Read/Write Signals Reconstruction Using Side Channel Analysis for Reverse Engineering*

Damien Marion **Antoine Wurcker**
damien.marion@unilim.fr antoine.wurcker@xlim.fr

XLIM-CNRS, University of Limoges, France

* *This work has been conducted under the framework of the MARSHAL+ french project*

COSADE 2013 - March 7, 2013

Outline

Outline

1. Introduction
2. Characterization Phase
3. Read/Write Signals Reconstruction Process
4. Conclusion and Further Work

Context

Side Channel Analysis for Reverse Engineering (SCARE):

- identify executed instructions

Our study done on:

- AES SubBytes function executed by an ASIC emultating 6502 microcontroller
- Consumption traces
- MODELSIM: values of several internal data (dint{1, 3}, flag_{c, v, z}, opcode, pc, sp, we, rd, reg_{a, x, y})

Notation

- $Y = (y_{i,j})_{\substack{0 \leq i < \text{nb\_of\_cycles} \\ 0 \leq j < \text{size\_of\_cycle}}}$ for the consumption traces.
- $X = (x_{i,j})_{\substack{0 \leq i < \text{nb\_of\_cycles} \\ 0 \leq j < \text{nb\_of\_internals}}}$ for the internal data.

Outline

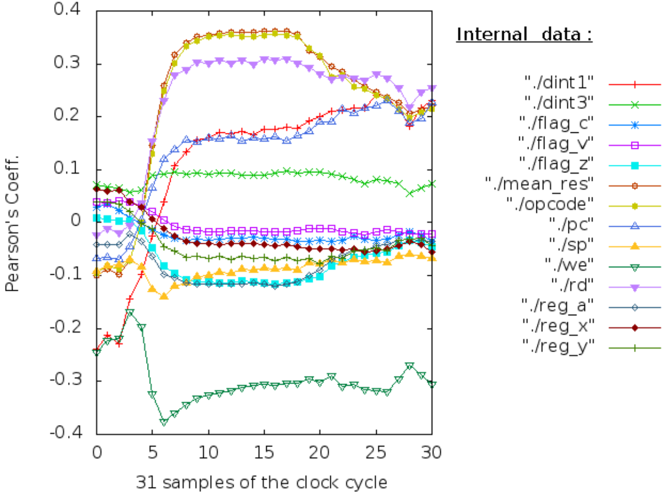
Horizontal CPA
Between Consumption Traces and Internal Data

Horizontal CPA computation

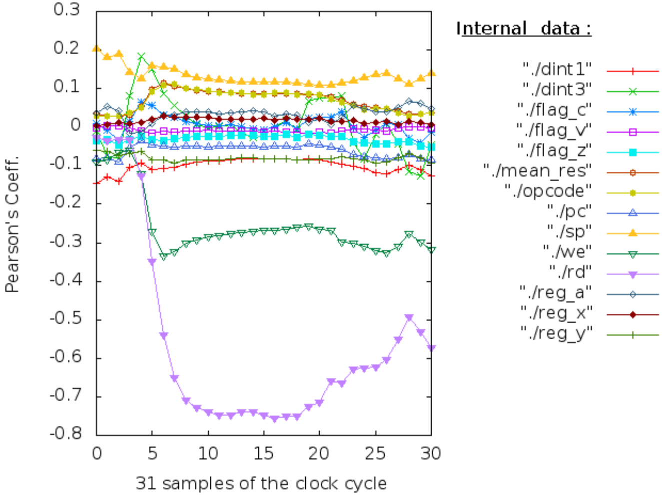
- $\forall j <$ nb_of_internals, we note $x = X(:, j)$
- $\forall j <$ size_of_cycle, we note $y = Y(:, j)$

$$Corr(x, y) = \frac{\sum_{i=1}^{N} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2}}$$

