A 4-fJ/b Delay-Hardened Physically Unclonable Function Circuit With Selective Bit Destabilization in 14-nm Trigate CMOS

Sudhir Satpathy, *Member, IEEE,* Sanu K. Mathew, *Senior Member, IEEE,* Vikram Suresh, *Member, IEEE,* Mark A. Anders, *Senior Member, IEEE,* Himanshu Kaul, *Senior Member, IEEE,* Amit Agarwal, *Member, IEEE,* Steven K. Hsu, *Member, IEEE,* Gregory Chen, *Member, IEEE,* Ram K. Krishnamurthy, *Fellow, IEEE,* and Vivek K. De, *Fellow, IEEE*

***Abstract*— This paper describes a full-entropy 128-b key generation platform based on a 1024-b hybrid physically unclonable function (PUF) array, fabricated in 14-nm trigate high-k/metal-gate CMOS. Delay-hardened hybrid PUF cells use differential clock delay insertion to favor circuit evaluation in the desired direction while leveraging burn-in-induced aging for selective bit destabilization enabling quick identification and masking of unstable cells, and subsequent temporal-majority-voting with soft dark-bit masking to reduce PUF bit error by 3.9 times to 1.45% resulting in ∼5 ppb failure probability. A stable full-entropy 128-b key is finally generated from the 1024 raw PUF bits using BCH error correction and AES-CBC-based entropy extraction. An all-digital design with compact PUF cell layout occupying 1.84 μm^2 achieves: 1) 4-fJ/b energy-efficiency with 3-μW leakage at 0.65 V, 70 °C; 2) peak operating frequency of 1 GHz resulting in 1.2-μs key generation latency; 3) robust operation with stable key generation across 0.55–0.75 V, and 25 °C–110 °C; 4) 14 times separation between intra/inter-PUF hamming distances with 0.99993 entropy ensuring cryptographic quality randomness and uniqueness; 5) 48% higher PUF stability with long-term aging by leveraging transistor degradation to reinforce favorable cell bias; and 6) resiliency to power cycling attacks with common centroid clock routing measured from 49.5% hamming distance between array's evaluation and wake-up states.**

***Index Terms*— Bit error, delay hardening, entropy extraction, key generation, physically unclonable function (PUF), soft dark-bit (DB) masking, selective bit destabilization.**

I. Introduction

PHYSICALLY unclonable functions (PUFs) constitute a key building block of modern security platforms owing to their ability to harvest manufacturing and process-induced variations for generating an embedded secret that is easy to verify but extremely challenging to predict [1]–[6]. Recently, PUFs have emerged as a strong contender as low-cost cryptographic primitive for generating a stable, repeatable device-specific signature [7]–[11]. This signature forms the foundation for many applications, such as key generation/storage, device anticounterfeiting, user authentication, IP protection, and hardware/software binding. Traditional one time programmable (OTP) fuse-based key storage technologies are prone to invasive probing attacks where an attacker decapsulates the chip and exposes internal circuit nodes to observe the stored key. The key stored in such fuses is also prone to manipulation on the manufacturing floor. In contrast, PUF-based approaches eliminate such possibilities by delaying generation of key until its time of consumption. They enable significant die cost reduction by eliminating additional manufacturing steps associated with fabricating fuses. In addition, absence of a copy of the key with the device manufacturer also alleviates counterfeiting risks. The volatile nature of PUFs provides a high-level of security and tamper resistance by raising the security bar versus fuses, and improves designer–manufacturer trust by enabling secret generation without any manual intervention. PUFs, thus represent a paradigm shift in physical security by transitioning from the conventional approach of explicitly programming digital IDs into fuses postmanufacturing, to a new approach of generating the IDs by exploiting the intrinsic characteristic of devices on the die [12]–[17]. Variation in these intrinsic characteristics during device manufacturing manifests as deviation in important transistor parameters such as threshold voltages that can alter circuit behavior. Although these variations are statistically characterized by device manufacturers, it is practically infeasible to accurately predict their manifestation in any unique sample. Furthermore, design choices to optimize PUF circuits usually lead to the exacerbation of the impact of these variations on circuit functionality, making it increasingly more difficult for an attacker to extract the key [18]–[21].

PUFs exploit die-to-die process variation to generate static entropy with sufficient repeatability in the face of temporal voltage, and temperature fluctuations and aging. A variety of PUF circuits based on bias generation, ring-oscillators,

Manuscript received August 8, 2016; revised October 14, 2016; accepted November 26, 2016. Date of publication January 23, 2017; date of current version March 23, 2017. This paper was approved by Guest Editor Makoto Ikeda.

The authors are with the Circuit Research Laboratory, Intel Corporation, Hillsboro, OR 97124 USA (e-mail: sudhir.k.satpathy@intel.com; sanu.k.mathew@intel.com; vikram.b.suresh@intel.com; mark.a.anders@intel.com; himanshu.kaul@intel.com; steven.k.hsu@intel.com; amit1.agarwal@intel.com; gregory.k.chen@intel.com; ram.krishnamurthy@intel.com; vivek.de@intel.com).

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/JSSC.2016.2636859

current mirrors, cross-coupled inverters, oxide breakdown, and SRAMs have been proposed in recent literature. Bias generator-based PUFs amplify the impact of random transistor threshold voltage variation by comparing the output voltages of a pair of identically designed proportional to absolute temperature generators [1]. This approach requires carefully designed comparators with postfabrication offset cancellation techniques to accurately generate the PUF output in the presence of temperature and voltage fluctuations. Oscillation collapse in a dual edge injected ring oscillator is another approach where the accumulating delay difference between two travelling edges determines the final settling voltage of the internal nodes, thus generating a PUF bit [2]. Such PUFs provide a promising solution for applications such as user authentication where device disambiguation can still be accomplished with minor deviations from the original PUF output by using multiple challenge-response pairs. However, they are not suitable for applications involving key generation where a PUF value has to be repeatedly created with 100% accuracy. Complementary current mirror-based monostable circuit is another recently proposed approach for PUF implementation [4]. This fully static design provides resiliency to environmental variations because of the absence of dynamic switching events, at the expense of higher standby power owing to bias currents. Hybrid variants of classical designs based on arbiter/delay chains and SRAM cells have also been recently used for PUF implementation [10]. These circuits exploit metastability resolution dynamics of matching cross-coupled inverters along with delay variations in clock paths to generate static entropy in the PUF bits. Techniques described earlier leverage random variation in transistor threshold voltages, and interconnect mismatch-induced delay variation as two common sources of entropy. Other sources of uncertainty such as time to die-electric breakdown, and capacitance variation in eDRAM can also be exploited to harvest static entropy. In [9], an approach of subjecting devices to elevated voltage stress at high temperature to randomly break the gate dielectric oxide of an array of transistors to generate an ID was proposed. In [7], the retention times of individual cells in an eDRAM array are used to generate a fail map based on a threshold. This bit fail map exhibits sufficient degree of randomness and stability for usage as Chip ID. Though NIST randomness and hamming distance tests qualify both of these approaches for PUF key generation, the need for multiple supply voltages and nonstandard manufacturing steps makes them less attractive for high volume manufacturing.

Raw IDs derived directly from the above listed PUF circuits are not inherently 100% stable. Insufficient device mismatch, high thermal noise, and changes in operating conditions can lead to high degree of instability in raw PUF bits. A key generation scheme must be resilient to changes in voltage, temperature, and device aging conditions to guarantee reliable operation of security oriented applications over the product life-time [22]–[25]. Schemes that reduce this inherent key instability usually involve conditioning steps to improve the stability of PUF cells and isolate easy to identify unstable cells. The double edge injected ring oscillator PUF uses the time duration for the initial oscillation to collapse as a metric to identify unstable bits [2]. In eDRAM PUFs, comparing the retention time against a lower and an upper bound can be used to indicate instability. In hybrid PUFs, repeated circuit evaluation serves as a simple and effective way to sequester unstable cells, also referred to as dark bits (DBs) that are corrected by applying a mask. However, the volatile nature of unstable cells makes it difficult to identify them accurately under varying operating conditions, with uncertainty in DB masking accounting for up to 95% of final PUF bit error [12]. Hence, techniques to improve the stability of DB mask generation hold promise for significant reduction in bit error.

