REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
18-07-2017	Final Technical	Jul 2015 - Jul 2017

4. TITLE AND SUBTITLE
NICOP - Thermally Optimized Paradigm of Thermal Management (TOP-M)

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N62909-15-1-2033

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Prof. Emeritus Yael Nemirovsky, Professor Ran Ginosar, and
PhD Student Maria Malits

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Technion R&D Foundation Ltd.
Technion City
Haifa 32000 ISRAEL

8. PERFORMING ORGANIZATION REPORT NUMBER
2021701

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
ONR Global
N62909
86 Blenheim Crescent
Ruislip MX HA47HB

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The main goal of this research was to present a New Thermal Management Approach, which combines thermally aware Very/Ultra Large Scale Integration (VLSI/ULSI) architecture design with liquid cooling design from the very early step of the demonstrator System on Chip (SoC) design. Thermal sensors, which were integrated on several parts of the SoC, measured the on-line the local temperature, thus enabling better operation of the SoC as well as better control over the liquid cooling by providing on-line feedback. Complementary Metal Oxide Semiconductor (CMOS) Single Photon Avalanche Diode (SPAD) image sensors were used to demonstrate the new thermal management approach.

15. SUBJECT TERMS
Thermal management, integrated temperature sensors, Vt extractor, SPAD, CMOS-SOI technology

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Yael Nemirovsky
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	34	97248294688

Thermally Optimized Paradigm Of Thermal Management (TOP-M)

Final Report 2015-2017

PI's: Prof. Emeritus Yael Nemirovsky and Professor Ran Ginosar

And

PhD Student Maria Malits

Project's methodology and objectives:

This project's main objectives are summarized below:
While addressing thermal effects in Very/Ultra Large Scale Integration (VLSI/ULSI) Complementary Metal Oxide Semiconductor (CMOS) chips, a new paradigm should be introduced:

1. **Thermal effects should be included at the first stage of ULSI/VLSI design.**
2. **Thermal cooling technology should be included as part of die architecture and packaging.**

The main goal of the proposal is to present the New Thermal Management Approach, which combines thermally aware Very/Ultra Large Scale Integration (VLSI/ULSI) architecture design with liquid cooling design from the very early step of the demonstrator System on Chip (SoC) design. Thermal sensors, which are integrated on several parts of the SoC, measure on-line the local temperature, thus enabling better operation of the SoC as well as better control over the liquid cooling by providing on-line feedback. Complementary Metal Oxide Semiconductor (CMOS) Single Photon Avalanche Diode (SPAD) image sensors are used for demonstrating the new thermal management approach.

1. Step 1 - Thermal characterization of Tower Jazz CMOS 0.18µm process

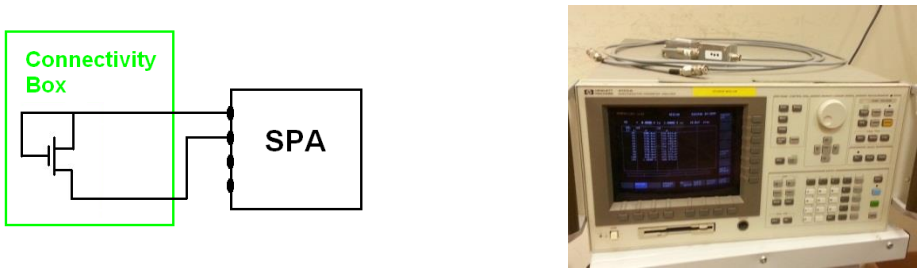
The first step for assessing the chip temperature distribution is to evaluate the chip thermal performance. This part is performed during electrical testing on single transistors with relatively small dimensions in order to avoid self-heating effects as well as coupled-heating, due to the low power consumption during operation. In such testing, the transistor temperature is well-known and regulated by an external temperature controller of the prober chuck. The resulting thermal performance can then be attributed to temperature-dependent process parameters such as threshold voltage, mobility, and non-ideality factor. Subsequently, by monitoring the on-line changes in V_t under actual operation, the area temperature rise due to self and joule heating is measured.

