Entropy Extraction in Metastability-based TRNG

Vikram B. Suresh
Dept. of Electrical & Computer Engineering
University of Massachusetts
Amherst, USA
vsuresh@ecs.umass.edu

Wayne P. Burleson
Dept. of Electrical & Computer Engineering
University of Massachusetts
Amherst, USA
burleson@ecs.umass.edu

***Abstract*— True Random Number Generators (TRNG) implemented in deep sub micron (DSM) technologies become biased in bit generation due to process variations and fluctuations in operating conditions. A variety of mechanisms ranging from analog and digital circuit techniques to algorithmic post-processing can be employed to remove bias. In this work we compare the effectiveness of digital post-processing using the XOR function and Von Neumann Corrector with circuit calibration technique for a meta-stability based reference TRNG design. The energy consumption per random bit is used as the metric for comparison of the different techniques. The results indicate that the calibration technique is effective for 12% larger process variation than the XOR function and extracts entropy comparable to the Von Neumann Corrector at 56% lesser energy/bit. The analysis thereby demonstrates that circuit calibration provides an efficient tradeoff between entropy and energy/bit for removing bias in lightweight TRNG.**

Keywords- TRNG, entropy, process variation

I. INTRODUCTION

On chip True Random Number Generators are used in cryptographic systems for various applications like multi-core processors, communication networks and RFID. TRNG provides random keys, device id and seed for Pseudo Random Number Generators (PRNG). The potential sources of entropy for a TRNG are cosmic rays, stray electromagnetic waves and thermal noise. Specific circuits are designed to harness the randomness from these sources. The entropy is extracted in the form of random clock jitter samples, power up state of memory cells, meta-stability of devices and chaos on deterministic analog signals. Implementations in deep submicron (DSM) technologies have made these circuits vulnerable to variations in the fabrication process and operating conditions (voltage and temperature) to generate unbiased outputs. The counter measures for removing the bias are mainly digital post-processing or calibration of the TRNG circuit. In lightweight and low power applications, the basic TRNG circuit by itself will be simple in construct and consumes low energy per bit. In such cases, the digital post-processing techniques may prove to be complex in implementation, inefficient for significant process variation or expensive in terms of energy overhead.

A number of publications have suggested novel circuits to extract entropy from random physical sources. In [1], P.Kocher *et al.* have described the extraction of random numbers from thermal noise across a resistor. The amplified thermal noise is fed to a voltage controlled oscillator, the output of which is sampled to obtain random bits. Ring oscillator based designs are prominently used as TRNG [2]-[5]. Sunar *et al*. [2] have proposed a TRNG design, using *r* ring oscillator chains with their output XORed and sampled in the transition zone. The random phased drift causing the jitter in the oscillator rings act as a mechanism to extract entropy. A variant of the ring oscillator based TRNG is presented in [6]. Meta-stable systems can be used to extract randomness that arises from thermal noise. C. Tokunaga *et al* .[7] have proposed a meta-stability based TRNG which does not directly extract the random bits, but estimates the random noise present in the circuit using the resolution time for the meta-stable element to reach stability. Memory cells also provide a means to generate randomness. In [8] D. Holcomb *et al.* have proposed a novel technique of generating random bits using the power up state of SRAM cells. During the power up process, the SRAM cells resort to either of the two states based on random noise present in the design, providing random bits. The design makes use of the already present SRAM as TRNG. DRAM cells do not allow access during the refresh cycles. A memory based design has been described in [9] to use the random access time due to collision between the DRAM accesses and refresh cycles as a source of entropy. A reference program is run multiple times and the random variation in the DRAM access time is measured and converted to random number.

Variations in fabrication process and operating conditions impact the behavior of the TRNG circuits, biasing the output bits. With increasingly more number of implementations in DSM technologies, these problems are aggravated. The counter measures include digital post-processing using the XOR function [2, 3, 12, 13], Von Neumann entropy extraction [1, 2, 6] and hash extractors like universal hash function [8] or Secure Hash Algorithm (SHA-1) [6]. Circuit design techniques like charge injection [7] and configurable devices [10] are also used to tune the TRNG against bias. Each of these techniques provides varied degree of improvement in the randomness of the generated bits at different energy overhead. The bias removal technique has to be appropriately chosen based on the application, entropy requirements, design constraints and energy budgets.

