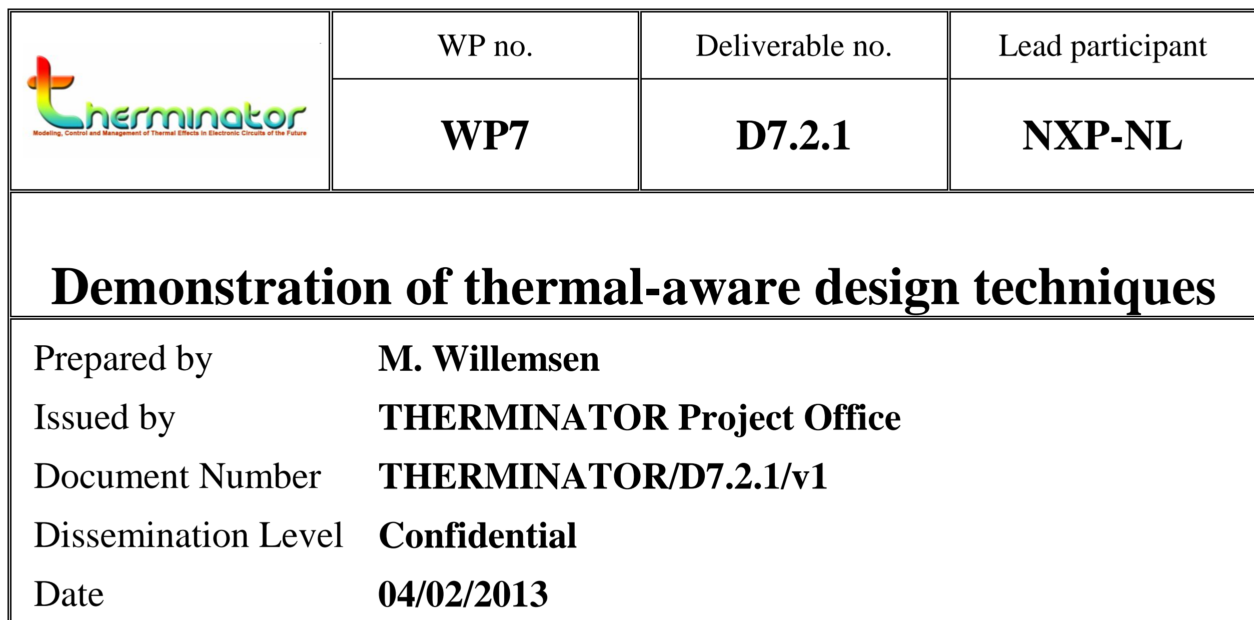
ICT-2009.3.2-248603-**IP**

Modelling, Control and Management of Thermal Effects in Circuits of the Future



	WP no.	Deliverable no.	Lead participant
	WP7	**D7.2.1**	**NXP-NL**

Demonstration of thermal-aware design techniques

Prepared by	**M. Willemsen**
Issued by	**THERMINATOR Project Office**
Document Number	**THERMINATOR/D7.2.1/v1**
Dissemination Level	**Confidential**
Date	**04/02/2013**

© Copyright 2010-2013 STMicroelectronics, Intel Mobile Communication, NXP Semiconductors, GRADIENT DESIGN AUTOMATION , MUNEDA, SYNOPSYS , BUDAPESTI MUSZAKI ES GAZDASAGTUDOMANYI EGYETEM , CSEM, FRAUNHOFER , IMEC, CEA-LETI, OFFIS, Politecnico di Torino, ALMA MATER STUDIORUM -Universita' Di Bologna, ST-Polito.

This document and the information contained herein may not be copied, used or disclosed in whole or in part outside of the consortium except with prior written permission of the partners listed above.

Document

Title	Demonstration of thermal-aware design techniques
Type	Deliverable CO
Ref	D7.2.1
Target version	V1_1
Current issue	V0_1
Status	Released
File	
Author(s)	D. Lorenz (IMC), C. Pacha (IMC), G. Greco (ST), S. Caporale (UNIBO), A. Macci (POLITO), A. Calimera (POLITO), R. Canegallo (ST), J.-L. Nagel (CSEM), A. Valentian (CEA-LETI)
Reviewer(s)	V. Melikyan (SNPS-AM), D. Rossi (UNIBO)
Approver(s)	G.Gangemi (ST)
Approval date	04/02/2013
Release date	04/02/2013

Distribution of the release

Dissemination level	CO
Distribution list	

History

Rev.	DATE	Comment
0.1	30-11-2012	Initial version
1.0	30-01-2013	Revised version, approved by all partners
1.1	04/02/2013	Check and ship out

References

[1] M. Ciappa, F. Carbognani, W. Fichtner: "Lifetime Prediction and Design of Reliability Tests for High Power Devices in Automotive Applications ", IEEE TRANSACTIONS ON DEVICE AND MATERIALS RELIABILITY, Vol 3, N0. 4,December 2003.

[2] R. Letor, S. Russo, R. Crisafulli "Life time prediction and design for reliability of Smart Power devices for automotive exterior lightining "CIPS2008" proceedings"

[3] H. Black, "Modulation Theory", D. V. Nostrand, Ed., 1953.

[4] B.F. Logan, "Information in the zero crossings of bandpass signals," *Bell System Technical Journal*, vol. 56, pp. 487–510, 1977.

[5] Z. Song and D. V. Sarwate, "The frequency spectrum of pulse width modulated signals," *Signal Process.*, vol. 83, pp. 2227–2258, October 2003.

[6] S. Caporale, R. Rovatti, G. Setti, "Representation of PWM signals through time warping," *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* , vol., no., pp.3589-3592, 25-30 March 2012.

[7] Thomas Baumann; Doris Schmitt-Landsiedel, and Christian Pacha, "Architectural assessment of design techniques to improve speed and robustness in embedded microprocessors," *Design Automation Conference, 2009. DAC '09. 46th ACM/IEEE*, vol., no., pp.947-950, 26-31 July 2009

[8] Neil H. E. Weste, David Money Harris "CMOS VLSI Design: A Circuits and Systems Perspective", Addison Wesley, 2011, pp.194-199

[9] Ali Dasdan, and Ivan Hom, "Handling inverted temperature dependence in static timing analysis", *ACM Trans. Des. Autom. Electron. Syst.* 11, 2, pp. 306-324, April 2006

[10] Therminator D6.3.2 - Presentation and evaluation of thermal-aware design techniques for 3D SiP stacks and 2D SoCs, Therminator Consortium, 2012

[11] Therminator D6.2.2 - Presentation and evaluation of an all level thermal simulator of 3D SiP stacks and 2D SoCs, Therminator Consortium, 2012

[12] Therminator D6.3.3 - Report on integration of the individual optimization techniques, Therminator Consortium, 2012

[13] M. Weiser, B. Welch, A. J. Demers, and S Shenker. Scheduling for reduced cpu energy. In *Proc. 1st. Symposium on Operating Systems Design and Implementations*. USENIX Association, 1994

This page was intentionally left blank.

Contents

Document 2

Distribution of the release 2

References 3

1 Introduction 7

2 Discrete Power Devices Design Framework and Design Improvement (ST) 8

2.1 Introduction 8

2.2 Design Framework and Design Improvement 9

2.3 Spared Time Estimation on the Whole Implementation Time 11

2.4 Reliability 12

2.5 Integrability 12

2.6 Conclusions 13

3 Class-D amplifier Design (UNIBO) 14

3.1 Introduction 14

3.2 Technical Results 15

3.3 Concluding Remarks 18

4 Comparative Analyses on Test-chips (IMC and POLITO) 19

4.1 Introduction 19

4.2 Comparative Analysis by IMC 19

4.2.1 The Demonstrator Chip 19

4.2.2 The Measurement Setup 22

4.2.3 Technical results 23

4.3 Comparative analysis by POLITO 27

4.3.1 Benchmark Description 27

4.3.2 Comparative Analysis Framework 29

4.3.3 Technical Results 30

4.4 Conclusions 38

5 Evaluation of a Thermal-aware 3D Physical Design Flow (ST, together with UNIBO) 39

5.1 Introduction 39

5.2 Design Flow Case Study 41

5.2.1 Memory and Logic Test-chip Fabricated in CMOS 65nm ST Process 41
5.3 Design Flow Validation Case Study 42
5.3.1 3D Stacking Configuration of DIE 1 and DIE 2 42
5.3.1.1 Stacking of DIE on Silicon Interposer 42
5.3.1.2 Stacking of DIE in a Face to Face (F2F) Configuration 44
5.3.2 Experimental electrical results: 46
5.4 Conclusion 47
5.5 Measurable Objectives (MO7.2.7) 47
6 Dynamic Thermal Management of Multi-Processor System-on-Chip 50
6.1 Introduction 50
6.2 Implementation of Genepy 51
6.2.1 Temperature Sensors 51
6.2.2 Activity Sensors 52
6.2.2.1 SMEP Cluster 52
6.2.2.2 IcySMEP Cluster 53
6.3 Correlation between Temperature Measurement and Activity Monitoring 55
6.4 Conclusion 56
7 Conclusions 58
8 Measurable objectives 59
9 Publications and presentations 61

1 Introduction

The main objective of WP7 is to validate the models, design techniques, and tools developed within the Therminator project. WP7 is divided into three tasks. In T7.1, validation of the thermal models for new devices, materials, and technologies will be done. The focus will be on the device level, and the level of elementary building blocks to be used in large(r) circuits. The effectiveness and usability of design techniques is addressed in T7.2. In T7.2, larger building blocks, parts of circuits, and test chips are used as test cases. Finally, benchmarking and demonstration of the effectiveness of the developed EDA tools is done in T7.3. In this task, the validation addresses test chips and prototypes.

The demonstration of design techniques on test-chips in T7.2 is reported in this deliverable. In the demonstration activities, the tools and design flows of WP2, 3, 5, and 6 of Therminator are used. The test-cases are taken from WP1, and examples provided by individual partners. Besides the aforementioned links to other WPs within Therminator, some of the results are also related to the other WP7 tasks T7.1 and T7.3. ST will demonstrate power devices with improved thermal behaviour using the thermal-aware design framework. The thermal aware design framework itself was validated on existing material in T7.1. The PSP-based modelling of advanced CMOS from IMC is another example of a design tool verified in T7.1 that is now used at circuit level. On the other hand, POLITO compares measurements of designs in advanced CMOS to simulations done with their thermal-aware design flow. The thermal-aware design flow will be used for optimizations in T7.3. Another example is the work on thermal management of Multi-Processors on System-on-Chip by CSEM and CEA-LETI. The measurement results will be used for validation of the tools from OFFIS, developed in Therminator, in task T7.3.

In summary, the thermal aware design framework of ST is used in chapter 2 to improve the design of power devices resulting in a lower peak temperature during operation. In chapter 3, UNIBO presents novel design techniques for class-D amplifiers, leading to reduced power consumption. IMC demonstrates the usefulness of the PSP based design flow for advanced CMOS by comparing simulations and measurement of performance and power consumption on a generic critical path structure in chapter 4. In chapter 5, ST demonstrates their thermal aware 3D physical design flow developed in WP6. The tools are evaluated on prototypes of 3D stacked devices. CSEM and CEA-LETI present work on thermal management of Multi-Processors on System-on-Chip in chapter 6. A summary of all of the results is given in the conclusions, i.e., chapter 7. The results of T7.2 are quantified in terms of measurable objectives. An overview of all of these measurable objectives is given in chapter 8 of this report. In this chapter, the measurable objectives are also linked to Therminator's project objectives. The novelty of the work is addressed in chapter 9, where all of the output in terms of journal papers and conference contributions is collected.

2 Discrete Power Devices Design Framework and Design Improvement (ST)

2.1 Introduction

The full implementation of a modern discrete power device imposes a series of steps that require a significant amount of time to be accomplished.
This implementation flow could be briefly described with the following diagram flow:

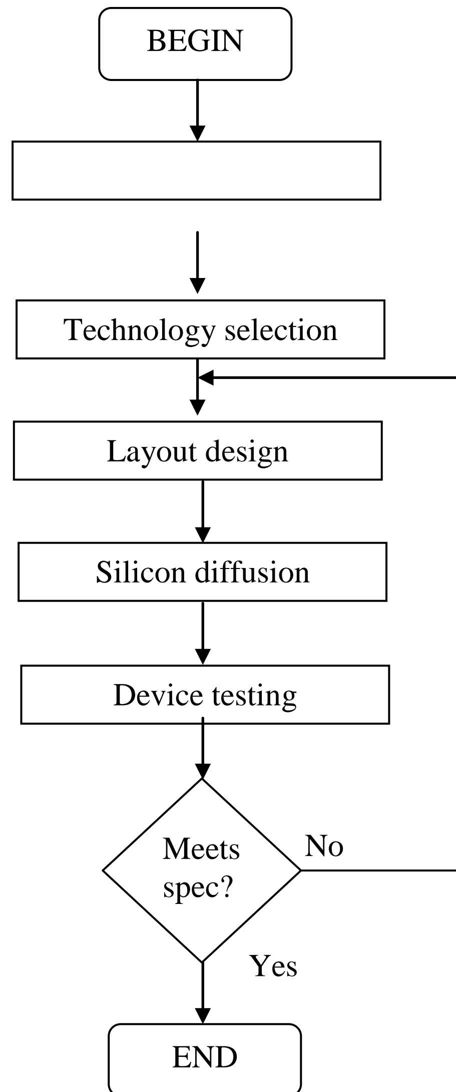


Figure 1 - Device implementation flow

It is evident that the weak point of this procedure is represented by the possibility of repeating several time part of the flow if the device does not meet the initial specifications. In particular, the item defined "Layout design" is the one that will be redefined since the good electrical and thermal performance of a device will depend on the characteristics of its layout (die area, gate pads, gate fingers, bonding, etc.). As already explained in deliverable 5.1.1 and 5.1.2, the lack of valid commercial tools that could help designers in testing in pre-silicon phase the goodness of the layout structure of a discrete power device lead consortium partners to

implement the thermal design framework developed in WP5. Thanks to this innovative design approach designers will be able to reduce sensitively the tentative silicons and therefore significantly reduce the device implementation time.

2.2 Design Framework and Design Improvement

This part of the document describes how the step flow named "Layout design" has been enhanced with a sub-flow which implies the use of the WP5 design framework (DF) (Figure 2).

Layout design
Tentative layout solution
DF Layout simulation
Meets spec?
No
Yes
Set up device specificat ion
Layout design

Figure 2 – Enhanced layout design flow

In order to meet required electrical and thermal specifications and, therefore, verify the assigned mission profile related to the application, a new layout solution has been analyzed where the three bond wires have been substituted with a three-ribbon solution. This approach would allow increasing the focus area of the current because of a bigger footprint area by obtaining average temperatures values lower than the one retrieved in the former case. Ribbons have a larger metallic mass than bond wires and therefore a larger footprint. This increase in terms of mass would increase the thermal capacitance of the system. A reduction of the device R_{on} will also be expected through this solution.
A new CAD layout has been drawn (only abstraction layers) and the related numerical matrix (Figure 3) has been generated by the Matrix Extractor Module.

```
0000000000000000000000000000000000000000000000000000000000000
0000000888888888888888888888888888888888888888888888800000000
0000004222222222222222222222222222222222222222222222240000000
0000422222222222222222222222222222222222222222222222222240000
0004222222222222222222222222222222222222222222222222222224000
0042222222222222222222222222222222222222222222222222222224000
0042222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222999999999999999999999992222400
```

```
0422222222222222222222222222299999999999999999999999222222400
0422222222222222222222222229999999999999999999999999222222400
0422222222222222222222222299999999999999999999999922222222400
0422222222222222222222229999999999999999999999999222222222400
0422222222222222222222299999999999999999999999992222222222400
0422222222222222222299999999999999999999999992222222222222400
0422222222222222222999999999999999999999999922222222222222400
0422222222222222229999999999999999999999992222222222222222400
0422222222222222999999999999999999999999922222222222222222400
0422222222222999999999999999999999999999222222222222222222400
0422222222222999999999999999999999999922222222222222222222400
0422222222229999999999999999999999999222222222222222222222400
0422222222299999999999999999999999922222222222222222222222400
0422222222999999999999999999999922222222222222222222222222400
0422222229999999999999999999999222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222999999999999999999999922222400
0422222222222222222222222222299999999999999999999999222222400
0422222222222222222222222229999999999999999999999999222222400
0422222222222222222222222299999999999999999999999922222222400
0422222222222222222222229999999999999999999999999222222222400
0422222222222222222222299999999999999999999999992222222222400
0422222222222222222299999999999999999999999992222222222222400
0422222222222222222999999999999999999999999922222222222222400
0422222222222222229999999999999999999999992222222222222222400
0422222222222222999999999999999999999999922222222222222222400
0422222222222999999999999999999999999999222222222222222222400
0422222222222999999999999999999999999922222222222222222222400
0422222222229999999999999999999999999222222222222222222222400
0422222222299999999999999999999999922222222222222222222222400
0422222222999999999999999999999922222222222222222222222222400
0422222229999999999999999999999222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222999999999999999999999922222400
0422222222222222222222222222299999999999999999999999222222400
0422222222222222222222222229999999999999999999999999222222400
0422222222222222222222222299999999999999999999999922222222400
0422222222222222222222229999999999999999999999999222222222400
0422222222222222222222299999999999999999999999992222222222400
0422222222222222222299999999999999999999999992222222222222400
0422222222222222222999999999999999999999999922222222222222400
0422222222222222229999999999999999999999992222222222222222400
0422222222222222999999999999999999999999922222222222222222400
0422222222222999999999999999999999999999222222222222222222400
0422222222222999999999999999999999999922222222222222222222400
0422222222229999999999999999999999999222222222222222222222400
0422222222299999999999999999999999922222222222222222222222400
0422222222999999999999999999999922222222222222222222222222400
0422222229999999999999999999999222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222222222222222222222222222222222222222400
0422222222222222222222211111111111111222222222222222222222400
0422222222222222222222211111111111111222222222222222222222400
0042222222222222222222211111111111111222222222222222222222400
0042222222222222222222211111111111111222222222222222222224000
0004222222222222222222211111111111111222222222222222222224000
0000422222222222222222211111111111111222222222222222222240000
0000004222222222222222211111111111111222222222222222224000000
```

00000008888888888888888333333333333338888888888888888880000000
00000000000000000000000000000000000000000000000000000000000

Figure 3– Matrix for the DUT with three ribbons

To perform the simulation by using the same test bench and the same active models already used in T7.1, we obtain the result reported in Fig.4 where a significant decrease of peak temperature is retrieved by passing from 55.68 °C to 47.46 °C with an improvement of about 8 °C.

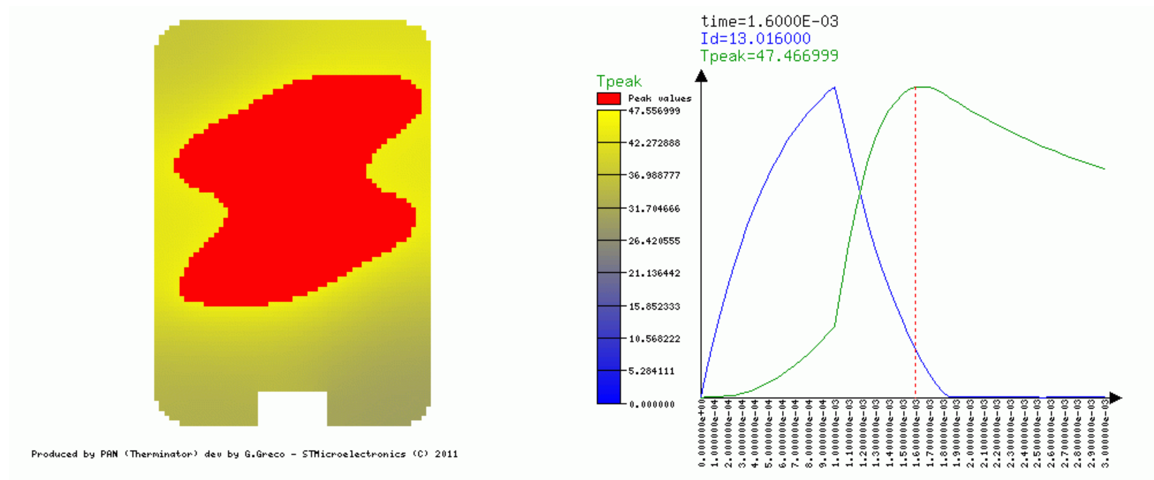


Figure 4 – Drain current (blue), temperature peak in (green) and thermal map (left) during a thermal cycle of improved DUT.

The strength of this approach is that designer could use it not only for guaranteeing performances within specific ranges but also for optimizing results to center it as shown in this last simulation.

2.3 Spared Time Estimation on the Whole Implementation Time

If n_1 is the average number of silicon tentatives with standard design flow and n_2 the number of silicon tentatives by supposing the use of the Design Framework, the spare time could be estimated as follows:

$$SpT = \frac{T_{flow} \cdot (n_1 - n_2)}{T_{flow} \cdot n_1} \cdot 100 \quad => \tag{1}$$

$$SpT = \frac{(n_1 - n_2)}{n_1} \cdot 100 \tag{2}$$

Where:

SpT is the spared time in percentage and T_{flow} the time necessary to complete a cycle of the flow in both cases (old flow and newer one) by assuming that the weight of simulation phase is negligible with respect to the whole production time where, i.e., only the diffusion phase could also take a couple of months. The worst case is obtained when $n_1 = 2$ (only two silicon tentatives) and n_2=1 (silicon working at first implementation) that is:

$$SpT=50\%$$

It means that, in this case, the whole reduction time is about 50% that is a very interesting result. If we suppose that the number of silicon tentatives could be greater than 2 (actually not a frequent case) the spared time would increase in a more significant way.

2.4 Reliability

In an automotive context, the main objective of the developed Design Framework is to reduce failures on applications field. As it turns out [1][2], due to the strong relationship between junction temperature peak and number of cycles that a device could guarantee (see Figure 5), this approach will help designers to manage aspects related to reliability in a better way.

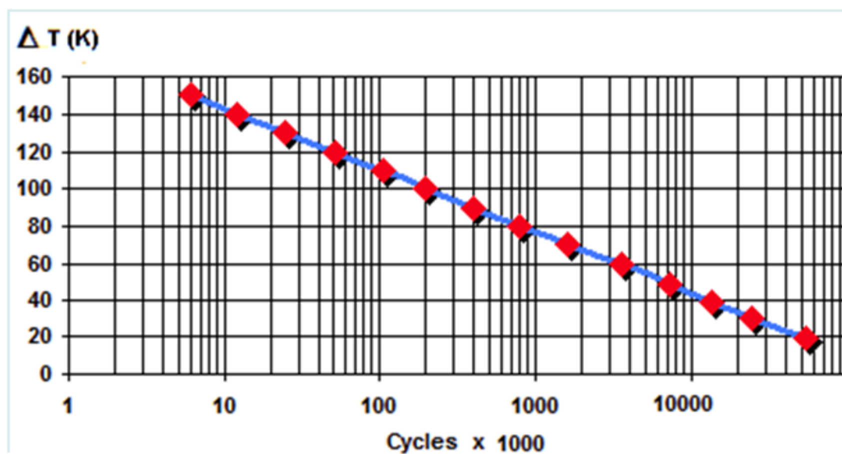


Figure 5 - Cycle lifetime estimation versus junction temperature variation

2.5 Integrability

The implemented flow is already deployed in STMicroelectronics design flow. Wp5 discrete power device modeling flow is based on Cadence DFII platform, an industrial standard in microelectronics companies.