In contrast to prior work that is limited to only identifying unstable bits and generating a mask to isolate them, we describe a method to improve the probability of accurately recreating the mask under varying operating conditions. Stable bits are further stabilized by reinforcing their existing differential bias, while cells contributing to bit error are conditioned toward higher instability to ensure easy identification and masking. This is accomplished using a selective bit destabilization technique that is applied in conjunction with the burn-in-based stabilization scheme. This is the first reported work that demonstrates significant improvement in the stability of PUF key by opportunistically destabilizing certain cells within the PUF array. The destabilization scheme is applied concurrently along-side stabilization techniques using existing silicon infrastructure without impacting the array's performance, area, and energy-efficiency. In addition, a delay hardened PUF circuit is proposed to maximize the impact of this concurrent stabilization and selective destabilization technique. A 1024-b PUF array fabricated in 14-nm trigate CMOS demonstrates the reliability of this approach by reducing overall bit error to 1.4% across 25 °C–110 °C over 200 mV supply fluctuation, resulting in a 100% stable and secure 128-b final key generated using subsequent BCH error correction and AES-CBC-based entropy extraction.

The remainder of this paper is organized as follows. The key generation datapath microarchitecture that includes the 1024-b PUF array, conditioning circuits, and other building blocks is presented in Section II. Section III explains the concept of delay-hardening and describes the PUF circuit aging phenomenon that aids array stability. In Section IV, we introduce the selective destabilization technique and show measurement results to demonstrate its efficacy in further improving overall PUF stability. Impact of long term aging and random telegraphic noise on PUF behavior is discussed in Section V. Measurement results showcasing the resiliency of the proposed circuit against power-ON attacks is presented in Section VI, and finally, we summarize this paper in Section VII.

II. PUF Datapath Organization

The key generation datapath consists of a PUF array that serves as the source of raw entropy, conditioning circuits that compute temporal majority voting (TMV) and DB mask for excluding unstable bits, and an error correction (ECC) unit to rectify left over instability, followed by AES-CBC-MAC-based entropy extraction to derive the

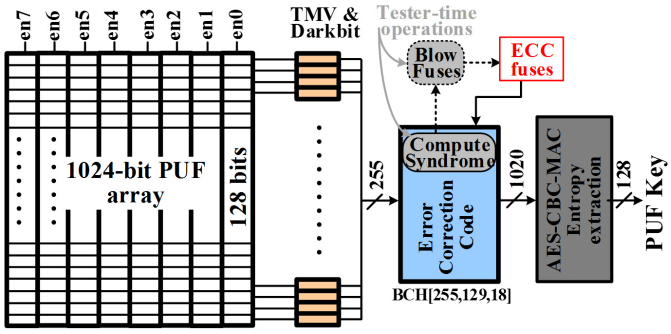


Fig. 1. PUF key generator organization.

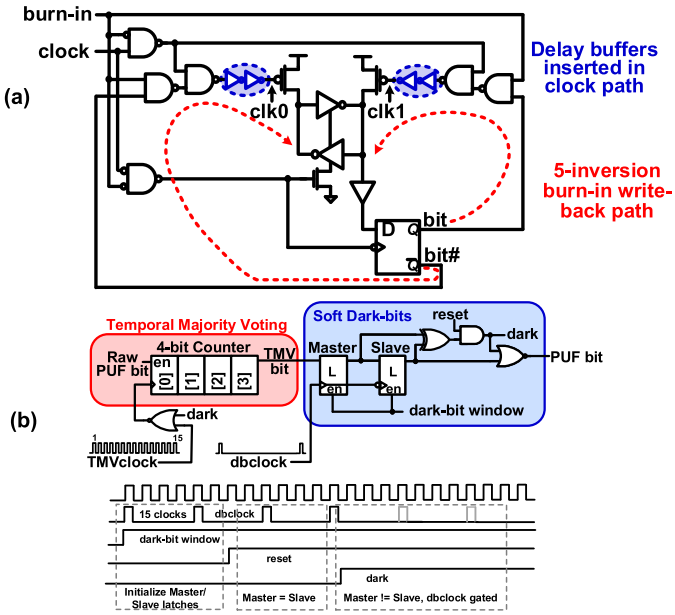


Fig. 2. (a) Delay-hardened PUF circuit. (b) Temporal majority voting and soft DB masking conditioning circuit.

final key (Fig. 1). The 1024-b PUF is organized as eight-entry column-multiplexed array of 128 b with one column read every cycle that subsequently undergoes temporal-majority voting and DB masking. In the first evaluation of this array during tester-time operation, a golden key is generated. An ECC signature, computed from this golden key and stored on-die using OTP fuses, is used to recreate the original key with 100% accuracy in subsequent in-field evaluations. A PUF circuit that produces an output bit with a high degree of stability is desirable to minimize ECC overheads.

The PUF array is based off of a hybrid cross-coupled inverter circuit with a pair of precharge transistors that initialize the internal nodes to an unstable state [Fig. 2(a)]. During the positive phase of the clock, the circuit evaluates and snaps to one of two stable states. Resolution toward a stable voltage from the initial unstable state is determined by the relative strengths of variation-impacted minimum sized inverters in the cross couple. Additionally, random variations in precharge transistors and devices along clock path produce mismatches in clock rise and arrival times introducing a transient dimension of uncertainty into PUF resolution dynamics. A flip-flop captures the evaluated PUF output at the negative clock edge. Since, the PUF value is not directly read from the cross couple, this allows gating the clock to precharge the cross-coupled inverter nodes back to VCC until the next evaluation phase. In its precharged state, the corresponding transistors in either legs of the PUF circuit are subjected to identical voltage conditions enable isoaging of the circuit, thereby limiting long-term stability degradation.

PUF cells that have insufficient within-die random variation generate unstable bits that resolve to "0" or "1" based on thermal noise, or voltage and temperature conditions. Such noisy bits undergo two conditioning steps to reduce overall PUF bit-error rate. The first conditioning circuit, called temporal majority voter (TMV) computes the quantized mean of PUF responses within a voting window. This is implemented using a 4-b counter that tracks the number of times a cell that evaluates to "1" during a voting window of 15 cycles [Fig. 2(b)]. Stable cells that result in count values above 8 are considered to be 1's and those with counts 7 or below are counted as 0's. Mildly unstable bits with error probabilities <8% use TMV-15 to reduce their effective error rate to 10^{-6}. Cells with counts around the threshold of 6–9 represent highly unstable bits that are unaffected by TMV, and require a secondary conditioning scheme. Most of these bits are identified as those that change values at least once within a DB window of 100 TMV evaluations (1500 clock cycles). DB window is defined as the number of times the PUF array is consecutively evaluated to identify an unstable cell that flips at least once. The change in value is detected by comparing master and slave latch values at the end of each TMV cycle. The presence of complementary values in these latches indicates a bit that was unstable even after TMV. This bit is marked as a DB and masked off, thus excluding it from participating in key generation. In contrast to the conventional approach of storing DB locations as a hard mask in NVM or fuses, a soft masking scheme is used that does not require any explicit mask storage. This not only saves area, but also provides higher degree of key security by removing a potential tamper-point from an attacker's access domain. This volatile mask is generated at every start-up to identify and exclude unstable cells.