1.1 Set-up and Instrumentation

Usually for semiconductor devices, a small number of measurements are sufficient for deriving all the parameters that are needed for their electrical and thermal characterization. In this study, I-V (current-voltage) curves were extensively measured with long integration times (DC) and then analyzed using Matlab. Additional information was obtained from measuring the I-V curves of devices with different W/L ratio at various temperatures regions, from which the self-heating of the devices is evaluated.

When the kind of measurement allows it, it is generally preferable to work with dies that are bonded to a package (e.g. a ceramic DIP) in order to make the connectivity simpler. The thermal performance, however, had to be evaluated on bare dies since a good thermal coupling with the temperature-controlled chuck was required.

The Semiconductor Parameter Analyzer (SPA) is an instrument that, among other functions, allows the user to force a voltage and to measure the series current by using just a single pair of wires. The precision reached by this instrument is on the order of 1 pA, but triax wires must be used all the way down to the sample to avoid measuring the leakage currents of the wire together with the signal. The measurements presented here were all made with an HP-Agilent 4145B SPA, connected as shown on Figure 1 (a)-(b).



(a) (b)

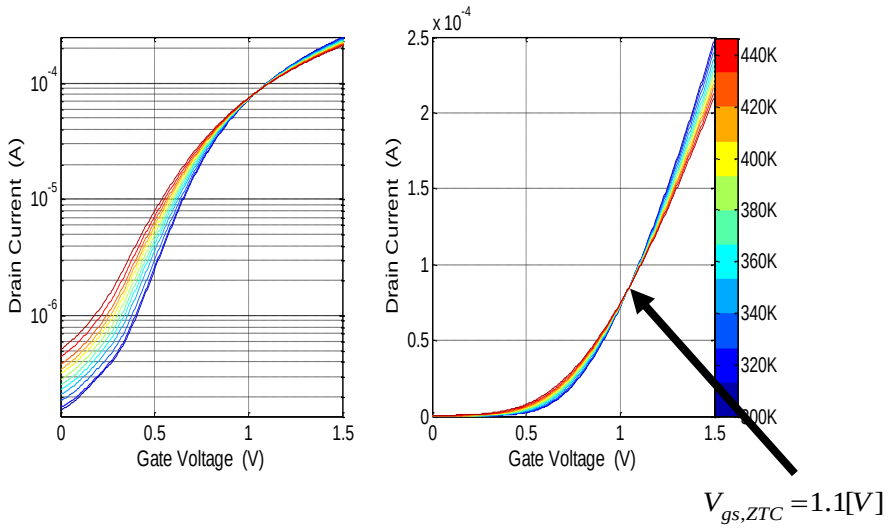
Figure 1: (a) electrical connection scheme for DC and thermal measurements (b) Agilent 4145B SPA front panel view.

When the thermal response was characterized, the SPA was connected to a probe station, making contact directly with the bonding pads on the dies. For measurements at room temperature and above, a chuck with a Joule heater and a thermal sensor was operated through an Electrotherm temperature controller.

1.2 Current- voltage Characteristics

In the calibration process the current-voltage characteristics of n-type and p-type CMOS transistors with several gate dimensions (W/L) were extensively measured as a function of temperature in the range of 300-450K. The I-V characteristics were measured in a wide range of temperatures in order to determine the gate voltage corresponding to the zero temperature coefficients (ZTC) point, as outlined below. At each temperature set point, a gate voltage sweep was conducted in the range of 0-1.5V with steps of 10 mV and the drain voltage was fixed at 2V, corresponding to transistor operating in subthreshold and saturation. The typical dimensions of the transistors under study correspond to long channel behavior (L is at least 0.5 μm and W is more than 1μm); thus, short channel effects may be ignored.

Figure 2 presents results for I-V sweep performed on a long channel NMOS device having a size of 10μm/0.6μm. Due to the small dimensions of this device, self-heating effects may be neglected and it is used as "reference" for the calibration process.



(a) (b)

Figure 2: I-V sweep performed from room temperature till 440K on an NMOS transistor with W/L= 10 μm /0.6 μm: on a semi-log scale (a) and on a linear scale (b).

This kind of measurements yield the threshold voltage (Vt) and its temperature dependence, the subthreshold slope (swing) and the non-ideality factor of 0.18 μm CMOS process under study. Then, we used this

thermal characterization to determine the temperature rise of the chip during operation using "Vt extractor" circuits.

