Increasing DPA Success on GALS Cryptographic Architectures through DSP Techniques

Luciano Loder, Marcelo Fay, Rafael Soares
Universidade Federal de Pelotas - UFPEL
Rua Gomes Carneiro, 1 Pelotas – RS Brazil
{llloder,mlcfay,rafael.soares}@inf.ufpel.edu.br

Adão de Souza Jr
Instituto Federal Sul Rio Grandense - IFSUL
Praça 20 de Setembro, Pelotas – RS Brazil
adaojr@gmail.com

ABSTRACT
Differential Power Analysis (DPA) has been effective in revealing cryptographic key on cryptosystems. In this sense, several methods to prevent the action of these attacks have been proposed in order to avoid information leakage. Random delay insertion (RDI) causes misalignments in the time domain hindering the action of DPA. However, some digital DSP techniques have being proven efficient as countermeasures to improve DPA rate of success. The GALS pipeline architectures uses random clock frequency and simultaneous processing of two or more pipeline stages to hide leakage information. This paper presents a preliminary investigation of DSP techniques to perform frequency sorting, realignment of power traces, noise reduction and downsampling traces before DPA attacks. The results show that applying digital filters and correlation phase adjustment may be efficient to improve DPA attacks.

Categories and Subject Descriptors
D.4.6 [**Security and Protection**]: Non-Invasive software attacks (DPA, CPA, DEMA, CEMA).

General Terms
Security.

Keywords
Cryptography, Security, Side Channel Attacks, DPA, CPA, CEMA, DSP.

1. INTRODUCTION
A major concern of the cryptographic circuit designers are Side Channel Attacks (SCA), since it was presented by Kocher [1]. SCA exploits leakage information on side channels such as power consumption, electromagnetic radiation, or time propagation to derive confidential information, specifically secret keys used in cryptographic systems. DPA uses simple statistical techniques that are almost independent of the cryptographic algorithm implementation to correlate data and power consumption. Differential Electromagnetic Analysis (DEMA) follows the same principle, but measures the electromagnetic radiation of circuits, according to Gebotys el al. [2]. Brier et al. [4] proposed an improvement in DPA using a power model so called Correlation Power Analysis (CPA) to be DPA more effective. A similar attack when applied on electromagnetic radiation is known as CEMA.

DPA evaluates the power signature caused by execution of an intermediate operation of the cryptographic system for a specific plaintext data input and compares all power signatures obtained by the execution of a set of different plaintexts. As the hardware device is designed according to synchronous paradigm, all operations are executed sequentially in the same order and spending a very close runtime. For a successful DPA, the power traces values at each point time caused by the same operation have to be captured correctly aligned in the time domain and then the attacker is able to compare them.

Since then, efforts are increasing to improve the security of cryptographic circuits against DPA. In this sense, a lot of proposes have been presented [3][4][10], commonly known as countermeasures. Moreover, several works are presented to improve the efficiency of SCA and find vulnerabilities in systems protected with some kind of countermeasure [8][9].

One strategy of these countermeasures to avoid DPA consists in execute the cryptographic algorithm at different time instants aiming to scatter power traces waveforms by insertion random delays. In [3], RDI is applied in software level, interleaving the cryptographic algorithm with dummy instructions. In [4], RDI is applied in hardware level by the addition of logic gates to the datapath. RDI can be implemented by driving the system at different clock frequencies as mentioned by [7].

Nagashima et al. in [9] prove to be possible unveil the secret key on cryptographic systems protected by RDI countermeasure through phase only correlation (POC). Generally pre-processing with signal processing techniques can be effective to align power traces before apply DPA [8].

Combining RDI with extra hardware to compute a cryptographic algorithm is effective against DPA. In [10] the authors proposed an architectural countermeasure that implements a hardware pipeline using GALS design style able to compute each pipeline stage driven by random clock frequencies.

In the literature, there is nothing related about the use of DSP techniques to improve DPA on GALS architecture. In this paper is presented a preliminary evaluation of GALS pipeline architecture against CEMA attacks using pre-processing techniques such as POC, filters and subsampling. This article is structured as follows: Section II explains the DPA analysis. Section III shows the GALS pipeline architecture. An overview of digital signal processing techniques is presented on Section IV. The experiments and results are discussed on Section V and some conclusions are presented on Section VI.

