CMOS unclonable system for secure authentication based on device variability

Daniele Puntin
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni,
Università di Pisa
Stefano Stanzione
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni,
Università di Pisa
Giuseppe Iannaccone
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni,
Università di Pisa

D. Puntin, S. Stanzione, G. Iannaccone, *CMOS unclonable system for secure authentication based on device variability*, 34th European Solid-State Circuits Conference, pp.130-133, Edinburgh 2008.

CMOS Unclonable System for Secure Authentication Based on Device Variability

D. Puntin, S. Stanzione and G. Iannaccone
Università di Pisa
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni
Via Caruso 16, 56122, Pisa, Italy
daniele.puntin@gmail.com, {stefano.stanzione, g.iannaccone}@iet.unipi.it.

Abstract— **An unclonable system for product authentication in anti-counterfeiting has been implemented in standard 90 nm CMOS technology. The circuit exploits the intrinsic variability of the electrical characteristics of minimum size MOSFETs, in order to generate a physical one-way function that univocally identifies each particular IC. Effects of temperature, voltage supply and process variations have been internally compensated to obtain a robust and reliable behavior. Experimental measurements show that the circuit exhibits 30 μW power consumption, a bit error rate in response to a challenge smaller than 0.4% at 125°C or with a 10% voltage supply variation. Accelerated aging tests provide an estimate of a lifetime much in excess of the ten-year requirement. The very low power consumption makes the circuit also feasible for integration in RFID transponders.**

I. Introduction

Counterfeiting severely affects different industrial sectors, including the pharmaceutical, the aircraft, the automotive, and the luxury goods Industry. Recent solutions for secure authentication based on a secret key embedded in an RFID transponder are emerging [1].

One promising authentication method is the use of Physical One-Way Functions (POWF) [2] or Physical Unclonable Functions (PUF), that are functions based on the physical properties of real objects, too complex to be determined through brute force attacks, and impossible to clone because unknown even to the manufacturer. A significant number of proposed PUFs is based on optical properties [2,3]. They are obtained from the interference pattern of a transparent material containing randomly distributed scattering particles illuminated with a laser beam from a specified angle. Their unclonability derives from the uniqueness and unpredictability of speckle patterns resulting from multiple scattering of laser light in a disordered optical medium and from the impossibility of extracting the response function from a finite number of readings.

Silicon PUFs have been proposed, in which the random variability of the delays of logic gates [4-7] or of transistor threshold voltages [8] are used to generate a string univocally identifying a given circuit.

In the present work we propose a circuit – the "silicon nanokey" for generating a PUF to be used in a challenge-response authentication scheme. The circuit exploits the variability of the electrical parameters of minimum size MOS transistors, in particular of the threshold voltage, in order to generate a unique unclonable and reliable digital response to any digital challenge provided as input to the system. The nanokey is also robust to invasive attacks because a direct physical access would alter its properties.

Since the response of the nanokey to any challenge is unknown even to the manufacturer, we envisage a scheme in which an authenticating entity tests each nanokey before use, and collects and stores a large number of valid challenge-response pairs. In use, when one wants to authenticate the nanokey, it has to ask the authenticating entity a challenge, and then has to send the response back to the authenticating entity which matches the response to the stored one. Maximum security is obtained if each challenge-response pair is used only once.

II. Circuit Description

Each bit of the response is obtained as a function of the analog output of a number of inverters biased around the point of maximum gain and realized with nMOSFETs of minimum length and width, in order to maximize threshold voltage variability and therefore output voltage dispersion. The block diagram of the nanokey is shown in Fig. 1. The core of the structure is the two-dimensional array of inverters. The digital challenge is a sequence of N inverter addresses. Each of the N-M+1 response bits is obtained from a group of M consecutive addresses, through a bit extractor circuit. In the simplest example realized we have M=2 and the bit extractor is an autozero comparator of the outputs of pairs of inverters. If the nanokey array and N are large enough the space of possible challenges can be easily made too large for a brute force attack.

