A Process and Temperature Tolerant Oscillator-Based True Random Number Generator

Takehiko AMAKI††, Nonmember, Masanori HASHIMOTO††a, and Takao ONOYE††b, Members

SUMMARY This paper presents an oscillator-based true random number generator (TRNG) that dynamically unbiases 0/1 probability. The proposed TRNG automatically adjusts the duty cycle of a fast oscillator to 50%, and generates unbiased random numbers tolerating process variation and dynamic temperature fluctuation. A prototype chip of the proposed TRNG was fabricated with a 65 nm CMOS process. Measurement results show that the developed duty cycle monitor obtained the probability of ‘1’ within 50 ± 0.007% in five chips in the temperature range of 0°C to 75°C. Consequently, the proposed TRNG passed the NIST and DIEHARD tests at 7.5 Mbps with 6,670 µm² area.

A Process and Temperature Tolerant Oscillator-Based True Random Number Generator

1. Introduction

Random number generation is an underlying technology to accomplish highly secure systems. For example, secret and public-key generation and challenge-response authentication require unpredictable random number. For such systems, a true random number generator (TRNG), which produces random number from physical random sources, is suitable due to its inherent unpredictability of output bits.

Probability of ‘1’ occurrence is the most significant criterion indicating the randomness of the output bit stream. The probability of ‘1’ occurrence, \( p \), should be approximately 50% to pass the frequency test of NIST tests [1], and the range of acceptable probability is only 0.250%. Process variation, temperature fluctuation, and power supply noise easily make the probability exceed the acceptable range. For mitigating this, recent literatures presented methods to reduce the bias of probability of ‘1’. Figure 1 illustrates the basic concept of bias correction. The bias monitor measures the bias of the TRNG, the information of the bias is feed to the bias corrector through the controller, and the corrector compensates the bias of the probability of ‘1’. Let us show an example. Srinivasan et al. [2] proposed a TRNG employing a metastable cross-coupled inverter. The TRNG tuned the configuration every cycle according to the output bits, and then, the TRNG attained tolerance to PVT-variation.

Oscillator-based TRNG, which utilizes jitter of oscillators as a random source, is a popular method for generating true random number [3]–[5]. This is because it simply consists of digital circuits, i.e. oscillators and a sampler and hence the oscillator-based TRNG is easy to implement, process-portable, and process-scalable. In addition, the TRNG is inherently tolerant to deterministic supply noises [6].

This paper proposes an oscillator-based TRNG which dynamically adjusts the duty cycle of the oscillator for unbiasing the output random bits. For this purpose, we have developed a duty cycle monitor and duty cycle corrector. Thanks to this unbiasing, the proposed TRNG attains robustness to the process and temperature variations in addition to the above-mentioned tolerance to the deterministic noises.

The rest of this paper is organized as follows. Section 2 describes the structure of the proposed TRNG and explains its operation. Section 3 presents measurement results of the proposed TRNG fabricated with a 65 nm technology. Finally, concluding remarks are given in Sect. 4.

2. Proposed Oscillator-Based TRNG

This section presents the overall structure of the proposed TRNG and explains the duty cycle monitoring and correction in the proposed TRNG.

2.1 Overall Structure

Figure 2 shows the proposed oscillator-based TRNG with the duty cycle correction. The basic components are fast and slow oscillators and a sampler. The fast oscillating signal is sampled with the jittery slow clock. An output being determined by a time when the clock rises, the randomly fluctuating rise edges of the slow signal result in a random bit stream. The duty cycle of the fast oscillator then determines the probability of ‘1’ occurrence. In the proposed TRNG, the duty cycle correction is enabled by the duty cycle mon-
itor, the duty cycle corrector and the corrector controller. Temperature information, which is provided by an external thermal sensor, is also given to the corrector controller.

The duty cycle monitor receives the oscillating signal of the fast oscillator and outputs a digitized number representing the duty cycle. The controller receives the digitized monitor output, and looks up the target monitor output from the table whose key is temperature. The outputs of the monitor are digital signals, which enables easy processing in the corrector controller. Then, the controller outputs a control signal to the duty cycle corrector so that the monitor output gets close to the target output.