In modern cryptographic applications, metastability based TRNGs are increasingly used since they are simpler to design and consume less energy as compared to other TRNG circuits. In this paper we analyze the sources of bias and the techniques

This work is partially supported by National Science Foundation: NSF CNS-0831133

to remove the same for reference metastability based TRNG. The rest of this paper will discuss, in detail, the effect of variation on TRNG circuits in section II and different bias removal mechanism in section III. Then, the implementation of the reference design, results from simulations and analysis of the same will be presented along with the conclusion and scope for future work in section IV and V respectively.

II. Effect of Variation on Bias

TRNG circuits depend on random physical variations to derive entropy. But, variations in the process, supply voltage and operating temperature may bias these circuits towards generating bits of one polarity at a higher probability than the other. Such effects render the TRNG statistically inefficient in generating random bits. With reducing feature size and increasing fabrication process complexity, TRNG designed in DSM technology may be subjected to a variety of random or systematic process variations that have the potential to introduce bias. Variation in on-chip feature size as compared to the drawn lengths, device mismatches in the form of threshold voltage (Vt), varied oxide thickness and contact resistances are some of the major sources of bias. Imperfections caused in the circuits due to prolonged usage or aging also contribute to bias and correlation in TRNG circuits.

Variation of operating voltage and temperature also affect the randomness of the TRNG output. In [10] a variation in randomness of upto 6% for varied temperature range of 100^{o}C and supply voltage range of 0.6V is observed. Noise and IR drop on the power grid may have a non-common mode effect on the different devices in a TRNG circuit. Such variations would tune parts of the circuits to work faster than the other causing disparity in the output generated. Variations in temperature also have similar effect on the performance of the TRNG. Power supply and temperature variations also provide a side channel for active attack on TRNG circuits. Controlled variation in power supply or temperature fluctuations are used to attack the TRNG devices and control the randomness of the bits generated. In [11], R. Santoro *et al.* have demonstrated an attack on ring oscillator based TRNG used in a smart card by introducing noise on the power supply through electromagnetic injection. Such an attack reduced the entropy of the TRNG and hence made the smart card unsafe for use.

III. Bias Removal Techniques

There are a number of techniques used to remove bias in the TRNG and make the output bits more random. These circuits may be broadly classified as digital post-processing techniques and calibration techniques, Fig. 1. Each of these mechanisms has specific effect in terms of efficiency in enhancing the randomness and the implementation overhead.

The XOR function is a commonly used entropy extractor. Outputs of two or more TRNG circuits are XORed to improve the entropy of the output as shown in Fig. 2. Bias in one of the circuits is masked by the other TRNG circuits. This technique also provides tolerance against device wear out or side channel attacks. Although the XOR function provides a simple implementation for improving the entropy of the design, it leads to overhead due to need for multiple TRNG circuits.

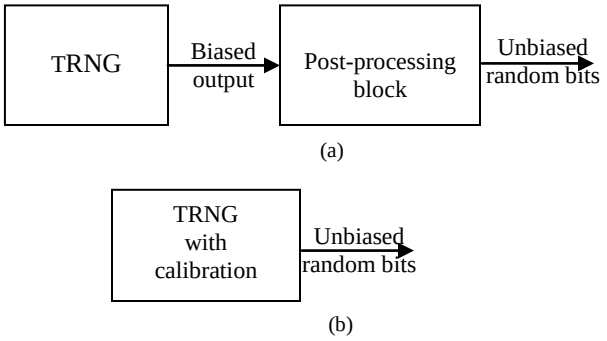


Figure 1: Bias removal techniques (a) Digital post-processing (b) Circuit calibration

