

Experiences of Low Power Design Implementation and Verification

Shi-Hao Chen

Global Unichip Corp.
Hsin-Chu Science Park,
Hsin-Chu, Taiwan 300
+886-3-564-6600
hockchen@globalunichip.com

Jiing-Yuan Lin

Global Unichip Corp.
Hsin-Chu Science Park,
Hsin-Chu, Taiwan 300
+886-3-564-6600
jylin@globalunichip.com

Abstract – In this paper, we present the experiences of some low power solutions that have been successfully implemented in 90nm/65nm production tape-outs. We also focus on power gating design, an effective low leakage solution, and present the experiences of power switch planning, optimization, and verification. Dynamic IR drop is an important issue in low power design, which may reduce the logic gate noise margins and result in functional or timing failures. We will present a low cost but effective methodology for dynamic IR drop prevention and fixing.

I. Introduction

As processes shrink to nanometer technology nodes, both of dynamic and static leakage power have become a significant design issue. To resolve the issue, power optimization could be achieved through many efforts contributed from various fields together from software policy, OS, system architecture, logical design, physical implementation, IP/library support and process technology, and so on, as shown in Fig. 1. In general, starting from higher level or earlier stage to set up the power saving plan will have more opportunity to save more power [1].

Dynamic power is generally proportioned to frequency, switching activity, capacitance, and square of supply voltage. The most effective way to reduce dynamic power is to reduce supply voltage because of the quadratic dependence of power

on voltage. Over recent years, some techniques have been developed to take the advantage of lower voltage for power reduction. These techniques include power shut off (PSO), multiple supply voltage (MSV), dynamic voltage scaling (DVS), adaptive voltage scaling (AVS), and so on. Reducing frequency and switching activity can also benefit dynamic power saving significantly. A module may slow down the frequency for some time while it is not required to operate at higher performance. This technique is usually employed together with voltage scaling to optimize the tradeoff between frequency and power by varying voltage. For switching activity reduction, RTL clock gating and architecture clock gating are widely used to restrict the distribution of clocks to those portions that are actually inactive at that time. The total capacitance reduction could come from the process shrinking and well physical implementation such as gate sizing and wire reduction.

Leakage power is increased dramatically while the device feature size shrinks. Designers have to put many efforts on leakage power reduction for extending the stand-by time. Several techniques such as multi-Vt optimization, back biasing, power gating, etc., have been developed for leakage power reduction for several years. Power gating becomes very popular in recent years [2]-[5]. It tends to turn off blocks that are not being used through voltage regulators or power switch cells. As shown in Fig. 2, power gating can be implemented through off-chip control or on-chip control. The off-chip power gating turns off the power sources supplied to specific power domains of a chip by a voltage regulator on board. This approach is suitable for long-term power shut-off because it may take long time to restore the power to the gated blocks. The on-chip power gating can turn off the power sources through switch-able power pads or power switch cells. Turning off a power source by switch-able power pads is quite simple, but it needs extra IO space allocated for power switch-able pads, and is not suitable for pad limited design. In addition, it is inflexible to control power-up ramp up time and rush current of the turned off blocks. For the power gating through power switch cells, MTCMOS is a good solution and is widely used recently, where sleep transistors, usually being high Vt transistors, are controlled by a power management unit to switch off the powers supplies to the gated blocks [2]. When the gated blocks are turned-off in standby mode, only a few leakage power consume due to the blocks being gated by high Vt sleep transistors. The implementation and verification