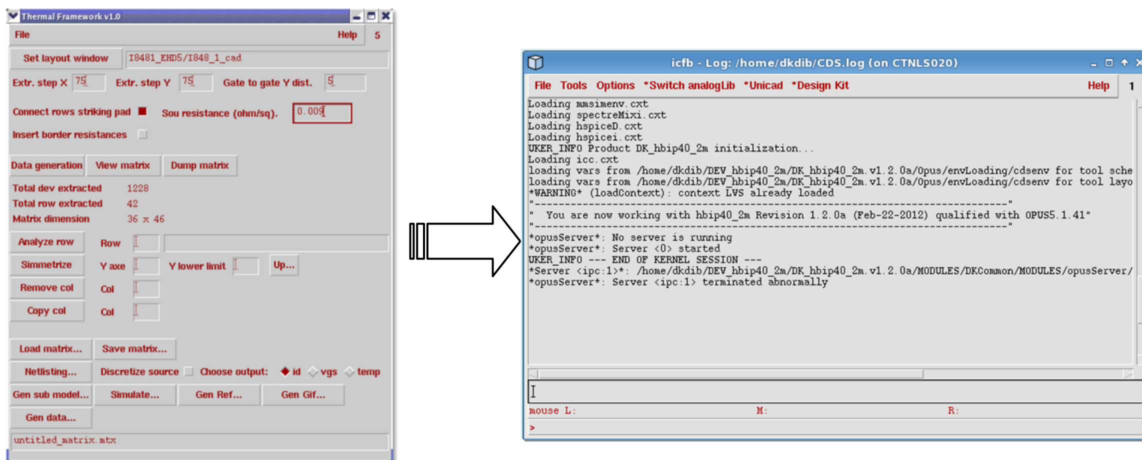


Figure 6 The Thermal Framework is fully integrated in Cadence Design Framework II, an industrial standard design platform.

2.6 Conclusions

As shown, the design phase represents a very critical aspect of the flow reported in Figure 1. Differently from the design of silicon integrated circuits where a lot of commercial electronics design automation tools are available, in the circumscribed field of discrete power devices the lack of consolidate design tools make the layout phase very hard to be accomplished and often the designer “experience” is the only available tool.

Thanks to the design framework, now it is possible to predict the electrical and thermal behaviour of discrete power devices (MOSFET, IGBT, Bipolars, IPEM, etc.) by having a precise estimation on possible layout weak points.

Testing the pre-silicon device internal thermal behaviour directly on the customer application is now possible and if weaknesses are spotted consequent opportune actions on layout structure could be carried out. Through the implemented design flow tentative silicons due to thermal design weakness could be almost wiped out by obtaining a reduction of costs and implementation time that can easily be of 50% (MO7.2.1).

3 Class-D amplifier Design (UNIBO)

3.1 Introduction

In this work we demonstrate the possibility of increasing the efficiency of class-D amplifier, thus reducing thermal effects and power consumption, by designing and implementing a new approach to the Pulse Width Modulation (PWM) which is the intrinsic processing of class-D amplification.
According to PWM, a source signal whose bandwidth is *B* is converted to a two-level signal such that the information is coded in the transitions instants, that is zeros positions, rather than in the signal amplitude. The modulation is obtained by first comparing the signal to a reference wave, i.e. a saw-tooth or triangle wave oscillating at a proper frequency which must exceed *B*, and then considering the sign of the resulting signal. The resulting average switching frequency is equal to the reference wave frequency, whose value normalized to *B* will be referred to as *R*. Within *B*, the obtained PWM signal is approximately equal to the modulating signal, while non-linear modulated components arise outside *B* which also slightly interfere inside *B* **Error! Reference source not found.Error! Reference source not und.Error! Reference source not found.**.
Since the class-D amplifier always works in saturation mode except when the PWM signal switches, power which is not delivered to the charge, i.e. dynamic power, is consumed only during transitions. Hence, reducing *R* might be a strategy to decrease thermal effects. Nevertheless, PWM in-band non-linear interfering components grow as *R* decreases, so *R* is generally chosen as a trade-off between signal quality and power consumption.
Although PWM implementations can be either analog or digital, nowadays the digital realization is always considered because of difficulties in generating a precise reference wave. Nevertheless, in digital implementations transitions are quantized in time so the resulting PWM signal suffer from aliasing which might be attenuated by some strategies like sigma-delta modulations. Moreover, *R* cannot be arbitrarily because of stability constraints.
Here we propose a new paradigm for balancing power consumption and interference which allows for a thermal-aware design. In case the source signal power is decreases, which may happen because of an intrinsic variation or because of a power control operated by a device or a user as represented in Figure 7, the in-band interfering components decrease non-linearly, so that *R* can be decreased without affecting the signal accuracy. In order to realize this control, ordinary digital PWM implementations are not suitable, so a new PWM mixed digital-analog implementations is also proposed.

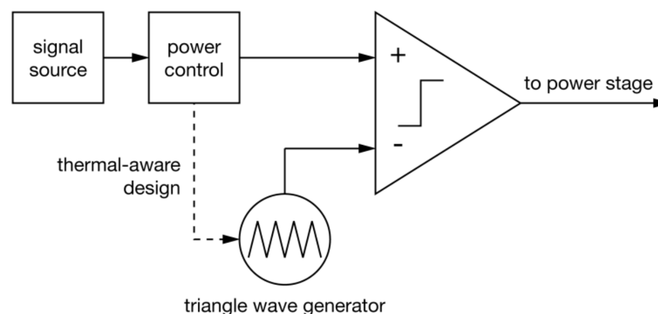


Figure 7 thermal-aware design of a class-D amplifier

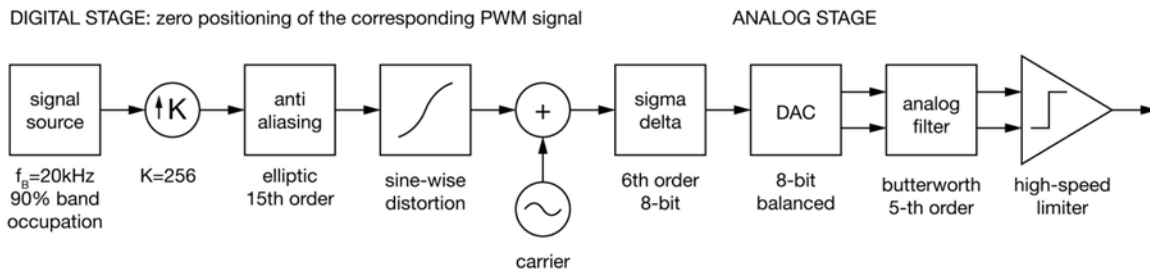


Figure 8: proposed implementation of PWM

3.2 Technical Results

The proposed PWM scheme has been reported in Figure 8 and will be detailed in the present section. The basic idea behind this approach is to digitally obtain a signal whose analog realization has the same zero positions as the difference between the source signal and the reference wave. This signal could be obtained in many ways, nevertheless we look for the implementation employing the minimum effective bandwidth, so that the approximations due to working in the discrete-time domain are minimized. Then the obtained signal is first converted to the analog domain and then is fed into a comparator which gives back a PWM signal which is theoretically equal to the PWM signal obtained by comparing to the reference wave.

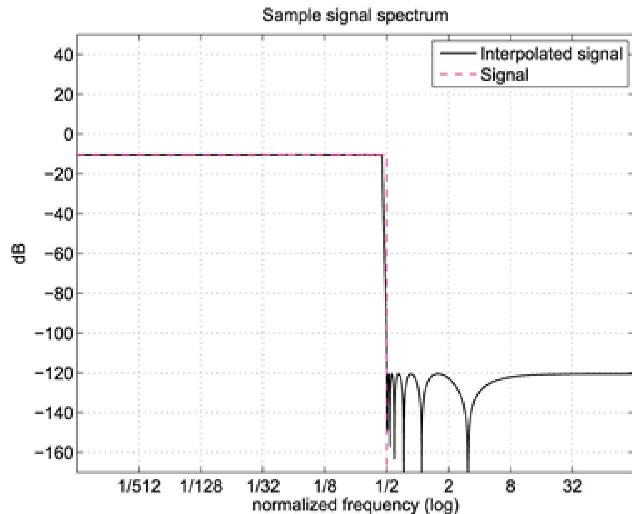


Figure 9: spectrum of the test interpolated source signal

In order to accomplish this aim, we considered that the crossings between two different signals, that is the zero positions of their difference, are not altered in case the same distortion is applied to both. Moreover, a proper scaled sine-wise distortion applied to a triangle wave gives back just a single harmonic on R. On the contrary, when the distortion is applied on the signal, the effective bandwidth is increased because of the addition of harmonics.

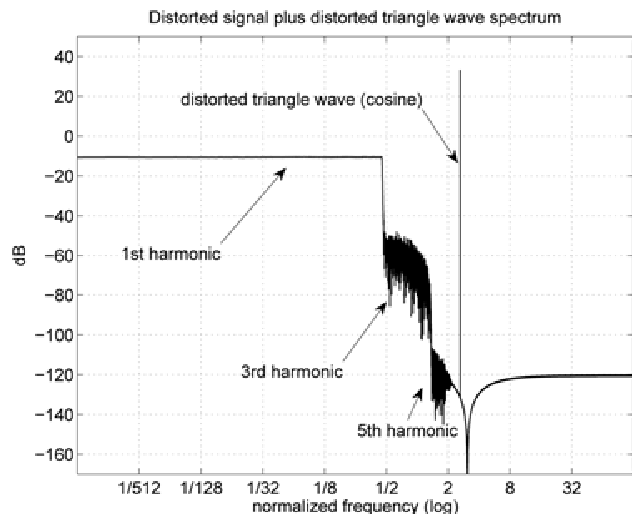


Figure 10: signal spectrum after sine-wise distortion

So, first the input signal is upsampled such that the forthcoming distortion is not causing aliasing. In Figure 9 we considered a flat spectrum signal with 100% band occupancy which, according to our design, is upsampled by K=256

and filtered by an elliptic filter of 15-th order which guarantees a 10% transition bandwidth. To preserve accuracy, the transition must not exceed *B*. Since the specification is to keep the noise floor below -100 dB, the filter attenuation is set to -120 dB. *K* also takes into account following processing steps. In Figure 10 the effects of the sine-wise distortion are shown.

To convert the signal to the analog domain a sigma-delta converter is employed. The design has to take into account many constraints:

- the bandwidth to be preserved after distortion is roughly 8*B*, such that the resulting OSR is *K*/8=32;
- by considering an audio bandwidth B=20kHz and *K*=256 the output rate is approximately 10Msample/s which cannot further increased;
- to control the out of band peaks after analog filtering (see Figure 12) the OSR has to be further reduced to 5, thus requiring a high order sigma-delta order;
- to limit the in band noise to -100 dB a 8-bit quantization is needed.

The spectrum after sigma-delta modulation is shown in Figure 11 where it is visible that the converted bandwidth exceeds requirements. By doing so, the steep increase of the NTF is moved on a band where it can be attenuated by the further analog filtering, as it can be observed in Figure 12, where the quantization noise major peak is limited to -120 dB. A 5-th order Butterworth filter has been employed. The filter order has to be even since a current steering DAC has been considered (THS5641).

Because of the real implementation of the analog filter, the match between the distorted reference wave and distorted source signal is slightly corrupted. In order to evaluate this mismatch, the

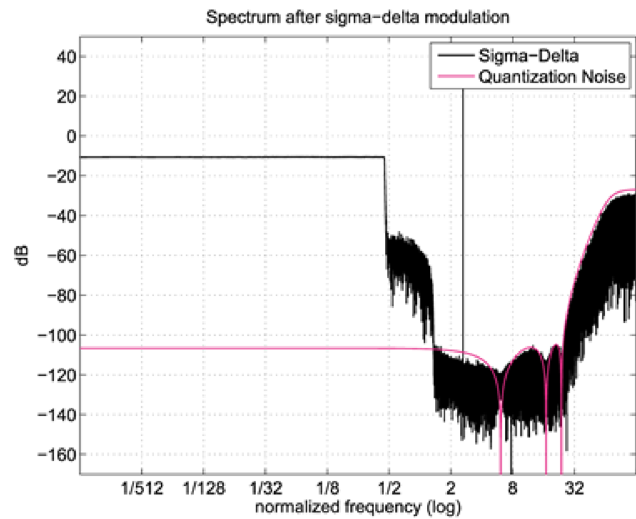


Figure 11: signal spectrum after sigma-delta modulation

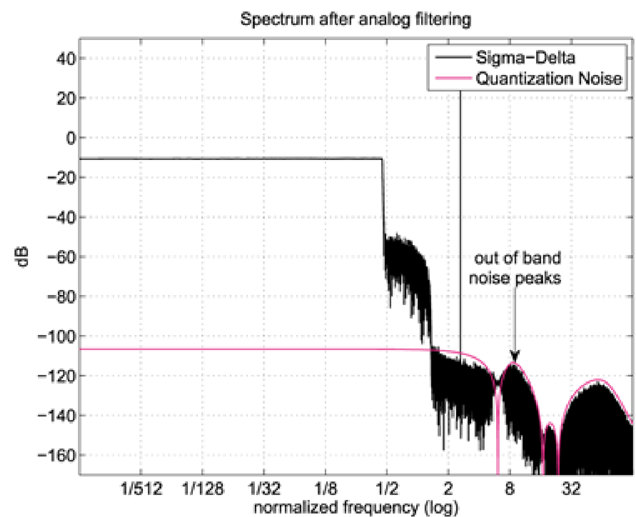


Figure 12: computed spectrum after analog filtering

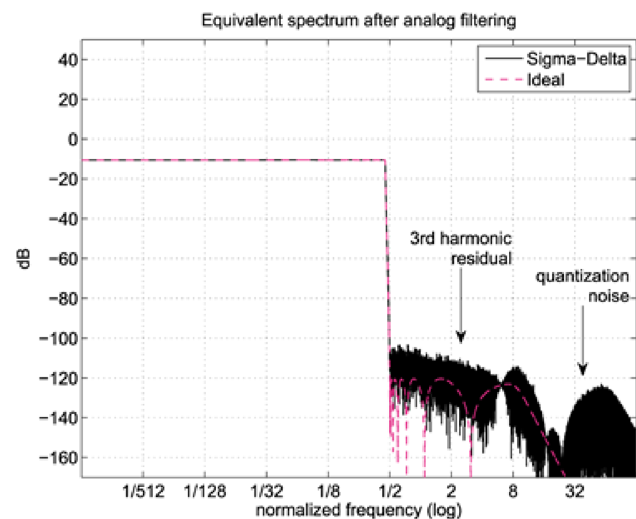


Figure 13: computed effective signal spectrum

equivalent undistorted signal as it was compared to an ideal triangle wave can be computed, as shown in Figure 13. As expected, by inverting the distortion a residual non-linear interference arises, which is anyway kept below -100 dB.

Finally, by comparison to an ideal PWM modulation performed on the analog converted source signal, the error introduced by the proposed scheme can be evaluated. As illustrated in Figure 14, the resulting error is compliant to the accuracy specification of -100 dB.

We point out that in the considered design maximum R has been set to approximately 5 which is about half the value of most common PWM digital realizations. Thanks to the proposed thermal-aware design, R can be further reduced when the signal power decreases. In more detail, the fulfilment of the measurable objective MO7.2.2 of reducing the driving frequency by 50% is demonstrated through Figure 15 and Figure 16. In these figures the PWM spectrum has been split in the linear part representing the signal, as it can be observed by comparison with Figure 13, and non-linearities added by PWM modulation. In Figure 15 the system is operating at maximum frequency and in-band noise is mainly conditioned by implementations imperfections since the noise floor in the signal is larger than the noise floor in non-linearities. Instead, in Figure 16 the signal power is 15 dB smaller than its maximum, hence the 4th power modulated on R is smaller than 100 dB and R can be reduced by 50%. At the same time, when decreasing R, the system is less sensitive to non-idealities, so the noise floor in the signals is smaller than the noise floor in non-linearities.

Finally, we note that such a driving frequency reduction is possible because the zero positioning is depending on the sigma-delta only and is performed with an accuracy in order of nanosec,

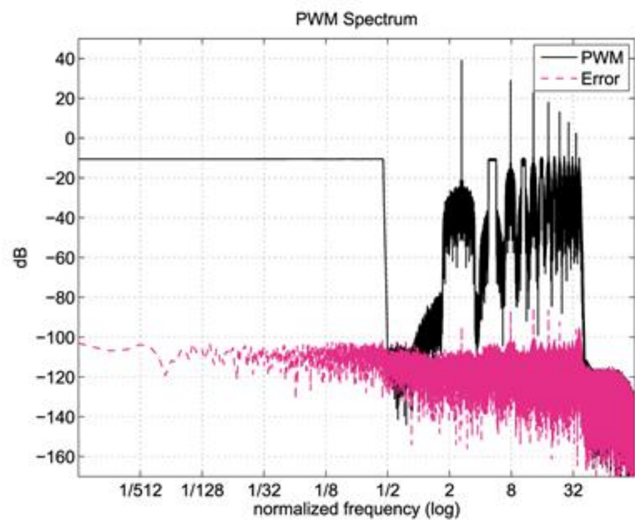


Figure 14: computed PWM spectrum and error

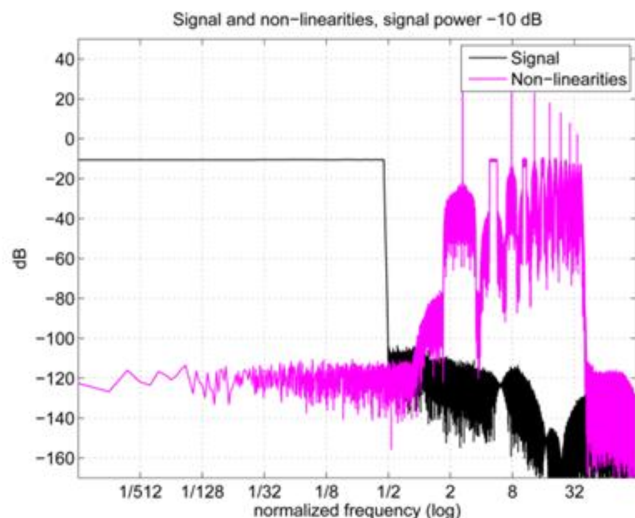


Figure 15: R=5, maximum needed driving frequency

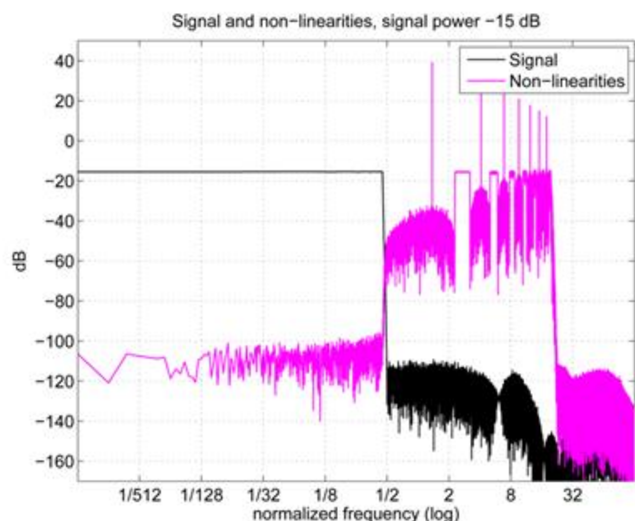


Figure 16: R=2.5, 50% reduced driving frequency

meeting the measurable objective MO7.2.3 of realizing an error in zero positioning in the order of nanosec.

3.3 Concluding Remarks

The proposed PWM modulation scheme has been implemented over a FPGA. A custom board featuring the current steering DAC converter THS5641, the analog filtering and the high-speed comparator Linear Technologies LT1719 has been realized, which is shown in Figure 17.

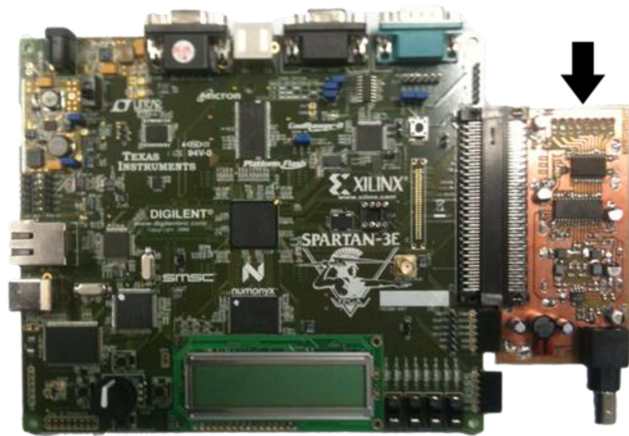


Figure 17: FPGA and custom board emplyed for tests

Some preliminary experimental measurements have been acquired by a digital signal analyser. A band-pass half-band signal has been synthetized in order to make in band non-linear interference visible. The employed instrument dynamic range is limited to 90 dB, hence it was not possible to verify the implementation accuracy, nevertheless obtained measurements are compliant to the expected results, as it can be observed in Figure 18.

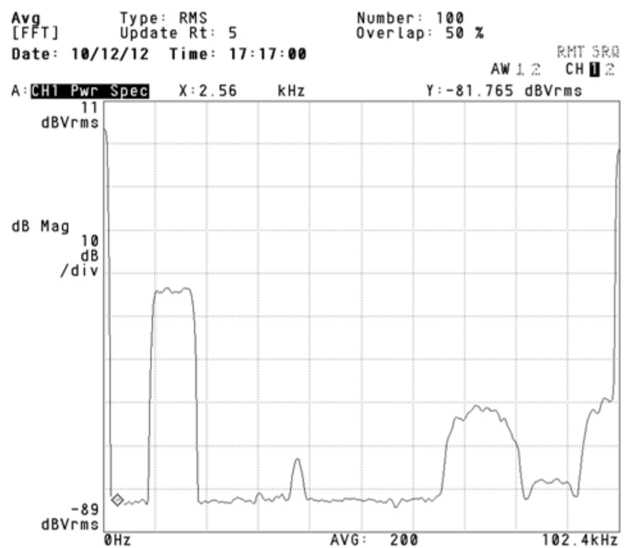


Figure 18: experimental spectrum by a digital signal analyser

Further measurements are planned to be performed by high resolution analog instrumentation.

In conclusion, the measurable objective MO7.2.2 of reducing the driving frequency by 50% has been demonstrated in Figure 16, while the measurable objective MO7.2.3 of positioning zeros with a tolerance in order of nanosec is implied by the achievement of 100 dB SNR or 16 bit precision, since $1/(2^{16}*B) \sim 1$ nanosec.

4 Comparative Analyses on Test-chips (IMC and POLITO)

4.1 Introduction

Objective of this document is the assessment, by POLITO and IMC, of the validation on test chips of the activities done in WP3 (T3.2). In particular, the comparative analysis done in the Deliverable D3.2.2 on different technology libraries has been carried out by IMC on their own demonstrator (See Section 4.2) and by POLITO on the test case 3 defined in the Task T1.3 (See Section 4.3).

