International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-9S3, July 2019

Accounting for PVT Variations in Design Timing Closure

Vasanth Mundargi, Siva Yellampalli

Abstract*: The On Chip Variation (OCV) refers to changes in the behavior of parameters like process, voltage and temperatures on a chip. In this paper, we go through different approaches followed to compensate for PVT variations on chip during design timing closure. We review the dominant approaches used for accounting such variations. We also review the advantages and disadvantages of these approaches used based on the ease of use, implementation, power, area, and the overheads involved in adopting them.

***Index Terms*: Delays, PVT variation, Static Timing Analysis, On-Chip Variation, Derating, AOCV, LOCV, GBA, PBA, IR Drop, Voltage Aware STA**

I. INTRODUCTION

Static Timing Analysis is one of the most widely used amongst many techniques used for timing analysis of digital design. Timing simulations is an alternative approach used. Simulation helps us to do timing analysis along with the analyzing the functionality of the design. STA is static since the analysis is independent of input data values and its main purpose is to validate if a design can meet the timing specifications with a given set of definitions for clock and external environment of the design. Few examples of timing checks that are done in STA are setup and hold checks. The most important aspect of STA is complete design is analyzed at once. The required timing checks are also performed for all possible paths. This makes STA an exhaustive and reliable method for timing verification of a digital design.

For digital circuit design power, timing and area are very important parameters. Designs need to operate within the specified power limit and must be restricted to the area budget of the board/system. Static timing analysis is typically performed at a specific Voltage, Process and Temperature (PVT). Delays of the interconnects and the standard cells are computed based on the PVT condition specified.

Process is a representation of the process models provided by foundries. Usually categorized as slow, typical and fast process models. Ideally, we would want the operating conditions to stay uniform across the die/die-lot, but in reality, it isn't uniform. These are further classified as intra-die and inter-die variations [2]. Besides the variations in the process parameters, different parts on a chip may also see different effective supply voltage. This essentially means two identical transistors or metal interconnects may behave differently on the same die.

The variations can arise due to many factors of which few of them are [1]:

- Power supply variation due to IR drop along the chip which will affect the local power supply.
- Threshold voltage, device dimension and doping concentration variation of the devices during fabrication.
- Spot temperature variations caused by hot spots.
- Interconnect metal thickness variations during fabrication impacting the interconnecting capacitance or resistance [1].

While these effects can be considered negligible at higher technologies, they start to play a very important role at lower deep submicron technologies [4]. So, considering the PVT variations or On-Chip variations in STA becomes very important. For reliability the design must be tested at the extremes of process, temperatures and power supply.

Many approaches have been proposed and are being used to account for the above-mentioned variations. We go through few of the approaches, discussing the implementation, benefits and drawbacks. The rest of the paper is organized as: Section 2 describes the effects that occur due to PVT variations. Section 3, 4, 5 explains different approaches based on method, their advantages, and disadvantages followed by conclusions and future work in section 6.

II. EFFECTS OF PVT VARIATIONS

A. Process Variation

The problem of intra-die variations is amplified with the scaling of technology. The intra-die channel length variation is estimated to deteriorate from 35% of the total variation for 0.13um to approximately 60% in 70nm. The intra-die routing variation will raise by approximately 35% in 70 nm [2].

Process variation are of the following type

- Random
- Systematic
- Intra-die (within die or chip)
- Inter-die (die-to-die or lot-to-lot) [3]

Fabrication of the devices has immense effect on the process variation observed in the design. Random process variations tend to have no spatial or layout correlation which may result in different transistor characteristics even for neighboring devices [4].

Systematic variations involve variations in exposure pattern variation during lithography process. Silicon surface flatness variations is caused by layout pattern density during the Chemical Mechanical Planarization. Differences in the process parameters may result in variation of sheet resistance, threshold voltage in turn impact the performance of the device.

Revised Manuscript Received on July 22, 2019.