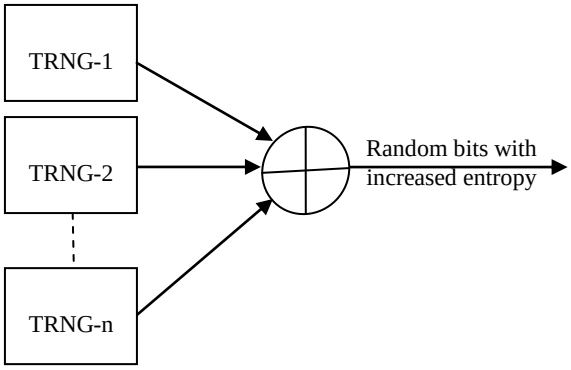


Figure 2: XOR function as entropy extractor

From the law of averaging the implementation would produce better results as more number of TRNG circuits is used. But, this would lead to additional overhead in terms of circuit area and power. Care should be taken to place the circuits in close vicinity to avoid non common mode variation in the operating conditions.

The most prominently used post-processing technique is the Von Neumann corrector or entropy extractor. The Von Neumann corrector function, as shown in Table 1, produces a balanced distribution of ones and zeros by converting bit pairs [0,1] from the TRNG into an output 1 bit and the pairs [1,0] into an output bit 0. The pairs of bits [0,0] and [1,1] are discarded. The Von Neumann corrector, Fig. 3, is very efficient in terms of producing an equal distribution of 1s and 0s. But, since the output rate of the Von Neumann corrector is not constant, the generated bits need to be stored in a shift register before using for further processing. Even with very high entropy TRNG, the maximum bit rate achievable is half the bit rate of the TRNG.

Hash functions are also used to extract entropy from the random bits generated. The universal hash function and the Secure Hash Algorithm (SHA-1) are used to increase the randomness of the bits obtained. Passing the TRNG output through a stream cipher or a block cipher also enhances the randomness of the bits. But, these digital post-processing techniques introduce design complexities in hardware and overhead in the form of increased energy consumption and reduced bit rates. Another algorithmic approach to enhance the entropy is the software implementation of resilient functions using redundant logic expressions or cyclic codes.

TABLE 1: VON NEUMANN FUNCTION

Input bit pairs (from TRNG)	Output from Von Neumann corrector
0 0	No output
0 1	1
1 0	0
1 1	No output

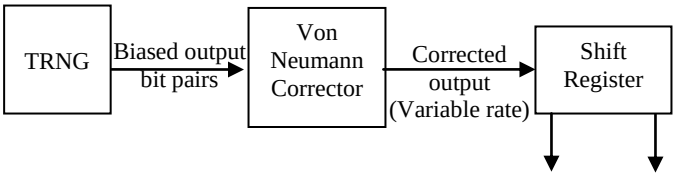


Figure 3: Von Neumann corrector

The TRNG circuits can also be physically calibrated to reduce bias by variable loading or tuning the effective device sizes by activating parallel source and sink current paths. In [12], a controller circuit calibrates the phase of the clock signal generated from a signal generator based on the quality of randomness in the bits generated. C. Tokunaga *et al.* [7] inject charge into the meta-stable circuit to counter any mismatch in the devices. The charge injection is done through an array of capacitors that are conditionally charged based on the amount of bias that needs to be corrected. The TRNG circuit described in [10] uses two levels of calibration to provide coarse and fine levels of granularity in the tuning using parallel PMOS and NMOS paths to tune the device drive strengths.

IV. IMPLEMENTATION, SIMULATION RESULTS AND ANALYSIS

In this section we present our work on the implementation of a reference TRNG circuit with three different bias removal techniques, the simulation results for the different scenarios and analysis of the same.

The reference design chosen for this work is the meta-stability based cross coupled inverter TRNG design proposed by S. Srinivasan *et al.* [10]. The circuit, shown in Fig. 4, consists of a pair of cross coupled inverters. The input of both inverters is pre-charged high during the negative half cycle of the clock. During the positive half of the clock cycle, the circuit is allowed to settle down to a stable state from a metastable condition. If both the inverters are identical, the random thermal noise on the input of the inverters decides how the contention is resolved. Accordingly, a bit 0 or bit 1 is generated each cycle. Under unbiased operation, the thermal noise acts as the source of randomness for the TRNG circuit. The circuit provides a simple platform to implement the different bias removal techniques and estimate the cost of improving the entropy.

