Activation-Aware Slack Assignment for Time-to-Failure Extension and Power Saving

Yutaka Masuda, *Student Member, IEEE*, Takao Onoye, *Senior Member, IEEE*, and Masanori Hashimoto, *Senior Member, IEEE*

***Abstract*— This paper proposes a mean time-to-failure (MTTF) aware design methodology for minimizing power dissipation while satisfying target chip lifetime. The key contributions of the proposed design methodology are to explicitly introduce MTTF as a design constraint and optimize the design with activation-aware slack assignment (ASA). Conventionally, the gates included in nonintrinsic critical paths are downscaled or replaced with high-Vth gates for power savings, where the nonintrinsic critical paths are timing paths which originally had large timing slacks before the downscaling and replacement. On the other hand, ASA gives timing slacks to nonintrinsic critical paths and reduces the number of active paths whose delays are very close to those of intrinsic critical paths whose timing slacks cannot be increased by resynthesis and sizing. The proposed optimization includes both pre-ASA circuit design and ASA implementation. The former pre-ASA design prepares several design candidates that laid out with different timing constraints and selects the most promising candidate regarding power. For this selection, every candidate is analyzed to estimate minimum supply voltage after ASA that can achieve the target MTTF. Then, the proposed methodology selects a set of flip-flops for ASA using integer linear programming, such that it reduces the sum of gatewise failure probability maximumly, and performs ASA. We evaluate MTTF of circuits with and without ASA and examine how much power saving can be obtained while satisfying the target MTTF, e.g., 10 years. Evaluation results show that the circuits with ASA achieve up to 49.6% power saving.**

***Index Terms*— Activation-aware slack assignment (ASA), integer linear programming (ILP), mean time to failure (MTTF), power saving, stochastic timing error rate estimation.**

I. Introduction

In A synchronous sequential circuit, a timing error occurs when the signal propagation time through the combinational circuit exceeds the clock cycle time. The signal propagation time varies depending on manufacturing variability, environmental fluctuation such as supply noise and temperature gradation, and aging effects. Manufacturing variability includes variations of a threshold voltage, gate length, oxide thickness, and so on which vary gate delay and wiring delay. Aggressive device miniaturization due to technology scaling has made the manufacturing variability more and more significant, and lower supply voltage makes circuits sensitive to supply noise and temperature variation. Furthermore, aging degrades performance, and one of the representative aging phenomena is negative-bias temperature instability (NBTI) [1], [2]. NBTI changes pMOS threshold voltage due to the gate oxide degradation induced by the negative bias voltage. Wang *et al.* [3] reported that circuit speed could degrade by about 8% after 10-year operation. An important point here is that the above-mentioned delay variations directly lead to circuit performance degradation and shorten the time to failure (TTF) of the chip.

For avoiding timing errors due to manufacturing variability, environmental fluctuation, and aging, a conventional worst case (WC) design adds design and operation margins in design time and field operation, respectively. However, as the performance degradation becomes significant, such margins tend to be too painful for designers. Timing closure becomes more and more challenging and time-consuming, and sometimes, it is infeasible. To set the necessary and sufficient margins while taking into account all the variation sources including aging effects, we need to estimate the chip TTF in design time.

Fig. 1 shows the TTF variation originating from the stochastic properties of manufacturing variability and aging process. Statistical characterization of manufacturing variability is studied comprehensively in the last decades, and its statistical modeling is now a common practice. Also, threshold voltage variation due to NBTI fluctuates statistically. Due to these statistical properties, the time when the circuit delay exceeds the clock cycle time, which corresponds to TTF, varies as shown in Fig. 1. A naive approach to calculate the TTF is to execute the gate-level simulation in the huge variational parameter space repeatedly. However, the probability that actual timing errors occur is quite low,[1] and hence, the simulation time required to reproduce these errors is prohibitively long. For example, when we evaluate the rate of timing errors that occur once per one month, the simulation time exceeds 10^8 years [4].

Recently, on the other hand, a stochastic framework that estimates mean TTF (MTTF) is proposed in [4] and [5]. Iizuka *et al.* [4] model circuit operation under dynamic delay variations as a continuous-time Markov process.

Manuscript received February 13, 2018; revised June 12, 2018; accepted July 18, 2018. Date of publication August 16, 2018; date of current version October 23, 2018. This work was supported by STARC and ICOM Foundation, Japan. *(Corresponding author: Yutaka Masuda.)*

The authors are with the Department of Information Systems Engineering, Graduate School of Information Science and Technology, Osaka University, Suita 565-0871, Japan (e-mail: masuda.yutaka@ist.osaka-u.ac.jp; hasimoto@ist.osaka-u.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2018.2862154

[1] Otherwise, such circuits with frequent error occurrence are useless.

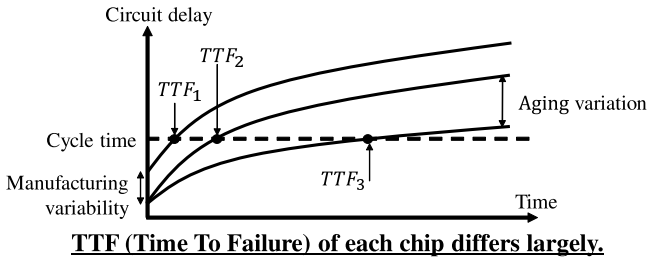

Fig. 1. TTF variation due to manufacturing variability and aging variation.

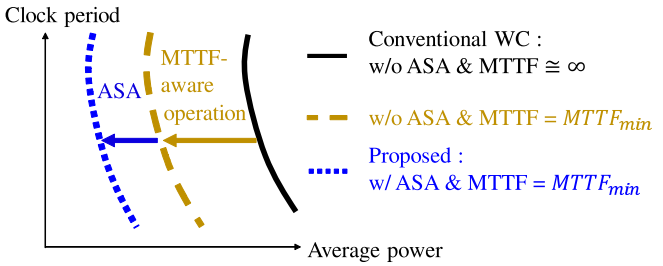

Fig. 2. Proposed MTTF-aware design reduces power dissipation due to MTTF-aware operation and design optimization with ASA.

The continuous-time Markov process modeling enables us to estimate the MTTF in a reasonable time. In a test case, MTTF is estimated 10^{12} times faster than a logic simulator. Reference [5] extended the framework proposed in [4] to consider manufacturing variability, temporal environmental fluctuation, and aging in the MTTF estimation. Also, [5] takes into account workload-dependent path activation probabilities. With this framework, it becomes possible to know, for example, the tradeoff between MTTF and supply voltage.

This paper proposes a design methodology that minimizes power while satisfying the given MTTF specification. The proposed methodology explicitly introduces MTTF as a design constraint and optimizes the design with activation-aware slack assignment (ASA). MTTF constraint helps explore a set of necessary operating conditions, such as clock period and supply voltage, and reduces the operation margin from a WC design while keeping the target MTTF. This margin reduction directly leads to supply voltage reduction. ASA, meanwhile, gives timing slacks to nonintrinsic active critical paths by engineering change order (ECO), where nonintrinsic critical paths are timing paths whose slacks were originally large but are reduced by downsizing and replacement to high-Vth cells for power savings. Thus, ASA reduces the number of active critical paths whose delays are very close to those of the intrinsic critical paths, i.e., timing paths whose slacks cannot be reduced by resynthesis, replacement to low-Vth cells, and sizing. In this case, we can expect that circuits with ASA have fewer paths where timing errors are likely to occur. This TTF extension can be converted to supply voltage reduction since the circuit can achieve the target MTTF at the lower supply voltage. Thus, ASA can further improve the performance from the MTTF-aware operation. In this paper, we assume that the supply voltage can be set for each chip individually to exploit the chip-dependent margin for power minimization. We hereafter call this situation as chipwise voltage assignment.

In this paper, we construct a design methodology for MTTF-aware ASA design. The design methodology needs to prepare a pre-ASA circuit, choose paths to which ASA is applied, and determine timing slacks for each path. In this paper, to save power maximally by ASA, we first propose a selection method of the pre-ASA circuit from several design candidates. Note that, with chipwise voltage assignment, the optimal design is not obvious since voltage scaling varies power and speed and the impact of ASA depends on the pre-ASA circuit. For each candidate, the proposed method estimates the minimum supply voltage after ASA ($V_{\min}$) at which the circuit can achieve the target MTTF, and evaluates the power dissipation of circuit at $V_{\min}$. Thus, we can choose the circuit whose estimated power is minimum. Second, we propose flip-flop (FF)-based ASA that assigns timing slack to each FF. We develop an FF selection method using integer linear programming (ILP) that maximizes the sum of gatewise failure probabilities aiming to improve MTTF maximally. Third, for each target FF, we extract necessary timing slacks to sustain the target MTTF at $V_{\min}$ and give these timing slacks as constraints to place and route (P&R) ECO.

The main contributions of this paper include: 1) design optimization with ASA that explicitly introduces MTTF as a design constraint and 2) quantitative evaluation of power saving effects for practically long MTTF. To the best of our knowledge, this is the first work that optimizes operating conditions and design under the constraint of MTTF in units of several years and demonstrates the impact of MTTF-aware design regarding power saving. Fig. 2 shows the expected power savings. The right black curve represents the conventional WC design that adds timing margins assuming the worst process, voltage, temperature, and aging (PVTA) condition. The middle yellow curve is also a WC design, but it optimizes the operating conditions such as supply voltage and clock period so that the design satisfies the target MTTF. The left blue curve corresponds to the proposed ASA with the MTTF-aware operation. The proposed ASA is expected to attain a better tradeoff between power dissipation and clock period. This paper will experimentally demonstrate these power saving effects in an embedded processor and a cipher circuit.

Preliminary results of ASA were reported in [6], where ASA was called critical path isolation. This paper establishes the design methodology by extending [6]. The main extensions are the following.

1) This paper includes the design optimization of pre-ASA circuit, whereas [6] assumes that the pre-ASA circuit is given and does not provide how to prepare the pre-ASA circuit. The power and area reduction thanks to the pre-ASA circuit selection will be shown in Section VI-C1.
2) This paper performs ASA as P&R ECO process, whereas [6] performs ASA in logic synthesis. Therefore, the previous work did not consider the impact of ASA on a physical layout design.