Horizontal CPA with Hamming Distance



Horizontal CPA with Hamming Weight



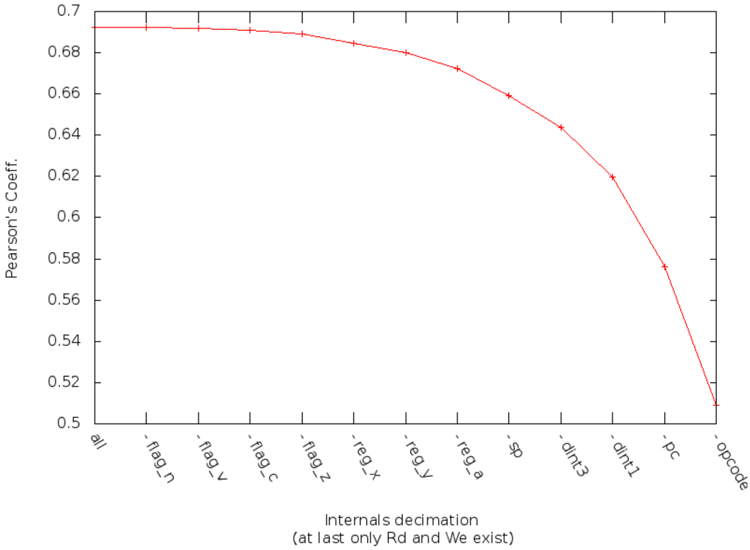
Multivariate Linear Correlation
Between Consumption Traces and Set of all Internal Data

Multivariate Linear Correlation

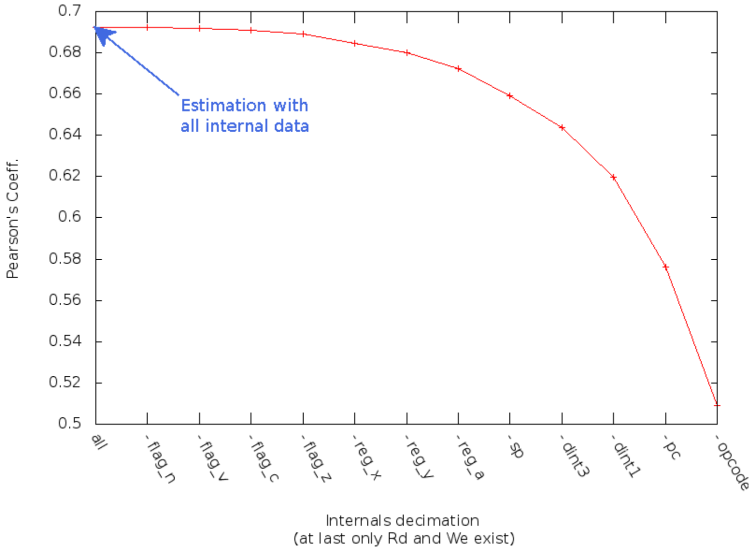
- $\forall i <$ nb_of_internals, we note $X = X(:, 0 : i)$
- we note $Y = Y(:, 9)$ (the 9^{th} sample gives the best results)

In the model $Y = X.A + \epsilon$,
we used **Least-Squares Estimation** to estimate A
then we evaluate the estimation with Pearson's correlation coefficient.

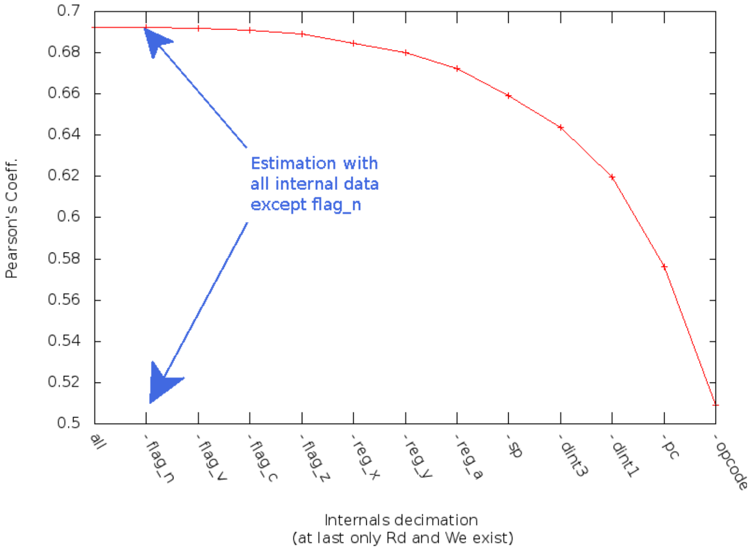
Multivariate Linear Correlation Hamming Distance