The structure of a row of the array is shown in Fig. 2. Note that a single pMOSFET is used for each row, in order to reduce the total area occupation on the silicon die. This solution is especially useful if the number N of inverters in the array is large. In fact, in this way, it is possible obtain N^2

978-1-4244-2362-0/08/$25.00 ©2008 IEEE.

inverters by using only N pMOSFETs. Furthermore, pMOSFETs are large enough to exhibit negligible mismatch in comparison to the minimum sized nMOSFETs, so that they do not introduce a systematic shift of the output voltages on the same row.

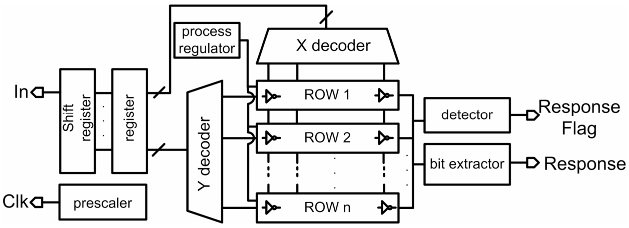

Figure 1. Block diagram of the nanokey chip.

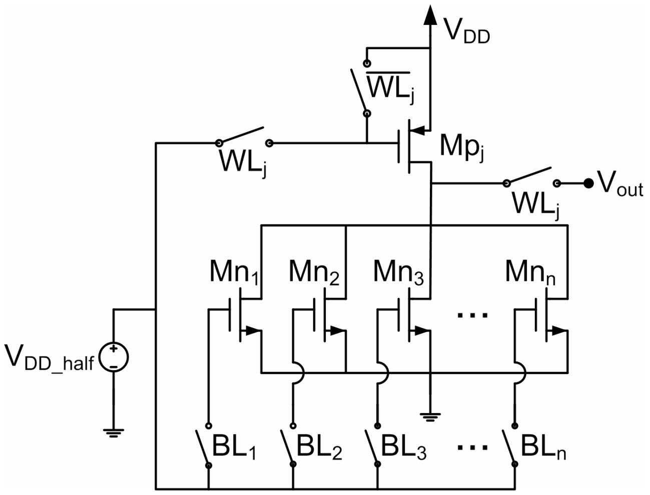

Figure 2. Circuit of a row of the nanokey array. Bit lines BL_i are the outputs of the X decoder, Wordlines WL_j are the outputs of the Y decoder shown in Fig. 1.

It is important to bias the nanokey inverters in the maximum gain region, in order to maximize the inverter output dispersion, and to make the output response robust and reproducible in the presence of temperature, supply voltage, and process variations. To this aim, the input voltage of the inverters, V_{DD_half}, is provided by a process regulator circuit, shown in Fig. 3: The two pairs of nMOS and pMOS are perfectly matched and amplify the differential voltage (V_X-$V_{DD}/2$). These MOSFETs have a large size, in order to be mismatch insensitive.

The circuit labeled with A in Fig. 3 consists of 6 inverters identical to those forming the nanokey array connected in parallel and is equivalent to a single inverter of the array biased in the gain region but less sensitive to mismatch. Note that a larger number of inverters in parallel in the circuit A could reduce its sensitiveness to mismatch. On the other hand, since inverters biased in the maximum gain region require a considerable bias current, the circuit A absorbs a significant portion of the power budget, and therefore the number of inverters in A has to be minimized.

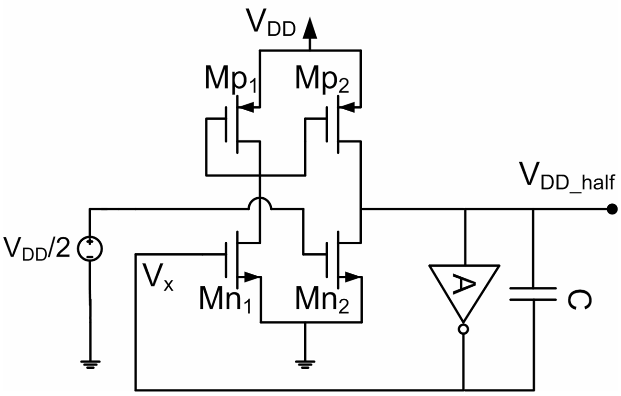

Figure 3. Circuit of the process regulator.

