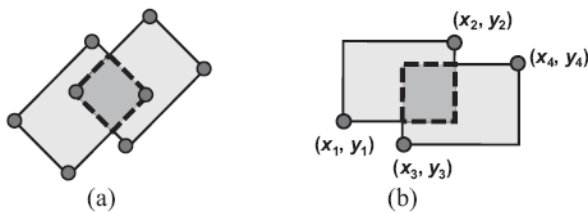


Fig. 4. (a) Overlapping region of two diamond shapes. (b) Rectangular shapes obtained by rotating the diamond shapes in (a) by 45° .

The equations used to transform coordinate system are shown in (1) and (2).

$$x' = \frac{x+y}{\sqrt{2}} \Rightarrow x'' = \sqrt{2} \cdot x' = x+y \quad (2)$$

$$y' = \frac{-x+y}{\sqrt{2}} \Rightarrow y'' = \sqrt{2} \cdot y' = -x+y. \quad (3)$$

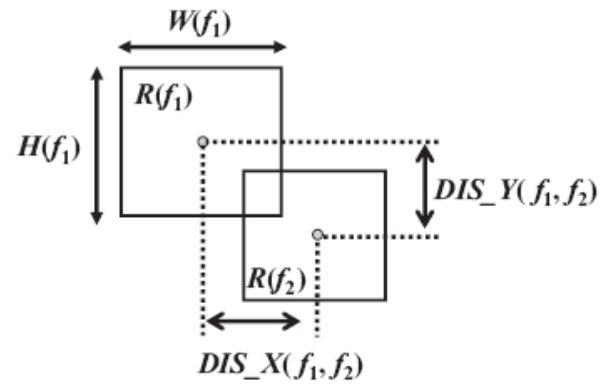


Fig. 5. (a) Overlapping region of two diamond shapes. (b) Rectangular shapes obtained by rotating the diamond shapes in (a) by 45° .

Then, we can find which flip-flops are mergeable according to whether their feasible regions overlap or not. Since the feasible placement region of each flip-flop can be easily identified after the coordinate transformation, we simply use (3) and (4) to determine whether two flip-flops overlap or not.

$$DIS_X(f_1, f_2) < \frac{1}{2} (W(f_1) + W(f_2)) \quad (4)$$

$$DIS_Y(f_1, f_2) < \frac{1}{2} (H(f_1) + H(f_2)) \quad (5)$$

where $W(f_1)$ and $H(f_1)$ [$W(f_2)$ and $H(f_2)$] denote the width and height of $R(f_1)$ [$R(f_2)$], respectively, in Fig. 8, and the function $DIS_X(f_1, f_2)$ and $(DIS_Y(f_1, f_2))$ calculates the distance between centers of $R(f_1)$ and $R(f_2)$ in x -direction (y -direction).

B. Build a Combination table

If we want to replace several flip-flops by a new flip-flop f_i' (note that the bit width of f_i' should equal to the summation of bit widths of these flip-flops), we have to make sure that the new flip-flop f_i' is provided by the library L when the feasible regions of these flip-flops overlap. Now a combination table is to be built, which records all possible combinations of flip-flops to get feasible flip-flops before replacements. Thus, we can gradually replace flip-flops according to the order of the combinations of flip-flops in this table.

Library L	
Type 1	Type 2
1 bit	4 bit

Combinational table	
n1	n2
1 bit	4 bit

Fig.6(a) Initialize the library L and the combination table T

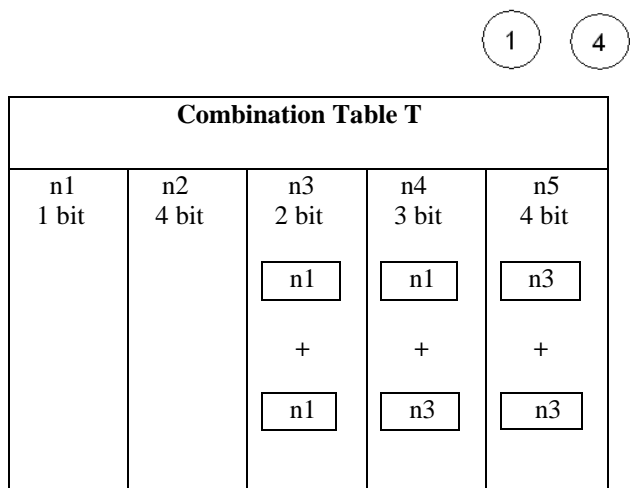


Fig.6(b) Pseudo types are added into L , and the corresponding binary tree is also build for each combination in T .