Vasanth Mundargi, post graduate student, UTL Technologies Ltd, VTU Extension Center, Bengaluru 560022,

Dr. Siva Yellampalli, Principal, UTL Technologies Ltd, VTU Extension Center, Bengaluru 560022.

Table 1. Process variation effect on path delay [3]

Process/Component	Leff	tox	Vth	Path Delay
Fast	↓	↓	↓	Minimum
Typical	-	-	-	
Slow	↑	↑	↑	Maximum

B. Voltage variation

Saturation current of the device is given by the equation:

$$I_{DS} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_T)^2(1+\lambda V_{DS})$$

Fig 1 Saturation current equation for MOSFET

Cell delay is dependent on the saturation current. In this way, the power supply varies the propagation delay of a cell. So as VDD decreases, driving capability of cell decreases and results in large delays and higher power supply makes a cell faster and hence reduced cell delays [3] [5].

C. Temperature variation

Due to power dissipation in MOS transistors the operating temperature in the chip varies according to the switching activity of the circuit. The dissipated power will increase the surrounding temperature. Temperature variation affects speed, power dissipation and reliability of the circuit by altering threshold voltage, and saturation velocity of the MOSFET. MOSFET mobility, threshold voltage, and saturation velocity are related to temperature according to the following expressions:

$$\mu(T) = \mu_0(T/T_0)^{\alpha_\mu}$$
$$V_{th}(T) = V_{th0}[1 + \alpha_{V_{th}}(T - T_0)]$$
$$v_{sat}(T) = v_{sat0}[1 + \alpha_{v_{sat}}(T - T_0)]$$

Parameter Symbol	Parameter	Comment
T ; T_0	Temperature ; Normal Temperature (25 ^{0}C)	
μ; V_{th}; v_{sat}	Mobility, Threshold voltage and Saturation velocity at temperature T	
μ_0 ; V_{th0}; v_{sat0};	Mobility, Threshold voltage and Saturation velocity at temperature T_0	
α_μ; $\alpha_{V_{th}}$; $\alpha_{v_{sato}}$	Mobility temperature coefficient (-1.3); Threshold Temperature coefficient (-3 mV/ ^{0}C); Velocity temperature coefficient (- 97 m/(s ^{0}C))	Negative Temperature coefficient

Fig 2. MOSFET mobility, threshold voltage and saturation velocity as a function of temperature [5]

The mobility of electron and hole is temperature dependent. The mobility decreases with increased temperature. As the mobility of the charge carriers decrease the propagation delay increases. Hence, with an increased temperature the propagation delay increases.

The threshold voltage is also seen to be dependent of the temperature. An increase in temperature will decrease the threshold voltage. A lower threshold voltage results in a higher current and hence a better delay performance. There is race between these two effects and usually the mobility effect wins [3].

III. MARGINS

A. Specifying margin in uncertainty [3]

The uncertainty in the timing of the clock edge is to account for several factors such as clock period and additional margins used for timing verification. The additional margins are included to estimate timing of a design for the following scenarios:

- Modeling and analysis errors: Device models inaccuracy, or timing analysis algorithms approximations.
- PVT variations: Uncertainty in the parameters of a fabricated devices and interconnects from die-to-die or within the same die.
- Operating context variations: Uncertainty in the operating environment of a device in its lifetime, like temperature, voltage.

Consider the equations for setup and hold checks:

$$T_{launch} + T_{ck2q} + T_{dp} < T_{capture} + T_{cycle} - T_{setup}$$

$$T_{launch} + T_{ck2q} + T_{dp} > T_{capture} + T_{hold}$$

Fig 3. Equations for setup and hold checks [1].

When this extra margin is specified in the uncertainty the timing analysis tool will subtracts these numbers from the required time for setup check and add this extra margin for the hold check. This simple approach puts a blanket value of uncertainty to the timing checks. There are three important drawbacks with this approach.

- Picking up the uncertainty values either requires some data from the foundries or needs prior experience with the technology node/design/foundry [6].
- Depth of the path/logic doesn't play any role. Both shorter paths and longer paths get applied with the same pessimism.
- A large timing uncertainty will add immoderate pessimism into the timing budget which translates into more area and unwanted leakage power which is undesirable [1].