TMV and DB condition circuits enable an almost stable PUF value with a final bit error measuring 1.4%. These residual "hard to stabilize" and "hard to identify" cells are rectified using error-correction code (ECC). Given an array of "N" cells where each cell has a failure probability of "P," an ECC logic that can fix at most "K" errors results in an overall probability of correct key generation, $S = \sum_{x=0}^{K} (N/x) P^x (1-P)^{N-x}$. Fig. 3(b) shows a plot of overall error probabilities for the corresponding number of maximum erroneous bits that ECC can correct with $N = 255$, and $P = 0.014$. A target failure rate of ~5 ppb (5 part in 10^9 samples) requires the capability to correct a maximum of 18 out of 255 b. Hence, BCH (255, 131, 18) code that can correct up to 18 errors in 255 b was selected for this design. During enrollment on the tester, the syndrome generator computes a 129-b syndrome tag (124-b ECC helper data +1-b parity +4-b hash) that is stored in on-die OTP fuses [Fig. 3(a)]. This tag is subsequently accessed and XORed with the raw

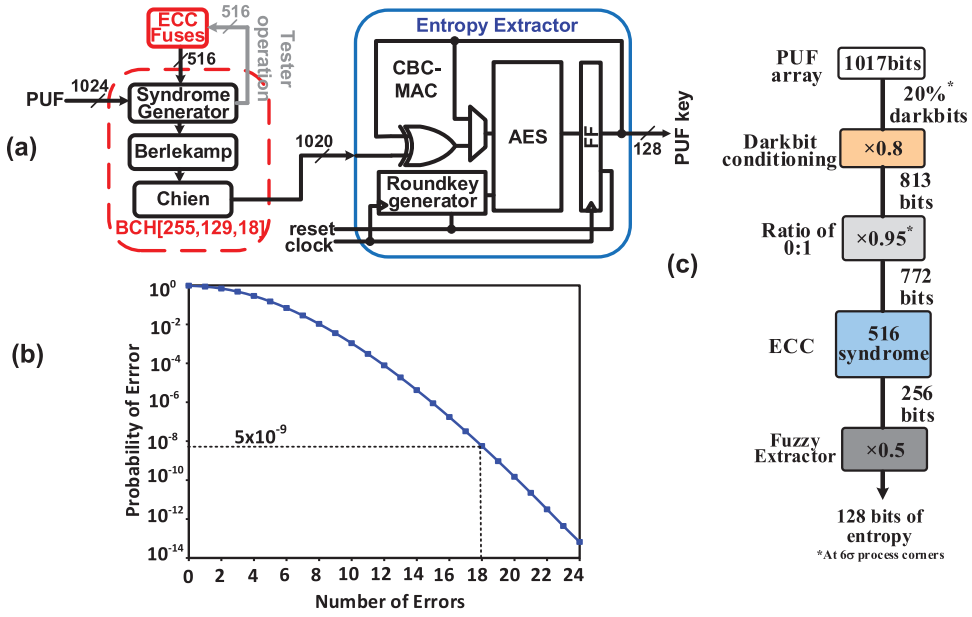


Fig. 3. (a) Error correction and AES CBC-MAC-based entropy extraction. (b) Key failure probability versus number of erroneous bits (out of 255) corrected with BCH code. (c) Chain entropy loss.

PUF values followed by syndrome regeneration, Berlekamp computation and Chien's search during in-field operation, resulting in a latency of 255 cycles to recover 255 PUF bits. This correction step is repeated four times to generate the 1024-b PUF value. At the end of ECC, the golden PUF value is regenerated with 100% accuracy. Corrected PUF values are sent to the entropy extractor, where an AES-CBC-MAC circuit generates a 128-b full-entropy key at the end of 80 cycles.

Entropy leakage occurs at each step of PUF conditioning and postprocessing [Fig. 3(c)]. To guarantee a minimum level of entropy/bit at the output, these losses have to be accounted for to ensure that the PUF array feeds sufficient amount of entropy into the chain. The goal of this design is to generate a full entropy PUF key with 1 b of entropy/key-bit, thus resulting in 128 b of entropy, uniformly distributed among the output bits. To generate such an output, the left-over hash lemma requires atleast twice the amount of entropy at the input of the AES-CBC-MAC-based extractor [26]. These 256 b of left-over entropy after ECC processing could be nonuniformly distributed among the input bits. The BCH (255, 131, 18) code requires 516 (129 × 4) syndrome bits stored in nonvolatile OTP fuses. Since these fuses are accessible to a malicious attacker, the entropy leaked in the 516-b syndrome needs to be discounted. Accounting for this leakage, upstream entropy requirements at ECC input increases to 772 b. To account for nonideal ratios of 0's and 1's due to systematic biases in the PUF array, an additional 5% margin is incorporated, increasing entropy requirement to 813 b. Finally, masking of DBs to a predetermined value destroys the entropy contained in those bits. A worst case scenario that masks 20% of total bits, requires a PUF array with at least 1017 b to guarantee a full-entropy key of 128 b.

III. Stability Improvement With Delay-Hardening

The hybrid PUF cell used to construct the 1024-b entropy source array combines the stability and compactness properties

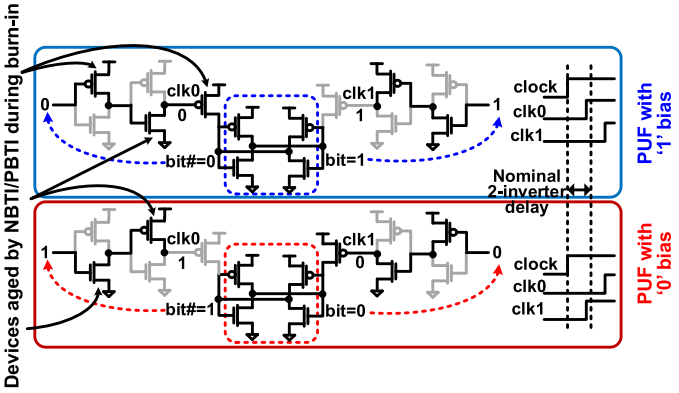


Fig. 4. Clock bias injection using delay-hardening during burn-in.

of SRAM-based PUFs, with added resiliency to invasive probing attacks that delay-based PUFs offer [9]. In contrast to prior approaches, the delay hardened PUF introduces two additional clock delay inverters (Fig. 4) into each precharge path (clk0 and clk1). These minimum-sized inverters introduce additional variation into PUF metastability dynamics, resulting in a maximum BER of 5.76%, a 15% reduction over the basic hybrid cell. However, a more significant impact of the clock delay buffers on the PUF cell operation is observed during burn-in. Burn-in is a process that uses accelerated aging of semiconductor devices to eliminate early failures by operating the chip at elevated temperature and voltage conditions. During accelerated aging (burn-in = 1), the complementary value (bit#) is differentially written back into the cell, biasing inverter devices in a direction such that NBTI/PBTI aging reinforces preexisting biases and improves cell stability. The clock delay inverters in the delay hardened cell are also differentially biased during burn-in such that a cell with an initial bias toward "1" self-biases and ages the appropriate clock devices to push out clk1 rise delays, while pulling in clk0 rising edges relative to a nominal two inverter delay (Fig. 4).