The operation of the process regulator circuit is the following: if for any reason a change is induced in the value of V_X, the negative feedback of the circuit will compensate it, changing the input voltage V_{DD_half} of the matrix inverters. So, the voltage V_{DD_half} is regulated in order to keep A in the gain region even if undesired variations occur. Since A works in the maximum gain region, the gain loop of the circuit is high and V_X is very close to $V_{DD}/2$. The capacitor C of 250 fF allows Miller compensation to achieve loop stability. Given that A exhibits the same process dependence of the nanokey inverters, their average outputs will be compensated with respect to process, temperature, and voltage supply variations.

Let us describe now the authentication algorithm in the case M=2: every response bit is obtained by the sign of the difference between consecutive array outputs. So, the response bits can be easily extracted using an autozeroing comparator, as shown in Fig. 4. During the phase Φ_n the switches are closed and the inverters work on their trip point. Afterwards, during the phase Φ, the output array commutation occurs and, being the comparator inverters biased in the maximum gain region, the slightest variation of the array output unbalances the comparator inverters. The sensitivity to the array output variations is further increased by the use of large capacitances C (1pF), if compared with the input capacitances of the inverters. In this way the digital bits are correct only during the phase Φ. In order to complete the AD conversion, a Flip Flop D-latch has been used.

An additional measure to make the response robust and reliable is to discard those bits more likely to change as a result of noise, operation condition variations, or aging. To this aim a response flag is attached to each response bit upon the first reading, when the challenge-response pair is collected and stored: if the response flag is 1 the bit is unreliable (not valid), and is not considered in the validation of the response. In the case of M=2, the response flag is set to one for those bits obtained when the difference of inverter outputs is smaller than a “decision margin” V_{DM}.

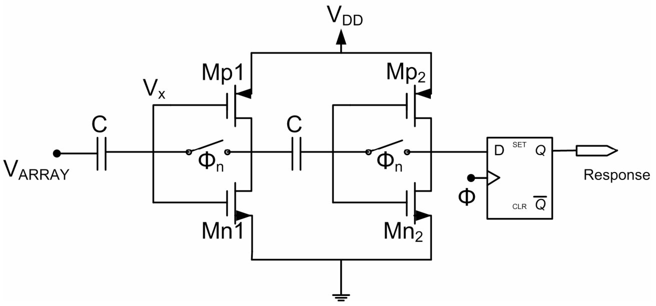

Figure 4. Bit extractor.

The detector circuit, shown in Fig. 5, produces the response flag (*RF*) bits, that are function of two voltages, V_{DM} and V_{ARRAY}, as results from the following relation:

$$RF = \begin{cases} 0\ ,|\Delta V_{ARRAY}| > V_{DM} \\ 1\ ,|\Delta V_{ARRAY}| < V_{DM} \end{cases} \quad (1)$$

The up and down chains simply evaluate if $\Delta V_{ARRAY} > V_{DM}$ and $\Delta V_{ARRAY} > -V_{DM}$ respectively. This means that a simple *XOR* operation between the two outputs of the chains evaluates if the absolute value of ΔV_{ARRAY} is lower than the decision margin V_{DM}.