The circuit behaves like a fair dice if the two inverters are perfectly matched. But, a relative variation in the device features or threshold voltage reduces the randomness of the bits generated. In such a case, some form of post-processing or calibration technique needs to be used to obtain a uniform distribution of bits.

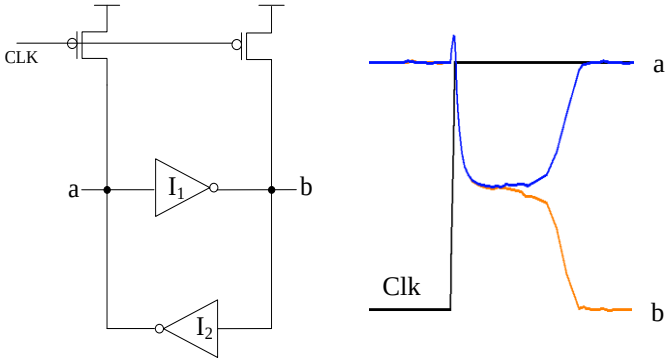


Figure 4: Metastability based TRNG

The circuit is implemented using 45nm transistor models and simulated in HSPICE. Monte Carlo simulations are run for process variation with a 3-sigma variation in the PMOS device lengths of 20% and the circuit generating bits at the rate of 1Gbps. Although simulation has clear limitations in terms of modeling physical randomness, HSPCE provides an effective simulation platform to analyze the TRNG and bias removal techniques for a wider range of variations in process and operating conditions.

Table 2 provides a comparison of the energy per random bit for the different calibration techniques. Note that Table 2 does not account for process variations and the resulting entropy and bit-rate of the various techniques. These are analyzed in the subsequent section.

A. *TRNG without correction*

The basic TRNG circuit without any post-processing or calibration is simulated for varying effective lengths of the transistors in the two inverters. The results, Fig. 5, show a steep decrease in the randomness of the bits generated with increase in relative variation between the two devices. DSM technologies at 65nm and below are inherently imperfect and such variations cannot be avoided. Hence, implementation of TRNG circuits at such low technology nodes will need additional post-processing or calibration to boost the entropy.

B. *TRNG with XOR function as entropy extractor*

As indicated earlier, XOR function can be used for entropy extraction. In this implementation, two TRNG circuits are run in parallel and their outputs are XORed to average out any variations in one of the TRNGs and obtain bits with improved randomness. The plot, Fig. 6, depicts the improvement in the

TABLE 2: COMPARISON OF ENERGY/RANDOM BIT FOR DIFFERENT BIAS REMOVAL TECHNIQUES

Bias removal technique	Average energy/random bit (pJ)
TRNG without correcttion	0.001
TRNG with XOR function	0.006
TRNG with Von Neumann corrector	0.282
TRNG with calibration	0.124

A wrong version of this paper was uploaded in the IEEE database.