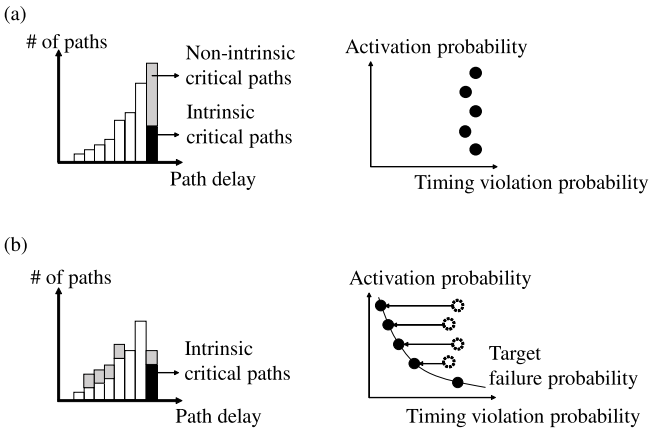

Fig. 3. Path delay distributions (left) and the activation probability and timing violation probability of nonintrinsic critical paths (right) of circuits. (a) Conventional design without ASA. (b) Proposed ASA.

3) This paper performs ASA so that timing slacks of FFs are adjusted to satisfy the target MTTF at $V_{\min}$. On the other hand, [6] increases timing slack as much as possible. This excessive slack increase leads to a useless increase in area and the number of low-Vth cells, which will be shown in Section VI-C3.

The rest of this paper is organized as follows. Section II introduces the concept of ASA and formulates the problem of ASA circuit design including the pre-ASA circuit design. Section III describes the overview of the proposed design methodology which is composed of pre-ASA circuit design and ASA implementation. Section IV introduces the selection method of pre-ASA candidates and identifies the most promising one in terms of power after ASA. Section V applies ASA to the selected pre-ASA circuit. Section VI experimentally evaluates the performance improvement thanks to MTTF-aware design regarding power. Last, concluding remarks are given in Section VII.

II. Activation-Aware Slack Assignment and Problem Formulation

This section, first, explains the concept of ASA. Next, we formulate the ASA that aims to save power while keeping MTTF as an optimization problem.

A. ASA

Before introducing ASA, let us first explain the conventional design. The left of Fig. 3(a) illustrates the path delay distribution of a conventionally designed circuit, and the right shows the pair of the activation probability and timing violation probability of nonintrinsic critical paths. In the conventional circuit design flow, cell instances included in noncritical paths are replaced with smaller cells and high-Vth cells for reducing power dissipation and area. Therefore, the number of paths whose delays are close to the critical path delay increases. On the other hand, this replacement decreases timing margin of the paths that go through the replaced instances and may increase the timing error occurrence probability under variations. In other words, more instances are prone to cause path delay variations.

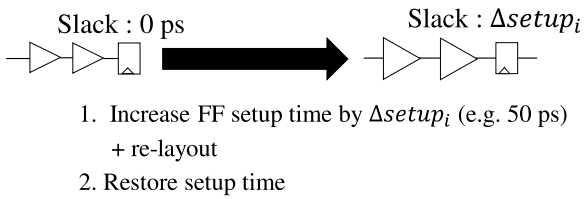

Fig. 4. Example of FF-based ASA.

On the other hand, ASA increases timing slacks of highly activated nonintrinsic critical paths. The left of Fig. 3(b) exemplifies the path delay distribution of the ASA circuit. As ASA enforces larger slacks on highly activated paths, highly activated paths sustain timing margin even when gate delay varies. Accordingly, as shown in the right of Fig. 3(b), timing violation probability in these paths is dramatically reduced compared to the conventional circuit, which is the main advantage of the ASA. These reductions extend MTTF and consequently save power, as mentioned in Section I. Here, it should be noted that ASA partially loses the power and area reduction acquired by the conventional design optimization. From this sense, we need to find a better tradeoff relation between the timing error occurrence probability and power. For pursuing the better tradeoff, the proposed ASA adjusts failure probability of the path to target failure probability as shown in the right of Fig. 3(b). In other words, the amount of slack increase is assigned to reduce power and area overheads while satisfying the target MTTF. Thanks to this assignment, the proposed ASA can save the overhead while extending MTTF and saving power. Note that the failure probability is defined as the product of activation probability and timing violation probability of a path, and the target failure probability can be calculated from the target MTTF, where the detail is given in Section IV-B. The proposed design methodology of ASA will be explained in Section V.

B. Problem Formulation

The concept of ASA was explained previously using the path delay distribution shown in Fig. 3. However, the path-based design optimization for ASA circuits is not efficient since the number of paths in a circuit is huge. Instead, we choose FF-based design optimization for ASA circuits. Fig. 4 exemplifies two-step FF-based ASA: 1) increase setup time of the target ith FF by Δsetup_i artificially and relayout the design as an ECO process and 2) restore the original setup time for the successive analysis process. It should be noted that modifying the setup time is just one implementation and there are other ways to perform FF-based ASA. For example, we manipulate timing derate factors for each FF expecting the same result of FF-based slack assignment. With this FF-based ASA, we enforce the paths ending at the target FF to have the slack that is larger than Δsetup_i.

Note that if there are intrinsic critical paths whose path delays cannot be shortened, such paths cannot have the slack of Δsetup_i. After the ASA, the circuit area increases since conventional designs exploit such slacks for area reduction. ASA circuits have more timing margin but involve the larger area.

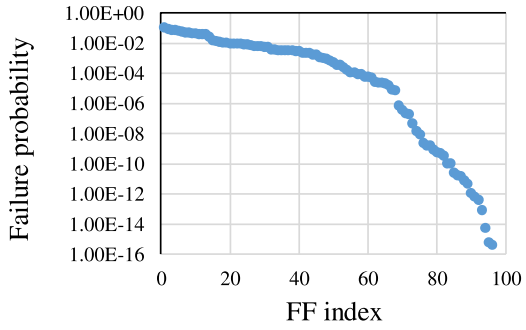

Fig. 5. Failure probabilities of FFs are largely different.

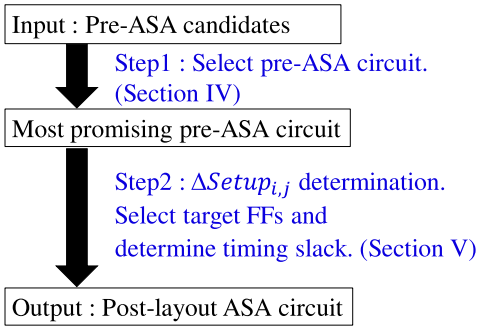

Fig. 6. Proposed design methodology with two-step procedure: 1) select most promising pre-ASA circuit in terms of power and 2) perform ASA to selected one.

An important observation in this paper is that all the FFs do not have the same contribution to MTTF. Fig. 5 shows failure probabilities, i.e., timing error occurrence probabilities, of FFs in an OR1200 OpenRISC processor, which will be used in our experiments. We evaluated them by calculating the joint probability of timing violation probability and activation probability. Fig. 5 shows that several FFs have high failure probabilities, which dominantly determine the MTTF. This result motivates us to smartly select a small number of target FFs that impact the MTTF. In this case, the area overhead of ASA can be mitigated.

Based on the above discussion, we formulate the problem of ASA circuit design as follows.

1) *Input:*
   N_{CKT} pre-ASA candidates.
2) *Output:*
   One ASA circuit.
3) *Objective:*
   Minimize: Power $= \min(\mathrm{Power}_1, \ldots, \mathrm{Power}_{N_{\mathrm{CKT}}})$.
4) *Constraints:*
   $\mathrm{MTTF}_j \geq \mathrm{MTTF}_{\min} (1 \leq j \leq N_{\mathrm{CKT}})$
   $\mathrm{Area}_j \leq \mathrm{Area}_{\max} (1 \leq j \leq N_{\mathrm{CKT}})$
   $N_{\mathrm{LVth}_j} \leq N_{\mathrm{LVth}}^{\max} (1 \leq j \leq N_{\mathrm{CKT}})$.
5) *Variables:*
   $\Delta\mathrm{setup}_{i,j}$ $(1 \leq i \leq N_{\mathrm{FF}}, 1 \leq j \leq N_{\mathrm{CKT}})$.

The inputs of this problem are N_{CKT} pre-ASA candidates, and the output is one ASA circuit. The objective of this problem is to minimize the power of the ASA circuit. The ASA circuit is constrained by MTTF ($\mathrm{MTTF}_{\min}$), circuit area ($\mathrm{Area}_{\max}$), and the number of low-Vth cells ($N_{\mathrm{Lvth}}^{\max}$). Chipwise voltage assignment adjusts the supply voltage for minimizing power dissipation while satisfying the target $\mathrm{MTTF}_{\min}$. The variables $\Delta\mathrm{setup}_{i,j}$ are the slacks given to FFs in the jth pre-ASA circuit, where $\Delta\mathrm{setup}_{i,j}$ is given to the layout ECO as an intentional increase in setup time of the ith FF_i in the jth pre-ASA circuit. N_{FF} is the number of FFs in the circuit, and it is identical in all the pre-ASA circuits. When $\Delta\mathrm{setup}_{i,j} = 0$, ith FF_i is not included in the set of target FFs of the jth pre-ASA circuit. Thus, the number of target FFs, i.e., N_{ASA}, is expressed as the number of FFs whose $\Delta\mathrm{setup}_{i,j}$ is larger than 0. Here, MTTF_j depends on $\Delta\mathrm{setup}_{i,j}$ and supply voltage, and these relations are evaluated by the stochastic error rate estimation method [5]. Area_j and N_{LVth_j} depend on $\Delta\mathrm{setup}_{i,j}$, and it is given by the layout tool after P&R ECO.

III. Overview of Proposed ASA

A difficulty to solve the formulated problem is the nonlinear relations among MTTF_j, Area_j, N_{LVth_j}, and $\Delta\mathrm{setup}_{i,j}$. Also, the evaluations of MTTF_j, Area_j, and N_{LVth_j} need relatively long CPU time, and hence, an explicit optimization is not efficient concerning CPU time. Thus, to determine the set of $\Delta\mathrm{setup}_{i,j}$ efficiently, we propose a two-step procedure shown in Fig. 6.

Fig. 6 shows the overview of the proposed design methodology which includes both pre-ASA circuit design and ASA implementation. The first procedure screens the pre-ASA candidates using the tradeoff analysis between MTTF and power, and identifies the most promising candidate that is expected to achieve the lowest power operation after ASA, and this candidate is given to the second step. After this candidate selection, circuit parameter of j is fixed, and the following second step of ASA implementation will determine $\Delta\mathrm{setup}_{i,j}$. With this screening process, we decouple the determination of the jth circuit and the ith FF. The detail of the screening procedure will be explained in Section IV.