1.3 Extraction of the process threshold voltage (Vt)

A basic requirement for our approach is finding an accurate and reliable method to extract the threshold voltage. The characterization process is conducted on devices operating under high drain voltage. Therefore, we adopted the square root extrapolation method. This method is based on I_{ds} values measured under strong inversion and thus, is highly sensitive to mobility-degradation and series-resistance effects, which is a major restriction. As a result, the plot of $\sqrt{I_{ds}}$ vs. V_{gs} is non-linear and the extrapolated value is strongly dependent on the voltage where linear extrapolation is performed. In order reduce the temperature dependence of the V_t extraction methodology we have used the zero temperature coefficient (ZTC) bias point as the gate voltage for extrapolation.

Figure 3 presents a graphical representation of the calculation procedure of the threshold voltage for the "reference" device shown in figure 2 at room temperature.

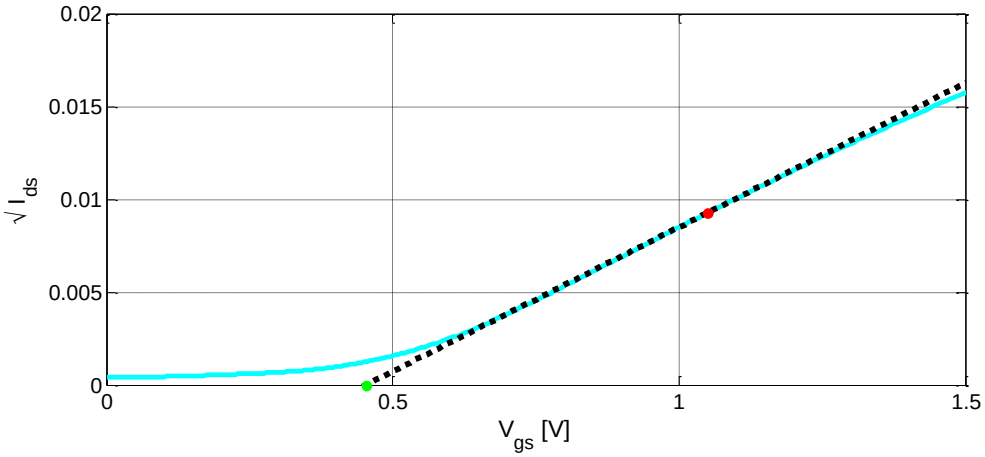


Figure 1: A graphical representation of the calculation procedure of the threshold voltage for the device shown in figure 2 at room temperature

One of the most important MOS transistor's parameters is the zero temperature coefficient (ZTC) point. This parameter has been pointed out to be of fundamental importance for the stable circuit operation over a wide temperature range. The drain current at this point is usually well defined up to approximately 200C, making it an optimal operation point for analog high- temperature CMOS ICs.

By definition, the ZTC point represents the gate bias which ensures the drain current remains constant with temperature variations, thus $\frac{dI_{ds}}{dT} = 0$. If the device current is described by an ideal square low, the ZTC bias point is given by

$$V_{gs}\big|_{ZTC} = V_t(T) + 2\mu(T,V_{gs|_{ZTC}})\left(\frac{d\mu(T,V_{gs|_{ZTC}})}{dT}\right)^{-1} \cdot \frac{dV_t}{dT}$$

It appears that gate voltage at ZTC is indeed a suitable reference point for V_t extraction. Only at this point the mobility and gate voltage temperature dependencies cancel the threshold voltage temperature dependence, resulting in a constant bias point for all measured temperatures. We used this unique independency in temperature to extract the threshold voltage and to eliminate the influence of temperature on the extrapolation method.

Accordingly, an important step of our method is to determine the ZTC voltage point of the process under study from the linear Ids (T) vs. Vgs characteristics (Figure 2). This is done by applying a MATLAB fitting procedure, which determines ZTC from the measured data as the point where the sum of the squared distances between the different temperature dependent characteristics for each Vgs is minimal (the least squares method). The extracted $V_{gs,ZTC}$ for the test devices is shown in Table 1.