Here, let us explain why the target monitor output is selected according to the temperature. Temperature affects not only the fast oscillator but also the duty cycle monitor and the sampler, and therefore the same digitized monitor outputs do not always mean the same probability of ‘1’ taking into account the temperature change. Accordingly, this work preliminarily measures the monitor outputs corresponding to the target probability of ‘1’ at various temperatures, and stores those outputs as a table of the target monitor outputs in the non-volatile memory. The duty cycle corrector varies the duty cycle of fast oscillating signal according to the control signals. This process can also eliminate effects of manufacturing variability on the quality of TRNG output.

The key components of the proposed TRNG are the duty cycle monitor and correctors, and they will be explained in the following.

2.2 Duty Cycle Monitor

Figure 3 illustrates a concept of the proposed duty cycle monitor. Output bit sampling, which is a method of duty cycle estimation, is also shown for comparison. Output bit sampling, whose principle is similar to the random sampling method [7], gathers the output bits from the sampler and calculates the proportion of ‘1’ as the duty cycle. Note that output bit sampling uses jittery slow oscillator as the random clock while the random sampling method [7] employs the chaotic oscillator, and then output bit sampling requires several cycles of the slow oscillator. On the other hand, the proposed monitor directly measures time differences between HIGH and LOW of the fast signal and obtains the duty cycle. The monitor estimates the duty cycle within one cycle of the slow oscillator, which is much faster than output bit sampling.

Figure 4 shows a block diagram of the duty cycle monitor. The monitor circuit mainly consists of two (P-type and N-type) duration-to-delay converter (DDC) and a Vernier time-to-digital converter (TDC) [8]. While the P-type and N-type DDCs receive the same oscillating signal from the fast oscillator through the duty cycle corrector, the rise transition timings at the outputs are different. The time difference between the two rising edges of the P-type DDC and the N-type DDC becomes larger as the duty cycle increases, where the edge from the N-type DDC is earlier than the P-type. The detailed circuit-level operation of DDCs will be explained later. An important point here is that the time difference represents the duty cycle. The time difference is digitized by the Vernier TDC and then the digitized information of the duty cycle is transmitted to the controller. In this work, a 63-bit Vernier TDC is implemented.

The DDC includes serially-connected DDC units, and in this work each DDC consists of four DDC units (Fig. 4). The schematics of P-type and N-type DDC units are illustrated in Figs. 5 and 6. Here, we explain the behavior of the P-type DDC unit only, but a similar explanation is valid for N-type DDC unit. The P-type DDC unit receives the signal from the fast oscillator (fast) and positive and negative resets (rst and xrst), where rst and xrst are generated from the signal of the slow oscillator. In the first DDC unit, the voltage of the negative input node (xin) is equal to rst, and in the other DDCs, xin is connected to the output node (out) of the previous unit. Each P-type DDC unit consists of five transistors. We specially illustrate Cm2 and Cm5 which represent the gate capacitances of M2 and M5, since their gate
areas are larger than other transistors.

Figure 7 exemplifies waveforms at the nodes in the P-type DDC unit. $V(C_{M5})$ and $V(C_{M2})$ mean the input gate voltages to $M5$ and $M2$. In an initial state, $xrst$ is LOW, and $rst$ and $xin$ of the first unit are HIGH, and thus $C_{M2}$ is charged through $M4$ and $C_{M5}$ is discharged through $M3$. After that, $xrst$ changes to HIGH, and $rst$ and $xin$ of the first unit become LOW. Note that the fast oscillating signal is always input to fast. In this situation, $M1$ and $M2$ charge $C_{M5}$ while fast is LOW. The transistors of $M1$ and $M2$ cannot fully charge $C_{M5}$ within a cycle of fast, and the voltage at node $n_m$, which is the gate voltage of $M5$, increases gradually rather than stepwise, because the frequency of fast is high and the RC product of $C_{M5}$ and the series on-resistances of $M1$ and $M2$ is much larger than the cycle time of the fast oscillator. This gradual charging enhances the time resolution of the proposed monitor since even a small difference between the durations of HIGH and LOW is amplified in time by accumulating the difference across several cycles and consequently a larger temporal fluctuation of the rise transition at node $n_m$ can be obtained. In the meantime, the voltage at $n_m$ exceeds the threshold voltage of $M5$, and $M5$ turns on. $C_{M2}$ is discharged through $M5$, and then, LOW signal is output to the next DDC unit. This circuit operation repeats sequentially in the series-connected four DDC units. Finally, the NOT gate in Fig. 4 receives the fall edge from the last DDC unit and generates a rise signal as an output of the P-type DDC. In the N-type DDC, a buffer instead of the NOT gate is connected at the last stage.