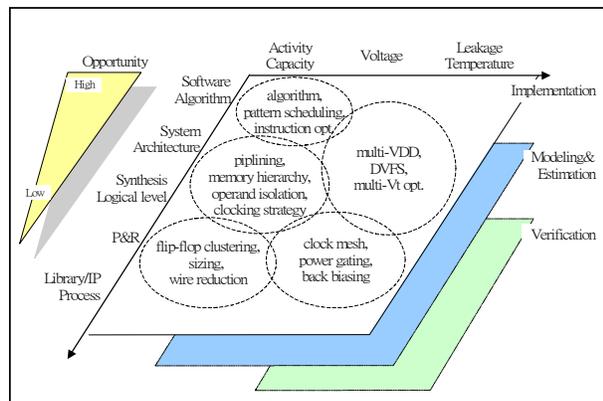


Fig. 1. Power management

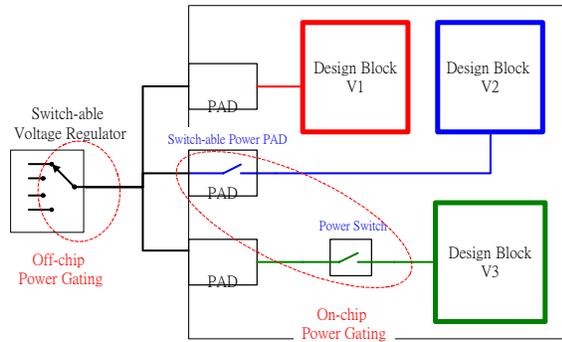


Fig. 2. Power gating methodology

of MTCMOS power gating are more complicated than those of power gating by switch-able power pads. Several effects such as power-up sequence, rush current, ramp-up time, dynamic IR, etc., should be taken into account for analysis [3].

The verification of low power design is a big challenge to success. For example, PSO and MSV may fail if there are structural errors such as missing isolation cell or level shifter, incorrect propagation of sleep control, incorrect power domain connection, and so on. Comprehensive low power verification should include the decision making and design quality check prior to the logic implementation. In addition, it should also include electrical implementation check, power aware formal verification, functional correctness of sleep control, timing closure among multiple corners and multiple modes, IR and EM analysis, etc.

The verification of dynamic IR drop becomes increasingly important for designs at 90nm and below, because simultaneous switching currents may induce peak IR drop in power/ground network and lower the voltage supplied to logic gates, which may reduce logic gate noise margins and result in function or timing failures. Dynamic IR drop is mainly dependent on the peak current of signal switching. The best approach to do the analysis is to simulate the design with peak power simulation patterns and then identify the hot spots. However, it is impractical to get the peak power simulation patterns at early stage due to long time preparation and verification. There is a vector-less heuristic approach proposed to solve this issue, but experimental results show that the approach cannot identify the hot spots correctly, and still need to improve the accuracy. In addition to verification, how to do dynamic IR drop prevention and fixing is another important topic to be resolved.

The remainder of this paper is organized as follows. In Section II, we present some low power solutions that have been implemented in GUC design flow for production. In Section III, we focus on on-chip power gating MTCMOS technique, and present our solutions from power switch planning, optimization, and verification for ramp-up time and rush current tradeoff. In Section IV, a heuristic approach for dynamic IR analysis is proposed. A dynamic IR prevention and fixing flow is also presented in this section. Finally, we

present the conclusion in Section V.

II. Low Power Chip Implementation

Whenever the industry moves from one technology node to another advance one, the battery-powered portable devices are driving the demand for more dynamic and static power reduction. New low power methodologies have to be developed to lower power consumption and also resolve the side effects such as dynamic IR. Low power design becomes more challengeable due to the extremely large and complex designs and increasingly complexity of design methodologies and verification.

In the following, we will present some low power chip implementation techniques that have been used in our production flow.

A. Clock Gating

Clock gating can save dynamic power for both of the registers being gated and the clock network between clock gated cell and the registers. As usual, a block needs clocking only when it requires to active, and the clock could be gated off in stand-by mode when the scheduled tasks are completed. From chip designer point of view, the challenge is to identify suitable control signal for the clock gating and have a good clock management. For chip implementation, engineers should use the clock buffers with proper driving strength to avoid unnecessary power consumption of clock network. Sometimes, they have to trade timing margin for power reduction. For example, de-cloning gated cells helps to reduce dynamic power, but cloning gated cells helps to reduce latency and skew.