Temperature region	NMOS 10/0.6 [V]	NMOS 10/10 [V]	PMOS 10/0.5 [V]	PMOS 10/10 [V]
300K -450K	1.1	1.1	1.2	1.3

Table 1: Measured $V_{gs,ZTC}$ for several test NMOS and PMOS transistors at room temperature and above

We can see from Table 1 that $V_{gs,ZTC}$ is independent of device dimensions and approximately the same for PMOS and NMOS transistors. Therefore, once the ZTC bias point of the process is determined the process threshold voltage and its temperature sensitivity (dVt/dT) can be extrapolated.

1.4 Process temperature dependent parameters

The second step or our thermal characterization process is to determine the "thermal quality" of the process under study in order to better understand the behavior of the devices and allowing future optimization of the design .This is done by extracting the process temperature dependent parameters from the I-V-T curves presented in section 1.2.

Threshold voltage

Fig. 4 presents calibration curves of Vt as function of applied temperature for several test devices manufactured using the processes under study. This calibration can be used to determine the chip local temperature. The chip by using the Vt extractor circuit (see section 2).

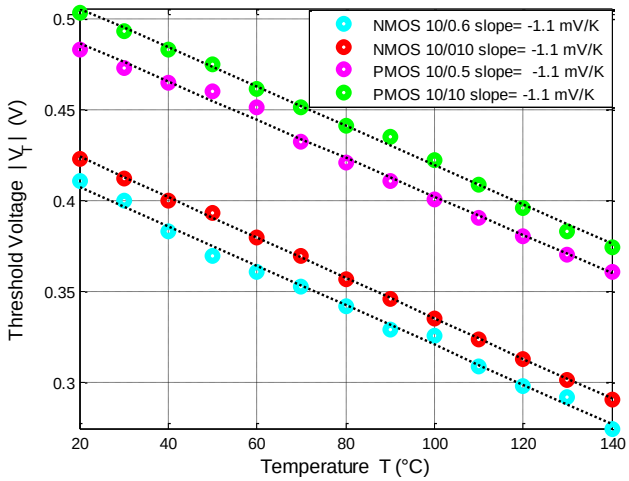


Figure 2 : Extracted threshold voltage of reference devices as a function of the chuck temperature

As shown in Figure 4 the threshold voltage decreases with temperature. This is caused due to the increase in intrinsic carrier concentration, resulting in a shift of the silicon Fermi level towards the intrinsic Fermi level. The calibration curves are linear over the entire temperature range with a constant slope of ~ -1.1 mV/K. The same slope is measured for both PMOS and NMOS devices due to the similar doping levels of the n-well and p-well.

Subthreshold swing

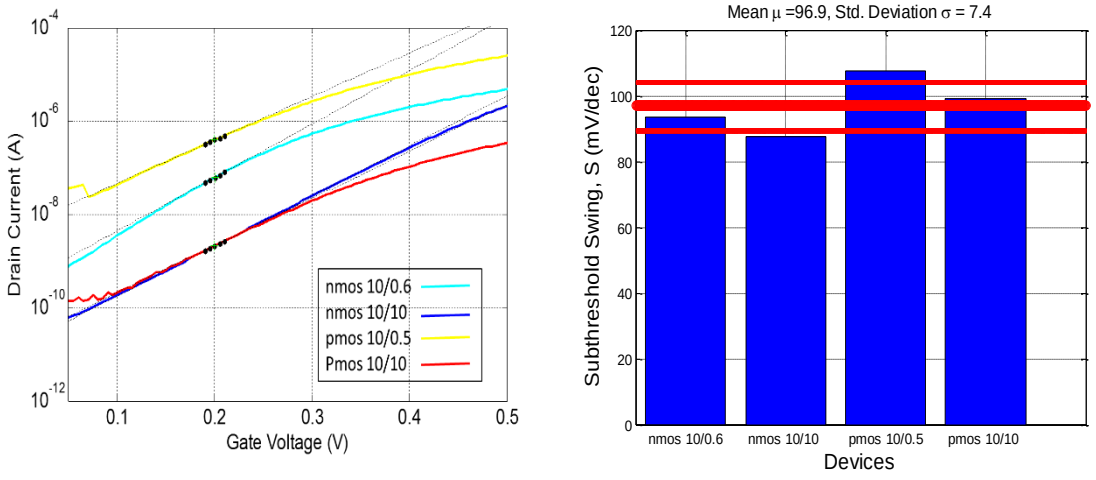
The subthreshold swing S [mV/decade] is defined in the weak inversion regime, as the variation of gate voltage necessary for producing one

decade change in the drain current. S is obtained from the reciprocal of the subthreshold slope:

$$S = \left(\frac{d \log I_{ds}}{dV_{gs}} \right)^{-1} = \ln(10) \cdot \frac{k_B T}{q} \cdot n$$

The dependence of device subthreshold current on gate voltage is exponential making the slope of the logarithm of the current in subthreshold approximately constant. Moreover, a linear dependence of S upon the transistor's temperature is theoretically expected.