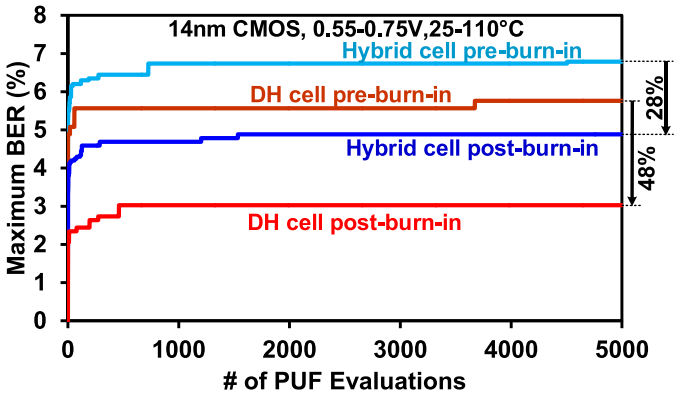


Fig. 5. Hybrid versus delay-hardened PUF BER measurement.

In contrast, a cell with a preexisting bias toward "0" pushes out clk0 delays relative to clk1. This extra bias reinforcement in the clock delay paths in addition to the directed aging of devices in the cross couple enables better postburn-in cell stability. Both the hybrid and delay-hardened PUF arrays were subjected to burn-in hardening at elevated supply voltage and temperature for 2 h. Measurements comparing preburn-in and postburn-in bit-error rates show that injection of differential clock bias is more effective in the delay-hardened cell that provides a 48% reduction in bit error from 5.76% to 3.03%, as opposed to the hybrid cell that reduces BER by 28% (Fig. 5). BER measurement involves evaluating the PUF array 5000 times across a supply and temperature range of 0.55–0.75 V, and 25 °C–110 °C, and comparing it against a golden value generated at 0.65 V and 70 °C to compute the percentage of bits that differ in the worst case scenario. TMV followed by DB masking with an evaluation window of 100 cycles further reduces final BER to 2.35%, providing net 59% BER reduction over the raw PUF BER. Achieving the target key failure probability of ∼5 ppb with a BER of 2.35% requires a BCH code with a 169-b syndrome to correct a maximum of 24 errors among 255 b. A lower BER requires a smaller ECC syndrome resulting in reduced fuse area overhead. Hence, further stabilization of the PUF array is necessary to make this solution attractive for area constrained applications.

IV. Selective Bit Destabilization

Analysis of the unstable cells that contribute to post-TMV and DB conditioning 2.35% BER reveals that 92% of the BER can be attributed to mask mismatch. A preburn-in mask is generated at the very first evaluation of the PUF array during tester operation under golden conditions of 0.65 V and 70 °C. Although this mask is not physically stored on-die, it is used to compute the golden PUF value and the corresponding ECC syndrome that is stored in fuses. Every subsequent PUF evaluation during run-time operations in the field requires the regeneration of a fresh DB mask. This mask is created at conditions varying from 0.55–0.75 V and temperature across 25 °C–110 °C. DB measurements on the 14-nm PUF array with 100 cycle evaluation window identifies 52 out of 1024 b as dark [Fig. 6(a)]. These 52 b evaluated to a different value at least once in 100 evaluations. However, only 32 of these bits overlap with the mask generated previously at golden conditions. Furthermore, 24 b marked as dark in the golden mask remain unidentified in the field mask. These 44 disjoint mask bits are forced to "0" during gold or field conditions, but not both, thus contributing to bit error. Measured DB-induced BER accounts for 92% of total BER (2.15% out of 2.35%). The speckle pattern in Fig. 6(a) shows the locations of all DBs in the 1024-b array laid out in a 16 × 64 grid. The "easy to identify" 32 b that are dark at both field and gold condition result in a net reduction of BER, while the remaining 20 and 24 b that show up as dark in one mask but not the other and are "hard to identify" worsen BER.

TMV conditioning not only enables the stabilization of mildly unstable PUF cells, but also provides a metric to quantify the degree of instability of all cells within the PUF array. Hence, examining the TMV counts allows characterization of DBs to analyze their behavior in field and golden conditions. Fig. 7 shows histogram plots of TMV Counts for all 1024 b in the PUF array after 75 000 evaluations expressed as a percentage of the maximum TMV count. The 948 stable bits exhibit TMV counts consistently in the range of 0%–10% or 90%–100% as shown in Fig. 7(a). DBs on the other hand are scattered across the entire range of 0%–100%. The desirable "easy to identify" DBs that are unstable at both field and golden conditions have TMV counts clustered in a close range centered around 50%. A separate plot [Fig. 7(b)] shows the TMV count spread for "hard to identify" disjoint DBs that are dark only at one of field or golden conditions, but not both. The TMV counts for these bits span the entire range of 0%–100%. Prior burn-in hardening techniques attempt to stabilize all cells in the PUF array by reinforcing existing bias [10] in all cells with TMV count <50% toward "0" and the remaining toward "1" improving overall BER by hardening mildly unstable cells. However, some of the highly unstable cells get partially stabilized making them hard to detect during DB masking, detrimentally affecting net BER. As a solution, we propose a TMV count-based winnowing technique that selectively hardens each cell toward higher or lower stability to improve overall BER. All cells with TMV counts ranging from 0%–10% to 90%–100% were stabilized by writing back the complementary PUF value into the cross couple, while the remaining cells with TMV counts from 10% to 90% were destabilized by writing back the original value. This approach destabilizes 80% of all DBs, thereby increasing their probability of remaining dark at both field and gold conditions and reducing mask mismatch-induced bit error.

Fig. 6(b) compares raw and postburn-in measurements of the delay-hardened PUF array. Overlap in DB masks constitutes 40% of the entire DB population with selective destabilization. In contrast burn-in without this TMV-based destabilization results in 2.4 times reduction in mask overlap, with only 16% mask overlap between field and gold conditions [Fig. 6(a)]. The TMV-based destabilization of DBs reduces mask mismatch-induced BER to 1.07%, a 50% reduction compared with the preburn-in values. Along with conditioning, this approach reduces PUF BER to 1.46%, a 3.9 times improvement over raw bit error. The lower BER provides 31% savings in fuse area by enabling ECC with a 129-b syndrome

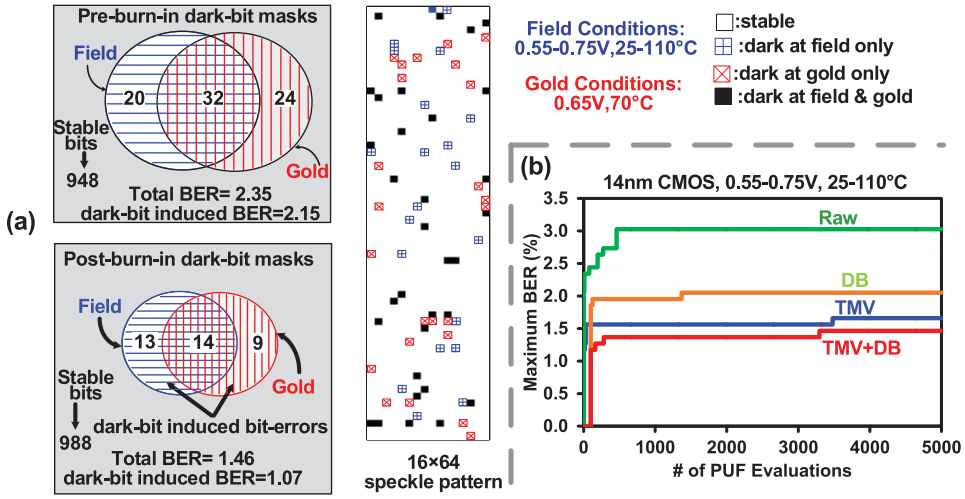


Fig. 6. (a) 1024-b PUF array DB mask distributions, and speckle pattern. (b) BER measurements after TMV and DB conditioning.