IV. DERATING METHODS

On chip variation (OCV) is accounted in the digital design during STA by changing/derating the delays of specific paths. The delays in the path is either reduced or increased and then the behavior of the design is validated with these variations. The cell delays in data and/or clock path, wire delays and the requirements from the libraries can be changed to model the effects of OCV. Derating can also be combined with the extra timing margin specified for the checks.

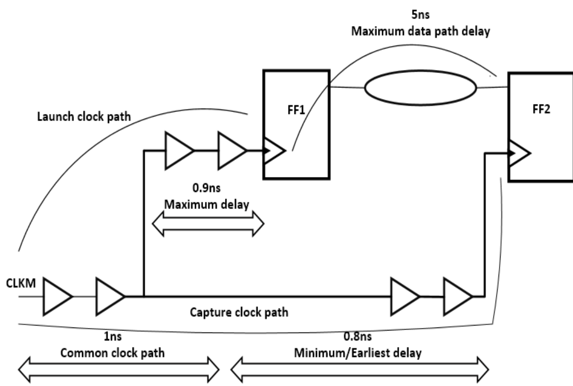

Fig.4 Timing path and delays

The value used for derating can be a fixed number or varying based on the approach followed. Derated can be applied to the launch clock paths, data paths, or capture clock paths (any or all the paths). For setup check, pessimism can be added by either delaying the launch path (data and/or clock), i.e., derating the path with a factor greater than 1 and/or by making the capture path faster, i.e., derating it with a factor lesser than 1. Similarly, derating can be applied to hold check by delaying capture and making the launch path faster. The discrepancy arises in the clock tree delay which has a common clock path. The path which is part of both launch and capture path and the same cell cannot behave differently with the same conditions at a given moment of time in operation. This pessimism is called Common Path Pessimism (CPP) or Clock Reconvergence Pessimism (CPR) which should be removed during analysis [1].

A. Single Derate OCV

For single derate OCV methodology, a single value is used to derate the paths for analysis. The derating factor is automatically obtained from the given spreadsheet, assuming the inputs are known.

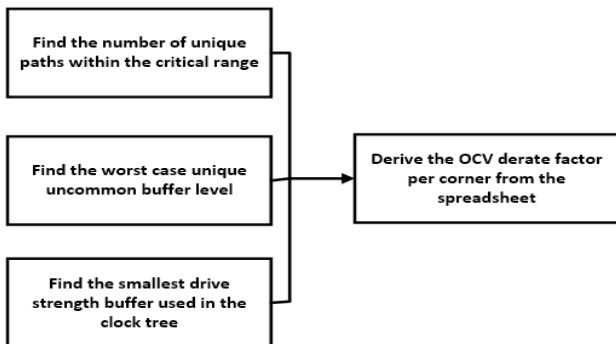

Fig.5 Single derate methodology [2]

All the inputs can be figured out from the STA with the post layout netlist through scripting or from the CTS report. The OCV spreadsheet is usually obtained from the foundry. An example of the OCV spreadsheet outline is shown:

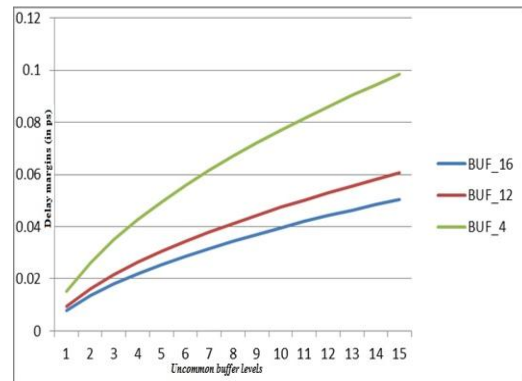

Fig.6 OCV Spreadsheet Outline [2].