To sum up, the propagation delay of the P-type DDC unit depends on the time necessary to charge $C_{M5}$ up, and therefore, the delay gets larger as the duty cycle of fast increases. In contrast, the delay of the N-type DDC gets smaller as the duty cycle increases. Therefore, the time difference between the rise transitions of P-type and N-type DDCs increases as the duty cycle becomes higher. This time difference is digitized by the TDC and its information is given to the corrector controller. Please remind that the corrector controller selects the target monitor output from the table stored in the non-volatile memory according to the temperature, and this target monitor output in the table was determined based on the preliminary measurement for calibration as mentioned in Sect. 2.1. This calibration eliminates effects of manufacturing variability on DDC and TDC and impacts of through-current and gate-leakage current on DDC performance, and makes accuracies of, for example, $C_{M5}$, less important.

It should be emphasized that the proposed duty cycle monitor estimates the duty cycle for every random bit generation, i.e. the monitor outputs the estimate every cycle of the slow oscillator. On the other hand, when the duty cycle is estimated from the output bit sequence, we need to gather a number of bits to obtain statistically accurate estimation. This means that the proposed TRNG can cope with faster dynamic environmental fluctuation.

### 2.3 Duty Cycle Corrector

The duty cycle corrector is a programmable delay cell whose propagation delay for rising input is controllable while the delay for falling input is constant. Figure 8 shows the behavior of the delay cell. An oscillating signal is input to in and is output to out being delayed through the cell. The time of HIGH and LOW of in are $T_{HIGH}$ and $T_{LOW}$, and those of out are $T'_{HIGH}$ and $T'_{LOW}$, respectively. The propagation delay for fall edge is $D_{fall}$ and that for rise edge is $D_{rise}$, which is controlled with ctrl. Here, $T'_{HIGH} = T_{HIGH} - D_{rise} + D_{fall}$ and $T'_{LOW} = T_{LOW} + D_{rise} - D_{fall}$ as illustrated in Fig. 8. Then, the duty cycle of in, $d_{in}$, and that of out, $d_{out}$, are expressed as follows:

\[
d_{in} = \frac{T_{HIGH} - T_{LOW}}{T_{HIGH} + T_{LOW}}
\]

\[
d_{out} = \frac{T'_{HIGH} - T'_{LOW}}{T'_{HIGH} + T'_{LOW}}
\]
The programmable delay cell tunes \( d_{\text{out}} \) by changing \( D_{\text{rise}} \).

Figure 9 shows a schematic of the programmable delay cell, namely the duty cycle corrector. An oscillating signal from the fast oscillator is input to \( \text{in} \), and the output signal of \( \text{out} \) is given to the duty cycle monitor and the sampler. The duty cycle corrector consists of two correctors; a coarse corrector and a fine corrector. Both the correctors have the same circuit topology, but their resolutions of duty cycle correction are differently designed via transistor sizing. The coarse corrector is configured statically, and it compensates the bias originating from process variation. On the other hand, the fine corrector is automatically controlled by the corrector controller, and it corrects the bias due to dynamic temperature fluctuation.

We next explain the behavior of the corrector. Here, only the coarse corrector is explained since the fine corrector operates in the same manner. The programmable inverter consists of two inverters connected in parallel, one of which includes gating transistors to control NMOS drive strength. A rising signal of \( \text{in} \) propagates through the programmable inverter by discharging \( C_{\text{inv}} \) with the NMOSs, and the discharging current depends on the number of enabled transistors. The control signal, \( C \), thus changes the delay for a rise input signal. On the other hand, the PMOS drive strength is constant. Consequently, the duty cycle corrector adjusts the duty cycle by changing the rise propagation delay (Eq. (2)).

\[
d_{\text{in}} = \frac{T_{\text{HIGH}}}{T_{\text{HIGH}} + T_{\text{LOW}}},
\]
\[
d_{\text{out}} = \frac{T'_{\text{HIGH}}}{T'_{\text{HIGH}} + T'_{\text{LOW}}} = \left( \frac{d_{\text{in}} + D_{\text{fall}}}{T_{\text{HIGH}} + T_{\text{LOW}}} \right) + \frac{D_{\text{rise}}}{T_{\text{HIGH}} + T_{\text{LOW}}},
\]

The proposed TRNG was fabricated with a 65 nm CMOS process. Figure 10 shows a chip photo of the implemented TRNG. The oscillators, the duty cycle monitor, the duty cycle corrector and the sampler were implemented on the chip, and the total macro cell area is 6,500 \( \mu \text{m}^2 \). The corrector controller and the table of the target monitor outputs were implemented with FPGA (Cyclone II), and a discrete thermal sensor (MAXIM DS1621) was employed.