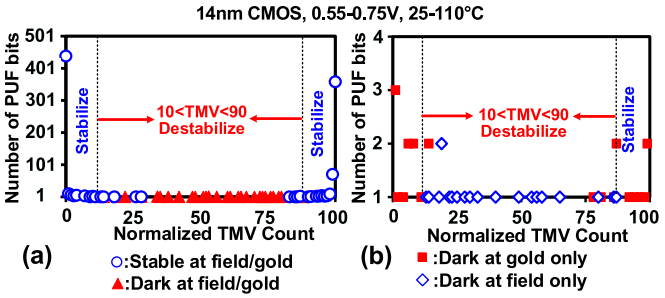


Fig. 7. TMV count distribution of 1024 PUF array bits.

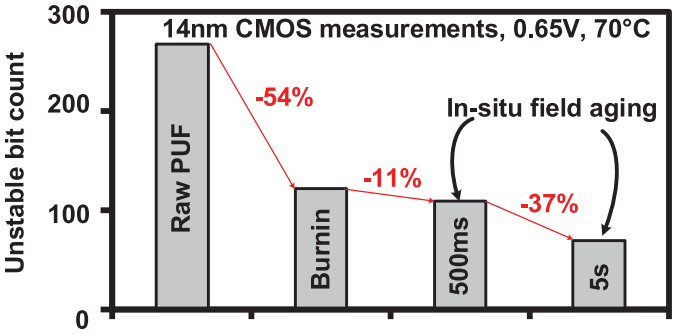


Fig. 8. Aging effect on PUF stability.

in contrast to the prior 169-b syndrome required to achieve a target 5 ppb key failure probability.

Stability analysis using a conventional hard DB masking scheme shows that PUF BER reduces from 1.4% to 0.4% when the mask generated during enrollment is stored in NVM and subsequently used during in-field operation, thereby eliminating mask mismatch-induced bit error. This reduces ECC syndrome tag to 84 b (from 129 b for soft-masking) to correct 11 errors (instead of 18 for soft masking) to achieve the target 5 ppb failure rate. A smaller syndrome leads to lower entropy leakage and reduces the PUF array size to 778 b (down from 1024 b). However, this scheme would require a total of 1886 b of NVM to store the syndrome and DB locations as opposed to the current scheme that requires 516 b for the PUF array. Analysis of total storage requirement including the PUF array shows that the proposed soft-masking approach provides 43% area savings over the conventional hard-masking scheme at identical key failure probability.

V. Aging and Random Telegraphic Noise

PUF stability characterization with long-term aging is necessary to guarantee reliable operation over a product life cycle. To emulate aging behavior, we subjected the array to high voltage stress for durations ranging from 500 ms to 5 s. Fig. 8 compares unstable bit count under different measurement conditions. TMV-based delay-hardening with selective destabilization reduces unstable bit count by 54% over raw measurements. Further improvements in stability are achieved by taking advantage of the write-back paths available in the cell. Enabling these write-back paths during field operation of the PUF array selectively switches ON and OFF devices for asymmetric degradation within the cross couple and clock paths, providing favorable *in situ* field aging that reduces unstable bit counts by up to 48%. This technique improves PUF array stability over the lifetime of the die, thus guaranteeing reliable operation resulting in ∼5 ppb key failure probability with ECC guard bands to correct a 7% worst case BER.

Instability analysis provides insight to understand the impact of random telegraphic noise on PUF behavior. The PUF array is evaluated 200,000 times with its output sampled every 400 cycles at frequencies ranging from 2 GHz to 100 MHz. Unstable bit count increases by four times at frequencies below 800 MHz [Fig. 9(a)] due to increasing influence of $1/f$ noise as PUF devices remain in precharged state for longer durations. Unstable PUF cells are most sensitive to temporary threshold voltage shifts, and are excellent candidates to observe and quantify the magnified impact of $1/f$ noise. Bit stability degradation mechanisms are further examined using 500-point FFT spectral analysis of 167 most unstable bits from multiple tiles indicating ten times higher bit transitions, with 9.3 times increase in mean-square FFT magnitude below 500 MHz [Fig. 9(b). FFT magnitude in the plot indicates the sum of energy contained in all nonzero frequency components exhibited in PUF bit evaluation pattern.

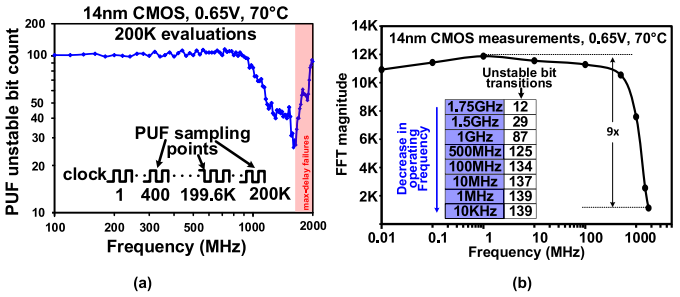


Fig. 9. Measurements showing effect of $1/f$ noise.

Low frequency operation increases probability of device threshold voltage shift, which results in higher bit flips during PUF evaluations. To account for the impacts of $1/f$ noise on PUF cell behavior, all BER metrics reported in this paper are measured at 100 MHz.

VI. Key Uniqueness and Resiliency to Probing Attack

Cryptographic quality randomness/uniqueness of the key along-with stability is necessary for security applications. Measured intra-PUF and inter-PUF peak hamming distances of 34.8 and 498 b, respectively, in the 1024-b array measured over 5000 samples indicate 14 times separation (8.2 times worst case separation) guaranteeing key uniqueness [Fig. 10(a)]. The distribution of run lengths of "1"s and "0"s for counts from 1 to 20 [Fig. 10(b)] over multiple evaluations of the PUF array match the expected theoretical distribution obtained from a random stream without any TMV/DB/entropy-extraction conditioning. Furthermore, 0.99993 measured entropy ensures that generated PUF bits are sufficiently random without bias toward any particular pattern. The PUF cell dissipates 3 μW of leakage power with an energy efficiency of 4 fJ/b measured at 0.65 V, 70 °C while operating at 1 GHz, with a layout occupying 1.84 μm^2 making this an excellent candidate for energy and area constrained applications.

The PUF resolution dynamics in the proposed design are determined not only by transistor threshold voltage mismatches in the cross-coupled inverters, but also due to variations in clock rise and arrival times at the precharge transistor nodes owing to differences in their propagation path delays. This complex dependency because of clock related sources of entropy makes the circuit resilient to dc probing and device parameter extraction attacks. Furthermore, mirrored layout of the PUF circuit with common centroid balanced clock routing at upper metal layer ensures that access to the common clock source node is disrupted in an event when an attacker tries to evaluate the circuit by grinding down the die to probe the internal nodes.

The impact of these clock related transients on overall circuit operation was examined by comparing the PUF array's run-time output against its wake-up state. The wake-up state was evaluated by disabling the clock nodes, and subsequently power cycling the array to read out its resolved state. In this mode of operation, the PUF internal nodes settle down to voltages that are determined by the relative strengths

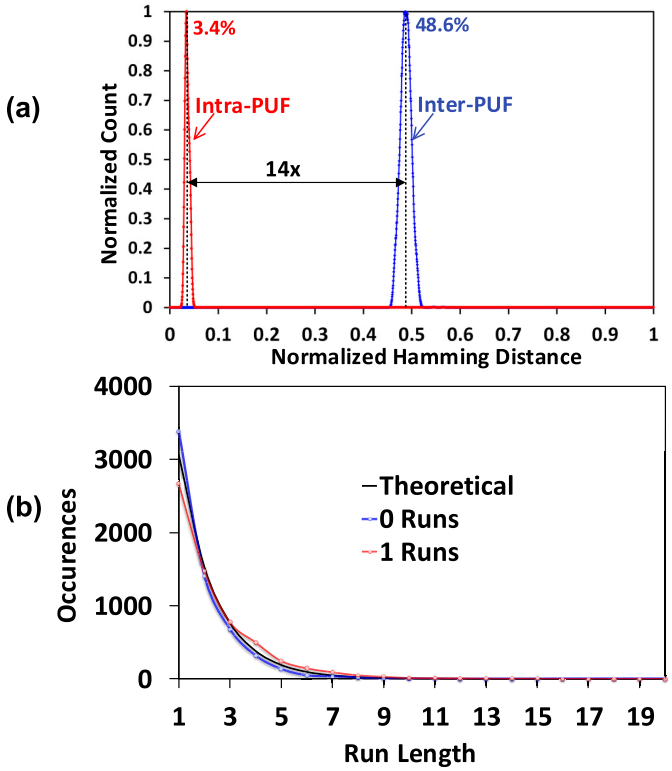


Fig. 10. (a) Measured inter-PUF and intra-PUF hamming distances. (b) Run lengths of "1"s and "0"s in measured PUF bitstream.