For example, 40ps delay margin is required if the clock network has a smallest clock cell of drive strength 4 with uncommon buffer level of 4. The definition of uncommon path is shown below:

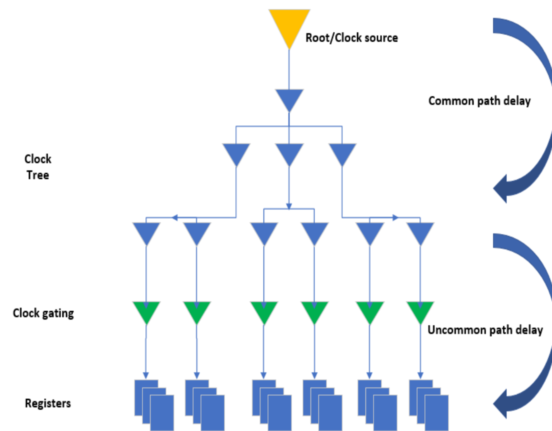

Fig.7 Uncommon Path definition. [2]

From the information above in fig 6. different combinations of derating factor can be derived and timing margin to meet the delay margin required can be obtained. Higher the timing margin used lower the derating factor that will be required.

The advantages this methodology brings in is that the foundry data of variation of the cell behavior is taken into account to calculate the derating factor which brings this closer to realistic scenario. Also, the derating is now dependent on the depth of the path. Longer the path, larger the pessimism getting added to account for variations.

The drawbacks that single derating brings are listed below:

- It is over-optimistic for timing paths with shorter timing depth and over pessimistic for paths with huge depth.
- The derating factor used varies with the design and this does not allow the reuse of the factor across designs for the same technology node.

B. Advanced OCV or Location based OCV [2]

Advanced or location based OCV is an upgraded version of constant derating approach. It helps us reduce the iterations and the number of fixes that go into the design. It is an optimal solution which falls somewhere in between OCV and SSTA (mentioned in the later section). The factor user for derating is not fixed in AOCV.

It uses pre-characterized derating tables for each cell in the form of a lookup table for both early and late analysis. These lookup tables are characterized with the information available from the foundry. The lookup tables come with the indices as the depth of the logic and the location of the cell and contents of the table are derates. The derate to be applied for delay calculation is now picked for each path independently based on the actual depth of the path and the distance calculated based on an actual bounding box that encloses the registers as seen in the figure below:

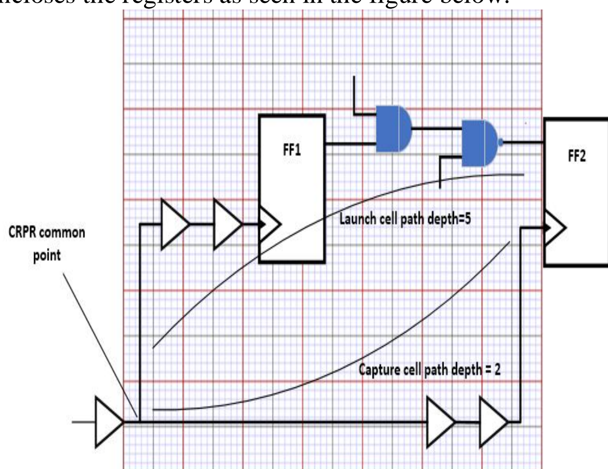

Fig.8 AOCV depth and distance definition. [2]

As show in the figure, the depth is the levels of logic cells in the path off the CRPR common point. Launch path and the capture paths use the depth values based on the actual connectivity in the design. These derate values are applied over any fixed margin that might be chosen as required for closure.

Margin values are further as compared to fixed derate and margin-based approaches, eliminates unwanted pessimism that would be been added. Since information from foundry about the variations seen on silicon is taken into account, this is closer to silicon behavior. These are some of the advantages this approach brings in.

Drawbacks for this approach are listed below:

- AOCV table characterization requires extra effort, resources and cost.
- With AOCV characterized data, the STA runtimes take a hit as there is greater details which are taken into account for delay calculation.