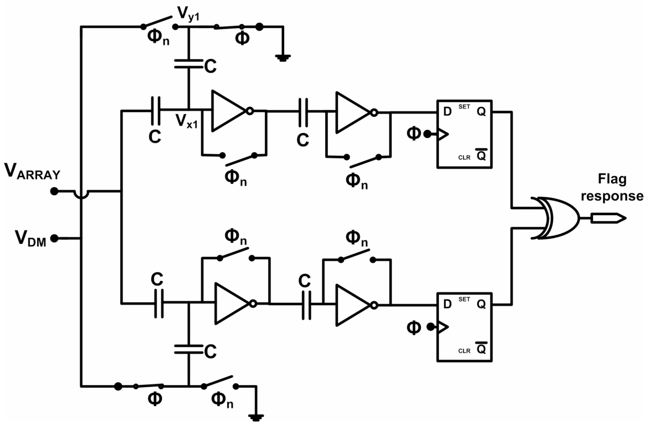

Figure 5. Detector.

It is important to observe that V_{DM} is a result of a tradeoff: decreasing V_{DM} results in an increased bit error rate, increasing V_{DM} results in an increased number of bits to be rejected from the response. The False Rejection Rate (FRR) and the False Acceptance Rate (FAR) are a function of N, M, V_{DM} , the number of wrong bits accepted n_e in the response, and of system properties (noise, disturbs, and aging). Their evaluation is beyond the scope of the present paper. For the moment let us stress the fact that N, M, V_{DM} and n_e are obtained as a tradeoff between specifications on FRR and FAR, and that an increase of N and M has a much stronger effect suppressing FAR than increasing FRR.

III. Experimental Results

The described nanokey has been implemented in a standard 90 nm CMOS process with an array size of 256 inverters and M=2, with the primary purpose of testing the robustness and reliability of the concept. At room temperature and with supply voltage of 600 mV the circuit absorbs 30 μW. The correlation between responses to the same input test challenge of 1500 addresses is essentially unitary if the same chip is considered and smaller than 1% if different chips are considered.

An important parameter of a nanokey is the Bit Error Rate (BER), that is the error probability of a valid response bit, that is a response bit not rejected in the first reference measurement. By performing measurements for various values of the decision margin V_{DM}, the behavior of the BER as a function of V_{DM} has been obtained. As shown in Fig. 6, the BER is a decreasing function of V_{DM}, and the fraction of non-valid bits in the response increases with V_{DM}, as shown in Fig. 7. A good tradeoff is represented by a decision margin of 35 mV, for which the Bit Rejection Rate is lower than 20% and the BER for V_{DD}= 600 mV at room temperature is lower than 0.07%.

Measurements have been performed also varying the operating conditions. Tests have been performed with operation at 125°C or with a 10% voltage supply variation. Although CMOS inverter outputs are strongly dependent on temperature and voltage supply, the process regulator makes the circuit robust to such of variations. For a V_{DM} of 35 mV, in the worst case, the BER is lower than 0.4%. This is comparable with the best result presented in literature, related to a different identification circuit [5].

Finally, accelerated aging tests have been performed. Circuits with no power supply have been put in an oven and periodically extracted, cycled to room temperature and tested. No observable degradation in their BER has been observed after 142 hours at 250°C, as shown in Fig. 6.

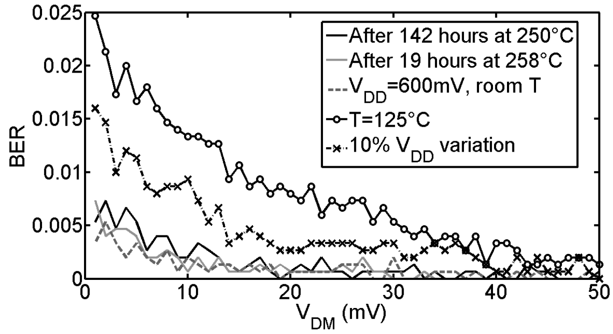

Figure 6. BER as a function of the decision margin V_{DM}.