2. DPA ANALYSIS
DPA attacks are the most popular type of power analysis attacks. Essentially, two dependencies of the power consumption are exploited: the data-dependency and the operation dependency. This is due to the fact that DPA attacks do not require detailed knowledge about the attacked device. Furthermore, they can reveal the secret key even if the recorded power traces are extremely noisy. DPA requires a large number of power traces recorded when the device encrypts or decrypts different data. DPA analyzes how the power consumption at fixed moments of time depends on the processed data.

The strategy of DPA consists of 5 steps. (i) Choose an intermediate result of the executed algorithm. It needs to be a function *f(d,k)*, where *d* is a known non-constant data and *k* is a small part of the key. (ii) Measure the power consumption of the cryptographic device while it encrypts *D* different data and to know the *d* value of the (i). (iii) Calculate a hypothetical intermediate value for every possible choice of *k*. (iv) Map the hypothetical intermediate values to a hypothetical power consumption values; (v) Compare the hypothetical power consumption values with the power traces. The correct hypothesis of key is that has the highest peak as depicted on Figure 1.

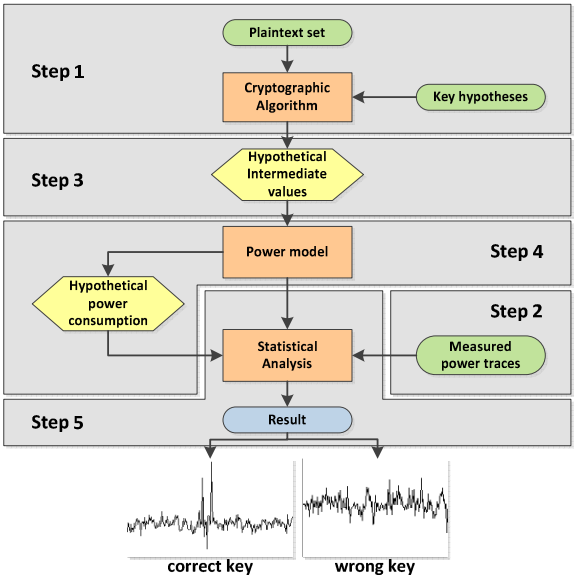

Figure 1. The strategy of DPA to unveil a secret key.

2.1 CORRELATION ANALYSIS

Correlation coefficient is the most common way to determine linear relationships between data. There is a well-established theory for the correlation coefficient that can be used to model statistical properties of DPA attacks. In DPA, the correlation coefficient is used to determine the linear relationship between the hypothetical intermediate values and the power traces. This coefficient reduces the wrong results in DPA caused by noise present on power traces. It is called Correlation Power Analysis (CPA) [4]. The same analysis can be done on electromagnetic radiation traces and so it is called CEMA.

3. GALS PIPELINE ARCHITECTURE

The GALS pipeline presented by [10] is an architecture that implements a cryptographic algorithm in hardware according to globally asynchronous locally synchronous design style to avoid leakage information. This architecture is composed by synchronous logic blocks that communicate each other using asynchronous interfaces as depicted in Figure 2, also known as synchronous islands.

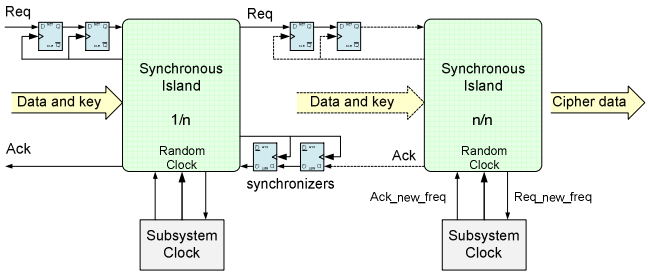

Figure. 2 Structure of the GALS pipeline architecture.

Each island is able to process one or more rounds of the algorithm. The islands can be driven by a global or a local clock signal. A subsystem clock is responsible to generate clock signals with different frequencies. A Linear Feedback Shift Register (LFSR) defines the clock signal that drives the island at end of data encryption. Thus, the architectural proposal avoids SCA through execution of pipeline stages on different frequencies in addition to the simultaneous executions of the islands causing amplitude distortions in the power traces [10]. The processing in different clock frequencies causes misalignment in the time domain besides provoke a distortion in the amplitude.