V. DELAY CALCULATION OR ANALYSIS METHODS

A. Graph based delay calculations.

Delay of a standard cell is a function of input transition and output load. In the figure shown below, we see that there can be multiple transitions reaching the pins of an instance.

If we categorize them as fast (best) slew and slow (worst slew), while performing the delay calculation the timing engine computes the delays using the worst-case slews for all the standard cells for all the inputs of the gate. Fast slew and the slow slews are worst case slews for early and late paths respectively. So, the fast slew is used for all the pins of the OR gate when the delay is calculated for a hold check and slow slew is used for delay calculation for setup check. If the actual slews are used for delay calculation from respective pins to output, it is termed as path-based analysis.

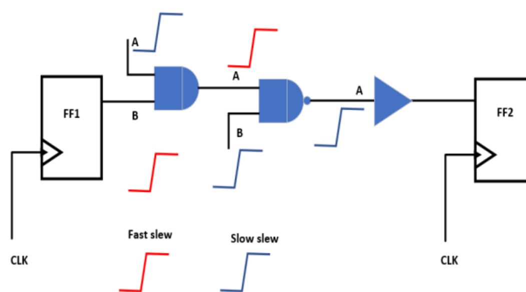

Fig.9 Graph based analysis and path-based analysis [8]

This method of timing computation is faster as the engine has to only report the delay values for each cell once. One of the major drawbacks this approach brings is that there is clear demarcation on the composition accounted for timing engine's computation approximations and PVTs variation.

B. Voltage aware delay calculations.

There has been significant reduction in the supply voltages because of technology scaling and low power requirements. At lower supply voltages, the ICs become highly sensitive to supply voltage fluctuations. The power is supplied to the standard cells through a power mesh or grid. The power supply to the cells can fluctuate because of the supply source or because of the IR drop on the power grid. The power supply related fluctuations are usually taken care of with proper analog design and simulations. IR drop is the voltage drop seen on the power grid before finally reaching the standard cell instances. There two possible reasons for IR drop, a weak power grid with high resistance, or high switching activity on the chip or in certain areas of the chip which results in drawing higher currents through the grid.

A voltage profile with IR drop is shown in the figure.

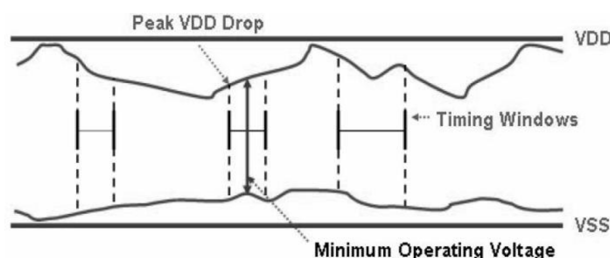

Fig.10. Voltage profile with IR drop. [9]

The delay, slew variations because of this effective voltage needs to be accounted. With either the global activity in percentage or an activity file dumped from a verification environment for a worst case switching test case the effective voltage seen on each standard cell can be extracted from power analysis tools. This effective voltage can be used as an input in STA to calculate more accurate delays.

There are some minor differences between the usual STA flow and the voltage aware STA flow. Voltage drop aware STA requires multiple libraries characterized with different operating voltages. Delay calculation engine computes an instance delay by interpolating delays between two libraries that are closer to the effective instance

voltage. For ex, if we use libraries characterized for 1.0V, 0.9V and 0.8V. Delay engine uses 1.0V and 0.9V for all the instances with effective voltage < 1.0V and 0.9V. The voltage range characterized should cover minimum and maximum IR drops. Along with multiple libraries the instance voltage mapping information is provided to the tool in a file.

As an added extension, minimum and maximum voltages for the same instance during a time period can be passed in for early and late analysis for pessimistic analysis. This approach can further used along with one or more approaches discussed above.

Accurate and reliable timing results for the worst-case IR drop or voltage variation can be obtained with this approach. The disadvantages this approach brings in are:

- Worst case scenario of the operating conditions usually covers the voltages within IR drop limit, hence this check can become redundant with some corner case exceptions.
- Requires extra resource in terms of multiple characterized libraries for the same process and temperature parameters, power analysis setup/results.
- Global activity information is usually approximate, which can add unwanted pessimism. While, the test case with worst IR drop might not be the worst scenario for timing.
- Runtimes are longer because of extensive calculation requirements.