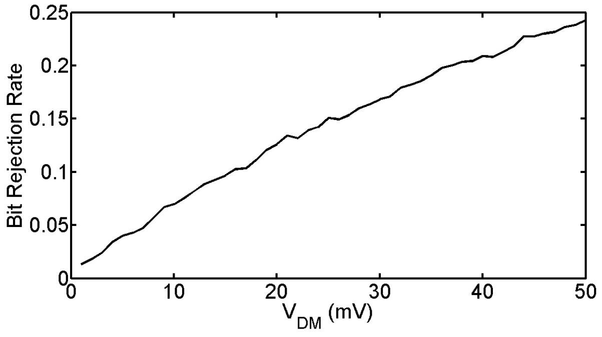

Figure 7. Bit Rejection Rate as a function of the decision margin V_{DM}.

In order to extract some additional information on their aging, we have evaluated the standard deviation of the difference between the comparator output voltages and those of the first reference measurements. Such value slowly increases with aging and is 7.8 mV after 142 hours at 250°C. Let us stress that this is a symptom of aging even if the circuit still works perfectly. At 258°C the same condition is reached after 19 hours. In Fig. 8 is shown the dispersion of the output voltage difference values with respect to those of the reference measurements as a function of the time during the accelerated aging tests. Higher temperatures could not be tested because of the large thermal time constants of the oven.

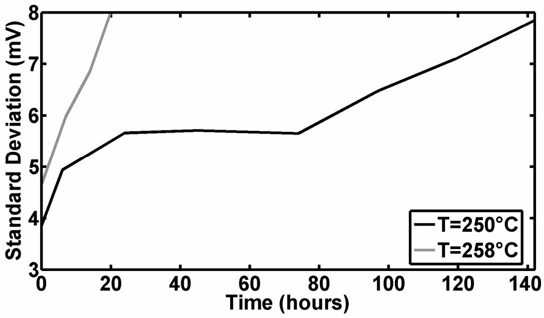

Figure 8. Standard deviation of the difference between comparator voltage outputs and those of the first reference measurement for the same challenge.

Such a strong temperature dependence is very promising from the point of view of the lifetime at room temperature: assuming an Arrenius-type aging process, it provides a lifetime well in excess of 10 years even at 120°C.

IV. Conclusion

A circuit for secure authentication, exploiting the intrinsic variability of the electrical characteristics of nanometric size MOS transistors, has been proposed and demonstrated. An internal compensation ensures a very good robustness to process, temperature and voltage supply variations. Measurements on the proposed IC have shown a 30 mW power supply and BER lower than 0.4% with operation at 125°C or with a 10% voltage supply variation. Finally, performing accelerated aging tests, a lifetime much in excess of the common ten-year requirement has been estimated. This work has been partially supported by Fondazione Cassa di Risparmio di Pisa and by Cassa di Risparmio di Pisa, Lucca e Livorno.

References

[1] P. Tuyls, L. Batina, “RFID-Tags for Anti-counterfeiting”, Topics in Cryptology – CT RSA 2006, Springer Berlin, pp. 115-131, 2006.

[2] Pappu, B. Recht, J. Taylor, N. Gerschen-Feld, “Physical one-way functions”, Science, vol. 297, pp. 2026-2030, 2002.

[3] B. Škorić, P. Tuyls, and W. Ophey, “Robust key extraction from physical uncloneable functions”, Proc. ACNS 2005, pp. 407-422, 2005.

[4] J. W. 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”, Proc. IEEE 2004 Symposium On VLSI Circuits, pp. 176-179.

[5] G. E. Suh and S. Devadas, “Physical unclonable functions for device authentication and secret key generation”, Proc. DAC 2007, pp. 9-14, June 4-8, 2007.

[6] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas, “Silicon physical random functions”, in Proc. Computer Communication Security Conf., Nov. 2002, pp. 148-160.

[7] D. Lim, J. W. Lee, B. Gassend, G. E. Suh, M. van Dijk, and S. Devadas, “Extracting secret keys from integrated circuits”, IEEE Trans. VLSI Systems, Vol. 13, no. 10, pp. 1200-1205, October 2005.

[8] K. Lofstrom, W. R. Daasch, D. Taylor, “IC identification circuit using device mismatch”, IEEE International Solid-State Circuits Conference, Digest Tech. Papers, pp. 372-373, 2000.