After pre-ASA circuit selection, the proposed methodology implements ASA to the selected circuit and determines $\Delta\mathrm{setup}_{i,j}$. For various N_{ASA}, i.e., the number of FFs that ASA is applied to, we decide the set of target FFs and their $\Delta\mathrm{setup}_{i,j}$ aiming at MTTF maximization. Here, we are expecting that a circuit with longer MTTF has a larger room for power saving and N_{ASA} is related to the increase in area and the number of low-Vth cells. Then, for each set of $\Delta\mathrm{setup}_{i,j}$, we perform P&R ECO to obtain Area_j and N_{LVth_j} and evaluate the tradeoff relation between the supply voltage and MTTF_j using the stochastic error rate estimation method. From the evaluation results, we find the set of $\Delta\mathrm{setup}_{i,j}$ that minimizes power while satisfying the constraints of MTTF_j, Area_j, and N_{LVth_j}.

Taking this approach, for each N_{ASA}, we need to select N_{ASA} FFs and determine $\Delta\mathrm{setup}_{i,j}$ of the selected FFs. Section V-B explains how to select N_{ASA} target FFs, and Section V-C presents how to determine $\Delta\mathrm{setup}_{i,j}$.

IV. Design of Pre-ASA Circuit

The important consideration in this paper is how to design the pre-ASA circuit to obtain the better ASA circuit. Our preliminary work [6] prepares a pre-ASA circuit that is

designed at the maximum operating frequency (FMAX) and performs ASA. This pre-ASA circuit tends to include low-Vth cells and large-area cells and consequently increases dynamic and static power. On the other hand, the circuit designed at looser frequency may be flexible for an additional design change in ECO compared with the FMAX design, and hence, ASA may provide better optimization results. Here, please remind that chipwise voltage assignment compensates the frequency difference in design time with voltage scaling after fabrication. Therefore, it is not apparent which pre-ASA circuits achieve the minimum power operation after ASA and postfabrication voltage assignment while satisfying the target MTTF at a given clock period.

This section proposes a method to select the pre-ASA circuit that is expected to be the most power-efficient from candidates. This supposes that pre-ASA candidates are synthesized and laid out with various clock periods since the clock period has the largest impact on the tradeoff between speed and area/power. The other constraints, such as maximum transition time, area, and power, are kept unchanged for simplicity. The proposed selection method first estimates the minimum supply voltage after ASA (V_{min}) at which the circuit can achieve the given target MTTF for each candidate. Then, our method evaluates the circuit power with each V_{min} and selects the circuit whose power is minimum among the candidates. The power evaluation and comparison can be performed using EDA tools with relatively short CPU times. On the other hand, for V_{min} estimation, the explicit computation to find V_{min} cannot be conducted regarding CPU time, since the solution space of ASA is huge. When the pre-ASA circuit has N_{FF} FFs, the total combinational number of FF selection for ASA, N_{comb}, is $\sum_{N_{\mathrm{ASA}}=1}^{N_{\mathrm{FF}}} C(N_{\mathrm{FF}}, N_{\mathrm{ASA}})$. In case of $N_{\mathrm{FF}} = 1000$ and $N_{\mathrm{ASA}} = 100$, for example, N_{comb} reaches 7.18×10^{139}.

To tackle this issue, we focus on the MTTF-dominant FF, which is expected to cause a timing error at the highest supply voltage in the circuit. If we can find the MTTF-dominant FF efficiently, the CPU time of V_{min} estimation can be dramatically reduced. In this case, the iteration times of ECO for finding the MTTF-dominant FF are limited to the number of FFs (N_{FF}), and this is much smaller than N_{comb}. From the above, to estimate V_{min} of each candidate circuit efficiently, the proposed method executes the following two steps for each candidate: 1) finding the most MTTF-dominant FF after ASA and 2) calculating V_{min}.

A. Finding the MTTF-Dominant FF

For finding the most MTTF-dominant FF after ASA, for each FF, the proposed method increases the timing slack as much as possible, and calculates the failure probability of the FF of interest, where the failure probability is the joint probability of the timing violation probability and activation probability. We calculate the timing violation probability by performing SSTA and derive the activation probability of each path by associating the signal transition time in logic simulation and the path delay in STA, as shown in Fig. 7. For the details of computation, please see [5].

In this paper, we regard the FF with the highest failure probability as the MTTF-dominant FF. Here, the most MTTF-dominant FF varies depending on the supply voltage. Therefore, we evaluate failure probabilities at various supply voltages and utilize them for V_{min} estimation in Section IV-B.

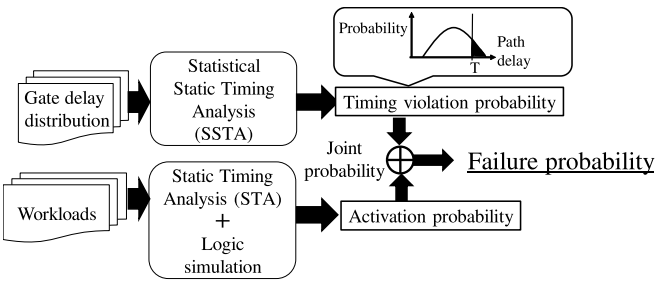

Fig. 7. Failure probability calculation.

B. Calculating V_{min}

Next, the proposed method estimates V_{min} by comparing the failure probability with the target one at each supply voltage. This paper calculates the target failure probability from the target MTTF with the WC assumption in which all the FFs have the identical highest failure probability, and timing error occurrences at all FFs are uncorrelated

$$P_{\mathrm{fail\_FF}}^{\mathrm{max}} = \frac{1}{N_{\mathrm{FF}} \times \mathrm{MTTF}_{\mathrm{min}}}. \tag{1}$$

In (1), $P_{\mathrm{fail\_FF}}^{\mathrm{max}}$ is the upper bound of the failure probability of FFs, which is the target failure probability. N_{FF} is the number of FFs, and $\mathrm{MTTF}_{\mathrm{min}}$ is the lower bound of the MTTF, i.e., target MTTF. In other words, (1) calculates the target failure probability to meet $\mathrm{MTTF}_{\mathrm{min}}$ even when all the N_{FF} FFs have the target failure probability.

Equation (1) is derived from the following equation with the Maclaurin expansion:

$$\mathrm{MTTF}_{\mathrm{min}} = \frac{1}{1 - \left(1 - P_{\mathrm{fail\_FF}}^{\mathrm{max}}\right)^{N_{\mathrm{FF}}}} \tag{2}$$

where $1 - P_{\mathrm{fail\_FF}}^{\mathrm{max}}$ represents the lower bound of the probability that no errors occur at an FF during a unit time, where the unit time is defined as a clock cycle in this paper. Thus, $(1 - P_{\mathrm{fail\_FF}}^{\mathrm{max}})^{N_{\mathrm{FF}}}$ is the lower bound of the probability that no errors occur in all the FFs during a clock cycle. Therefore, $1 - (1 - P_{\mathrm{fail\_FF}}^{\mathrm{max}})^{N_{\mathrm{FF}}}$ is the upper bound of the probability that timing error occurs during a clock cycle. Note that (2) simply multiplies the failure probabilities of FFs, which means that timing error occurrences at different FFs are treated as uncorrelated events.

Next, (2) is derived. Here, the MTTF-dominant FF has the highest failure probability among N_{FF} FFs. Then, if the failure probability of the MTTF-dominant FF is smaller than $P_{\mathrm{fail\_FF}}^{\mathrm{max}}$, the probability that no timing error occurs in the circuit during a clock cycle, i.e., $\prod_{i=1}^{N_{\mathrm{FF}}}(1 - P_{\mathrm{fail\_FF}_i})$ is larger than $(1 - P_{\mathrm{fail\_FF}}^{\mathrm{max}})^{N_{\mathrm{FF}}}$ since $P_{\mathrm{fail\_FF}}^{\mathrm{max}}$ is larger than $P_{\mathrm{fail\_FF}_i}$ for each i $(1 \leq i \leq N_{\mathrm{FF}})$. Note that $P_{\mathrm{fail\_FF}_i}$ denotes the failure

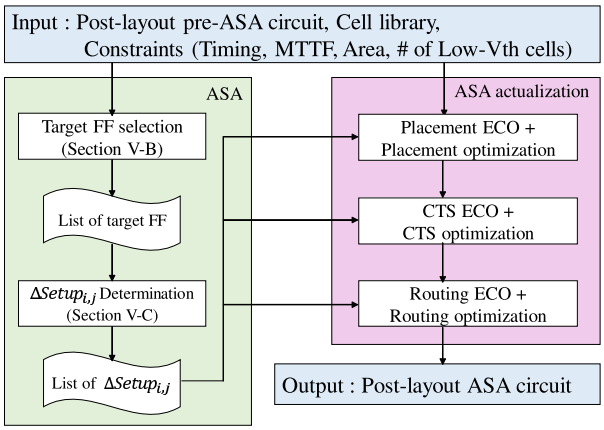

Fig. 8. Proposed ASA design flow. Proposed ASA first selects target FFs (Section V-B) and determines $\Delta\text{setup}_{i,j}$ for target FFs (Section V-C). Then, the proposed ASA performs placement, CTS, and routing ECO.

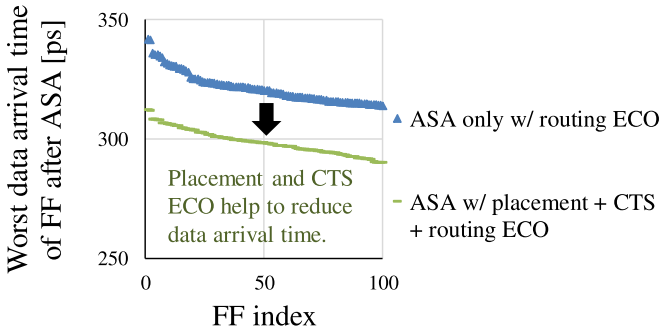

Fig. 9. Largest data arrival time of each FF after ASA at typical PVTA condition in the AES circuit.

probability of the ith FF. In this case,

$$\text{MTTF}_{\min} = \frac{1}{1-\prod_{i=1}^{N_{\text{FF}}}(1-P_{\text{fail\_FF}_i})} \tag{3}$$

$$> \frac{1}{1-\left(1-P_{\text{fail\_FF}}^{\max}\right)^{N_{\text{FF}}}}. \tag{4}$$

From the above, when the failure probability of the MTTF-dominant FF is smaller than $P_{\text{fail\_FF}}^{\max}$, the circuit satisfies the target MTTF of $\text{MTTF}_{\min}$. Therefore, the proposed method finds $V_{\min}$ where the failure probability of the MTTF-dominant FF is smaller than $P_{\text{fail\_FF}}^{\max}$ obtained from (1).