VI. CONCLUSION & FUTURE SCOPE

In this paper, our focus is to review the dominant methods to model/account for PVT variations in STA. As an introduction, the reasoning for each of these variations have been discussed. A dedicated section describes the effects of PVT variations to delay and timing. In the following sections few methods are discussed which are used to model/account for such variations segmented based on the approach.

The vast majority of the industry uses a mixed approach which includes a small fixed margin and either fixed/variable derating with GBA. Voltage based STA hasn't gathered enough takers yet, to be classified as a signoff check or approach.

There are other approaches which are proposed, like Parametric OCV and Statistical STA. POCV uses a statistical approach, where it models local and global random variations for cells and nets. It calculates delay variation by modeling intrinsic cell delay and load parasitic to determine mean and sigma or variation of the logic stage. POCV was put forward and supported by Extreme DA, which is now acquired by Synopsys. So, there is no clear roadmap available for the usage of this approach. Statistical Static Timing Analysis receives a strong interest in academia for reduced pessimism, faster timing closure. While still hasn't been widely adopted in the industry for its nontrivial effort in understanding of the timing results and few other drawbacks.

REFERENCES

1. J. Bhasker and Rakesh Chadha, Static Timing Analysis for Nanometer Designs: A Practical Approach, New York: Springer, 2009.
2. Ang Boon Chang, "ASIC design margin methodology," TENCON Spring Conference, April 2013, **doi:** 10.1109/TENCONSpring.2013.6584434
3. http://www.idc-online.com/technical_references/pdfs/electrical_engineering/Process_Voltage_Temperature_PVT_Variations_and_Static_Timing_Analysis.pdf
4. Franco Stellari at al, "Mapping systematic and random process variations using light emission from Off-Stage Leakage", Reliability Physics Symposium, April 2009, doi: 10.1109/IRPS.2009.5173323
5. Vinod Kumar Singh, "Improvement in Static Timing Analysis for Early TA-Signoff Closure in SOC flow at Very Deep Submicron Nodes", M.Tech thesis, Dept of Elec. & Comm., IIIT, New Delhi, 2014 https://repository.iiitd.edu.in/jspui/bitstream/123456789/174/1/MT12110.pdf
6. Susumu Kobayashi and Kenichi Horiuchi, "An LOCV-based static timing analysis considering spatial correlations of power supply variations", Design, Automation & Test in Europe Conference & Exhibition, March 2011, doi : 10.1109/DATE.2011.5763283
7. Pua Siaw Fuang and Nor Muzlifah Mahyuddin, "Implementation study of path-based AOCV model on pessimism reduction for 20nm technology", ICCSCE 2014, Nov 2015, doi : 10.1109/ICCSCE.2014.7072693
8. Naman Gupta, "Timing Analysis: Graph Based v's Path Based", July 2014, http://vlsi-soc.blogspot.in/2014/07/timing-analysis-graph-based-vs-path.html
9. IR Drop Analysis, Cadence Design Systems, Inc., San Jose, USA

AUTHORS PROFILE

Vasanth Mundargi is a M. Tech student in VLSI Design and Embedded Systems at VTU Extension Centre, UTL Technologies Ltd. Bengaluru. He has obtained his B. E. from KLE Society's College of Engineering and Technology, VTU Belagavi.

Dr. Siva Yellampalli obtained his MS & Ph.D from Louisiana State University. He is currently with VTU Extension Centre, UTL Technologies Ltd. He worked on a broad range of research topics including Very Large Scale Integration (VLSI), mixed signal circuits/systems development, micro electromechanical systems (MEMS), and integrated carbon nanotube based sensors. He has published a book in the area of mixed signal design, and edited two books on carbon nano tubes. He also published multiple journal papers & IEEE Conference papers in these areas of research.