Fig. 5(a) presents the calculation method of the subthreshold swing and Fig. 5(b) shows S values for the test devices at room temperature.



(a) (b)

Figure 5: (a) Graphical representation of the calculation S; (b) The calculated values of S for the test devices at room temperature

At room temperature, these results are in good agreement with measured values of state-of-the art CMOS transistors where the swing is usually in the range of 70-100 mV. The small differences between devices can be attributed to process variations.

The subthreshold swing is also used to calculate the non-ideality factor of the process under study (n). The "ideality" parameter n is a figure of merit of the processing technology and it indicates the concentration of defects in a transistor's channel; in fact, n expresses the coupling between the gate and the bulk. In state-of-the art CMOS transistors, where the swing is usually in the range of 70-100mV, n roughly corresponds to the range

1.2-1.5. The extracted values of the non-ideality factor n for the "reference" NMOS and PMOS transistors are presented in Table 2. The measured n is in good agreement with measured values of state-of-the art CMOS transistors.

Non-ideality factor	NMOS 10/0.6 [V]	NMOS 10/10 [V]	PMOS 10/0.5 [V]	PMOS 10/10 [V]
n	1.55	1.45	1.78	1.64

Table 2: Extracted non-ideality factor n at room temperature for the small test devices

2. Step 2 - Design, simulation and optimization of Vt extractor circuit

An important part of this study was to determine the optimal architecture for the Vt extractor circuit used to extract the MOSFET threshold voltage. Our requirements for an optimal Vt extractor circuit are:

- High accuracy of the extracted threshold voltage
- Low current (power) consumption
- Small chip area
- Independence of the output from the supply voltage
- Robustness (tolerance to transistor's mismatch and different processes)

The optimal architecture of a Vt extractor circuit we have chosen to implement is presented in Figure 6. We chose this design because it combines a simple low voltage Vt extracting block and feedback, to achieve independence of the output from the supply voltage, low current consumption (therefore, low power consumption), accuracy of the extracted threshold voltage toward supply voltage variations and compensation of body effect and transistor mismatch.