V. ASA Implementation

This section proposes a design method that applies ASA to the pre-ASA circuit designed in Section IV.

A. Design Flow

Fig. 8 shows the proposed design flow of the ASA circuit. First, the proposed design method selects target FFs for ASA and determines $\Delta\text{setup}_{i,j}$ for these FFs. Section V-B explains how to select N_{ASA} target FFs, and Section V-C presents how to determine $\Delta\text{setup}_{i,j}$. Then, the proposed method modifies the circuit to actualize $\Delta\text{setup}_{i,j}$ through ECO processes.

First, let us explain why we focus on the layout ECO-based ASA, not logic synthesis-based ASA [6]. The synthesis-based ASA has more flexibility in changing circuit structure compared with ECO-based ASA. On the other hand, there are large differences in path delays before and after layout design since the interconnects have a large impact on timing. Our final goal is to have a layout that reflects our ASA idea. When we apply ASA in logic synthesis, the intention of ASA may disappear in the P&R process. Consequently, we manipulate timing slacks in ECO phase for making sure that the intention of ASA is reflected in the final layout.

Then, let us explain why the proposed ASA gives $\Delta\text{setup}_{i,j}$ to three ECO processes from the placement stage to the routing stage. When we enforce the target FFs to increase timing slack, some other FFs and combinational cells need replacement and rerouting. Also, clock tree resynthesis might be necessary. Fig. 9 shows such a tendency. For each FF, we extract the largest data arrival time at path endpoints and compare the data arrival time between the following two ASA conditions: 1) perform ASA only with routing ECO and its optimization and 2) perform ASA with placement, clock tree synthesis (CTS), and routing ECO and their optimizations. Here, we use an advanced encryption standard (AES) circuit, where the detail will be explained in Section VI. We obtained the worst data arrival time of each FF by STA at the typical PVTA condition. Fig. 9 indicates that three ECO processes at the placement, CTS, and routing achieve the smaller path delay than a single ECO process at routing. Therefore, the proposed ASA performs that three ECO processes to more precisely achieve $\Delta\text{setup}_{i,j}$ in the layout modification.

B. Target FF Selection

This section presents a selection method of target FFs aiming at MTTF maximization since MTTF extension can be converted to power saving. An approach of FF selection is to choose target FFs with the descending order of failure probability. This approach is based on the idea that the FFs with high failure probability are more likely to cause a timing error. Thus, increasing slack of such FFs improves MTTF efficiently. However, this selection strategy does not take into account the common paths between FFs, e.g., how many instances are shared between paths and how much paths are shared between FFs. If a set of FFs shares the most of paths, increasing timing slacks of the small number of FFs may be enough for the set, which contributes to reducing the number of target FF selection. In the VLSI circuit, there are many sets of FFs that share the clock path and data path. If we ignore the common path and choose FFs according to their failure probabilities only, the selected FFs may share the large part of clock and data paths. In this case, the most of the selected FFs could be redundant and thus wasteful in terms of the circuit area and power. To overcome this issue, on the other hand, the proposed method introduces gatewise failure probabilities. Gatewise failure probability denotes how much the instance contributes to the timing error. The proposed method first distributes the failure probability from endpoint FF to instances at the upper stream of the FF as gatewise failure probability. Then, our method selects target FFs by

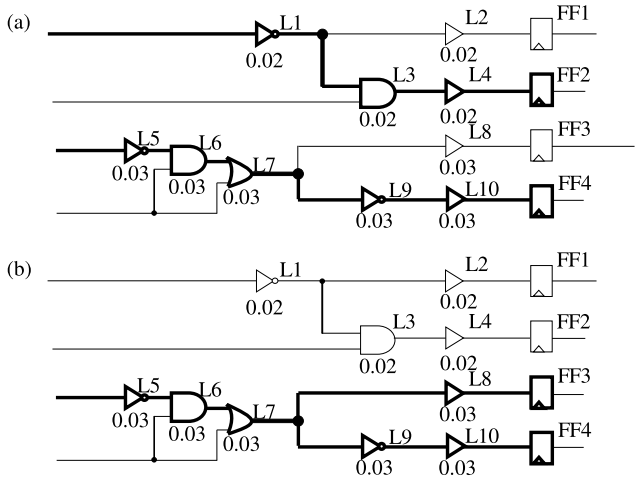

Fig. 10. Example to select target FFs. (a) Proposed ILP selects FF2 and FF4, and expected probability of error reduction is 0.21. (b) If we select FF3 and FF4, expected probability of error reduction is 0.18.

solving the covering problem of instances weighted with the gatewise failure probability for maximizing the sum of gatewise failure probabilities aiming at MTTF maximization.

Fig. 10 shows a simple example, where the circuit is composed of 10 combinational logic cells and 4 FFs. Fig. 10(a) exemplifies the proposed FF selection, and Fig. 10(b) chooses FFs with the descending order of FF failure probability. The numbers attached to each gate are the gatewise failure probabilities, where their computation is explained later. Let us suppose $N_{\mathrm{ASA}} = 2$ in the following.

When the slack times of FF2 and FF4 are increased, the slack times of L1, L3, L4, L5, L6, L7, L9, and L10 are also increased. In this case, even if a delay variation occurs at one of L1, L3, L4, L5, L6, L7, L9, and L10, the variation might be concealed by the increased slack. The expected probability of error reduction corresponds to the sum of gatewise failure probabilities and it is 0.21 ($= 0.02 + 0.02 + 0.02 + 0.03 + 0.03 + 0.03 + 0.03 + 0.03$). On the other hand, if we choose FF3 and FF4 like Fig. 10(b), the slack times of L5, L6, L7, L8, L9, and L10 are increased. In this case, the reduced failure probability is 0.18($= 0.03 \times 6$) and this amount of reduction is smaller than the previous one. In this case, TTF tends to be shorter.

In this paper, we propose an ILP-based FF selection method. We formulate the FF selection problem (or instance covering problem) as follows.

1) *Objective:*
   Maximize: $\sum_{k=1}^{N_{\mathrm{inst}}} (P_{\mathrm{fail\_inst}_k} \times B_{\mathrm{inst}_k})$.
2) *Constraints:*
   $0 \le B_{\mathrm{inst}_k} \le 1 \ (1 \le k \le N_{\mathrm{inst}})$
   $0 \le B_{\mathrm{FF}_i} \le 1 \ (1 \le i \le N_{\mathrm{FF}})$
   $\sum_{i=1}^{N_{\mathrm{FF}}} B_{\mathrm{FF}_i} \le N_{\mathrm{ASA}}$
   $B_{\mathrm{inst}_k} \le \sum_{i=1}^{N_{\mathrm{FF}}} (B_{\mathrm{FF}_i} \times B_{\mathrm{FF}_i\_\mathrm{inst}_k})$.
3) *Variables:*
   $B_{\mathrm{FF}_i} \ (1 \le i \le N_{\mathrm{FF}})$.

The number of instances in the circuit is N_{inst}. The objective of this ILP problem is to maximize the sum of $(P_{\mathrm{fail\_inst}_k} \times B_{\mathrm{inst}_k})$. $P_{\mathrm{fail\_inst}_k}$ is the gatewise failure probability of the kth instance, representing how much kth instance contributes to timing error. B_{inst_k} is a binary variable and it becomes 1 when the kth instance is located in the upper stream of any target FF. Therefore, the sum of $P_{\mathrm{fail\_inst}_k} \times B_{\mathrm{inst}_k}$ represents the gatewise failure probability reduction. In this problem, we assign binary variables B_{FF_i}, where B_{FF_i} becomes 1 when the ith FF is selected as target FFs.

The first and second constraints are given to restrict B_{inst_k} and B_{FF_i} to binary numbers. The third constraint means that the number of target FFs for ASA should be equal or less than N_{ASA}. The fourth constraint is a key constraint that defines the relation between B_{inst_k} and B_{FF_i}. $B_{\mathrm{FF}_i\_\mathrm{inst}_k}$ is a binary constant which is determined by the circuit topology, and it becomes 1 when the kth instance is included in the paths ending at the ith FF. The product term of $B_{\mathrm{FF}_i} \times B_{\mathrm{FF}_i\_\mathrm{inst}_k}$ becomes 1 when both B_{FF_i} and $B_{\mathrm{FF}_i\_\mathrm{inst}_k}$ are 1. B_{inst_k} becomes 0 only when the product of B_{FF_i} and $B_{\mathrm{FF}_i\_\mathrm{inst}_k}$ is 0 for all the FFs. On the other hand, if the kth instance is included in the paths ending at target FFs, at least one of the products of B_{FF_i} and $B_{\mathrm{FF}_i\_\mathrm{inst}_k}$ becomes 1. In this case, B_{inst_k} can be 1. In this ILP formulation, we are maximizing the sum of $(P_{\mathrm{fail\_inst}_k} \times B_{\mathrm{inst}_k})$, and hence, B_{inst_k} is necessarily assigned to be 1.

The remaining issue is $P_{\mathrm{fail\_inst}_k}$ calculation. The failure probabilities at individual FFs, $P_{\mathrm{fail\_FF}_i}$, can be computed referring to Fig. 7. Now, we calculate $P_{\mathrm{fail\_inst}_k}$ using $P_{\mathrm{fail\_FF}_i}$ as follows:

$$P_{\mathrm{fail\_inst}_k} = \max \left\{ \frac{P_{\mathrm{fail\_FF}_i}}{\sum_{k=1}^{k_{\max}} (B_{\mathrm{FF}_i\_\mathrm{inst}_k})} \right\} (1 \le i \le N_{\mathrm{FF}}). \quad (5)$$

Equation (5) assumes that each instance included in the fan-in cone of FF_i has the same contribution to a timing error for simplicity, and hence, $P_{\mathrm{fail\_FF}_i}$ is divided by the number of instances in the fan-in cone of FF_i. When we need to consider the different contributions of each instance due to, for example, different intrinsic variation sensitivities of the instances themselves, we may distribute $P_{\mathrm{fail\_FF}_i}$ to each gatewise failure probability taking into account the different sensitivities. We also note that an instance can be included in the fan-in cones of multiple FFs. For coping with this, a max operation is performed in (5).