For example, consider a library L that provides two types of flip-flops, whose bit widths are 1 and 4 (i.e., $b_{min} = 1$ and $b_{max} = 4$), in Fig. 6(a). We first initialize two combinations $n1$ and $n2$ to represent these two types of flip-flops in the table T [see Fig. 6(a)]. Next, the function **InsertPseudoType** is performed to check whether the flip-flop types with bit widths between 1 and 4 exist or not. Thus, two kinds of flip-flop types whose bit widths are 2 and 3 are added into L , and all types of flip-flops in L are sorted according to their bit widths [see Fig. 6(b)]. Now, for each combination in T , we would build a binary tree with 0-level, and the root of the binary tree denotes the combination. Next, we try to build new legal combinations according to the present combinations. By combing two 1-bit flip-flops in the first combination, a new

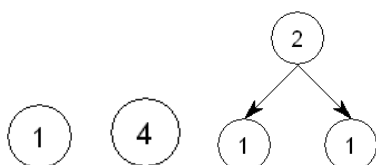
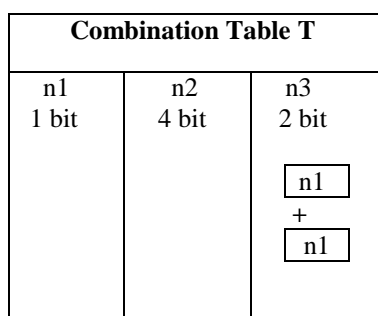
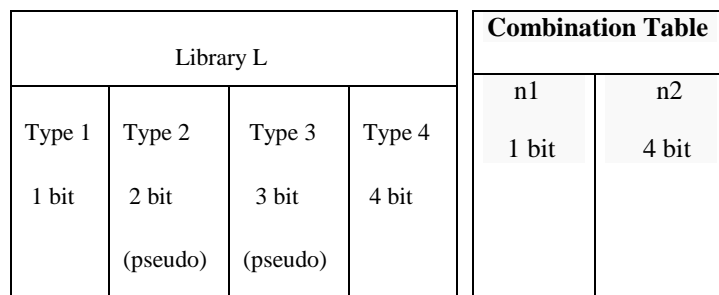


Fig. 6(c) New combination $n3$ is obtained from combining two $n1$ s.

Combination $n3$ can be obtained [see Fig. 6(c)]. Similarly, we can get a new combination $n4$ ($n5$) by combining $n1$



and $n3$ (two $n3$'s) [see Fig. 6(d)]. Finally, $n6$ is obtained by combing $n1$ and $n4$.

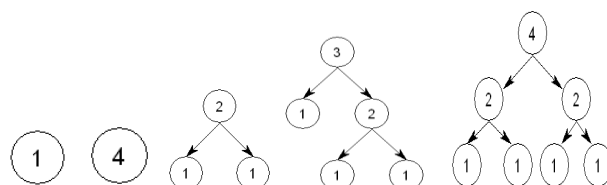


Fig. 6(d) New combination $n4$ is obtained from combining $n1$ and $n3$, and the combination $n5$ is obtained from combining two $n3$ s.

All possible combinations of flip-flops are shown in Fig.6(e). Among these combinations, $n5$ and $n6$ are duplicated since they both represent the same condition, which replaces four 1-bit flip-flops by a 4-bit flip-flop. To speed up the process, $n6$ is deleted from T rather than $n5$ because its height is larger. After this procedure, $n4$ becomes an unused combination [see Fig. 6(e)] since the root of binary tree of $n4$ corresponds to the pseudo type, $type3$, in L and it is only included in $n6$. After deleting $n6$, $n4$ is also need to be deleted. The last combination table T is shown in Fig. 6(f).