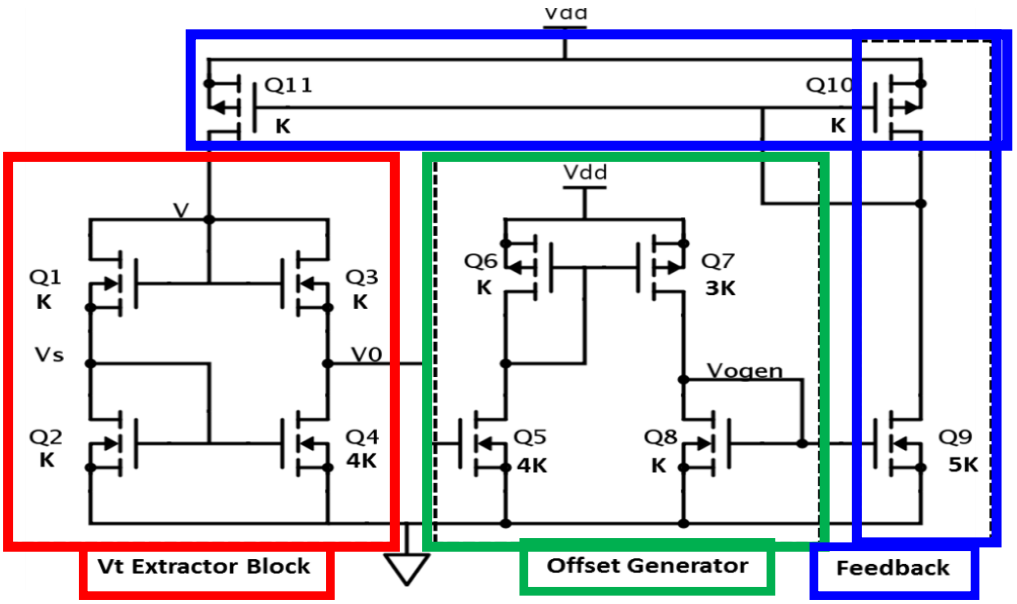


Figure 6: The tested Vt Extractor circuit

The proposed extractor consists of three blocks: 1) a simple Vt extracting block; 2) an offset generator; and 3) the current feedback loop.

2.1 Principle of operation

Vt extractor block*:*

According to the transistor size ratios described in Fig. 6 ($K_1 = K_2 = K_3 = K_4 / 4 = k$), assuming all transistors operate in saturation ($4V_{TN0} > V > 2V_{TN0}$) and $V_{BSi} = 0$, the following relations apply when second order effects are neglected:

$$I_{Di} = K_i (V_{GSi} - V_{TN0})^2$$
$$I_{D1} = I_{D2} \Rightarrow V_s = V / 2$$
$$I_{D3} = I_{D4} \Rightarrow V_o = V_{TN0}$$

The undesired sensitivity of the output to second-order effects (channel length modulation, geometric mismatches, body effect and mobility reduction) could be compensated by appropriate feedback $V=2V_o$. However, to ensure all transistors operate in strong inversion and validating the current equations the voltage V needs to be greater than 2Vo by a magnitude of several $U_T=kT/q$.

Offset generator:

The purpose of the offset generator block is to take V_o as input and add an offset of several U_T on this value. Considering that V_o is very near to V_{TN0}, then V_{GS5}-V_{TN0} < 3kT/q, subthreshold drain current equations are more realistic to use in order to calculate this block's output - V_{OGEN}.

$$aI_{D5} = I_{D8} \Rightarrow aI_0 \exp(V_o / U_T) = \frac{I_0}{b} \exp(V_{OGEN} / U_T)$$
$$\Rightarrow V_{OGEN} = V_o + U_T \ln(ab)$$

where $a = K_7/K_6$ and $b = K_5/K_8$. Hence, V_{OGEN} includes the necessary offset from V_o. The accurate value of V_{OGEN} is not an issue due to the flexibility in the feedback value

$$V = 2V_o + \frac{nk_BT}{q}; \quad 2 < n < 8$$

Feedback block:

Transistors Q9 and Q4 in Fig. 6 are 5 and 4 times wider than Q2, respectively, and due to the current mirror formed by Q10 and Q11, V_{OGEN} is fed back to V_S as follows:

$$I_{D9} = I_{D2} + I_{D4} \Rightarrow V_s = V_{OGEN} \Rightarrow$$
$$V = 2V_o + 2U_T \ln(ab)$$

So, the proposed feedback implements the needed result for the circuit proper operation.

2.2 Simulation Results

The circuit was implemented in Cadence Virtuoso Schematics using CMOS 0.18µm process. The performed simulations compared the accuracy of the circuit output with the nominal threshold voltage calculated by SPICE MOSFET models at different temperatures. For supply voltage of 1.8V, an optimal performance was achieved for thin oxide transistors with body contact shortened to the source (VBS =0 V) and W/L= 0.52/1 µm/ µm. These dimensions also gave the optima in aspects of chip area and circuit power consumption.

Fig. 7 presents the results of the circuit DC simulations at room temperature (~300K). As seen from this simulation the circuit operates at low current (~ 5.5µA) resulting in very low power consumption of the

circuit (~20µWatt). This allows to place the circuit is several locations on the chip without consuming much of the chip resources.