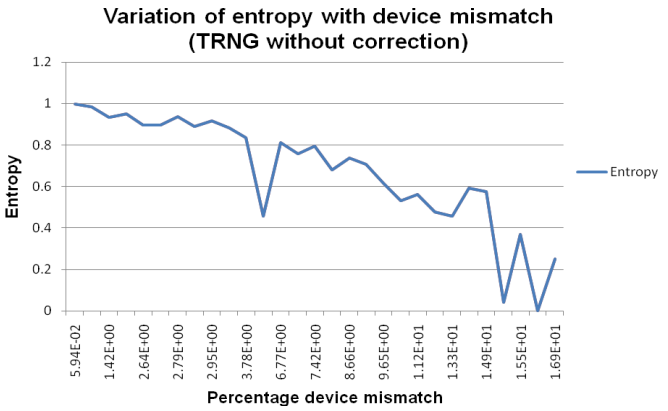


Figure 5: Variation of entropy of basic TRNG with device mismatch

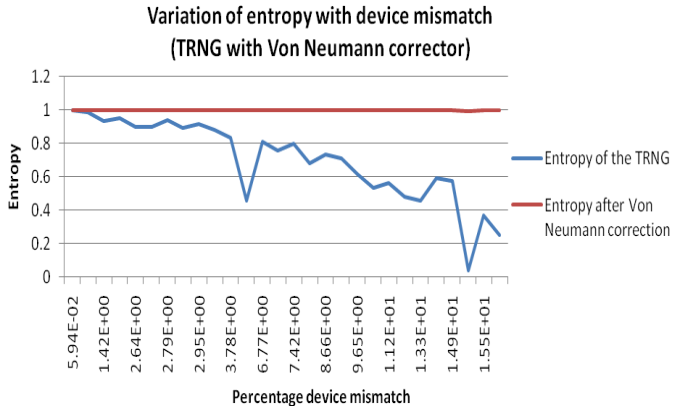


Figure 7: Enhancement of entropy using the Von Neumann corrector

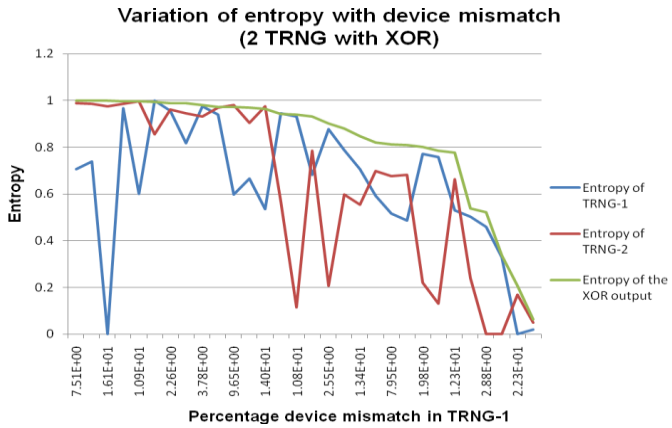


Figure 6: Enhancement of entropy using the XOR function

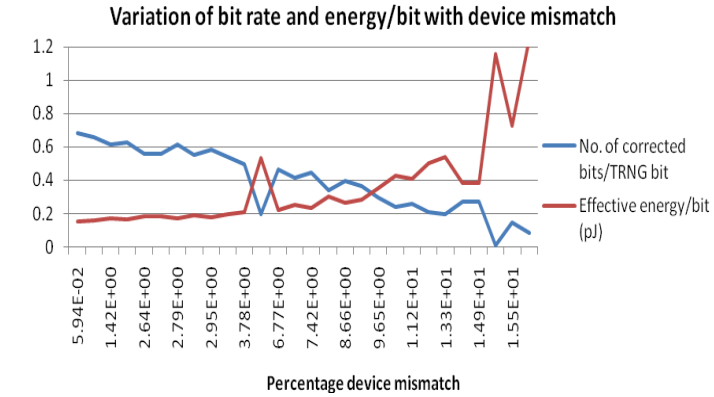


Figure 8: Variation of bit rate and energy/bit with device mismatch

randomness of the bits obtained using the XOR function. Although the XOR function provides a fair degree of enhancement, it is highly data dependent. A dip in the entropy of one of the TRNG will make the output dependent entirely on the other TRNG. As a result, with increased process variation, a significant fall in the randomness of the bits generated at the XOR output is observed. The solution to this issue is the use of multiple TRNG which would add to the overhead.