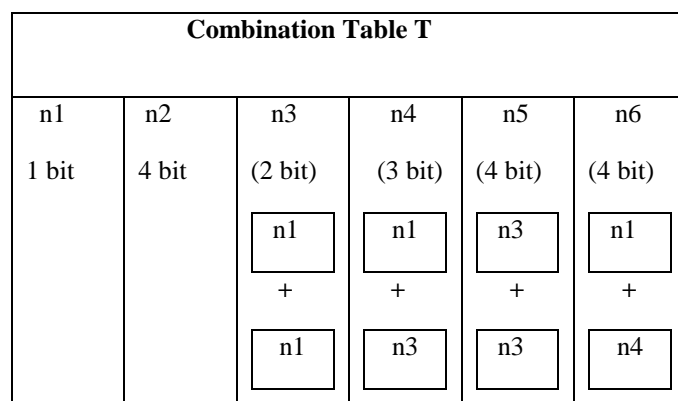


Fig. 6(e) New combination $n6$ is obtained from combining $n1$ and $n4$.

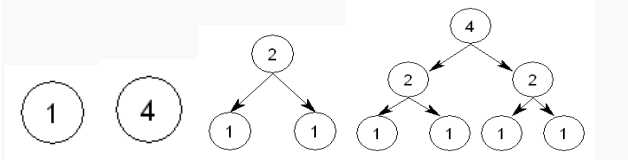


Fig.6(f) Last combination table is obtained after deleting the unused combination in (e).

C. Merge Flip-Flops.

After the combination table is built the combination of flip-flops are used for merging and replacing. To reduce the complexity the whole placement region is divided into several sub regions. Then, several subregions are combined into a larger subregion and the flip-flops are replaced again so that those flip-flops in the neighboring subregions can be replaced further. Finally, those flip-flops with pseudo types are deleted in the last stage because they are not provided by the supported library.

1) Region Partition: To speed up our problem, we divide the whole chip into several subregions. By suitable partition, the computation complexity of merging flip-flops can be reduced significantly. As shown in Fig. 11, we divide the region into several subregions, and each subregion contains six bins, where a bin is the smallest unit of a subregion.

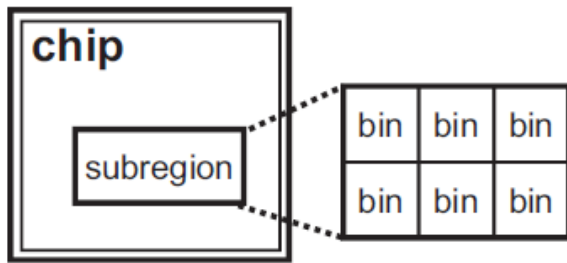


Fig. 7 Example of region partition with six bins in one subregion.

2) Replacement of Flip-flops in Each Subregion: Before illustrating the procedure to merge flip-flops, first an equation is given to measure the quality if two flip-flops are going to be replaced by a new flip-flop as follows:

$$\text{cost} = \text{routing_length} - \alpha \times \sqrt{\text{available_area}} \quad (6)$$

where routing_length denotes the total routing length between the new flip-flop and the pins connected to it, and available_area represents the available area in the feasible region for placing the new flip-flop. α is a weighting factor. The cost function includes the term routing_length to favor a replacement that induces shorter wirelength. Besides, if the region has larger available space to place a new flip-flop, it implies that it has higher opportunities to combine with other flip-flops in the future and more power reduction. Thus, we will give it a smaller cost. Once the flip-flops cannot be merged to a higher-bit type we ignore the available_area in the cost function, and hence α is set to 0. First the flip-flops are linked below the combinations corresponding to their types in the library.