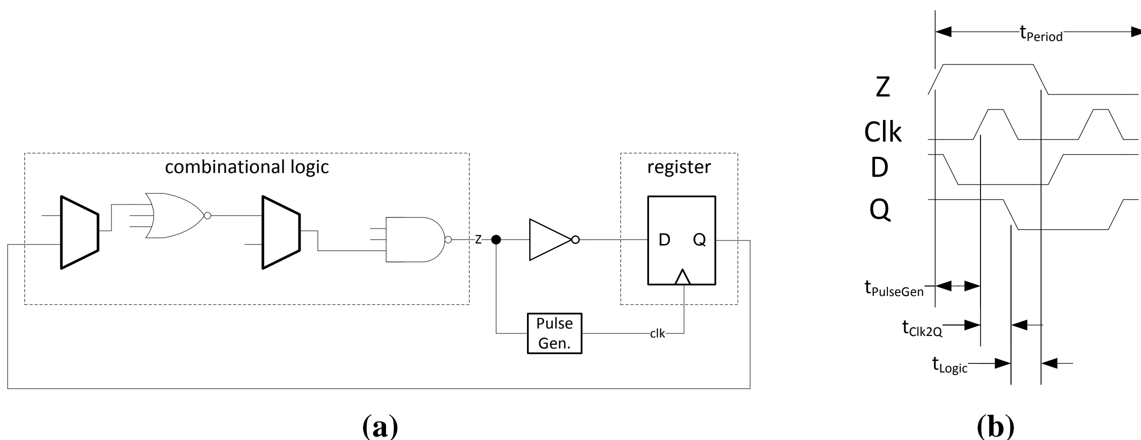
4.2 Comparative Analysis by IMC

Accurate models are crucial for the design of digital CMOS circuits to exploit the full potential of a new silicon technology node. The models must be sufficient accurate throughout the complete operating range spanned by temperature and supply voltage. Otherwise performance is wasted, the power consumption is higher than necessary and the circuit is not as competitive as it might be with more accurate models. In Therminator we turn our attention particularly on the correct modeling of the temperature dependence of the transistor models (and parasitic models as well) which are required for circuit-level simulations. These models must be accurate since the more abstract models (e.g. timing model for standard cells used for a static timing analysis) are based on circuit-level simulations.

4.2.1 The Demonstrator Chip

The contribution of Intel Mobile Communications (IMC) to task T7.2 of work package WP7 is the measurement of performance and power consumption (dynamic as well as leakage) on a generic critical path structure. The critical path structure is fabricated on a test chip in a 28nm foundry technology. The results are compared to post-layout simulations to verify the accuracy of the underlying models.

The task contributes to the third project objective[1]. The new transistor models are integrated into the existing mixed-signal design flow and their accuracy is validated by comparing simulation results against measurements of the generalized critical path on the demonstrator.



(a) (b)

[1] "Demonstration of the accuracy and ease of integration within existing design flows of the new models by validation against measured data obtained on ad-hoc silicon structures".

Figure 19: (a) Principle of the ring oscillator configuration of the generalized critical path. (b) Corresponding signal waveforms

The critical path structure of the Demonstrator is a generalized register-logic-register path. The generalization has been done based on a standard PrimeTime SI static timing analysis of an industrial ARM1176 micro-processor design followed by an in-depth post-processing to extract critical paths with high circuit sensitivities [7]. The selected register-logic-register path comprises simple NAND/NOR gates, complex CMOS logic gates and a register. All gates have typical device and interconnect loads. Compared to a generic path (e.g. path just consisting of simple inverters) such a critical path has the advantage that it better projects the behaviour of a real logic circuit (i.e. variations due to process and temperatures are more pronounced in complex CMOS logic gates with 2-3 transistors in a stacked configuration).

For a digital logic circuit the two most important properties are the maximal operating frequency and the power consumption. Hence, it must be checked whether those two properties of a circuit can be projected by the given transistor models.

The maximal operating frequency is determined by the delay of the critical path. For a synchronous logic circuit it must be guaranteed that within one clock period a signal can propagate from the sending register, through the combinatorial logic, to the receiving register. This means that the maximal operating frequency is determined by the clock-to-q delay of the sending register, the maximal delay of the combinatorial logic in between the sending and the receiving register and the setup time of the receiving register. The setup time is the minimal time the signal must be stable at the receiving register before the capturing clock edge in order to safely capture the signal. The overall contribution of the flip flop timing overhead in a critical path is on the order of 80-150ps (depending on PVT condition). Therefore, for an industrial high performance design with clock frequencies between 600MHz and 1GHz considering this contribution is essential.

To measure the performance of the generic critical path on the demonstrator test chip, the path is operated in a ring oscillator configuration. The ring oscillator configuration works as follows (see Figure **19**): The combinational logic path is connected to the data input of the register and via a pulse generator (PG) to the clock input of the register as well. The output of the register is fed back to the combinational path. The purpose of the pulse generator is to generate a pulse for every transition at its input. The result is a free running oscillator, since the clock signal for the register is generated from the path itself. That way it is possible to measure the delay of the generalized critical path including the clock-to-q delay of the register and some offset delay for the pulse generator.

In technologies above 90nm a circuit becomes slower as soon as the temperature increases due to the dissipated power. For instance, CPUs are cooled to achieve fault free operation. The reason is that the mobility of the carriers in the transistor channel is reduced with increasing temperature and the gate delay is inversely proportional to the carrier mobility. In modern technologies the situation looks a bit different. Dependent on the supply voltage the performance either increases or decreases with higher temperature. At high supply voltages, like in former technologies, the delay increases with increased temperature. However, for low supply voltages the dependence of the temperature on the performance is inverted [9] – the delay decreases with increased temperature. The reason is that now a second effect dominates. The threshold voltage of a transistor is reduced with higher temperature. The gate delay is directly proportional to the threshold voltage. For low supply voltages the overdrive voltage

(supply voltage minus threshold voltage) is small and the reduced threshold voltage dominates. In the result section it is shown that below a certain voltage the performance increases with increased temperature, this voltage is called the temperature inversion point.

The power dissipation of a logic circuit can be divided into static power consumption and dynamic power consumption. Static power is consumed even if the circuit is not switching. It is caused by leakage currents of transistors that are in off-state (gate leakage, sub-threshold leakage, junction leakage) [8]. The leakage current can be measured when the demonstrator is connected to the supply voltage and the oscillation of the generic critical path is disabled. Prior to 90nm power dissipation by leakage currents could be neglected, unfortunately in current technologies this is no longer the case. The leakage currents are strongly dependent on the temperature as can be seen in the results section.

The dynamic power consumption is mainly caused by switching currents that charge the output load of a gate. First, the load capacitance is charged to V_{DD} by the pull-up network of a gate. At the next output signal transition, the charge stored at the capacitance flows through the pull-down network to ground. $P_{switching}$ is given by the following formula:

$$P_{switching} = \alpha \cdot f_{CLK} \cdot {V_{DD}}^2 \cdot C_L$$

The switching Power is dependent in the activity α, the clock frequency f_{CLK}, the supply voltage V_{DD} and the load capacitance C_L. Since the power dissipation is caused by charging and discharging a capacity, there is almost no temperature dependence visible. By measuring the current consumption of the demonstrator during oscillation, the total power consumption can be determined. To get the dynamic power consumption, the static power has to be subtracted from the total power.

The size of the demonstrator macro is 52 x 25 µm² (see Figure **20**). It has a complexity of 1k gate equivalences and is fabricated in a 28nm high-k metal gate (HKMG) foundry process. The macro consists of 4 identical instances of the generalized critical path, input buffers and output divider stages. The performance of all four instances can be measured separately to collect some information about within-die process variation. On a die 4 instances of the macro are placed with different technology options (2x standard threshold voltage, 1x low threshold voltage, and 1x high threshold voltage). The die is packaged into a ceramic dual in-line package with 24 pins with one macro bonded to the pins.

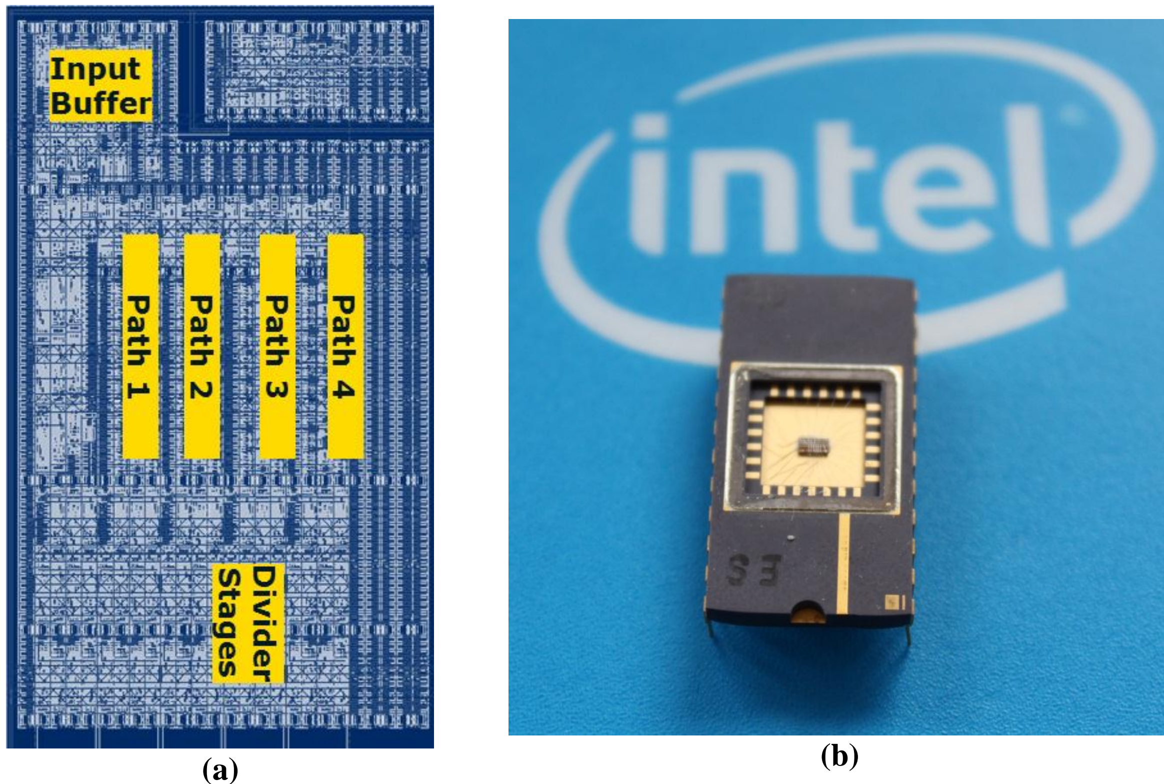


Figure 20: Layout of the demonstrator macro (a) and packaged die (b).

4.2.2 The Measurement Setup

The measurement setup is depicted in Figure **21**. To alter the temperature, the packaged demonstrator is placed into an oven. The oven can achieve temperatures of more than 150°C. The demonstrator inputs are connected to a controller board. The controller enables the oscillation of the generalized critical paths. It has a DC voltage source to set the desired supply voltage and an analog-to-digital converter to measure the current consumption. The outputs of the generalized critical paths are connected to an oscilloscope to detect the oscillation frequencies. A laptop controls the heater, the controller and the oscilloscope. It also collects the data. The complete measurement is automated to a large extend, since it takes up to one hour until a stable temperature is reached after the temperature is changed.

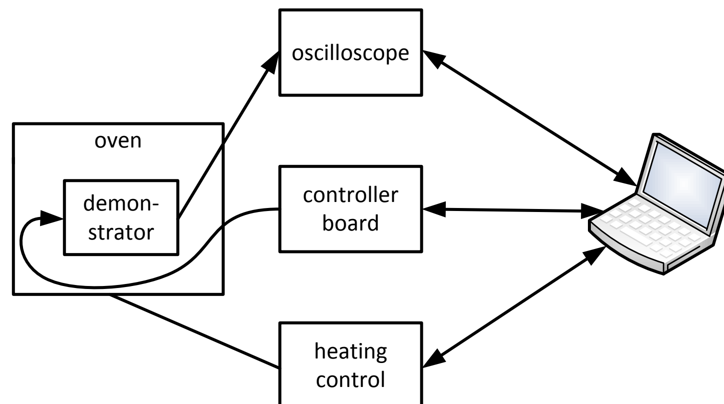


Figure 21: Measurement setup

4.2.3 Technical results

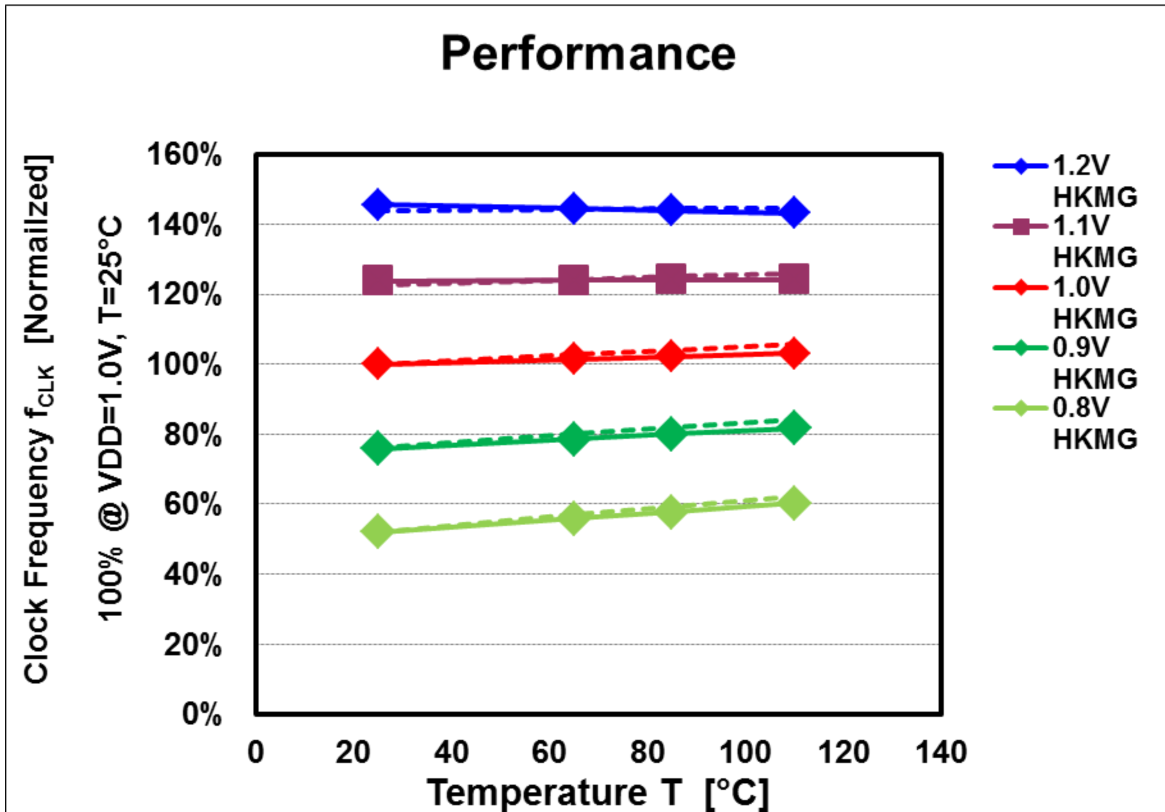


Figure 22: Critical path performance over temperature for different supply voltages

To verify the accuracy of the transistor models and the correct modeling of the temperature dependence, the measured data is compared to post-layout simulations on circuit level. The simulated netlist included extracted parasitic resistors and capacitances. The generalized critical paths were measured and simulated with standard threshold voltage transistors. All the shown data is normalized to 25°C and 1.0 V.

The first graph (see Figure **22**) shows the oscillation frequency over temperature for different supply voltages. The error between the measured data (solid lines) and the simulated results (dashed lines) is less than 3%. This indicates the excellent quality of the transistor models. One can see that for the typical active operation modes, which are between 1.0V and 1.1V, the impact of temperature on the performance is very low. The performance changes less than 6% in a temperature range between -30°C and 125°C. For low voltage operating modes (supply voltage ≤ 0.8V), the circuit performance is more sensitive to the temperature. The performance changes more than 14% between -30°C and 125°C. For those low voltage modes the performance increases with increased temperature due to the inverted temperature effect.

The inverted temperature effect can be better observed by plotting performance over supply voltage for different temperatures (see Figure **23**). The temperature inversion point is at 1.1V, since the performance is independent of the temperature (the series for different temperatures cross). For supply voltages below the temperature inversion point, the performance increases with increased temperature. For supply voltages above the temperature inversion point, the situation is vice versa. Now the performance decreases with increased temperature.

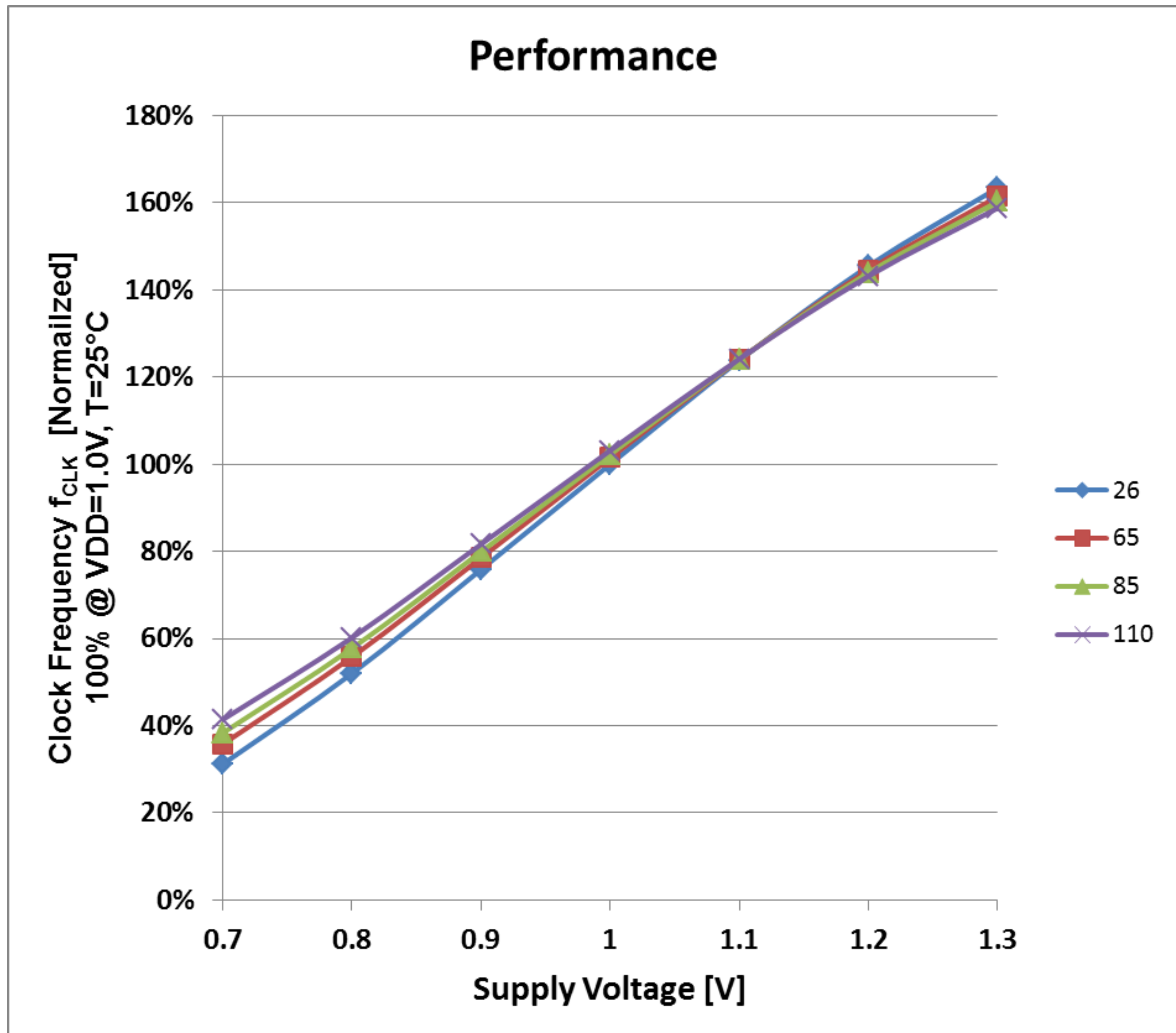


Figure 23: Critical path performance over supply voltage for different temperature

Figure **24** compares the measured performance of the demonstrator (solid lines) with the measured performance of a critical path in 40nm *SiON* technology (dashed lines). It should be noted that there is no fundamental difference between 40nm and 28nm. Both technologies show comparable temperature sensitivities for the same supply voltage. From this we can deduce that in 28nm the classical temperature corner design is still feasible for high performance mobile applications – like it has been in 40nm. Figure **24** also demonstrates our measurable objective MO7.2.5 (40nm CMOS: <5% performance impact over T=[-30, 85]°C range for typical wireless baseband operation (ref point:1.2V, T=25°C, nom =100% performance) [IMC] 28nm CMOS: <10% performance impact over T=[-30,85]°C range for typical wireless baseband operation (ref point: 1.1V T=25°C, nom =100% performance)[IMC]), since the performance impact in the given temperature range is below 5%.

Besides performance also the power consumption of the demonstrator is measured and simulated. Figure **25** shows the total power consumption when the generalized critical paths are oscillating. The agreement between model and experiment is again very good, except for high temperatures. For high temperatures the impact of the leakage current on the power dissipation becomes visible. However, the deviation is within expectations considering the

process variations of the measured die (simulations are done using nominal corner models, however the demonstrator die is a bit shifted to the fast corner).