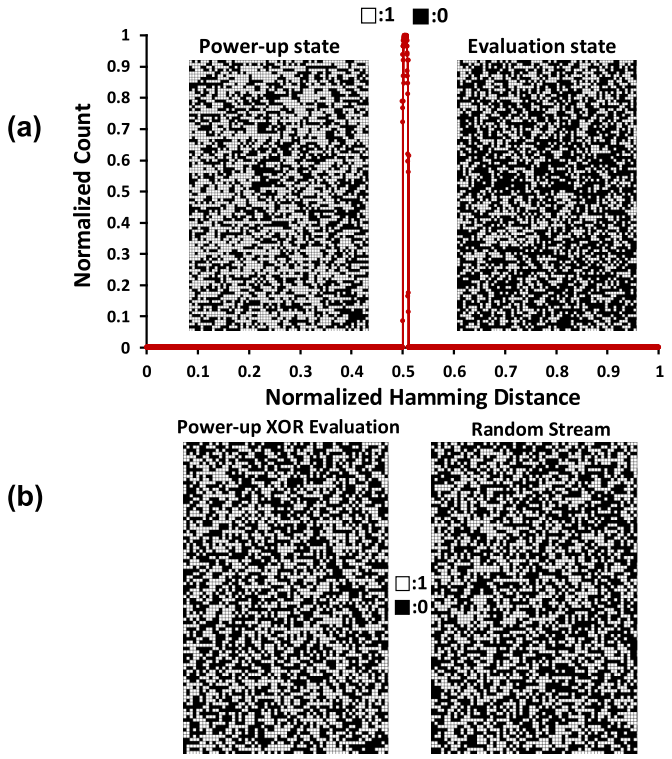
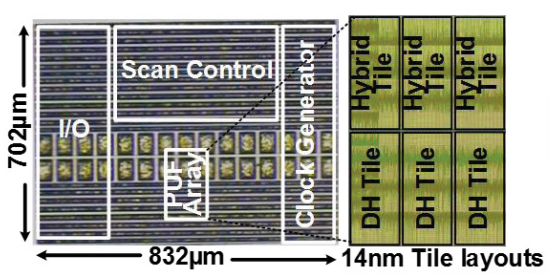
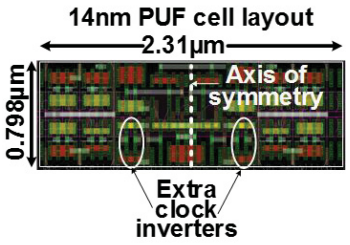


Fig. 11. (a) Hamming distance separation and speckle patterns of PUF evaluation and power-up states. (b) Speckle pattern showing mismatching bit locations and random stream.

of variation impacted legs in the cross couple. The speckle patterns in Fig. 11(a) shows 6144 PUF bits (obtained from six arrays) while operating in these two contrasting modes.





Process	14nm tri-gate CMOS
Die area	0.58mm^2
Cell area	1.84µm^2
Array size	64x16

Energy/entropy measurements

Energy/bit	4fJ, 0.65V,70°C
Entropy	0.99993
Leakage power	3µW, 0.65V,70°C

BER measurements

Raw	5.76
After Burn-in	3.03
With TMV15 only	1.66
With Darkbits only	2.05
TMV15 & Darkbits	1.46

Fig. 12. 14-nm die micrograph, layout and measurement summary.

	[10]	[4]	[1]	[3]	This work
Technology	22nm	65nm	65nm	45nm	14nm
Topology	Hybrid cell	Current mirror	PTAT cell	Biased NAND	Delay hardened Hybrid cell
Bitcell size(um^2)	4.7	3.42	3.07	5.30	1.84
VDD(V)	0.7-0.9	1	0.6-1.2	-	0.55-0.75
Temperature(^{0}C)	25-50	25-85	0-80	-40-85	25-110
BER	0.97	1.73	2%	1.69%	1.46%
Energy/bit (fJ)	13	15	1100	-	4

Fig. 13. Comparison with prior work.

Hamming distance measurement between the two states showcase their uncorrelated behavior as indicated by 49.5% of the bits resolving to opposite values. This affirms the significant role of clock transients in PUF circuit behavior, and hence its resiliency to dc probing.

Location of mismatching bits within the PUF array was further analyzed by XORing the two PUF states. In Fig. 11(b), the matching and differing PUF cells are indicated by white and black squares, and the resulting pattern is shown alongside a randomly generated stream of "1"s and "0"s. Similarity between the two speckle patterns indicate that clock transient-induced entropy is evenly distributed across the entire array, making it extremely challenging to predict the PUF's in-field run-time behavior from its power-up state.

Relative contribution of addition of extra inverters in the delay-hardened circuit toward probing resiliency was further explored using similar analysis on the basic and delay-hardened arrays. Hamming distance comparison of both array's run-time evaluation values against their corresponding wake-up states indicate that addition of the inverter pair improves average hamming distance separation measured over 5000 samples by 9% from 45.5% to 49.5%.

VII. Conclusion

The design of a 1024-b PUF array fabricated in 14-nm trigate high-*k*/metal-gate CMOS [27] targeted for low-latency energy-efficient secure key generation has been described. The array is organized around a delay-hardened hybrid PUF circuit that harvests static entropy from manufacturing variations to generate a stable and secure 128-b key. Burn-in-induced aging with preferential bias reinforcing write-back and selective destabilization scheme, temporal-majority-voting (TMV), soft DB masking, BCH error correction, and AES-CBC-MAC-based entropy extraction result in 100% stable full-entropy 128-b key generation from the 1024 raw PUF bits. These techniques reduce overall PUF bit error by 3.9 times down to 1.46% with a compact cell layout occupying 1.84 μm^2. (Fig. 12). The design achieves 4-fJ/b energy-efficiency (3.2 times higher than prior work), dissipating 6.06 pJ for generating the 128-b key (1515 evaluations of PUF array) with 3-μW leakage measured at nominal supply of 0.65 V, 70 °C with peak operating frequency of 1 GHz (Fig. 13). Robust operation with stable unique key generation is demonstrated across a supply variation of 0.55–0.75 V with temperature across 25 °C to 110 °C. Directed *in situ* field aging provides 48% higher long-term PUF stability by leveraging transistor degradation to reinforce favorable cell bias. Resiliency to power cycling attacks with common centroid routing and clock delay-dependent PUF resolution dynamics is demonstrated with 49.5% hamming distance separation between evaluated key and the PUF array's wake-up state.

Acknowledgment

The authors would like to thank M. Haycock, M. Mayberry, J. Tschanz, S. Iyengar, A. Rajan, R. Parker, and P. Munguia for their valuable discussions and encouragement.