In this paper, we use ILP to derive an exact solution. For this meaning, other techniques, such as SAT, could be used for the FF selection. Besides, ILP has proven to be NP-hard [7] in general, and thus, the ILP may not be suitable for large-scale optimization problems due to computational cost. In this paper, for mitigating the computational cost, we only consider timing-critical and activating paths in ASA, which will be explained in Section VI-A, and reduce the size of design problem. Although the weakness of the ILP is not completely solved, we can solve the FF selection problem of the two benchmark circuits, which will be explained in Section VI-A, in a few seconds thanks to the problem size reduction. When the circuit size becomes larger and the CPU time is unacceptable, we need to, for example, find an approximate solution or partition the circuit into subcircuits for problem size reduction.

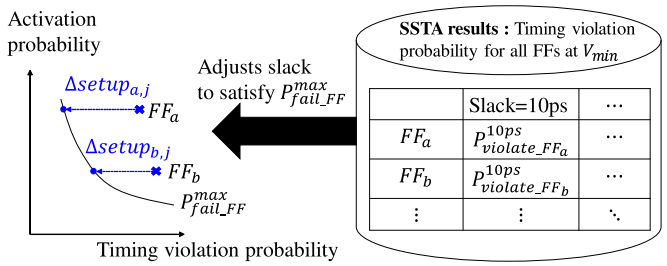

Fig. 11. $\Delta\text{setup}_{i,j}$ determination. Proposed ASA adjusts the setup slack so that each FF satisfies $P^{\max}_{\text{fail\_FF}}$.

C. $\Delta setup_{i,j}$ Determination

Next, we determine $\Delta\text{setup}_{i,j}$ for the set of target FFs selected in Section V-B. Fig. 11 shows our ASA strategy. The proposed method gives timing slacks for each target FF so that the failure probability of each FF is equal to or smaller than the target failure probability, i.e., $P^{\max}_{\text{fail\_FF}}$. Note that ASA cannot change the activation probability but can adjust the timing violation probability. Therefore, we adjust the timing violation probability of each FF and thus set the failure probability to the $P^{\max}_{\text{fail\_FF}}$. In this paper, we obtain the relation of the timing violation probability and the timing slack from SSTA.

In summary, the proposed design needs tasks of logic simulation and STA for deriving activation probability of paths/FF, SSTA for deriving timing violation probability, pre-ASA circuit selection, ILP for selecting target FFs for ASA, ASA with layout ECO, and MTTF calculation, whereas the conventional WC design requires only STA or SSTA. In case of MTTF-aware design without ASA, logic simulation, STA, SSTA, and MTTF calculation are necessary.

VI. Experimental Evaluation

This section experimentally evaluates the performance improvement thanks to proposed MTTF-aware design in comparison to the conventional WC design. Section VI-A explains the evaluation setup and Section VI-B demonstrates the power saving effects by MTTF-aware operation and ASA individually. Last, Section VI-C examines the power saving effects thanks to the proposed MTTF-aware design and discusses the effectiveness of the proposed ASA.

A. Evaluation Setup

In this paper, we used the AES circuit and OR1200 OpenRISC processor, which is a 32-bit RISC microprocessor with five pipeline stages, as target circuits. These two circuits were laid out by a commercial P&R tool with a 45-nm Nangate standard cell library. Also, standard cell memories [8], [9] were used as external main memories with OpenRISC processor. The minimum clock period of postlayout circuits at 1.20 V in the typical PVTA conditions and the WC were 3150 and 3800 ps in OpenRISC and 370 and 480 ps in AES, respectively. Hereafter, the target clock period was set to 3800 ps in OpenRISC and 480 ps in AES, and then, ASA optimized the timing slack of FF/path for these target clock periods.

The postlayout circuits included 23 247 combinational logic cells, 2504 FFs, 2 macro cells of standard cell memory in OpenRISC, and 17 948 combinational logic cells and 530 FFs in AES. Thus, sets of N_{inst} and N_{FF} were 23 249 and 2504 in OpenRISC and 17 948 and 530 in AES, respectively. We used Gurobi Optimizer 7.0 to solve the ILP problem defined in Section V-B. The solver was executed on a 2.4-GHz Xeon CPU machine under the Red Hat Enterprise Linux 6 operating system with 1-TB memory. The required CPU times for solving the proposed ILP problem with the Gurobi optimizer were at most 2.56 s in AES and 0.53 s in OpenRISC. For calculating meaningful MTTF, practical delay variations should be considered. Our evaluation took into account the following variations.

1) Manufacturing variability, which is assumed to consist of intradie random variation and interdie variation. Both intradie and interdie variations include the nMOS and pMOS threshold voltage variation of $\sigma = 10$ mV.
2) Dynamic supply noise, which is assumed to fluctuate between -50 and 50 mV $\times$ 10 mV with 11 steps.
3) For an MTTF analysis, six degradation states of 0, 0.5, 1, 5, 10, and 15 mV are prepared. The necessary information is obtained as follows. NBTI aging, whose model is obtained by fitting the measured data in [10] to a trapping/detrapping model [11]. The fit model tells us that, for example, 18.9-year operation at 1.10 V induces 15-mV degradation of threshold voltage and at 1.20 V 3.2-year operation does. Note that we used the NBTI degradation data with the stress probability of 100% in [10] as a WC. Therefore, our experimental setup gives the most pessimistic MTTF regarding NBTI effects. Activation probability aware analysis and optimization are included in our future work. We also note that the impact of BTI on timing and the degradation speed in our experimental setup are on the same level with the 45-nm result reported in [12]. Namely, 15-mV pMOS threshold degradation after 3.2-year 1.20-V operation causes 2.0% path delay increase in our setup, whereas the performance of the system degraded by 1.5% after 500-day operation in [12].

We performed SSTA with the following processes. First, we generated the probability density functions of gate delay variability according to the assumed variations. Second, we executed sensitivity-based SSTA (see [13], [14]) with common path pessimism removal (see [15], [16]) to obtain the canonical-form expression of the timing violation probability. Third, we calculated the timing violation probability by integrating the canonical-form expression with MATLAB 2016b.

As for workload in OpenRISC, we selected three benchmark programs (CRC32, SHA1, and Dijkstra) from MIBench-mark [17]. For each program, 30 sets of input data were prepared for MTTF estimation. Totally, we used 90($= 3 \times 30$) workloads. In AES, 1000 random test patterns were used. The number of activated paths, i.e., $N_{\text{act\_path}}$, was 167 626 in OpenRISC and 81 829 in AES. Fig. 12 shows the distributions of FF activation probability in AES and OpenRISC. We can see that FFs in OpenRISC are less activated and their activation

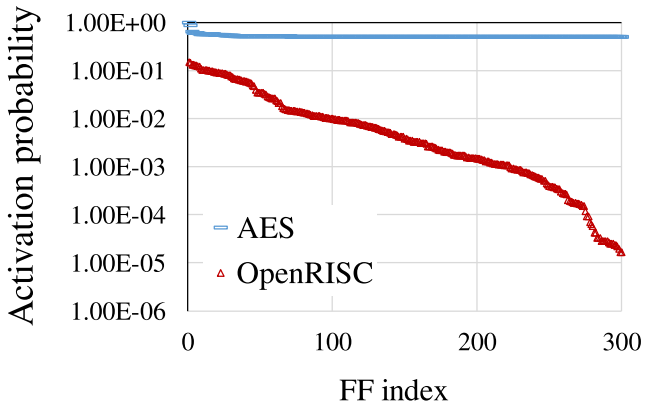

Fig. 12. Activation probability of FFs.

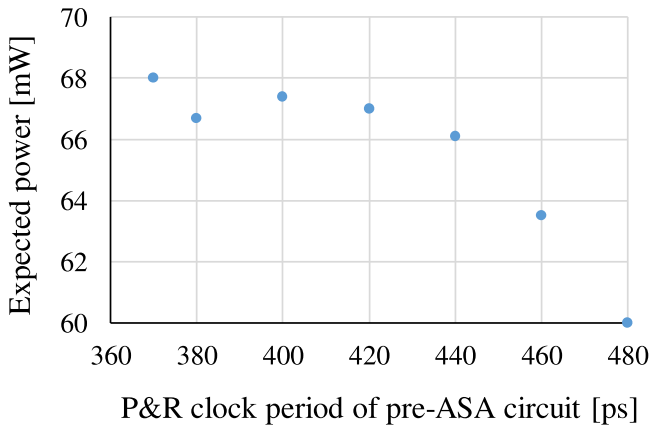

Fig. 13. Expected minimum power after ASA in AES.

probability is widely spread, which suggests the ASA is more effective to OpenRISC.

In our evaluation, we set MTTF of 10 years in OpenRISC and 1.6 years in AES, as $\mathrm{MTTF}_{\mathrm{min}}$. Note that these MTTF constraints are just examples and we can set any MTTF, such as 20 years, and derive it similarly. From (1), we set the upper bound of failure probability of an FF, i.e., $P^{\mathrm{max}}_{\mathrm{fail\_FF}}$, to 3.99×10^{-21} [1/cycle] in OpenRISC and 1.88×10^{-20} [1/cycle] in AES. These $P^{\mathrm{max}}_{\mathrm{fail\_FF}}$ were used in $\Delta\mathrm{setup}_{i,j}$ determination as explained in Section V-C.

In the MTTF evaluation, we placed emphasis on not to miss paths which affect MTTF, i.e., we should take a conservative approach. On the other hand, considering all the activated paths may lead to prohibitively long CPU time. Also, as previously shown in Fig. 5, several FFs may have high failure probabilities and dominantly determine the MTTF, i.e., taking into account these paths or FFs may be enough to evaluate MTTF. Motivated by these considerations, we selected the paths for MTTF evaluation with the following two steps. First, we calculated the maximum standard deviation of path delay variation for all the activate paths and extracted the largest standard deviation of the path delay variation. Second, we calculated timing violation probability assuming that all the paths had the above largest standard deviation and selected the paths whose failure probabilities were higher than the target failure probability. Note that we calculated the target failure probability of paths $P^{\mathrm{max}}_{\mathrm{fail\_path}}$ with the following equation, which is similar to (1):

$$P^{\mathrm{max}}_{\mathrm{fail\_path}} = \frac{1}{N_{\mathrm{act\_path}} \times \mathrm{MTTF}_{\mathrm{min}}}. \tag{6}$$

The paths whose failure probability is lower than $P^{\mathrm{max}}_{\mathrm{fail\_path}}$ do not contribute to the violation of target MTTF even with the worst delay variation, and hence, we excluded these paths from $N_{\mathrm{act\_path}}$ paths in the MTTF evaluation. Thus, we selected 1227 paths from 167 626 paths in OpenRISC and 21 067 paths from 81 829 paths in AES, and evaluated MTTF.