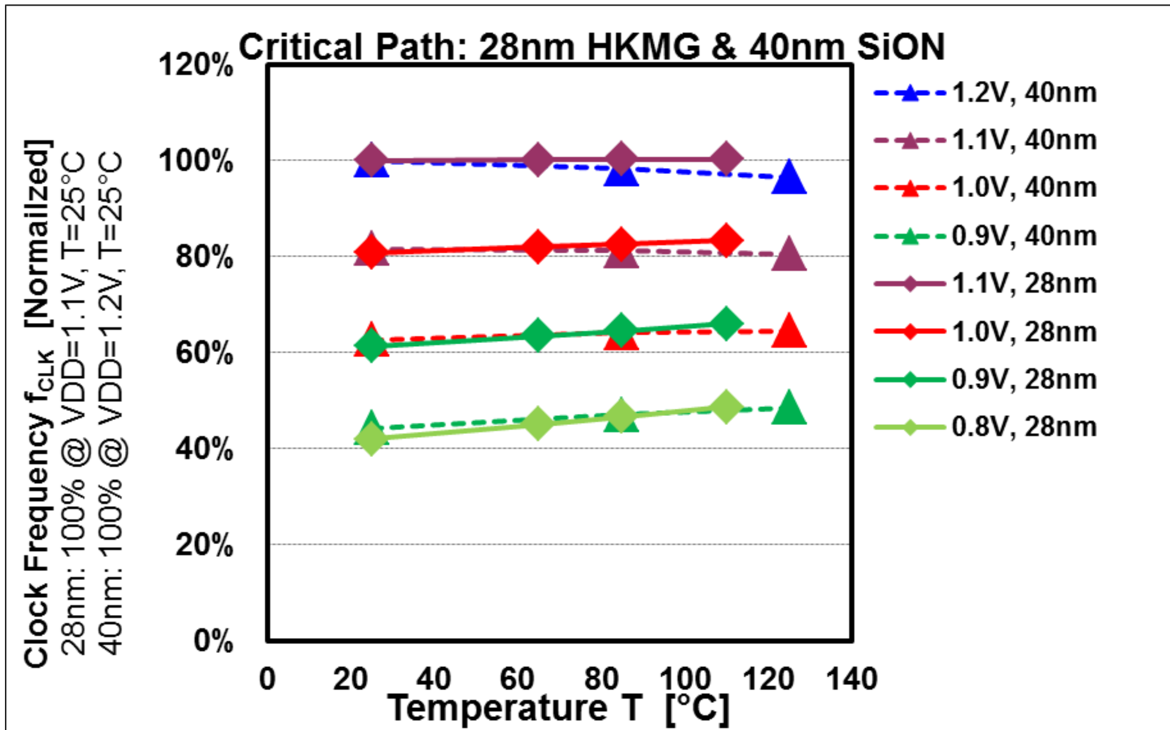


Figure 24: Comparison of 28nm HKMG and 40nm SiON critical path performance

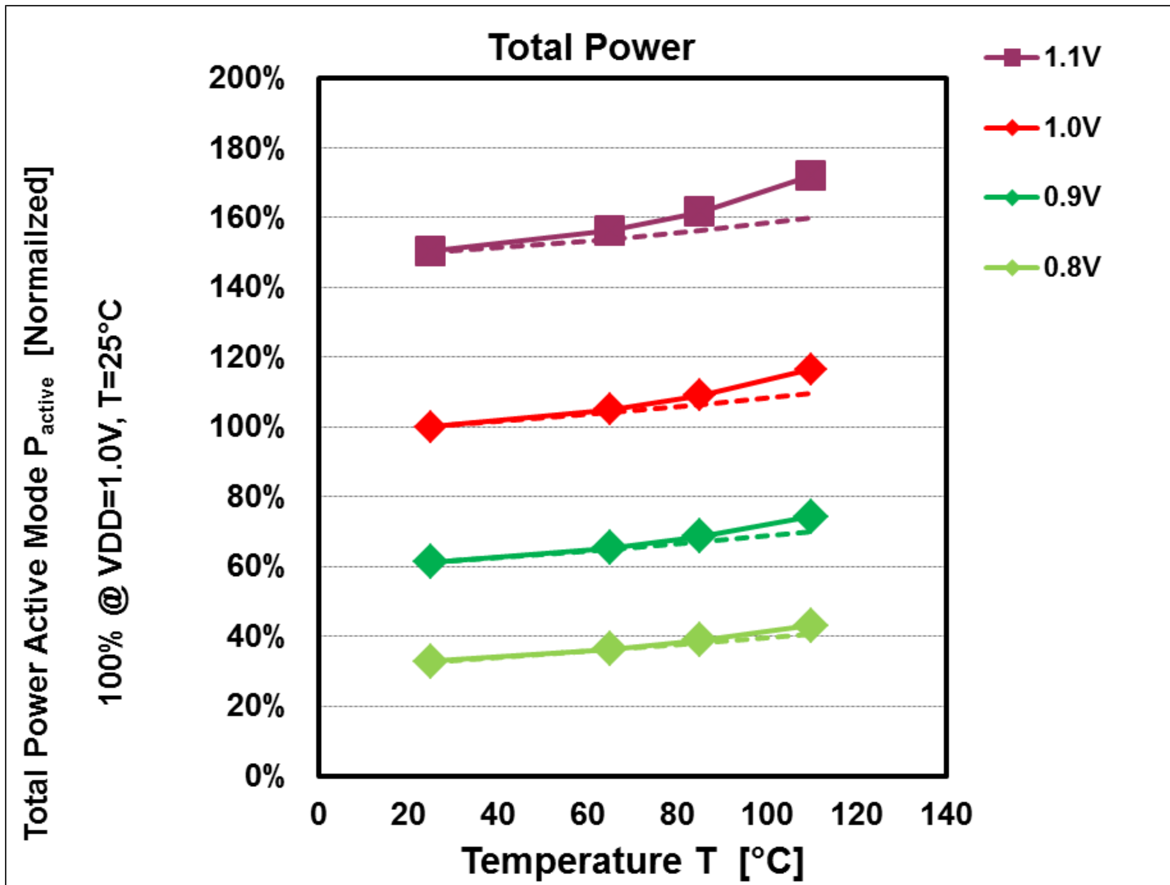


Figure 25: Total power consumption over temperature for different supply voltages

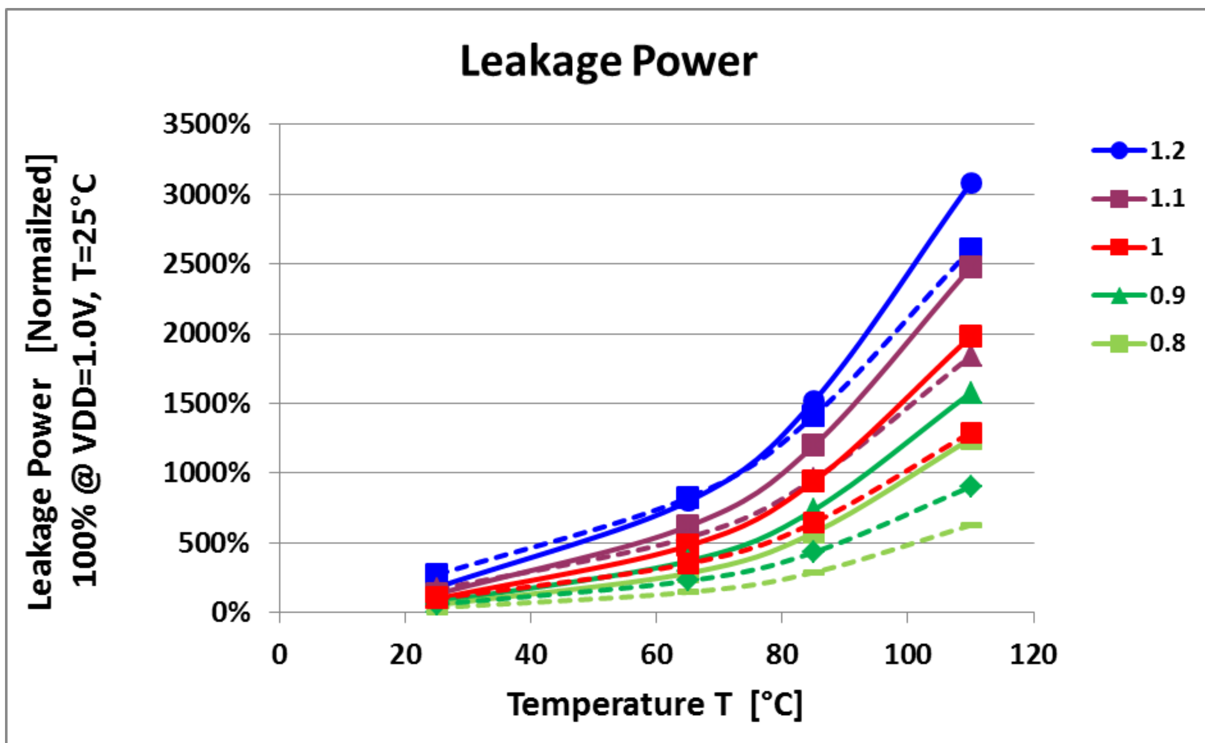


Figure 26: Leakage power over temperature for different supply voltages

Figure **26** shows the leakage power dissipation over temperature. As mentioned in the previous paragraph, the error between simulation and measurement is quite large but can be explained by the discrepancy between the nominal simulation and the demonstrator die which is shifted to the fast corner.

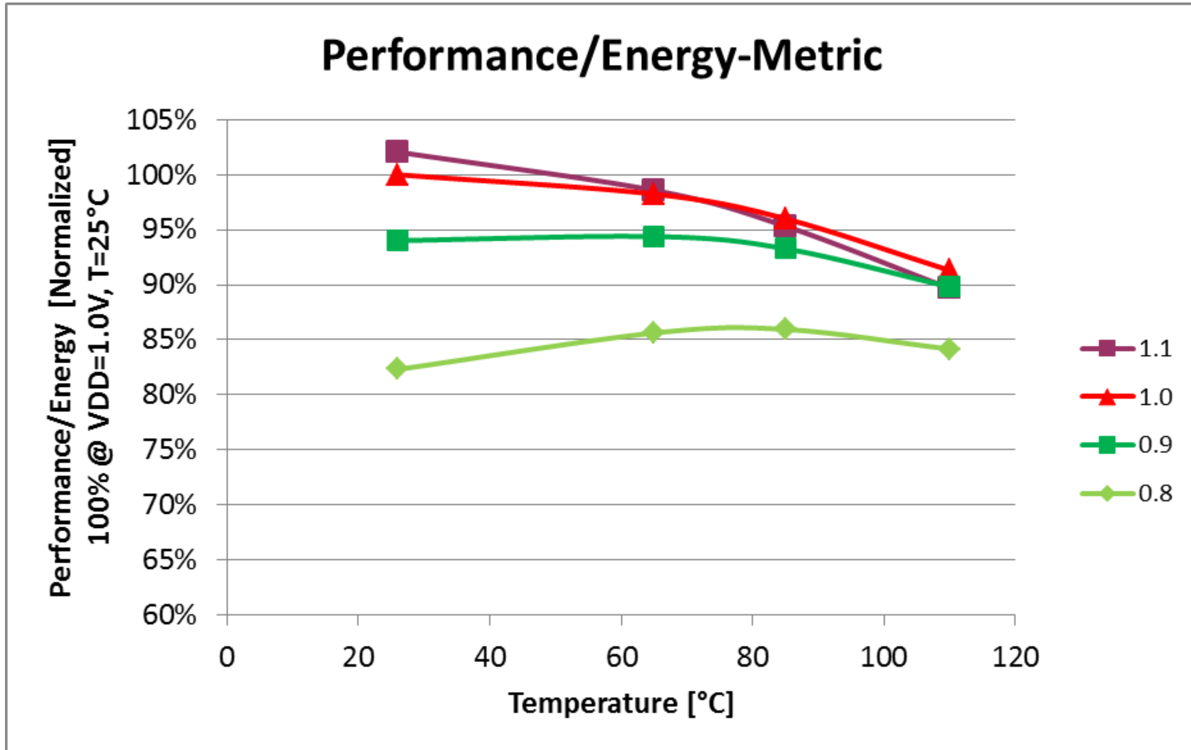


Figure 27: Performance/Energy-Metric over temperature for different supply voltages

For mobile applications with their limited battery capacity it is important to use the available energy as efficiently as possible. One metric to optimize a circuit for mobile applications is the performance-energy-quotient (PEQ):

$$\mathrm{PEQ} = \frac{\mathrm{Performance}}{Total\ Enery}$$

This metric should be maximized. Figure 27 shows the calculated PEQ from the measurement results over temperature for different voltages. For higher temperatures, the increase of the leakage power is dominant compared to the performance increase resulting in a drop of the efficiency. The quotient is maximal for supply voltages between 1.1V and 1.0V. One can see that for mobile applications one should not always aim for lower and lower supply voltages, since the available energy is used more efficiently at higher supply voltages. At very low supply voltages one saves more power, however, this is not very efficient since the operating frequency is disproportionally reduced.

4.3 Comparative analysis by POLITO

4.3.1 Benchmark Description

The test-case provided by ST is a subset skeleton of a typical MCU suitable for a wide range of applications such as motor drives, application control, medical and handheld equipment, industrial applications, inverters, printers, etc. It includes several general purpose IPs (DMA, I2C, Timers, USART, SPI, USB, I/Os, etc.) and all the interconnect infrastructure, plus memory controllers for embedded Flash, embedded RAM and external NVM (Table 1). ST

delivered a gtech representation of the overall design (generic technology, a technology independent netlist, good enough to apply any kind of synthesis flow on it) without the analog IPs, not relevant for the purpose of the test-case.

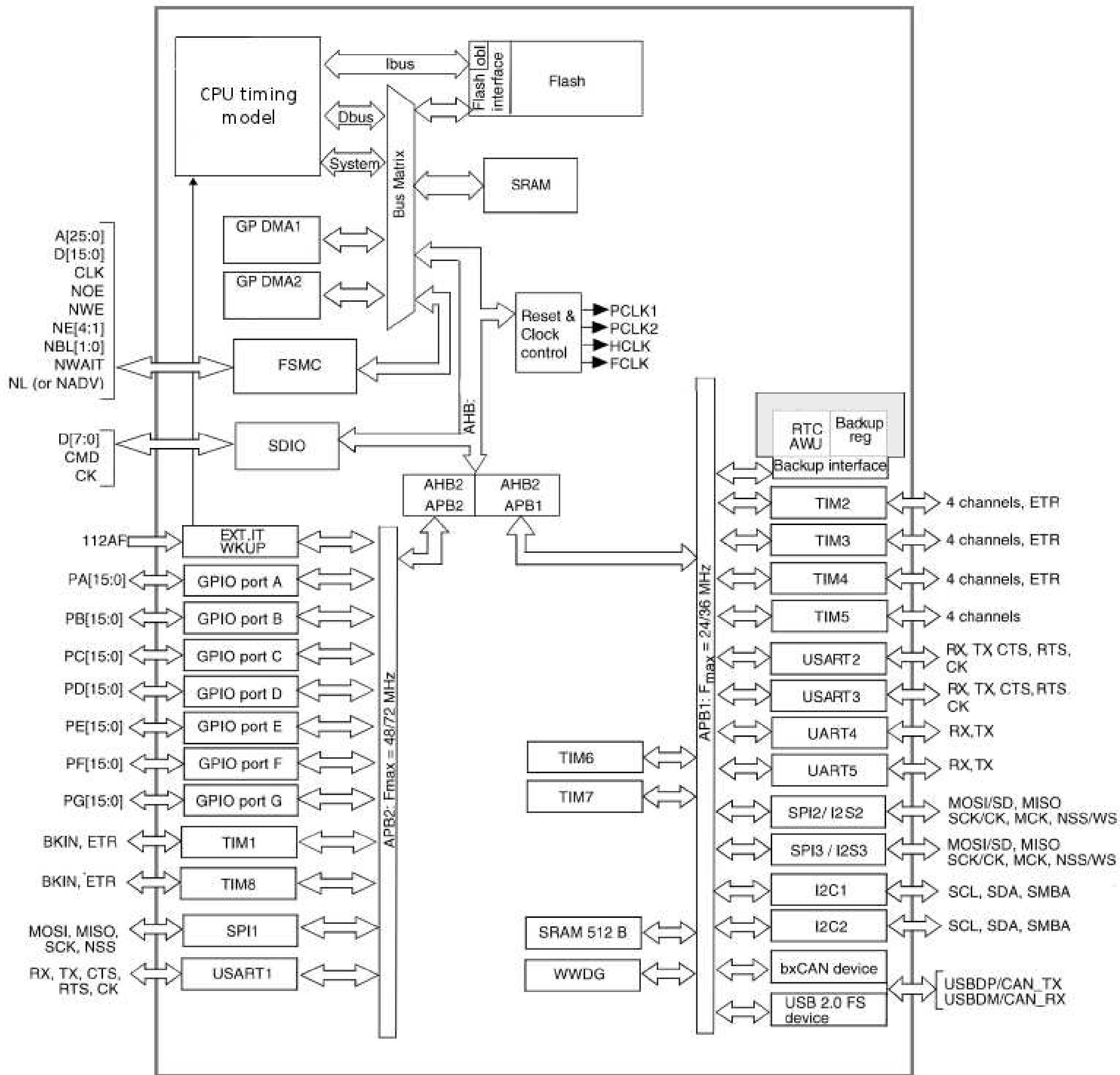


Table 1: Test Case Block Diagram

All the external and internal pins are modelled in terms of input and output delays. The test-case is made of a single voltage domain but support various power modes, depending on the external power manager, and clock domain controls; a test aiding logic is also present as part of the normal IP set of the MCU.

The main clock domain (see Table 2) is a fast AHB one, running at full speed, pacing the CPU, DMAs, embedded memories, external memory controllers and the Clock Controller itself; this is interfaced to two APB domains (by means on two bridges) running at a lower (ratio n) speed where all the remaining IP are instantiated.

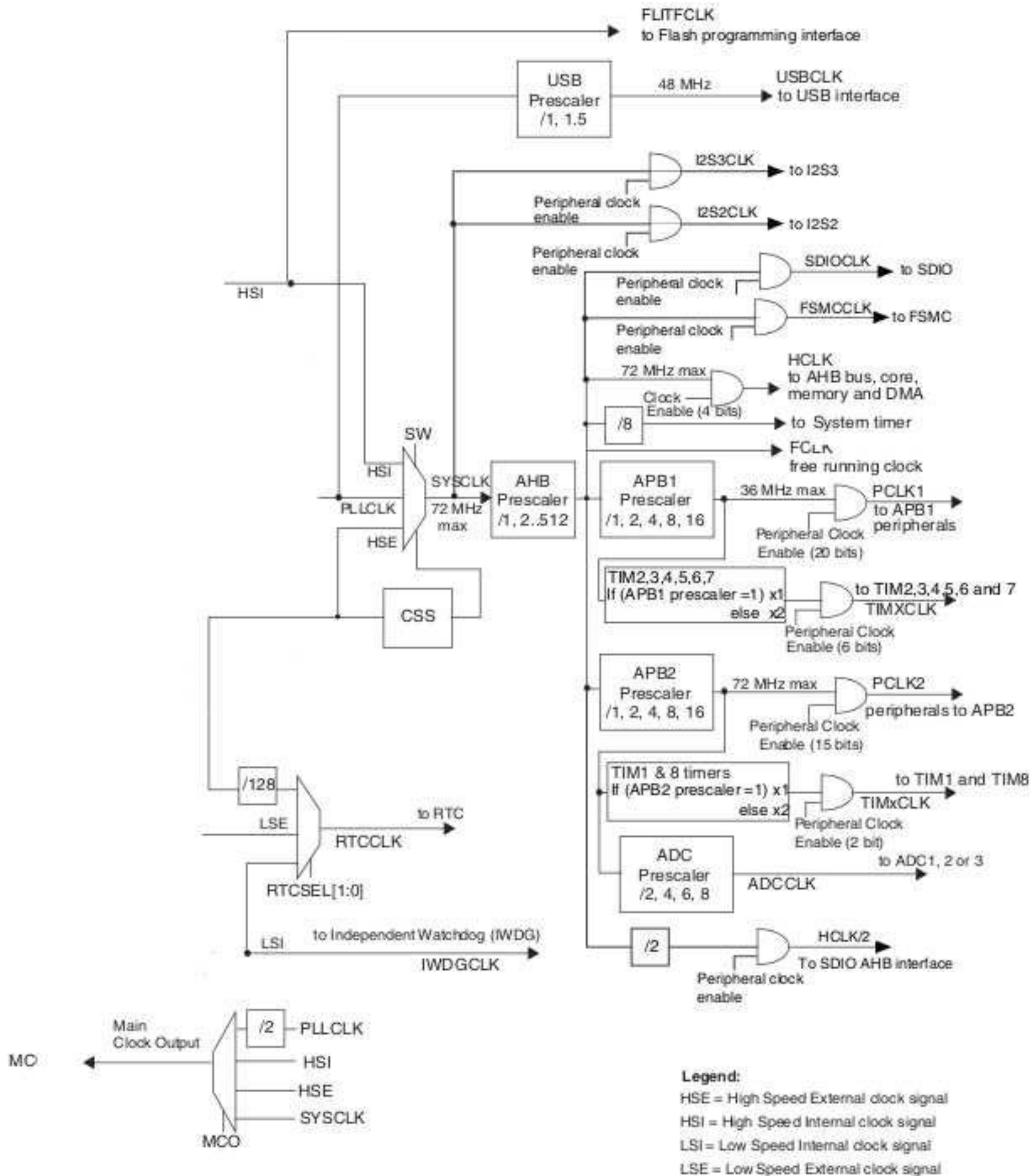


Table 2: Test Case Clock Tree

4.3.2 Comparative Analysis Framework

In this section a brief description of the Timing/Power evaluation flow provided by POLITO (reported in more details in the Deliverable D3.2.1) will be given.

The evaluation procedure for analyzing the performance of the test case circuit, in terms of power and timing, is new and different from the state-of-the-art procedures since it does not ignore the fact that on a die (in deep sub-micron technologies, multi-Gigahertz designs and ultra-high power densities) thermal gradients may occur.

Table 3 shows the methodology to evaluate the temperature effect on timing and power for a test case design. This is used for generating a comparative analysis on the test case.

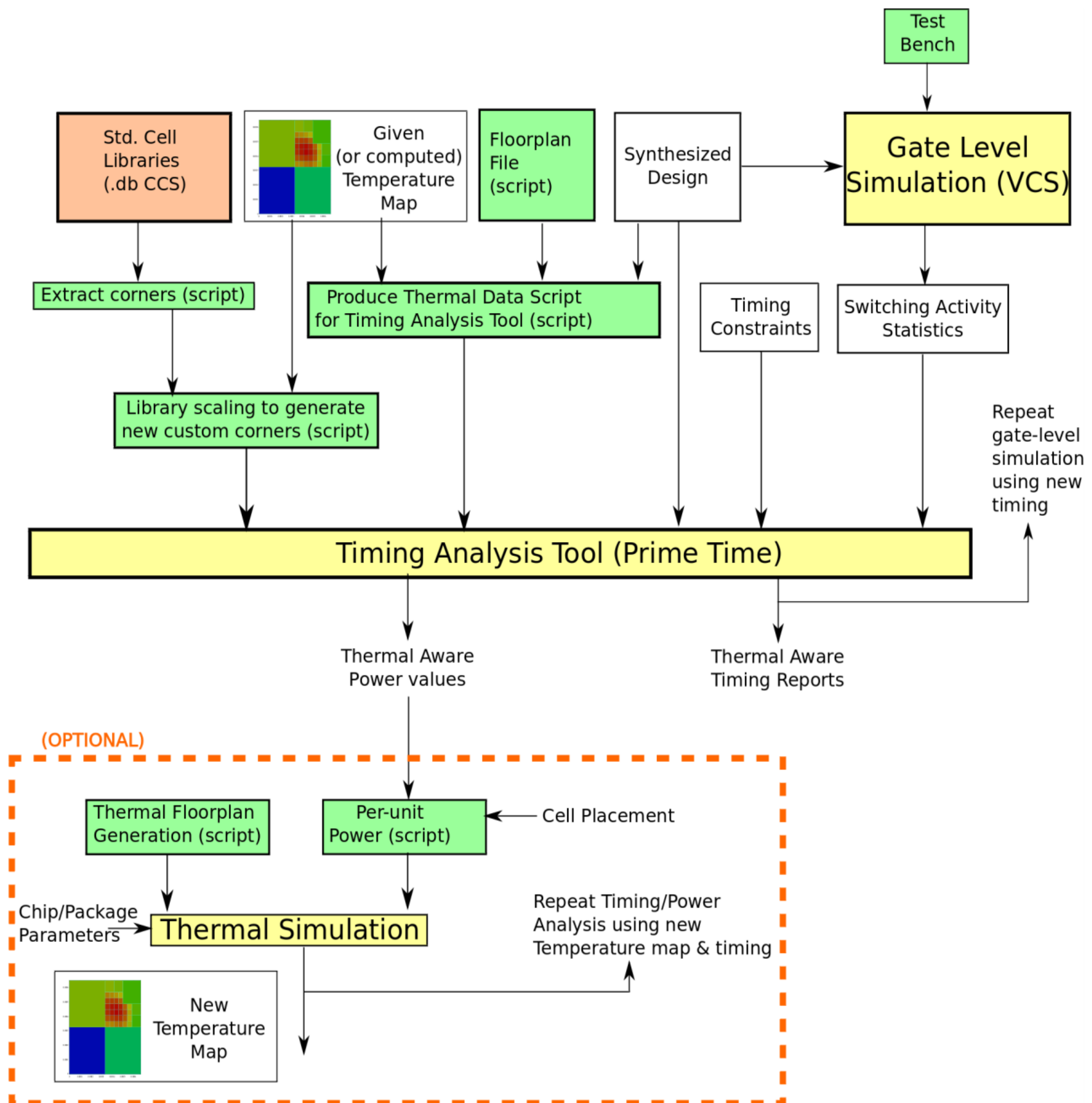


Table 3: Timing/Power evaluation flow

The depicted flow consists of a loop in which timing/power analysis on the design is repeatedly performed. The accuracy of the results is gradually improved at each iteration. This is done until an acceptable convergence level on the results is achieved.