### 3. Measurement Results

This section shows measurement results of the proposed TRNG and gives comparisons with TRNGs in literatures.

#### 3.1 Prototype Chip

This subsection evaluates the quickness of the duty cycle estimation. For comparison, output bit sampling method is considered here, which gathers a bit sequence from the TRNG output and calculates the proportion of ‘1’ occurrence.

Figure 11 shows the accuracies of probability estimation with the duty cycle monitor and with the output bit sampling. The y-axis represents the standard deviation of the measured 100 duty cycles, which corresponds to the statistical accuracy of the duty cycle estimation. The x-axis is the time spent to measure a duty cycle. The time unit is the number of cycles of the slow oscillator. As the measurement time becomes longer, the output bit sampling method has more 0/1 samples and the proposed monitor can output more estimates of duty cycle. In this case, the statistical uncertainty decreases. From another viewpoint, the x-axis can be regarded as the time required to attain the corresponding accuracy. VDD for the DDC was 0.9 V and VDD for the others were 1.2 V.

Note that the measured output value of the monitor \( x \) was digital data whose bit length was 6 bits, and the value was converted into the duty cycle. The first order
approximation of the probability of ‘1’ at 25°C was \( y = 0.16x + 44.22 \), and then, the measured output value \( x \) was multiplied by 0.16. For example, when a standard deviation of measured output values is 3.0, the converted accuracy of duty cycle is \( 3.0 \times 0.16 = 0.48\% \).

Figure 11 shows that the measurement time of the proposed monitor is 4,100 times as short as that of the output bit sampling when the pass mark of accuracy is set to 0.25% as an example. From another point of view, the proposed monitor attains 70 times higher accuracy when the measurement time is one cycle of the slow oscillator. That is, in this case, the error of the probability of ‘1’ from 50% is estimated to be 0.7% with the proposed duty monitor while the error is 46.1% with the output bit sampling method. For reference, the TRNG in [2] adjusts the bias every cycle according to the output bit, which corresponds to the probability estimation by the output bit sampling method with one cycle.

### 3.3 Duty Cycle Correction

This subsection shows the utility of the dynamic duty cycle correction. 2-Mbit streams were measured with a logic analyzer, and the probability of ‘1’ and entropy were calculated. Five TRNGs on different chips were measured. Temperature was changed from 0°C to 75°C.

Figure 12 (a) shows the maximum errors from 50% with and without the bias correction. The proposed TRNG effectively reduces the error to 0.07%. Figure 12 (b) shows the minimum entropy with and without the correction. The proposed TRNG attained more than 0.999999 of entropy in the five chips even under temperature variation, while the entropy without bias correction degraded significantly.

For ensuring high randomness even under process variation and temperature fluctuation, the proposed TRNG introduced a feedback loop to control the probability of ‘1’ occurrence mainly consisting of duty cycle monitor, duty cycle corrector and corrector controller, since the duty cycle of the fast oscillator mainly determine the probability of ‘1’ occurrence. The experiment confirmed that the proposed TRNG attained process and temperature tolerance.

### 3.4 Randomness Test

The proposed TRNG was evaluated with NIST [1] and DIEHARD [9] tests. NIST and DIEHARD tests evaluated 100 M and 80 M random bits obtained from the TRNG, respectively. Temperature was set to 25°C.

The oscillator-based TRNG exploits the jitter of oscillators as randomness source. As the oscillating frequency of the slow oscillator, which is related to the TRNG throughput, becomes higher, the cycle time of the oscillator becomes shorter and consequently the absolute value of the jitter decreases. To generate good random numbers, the absolute value of the jitter should be several times larger than the cycle time of the fast oscillator [10]. Therefore, there is the maximum throughput for the oscillator-based TRNG to have enough randomness and pass the randomness tests. In this experiment, a bitrate controller which was attached to
the TRNG output discarded a half of the sampled bits. The frequency of the slow oscillator was 15.1 MHz, and then, the TRNG throughput was 7.5 MHz. Above this throughput, the TRNG cannot pass NIST tests.