$\mathrm{Area}_{\mathrm{max}}$ and $N^{\mathrm{max}}_{\mathrm{LVth}}$ were set to 2.02 mm^2 and 4494 in OpenRISC and 0.05 mm^2 and 17 231 in AES. These values come from the area and the number of low-Vth cells of the pre-ASA circuits designed at 3150 ps in OpenRISC and 370 ps in AES.

With this setup, we performed ASA to both AES and OpenRISC. The number of pre-ASA candidate circuits was seven in AES, where P&R clock periods of these pre-ASA circuits were 370, 380, 400, 420, 440, 460, and 480 ps. As for OpenRISC, two candidates with 3150 and 3800 ps were given. Then, we evaluated MTTF, average supply voltage, and average power under PVTA variation by the stochastic MTTF estimation framework [5]. We prepared seven supply voltages from 1.20 to 0.90 V with 50-mV interval and swept clock period with 10-ps interval from 300 to 1000 ps in AES and from 3000 to 8000 ps in OpenRISC. We calculated power dissipation for each pair of supply voltage and clock period with a commercial power estimation tool, which reports dynamic and leakage power separately. Note that dynamic power is much higher than leakage in both OpenRISC and AES with our evaluation setup.

B. Evaluation Results

1) Selection of Pre-ASA Circuit and ASA Implementation: Fig. 13 shows the estimation results of the expected minimum power after ASA for each pre-ASA candidate in AES. From Fig. 13, we can see that the pre-ASA candidate designed at 480 ps is the most promising one regarding power. We then select the pre-ASA circuit that is laid out at 480 ps. Note that the expected minimum supply voltage that satisfies the target MTTF after ASA, V_{min}, is 0.90 V in all the pre-ASA candidates. Similar to AES, we evaluate the expected minimum power of OpenRISC and select the circuit that is laid out at 3800 ps.

Next, we perform ASA to the chosen candidate. Figs. 14 and 15 show the area and the number of low-Vth cells in the ASA circuits. In both the figures, the area and the number of low-Vth cells are normalized by $\mathrm{Area}_{\mathrm{max}}$ and $N^{\mathrm{max}}_{\mathrm{LVth}}$, respectively. Taking into account the constraints of Area_{max} and $N^{\mathrm{max}}_{\mathrm{LVth}}$, we set N_{ASA}, which is the number of FFs to which ASA is applied, to 300 in both AES and OpenRISC. An interesting observation is that the proposed ASA circuits of AES and OpenRISC have a smaller area and a smaller number of low-Vth cells compared with the pre-ASA circuit. For example, when N_{ASA} equals 300, circuit area is reduced from $\mathrm{Area}_{\mathrm{max}}$ to $0.986 \times \mathrm{Area}_{\mathrm{max}}$ by 1.4% in OpenRISC and

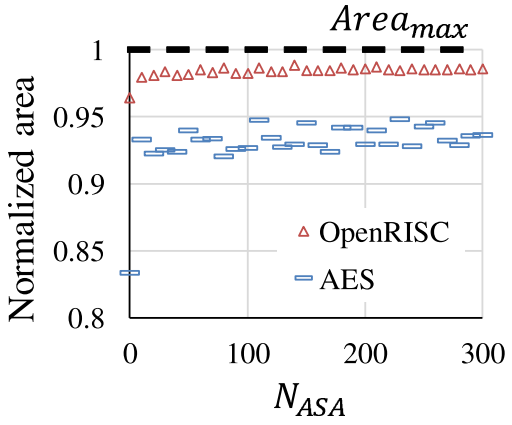

Fig. 14. Normalized area of ASA circuits Y-axis is normalized by Area_{max}.

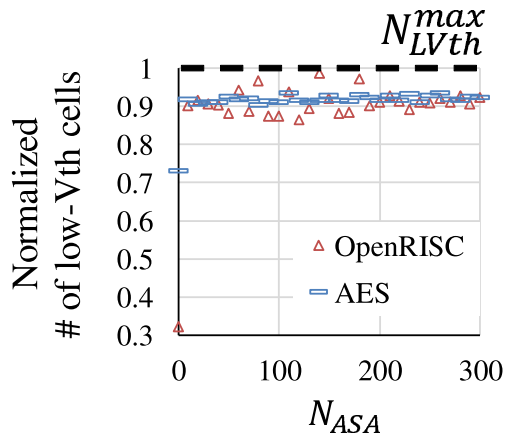

Fig. 15. Normalized number of low-Vth cells in ASA circuits. Y-axis is normalized by $N_{\text{LVth}}^{\text{max}}$.

by 6.4% in AES. We will discuss this observation in Section VI-C1.

2) Power Saving Effects: Fig. 16 shows the tradeoff curves between the power dissipation and the clock period under the MTTF constraint of 10 years in OpenRISC and 1.6 years in AES. The black quadrilateral plots represent the conventional WC design with guard banding for PVTA variation. The yellow circular plots are also WC design, but the MTTF-aware chipwise voltage assignment is performed. The blue triangular plots correspond to the proposed ASA with the MTTF-aware operation. In this section, we examine individual contributions of MTTF-aware operation and the proposed ASA in addition to the overall power saving.

First, we compare the black quadrilateral and blue triangular plots. Fig. 16 shows that the proposed MTTF-aware design reduces average power while satisfying the MTTF constraint. In Fig. 16(a), at a clock period of 3800 ps, the proposed design achieves the target MTTF with an average power of 13.2 mW, whereas the conventional WC design consumes 23.0 mW. The power saving due to the proposed design is 42.3%. Similarly, in Fig. 16(b), at 480 ps, the proposed design achieves 49.6% power saving from 132.0 to 66.5 mW. We experimentally confirmed that the proposed MTTF-aware design made the significant power saving both in AES and OpenRISC while reducing circuit area by 1.4% in OpenRISC and 6.4% in AES.

Second, we compare the conventional WC with and without MTTF-aware operation, i.e., black quadrilateral and yellow circular plots. Fig. 16 shows that the MTTF-aware operation improves the performance from the conventional WC. For example, at 3800 ps, the MTTF-aware operation achieves 26.0% power saving from 23.0 to 17.0 mW in OpenRISC and 41.4% power saving from 132.0 to 77.3 mW in AES. This power saving effects reveal that the MTTF-aware operation can significantly reduce the excessive operation margin while satisfying the target MTTF without any circuit modification.

Third, we compare the yellow circular and blue triangular plots to clarify the performance improvement by the proposed ASA. Fig. 16 shows that the proposed ASA further improves performance from the conventional WC with MTTF-aware operation. For example, at 3800 ps, the MTTF-aware operation achieves 22.3% power saving from 17.0 to 13.2 mW in OpenRISC and 13.9% power saving from 77.3 to 66.5 mW in AES. Here, it should be noted that AES has many FFs with higher activation probability as shown in Fig. 12, which means that paths having the slack of 0 or close to 0 tend to have high failure probability. Thus, the effectiveness of the ASA is smaller in AES than in OpenRISC. We also observe that the performance improvement thanks to ASA is the largest around the target clock periods of 3800 ps in OpenRISC and 480 ps in AES and it becomes smaller as the period goes away from the target one since ASA optimized the circuit at the target clock period under MTTF constraint. There could be a room for improvement at different clock periods.

C. Discussion

This section first examines the power saving in terms of V_{dd}, area, and the number of low-Vth cells, and second discusses the effectiveness of the proposed ASA regarding FF selection and slack determination.

1) Reduction of V_{dd}, Area, and the Number of Low-Vth Cells: Performance evaluation results in Section VI-B2 showed that the proposed design saved power significantly. Let us investigate its reason.

First, we examine the supply voltage reduction effects by the proposed MTTF-aware design. Fig. 17 shows the tradeoff curves between the average supply voltage and the clock period under the MTTF constraints of 10 years in OpenRISC and 1.6 years in AES. We can see that the proposed design, which corresponds to the blue triangular plots, achieves the target MTTF at a lower supply voltage compared with the conventional WC design, i.e., black quadrilateral plots. For example, in Fig. 17(a), at a clock period of 3800 ps, the proposed design achieves the target MTTF at an average supply voltage of 0.90 V, whereas the conventional WC design requires 1.20-V operation, which means that the proposed design achieves 25.0% V_{dd} reduction from 1.20 to 0.90 V. Thanks to this V_{dd} reduction, the circuit power dissipation is dramatically reduced, as shown in Fig. 16. Also in AES, our MTTF-aware design reduces the supply voltage from

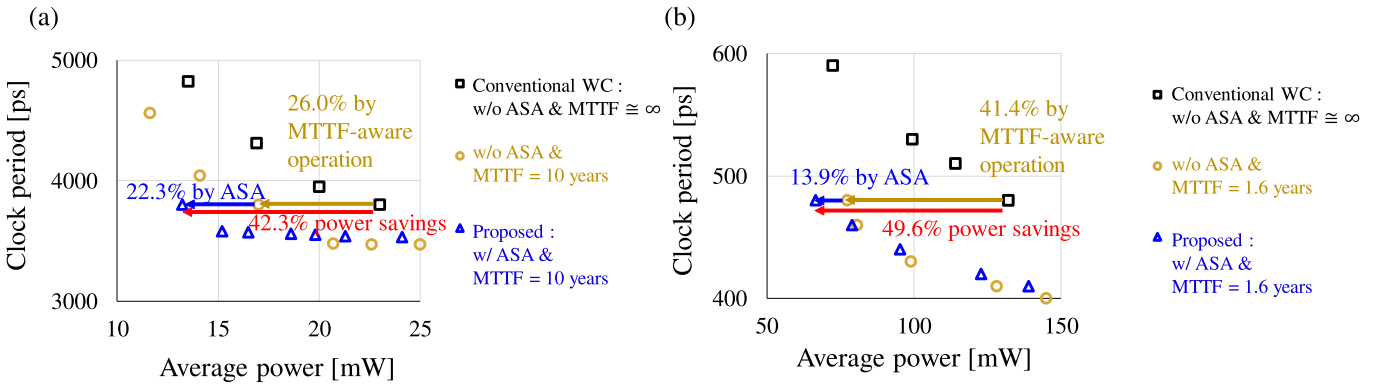

Fig. 16. Tradeoff curves between clock period and average power. (a) OpenRISC. (b) AES.