C. TRNG with Von Neumann corrector

Von Neumann corrector is a very widely used corrector or whitener to enhance the entropy of RNGs. The results, Fig. 7, for the simulation of the TRNG with Von Neumann post-processing reflect a highly robust correction mechanism in terms of improvement in entropy. The output random bits from the Von Neumann corrector are not generated at a constant rate. This variable rate of bit generation necessitates the use of a shift register, adding to the overhead. Further, in applications that require a constant flow of random bits, the Von Neumann corrector may prove to be inefficient. Fig. 8 shows a plot of the number of corrected bits generated for every TRNG bit generated with mismatch in device features. With increasing intra-die variations, the effective bit rate of the circuit decreases due to a decrease in the entropy of the basic TRNG circuit. As a result, more number of TRNG bits have to be generated per Von Neumann corrected bit. This results in an increase in the energy per random bit of the system.

D. TRNG with calibration

The calibration technique used for the reference design consists of two stages as shown in Fig. 9. The coarse grain calibration circuit consists of the programmable inverter, with tunable pull up and pull down transistors to counter the increase in drive capability of the other device because of variation. The fine grain calibration tunes the circuit by varying delays in the pre-charge clock path to remove bias due to small variations. The circuit for the coarse and fine grained tuning is shown in Fig. 10.

By using a combination of coarse and fine grain configuration, the circuit can be calibrated against variations like mismatch in feature size and threshold voltage. A comparison of the different bias removal techniques is shown in the Fig. 11. As seen from the simulation results, the circuit calibration technique provides significantly better correction in entropy as compared to the XOR function, with increasing device mismatch and improvement comparable to the Von Neumann technique, but at a constant bit rate.

A very crucial factor in choosing the bias removal technique is a tradeoff between the enhancement in entropy and the energy cost per bit. Table.2 shows the average energy/bit

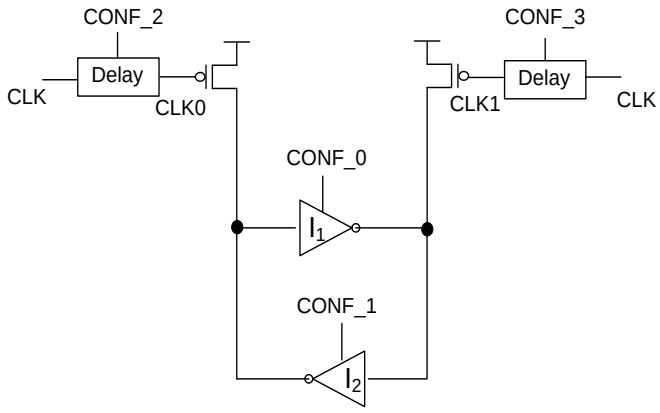


Figure 9: TRNG with calibration [10]

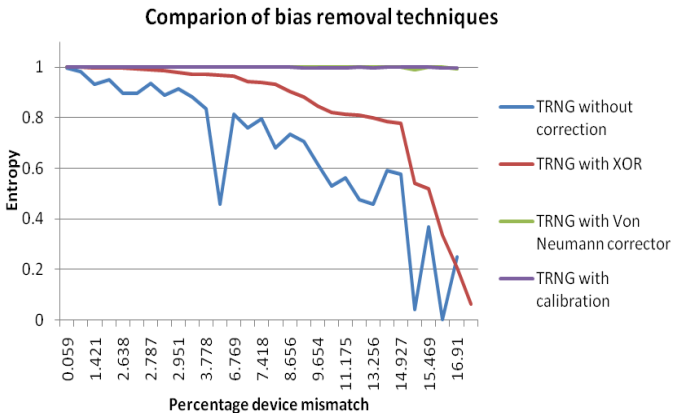
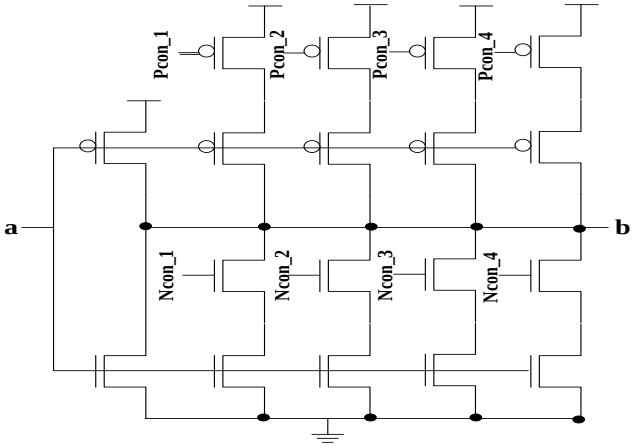
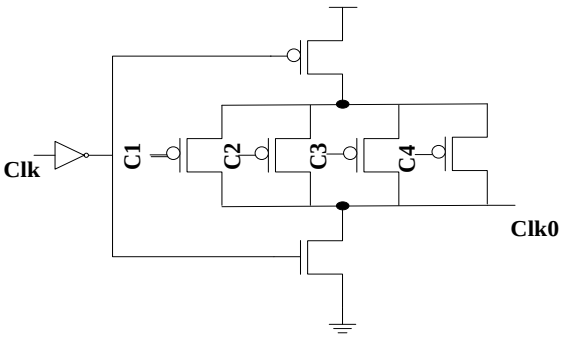