Table 1 lists the results of NIST randomness test. The implemented TRNG with the proposed 0/1 bias correlation successfully passed all the tests.

DIEHARD test returned 221 p-values, and the smallest and largest values were 0.0046 and 0.9995 as listed in Table 2. According to [11], a TRNG passes the entire test suite with a 95% confidence interval when the p-values are between 0.0001 and 0.9999. We thus conclude the proposed TRNG passed DIEHARD test.

3.5 Comparison

We finally compare the proposed TRNG to others. The area of our TRNG is estimated as follows. The area for the oscillators, the sampler, and the duty cycle monitor and corrector is estimated to be 6,500 $\mu$m$^2$ from the test chip layout. The corrector controller, which was implemented on an FPGA in the measurement, was synthesized for the same 65 nm process by a logic synthesis tool and the estimated area is 170 $\mu$m$^2$. A temperature sensor was assumed to already exist for other purposes. The total area is therefore 6,670 $\mu$m$^2$.

Table 3 shows the performance comparison to other recent TRNGs. The proposed TRNG passed NIST test with a small area without using post processing. In addition, this work has a significant advantage that the TRNG is tolerant to process variation and temperature fluctuation in addition to the robustness to deterministic noises [6], whereas these tolerance and robustness are not clearly demonstrated in most previous publications. For example, TRNGs in [3], [12] have a feedback loop based on output bit sampling, but its effectiveness for the tolerance improvement is not discussed in experiments. In addition, these TRNGs use post processing, and hence the unbiasing performance of the feedback loop, such as accuracy and response time, could be moderate. In [13], variation tolerance is not discussed. On the other hand, a bias correction method for metastable-based TRNG and its effectiveness are presented in [2]. This work adjusts metastable condition of cross-coupled inverters to make the occurrence probability of ‘1’ 50% according to the result of output bit sampling.

4. Conclusion

This paper presented an oscillator-based TRNG with dynamic 0/1 bias correction. The proposed duty cycle monitor and corrector make the TRNG tolerant to process and temperature variations, and their contribution to the randomness improvement was clarified with 65 nm test chips.

Acknowledgment

The VLSI chip in this study has been fabricated in the chip fabrication program of VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with STARC, e-Shuttle, Inc., and Fujitsu Ltd. This work is supported by SCOPE program of Ministry of Internal Affairs and Communications.

References


<table>
<thead>
<tr>
<th>Test name</th>
<th>Frequency</th>
<th>BlockFrequency</th>
<th>CumulativeSums</th>
<th>Runs</th>
<th>LongestRun</th>
<th>Rank</th>
<th>FFT</th>
<th>NonOverlappingTemplate</th>
<th>OverlappingTemplate</th>
<th>Universal</th>
<th>ApproximateEntropy</th>
<th>RandomExcursions</th>
<th>RandomExcursionsVariant</th>
<th>Serial</th>
<th>LinearComplexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-value</td>
<td>0.0003</td>
<td>0.0083</td>
<td>0.0019</td>
<td>0.4190</td>
<td>0.1296</td>
<td>0.1296</td>
<td>0.7399</td>
<td>0.1626</td>
<td>0.0401</td>
<td>0.4012</td>
<td>0.4012</td>
<td>0.2757</td>
<td>0.9241</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass rate</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
<td>1.00</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test name</th>
<th>Test name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>ALL PASS</td>
</tr>
</tbody>
</table>

| Minimum p-value | 0.0046 |
| Maximum p-value | 0.9995 |
| Result         | PASS   |

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Comparison with other TRNGs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>Direct amp.</td>
</tr>
<tr>
<td>Area</td>
<td>180 nm</td>
</tr>
<tr>
<td>Area normalized to 45 nm</td>
<td>25,000 $\mu$m$^2$</td>
</tr>
<tr>
<td>Throughput</td>
<td>40 Mbps</td>
</tr>
<tr>
<td>Randomness</td>
<td>FIPS140-1</td>
</tr>
<tr>
<td>Assessment</td>
<td>Entropy eval.</td>
</tr>
<tr>
<td>Post processing</td>
<td>XOR</td>
</tr>
</tbody>
</table>
tors for cryptographic applications,” NIST, pub.800-22rev1a, 2010.