4.3.3 Technical Results

In this section we report the effects of the temperature variations on key design metrics as power (static and dynamic) and delay related to the test-case described in Section 4.3.1.

We used two different technology nodes (40nm and 65nm) by TSMC for our experiments. We considered the temperature range [-40°C ÷ 125°C].

For each technology node we mapped the test-case considering Low Voltage Threshold (LVT), Regular Voltage Threshold (RVT) and High Voltage Threshold (HVT) cells and we made the comparison with different thermal maps and with different supply voltages.

The reported values are normalized with the case 65nm RVT. We considered four corner cases available from libraries characterization, namely LT, ML, WCL, WC (see Deliverable D3.2.2 for more details).

Table 4 shows the static and dynamic power and the clock frequencies (maximum) of the target design for the four corner cases.

Node	VTH	VDD	Corner	T=-40				Corner	T=125			
				Static	Dyn.	Freq.	Period		Static	Dyn.	Freq.	Period
40nm	LVT	0,81	WCL	2,48	1,29	1,08	0,93	WCL	871,96	1,28	1,57	0,64
	RVT			2,45	1,18	0,55	1,82		224,57	1,25	1,1	0,91
	HVT			2,77	1,22	0,14	7,15		42,58	1,32	0,44	2,28
65nm	LVT	0,9	WCL	1,17	1,02	1,66	0,6	WCL	513,09	1,06	1,77	0,56
	RVT			1	1	1	1		125,03	1,08	1,23	0,81
	HVT			1,18	1,08	0,28	3,53		22,2	1,16	0,51	1,97
40nm	LVT	1,21	LT	856,32	3,03	7,91	0,13	ML	134134	3,46	6,79	0,15
	RVT			69	2,46	6,24	0,16		17220	2,7	5,67	0,18
	HVT			54,99	2,98	4,22	0,24		2015,6	3,03	4,04	0,25
65nm	LVT	1,32	LT	1720,2	2,58	7,43	0,13	ML	149580	3	6,06	0,16
	RVT			34,83	2,42	5,81	0,17		18197	2,68	4,82	0,21
	HVT			20,07	2,74	3,73	0,27		826,77	2,79	3,39	0,3

Table 4: Static and Dynamic Power and Speed for the main characterized library corners (WCL , WC, ML, LT).

Table 4 shows the maximum running clock frequencies of final designs as well as consumed dynamic and static power for main different corner relativized to the 65nm RVT implantation. To perform timing/power analysis at each corner we use characterized cell library and extracted wire parasitic specific for that corner using the tool presented in Deliverable 3.2.2.

For each design we perform gate-level simulation and we apply a unique work-load to the test case design to obtain switching activities which will be used for power estimation. In Table 4 power values are represented as power densities. For dynamic power density, since it changes linearly with clock frequency, we show per-MHz density value.

As we can see, the fastest running frequency of the design is for 40nmLVT (7.91X compared to 65nm RVT). This happens at LT corner which has the lowest temperature and highest supply voltage. At WCL, we see slowest clock for both designs: 7.15X for 40nmHVT vs. 3.53X for 65nm HVT. Highest leakage power densities (~130kX for 40nmLVT and 140kX for 65nmLVT) can be seen at ML corner. The fact that for LVT chips the leakage current at ML and LT is slightly higher for 65nm compared to 40nm can be explained by the higher

supply voltage of 65nm design. For the rest of the cases, the leakage power density at 40nm is always approximately 2X bigger than 65nm. Considering *per-MHz* dynamic power density, for 40nm designs it is always larger than the corresponding 65nm design.

Based on the results in Table 4, the sensitivity of each design to changes in temperature and supply voltage is quantified. Table 5 shows the relative sensitivity of estimated power/delay values to changes in temperate.

Node	VTH	VDD	Sensitivity to Temperature		
			Static	Dynamic	Period
40nm	LVT	0,81	7,01	-0,05	1,55
	RVT		1,79	0,95	4,85
	HVT		0,32	1,35	25,94
65nm	LVT	0,9	4,13	0,55	0,21
	RVT		1	1	1
	HVT		0,17	1,05	8,3
40nm	LVT	1,21	1074,57	5,5	-0,11
	RVT		138,28	3,1	-0,09
	HVT		15,81	0,7	-0,06
65nm	LVT	1,32	1192,14	5,45	-0,16
	RVT		146,43	3,35	-0,19
	HVT		6,5	0,65	-0,14

Table 5: Sensitivity of static/dynamic power and delay to changes in temperature for fixed supply voltage.

Here for each of static/dynamic power density values and the clock period, we keep the supply voltage fixed, and we calculate the rate of change in the parameter by the increase in temperature (from -40°C to 125°C). In this table, the values are shown as percentages relative to each parameter's final value at 125°C. For static power, since it changes exponentially with temperature, we show the rate relativized to the 65nm RVT case.

Looking at clock period sensitivity, we see negative values for low supply voltage and positive ones for high supply voltage. Indeed due to temperature inversion effect, clock period decreases by the increase in temperature for lower supply voltages thus the design gets faster. At high supply voltage however, clock period grows by the increase in temperature. We also note that in low supply voltages, the sensitivity of clock period to temperature is orders of magnitude higher than for high supply voltage values. Among these, the HVT designs are always the most sensitive to temperature. 40nm designs are always more sensitive than 65nm at low supply voltages. At high supply voltages however, 65nm designs are more sensitive.

Considering static power, we clearly note that HVT designs have always the least sensitivity to changes in temperature. Looking at dynamic power, the relative sensitivity to temperature is less sensible to technology nodes than the other figures of merit.

Table 6 shows the impact of changes in supply voltage value. Here for each design, we keep the temperature fixed, and we calculate the rate of change in design parameter with the change of supply voltage value from its minimum to maximum. This table show the change for each parameter when supply voltage changes by 1.0 volt however the standard cell library has a practical voltage range of 0.4 volts.

Node	VTH	T	Sensitivity to Voltage		
			Static	Dynamic	Period
40nm	LVT	-40	26,5	1,29	1,02
	RVT		2,07	0,94	2,11
	HVT		1,62	1,3	8,76
65nm	LVT	-40	50,82	1,1	0,57
	RVT		1	1	1
	HVT		0,56	1,17	3,94
40nm	LVT	125	4136,61	1,61	0,62
	RVT		527,56	1,07	0,93
	HVT		61,25	1,26	2,57
65nm	LVT	125	4406,85	1,36	0,48
	RVT		534,26	1,13	0,73
	HVT		23,79	1,15	2,02

Table 6: Sensitivity of power and delay to changes in supply voltage for fixed temperature.

Considering clock period, as we see, sensitivity to voltage is always negative, meaning that by going to higher supply voltages the clock frequency will always increase. In contrast to static power, the sensitivity of clock to voltage is always higher in lower temperature for both of 40nm and 65nm.
Moreover, the sensitivity of 40nm designs is always higher than 65nm. Considering static power, we see at lower temperature the sensitivity to voltage is small however, it grows by orders of magnitudes at high temperature value. Looking at dynamic power, for RVT designs, this sensitivity is usually smaller than LVT and HVT.

Effect of Temperature Variation:
The effect of temperature variation on power and speed is quantified by creating the following three different test scenarios:

1- Temperature is uniform across entire die and sweeps from 0 to 100 degrees. Here we recognize how the design behaves across different temperature values.
2- Non-uniform on-die temperature distribution is a specific user-defined pattern. Here the error caused by not considering on-die temperature variation on estimated delay/power values is quantified.
3- Non-uniform temperature map is the output of thermal simulation done on the design. This represents a real testcase in which self-heating effect of the chip is taken into account while doing power/delay analysis and temperature estimation.

We evaluate the behaviour of each design configuration with the sweep in uniform on-die temperature value. Table 7 shows the results. Dynamic power, changes approximately linearly with temperature. Changes of static power density is exponential thus, we represent its absolute value at 0°C and 100°C.

Node	VDD	VTH	Dynamic			Static		T(S.=D.)	Frequency			Disp.
			0	100	rate	0	100		0	100	rate	%
40nm	0,81	LVT	1,3	1,3	0,0	4	95	0,5	1,1	1,4	2,2	0,0

		RVT	1,2	1,2	0,8	2	28	1,0	0,6	0,9	2,3	0,0
		HVT	1,2	1,3	1,3	1	6	4,7	0,2	0,4	1,2	0,2
	1,21	LVT	3,0	3,3	6,8	974	15984	0,1	7,2	6,9	-2,5	1,5
		RVT	2,5	2,6	3,0	81	1938	0,2	5,8	5,6	-1,6	0,0
		HVT	2,9	3,0	0,6	34	269	0,7	3,9	3,9	-0,3	0,0
65nm	0,9	LVT	1,0	1,0	0,7	2	55	0,6	1,6	1,7	0,5	0,1
		RVT	1,0	1,1	1,0	1	15	1,0	1,0	1,1	1,0	1,0
		HVT	1,1	1,1	0,9	1	3	4,4	0,3	0,4	1,0	0,0
	1,32	LVT	2,6	2,9	5,4	1594	18861	0,1	6,8	6,0	-6,0	0,1
		RVT	2,4	2,6	3,3	56	1906	0,2	5,3	4,7	-4,2	0,0
		HVT	2,7	2,7	0,6	13	108	0,9	3,5	3,3	-1,5	0,0

Table 7: Changes in power and delay with temperature sweep from 0c to 100c. A unique temperature is assumed for entire die.

To propose a measure on importance of static power in comparison with dynamic power in each configuration of the design, we calculate the temperature value in which these two power entities get equal. This is done by doing required interpolation (and if needed, extrapolation) in power density values and finding the collision point. Obtained numbers are just aimed to show the importance of static power in each design and usually fall outside the practical boundaries of chip operation temperature. All the results are normalized to the 65nm RVT design implementation.

As can be seen, the collision point always increases by going from LVT to HVT. The amount of increase for low supply voltages is higher compared to high supply voltage. For 40nm LVT (VDD=1.21) and 65nm (VDD=1.32) designs the collision happens at much lower temperature (1/10th of the reference one).

For each (T, VDD) test point, we obtain the physical on-die location of first 10 timing critical paths. In Table 7, we represent the amount of displacement for the first timing critical path. We must notice that the change in physical location highly depends on how the tool has performed the placement stage however, these changes emphasize on the fact that the leaf-cells of the design do not show similar behaviour to changes in temperature.

Figure 28 shows how the physical locations of first 10 timing critical paths change by temperature. For 40nm LVT a significant jump happens between 75°C and 100°C. For 65nm LVT the location of critical paths is different in 75°C compared to rest of the cases.

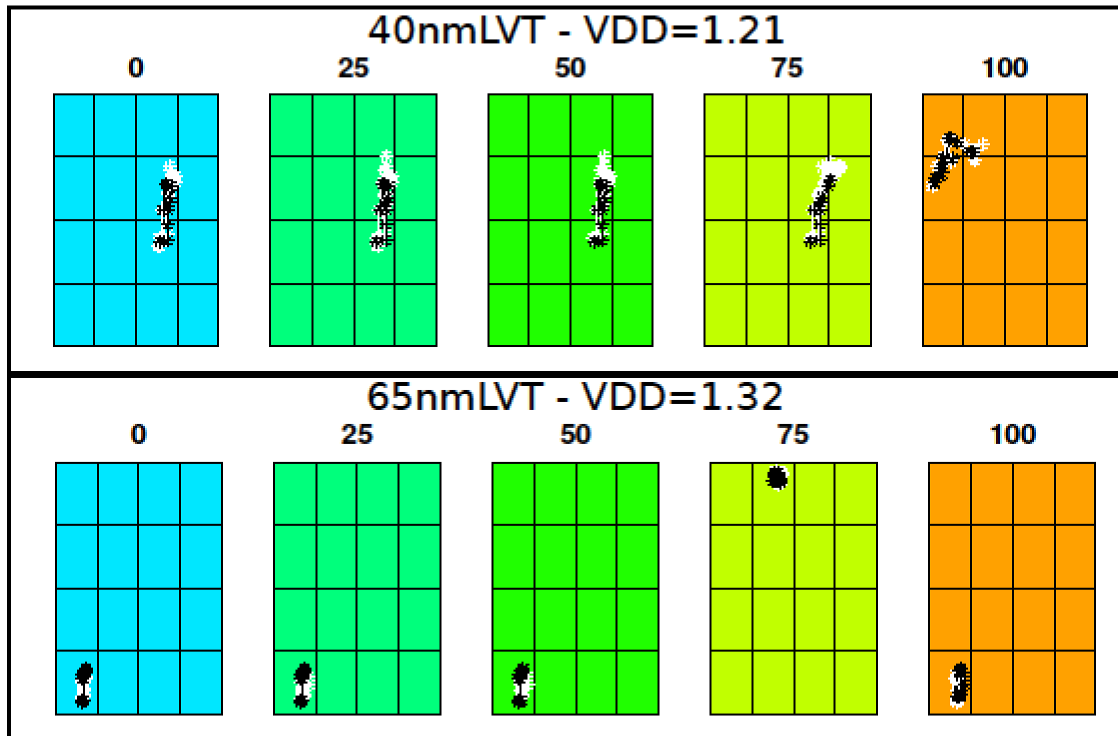


Figure 28: Physical location of first 10 timing critical paths of the design for different uniform temperature values. First path is shown in black.

We apply a collection of non-uniform temperature maps to our design configurations. The following temperature patterns are utilized:

1. Chess-board: we divide die area into equally sized rectangles. Two temperature values are selected and applied to them. This case is a worst case scenario
2. Gradient: we create horizontal gradients of chip temperature across the die. This case is representative of many real-world situations such as PCBs in which components with high power densities (such as DRAM modules) are located just at one side of the CPU or cooling units are placed asymmetrically.

For each of the mentioned patterns we represent the results for two test cases: G1 and G2.

1. Chess-board: G1 pattern uses 40°C and 60°C for adjacent blocks. G2 uses 20°C and 80°C. G10 and G20 patterns use the same temperature values as G1 and G2 but the location of blocks is reversed.
2. Gradient: G1 is a horizontal gradient from 35°C to 65°C and G2 is from 20°C to 80°C. For G10 and G20 the direction is reversed.

For all of the produced temperature maps, the average on-die temperature is always 50°C. It should be noted that, G2 and G20 patterns are not realistic, however we use them as an extreme case of non-uniform on-die temperature distribution to illustrate the operation of our solution and to show the effect of temperature variation on key design parameters. Table 8 shows the results of this experiment. In this table, the amount of error caused by assuming a uniform averaged temperature for entire die is quantified. The difference (relative error) between each estimated design parameter (considering temperature non-uniformity) and the same parameter with the assumption of uniform temperature of 50°C is calculated. For dynamic power densities, since it changes linearly with temperature the error values are very small (highest error: 0.7%). For static power, significant error values (> 20%) can be seen for

G2 and G20. This is due to exponential relationship of leakage and temperature. The average relative error value of G2 pattern for all 40nm designs is 15:89% and for 65nm is 13:03%. The higher error value for 40nm design emphasizes on the fact that by going into smaller scales of fabrication, the impact of temperature variation on design parameters gets more severe.

Pattern	Node	VDD	VTH	Static Power Error %				Clock Period Error %				First Timing Critical Path Disp. %			
				G1	G2	G1'	G2'	G1	G2	G1'	G2'	G1	G2	G1'	G2'
Chess	40nm	0,81	LVT	4	27	4	29	2	6	2	6	77	77	49	48
			RVT	3	19	1	16	4	11	4	11	65	25	64	62
			HVT	1	6	1	6	6	16	8	16	9	9	14	16
		1,21	LVT	2	20	3	21	0	0	0	1	0	1	0	55
			RVT	4	26	3	24	0	0	0	0	0	0	0	0
			HVT	1	12	1	12	0	0	0	0	0	0	0	0
	65nm	0,9	LVT	0	26	0	27	0	1	0	0	0	0	73	72
			RVT	0	17	2	20	1	4	0	3	45	37	41	41
			HVT	0	6	0	7	3	9	4	10	78	78	0	0
		1,32	LVT	3	12	3	12	1	3	1	3	67	67	0	2
			RVT	2	29	3	31	0	3	1	3	14	14	0	0
			HVT	0	12	1	13	1	0	1	2	0	23	0	0
Grad	40nm	0,81	LVT	7	20	2	12	1	2	3	6	1	0	47	47
			RVT	4	13	0	6	1	3	6	11	26	25	21	21
			HVT	2	6	1	0	5	0	10	17	14	70	14	0
		1,21	LVT	5	15	1	9	0	1	0	1	0	6	53	54
			RVT	5	17	2	12	0	0	0	0	0	1	0	0
			HVT	3	9	0	4	0	0	0	0	0	0	0	0
	65nm	0,9	LVT	0	14	1	15	1	1	0	0	0	3	73	72
			RVT	2	11	1	9	2	4	1	3	37	37	86	86
			HVT	0	2	0	3	5	10	4	8	0	0	73	73
		1,32	LVT	3	4	2	5	2	3	2	3	67	67	0	1
			RVT	4	19	3	17	1	2	1	3	84	84	37	37
			HVT	0	5	1	7	0	0	0	1	0	61	0	0

Table 8: Relative error values for estimated power, minimum clock period and physical location of first timing critical path, when considering non-uniform on-die temperature compared to assumption of uniform averaged temperature for entire die.

Considering clock period, error values of up to 16% can be seen for HVT designs at low supply voltages. For high supply voltage values however, LVT designs show higher percentage of error than the rest. This is in line with our previous observations (Table 4).

Considering the displacement of first timing critical path, it happens almost always for 40nm design at low supply voltage. For high supply voltage however, it happens mostly for LVT designs. Figure 29 shows how the on-die physical location of timing critical path changes due to different temperature patterns and different supply voltages. The figure is showing two designs: 40nmLVT and 65nmLVT. As we see, for low supply voltage, the critical path usually happens in colder areas of the chip however, with the same temperature pattern, at high supply voltage, it happens in hotter areas.

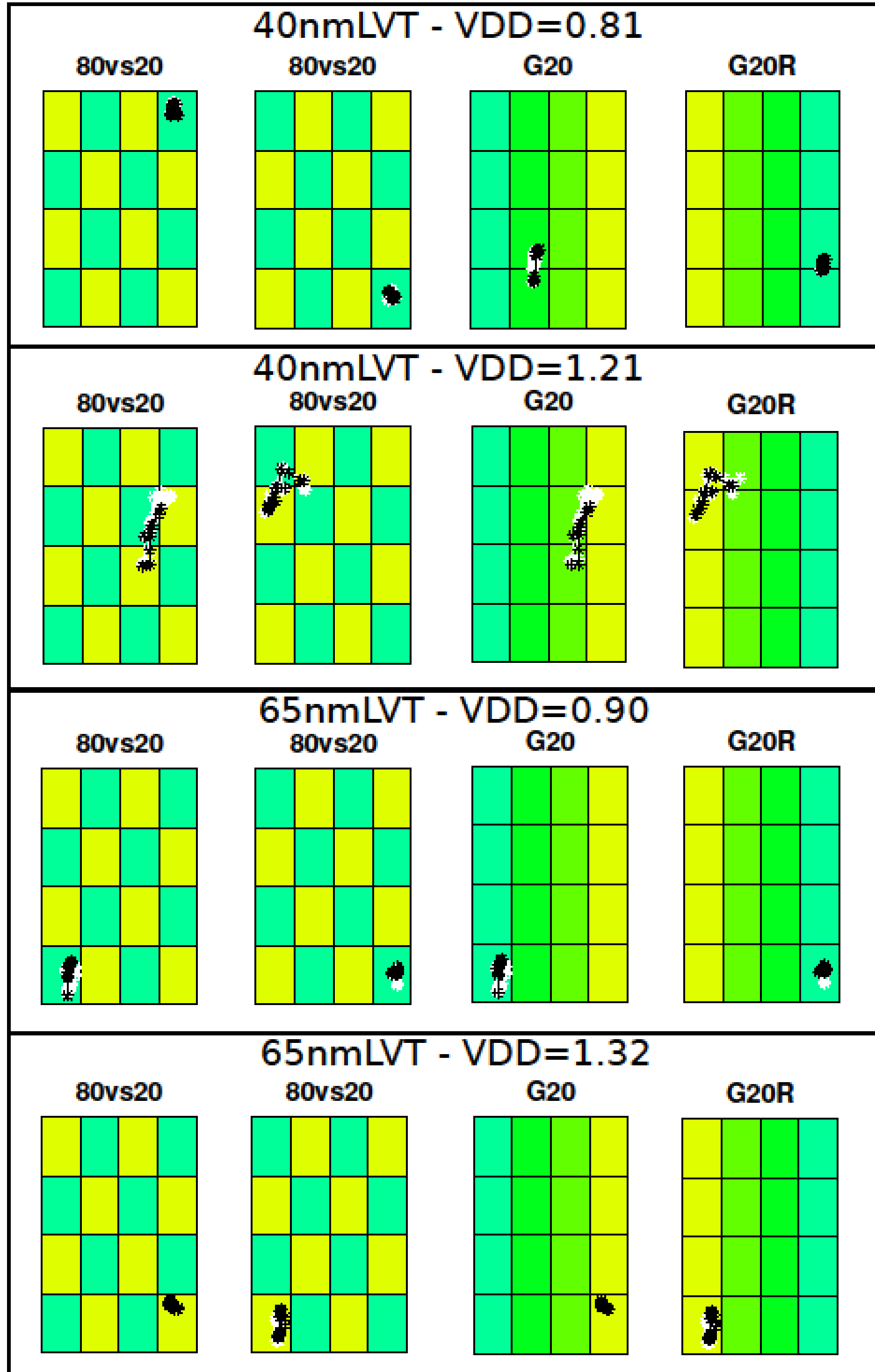


Figure 29: Location of first 10 timing critical paths for different temperature maps and supply voltages. First critical path is in black.

4.4 Conclusions

The simulations and measurements conducted in this work package task show a very good agreement between model and experiment. For the performance the error is less than 3% and for total power consumption the error is less than 4% expect for temperatures larger than 85°C. Since Intel Mobile Communications focuses on high performance products for the mobile market, temperatures above 85°C are not relevant for our applications.

In Section 4.2.3, we demonstrated our measurable objective[2] MO7.2.5 (see Figure **24**).The overall picture for CMOS logic circuits from 40nm to 28nm did not change significantly. Despite the fact that the gate stack changed from SiON to HKMG. Thermal effects for IMC's mainstream LTE/3G digital baseband ICs can be handled via high-quality circuit models and the existing semi-custom flow based on temperature corner models.

For low voltage modes (i.e. supply voltage below 0.8V) the impact of temperature on performance becomes a challenge. However, this topic is relevant for other, non-IMC applications (e.g. nodes of wireless sensor networks), which have very low performance requirements (e.g. processor speeds below 50MHz).