Figure 11: Enhancement in entropy for TRNG with calibration compared to other techniques



(a)



(b)

Figure 10: Calibration circuit (a) Coarse calibration (b) Fine calibration

for the basic TRNG and the different bias removal techniques. Although the XOR function incurs a very small overhead in the form of energy, its inefficiency in improving the entropy with increasing variability does not make it a suitable candidate for usage in encryption systems designed in DSM technologies. The Von Neumann corrector maintains the entropy very close to one but at a significant energy cost. With increase in device mismatch more number of TRNG bits is needed per effective random bit generated. Hence, the energy per random bit increases.

The physical calibration technique provides an effective entropy v/s energy trade off. It provides considerable improvement in randomness similar to the Von Neumann corrector, but at a significantly lower energy per bit. In systems that can tolerate a lower entropy random bit stream, the calibration technique can be implemented with fewer configuration bits and hence lower energy consumption. Fig. 12 and Fig. 13 show the impact of varying the number of configuration bits to the entropy of the circuit and the average energy per bit. For applications operable with lower entropy random bits or implementations not susceptible to large intra-die variations, fewer configuration bits can be used resulting in a reduced energy/bit cost.

Although the calibration technique provides a very flexible solution to trade off between entropy and energy/bit, certain cryptographic applications may need very high entropy random bits. In such scenarios, a combination of circuit calibration and algorithmic post-processing may have to be used to satisfy the need for randomness.

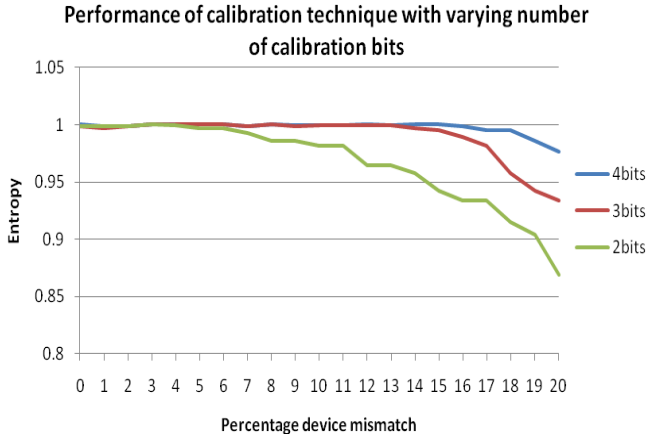


Figure 12: Performance with varying number of configuration bits

A wrong version of this paper was uploaded in the IEEE database.