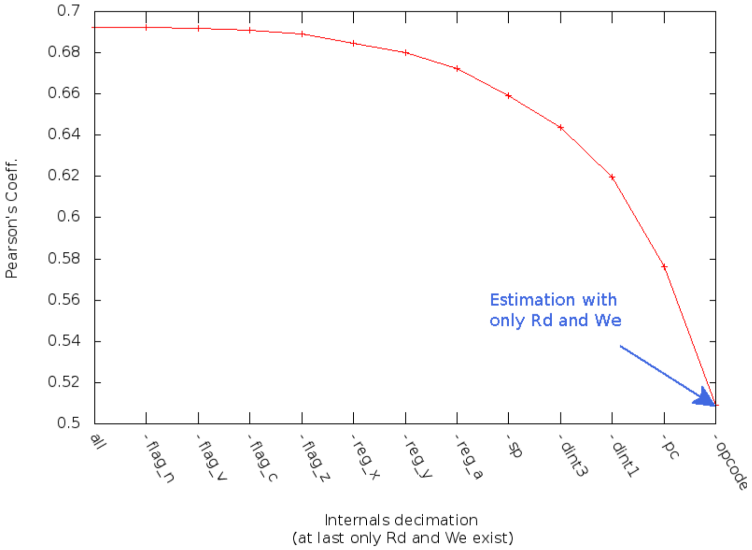
Multivariate Linear Correlation Hamming Distance



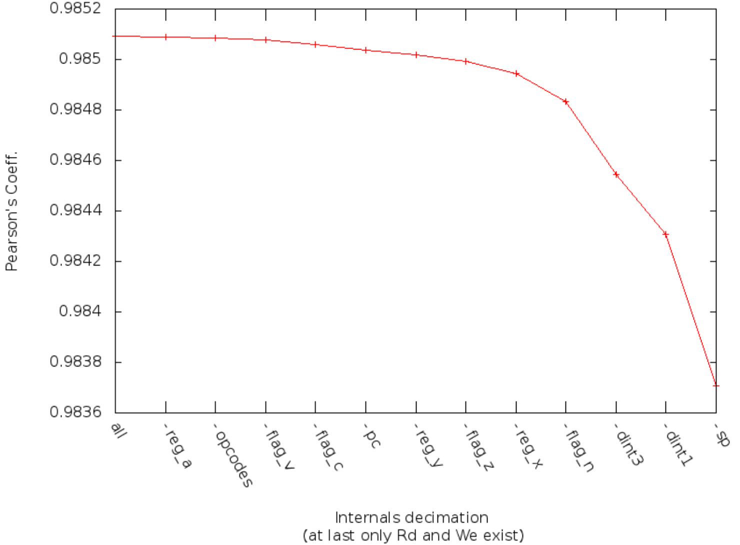
Multivariate Linear Correlation Hamming Distance



Multivariate Linear Correlation Hamming Distance



Multivariate Linear Correlation Hamming Weight



Opcodes Characterization

Each opcode has:

- a fixed length (number of cycles) that never changes
- a constant R/W signature, but some distinct opcodes have the same R/W signature

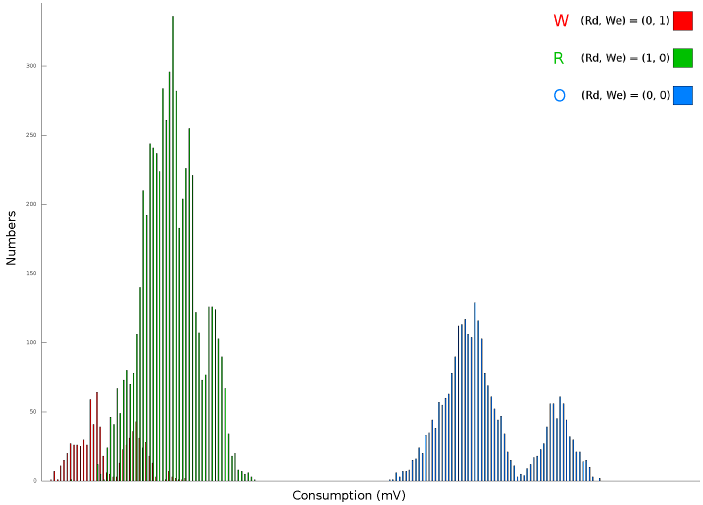
Use MODELSIM Simulation to Determine each Opcode R/W Signature