B. Clock Mesh

Clock mesh (CM) is another clocking scheme in addition to clock tree synthesis (CTS). It is usually used for high speed clock design due to low clock skew, but may have low power benefits if clock skew and hold time fixing are well controlled. Experimental results show that, after improving clock skew, CM may gain 100ps to 150ps timing margins with compared to CTS and get speed, power or area improvement by exploiting the timing margins. As an example of ARM1136 with 500MHZ, the CM design achieved around 80ps clock skew, as shown in Table I. With compared to CTS, there are around 100ps timing margin which was exploited for about 10% area reduction and 15% power reduction. Basically, lower clock skew could get better hold time fixing. In the experiment, the number of hold time violations was reduced from 6K to 1K in functional mode. Thus, the number of buffers used for hold time fixing was reduced accordingly for saving power. CM is usually implemented in single and large-scale power/clock domain, and is not suggested for multiple power domains due to high complexity. This approach could be helpful for dynamic voltage scaling (DVS) approach due to the less hold timing violations.

TABLE I
Power Consumption Comparison between CTS and CM

	CTS	CM	Difference
Clock Skew	180ps	80ps	
Function Mode	783mW	665mW	-15%
Clocked Cell Counts	1000	4000	
ICG Cell Counts	2000	1000	
Clocked Power	240mW	200mW	
Instance Counts	208k	184k	-11%
Hold Time Violations	6k	1k	-83%

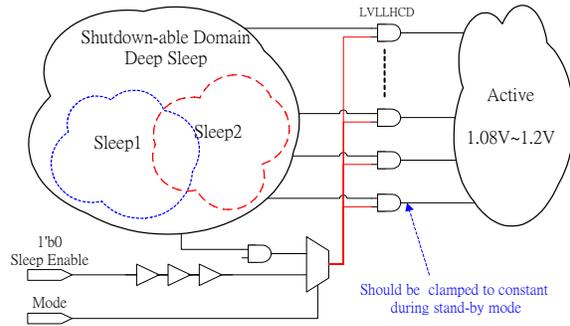


Fig. 3. Various depths of sleep modes

C. Multi-Voltage and Voltage Scaling

As mentioned earlier, reducing supply voltage is the most effective way for dynamic power reduction. By considering power consumption and performance tradeoff, one may apply different supply voltages to different blocks of a chip based on their performance requirements. The different supply voltages can be fixed at all, named multi-supply voltage (MSV) or be changed dynamically, named dynamic voltage scaling (DVS). A traditional DVS application such as ARM-IEM application may have the capability to monitor CPU workload from the command execution queue and adjust the voltage and frequency dynamically according to a predefined voltage/frequency combination table. The close loop solution most likely acknowledge as adaptive voltage scaling (AVS), which will dynamically monitor the current performance and then provide power management information to finely adjust supply voltage. Of course, the required performance of every next task should be predicted by software.

Level shifters are needed to pass the data signals from one voltage domain to another in MSV and DVS. As mentioned before, the complexities of implementation and verification are increasing, for instance, level shifter arrangement between different voltage domains, domain-aware buffering, routing, and equivalence check, power routing for level shifters, timing closure of multiple corners and multiple modes, etc.

D. Multi-Vt Libraries Optimization

From 0.13um and below technologies, library vendors started to offer multi-Vt libraries, including high Vt, normal Vt, and low Vt libraries, for leakage power and timing optimization. In general, increasing the threshold voltage (Vt) of a device will reduce the sub-threshold current effectively, but in contrast it may degrade the performance [2]. Therefore, for those non-critical timing paths, high Vt cells are exploited to lower leakage power. But, for those critical paths, low Vt cells may replace some high Vt or normal Vt cells to reduce cell delays and meet the performance target.

E. Power Gating with Various Suspend Modes

Power gating is the most effective method to manage leakage power recently, which tends to shut off the inactive blocks of a design by external voltage regulators or on-chip sleep transistors. However, power gating may incur power by

sleep transistors and have area penalty resulted from sleep transistors and extra decoupling cells [4]. It also has performance degradation issue due to IR drop across the sleep transistors, i.e., the IR drop at virtual VDD, and lower voltage supplied to the gates in gated blocks. Therefore, if the idle time of gated blocks is too short and sleep transistors have to be switched frequently, it is necessary to evaluate the benefits and penalties to make good decision prior to starting design.