REFERENCES

[1] J. Li and M. Seok, "A 3.07μm^2/bitcell physically unclonable function with 3.5% and 1% bit-instability across 0 to 80°C and 0.6 to 1.2V in a 65nm CMOS," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2015, pp. C250–C251.

[2] K. Yang, Q. Dong, D. Blaauw, and D. Sylvester, "A physically unclonable function with BER $<10^{-8}$ for robust chip authentication using oscillator collapse in 40nm CMOS," in *ISSCC Dig. Tech. Papers*, Feb. 2015, pp. 1–3.

[3] B. Karpinskyy, Y. Lee, Y. Choi, Y. Kim, M. Noh, and S. Lee, "Physically unclonable function for secure key generation with a key error rate of 2E-38 in 45nm smart-card chips," in *ISSCC Dig. Tech. Papers*, Jan. 2016, pp. 158–159.

[4] A. Alvarez, W. Zhao, and M. Alioto, "15fJ/b static physically unclonable functions for secure chip identification with <2% native bit instability and 140× Inter/Intra PUF Hamming distance separation in 65nm," in *ISSCC Dig. Tech. Papers*, Feb. 2015, pp. 1–3.

[5] S. Stanzione, D. Puntin, and G. Iannaccone, "CMOS silicon physical unclonable functions based on intrinsic process variability," *IEEE J. Solid-State Circuits*, vol. 46, no. 6, pp. 1456–1463, Jun. 2011.

[6] J. Guajardo, S. S. Kumar, G.-J. Schrijen, and P. Tuyls, "FPGA Intrinsic PUFs and their Use for IP protection," in *Proc. Int. Workshop Cryptogr. Hardw. Embedded Syst.*, 2007, pp. 63–80.

[7] D. Fainstein, S. Rosenblatt, A. Cestero, N. Robson, T. Kirihata, and S. S. Iyer, "Dynamic intrinsic chip ID using 32nm high-*k*/metal gate SOI embedded DRAM," in *Symp. VLSI Circuits Dig. Tech. Papers*, pp. 146–147, Jun. 2012.

[8] H. Fujiwara, M. Yabuuchi, H. Nakano, H. Kawai, K. Nii, and K. Arimoto, "A chip-ID generating circuit for dependable LSI using random address errors on embedded SRAMand on-chip memory BIST," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2011, pp. 76–77.

[9] N. Liu, S. Hanson, D. Sylvester, and D. Blaauw, "OxID: On-chip one-time random ID generation using oxide breakdown," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2010, pp. 231–232.

[10] S. Mathew *et al.*, "A 0.19pJ/b PVT-variation-tolerant hybrid physically unclonable function circuit for 100% stable secure key generation in 22nm CMOS," in *ISSCC Dig. Tech. Papers*, Feb. 2014, pp. 278–279.

[11] S. Mathew *et al.*, "A 4fJ/bit delay-hardened physically unclonable function circuit with selective bit destabilization in 14nm tri-gate CMOS," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2016, pp. 1–2.

[12] S. Satpathy *et al.*, "13fJ/bit probing-resilient 250K PUF array with soft darkbit masking for 1.94% bit-error in 22nm tri-gate CMOS," in *Proc. IEEE ESSCIRC*, Sep. 2014, pp. 239–242.

[13] Y. Su, J. Holleman, and B. Otis, "A 1.6pJ/bit 96% stable chip-ID generating circuit using process variations," in *ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 406–407.

[14] J. Lee, D. Lim, B. Gassend, G. E. Suh, M. van Dijk, and S. Devadas, "A technique to build a secret key in integrated circuits for identification and authentication applications," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2004, pp. 176–179.

[15] S. Okumura, S. Yoshimoto, H. Kawaguchi, and M. Yoshimoto, "A 128-bit chip identification generating scheme exploiting SRAM bitcells with failure rate of 4.45×10^{-19}," in *Proc. IEEE ESSCIRC*, Sep. 2011, pp. 527–530.

[16] D. Puntin, S. Stanzione, and G. Iannaccone, "CMOS unclonable system for secure authentication based on device variability," in *Proc. 34th Eur. Solid-State Circuits Conf.*, Sep. 2008, pp. 130–133.

[17] S. Rosenblatt *et al.*, "Field tolerant dynamic intrinsic chip id using 32 nm high-*k*/metal gate SOI embedded DRAM," *IEEE J. Solid-State Circuits*, vol. 48, no. 4, pp. 940–947, Apr. 2013.

[18] R. Maes, V. Rozic, I. Verbauwhede, P. Koeberl, E. van der Sluis, and V. Leest, "Experimental evaluation of physically unclonable functions in 65 nm CMOS," in *Proc. IEEE ESSCIRC*, Sep. 2012, pp. 486–489.

[19] D. Nedospasov, J.-P. Seifert, C. Helfmeier, and C. Boit, "Invasive PUF analysis," in *Proc. Workshop Fault Diagnosis Tolerance Cryptogr.*, 2013, pp. 30–38.

[20] J. Delvaux and I. Verbauwhede, "Side channel modeling attacks on 65nm arbiter PUFs exploiting CMOS device noise," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust*, Jun. 2013, pp. 137–142.

[21] M. Bhargava, C. Cakir, and K. Mai, "Attack resistant sense amplifier based PUFs (SA-PUF) with deterministic and controllable reliability of PUF responses," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust*, Jun. 2010, pp. 106–111.

[22] A. Garg and T. T. Kim, "Design of SRAM PUF with improved uniformity and reliability utilizing device aging effect," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Melbourne, VIC, Australia, Jun. 2014, pp. 1941–1944.

[23] R. Maes and V. van der Leest, "Countering the effects of silicon aging on SRAM PUFs," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST)*, Arlington, VA, USA, May 2014, pp. 148–153.

[24] M. Bhargava, C. Cakir, and K. Mai, "Reliability enhancement of bistable PUFs in 65nm bulk CMOS," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST)*, San Francisco, CA, USA, Jun. 2012, pp. 25–30.

[25] M. Bhargava and K. Mai, "An efficient reliable PUF-based cryptographic key generator in 65nm CMOS," in *Proc. Conf. Design, Autom. Test Eur. (DATE)*, Dresden, Germany, 2014, Art. no. 70.

[26] Y. Dodis, R. Gennaro, J. Håstad, H. Krawczyk, and T. Rabin, "Randomness extraction and key derivation using the CBC, Cascade and HMAC modes," in *Proc. 24th Annu. Int. Cryptol. Conf.*, Santa Barbara, CA, USA, 2004, pp. 494–510.

[27] C.-H. Jan *et al.*, "A 14 nm SoC platform technology featuring 2^{nd} generation Tri-Gate transistors, 70 nm gate pitch, 52 nm metal pitch, and 0.0499 um^2 SRAM cells, optimized for low power, high performance and high density SoC products," in *Symp. (VLSI Technol.) Dig. Tech. Papers*, Jun. 2015, pp. T12–T13.

Sudhir Satpathy (M'09) received the B.Tech. degree in electrical engineering from IIT Kanpur, India, in 2007, and the Ph.D. degree in very large scale integration from the University of Michigan, Ann Arbor, MI, USA, in 2011.

He is currently a Research Scientist with Intel's Circuits Research Lab, Hillsboro, OR, USA. He has authored 35 technical articles and holds 10 issued U.S. patents with 35 pending. His research interests include on-die switch fabrics, interconnects and hardware accelerators for high-performance computer arithmetic.

Sanu K. Mathew (M'99–SM'15) received the B.Tech. degree in electronics and communications engineering from the College of Engineering, Trivandrum, India, in 1993, and the M.S. and Ph.D. degrees in electrical and computer engineering from The State University of New York at Buffalo, Buffalo, NY, USA, in 1996 and 1999, respectively.