R = (Rd, We) = (1, 0) / W = (Rd, We) = (0, 1) / O = (Rd, We) = (0, 0)

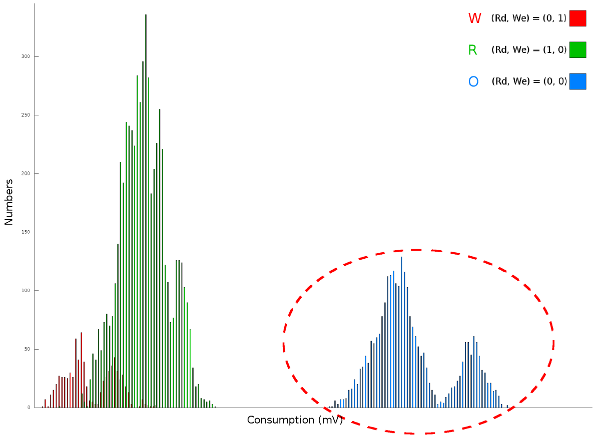
Signature	nb of opcodes with the same signature	Signature	nb of opcodes with the same signature
RR	21	RRWWWRR	1 *(brk)*
WRR	2 *(php, pha)*	ORRORRR	14
ORR	11	OROOWRR	12
ORRR	23	ORROWRR	2 *(sta)*
OORR	8	ORROORRR	1 *(reset)*
OWRR	6	OROOWRR	12
RORRR	28	RRWWWORR	2 *(s_nmi, s_irq)*
RWWRR	1 *(jsr)*	ORRROORR	1 *(rti)*
ROORR	1 *(jmp)*	ORROOORR	1 *(rts)*
ROWRR	5	RORROORR	1 *(jmp)*

Outline

Histogram consumption at the 9^{th} sample of each clock cycle, colors refer to the values of R/W signals.



First Histogram Characterization



ORRORRRWWRRWRRORRRRRORROWRROORRORRRRORROWRRORRRRRORROWR
O..O..........O.....O..O...OO..O....O..O...O.....O..O..

After this step, 28.49% of all R/W values are recovered
(all clock cycles at $(Rd, We) = (0, 0)$)

Examples of opcode signature rules

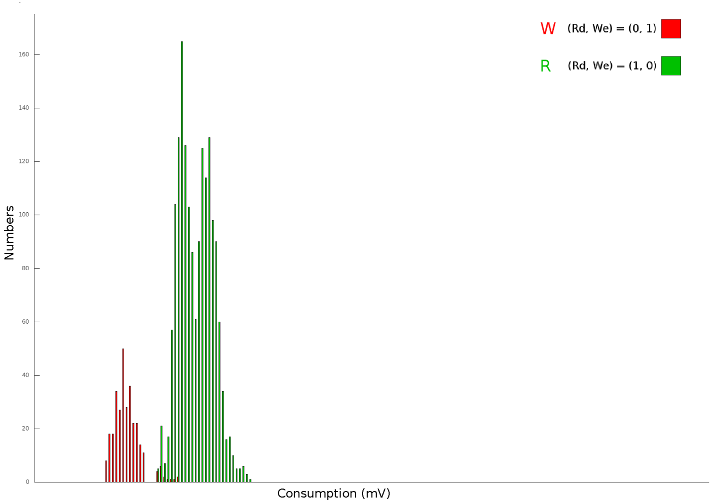
The knowledge of each the opcode's R/W signature permits to set up rules that the reconstructed signal should respect:

- "OOO" → "ORROOORR"
- "WWW." → "RRWWWRR"
- ".OO" → "R.OO.R"
- ".O" → "RO"
- ...

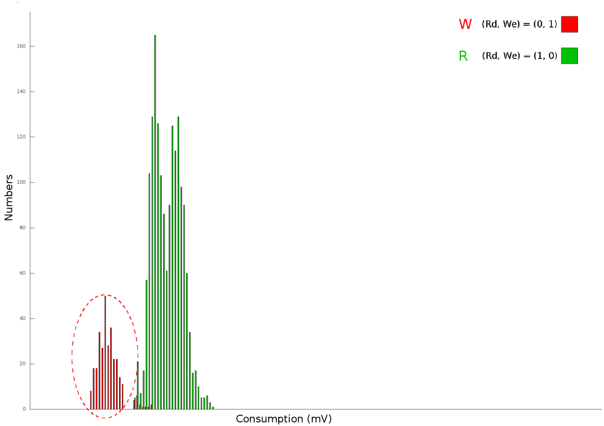
ORRORRRWWRRWRRORRRRRORROWRROORRORRRRORROWRRORRRRRORROWR
ORRO.R......RRO.R.RRORRO.RROORRORRRRORRO.RRO.R.RRORRO.R