There are several implementation and verification challenges of power gating designs. These include design of power gating controller, isolation cell insertion, sleep control propagation, retention register placement and routing, power domain-aware CTS, buffer insertion and routing, power-aware equivalence check, power routing verification, and timing analysis of multiple corners and multiple modes, and so on. In our design flow, we developed an in-house tool named UPDC, and collaborated with some commercial EDA tools to resolve above concerns. UPDC can check the correctness of sleep control propagation, the structure errors due to isolation cell missing, and especially the completeness of sleep mode control at various depths [5]. As the example shown in Fig. 3, there are three different sleep modes, sleep-1, sleep-2, and deep sleep mode. For different sleep modes, the clusters of logic gates to be turned off are also different, which makes the difficulties of sleep mode validation, power routing, isolation cell verification, etc.

III. MTCMOS Technology

MTCMOS power switch cells are widely used for on-chip power gating. In general, there are two approaches of MTCMOS power gating control: fine grain power gating and coarse grain power gating. In fine grain power gating, the sleep transistors are placed inside each standard cell, and thus have considerable area penalty. In contrast, coarse grain power gating shares sleep transistors to all of logic gates in the gated block and has very few area overhead. However, implementing coarse grain MTCMOS presents certain challenges to the design flow. These include power switch planning, power switch optimization, and power switch verification, which are to be discussed in the following.

A. Power Switch Planning

Achieving a power switch configuration is an optimization process subject to performance, area, power, ramp-up time, and peak current, as illustrated in Fig. 4. The number of switch cells and turn-on sequence of a design should be bounded by the constraints of IR drop and EM requirement, ramp-up time limitation and acceptable peak rush current. Basically, shorter ramp-up time may have higher rush current, and vice versa, and different power-up sequence may have different ramp-up time. As shown in Fig. 4, the worst case power-up sequence is one-by-one turn on, which takes the longest time to turn on all switches, but the peak current is the smallest. In the following, we will talk about the number of power switches, the placement of power switches, and the power up sequence.

- **Switch transistor modeling:** During power-up period, the switch transistors (sleep transistors) may stay at saturation region, and act as a non-linear voltage-dependent resistance. After the virtual rail is charged to the normal operating voltage, the switch transistors may change to the linear region. When we plan the number of switches during prototyping stage, the simple evaluation of current density can be applied first by using the resistance in linear region to determine the rough number of switches.
- **Switch transistor partitioning:** Power switches can be arranged as ring type or column type. Ring type is to place the power switches around the gated block, and column type is to place the power switches inside the gated block. Given the ramp-up time constraint, one should determine the maximum depth of the power switches either using ring type or column type. Then, all of the power switches should be partitioned and clustered into several banks to satisfy the maximum depth constraint.
- **Switch transistor assembly:** After confirming the power switch partition, the next step is to assemble the power switches in each bank and make sure the power-up

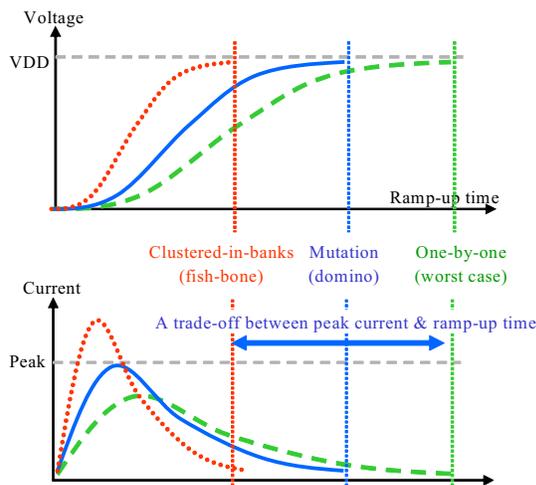


Fig. 4. Power switch planning

sequence can meet the requirements of ramp-up time, rush current, dynamic IR and EM, etc. In the worst case, if all of power switches are enabled simultaneously, there may have a surging current and have unacceptable peak current and IR drop. In practice, sleep transistors are not turned off or turned on simultaneously to reduce the associated large transient current and voltage drop noise. The power-on sequence can be formulated as a function subject to multiple objects: power consumption, power source location, the root position of sleep control signal, core cell IR degradation and etc. Power switches should be clustered, weighted and assembled with a heuristic, mostly like a two dimension weighting table, to minimize the dynamic IR during placement stage.