Then, for each combination n in T , we serially merge the flip-flops linked below the left child and the right child of n from leaves to root. Based on the binary tree, we can find the combinations associated with the left child and right child of the root. Based on the binary tree, we can find the combinations associated with the left child and right child of the root. Hence, the flip-flops in the lists named l_{left} and l_{right} , linked below the combinations of its left child and its right child are checked. Then, for each flip-flop f_i in l_{left} , the best flip-flop f_{best} in l_{right} , which is the flip-flop that can be merged with f_i with the smallest cost recorded in C_{best} , is picked. For each pair of flip-flops in the respective list, the combination cost is computed if they can be merged and the pair with the smallest cost is chosen. Finally, we add a new flip-flop f' in the list of the combination n and remove the picked flip-flops which constitutes the f' .

For example, given a library containing three types of flipflops (1-, 2-, and 4-bit), we first build a combination table T as shown in Fig. 8(a). In the beginning, the flip-flops with various types are, respectively, linked below $n1, n2,$ and $n3$ in T according to their types. Suppose we want to form a flipflop in $n4$, which needs two 1-bit flip-flops according to the combination table. Each pair of flip-flops in $n1$ are

Combination Table T			
n1	n2	n3	n4
1 bit	4 bit	2 bit	4 bit
		n1	n3
		+	+
		n1	n3

selected and checked to see if they can be combined. If there are several possible choices, the pair with the smallest cost value is chosen to break the tie. In Fig. 8(a), f_1 and f_2 are chosen because their combination gains the smallest cost. Thus, we add a new node f_3 in the list below $n4$, and then delete f_1 and f_2 from their original list [see Fig. 8(b)].

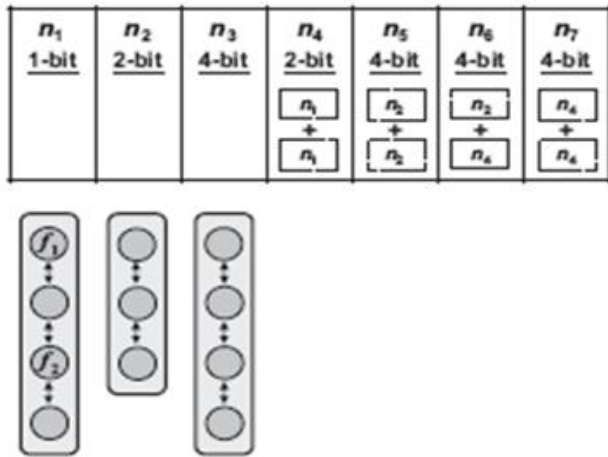


Fig. 8(a) Sets of flip-flops before merging.

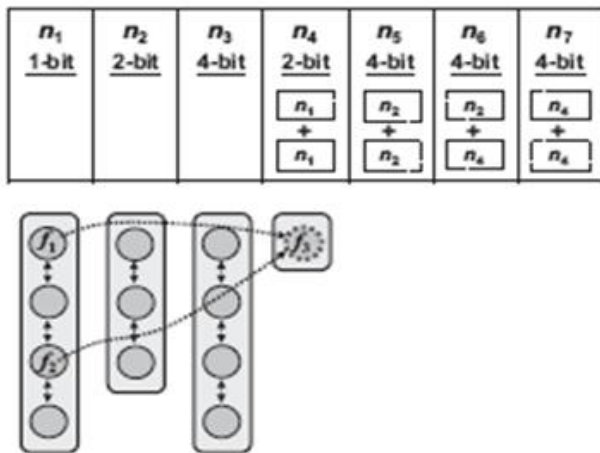


Fig. 8(b) Two 1-bit flip-flops, f_1 and f_2 , are replaced by the 2-bit flip-flop f_3 .

Similarly, f_4 and f_5 are combined to obtain a new flip-flop f_6 , and the result is shown in Fig. 8(c). After all flip-flops in the combinations of 1-level trees (n_4 and n_5) are obtained as shown in Fig. 8(d), we start to form the flip-flops in the combinations of 2-level trees (n_6 , and n_7).