[2] 40nm CMOS: <5% performance impact over T=[-30, 85]°C range for typical wireless baseband operation (ref point:1.2V, T=25°C, nom =100% performance) [IMC] 28nm CMOS: <10% performance impact over T=[-30,85]°C range for typical wireless baseband operation (ref point: 1.1V T=25°C, nom =100% performance)[IMC]

5 Evaluation of a Thermal-aware 3D Physical Design Flow (ST, together with UNIBO)

General remarks

The presented work was executed in the context of Task 7.2, demonstration of design techniques on test chips, which started at month M19 and will end in month M36. This chapter covers the activities carried out by ST and UNIBO In the context of package and system level design, validation of the thermal design techniques and thermal model for interconnections of the 3D stacked test cases will be carried out as a result of Task T6.3. This will involve prototypes including 3D stacks of macro-blocks with different functions (e.g. digital devices, analog, devices, memories and high speed RF devices) which will be assembled in a 3D-SiP configuration with two silicon dies. In order to monitor and demonstrate optimization of energy consumption the prototypes ST will placed on a dedicated printed circuit board adapted for the thermal measurement. During the measurement, both thermal properties of the interposer layers and thermal model of package used in 3D stack approach will be taken into account. Validation will be carried out using comparisons to system performance.

5.1 Introduction

In Task 6.2 and 6.3 ST and UNIBO developed a thermal aware 3D physical design flow and applied it to an automatic 3D architecture design flow (details of the design flow are reported in deliverables D6.3.2 and D6.3.3). Figure 30 illustrates the proposed thermal aware 3D physical design system. The design flow is composed of three major steps: floor planning, placement and routing.

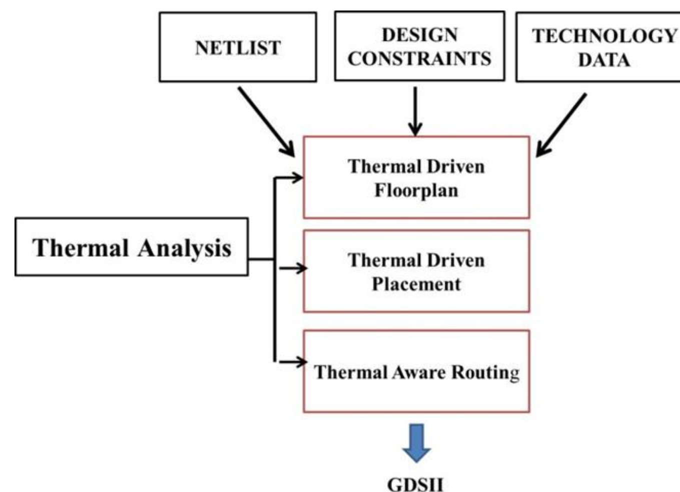


Figure 30: Thermal aware design flow

The thermal aware floor planning is accomplished with an accurate power model of each basic IP or macro blocks that composed each individual die of the 3D IC. In Figure **31** is proposed a real test case as driver to evaluate the effects of Thermal design flow on 3D Integration

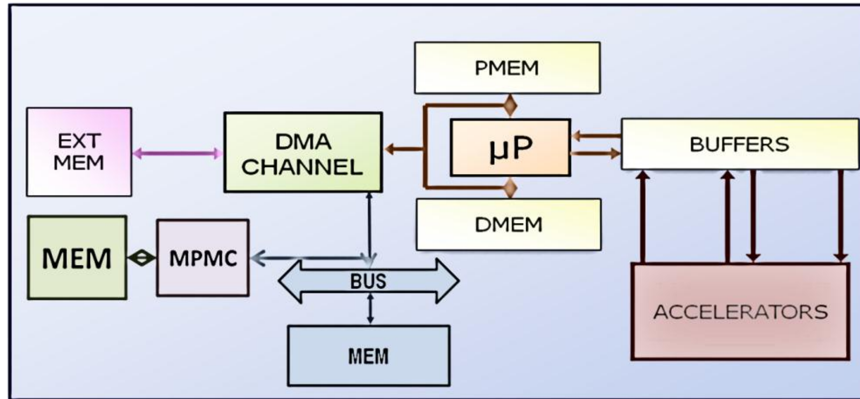


Figure 31: Block scheme of the Motion detection system test-case

The system includes a processor (µP), a direct memory access controller (DMAC), an external memory controller (MPMC), and a hardware accelerators subsystem and some logic macro cells used for testing purpose and additional buffers to allow external communication. The optimization of the macro-blocks floor plan in the system has been performed within two steps. On a first step the 3D partitioning of the blocks composing the system have been performed utilizing uniform power densities of the various blocks. The partitioning has taken advantage from the hierarchical implementation of the system. Thus, the estimation utilized for the macro-blocks composing the system are post-layout power estimations. On a second phase, the right placement of the system macro-blocks within each die of the system has been refined, utilizing more accurate power maps. Two different floor plan configurations of system have been considered to evaluate the design flow proposed. In the first design, the DIE 1 contains the DMAC, the MPMC, the hardware accelerator and memories cuts. On the other hand, the DIE 2 contains other SRAM memories cuts and the microcontroller unit. In the second design, the DIE 2 includes only memories, while the DIE 1 contains all the logic blocks composing the system and other logic. Thermal maps and thermal simulation have been evaluated and results are reported in Figure 32 and Table 9. It is interesting to notice that the placement of most power consuming blocks on the same die (i.e. the die attached to the package) is beneficial even in the “worst case” configuration, where the most power consuming blocks are closer to each other, with respect to the placement of these blocks on the two dies

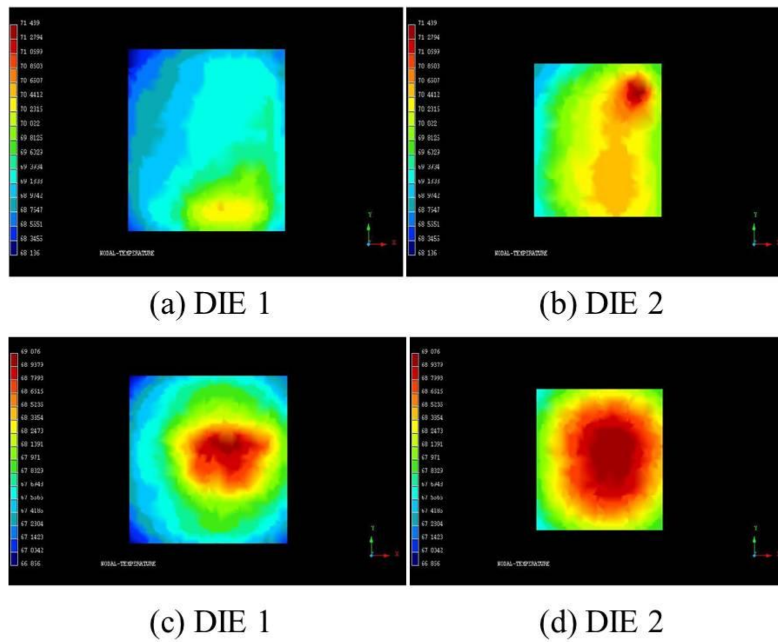


Figure 32: Thermal maps for 3D floor plan configurations in the Motion Detection system. (a) and (b) partitioning of most power consuming blocks among the two DIES (c) and (d) partitioning of most power consuming blocks within the same DIE.

FLOORPLAN	DIE1 peak temperature (°C)	DIE1 temperature gradient (°C)	DIE2 peak temperature (°C)	DIE2 temperature gradient (°C)
FLOORPLAN A	70,4	2,9	71,4	2,3
FLOORPLAN B	69,0	2,2	69,0	1,5

Table 9: Results of the thermal simulation with block-based power distribution

5.2 Design Flow Case Study

5.2.1 Memory and Logic Test-chip Fabricated in CMOS 65nm ST Process

The proposed design flow including thermal analysis has been verified, on real test case. Two test-chips in CMOS 65nm STMicroelectronics technology have been designed and fabricated. They enable to evaluate the functionality and the performance with respect thermal constrains, and in addition they are used to validate the design flow proposed in WP6. The test chip 1, called as DIE 1, is composed by the following macro cells:

- a microprocessor based on a standard RISC core architecture;
- two banks of 16Kbyte data and instruction SRAM memories;
- a DMA controller;
- a JTAG port for testing interface;
- a 32bit AMBA bus for data communication;
- an hardware accelerator designed using an ST proprietary metal programmable logic sea of gate library .
- Buffers and on chip clock generator for testing and to allow external communication
- IO pads from ST libraries

The test chip 2, called as DIE 2, is composed of:

- Cuts of SRAM memories single and double ports
- A dedicated testing interface.
- Buffers and logic macro blocks for testing purpose and to enable external communication
- IO pads from ST libraries

Figure 33 shows the die photograph of the test chips realized in CMOS 65nm process, 1V power supply, double threshold transistors.

Figure 33: die photo technology CMOS 65nm

5.3 Design Flow Validation Case Study

5.3.1 3D Stacking Configuration of DIE 1 and DIE 2

To validate the thermal analysis and the design flow developed during activities carried out in WP6 and reported in deliverables D6.3.2 and D6.3.3, 3D systems have been realized assembling two dies in different stacking configuration

5.3.1.1 Stacking of DIE on Silicon Interposer

The first assembly approach tested consists of two die that are assembled on a silicon interposer with face down as shown in Figure 34

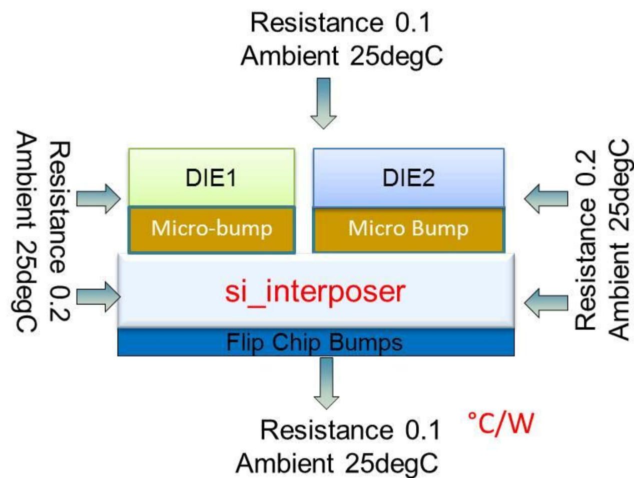


Figure 34: 3D IC stacking of two dies on silicon interposer

The dies are connected in a 3D configuration to the interposer through micro bumps and silicon interposer is packaged to a printed circuit board using flip chip approach and bumps. Thermal analysis, simultaneously solve thermal equations for all dies at the same time. The two dies generate heat due to power consumption. Heat generation from the top two dies exposes thermal profile on the silicon interposer. Silicon interposer has no active devices and no heat generation is produced on the die. Thermal maps of individual die and silicon interposer are shown in Figure 35 and Figure 36 and have been calculated using methodology developed in WP6 using thermal tools form EDA Company Apache.

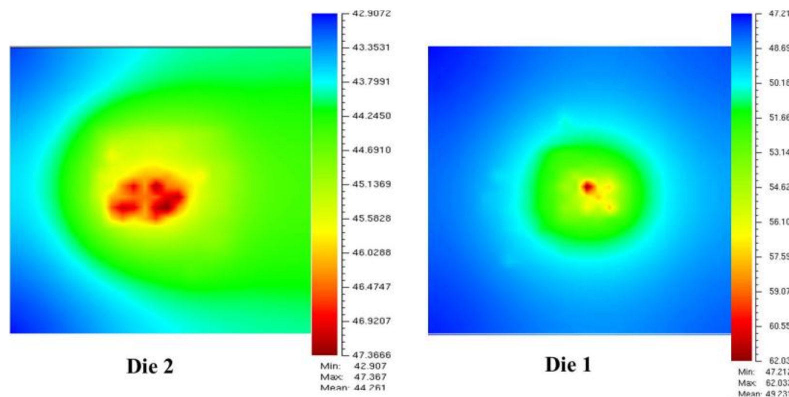


Figure 35 Thermal maps DIE2 and DIE1

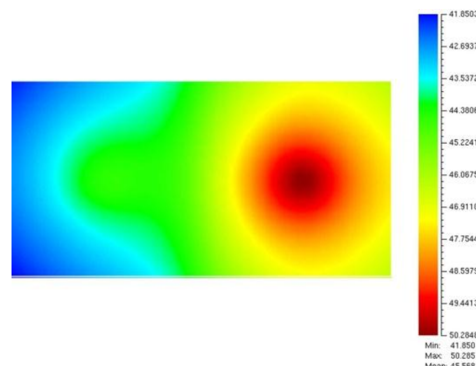


Figure 36: Thermal Profile on Si Interposer

Thermal performance of the micro-pad bonding layer between the two dies depends on geometry and thermal conductivity of the filler material and micro-pad metal. Functionality of the 3D system has been tested evaluating the propagation delay in the transmission of digital data from one device to the second one of the 3D Stack. The table 10 reports the total propagation delay of data from µP block on DIE 1 to the a memory cuts on DIE 2 through the micro bumps interconnections of 50µm x50µm, connected to silicon interposer by means of top metal layer in aluminum:

Experimental results	**Delay @ T = -40 °C (ns)**	**Delay @ T = 68.9 °C (ns)**	**Delay @ T = 125 °C (ns)**
DIE1 to DIE2	2.3132	2.6553	2.8399

Table 10: Total propagation delay from DIE 1 to DIE 2 at different temperatures.

The propagation delay has been measured at different corners operating condition of DIE at nominal voltage 1V. An increase of temperature @ T = 68.9 °C is measured during the transmission of digital signal, due to low thermal conductivity of micro-bumps layer and propagation delay is larger than delay at ambient temperature. The printed circuit board that includes the 3D stack was fabricated, for a mistake, with all the power supply pads connected together so it was not feasible makes individual measurement on power consumption.

5.3.1.2 Stacking of DIE in a Face to Face (F2F) Configuration

A prototype of a 3D system has been realized assembling two DIE in a front to front stacking configuration. This assembly concept is depicted in Figure 37, two dies are stacked in a face to face configuration with high precision alignment accuracy. Mechanical stability and inter-chip isolation are provided by a thin layer of acrylate adhesive layer sandwiched between the chips. I/Os are not covered in 3D stacking configuration in order to provide through wire bonding DC power supply and ground. Stacked chips are mounted on a printed circuit board (pcb) by using a standard chip on board approach: face up chip is wire bonded to the top side of the board while face down chip is wire bonded to the opposite side of board through a pcb slot.

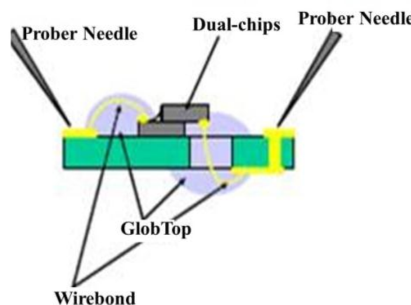


Figure 37: Concept of wire bonded assembly of 3D Die-to-Die stack indicating prober needles for testing purpose

Figure 38 shows the photograph of F2F 3D stack. The chip-stack is mounted with innovative approach by means of Chip-on-board technology. The bottom chip is wire bonded with conventional technique while the upper-one is bonded through a cavity to the opposite site of the underlying structure. In Figure 39 shows the printed circuit board fabricated with 4 layers and 3D stack is assembled in ceramic package and wire-wrapped to the board to test the prototype. This mentioned board realizes a flexible interface between the chip-stack and the measurement setup (pattern generator, oscilloscope, power supply, acquisition board). To test the system, patterns are generated by pattern generator equipment and using an acquisition board with a commercial FPGA devices.

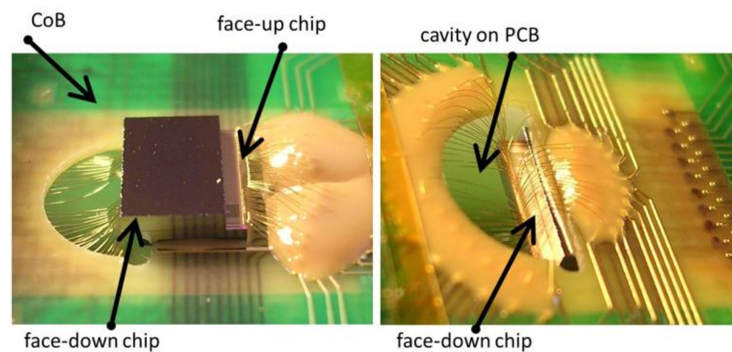


Figure 38: 3D stacking assembly face to face

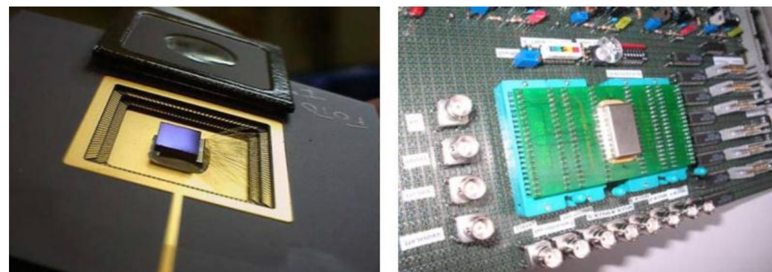


Figure 39: Package and PCB board for test of 3D system

The communication between the dies occurs through the face-to face (F2F) via. These TSV interconnects have been characterized and model following the approach developed in WP6 and described in deliverable D6.2.2. Thermal analysis has been performed following the flow proposed and described in Deliverables D6.3.2 and D6.3.3. 3D designs have the potential to suffer from significant thermal problems due to the higher thermal resistance of the adhesive layer used in the F2F assembly An accurate power analysis using EDA Apache tools (RedHawk) have been performed . Figure 40 shows power map of 3D F2F system:

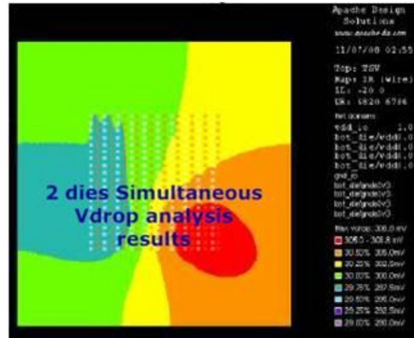


Figure 40: power map of 3D F2F system

Thermal analysis, performed with the methodology described in WP6, has been verified on the real test case fabricated. Figure 41 shows the correspondent thermal maps of individual dies and of the F2F stack achieved using CAD Apache tools.

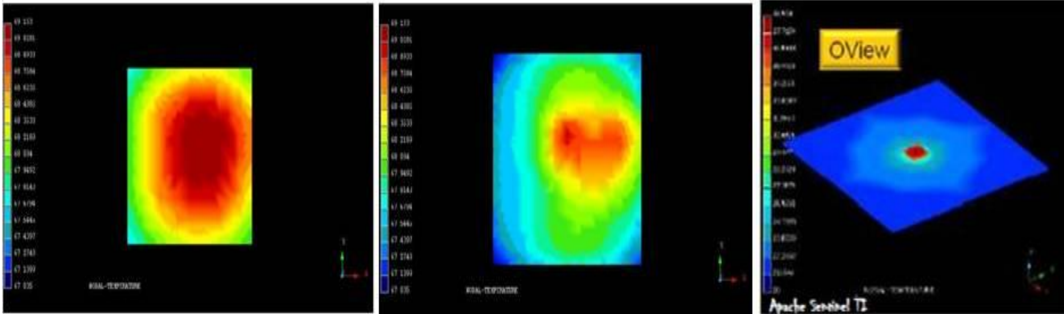


Figure 41: Thermal maps: Top Die, Bottom Die and 3D F2F system

5.3.2 Experimental electrical results:

A test to verify the functionality of the 3D system with respect the thermal constrains and to validate the design flow proposed and developed in WP6 and described in D6.3.2 and D6.3.3 has been performed. A digital signal has been transmitted through 3D TSV interconnect from the top die to bottom die of the stack. The results of transmission shown in Figure 42 on oscilloscope, demonstrate the fully functionality of 3D system, Data In from chip 1 is received by chip 2.

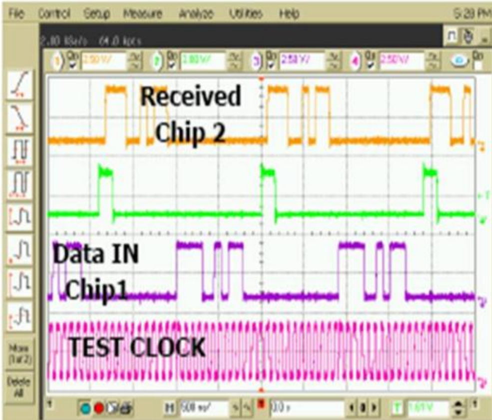


Figure 42: Data Transmission from Top Die to bottom Die in a 3D F2F system

A propagation delay of 5ns was measured in data transmission and the power consumption measured was 0.5mW @ 1MHz (Test clock) plus 6µW of static power (leakage currents). A peak temperature of 68.9 °C was estimated in the operating condition and with nominal supply voltage of 1V. Increasing frequency, the peak temperature increases and a degradation in propagation delay of 5-10% is observed. At higher frequency, the functionality of the device become worst because the temperature increase and a cooling must be introduced to avoid not working condition in the system.

5.4 Conclusion

During the activities carried out in Task 7.2 , ST developed and fabricated some prototypes of 3D stacking device in order to validate the methodology and the design flow that have been proposed in WP6. Two Test chips have been designed and manufactured in ST CMOS 65nm technology following design steps of floor planning, placement and routing that take in account thermal constrains. Thermal and power analyses have been performed with dedicated tools. Two different stacking configuration have been tested , the first with dies flipped down on a silicon interposer and connected through micro bumps, the second with dies stacked face to face and bonded on a printed circuit board with both chip on board and package assembly. This last stacking approach includes as 3D interconnects some TSV. The prototypes have been tested to verify the functionality and performance in term of propagation delay and power consumption.

5.5 Measurable Objectives (MO7.2.7)

In the previous section, different test cases of 3D staked devices have been presented to validate an accurate physical design flow, which includes floor planning, placement and routing and thermal model capabilities. This automated flow, developed in WP6, enables to study the impact of 3D integration on the performance and temperature of a real 3D design achieved coupling IC and Package/PCB in power-thermal co-analysis model at 65nm ST CMOS technology. The temperature variation across the chip is dependent on the power distribution, as well as thermal conductivity and geometry of different materials. The amount of heat removed and the amount of heat trapped in the system are dependent on the design of package, board, heat sink, and cooling system, as well as the temperature difference between the VLSI system and the ambient, All these factor have been considered during assembly of the 3D system prototypes. In the design drivers used as validation test case , the Processor chip DIE 1 and Memory chip DIE 2 are separated by a silicon spacer and 3d interconnect micro bump and TSVs have been included to study the thermal effects in 3D IC. The power and temperature map of single devices shown in the reporting figures are helpful to identify the temperature differences of the on-chip temperature sensors to the actual peak temperature on chip. The effect of 3d interconnect to peak temperature change in 3D system with silicon spacer very conductive are low. The temperature maps have been used for IC power, electro migration, and voltage drop evaluations. The experimental results shown that many of the thermal, electrical, and mechanical properties of the system are also temperature-dependent and they will change as the temperature in the system is changing. These effects have been seen for example, on leakage power in the DIE1 and DIE 2 chip that increases sharply with

temperature, and metal resistivity of 3D Interconnect that increases with temperature. The Thermal analyses and automated flow was demonstrated on measurement of performance of 3D system, power dissipation based on a homogeneous temperature distribution and critical path delay. We have observed that the automatic floor planner used in our work produces good quality floor plans, where blocks whose communication latency is critical for performance are placed close to one another in 2D as well as in 3D. This configuration was used for optimization and evaluation of propagation delay and clock timing in different temperature conditions. Below are summarized a list of parameters that are required for effective evaluate the benefits of thermal analysis on 3D IC design and provides a metric for measure the improvements of adding Thermal behaviour in a design flow.