- **Switch transistor verification:** As mentioned above, power switch partitioning and assembly should take the impact of IR drop, EM, and peak current into account for evaluation. Therefore, there may have back-and-forth analysis among partitioning, assembly, and verification. Currently, there already have commercial tools to support these analyses. In addition, it also needs to monitor the IR drop of virtual VDD. If the IR drop is too big, there may have serious performance degradation, and extra timing margins to compensate the IR drop may be needed for sign-off.

B. Power Switch Optimization

After power switch assembly and related verifications, power switch optimization can be applied to refine the number of power switches and the size of power switches to satisfy the constraints. There are three major ways to optimize power switch configuration. The first one is to size-up or size-down switch cells incrementally. The second one is to remove some redundant switch cells to save area and power. The third one is power-up sequence reordering. All of these optimizations should be base on the analyses of area, power, ramp-up time, rush current and dynamic IR drop, etc.

Different power up sequence may result in different dynamic IR drops. As an example shown in Fig. 5, the power

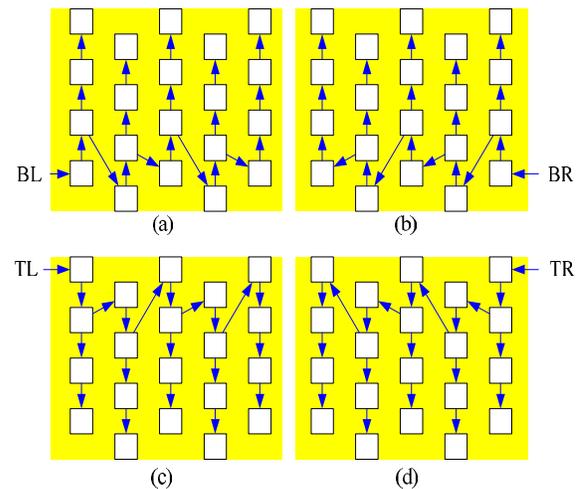


Fig. 5. Power switch configuration in domino fashion

TABLE II
Switch Configuration vs. Dynamic IR

Config/peak-IR (%)	VDD	VSS	Instance
BL	7%	6.5%	13.0%
BR	6%	5.5%	11.5%
TL	5.5%	5%	10.5%
TR	5.5%	5.5%	11.0%

switch partitions are the same for Fig. 5(a) to Fig. 5(d), but their assembly and power-up sequences are different. There are 2400 switch cells partitioned in 18 banks, and they are assembled in domino fashion. In the experiment, the ramp-up time requirement is 50 nano-seconds. In Fig. 5(a), the root position of sleep enable signal is at bottom-left corner, labeled as BL. Similarly, BR, TL and TR in Fig. 5 (b) to 5 (d) mean the root positions of the sleep enable signal being at bottom-right, top-left, and top-right corners, respectively. Table II indicates the dynamic IR drop effects of the power switch configurations shown in Fig. 5. From the results, we found the power-up sequence started at TL corner has smaller dynamic IR drop in whole chip including always-on domain. Therefore, the evaluation of the power-up sequence is very important.