4. PRE-PROCESSING TECHNIQUES

Since the power traces are misaligned in time and have different clock frequencies, DPA attack is not able to unveil a secret cryptographic key on GALS pipeline architectures. Some different pre-processing techniques were used in this work to improve DPA success rate: phase-based waveform matching, as presented by [9]; frequency classification, low pass filter de-noising and subsampling signals.

4.1 Phase based waveform matching

Consider two signals, *f(n)* and *g(n)*, where we assume that the index range is *n* = *-Z, ... ,Z* for mathematical simplicity, and hence the length of waveforms *L = 2Z + 1*. Let *F(k)* and *G(k)* denotes the Discrete Fourier Transforms (DFTs) of the two waveforms [9]. *F(k)* and *G(k)* are given by

$$F(k) = \sum_{n=-Z}^{Z} f(n) W_L^{kn} = A_F(k) e^{j\theta_F(k)} \quad (1)$$

$$G(k) = \sum_{n=-M}^{M} g(n) W_L^{kn} = A_G(k) e^{j\theta_G(k)} \quad (2)$$

where $W^N = e^{-j\frac{2\pi}{L}}$, $A_F(k)$ e $A_G(k)$ are amplitude components, and $e^{j\theta_F(k)}$ and $e^{j\theta_G(k)}$ are phase components [9]. The cross-phase spectrum (or normalized cross spectrum) $R_{FG}(k)$ is defined as

$$R_{FG}(k) = \frac{F(k)\overline{G(k)}}{\left|F(k)G(k)\right|} = e^{j\theta FG(k)} \quad (3)$$

where $\overline{G(k)}$ denotes the complex conjugate of *G(k)* and $\theta FG(k) = \theta F(k) - \theta G(k)$. The POC function $r_{fg}(n)$ is the Inverse Discrete Fourier Transform (IDFT) of *RFG(n)* and is given by

$$r_{fg}(n) = \frac{1}{L} \sum_{k=-Z}^{Z} R_{FG}(k) W_L^{-kn} \quad (4)$$

If there is a similarity between two waveforms, the POC function gives a distinct sharp peak. (When *f(n) = g(n)*, the POC function

becomes the Kronecker delta function.) If not, the peak drops significantly. The height of the peak can be used as a good similarity metric for the waveform matching, and the location of the peak shows the translational displacement between the two waveforms.

4.2 Frequency Separation

The FFT can be used in order to evaluate the clock frequency of a synchronous circuit. The FFT shows the frequency components of a given signal and since all operations in a synchronous circuit depends on the clock frequency, the highest peak of the FFT should be in the operating clock frequency. We realized some experiments and concluded that this assumption is correct. For a more accurate evaluation, boundary effects must be considered. To reduce the boundary effects, the power traces are processed with a window function before the FFT process. An important consideration is that this step is performed before the application of the moving average filter, because this filter attenuates the higher frequencies.

4.3 Moving Average Filter

To reduce the influence of the frequencies above the desired frequency, it has been used a moving average filter.

This filter consists of a low-pass Finite Impulse Response (FIR) filter that is used to remove the frequencies above the desired frequency to reduce the noise in order to improve the DPA analysis. This filter is defined by following equation

$$y[n] = \frac{1}{N}\sum_{k=0}^{N-1} x[n-k] \qquad (5)$$

The window used has n=100 taps.

4.4 Downsampling Signal

Downsampling is the process of reducing the sampling rate of a signal and to apply it is necessary to respect the Shannon-Nyquist sampling theorem [11]. Although it is commonplace to use sampling rates at least one order of magnitude above the Nyquist frequency, from systems identification perspective an oversampled signal can lead to a numerically ill-conditioned model. As samples get closer, successive samples are increasingly correlated to the point where extra samples are adding little information about the system dynamics. Also, since the computation time is related to the number of samples in each trace, it can be reduced by downsampling the traces without loss of information.