- ***Platform architecture*** :

Thermal integrity analysis must be based on a single-platform that performs power analysis, thermal analysis, circuit analysis, power/signal net extraction, reduction, and temperature-dependent cell characterization

  - Results achieved in WP6 with automated design flow
  - Quantification: accuracy, simulation time
  - Measurable results: 3D stack of dies power consumption and delay calculation measured at different temperature condition D7.2.1 (Table 9 and Table 10)

- ***Power Calculation***:

Accurate power model of each components and device considered in the design flow including leakage current

  - Results achieved in WP6 with automated design flow
  - Quantification: accuracy, simulation time, Process data from ST CMOS 65nm
  - Measurable results: 3D stack of dies power consumption measured , power maps and total power dissipation D7.2.1 (figure 32, figure 35, figure 36, figure 40, and figure 41)

- ***Thermal simulation***:

Simulation of the temperature profile across the chip, layer by layer, taking into account the thermal properties of different materials, packages, board, heat sinks, and cooling systems.

  - Results achieved in WP6 with automated design flow
  - Quantification: accuracy, simulation time, process data from ST CMOS 65nm , electrical and thermal model of interconnect D6.2.2
  - Measurable results: 3D stack of dies Thermal maps of individual devices D7.2.1 (figure 32, figure 35, and figure 41)

- ***Electrical analysis of the power/signal network and Thermal-aware library characterization***

Analysis must provide temperature-dependent R extraction, reduction, and circuit simulation of the on-chip power/signal network. The complete cell library should be characterized using a temperature-dependent and voltage-dependent library characterization flow based on an accurate Spice simulation engine.

- Results achieved in WP6 with automated design flow, flow is based on ST technology process that provide temperature characterization data for each devices and logic cell libraries
- Quantification: accuracy, simulation time, process data from ST CMOS 65nm
- Measurable results: 3D stack of dies based on CMOS 65nm process , standard EDA tools for automated digital design flow , libraries and IP characterization temperature dependent devices fabricated and shown in D7.2.1 (figure 4 die photo fabricated on ST process 65nm)

- ***Impact on Timing***:

For the timing impact analysis, temperature-dependent and voltage-dependent delay information of the devices is needed

- Results achieved in WP6 with automated design flow, flow is based on ST technology process
- Quantification: accuracy, simulation time, process data from ST CMOS 65nm
- Measurable results: 3D stack of dies based on CMOS 65nm process, standard EDA tools for automated digital design flow measured results on delay of critical path. (table 10 and section 5.3.2 experimental results)

6 Dynamic Thermal Management of Multi-Processor System-on-Chip

6.1 Introduction

In Task 3.4, CSEM and CEA-LETI worked on the dynamic thermal management of Multi-Processor System-on-Chip (MPSoC) circuits. The aim is to prevent thermal hot spots from building up in a chip. Two sets of tools were therefore designed: (1) sensors to track changes, namely temperature sensors and activity monitors – the latter being used for estimating power dissipation –; (2) a Decision Maker or controller, to determine the appropriate operating frequency and eventually supply voltage. This Decision Maker must be stable, fast and easily scalable with the number of processors. In Deliverable D3.4.1, it was shown that a distributed, localized control scheme scales much better with system size than a centralized global control, hence exhibiting smaller reaction times and less computations.
This is why we chose a distributed control scheme based on the Game Theory, where each player – here the individual local Decision Makers – tries to maximize its gains, i.e. tries to minimize the local temperature.

For demonstrating the effectiveness of such a technique, a test-chip called GENEPY was designed and fabricated in CMOS Bulk 65nm technology. It is a 30mm² circuit incorporating clusters from CSEM and CEA-LETI, interconnected by an Asynchronous Network-on-Chip. In this way, each cluster is an independent frequency island. In this implementation, only the frequency can be changed: this is due to the fact that in the target technology, the SRAM memories offer a very limited minimum operating voltage, only 200mV lower than the nominal supply voltage. The necessary hardware for exploiting Dynamic Voltage Scaling, namely DC-DC converters or VDD Hopping, level converters …, was deemed too costly for a small benefit.

The GENEPY test-chip came back late from foundry and the temperature sensors proved somewhat cumbersome to exploit. This deliverable shows preliminary measurement results. However, a number of lessons were learnt. First, it is possible to correlate temperature with activity monitoring, as it was already possible to estimate power dissipation from activity monitoring. However, as thermal effects are really slow by nature, it was found that a fine grain activity monitoring is not necessary: one can rely on a few high level signals, e.g. counting the number of times an FFT operation is executed.
Another lesson learnt, in this special implementation, is that since we used energy-optimized computing cores, the temperature elevation is moderate and is mostly homogenized across the whole chip. As a consequence, all clusters have to exchange their activity monitoring values in order to have a fair picture of the total power dissipation.
Finally, although a correlation between temperature and activity was demonstrated, the temperature sensors cannot be totally removed. Indeed, for a correct temperature estimation, the outside temperature, i.e. directly at the exterior of the package, needs to be known to calculate the thermal flux cooling the chip.

6.2 Implementation of Genepy

The GENEPY platform, designed as a joint work of CEA-LETI and CSEM, contains different types of low-power clusters. Having low-power cores in itself allows reducing the power consumption and thus somehow relaxes the thermal management of the MPSoC. Furthermore, the heterogeneity of a chip provides additional possibilities to trade-off performance, power and temperature.
In GENEPY, the heterogeneity is achieved by providing several types of controller and DSP cores. The IcySMEP core, from CSEM, provides a controller core (called icyflex2) and two DSP cores (called icyflex4). The SMEP cores, from CEA-LETI, provide a MIPS and two specialized FFT 1024 engines, called Mephisto. The floorplan of GENEPY can be seen in Figure 43.

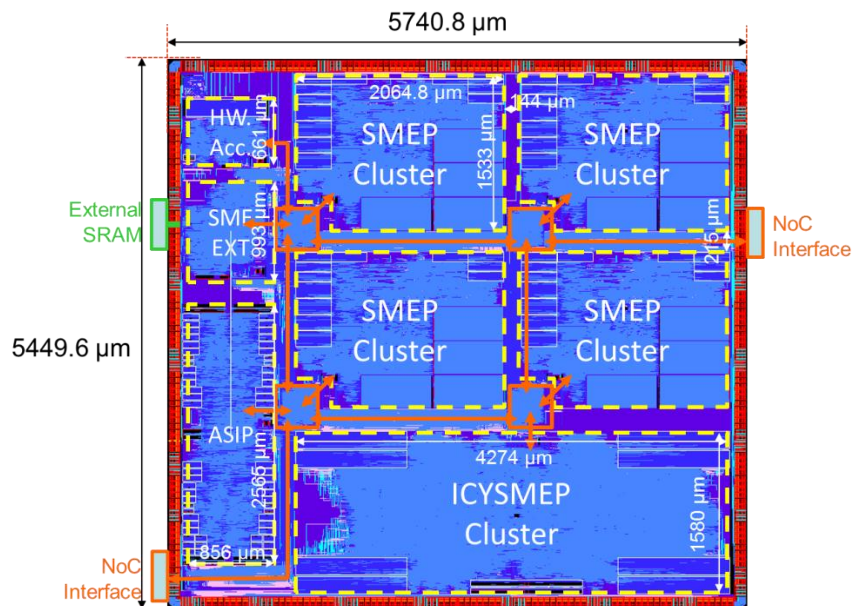


Figure 43 - Floorplan of Genepy.

6.2.1 Temperature Sensors

In the past, different strategies were devised so as to place the temperature sensor most effectively within a chip in order to minimize their number, while being able to identify hotspots on the chip. As described in Deliverable D3.4.1, the thermal resistance of the package is quite high, i.e. calculated close to 1000 K/W. The temperature of the die is therefore homogenized before heat can be evacuated, i.e. the temperature is pretty uniform across the whole die. This means that the temperature is almost the same for the 4 SMEPs and IcySMEP (although a single SMEP was actually running in the experiments).
For avoiding a too high computational complexity and memory usage of the temperature management policy, it was decided to map one thermal sensor per cluster, i.e. 5 temperature sensors are distributed over the GENEPY SoC. The temperature sensors were a-priori located close to the middle of the clusters during the floor planning stage, i.e. no detailed analysis was carried out in order to align the sensors on a DSP or controller core, nor to optimize the number of thermal sensors on the die.
As described in D3.4.1, the thermometers that were included on the MPSoC need a calibration phase. Due to the high sensitivity of the thermometers to intra-die variability and to inter-die variability, each thermal sensor has to be calibrated individually in order to give

an absolute temperature measurement. Furthermore, it was observed that the temperature sensors are very sensitive to the load of the SMEP/IcySMEP clusters. Consequently, a temperature measurement is always done while the DSPs are inactive.
The calibration was done by measuring the oscillation frequency of inverter chains in a temperature-controlled environment (stove). The IcySMEP cluster contains a single thermometer located approximately in the middle of the cluster. This thermometer was calibrated by placing the GENEPY MPSoC and board within a stove at constant and controlled temperature. After one hour at the reference temperature, the chip was powered-on and the temperature sensor was read as quickly as possible, i.e. before the chip had a chance to self-heat. The oven temperature was increased from 0°C to 90°C by steps of 10°C.
Figure 44 shows the sensor read-out vs. the stove temperature for a single die, with an estimation of parameters for a quadratic approximation.

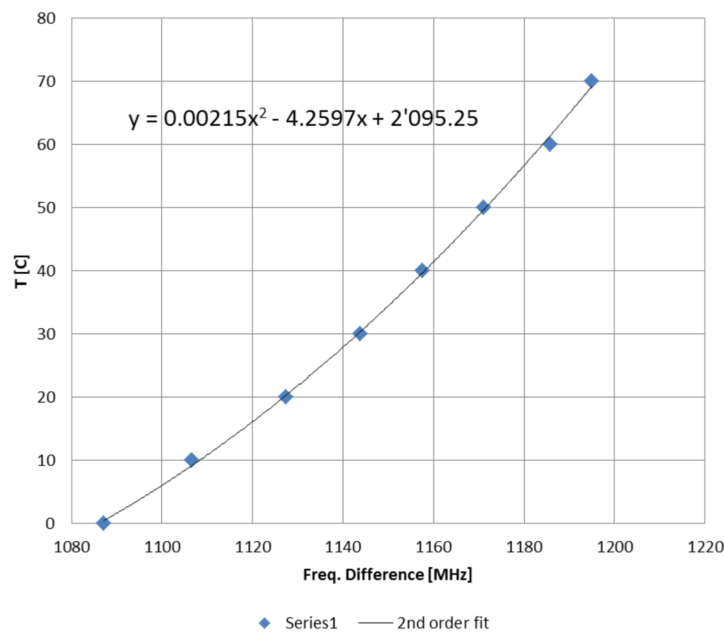


Figure 44 – Temperature vs. sensor frequency difference, with a quadratic fit.

6.2.2 Activity Sensors

6.2.2.1 SMEP Cluster

The SMEP clusters were instrumented with a set of hardware activity monitors. These monitors consist in counters sensitive to different hardware events relevant to the power consumption profile of hardware resources. In the context of thermal estimation, these monitors appear to be much too accurate, given that the temperature fluctuations are observed on a larger area than the one covered by a single monitor and with much longer time constants.
As a consequence, it seems more appropriate to evaluate the activity at a more coarse-grained level. The proposed solution is based on software counters. Indeed, the MIPS processor is in charge of controlling computation tasks on the MEPHISTO cores. It starts the tasks and receives an interruption when the tasks are done. If a counter is incremented for each MEPHISTO core each time a task is done, the MIPS is able to keep a trace on the number of single task execution on MEPHISTO cores.

In parallel, the clock frequency is controlled by the MIPS and the number of clock cycles required for a task is deterministic (in case there is no blocking access to data). The MIPS is then able to compute how much time each MEPHISTO core has spent in computing state and translate it into a temperature variation after synchronization with other SMEP cluster.

6.2.2.2 IcySMEP Cluster

There is no hardware monitor in the IcySMEP cluster, but as seen in the previous paragraph, it is still possible to perform a very rough activity estimation based on software monitoring. This can be done e.g. by exploiting the scheduler in a real-time operating system (RTOS), e.g. the task activity, the number of offload to accelerators and their duration, etc.
FreeRTOS was ported to the icyflex2 controller of the IcySMEP cluster. In a first experiment, only the icyflex2 controller executes code, at a fixed frequency (120 MHz). One task is monitoring the temperature at regular intervals, while another task is a computing intensive one (FIR filtering).
The idleness is defined as the percentage of time the FreeRTOS scheduler spends in the idle state. It also corresponds to the amount of spare processing capacity of the system, and gives a hint on how much the frequency could be reduced without altering the computing capacity (not taking into account real-time deadlines). To measure this idleness, an idle task hook is set up in FreeRTOS.
A correlation between the idleness of the RTOS and the temperature of the sensor would mean that the idleness could be used to perform thermal management. This technique is traditionally used at the level of a single core to modify its operating frequency [13].
The idleness of the system was zero in the first experiment, i.e. the FIR was running all the time on random data. The system was started at room temperature (20°C). It is observed (Figure 45) that the temperature rapidly reaches 23°C. It is however not clear whether this temperature rise is due to the GENEPY self-heating or due to other components on the board. The temperature increase is relatively small.

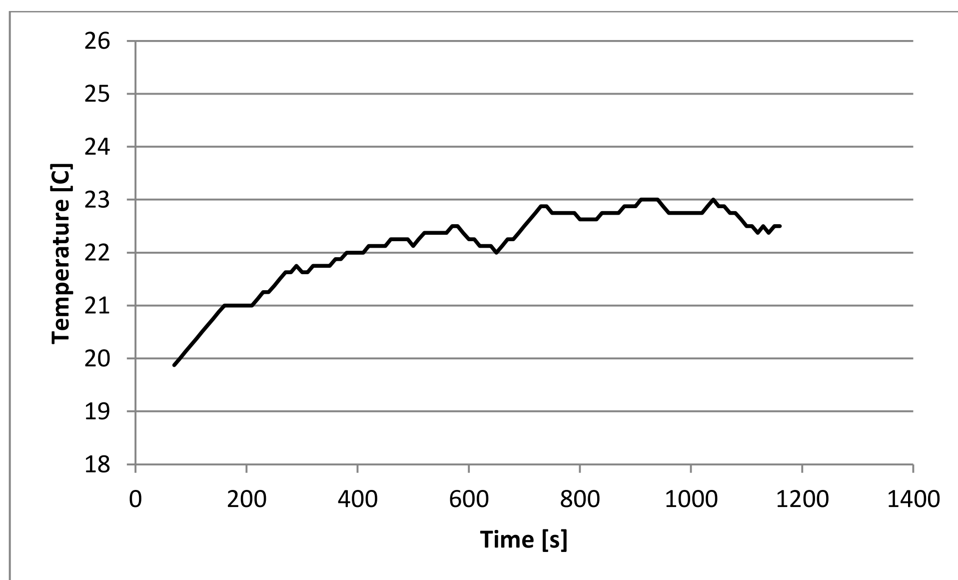


Figure 45 – Temperature when the icyflex2 executes an FIR filtering task at start up (cold start).

After a while, the measure temperature reaches a steady-state (Figure 46).

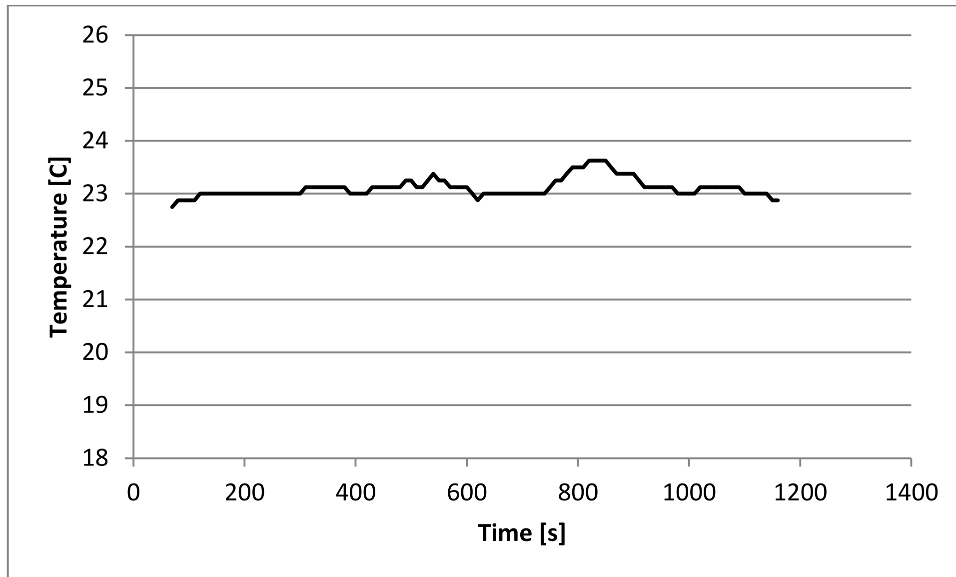


Figure 46: Temperature when the icyflex2 executes the FIR filtering (steady-state).

So far, the icyflex4 DSPs were not active. In the next experiment, the two DSPs also execute some FIR filtering tasks. Figure **47** shows the two temperature curves over time, when only one controller (icyflex2) or the controller and two DSPs (icyflex4) are active. The DSPs perform the same task as the controller (FIR filtering). A temperature increase of less than one degree is observed. This can be explained by the very low power consumption of the cores (in the order of a few mW).

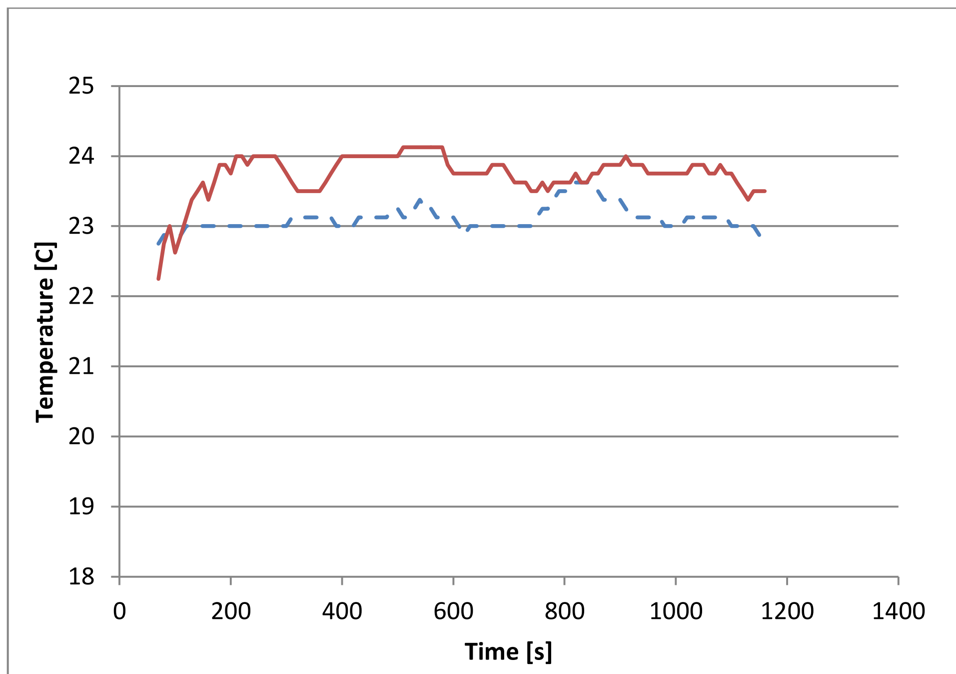


Figure 47: Temperature comparison for a single icyflex2 active (dashed blue line), and for an icyflex2 and two icyflex4 cores active (plain brown line).

6.3 Correlation between Temperature Measurement and Activity Monitoring

The SMEP clusters from CEA-LETI were fitted with activity monitors, but based on experiment, it was found that only high-level signal monitoring was enough for estimating the temperature. It was thus decided to use the MIPS to just count the number of FFT computations performed in a 10μs time frame.

The temperature measurement on the 4 SMEPs was carried out after the calibration of the temperature sensors in an oven. In the considered testbench, there are two FFT computations running simultaneously on SMEP_00, between time 600s and time 1200s. Measured values are given in Figure 48. The initial phase, between time 0 and 600s, shows the configuration of the circuit followed by a long period to reach a stable temperature value; then there is the actual FFT phase with a temperature rise of 1.5°C; and finally a third phase where there is no computation anymore. It can be seen that the temperature is pretty much uniform between the 4 SMEP clusters, with roughly a 0.5°C temperature difference: SMEP_00 is the hottest one, while SMEP_11, which is the cluster the farthest from it, is the coolest one.

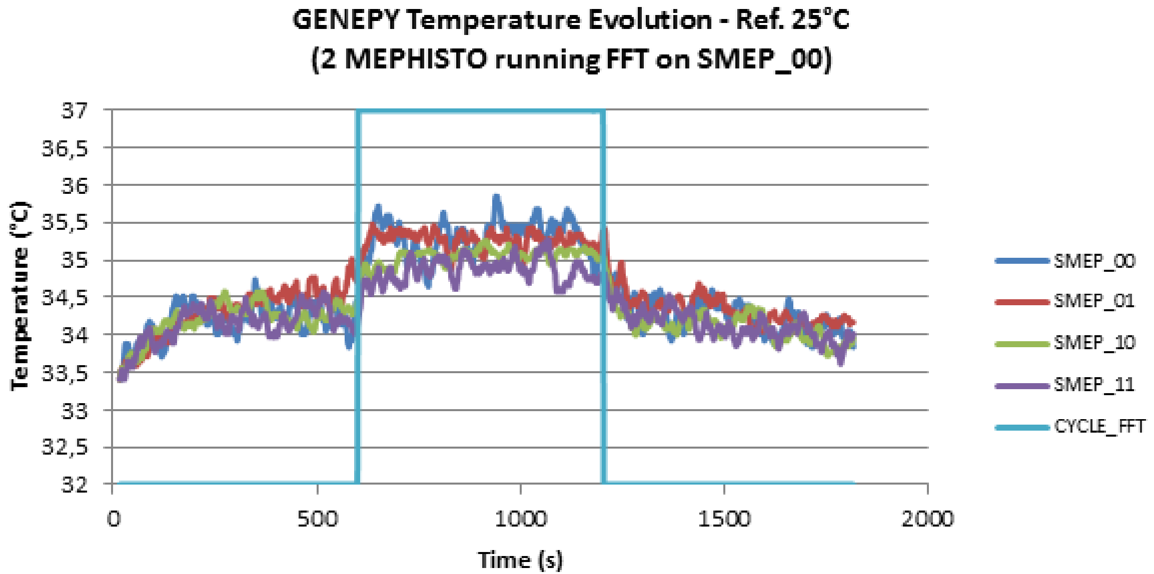


Figure 48 – Temperature measurement results.

For estimating the temperature from the activity monitoring, and thus the corresponding power dissipation, the equation to be solved is:

$$C\, {}^{dT(t)}\!/_{dt} + AT(t) = Pu(t) \qquad (1)$$

Where C is the thermal capacitance matrix ([n*n] diagonal matrix), A is the thermal conductivity matrix ([n*n] matrix), $T(t)$ and P are temperature and power vectors (of n-dimensional Euclidean space R^n) and $u(t)$ is the unit step function, *i.e.* the time step used by the MIPS to periodically read the activity monitors.

The temperature estimated by the MIPS at each time step is calculated thanks to the following equation:

$$T(t) = T\big(t - u(t)\big) + \frac{P - A * \big(T\big(t - u(t)\big) - T_{Ambient}\big)}{C} * u(t) \qquad (2)$$

The unit time step *u(t)* was chosen to be equal to 10µs. The temperature estimation is shown in Figure 49 and superimposed with the corresponding measurement results. It shows a pretty good correlation. For this curve, the following parameters were set:

- A = 0.082 W/m.K (thermal conductance);
- C = 8 J/K (thermal capacitance).