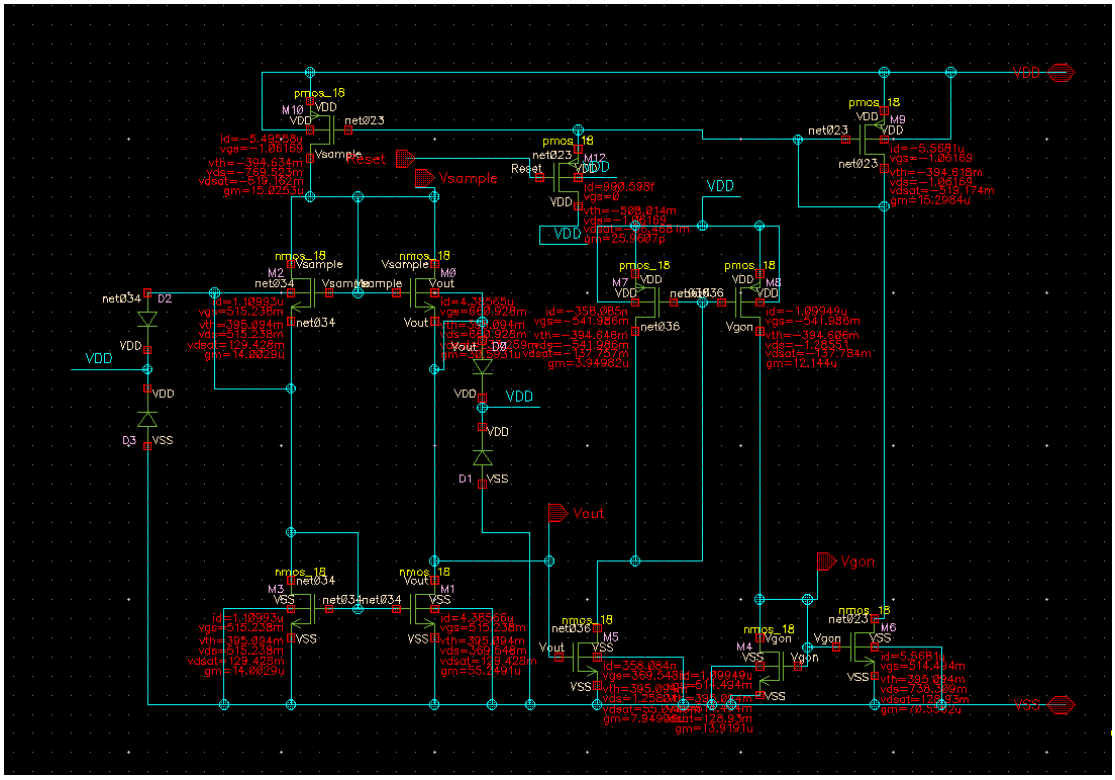
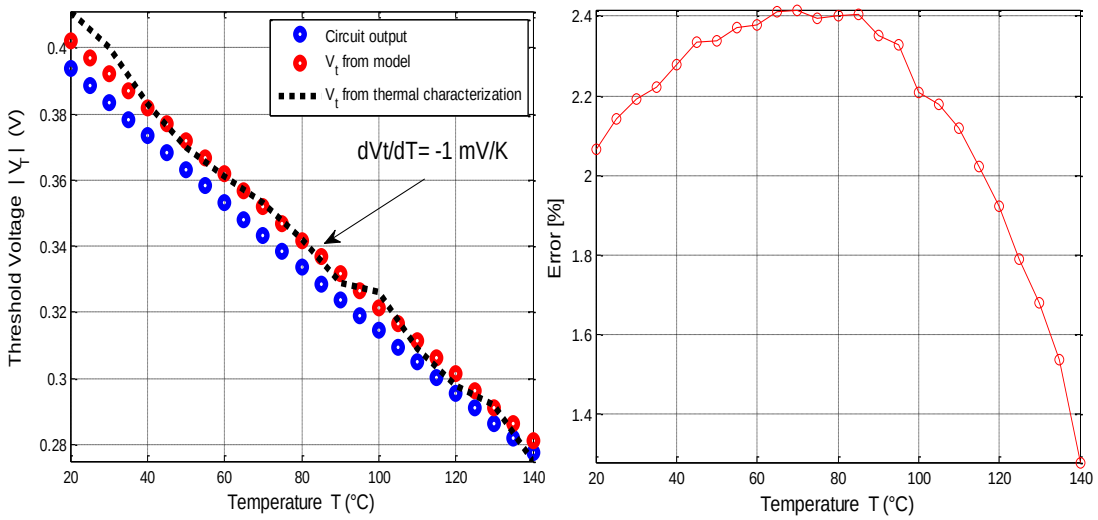


Figure 7: DC simulation results at 300K

Fig. 8(a) presents a comparison between the threshold voltage at the output of the Vt extractor circuit, Vt as extrapolated from the process characterization (section 1.4) and the threshold voltage calculated according to Spice model. Fig. 8(b) shows the circuit error as a function of temperature.