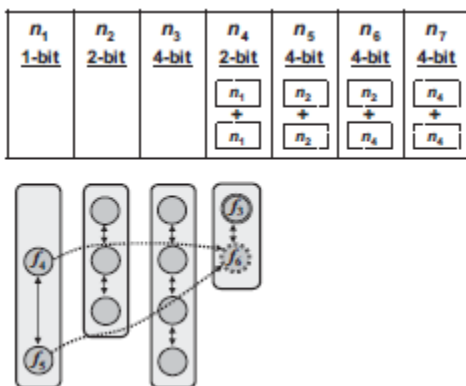


Fig. 8(c) Two 1-bit flip-flops, f_4 and f_5 , are replaced by the 2-bit flip-flop f_6 .

In Fig. 8(e), there exist some flip-flops in the lists below n_2 and n_4 , and we will merge them to get flip-flops in n_6 and n_7 , respectively. Suppose there is no overlap region between the couple of flipflops in n_2 and n_4 . It fails to

form a 4-bit flip-flop in n_6 . Since the 2-bit flip-flops f_3 and f_6 are mergeable, we can combine them to obtain a 4-bit flip-flop f_{10} in n_7 . Finally, because there exists no couple of flip-flops that can be combined further, the procedure finishes as shown in Fig. 8(f).

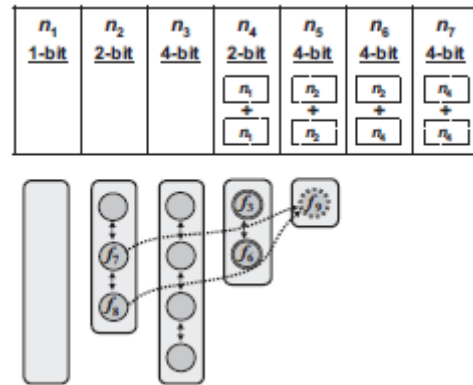


Fig. 8(d) Two 2-bit flip-flops, f_7 and f_8 , are replaced by the 4-bit flip-flop f_9 .

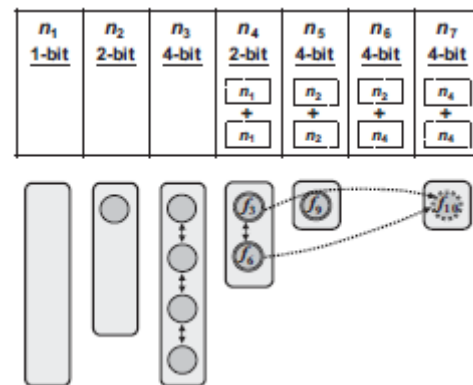


Fig.8(e) Two 2-bit flip-flops, f_3 and f_6 , are replaced by the 4-bit flip-flop f_{10} .

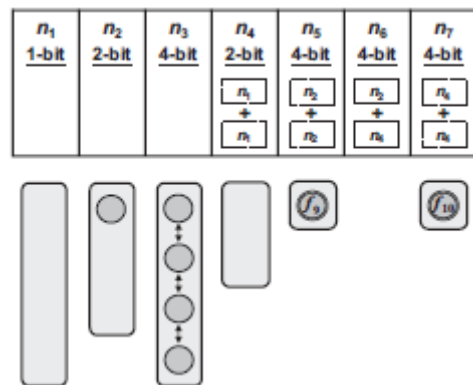


Fig.8(f) Sets of flip-flops after merging.

If the available overlap region of two flip-flops exists, we can assign a new one to replace those flip-flops. Once there is sufficient space to place the new flip-flop in the available region, the algorithm will perform the replacement, and the new generated flip-flop will be placed in the grid that makes the wirelength between the flip-flop and its connected pins smallest. If the capacity constraint of the bin, B_k , which the grid belongs to will be violated after the new flip-flop is placed on that grid, we will search the bins near B_k to find a new available grid

for the new flip-flop. If none of bins which are overlapped with the available region of new flip-flop can satisfy the capacity constraint after the placement of new flip-flop, the program will stop the replacement of the two flip-flops.

III. RESULTS AND DISCUSSIONS

In this algorithm, there exist two values which would affect our results: the first one is the dimension of a subregion since we would partition a chip into several subregions. The second one is the parameter used in the cost function of (6). Thus, we first do some experiments to explore better values for these two parameters.