A good sampling rate can be achieved with a two steps method: first the signal is sampled at a rate that is several times higher than the Nyquist rate; then a downsampling factor Δ is determinate analyzing the autocorrelation on the oversampled trace. Factor is chosen taking two autocorrelation functions: from the trace and from its first power series (9)(10) where k is the index for the delay in the autocorrelation function and n the absolute index in the series.

$$r_y(k) = E\{y[n]\cdot y[n-k]\} \qquad (6)$$

$$r_y^2(k) = E\{y[n]^2\cdot y[n-k]^2\} \qquad (7)$$

Downsampling factor will Δ be chosen as the smallest value of k that minimizes the autocorrelation in each function (lobe). Equation (10) allows to take into account some nonlinear behavior in the system dynamics first [11]. For the data set under analysis, the obtained subsampling rate was 50, indicating that the trace size could be reduced by 50 times from the original without loss information. To avoid spectral overlapping in the downsampling process it is necessary to remove frequencies above the final sampling rate [12]. This filter must be applied before the trace is downsampled in the same way one uses an anti-alias filter. In the present work a frequency domain low pass filter was employed [12].

5. EXPERIMENTATION

In this paper we used a collection of 100,000 electromagnetic radiation traces acquired in [10]. The architecture target of attack presents 2 stages pipeline, each one executing 8 rounds of DES algorithm and sequential execution without parallelism. Since the traces obtained had different clock frequencies, frequency classification was performed.

5.1 First experiment

The first experiment was to perform FFT on traces where the encryption was executed. As a result, we have each trace associated to the frequency operation of the first stage. Two clusters of frequencies clearly emerged: Group 1 composed of traces whose frequencies vary between 38MHz and 42MHz and Group 2 composed of traces whose frequencies vary between 55MHz and 60MHz.

The act of processing at different frequencies causes a change in the beginning of the execution of the algorithm. Thus, the starting point of the computation has been detected using an amplitude criterion. A threshold is defined to approximately distinguish computing and noise. In this context, POC-based waveform matching is used to align traces in each group accordingly with one reference trace.

The next step, attacks are performed on traces aligned and compared with attacks applied on traces misaligned. The CEMA attacks are used to evaluate the robustness of the GALS architecture. The results are summarized in Table 1. In the first column is presented the function target of attacks, the eight Substitutions Boxs (referred as Sbox) of the DES algorithm. The second column presents the group of traces addressed by the attacks. The others columns present the experiment applied and its respective minimal number of traces (*# T*) necessary for CEMA unveil a correct key and its associated rank. The rank sorts the probabilities of all possible keys for each Sbox. In a successful attack, the rank of the key guessed must be 1, otherwise the attack fails. The experiments performed are CEMA without pre-processing signal (WPP), CEMA with POC and CEMA with POC and moving average filter presented respectively by the 3rd, 4th and 5th columns in Table 1.

As expected, CEMA without pre-processing is not effective to guess a key in an architecture protected as shown in Table 1 column WPP. The hyphen (-) indicates an unsuccessful attack. Therefore, when POC is applied on traces the attacks are successful and only the sbox8 was not guessed. To improve even more the results, was used a moving average filter before CEMA attack, mitigating the high frequency noise. A significant reduction of number of traces was observed.

5.2 Second experiment

In this work is presented a preliminary experiment with downsampling. In this sense, only one experiment with the group 1 of traces was applied. A subsampling rate 50 is performed on traces. Next, the traces are submitted to a lowpass filter and finally a POC function to align the traces subsampled. The CEMA attack was executed and the results are applied in the Table 2. The (# T) is presents in row 2 and the rank on row 3 of Table 2.

Table 1. Results of some CEMA attacks.

		WPP		POC		POC+Filter	
Function	**Group**	**# T**	**Rank**	**# T**	**Rank**	**# T**	**Rank**
Sbox1	**1**	-	40	6943	1	1290	1
	2	-	29	2513	1	1382	1
Sbox2	**1**	-	21	22130	1	12932	1
	2	-	28	10468	1	2032	1
Sbox3	**1**	-	31	29752	1	18836	1
	2	-	15	3798	1	1077	1
Sbox4	**1**	-	8	12211	1	4790	1
	2	-	32	3510	1	2562	1
Sbox5	**1**	-	61	27238	1	21516	1
	2	-	12	-	2	16023	1
Sbox6	**1**	-	25	15987	1	15987	1
	2	-	37	9202	1	2294	1
Sbox7	**1**	-	32	33511	1	13886	1
	2	-	38	2187	1	2097	1
Sbox8	**1**	-	5	-	56	-	22
	2	-	22	-	2	13777	1

Table 2. CEMA with downsampling, POC and filter.