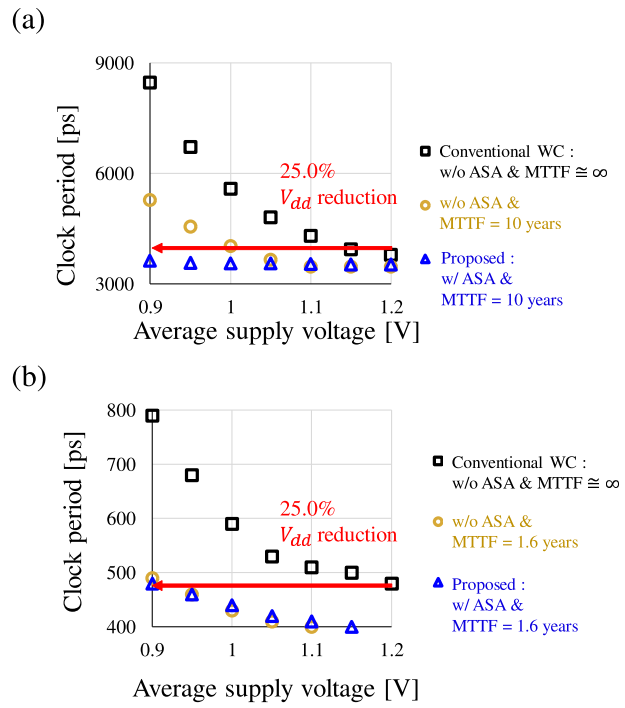

Fig. 17. V_{dd} reduction by proposed MTTF-aware design. (a) OpenRISC. (b) AES.

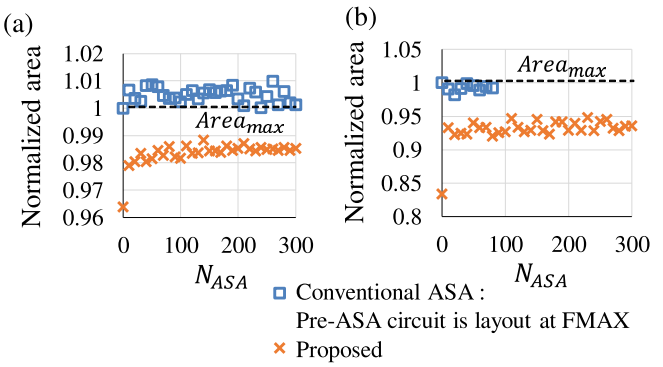

Fig. 18. Area of conventional and proposed ASA circuits. Y-axis is normalized by Area_{max}. Proposed ASA reduces area from pre-ASA circuits. (a) OpenRISC. (b) AES.

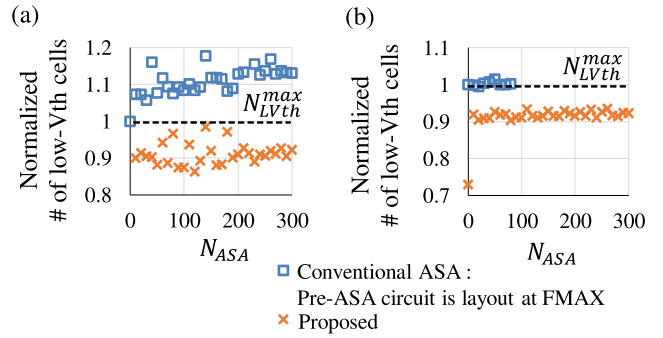

Fig. 19. Number of low-Vth cells of the conventional and proposed ASA circuits. Y-axis is normalized by $N_{\text{LVth}}^{\text{max}}$. Proposed ASA reduces the number of low-Vth cells from pre-ASA circuits. (a) OpenRISC. (b) AES.

1.20 to 0.90 V and achieves 25.0% V_{dd} reduction, as shown in Fig. 17(b).

Second, we investigate the area and the number of low-Vth cells of ASA circuits, which are partially shown in Figs. 14 and 15. Figs. 18 and 19 show the area and the number of low-Vth cells of the conventional and proposed ASA circuits. In both figures, the area and the number of low-Vth cells are normalized by Area_{max} and $N_{\text{LVth}}^{\text{max}}$ similar to Figs. 14 and 15. In this context, conventional ASA is supposed to use pre-ASA circuits that are laid out at FMAX, e.g., 3150 ps in OpenRISC and 370 ps in AES. On the other hand, the proposed methodology performs ASA to circuits laid out for longer periods of 3800 ps in OpenRISC and 480 ps in AES. Note that pre-ASA AES circuit designed at 370 ps has only 81 FFs whose failure probability is larger than 0, and hence, the maximum N_{ASA} is set to 81.

Figs. 18 and 19 show that the proposed ASA reduces the area and the number of low-Vth cells from the conventional ASA circuits and even from their pre-ASA circuits. For example, at $N_{\text{ASA}} = 300$, in OpenRISC, the proposed ASA reduces the area by 1.4% from Area_{max} to $0.986 \times \text{Area}_{\text{max}}$ and the number of low-Vth cells by 7.7%, while the conventional ASA increases the area by 0.1% and increases the number of low-Vth cells by 13.1%. Similarly, at $N_{\text{ASA}} = 300$ in AES, the proposed ASA reduces the area by 6.4% and decreases the number of low-Vth cells by 7.7%. These reductions directly decrease the dynamic and static power dissipation. Thus, we confirmed that the proposed pre-ASA circuit selection further contributes to improving power in addition to V_{dd} reduction by ASA.

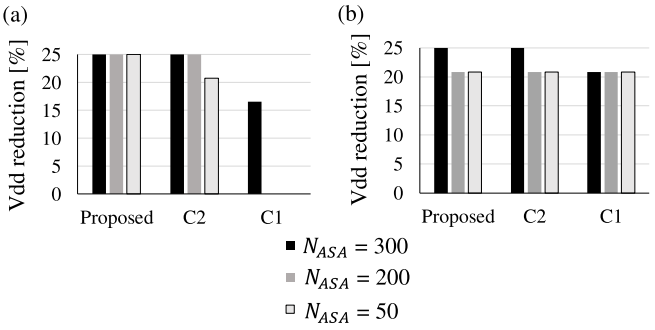

Fig. 20. Achieved V_{dd} reduction. (a) OpenRISC. (b) AES.

2) Effectiveness of FF Selection for ASA: Next, we compare the proposed methodology with the following two approaches focusing on the effectiveness of target FF selection for ASA.

1) *C1:* Choose FFs in an ascending order of slack time.
2) *C2:* Choose FFs in a descending order of failure probability.

The first approach of C1 supposes that timing-critical FFs are most likely to cause timing error. This method needs only STA or SSTA timing reports, and hence, this approach is more tractable. The second approach of C2 places importance on the failure probability. Remind that failure probability is defined as the joint probability of the timing violation probability and the activation probability. To calculate the activation probability, we need to perform a logic simulation with prospective workloads or to calculate signal transition rates analytically in, for example, [18].

Fig. 20 summarizes the achieved supply voltage reduction with the proposed methodology, C1 and C2. The proposed methodology achieves the largest supply voltage reduction, which is 25.0% from 1.20 to 0.90 V, in the cases of $N_{\mathrm{ASA}} = 50, 200$, and 300 in OpenRISC and $N_{\mathrm{ASA}} = 300$ in AES. Besides, C2 also achieves the 25.0% supply voltage reduction in the cases of $N_{\mathrm{ASA}} = 200$ and $N_{\mathrm{ASA}} = 300$ in OpenRISC and $N_{\mathrm{ASA}} = 300$ in AES. Here, the difference between C1 and C2 in AES is much smaller than OpenRISC. For example, at $N_{\mathrm{ASA}} = 50$, both C1 and C2 achieve 20.8% V_{dd} reduction in AES, whereas C1 could not reduce supply voltage and C2 attains 20.8% V_{dd} reduction in OpenRISC. A possible reason is that, in AES, the activation probability of FFs is much higher and more similar than OpenRISC as shown in Fig. 12, and hence, timing-critical FFs are likely to have high failure probability.

Fig. 21 shows the MTTF comparison, where $N_{\mathrm{ASA}} = 50$ and $V_{\mathrm{dd}} = 0.90$ V. From this figure, we can see that the proposed methodology attains the best tradeoff relation between MTTF and clock period. For example, from Fig. 21(a), at a clock period of 3800 ps, the proposed ASA improves MTTF compared with C1 from 3.32×10^{-14} to 9.56×10^{1} years by 15 orders of magnitude and compared with C2 from 9.07×10^{-9} years to 9.56×10^{1} years by 10 orders of magnitude. Similarly, from Fig. 21(b), at a clock period of 480 ps, the proposed ASA improves MTTF by six orders of magnitude from C1 and by five orders of magnitude from C2. Thus, the MTTF improvement of the proposed ASA is remarkable even while the proposed ASA also reduces the

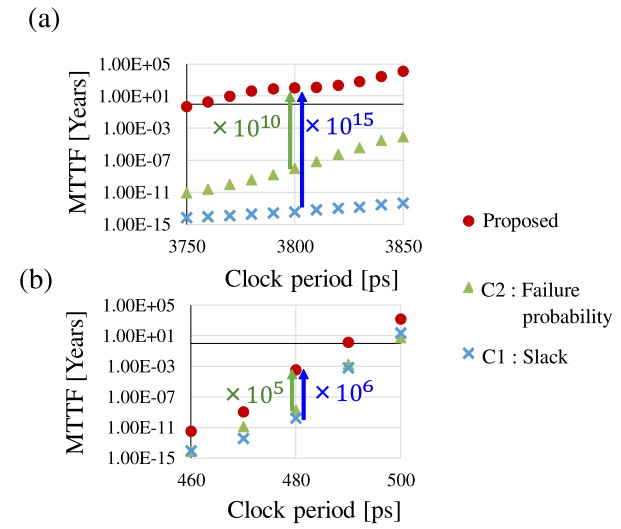

Fig. 21. MTTF comparison. Proposed FF selection improves MTTF significantly compared with C1 and C2. (a) OpenRISC. (b) AES.