After this step, 78.31% of all R/W values are recovered

Histogram consumption at the 9^{th} sample for all cycles that follow a $(Rd, We) = (0, 0)$ cycle, colors refer to the values of R/W signals.



Second Histogram Characterization



ORRORRRWWRRWRRORRRRRORROWRROORRORRRRORROWRRORRRRRORROWR
ORRO.R......RRO.R.RRORROWRROORRORRRRORROWRRO.R.RRORROWR

After this step, 81.80% of all R/W values are recovered
(only clock cycles at $(Rd, We) = (0, 1)$)

Pattern Matching for R/W Pattern Recognition

Process

- First step: calculate average consumption patterns.
- Second step: try to match the unknown with the known patterns.

```
ORRORRRWWRRWRRORRRRRORROWRROORRORRRRORROWRRORRRRRORROWR
ORRO.RRWWRRWRRO.R.RRORROWRROORRORRRRORROWRRO.R.RRORROWR
```

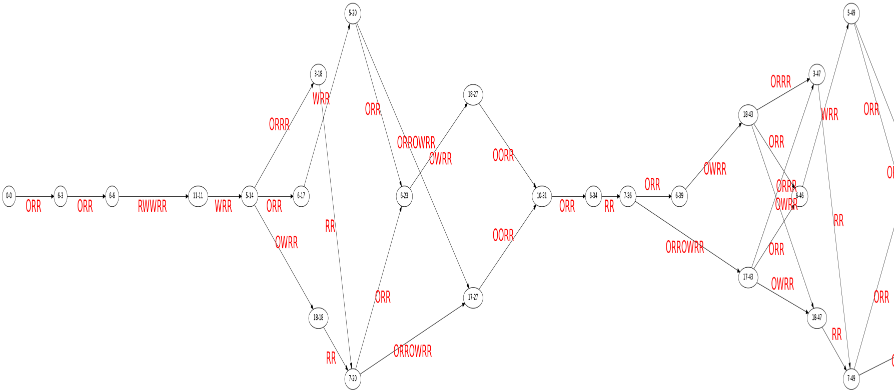
After this step, 87.32% of all R/W values are recovered

Exhaustive Search Using Tree Representation

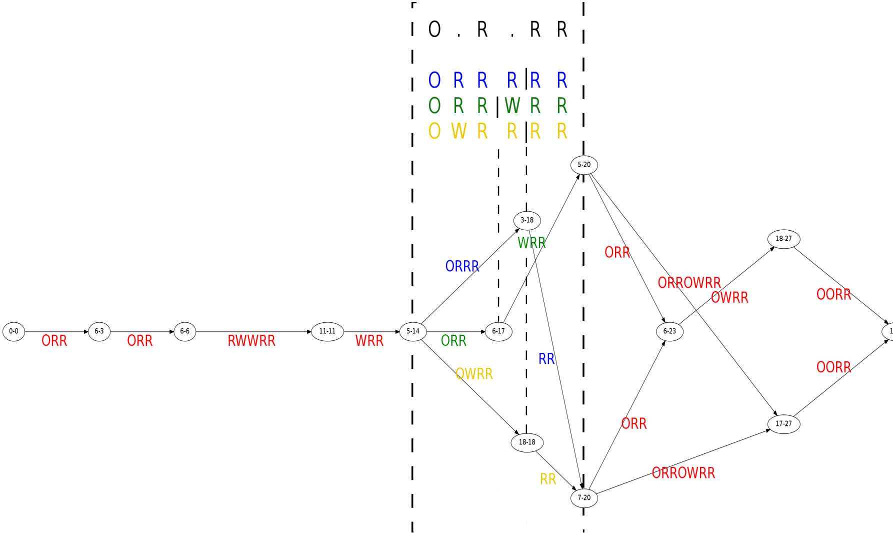
For this sequence:

ORRO.RRWWRRWRRO.R.RRORROWRROORRORRRORROWRRO.R.RRORROWR

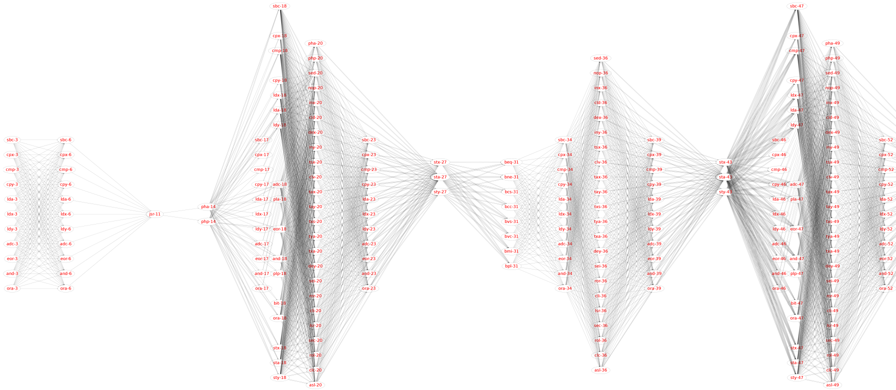
We obtain this tree representation using all opcode R/W signatures:



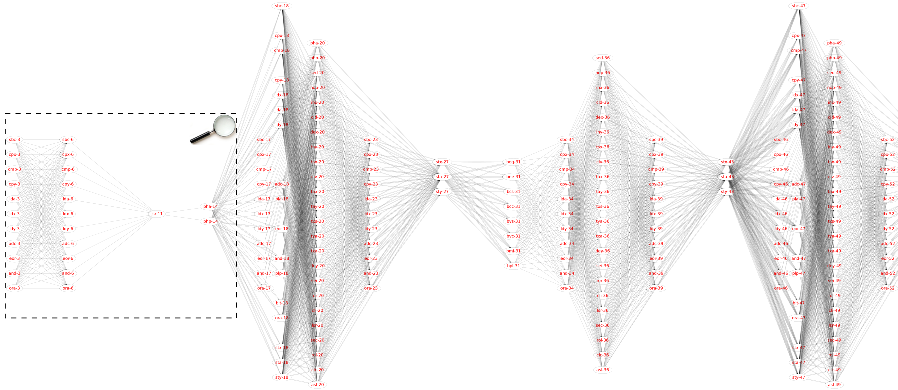
Exhaustive Search Using Tree Representation



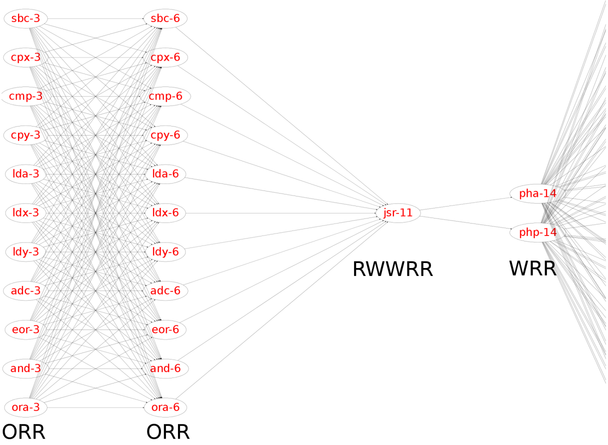
From the R/W Signature to Opcodes, the Tree Representation



From the R/W Signature to Opcodes, the Tree Representation



From the R/W Signature to Opcodes, the Tree Representation



Outline

1. Introduction
2. Characterization Phase
3. Read/Write Signals Reconstruction Process
4. Conclusion and Further Work

Conclusion

- Characterization of 6502 microcontroller.
- Process for R/W signal reconstruction.
- Extract possible opcodes' streams.

Further work

- Study the semantics of the language, to obtain probabilites on opcodes' sequences.
- Doing further experimentations on other executed codes.

Thank you for your attention.
Questions ?

Damien Marion **Antoine Wurcker**
damien.marion@unilim.fr antoine.wurcker@xlim.fr

XLIM-CNRS, University of Limoges, France

* *This work has been conducted under the framework of the MARSHAL+ french project*

COSADE 2013 - March 7, 2013