IV. Dynamic IR Aware Low Power Flow

Dynamic IR drop occurs when the simultaneous switching of on-chip components causes a big current on the power grid, which may reduce the logic gate noise margins and result in function or timing failures. So, it does make sense to perform dynamic IR prevention as early as possible to reserve required space for decoupling capacitance insertion. However, as

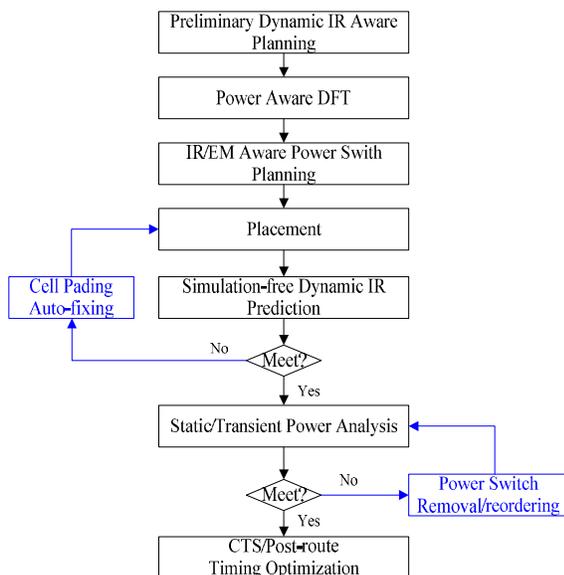


Fig. 6. Dynamic IR aware low power flow

mentioned earlier, it is impractical to do peak power pattern simulation (VCD-based analysis) in early design stage for checking dynamic IR hot spots due to the unavailability of simulation patterns and post-layout loading. Usually, dynamic IR checks are done in post-layout stage. But, it is often too late to find the problems, and needs to back to placement stage to resolve the problems by performing cell replacement or inserting decoupling capacitance.

In this session, we present some proven methods to resolve dynamic IR problems and provide a prevention flow to effectively shorten the lengthy loop in terms of physical implementation and verification. As the flow shown in Fig. 6, we propose a simulation-free dynamic IR prediction methodology to highlight the potential dynamic IR hot spots. Because this approach is based on cell placement information to do the prediction, it is easy to back to cell placement stage for hot spot fixing by cell padding and de-coupling cell insertion. As mentioned in Section III, power switch optimization needs to take dynamic IR drop into account, which has been integrated in this flow.

A. Preliminary Power Planning

Power planning must be part of early implementation phase including power calculation, power grid design, power grid analysis and refinement. It is obvious that technology trend towards more metal layers and flip chip design in recent years. Flip-chip design can reduce both of static and dynamic IR drops significantly, but the cost is much higher than wire bonding package. For wire binding, there may have an AP (Aluminum Pad) layer above top metal, which is usually used for RDL (re-distribution layer) routing and can be used as an additional power routing also. As the experimental results shown in Table III, both of static and dynamic IR-drop can be reduced dramatically by using AP layer.

Well planning of decoupling capacitor insertion can also prevent dynamic IR drop effectively. One can insert decoupling cells in the high switching activity area such as the outputs of clock buffer and flip-flops, and the area under macro power rings, etc.

B. Power Aware DFT

From our experiences, dynamic IR drop may result in serious yield loss in DFT mode other than function mode [6]. It is necessary to pay more attention to dynamic IR reduction in DFT mode. There are some proven methodologies such as macro clock gating in scan mode, functional clock gating in memory BIST mode, multi-power domain BSD (Boundary Scan Design), and location based memory grouping. In

TABLE III
Maximum IR-drop (VDD+VSS)

	Without AP		With AP	
	Static	Dynamic	Static	Dynamic
Case A	4.36%	24.25%	1.15%	10.36%
Case B	-	13.07%	-	6.99%
Case C	5.18%	-	1.86%	-
Case D	-	23.75%	-	7.17%

practice, sequential elements are clustered in banks to control test cycle and minimize the signal crossing between different power domains. Besides, it is better to create multiple test clocks with different shifted phase delays to avoid simultaneous switching of test clock and reduce scan mode dynamic IR.