Sbox1	Sbox2	Sbox3	Sbox4	Sbox5	Sbox6	Sbox7	Sbox8
1957	14036	30262	4709	50714	12352	12822	40307
1	1	1	1	1	1	1	12

6. CONCLUSIONS

Architectures GALS pipelines are proposed to hide information leaked through power consumption and electromagnetic radiation combining random frequency and noise addition to misalign and disturb power traces. Soares et al. [10] proved that this countermeasure is effective against DPA attacks. However, there are some digital signal processing techniques able to remove misalignment and noise present on power or electromagnetic traces exploring the remaining vulnerabilities. This paper presented a preliminary investigation of DSP techniques to perform frequency sorting, realignment of power traces, noise reduction and downsampling traces before DPA attacks. The results show that grouping traces produced with very close clock frequencies combined with phase correlation is effective to ensure DPA success. The attack has been able to reveal 7 out of 8 intermediate keys. In addition, the high frequency noise reduction was able to decrease the number of traces required to unveil the intermediate keys. Downsampling the traces allowed to keep the same DPA success rate with a significant reduction in the time of analysis.

Although the sbox8 key was not able to be revealed its position in the hypotheses ranking was improved from 56 to 22 when moving average filter was applied. This preliminary evaluation proves to be efficient to improve DPA attacks against random frequency misalignment traces. In future works, we intend to evaluate the robustness of the combination of two pipeline stages processing simultaneously with random clock frequencies to hide leakage information against DPA attacks.

7. ACKNOWLEDGMENTS

This work was supported by FAPERGS ARD-2011 Processo 1836-4 and program PIBIC/CNPQ.

8. REFERENCES

[1] P. Kocher. "Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and others Systems". In: 16th International Cryptology Conference on Advances in Cryptology (CRYPTO'96), Aug 1996, pp. 104-113.

[2] Gebotys, C.; Tiu, C.; Chen, X. "A Countermeasure for EM Attacks of a Wireless PDA". In: International Conference on Information Technology: Coding and Computing, 2005, pp. 544-549.

[3] Clavier, J. Coron, and N. Dabbous. "Differential Power Analysis in the Presence of Hardware Countermeasures". In: Cryptographic Hardware Embedded Systems, 2000, pp. 252 - 263.

[4] Brier, E.; Clavier, C.; Olivier, F. "Correlation Power Analysis with a Leakage Model", In: Cryptographic Hardware and Embedded Systems, 2004, pp. 16-29.

[5] Guilley, S.; Sauvage, L.; Danger, J.; Graba, T.; Mathieu, Y. "Evaluation of Power-Constant Dual-Rail Logic as a Protection of Cryptographic Applications in FPGAs". In: 2nd International Conference on Secure System Integration and Reality Improvement, 2008, pp. 19-23.

[6] Lu, Yingxi; O'Neill, Maire P. "FPGA implementation and analysis of Random Delay Insertion Countermeasure against DPA". In: International Conference on Field-Programmable Technology, 2008, pp. 201-208.

[7] Avirneni, Naga Durga Prasad; Somani, Arun K. "Countering power analysis attacks using reliable and aggressive designs". IEEE Transactions on Computers, vol. 99, pp. 1. 2013.

[8] Le, Tanh-Ha; Clédière, Jessy; Servière,Christine; Lacoume, Jean-Louis; How can signal processing benefit Side Channel Attacks?. In: Workshop on Signal Processing Applications for Public Safe, 2007, pp. 1-7.

[9] Nagashima, Sei; Homma, Naofumi; Imai, Yuichi; Takafumi, Aoki; Satoh, Akashi. "DPA using phase-based waveform matching against random-delay countermeasures". In: IEEE International Symposium on Circuits and Systems, 2007, pp 1807-1810.

[10] Soares, R.; Calazans, N.; Moares, F.; Maurine, P.; Torres, L. "A robust architectural approach for cryptographic algorithms using GALS pipelines. IEEE Design & Test of Computers, 28(5), 2011, pp. 62-71.

[11] Aguirre, L. A.; Mendes, E. E. "The least squares Padé method for model reduction on multivariable systems". In: International Journal of Systems Science, n. 26, vol. 4, pp.819-839, 1995.

[12] Schynk, J.J. "Frequency-domain and multirate adaptive filtering", Signal Processing Magazine, IEEE, vol. 9, n. 1, pp.14-37, January, 1992.