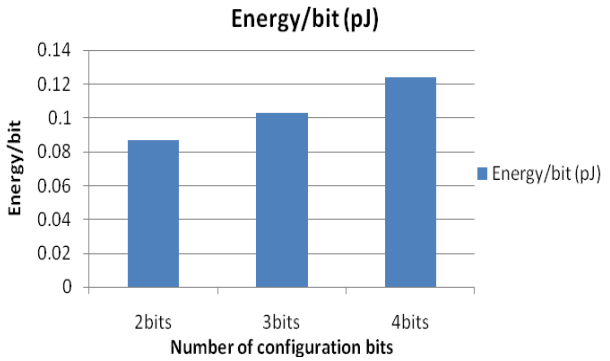


Figure13: Energy/bit with varying number of configuration bits

V. Conclusions and Future Work

Modern security systems need on-chip true random number generators. Applications in microprocessor systems will need design of efficient TRNG circuits in DSM technologies and operating at high frequency. With ever increasing random process variations and susceptibility to variation in operating conditions, robust calibration mechanisms need to be designed to counter bias and correlation. Conventional post-processing techniques are not efficient for simple TRNG, like meta-stability based circuits, for enhancing the entropy in terms of data rate and energy consumption per random bit. Physical calibration techniques are required to tune such TRNG circuits effectively at a lower energy overhead. Such techniques also provide a greater flexibility for trading off entropy for energy, specifically for low power and lightweight applications like RFID. The results and conclusions drawn in this paper can be expected to hold good for other metastability based TRNG designs that are based on similar underlying circuit operation. Apart from countering natural degradation of randomness, the TRNG circuits should also be protected against side channel attacks trying to determine the generated bits or have an invasive effect of biasing the circuit. Design of efficient TRNG circuits and effective calibration mechanisms for the same are proving to be a major challenge for circuit designers.

References

[1] B. Jun and P. Kocher, "The Intel Random Number Generator", in Cryptography Research Inc. White Paper Prepared For Intel Corporation, 1999

[2] B. Sunar, W. J. Martin, and D. R. Stinson, "A Provably Secure True Random Number Generator with Built-In Tolerance to Active Attacks", in IEEE Transactions On Computers, Vol. 56, No. 1, January 2007

[3] V. Fischer, F. Bernard, N. Bochard and M. Varchola, "Enhancing Security of Ring Oscillator-Based TRNG Implemented In FPGA", IEEE, 2008

[4] N. Bochard, F. Bernard and V. Fischer, "Observing the randomness in RO-based TRNG", in International Conference on Reconfigurable Computing and FPGAs, 2009

[5] O. Cret, A. Suciu, T. Gyorfi, "Practical Issues in Implementing TRNGs in FPGAs based on the Ring Oscillator Sampling Method", in 10th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 2008

[6] "Evaluation of VIA C3 Nehemiah Random Number Generator", by Cryptography Research, Inc, 2003

[7] C. Tokunaga, D. Blaauw and Trevor Mudge, "True Random Number Generator With a Metastability-Based Quality Control", in IEEE Journal Of Solid-State Circuits, Vol. 43, No. 1, January 2008

[8] D. E. Holcomb, W. P. Burleson and K. Fu, "Power-Up SRAM State as an Identifying Fingerprint and Source of True Random Numbers", in IEEE Transactions On Computers, Vol. 58, No. 9, September 2009

[9] C. Pyo, S. Pae and G. Lee, "DRAM as source of randomness", in ELECTRONICS LETTERS 1st January 2009 Vol. 45 No. 1

[10] S. Srinivasan, S. Mathew, V. Erraguntla, R. Krishnamurthy, "A 4Gbps 0.57pJ/bit Process-Voltage-Temperature Variation Tolerant All-Digital True Random Number Generator in 45nm CMOS", in 22nd International Conference on VLSI Design, 2009

[11] A. T. Markettos and S. Moore, "The Frequency Injection Attack on Ring-Oscillator-Based True Random Number Generators", in CHES 2009

[12] H. Istvan, A. Suciu, O. Cret, "FPGA based TRNG using automatic calibration", in IEEE 978-1-4244-5007-7, 2009

[13] K. Wold and C. H. Tan, "Analysis and Enhancement of Random Number Generator in FPGA Based on Oscillator Rings", International Conference on Reconfigurable Computing and FPGAs, 2008

[14] R. Santoro, O. Sentieys and S. Roy, "On-the fly evaluation of FPGA-based True Random Number Generator", in IEEE Computer Society Annual Symposium on VLSI, 2009

A wrong version of this paper was uploaded in the IEEE database.