1) Influence of Region Size on Performance: In this part, we first determine a suitable size for each subregion during partitioning. Since the execution time is actually dominated by the average number of flip-flops included in a subregion, we use the number of flip-flops in a single subregion to represent the size of a subregion, which can be obtained from multiplying the number of bins in a subregion by the average number of flip-flops in a bin. We sweep the number of flip-flops included in a subregion to observe its effect on power consumption and execution. While a subregion gets larger, the execution time becomes longer. However, the power consumption does not decrease proportionally. On the contrary, if the subregion size becomes very small, the power consumption will increase significantly.

To balance execution time and power consumption, we select 600 as the number of flip-flops in a single subregion (the normalized power and execution time are about 83% and 0.8% if the number of flip-flops in a single subregion is 600).

2) Influence of Weighting Factor α on Performance: Since the parameter α used by (6) would affect our results, it is necessary to find a suitable value for getting better results. In this experiment, we sweep α from 0 to 3 to get the data of power consumption and wirelength. While the value of α becomes larger, the power reduction ratio gets larger.

IV. SIMULATION RESULTS AND DISCUSSIONS

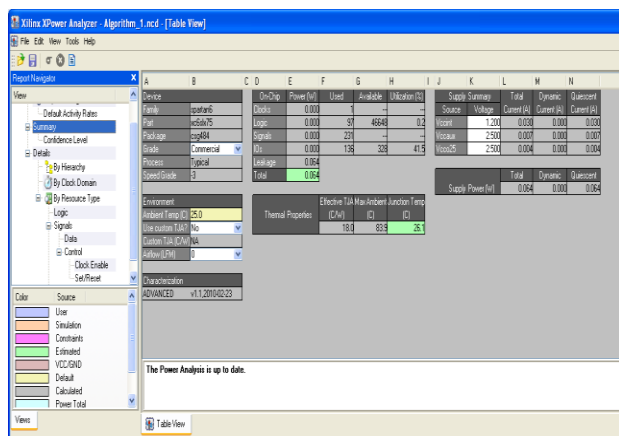


Fig. 9. Power consumed by the algorithm.