He joined Intel Corporation in 1999. He is currently a Principal Engineer with Circuits Research Laboratories at Intel Hillsboro, OR, USA, where he is responsible for developing novel high-performance and energy-efficient digital circuits for next-generation microprocessors. He also mentors Intel and SRC-funded research projects in leading universities. He holds 29 issued patents, has 51 patents pending, and has published over 64 conference/journal papers. His research interests are in the areas of high-speed/low-power computer arithmetic datapath circuits and special-purpose hardware accelerators for cryptography and security.

Dr. Mathew was a recipient of the ISSCC Distinguished Technical Paper Award in 2012. He has served on the program committees of the ARITH, ISLPED, DAC, and SOCC conferences.

Vikram Suresh (M'10) received the Bachelor's degree in engineering from Visvesvaraya Technological University, India, in 2007, the M.S. degree from the University of Massachusetts, Amherst, MA, USA, in 2012, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Massachusetts, in 2014.

He is currently a Senior Research Scientist at Intel Labs, Hillsboro, OR, USA. Prior to his graduate studies, he worked as a hardware design engineer at Mindtree Ltd., India, from 2007 to 2009. He also interned at Advanced Micro Devices in 2011 and Intel Labs in 2012. His research interests include variation-aware circuit design, impact of noise in nanometer CMOS, and hardware security.

Mark A. Anders (M'99–SM'15) received the B.S. and M.S. degrees in electrical engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1998 and 1999, respectively.

Since 1999, he has been with Intel Corporation's Circuit Research Laboratory, Hillsboro, OR, USA, where he is a Principal Engineer. He has published over 70 conference and journal papers, and holds 60 pending and issued patents. His research interests include high-speed and low-power computer arithmetic datapath, DSP, and on-chip interconnects and networks.

Mr. Anders has received the ESSCIRC Best Paper Award, the ISSCC Distinguished Technical Paper Award, and the Intel Achievement Award. He has been recognized as a top IEEE ISSCC paper contributor.

Himanshu Kaul (M'05–SM'16) received the B.Eng. degree (Hons.) in electrical and electronics engineering from the Birla Institute of Technology and Science, Pilani, India, in 2000, and the M.S. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2002 and 2005, respectively.

Since 2004, he has been with Intel Corporation's Circuits Research Laboratory, Hillsboro, OR, USA, where he is currently a Research Scientist. He has published over 50 papers in journals and conferences. His research interests include low-voltage and high-performance circuit design and on-chip communication.

Dr. Kaul received the Distinguished Technical Paper Award at ISSCC 2012 and the Best Paper Award at ESSCIRC 2012.

Amit Agarwal (S'04–M'06) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 2000, and the M.S. and Ph.D. degrees in electrical and computer engineering from Purdue University, Lafayette, IN, USA, in 2001 and 2005, respectively.

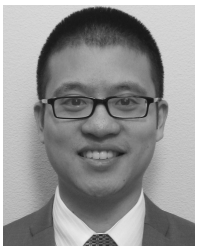
Since 2005, he has been with Intel Corporation's Circuit Research Laboratory, Hillsboro, OR, USA, where he is currently a Research Engineer with the High-Performance Circuits Research Group. He also mentors Intel and SRC-funded research projects in leading universities. He has published over 60 conference/journal papers, and holds 10 issued patents. His research interests are in the areas of high-speed, low-power, process-tolerant, ultralow-voltage circuits/memories, and reconfigurable fabric/architecture.

Dr. Agarwal received the SRC Technical Excellence Award from Semiconductor Research Corporation in 2005, the VLSI Transactions Best Paper Award from the IEEE Circuits and Systems Society in 2006, the ESSCIRC Best Paper Award in 2012, and the ISSCC Distinguished Technical Paper Award in 2012.

Steven K. Hsu (M'99) received the B.S., M.S., and Ph.D. degrees in computer engineering from Oregon State University, Corvallis, OR, USA, in 1999, 2001, and 2006, respectively.

Since 2001, he has been with Intel Corporation's Circuits Research Laboratory, Hillsboro, OR, USA, where he is currently a Research Engineer with the High-Performance Circuits Research Group. He also mentors Intel and SRC-funded research projects in leading universities. He has authored over 52 conference and journal papers, and holds 46 pending and issued patents. His research interests include high-speed, low-power, ultralow-voltage, and variation-tolerant circuit design for datapath, standard cells, and memory.

Dr. Hsu received the ESSCIRC Best Paper Award and the ISSCC Distinguished Technical Paper Award in 2012. He was recognized as a top IEEE ISSCC paper contributor in 2013.



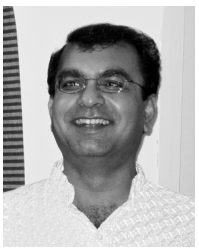
Gregory Chen (S'06–M'11) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2006, 2009, and 2011, respectively.

He is a Research Engineer in Intel Corporation's Circuit Research Lab, High Performance Circuits Group, Hillsboro, OR, USA. He has authored 33 conference and journal papers and holds 19 patents. His research interests include neuromorphic computing and networks-on-chip.

Ram K. Krishnamurthy (M'98–SM'03–F'11) received the B.E. degree in electrical engineering from the Regional Engineering College, Trichy, India, in 1993, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 1998.

He has been with Intel Corporation since 1998, where he is a Senior Principal Engineer and heads the high performance and low voltage circuits research group at the Circuits Research Lab in Hillsboro, OR, USA. In this role, he leads research in high performance, energy efficient and low voltage circuits for microprocessors and SoCs, and has made contributions to the circuit design of various generations of Intel products, including Intel® Itanium®, Pentium4®, Xeon®, Core®, Atom®, and Quark® line of microprocessors and SoCs. He holds 106 issued patents with over 50 patents pending and has published 150 conference/journal papers and 3 book chapters on high-performance energy-efficient circuits.

Dr. Krishnamurthy was a recipient of the 2012 ISSCC Distinguished Technical Paper Award, the 2012 ESSCIRC Best Paper Award, the Semiconductor Research Corporation Outstanding Industry Mentor Award, Intel Awards for most patents filed and most patents issued, the Carnegie Mellon University alumni Recognition Award, the MIT Technology Review's TR35 Innovators Award, and two Intel Achievement Awards for pioneering high-performance and energy-efficient microprocessor and accelerator circuit technologies on Intel products. He is recognized as a top ISSCC Paper Contributor. He served as the Technical Program Chair/General Chair of the 2005/2006 IEEE International System-on-Chip Conference, Guest Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS and an Associate Editor of the IEEE TRANSACTIONS ON VLSI SYSTEMS. He has served on the technical program committees of ISSCC, CICC, and SOCC conferences. He currently serves on the SRC Technical Advisory Board and the SOCC conference's steering committee. He is currently a Distinguished Lecturer of the IEEE Solid-State Circuits Society.



Vivek K. De (F'11) received the Ph.D. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, USA.

He is an Intel Fellow and Director of Circuit Technology Research in Intel Labs, Hillsboro, OR, USA. He is responsible for providing strategic technical directions for long-term research in future circuit technologies and leading energy-efficiency research across the hardware stack. He has 249 publications in refereed international conferences and journals and 208 patents issued, with 27 more patents filed (pending).

Dr. De received an Intel Achievement Award for his contributions to an integrated voltage regulator technology, the Best Paper Award at the 1996 IEEE International ASIC Conference, and nominations for Best Paper Awards at the 2007 IEEE/ACM Design Automation Conference (DAC) and the 2008 IEEE/ACM International Conference on Computer-Aided Design. One of his publications was recognized in the 2013 IEEE/ACM DAC as one of the "Top 10 Cited Papers in 50 Years of DAC."