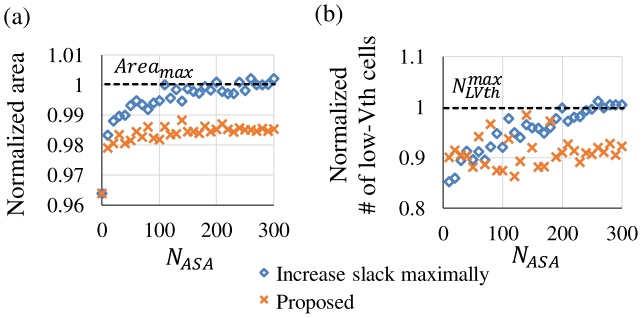

Fig. 22. Comparison of (a) area and (b) number of low-Vth cells in OpenRISC. Y-axis is normalized by $\mathrm{Area}_{\mathrm{max}}$ in (a) and $N_{\mathrm{LVth}}^{\mathrm{max}}$ in (b).

area and the number of low-Vth cells. The longer MTTF means fewer timing errors in field, which is also desirable for resilient circuit designs, such as Razor [19] and tunable replica circuit [20], and error prediction technique, e.g., timing error predictive FF [21]. With the ASA, the power dissipation of such resilient circuits could be reduced further and the reliability would improve.

3) Effectiveness of Slack Determination: Last, we investigate the importance of slack determination, i.e., $\Delta\mathrm{setup}_{i,j}$ determination, by ASA. Fig. 22 shows the comparison results of the area and the number of low-Vth cells between the proposed ASA and the conventional approach [6] that increases slack as much as possible. In this comparison, the identical pre-ASA OpenRISC circuit laid out at 3800 ps is given for clarifying the effectiveness of the $\Delta\mathrm{setup}_{i,j}$ determination. From Fig. 22, we can see that the proposed ASA saves both the area and the number of low-Vth cells. For example, at $N_{\mathrm{ASA}} = 300$, the proposed ASA saves the area by 1.6% and reduces the number of low-Vth cells by 8.1%. These reductions contribute to dynamic and static power reduction. These results show that the determination of $\Delta\mathrm{setup}_{i,j}$ in the proposed ASA more contributes to power saving compared with the conventional ASA that increases slack maximally. Note that, in AES, the proposed ASA also increases slack maximally, and thus, the area and the number of low-Vth cells are identical to those of the conventional ASA.

VII. Conclusion

This paper proposed an MTTF-aware design methodology. The key ideas of the proposed design methodology are MTTF-aware operation and design optimization with ASA, where the ASA gives timing slacks to nonintrinsic active critical paths and reduces the number of activated paths whose delays are very close to those of the inherent critical paths. The proposed optimization includes the pre-ASA circuit design and ASA implementation. In the pre-ASA circuit design, the proposed methodology selects the most promising one from candidates regarding power dissipation. For each candidate, this selection estimates the minimum supply voltage after ASA ($V_{\min}$) at which the circuit can achieve the target MTTF and evaluates the power at $V_{\min}$. Thus, we choose the circuit whose estimated power is minimum. Then, the proposed methodology selects a set of FFs for FF-based ASA using ILP so that it reduces the sum of gatewise failure probability maximally. We evaluated MTTF of circuits with and without ASA and examined how much power saving could be obtained while satisfying the target MTTFs of 10 years in OpenRISC and 1.6 years in AES. The evaluation results showed that the circuits with ASA achieved 49.6% power saving in the AES circuit and 42.3% power saving in the OpenRISC processor. Also, thanks to pre-ASA design in the proposed methodology, 6.4% of the area and 7.7% of low-Vth cells are reduced in the AES circuit, and 1.4% of the area and 7.7% of low-Vth cells are reduced in the OpenRISC processor.

References

[1] B. Zhang and M. Orshansky, "Modeling of NBTI-induced PMOS degradation under arbitrary dynamic temperature variation," in *Proc. Int. Symp. Qual. Electron. Design*, Mar. 2008, pp. 774–779.

[2] T. Wang and Q. Xu, "On the simulation of NBTI-induced performance degradation considering arbitrary temperature and voltage variations," in *Proc. 51st ACM/EDAC/IEEE Design Automat. Conf. (DAC)*, Jun. 2014, pp. 1–6.

[3] W. Wang *et al.*, "The impact of NBTI on the performance of combinational and sequential circuits," in *Proc. 44th Annu. Design Automat. Conf.*, Jun. 2007, pp. 364–369.

[4] S. Iizuka, M. Mizuno, D. Kuroda, M. Hashimoto, and T. Onoye, "Stochastic error rate estimation for adaptive speed control with field delay testing," in *Proc. Int. Conf. Comput.-Aided Design*, Nov. 2013, pp. 107–114.

[5] S. Iizuka, Y. Masuda, M. Hashimoto, and T. Onoye, "Stochastic timing error rate estimation under process and temporal variations," in *Proc. Int. Test Conf.*, Oct. 2015, pp. 1–10.

[6] Y. Masuda, M. Hashimoto, and T. Onoye, "Critical path isolation for time-to-failure extension and lower voltage operation," in *Proc. 35th Int. Conf. Comput.-Aided Design*, Nov. 2016, Art. no. 63.

[7] P. Cappello and K. Steiglitz, "Some complexity issues in digital signal processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-32, no. 5, pp. 1037–1041, Oct. 1984.

[8] P. Meinerzhagen, C. Roth, and A. Burg, "Towards generic low-power area-efficient standard cell based memory architectures," in *Proc. Int. Midwest Symp. Circuits Syst.*, Aug. 2010, pp. 129–132.

[9] A. Teman, D. Rossi, P. Meinerzhagen, L. Benini, and A. Burg, "Controlled placement of standard cell memory arrays for high density and low power in 28 nm FD-SOI," in *Proc. 20th Asia South Pacific Design Automat. Conf. (ASP-DAC)*, Jan. 2015, pp. 81–86.

[10] H. Awano, M. Hiromoto, and T. Sato, "Variability in device degradations: Statistical observation of NBTI for 3996 transistors," in *Proc. 44th Eur. Solid State Device Res. Conf. (ESSDERC)*, Sep. 2014, pp. 218–221.

[11] B. Velamala *et al.*, "Compact modeling of statistical BTI under trapping/detrapping," *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3645–3654, Nov. 2013.

[12] P.-F. Lu and K. A. Jenkins, "A built-in BTI monitor for long-term data collection in IBM microprocessors," in *Proc. IEEE Int. Rel. Phys. Symp.*, Apr. 2013, pp. 4A.1.1–4A.1.6.

[13] H. Chang and S. S. Sapatnekar, "Statistical timing analysis under spatial correlations," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 9, pp. 1467–1482, Sep. 2005.

[14] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, and S. Narayan, "First-order incremental block-based statistical timing analysis," in *Proc. 41st Design Automat. Conf.*, Jul. 2004, pp. 331–336.

[15] V. Garg, "Common path pessimism removal: An industry perspective: Special Session: Common Path Pessimism Removal," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2014, pp. 592–595.

[16] J. Zejda and P. Frain, "General framework for removal of clock network pessimism," in *Proc. IEEE/ACM Int. Conf. Comput. Aided Design*, Nov. 2002, pp. 632–639.

[17] M. R. Guthaus, J. S. Ringenberg, D. Ernst, T. M. Austin, T. Mudge, and R. B. Brown, "MiBench: A free, commercially representative embedded benchmark suite," in *Proc. IEEE Int. Workshop Workload Characterization*, Dec. 2001, pp. 3–14.

[18] F. N. Najm, "Transition density: A new measure of activity in digital circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 12, no. 2, pp. 310–323, Feb. 1993.

[19] S. Das *et al.*, "A self-tuning DVS processor using delay-error detection and correction," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp. 792–804, Apr. 2006.

[20] K. A. Bowman *et al.*, "A 45nm resilient and adaptive microprocessor core for dynamic variation tolerance," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 194–208, Jan. 2011.

[21] H. Fuketa, M. Hashimoto, Y. Mitsuyama, and T. Onoye, "Adaptive performance compensation with in-situ timing error predictive sensors for subthreshold circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 2, pp. 333–343, Feb. 2012.

Yutaka Masuda (S'14) received the B.E. and M.E. degree in information systems engineering from Osaka University, Suita, Japan, in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree in information science and technology.

His current research interests include low-power circuit design.

Mr. Masuda is a Student Member of IPSJ. Since 2018, he has been a Research Fellow with the Japan Society for the Promotion of Science.

Takao Onoye (S'93–M'95–SM'07) received the B.E. and M.E. degrees in electronic engineering and the Dr.Eng. degree in information systems engineering from Osaka University, Suita, Japan, in 1991, 1993, and 1997, respectively.

He is currently a Professor with the Department of Information Systems Engineering, Osaka University. His current research interests include media-centric low-power architecture and its system-on-chip implementation.

Dr. Onoye is a member of The Institute of Electronics, Information and Communication Engineers, Information Processing Society of Japan, and the Institute of Television Engineers of Japan.

Masanori Hashimoto (S'00–A'01–M'03–SM'11) received the B.E., M.E. and Ph.D. degrees in communications and computer engineering from Kyoto University, Kyoto, Japan, in 1997, 1999, and 2001, respectively.

He is currently a Professor with the Department of Information Systems Engineering, Graduate School of Information Science and Technology, Osaka University, Suita, Japan. His current research interests include the design for manufacturability and reliability, timing and power integrity analysis, reconfigurable computing, soft error characterization, and low-power circuit design.

Dr. Hashimoto was a recipient of the Best Paper Award from Asia and South Pacific Design Automation Conference in 2004 and the Best Paper Award of the IEICE Transactions in 2016. He was on the Technical Program Committee of international conferences, including Design Automation Conference, International Conference on Computer Aided Design, International Test Conference, Symposium on VLSI Circuits, ASP-DAC, and DATE. He serves/served as an Associate Editor for the IEEE Transactions on Very Large Scale Integration, IEEE Transactions on Circuits and Systems I, *ACM Transactions on Design Automation of Electronic Systems*, and *Microelectronics Reliability* (Elsevier).