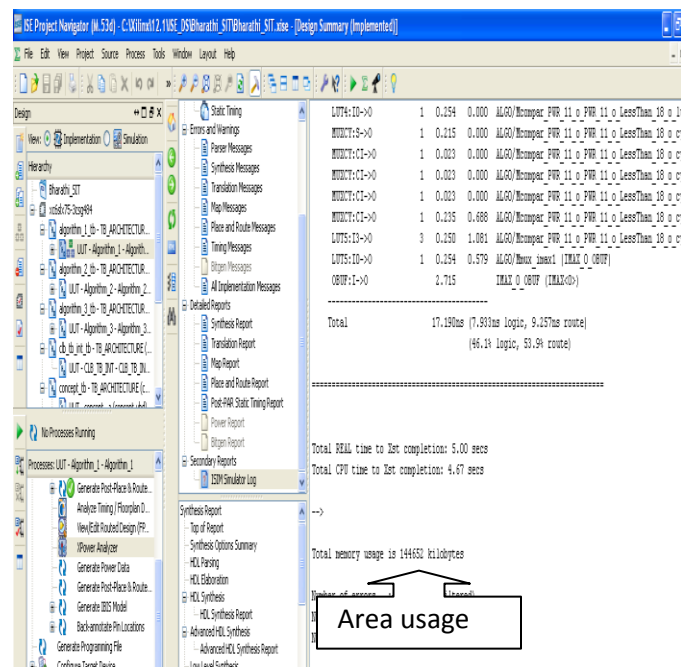


Fig. 10 Area Usage of the algorithm

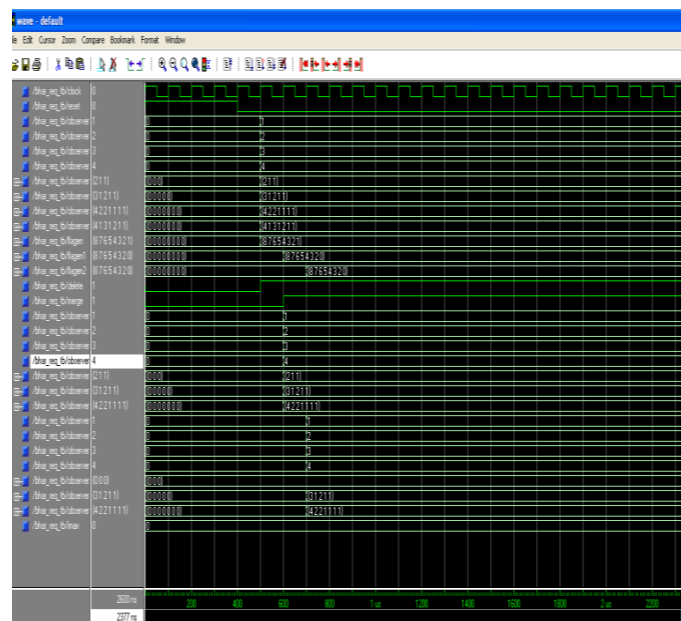


Fig. 10. Simulation result of the algorithm

V. CONCLUSION

This paper has proposed an algorithm for flip-flop replacement for power reduction in digital integrated circuit design. The procedure of flip-flop replacements is depending on the combination table, which records the relationships among the flip-flop types. The concept of pseudo type is introduced to help to enumerate all possible combinations in the combination table. By the guidelines of replacements from the combination table, the impossible combinations of flip-flops will not be considered that decreases execution time. The experimental results show that our algorithm can achieve

a balance between power reduction and wirelength reduction. Besides power reduction, the objective of minimizing the total wirelength is also considered.

References

- [1] L.-T. Wang, Y.-W. Chang, and K.-T. Cheng, Eds., *Electronic Design Automation: Synthesis, Verification, and Test*. Burlington, MA: Elsevier/ Morgan Kaufmann, 2009.
- [2] D. Duarte, V. Narayanan, and M. J. Irwin, "Impact of technology scaling in the clock power," in *Proc. IEEE VLSI Comput. Soc. Annu. Symp.*, Pittsburgh, PA, Apr. 2002, pp. 52–57.
- [3] P. Gronowski, W. J. Bowhill, R. P. Preston, M. K. Gowan, and R. L. Allmon, "High-performance microprocessor design," *IEEE J. Solid-State Circuits*, vol. 33, no. 5, pp. 676–686, May 1998.
- [4] L. Chen, A. Hung, H.-M. Chen, E. Y.-W. Tsai, S.-H. Chen, M.-H. Ku, and C.-C. Chen, "Using multi-bit flip-flop for clock power saving by DesignCompiler," in *Proc. Synopsys User Group (SNUG)*, 2010 [Online]. Available: <http://www.synopsys.com.cn/information/snug/2010/using-multi-bit-flip-flop-for-clock-power-saving-by-designcompiler>.
- [5] W. Hou, D. Liu, and P.-H. Ho, "Automatic register banking for lowpower clock trees," in *Proc. ISQED*, 2009, pp. 647–652.
- [6] Y.-T. Chang, C.-C. Hsu, P.-H. Lin, Y.-W. Tsai, and S.-F. Chen, "Post-placement power optimization with multi-bit flip-flops," in *Proc. EEE/ACM Comput.-Aided Design Int. Conf.*, San Jose, CA, Nov. 2010, pp. 218–223.
- [7] J.-T. Yan and Z.-W. Chen, "Construction of constrained multi-bit flipflops for clock power reduction," in *Proc. ICGCS*, 2010, pp. 675–678.
- [8] Y.-T. Chang, C.-C. Hsu, M. P.-H. Lin, Y.-W. Tsai, and S.-F. Chen, "Postplacement power optimization with multi-bit flip-flops," in *Proc. ICCAD*, 2010, pp. 218–223.
- [9] S.-H. Wang, Y.-Y. Liang, T.-Y. Kuo, and W.-K. Mak, "Power-driven flip-flop merging and relocation," in *Proc. ISPD*, 2011, pp. 107–114.