C. Dynamic IR Drop Failure Prevention

- Prediction: The hot spots of dynamic IR drop usually happen at the area with high simultaneous switching currents. Clock buffers and flip-flops may have high potential to make big simultaneous switching currents if too many of them are placed in a local area. Our methodology is to perform flip-flop density check after cell placement [7] and predict the hot spot area where the flip-flop density is over a certain percentage. This approach is very simple but effective in contrast to the expensive VCD-based dynamic IR analysis. In addition, this approach has been correlated with various VCD-based simulation results and silicon data. Fig. 7(a) illustrates the dynamic IR analysis result based on peak power pattern simulation (VCD-based analysis), where the serious dynamic IR hot spots are highlighted in red. Fig. 7(b) shows the prediction of dynamic IR hot spots using our flip-flop density check. The results show that our approach is highly correlated to the VCD-based analysis. But, our approach takes only few minutes in contrast to couple of hours to even days in VCD-based analysis.

- Fixing: Currently, we will be back to cell placement stage and resolve those high flip-flop density regions by applying cell padding. After performing flip-flop density check, we can grab the cell types of flip-flops within the violation windows, and specify proper placement padding on them. The placement padding is to make sure those flip-flops can be kept away in a certain distance during timing optimization and clock tree synthesis. Generally, this approach can achieve a reasonable quality for dynamic-IR reduction without significant timing degradation. Padding is a soft constraint and should be brought into clock tree synthesis stage. In our experiment, an ECO script will be generated automatically after performing the window-based analysis. Of course, one can also specify a proper padding for entire flip-flops, because it is just a soft constraint.

V. Conclusion

In this paper, we have presented some effective low power techniques such as clock gating, clock mesh, MSV and DVS, multi-Vth optimization, and power gating, etc., that have been widely used in our products. We also addressed the challenges of on-chip power gating and presented the solutions of MTCMOS power switch planning and optimization. For dynamic IR drop reduction, we presented the significant improvement of using AP layer and planning in DFT mode. We also presented a flip-flop density check methodology that is highly correlated to VCD-based dynamic IR analysis, but is much faster. The methodology can predict dynamic IR hot spots in cell placement stage, which makes the dynamic IR

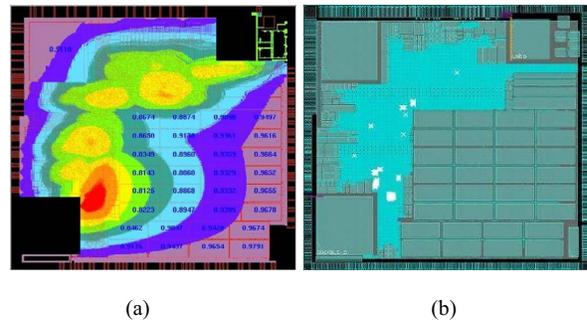


Fig. 7. Dynamic IR failure prediction

fixing much easy in a short iteration.

References

- [1] J. Frenkil, "Tools and methodologies for low power design", *Proc. Design Automation Conference*, Jun. 1997, pp. 76-81.
- [2] M. Anis, M. Elmasry, *Multi-Threshold CMOS Digital Circuits - Managing Leakage Power*, Kluwer Academic Publishers, 2003
- [3] K. Shi, D. Howard, "Sleep transistor design and implementation - simple concepts yet challenges to be optimum", *International Symposium on VLSI Design Automation and Test*, April 2006, pp. 1-4
- [4] H. Jiang, M. Marek-Sadowska, "Benefits and costs of power-gating technique", *Proc. IEEE International Conference on Computer Design: VLSI in Computers and Processors*, Oct. 2005, pp. 559-566.
- [5] K. Agarwal, H.S. Deogun, D. Sylvester, and K. Nowka, "Power gating with multiple sleep modes", *ACM/IEEE International Symposium on Quality Electronic Design*, pp. 633-637, 2006.
- [6] S.H. Chen, K.C. Chu, J.Y. Lin, "DFM/DFY practices during physical designs for timing, signal integrity, and power", *Proc. IEEE Asia and South Pacific Design Automation Conference*, 2007, pp. 232-237
- [7] A.B. Kahng, "Physical chip implementation methodology: hot spots and best practices", *Design Automation Conference*, Tutorial #5, June 14, 2002