As shown in Equation $T(t) = T\big(t - u(t)\big) + \frac{P - A * \big(T\big(t - u(t)\big) - T_{Ambient}\big)}{C} * u(t)$ (2), it is necessary to know the temperature outside of the package to be able to calculate the thermal flux to the heat sink, meaning that it is not possible to not rely on an absolute temperature value measurement. An on-die temperature measurement would probably be necessary also to safeguard against thermal runaway, due to a positive feedback between temperature and leakage currents.

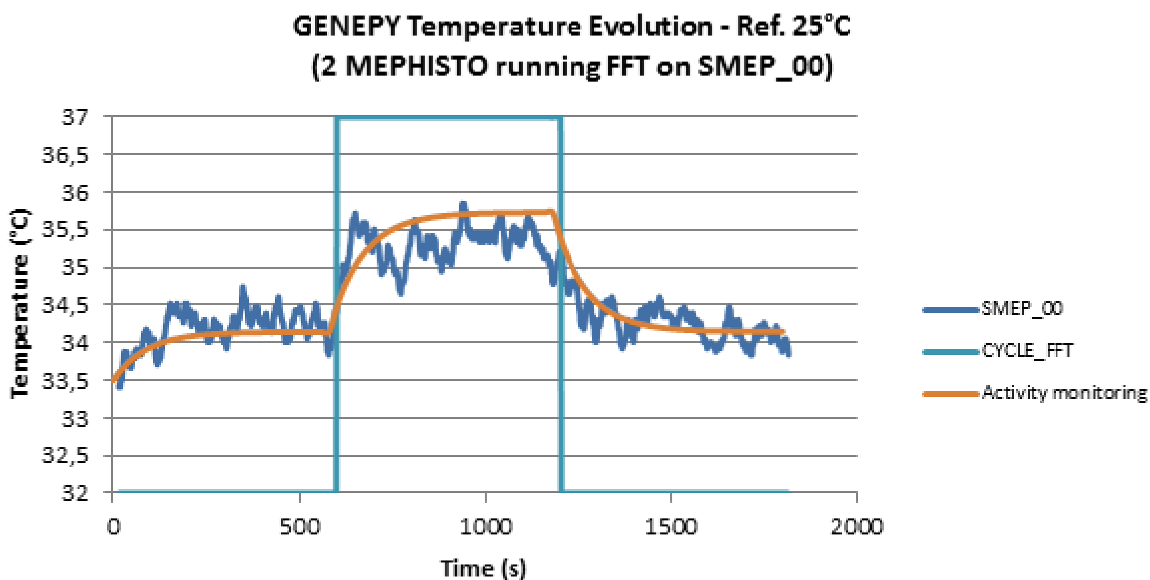


Figure 49 – Comparison of temperature estimation from activity monitoring with measurement results.

6.4 Conclusion

The design techniques developed in WP3, in order to dynamically manage the temperature of a Multi-Processor SOC, were demonstrated in the GENEPY test-chip. GNEPEY is an heterogeneous MPSOC incorporating computing clusters from CSEM and CEA-LETI: it is fitted with temperature sensors and activity monitors – implemented either in hardware or in software. A number of lessons were learnt from this experiment. First, since thermal effects are really slow by nature, a fine grain activity monitoring is not necessary: a good temperature estimation can be obtained by relying only on a few high level signals, e.g. counting the number of times an FFT operation is executed. Then, in computing platforms dedicated to mobile applications like GENEPY, and thus already energy optimized, the temperature is

pretty much homogenized, meaning that each core has to provide its activity data to all others in order to have a fair temperature estimation. Finally, a good temperature estimation cannot be obtained without an absolute temperature sensor, which might seem utterly paradoxical. The explanation is that the temperature outside the package must be known precisely in order to calculate the thermal flux cooling the chip, which is dependent on the package thermal properties but also on the temperature difference between the inside and the outside.

Concerning Measurable Objective MO7.2.6, it was demonstrated that the temperature of a chip can be properly estimated from its activity: it is thus met.

7 Conclusions

In this deliverable, the results of the demonstration of thermal-aware and energy-efficient design techniques on different electronic devices and technologies are presented. Examples are given in digital design, advanced CMOS, and designs using power devices. The presented work is closely related to the other WPs of Therminator. Test-cases of WP1 are used, together with examples provided by individual partners. Furthermore, the models and tools developed in WP2, 3, 5, and 6 are used in this task. Before the modelling tools of the industrial partners (ST and IMC) were used at the level of designs, validation on device level as part of T7.1 was carried out.

The exploration of the modelling tools on practical design has resulted in interesting results, varying from better understanding of thermal behaviour in circuits to improved devices. More specifically, the thermal-aware design framework developed by ST in Therminator has been used to improve the thermal behaviour of discrete power transistors under application conditions. The peak temperature of the power transistor in an operation cycle was reduced, and the impact on the cycle lifetime has been addressed. UNIBO have demonstrated interesting novel design techniques for class-D amplifiers. It was demonstrated that the driving frequency can be reduced by a factor of two, which in turn leads to 50% less energy consumption. IMC has presented results on the performance and power consumption (dynamic as well as leakage) of a generic critical path structure, a building block in their ICs. The critical path structure is simulated with the PSP-based design flow, developed in WP2 and validated in T7.1. The comparison to measurements, performed on the advanced 28 nm CMOS node, shows accurate agreement with the simulation results. In other words, the thermal effects for IMC's mainstream LTE/3G digital baseband ICs can be handled with high-quality circuit models. In addition, an extensive overview of the key figures of merit over a wide temperature range has been presented. POLITO tested the developed thermal-aware flow by comparison of simulations to measurements on structures designed in advanced CMOS nodes. Their tool will be used for optimization in T7.3. Furthermore, ST demonstrated the thermal 3D physical design flow developed in WP6. Several prototypes of 3D stacked devices have been fabricated. The functionality and performance of these prototypes was measured and analyzed. CSEM and CEA-LETI investigated the dynamical management of the temperature of a Multi-Processor System on a Chip using activity monitors and temperature sensors.

All of the results are quantified by measurable objectives, and a detailed overview of these objectives, linked to the general goals of Therminator is presented in chapter 8. Finally, the results obtained have been used for conference contributions and publications, as shown in chapter 9, demonstrating the novelty of the work.

8 Measurable objectives

The Therminator project objectives are

1. New modeling and simulation capabilities to support accurate circuit thermal analysis and simulation
2. Innovative thermal-aware design techniques, methodologies and prototype tools for controlling, compensating and managing thermally-induced effects on parameters such as timing, (dynamic and leakage) power, reliability and yield
3. Demonstration of the accuracy and ease of integration within existing design flows of the new models by validation against measured data obtained on ad-hoc silicon structures
4. Demonstration of the applicability and effectiveness of the new design solutions through manufacturing of test-chips featuring leading-edge silicon technology, as available from some of the project partners
5. Demonstration of the usability and effectiveness of the new design methodologies and tools by their application to industry-strength design cases made available by some of the project partners

In order to quantify the output of the demonstration work on design techniques presented in this deliverable, measurable objectives have been defined. In the table below, all the measurable objectives of T7.2 are summarized. In this table, it is also shown to which test case of WP1, and to which Therminator's project objectives the measurable objectives are related. Since T7.2 focuses on demonstration of design techniques, the measurable objectives are related to Therminator's project objectives 2-5.

Area	Measurable objective	Innovation Metric	Quantification	Test case	Project objective
Demonstration of design techniques on test chips	MO7.2.1	WP5 Validate thermal models of discrete power devices in typical automotive application with measurements	Design cycle time **30% Faster** [ST]	Testcase 2	5
	MO7.2.2	Adapt the driving signal frequency to the input signal power by taking advantage of non-linear effects	**50% reduction** of driving frequency, resulting in a 50% percentage reduction of energy consumption [UNIBO]	Class D amplifier design **Testcase 4**	2
	MO7.2.3	New design technique for generating signals with assigned zero-crossings	Error in zero positioning in order of nanosec [UNIBO]	Driver circuit of class D amplifier Testcase 4	4
	MO7.2.4	Quantify the thermally induced effects on advanced CMOS technologies	Presence of the Inverted Temperature Dependence (ITD) for CMOS circuits powered with nominal supply voltage. (yes/no) [POLITO]	Datapath Macrocell provided by IMC Testcase 3	2
	MO7.2.5	Modelling interaction of process variations, supply	40nm CMOS: <5% performance impact over T=[-30, 85]°C range for	measurement & simulation on	3

		voltage corner and temperature effects	typical wireless baseband operation (ref point:1.2V, T=25°C, nom =100% performance) [IMC] 28nm CMOS: <10% performance impact over T=[-30,85]°C range for typical wireless baseband operation (ref point: 1.1V T=25°C, nom =100% performance)[IMC]	the product critical path structure in IMC's WP 7.2 Testcase 3	
	MO7.2.6	Estimation of thermal dissipation from activity monitoring	Comparison of data provided by built-in activity monitors and temperature sensors (yes/no) [LETI/CSEM]	Genepy testchip Testcase 4	4
	MO7.2.7	Functionality and measurements of face to face 3D stacked structure	Quantify temperature impact on performance of a heterogeneous 3D stacked SiP (yes/no) [ST]	3D stacked System in Package Testcase 6	3

9 Publications and presentations

- A. Burenkov, J. Lorenz, "*Self-heating effects in nano-scaled MOSFETs and thermal aware compact models"*, THERMINIC, 17th International Workshop on Thermal investigations of ICs and Systems, Paris, 27-29 Sept. 2011, EDA Publishing, pp. 17-18.

- FHG, *"Tutorial on TCAD Simulations of Nano-CMOS Including Self-Heating"* was held at IMC in September 2011 at IMEC in March 2012

- F. Beneventi, A. Bartolini, L. Benini, *"Static Thermal Model Learning for High-Performance Multicore Servers"*, Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on, Issue Date: July 31 2011-Aug. 4 2011, On page(s): 1 – 6, Location: Lahaina, HI, USA, ISSN: 1095-2055 Print ISBN: 978-1-4577-0637-0, 2011 IEEE

- A. Bartolini, M. Cacciari, A. Tilli, L. Benini, *"A distributed and self-calibrating model-predictive controller for energy and thermal management of high-performance multicores",* Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011, Grenoble, France 14-18 March 2011, On page(s): 1 – 6, ISSN : 1530-1591, Print ISBN: 978-1-61284-208-0. IEEE Press 2011

- A. Bartolini, M. Sadri, F. Beneventi, M. Cacciari, A. Tilli, L. Benini, "*SCC Thermal Sensor Characterization and Calibration*", 3rd Many-core Applications Research Community (MARC) Symposium, Ettlingen, Germany, Issue Date : 5-6 June 2011, On page(s): 7-12, KIT Scientific Publishing 2011, ISBN 978-3-86644-717-2

- A. Bartolini, M. Sadri, F. Beneventi, M. Cacciari, A. Tilli, L. Benini, *"A System Level Approach to Multi-core Thermal Sensors Calibration"*, Integrated Circuit and System Design. Power and Timing Modeling, Optimization, and Simulation, Editor: Ayala J., García-Cámara B., Prieto M., Ruggiero M., Sicard G., Book Series Title: Lecture Notes in Computer Science, Page(s): 22- 31, Volume: 6951, Copyright: 2011, Publisher: Springer Berlin / Heidelberg, ISBN: 978-3-642-24153-6

- M. Sadri, A. Bartolini, L. Benini, "*Single-Chip Cloud Computer thermal model"*, Thermal Investigations of ICs and Systems (THERMINIC), 2011 17th International Workshop on, Paris, France 27-29 Sept. 2011, On page(s): 1 – 6, Print ISBN: 978-1-4577-0778-0, IEEE Press 2011

- A. Sassone, A. Calimera, A. Macii, E. Macii, M. Poncino, R. Goldman, V. Melikyan, E. Babayan, S. Rinaudo, *"Investigating the Effects of Inverted Temperature Dependence (ITD) on Clock Distribution Networks",* Proceedings of Design, Automation & Test in Europe (DATE'12) conference, Dresden, Germany, 2012.-P.165-167

- Wei Liu, V. Tenace, A. Calimera, A. Macii, E. Macii, M. Poncino, *"NBTI Effects on Tree-Like Clock Distribution Networks",* GLSVLSI-12: accepted for publication

- A. Sassone, W. Liu, A. Calimera, A. Macii, E. Macii, M. Poncino, *"Modeling of thermally induced skew variations in clock distribution network"*, THERMINIC-11: IEEE Thermal Investigations of ICs and Systems, 2011.

- L. M. de Lima Silva, A. Calimera, A. Macii, E. Macii, M. Poncino, *"Power Efficient Variability Compensation Through Clustered Tunable Power-Gating",* IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol.1, no.3, Sept. 2011
- M. Caldera, A. Calimera, A. Macii, E. Macii, M. Poncino , *"Minimizing temperature sensitivity of dual-Vt CMOS circuits using Simulated-Annealing on ISING-like models",* THERMINIC-10: IEEE Thermal Investigations of ICs and Systems, 2010.
- A. Calimera, A. Macii, E. Macii, S. Rinaudo, M. Poncino, *"THERMINATOR: Modeling, control and management of thermal effects in electronic circuits of the future"*, THERMINIC-10: IEEE Thermal Investigations of ICs and Systems, 2010.
- A. Calimera,R. Bahar, E. Macii, M. Poncino, *"Temperature-Insensitive Dual-Vth Synthesis for Nanometer CMOS Technologies Under Inverse Temperature Dependence",* IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 18, no. 11, 2010
- A. Timár, M. Rencz, "*Studying the influence of chip temperatures on timing integrity",* 12th IEEE Latin-American Test Workshop. Brazil, 27-30. 03. 2011.
- A Timár, M. Rencz, *"Studying the Influence of Chip Temperatures on Timing Integrity Using Improved Power Modeling",* JOURNAL OF LOW POWER ELECTRONICS 7: pp. 1-10. (2011)
- A Timár, Gy. Bognár, M. Rencz, "*Improved power modeling in logi-thermal simulation",* 17th International Workshop on Thermal investigations of ICs and Systems. Paris, Paris, France, 27-29. 09. 2011.
- Gergely Nagy, András Poppe, *"A Novel Simulation Environment Enabling Multilevel Power Estimation of Digital Systems",* Proceedings of the 17th International Workshop on THERMal INvestigation of ICs and Systems
- G. Gangemi, *"FP7-Funding Projects THERMINATOR, SMAC, MANON Overview"*, MUGM MunEDA User Group Meeting 2012, October 2012, Munich, Germany
- Z. Abbas, M. Olivieri, A. Ripp, G. Strube, M. Yakupov, "*Yield optimization for low power current controlled current conveyor"*, SBCCI 2012, September 2012, Brasília, Brazil
- Colaci, G. Boarin, A. Roggero, L. Civardi, C. Roma, A. Ripp, M. Pronath, G. Strube: *"Systematic Analysis & Optimization of Analog/Mixed-Signal Circuits Balancing Accuracy and Design Time*", SBCCI 2011 Brazil, September 2011, Sao Paolo, Brazil
- N. Seller, *"Optimization of a 2.133GHz level shifter in 28nm*", MUGM MunEDA User Group Meeting 2011, Munich, Germany
- U. Trautner, M. Pronath, *"Synopsys Custom and Analog Mixed-Signal Overview & MunEDA WiCkeD Integration*", MUGM MunEDA User Group Meeting 2010, Munich, Germany
- S. Coparale, R. Rvatti, G. Setti, "*Representation of PWM signals through time warping*", *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* , vol., no., pp.3589-3592, 25-30 March 2012.

- Reef Eilers, Malte Metzdorf, Sven Rosinger, Domenik Helms, Wolfgang Nebel, "*Phase space based NBTI model*", *Proc. of International Workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS), 2012*
- Sven Rosinger, Malte Metzdorf, Domenik Helms, Wolfgang Nebel, *"Behavioral-Level Thermal- and Aging-Estimation Flow", Proc. Of 12th Latin-American Test Workshop (LATW), p. 1-6, 2011*
- V. Melikyan, A. Gevorgyan, A. Baghdasaryan, H. Melikyan, *"Thermal Via's Placement Zones Identifying Using Voronoi Diagrams*", Proceedings of the 32th International Scientific Conference Electronics and Nanotechnology (ELNANO 2012), Kiev, Ukraine, 2012.-P.77-79
- V. Melikyan, Babayan E., Harutyunyan A., Melikyan N., Zargaryan G., "*Method of Reducing Thermal Dependence of Timing Delays of Digital Integrated Circuits",* Proceedings of 5th All-Russian scientific-technical conference "Problems of Developing Advanced Micro- and Nanoelectronic Systems -2012" (MES-2012), Moscow, Russia, 2012. –P409-412
- V.Sh. Melikyan, A.A. Durgaryan, A.H. Balabanyan, E.H. Babayan, M. Stanojlovic, A.G. Harutyunyan, *"Process-voltage-temperature Variation Detection and Cancellation Using On-Chip Phase-Locked Loop"*, Proceedings of the 56th Electronics, Telecommunications, Computers, Automatic Control and Nuclear Engineering (ETRAN) Conference, Zlatibor, Serbia, 2012.-P.EL1.2-1-4
- P. Bibilo, A. Solovev, V. Melikyan, A. Harutyunyan, E. Babayan, "*Estimation of Power Consumption of Digital CMOS Circuits Based on Logic Simulation of their Structural Descriptions"*, Proceedings of Engineering Academy of Armenia, Yerevan, Armenia, 2012. – P.600-610
- V. Melikyan, E. Babayan, A. Harutyunyan, *"Pattern-Based Approach to Current Density Verification*", Proceedings of the 4th Small Systems Simulation Symposium 2012, Nis, Serbia, 2012.-P.58-61
- V. Melikyan, A. Balabanyan, E. Babayan, A. Durgaryan, *"Decreasing of Frequency Variation in High-Speed Ring Oscillator using Bandgap Reference",* Proceedings of the 32th International Scientific Conference Electronics and Nanotechnology (ELNANO 2012), Kiev, Ukraine, 2012.-P.79-81
- R. Roldman, K. Bartleson, T. Wood, V. Melikyan, E. Babayan, "*Synopsys' Low Power Design Educational Platform*", Proceedings of the 9th European Workshop on Microelectronics Education (EWME 2012), Grenoble, France, 2012.-P.23-26
- V. Melikyan, E. Babayan, A. Harutyunyan, *"Pattern-Based Approach to Current Density Verification",*/ Electronics, Faculty of Electrical Engineering, University of Banja Luka, Volume 16, Number 1, Serbia, 2012.-P.77-82
- V. Melikyan, A. Harutyunyan, *"Modeling of IC Interconnects and Power Rails*", Chartarapet, Yerevan, 2012 (in Armenian)
- V. Melikyan, A. Durgaryan, A. Khachatryan, H. Manukyan, E. Musayelyan, *"Self-compensating Low Noise Low Power PLL Design*", Proceedings of IEEE East-West Design & Test Symposium (EWDTS'12), Kharkov, Ukraine, 2012.-P.29-33

- V.Sh. Melikyan, S.V. Gavrilov, V.K. Aharonyan, N.K. Aslanyan, A.S. Hovhannisyan, *"On-die CMOS Termination Resistor for USB Transmitter"*, RAs National Academy of Science and SEUA, Yerevan, RA, Vol. 65, N 3, Yerevan, 2012.-P. 295-304

- P. Magnone, C. Fiegna, G. Greco, G. Bazzano, E. Sangiorgi, S. Rinaudo, *"Modeling of Thermal Network in Silicon Power MOSFETs"*, Ultimate Integration on Silicon (ULIS), 14-16 Marzo 2011, Cork, Ireland.

- P. Magnone, C, Fiegna, G. Greco, G. Bazzano, S. Rinaudo, E. Sangiorgi, *"Numerical Simulation and Modeling of Thermal Transient in Silicon Power Devices"*, Ultimate Integration on Silicon (ULIS), pp. 153-156, 6-7 March 2012,. Grenoble (France).

- P. Magnone, C, Fiegna, G. Greco, G. Bazzano, S. Rinaudo, E. Sangiorgi, "*Numerical Simulation and Modeling of Thermal Transient in Silicon Power Devices*", ELSEVIER Solid-State Electronics, in press.

- H. Oprins, V. Cherman, B. Vandevelde, M. Stucchi, G. Van der Plas, P. Marchal, and E. Beyne, "*Steady state and transient thermal analysis of hot spots in 3D stacked ICs using dedicated test chips*", 27th Annual IEEE Thermal Measurement, Modeling and Management Symposium (SEMI-Therm), March 20-24, 2011, 131-137.

- H. Oprins, V. Cherman, B. Vandevelde, C. Torregiani, M. Stucchi, G. Van der Plas, P. Marchal, and E. Beyne, "*Characterization of the Thermal Impact of Cu-Cu bonds achieved using TSVs on hot spot dissipation in 3D stacked ICs*", Proceedings of ECTC, May 30- June 1, 2011, 861-868.

- H. Oprins, V. Cherman, *"Numerical and experimental characterization of hot spot dissipation in 3D stacks"*, Electronics Cooling Magazine, Vol. 18(2), 2012, pp. 18-23.

- D. Milosevic, H. Oprins, J. Ryckaert, P. Marchal, G. Van der Plas, "*DRAM-on-logic Stack – Calibrated Thermal and Mechanical Models Integrated into A Design Flow"*, IEEE Custom Integrated Circuits Conference (CICC), September 18-21 2011, San Jose, California, *invited*.

- H .Oprins, V. Cherman, B. Vandevelde, G. Van der Plas, P. Marchal, and E. Beyne, *"Numerical and experimental characterization of the thermal behavior of a packaged DRAM-on-logic stack*", 62nd Electronic Components and Technology Conference - ECTC, 2012, pp. 1081-1088.

- Gergely Nagy, László Pohl, András Timár, András Poppe, "*Yield enhancement by logi-thermal simulation based testing",* Proceedings of the 18th International Workshop on THERMal INvestigation of ICs and Systems (THERMINIC'12). Budapest, Hungary, 2012.09.25-2012.09.27. pp. 196-199. Paper 42.

- Gergely Nagy, András Timár, Albin Szalai, Márta Rencz, András Poppe, "*New simulation approaches supporting temperature-aware design of digital ICs",* Proceedings of the 28th IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM'12). San Jose, USA, 2012.03.18-2012.03.22. pp. 313-318.(ISBN: 978-1-4673-1109-0)

- A Timar, M. Rencz, "*Temperature dependent timing in standard cell designs",* Proceedings of the 18th International Workshop on THERMal INvestigation. Budapest, Hungary, 2012.09.25-2012.09.27. pp. 179-183.

- A Timar, M. Rencz, "*Real-time heating and power characterization of cells in standard cell designs",* MICROELECTRONICS JOURNAL (2012)IF: [0.919*]
- A Timar, M. Rencz, "*Acquiring real-time heating of cells in standard cell designs",* Proceedings of the 13th IEEE Latin-American Test Workshop (LATW'12). Quito, Ecuador, 2012.04.10-2012.04.13. pp. 121-125.
- Gergely Nagy, András Poppe*, "Simulation Framework for Multilevel Power Estimation and Timing Analysis of Digital Systems Allowing the Consideration of Thermal Effects"*, Proceedings of the 13th IEEE Latin-American Test Workshop (LATW'12). Quito, Ecuador, 2012.04.10-2012.04.13. pp. 1-5.