(a) (b)

Figure 8: (a) The output of the Vt extractor circuit and Vt according to Spice model as a function of temperature (b) the circuit "sampling" error

As seen from Fig. 8(a) the circuit output is linear with temperature, as expected, with a slope of ~-1mV/K. The output is highly accurate and produces an error of ~2% from the Spice nominal value which corresponded to an error of ~6K for temperature reading. This error can be minimized by increasing the transistors sizes and power consumption . The implemented Vt extractor circuit has a small error (2%) under nominal conditions, and reasonable error under extreme corners. The extracted value has a linear dependence on temperature, making the circuit applicable for temperature detection. A low dependence on process variation and supply voltage variation makes the design robust and process-independent.

Fig. 9 presents the circuit transient response to voltage changes due to circuit reset or local temperature changes. The circuit time constant is ~40 [nsec] making this architecture suitable for temperature measuring on-line because changes caused by temperature has much higher time constants.

Fig. 10 shows the circuit layout as it was implemented using Tower Jazz 0.18µm CMOS process. The total circuit area is ~ 44 x 35 [µm x µm].

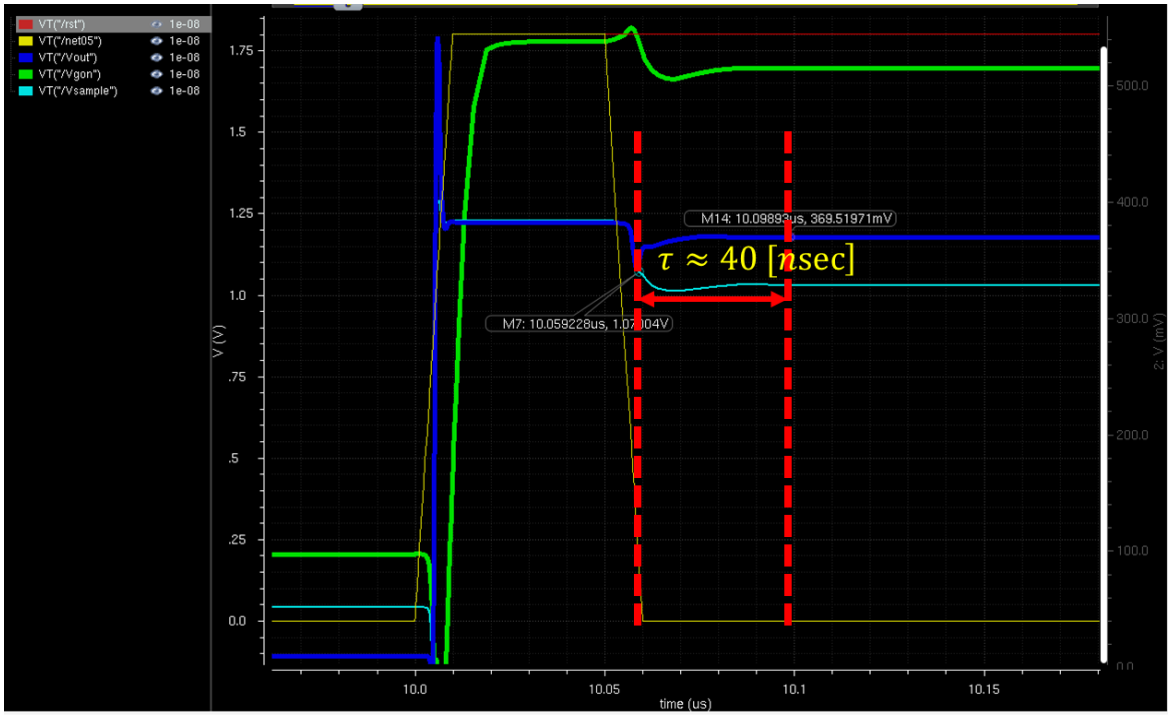


Figure 9: transient simulation results at 300K

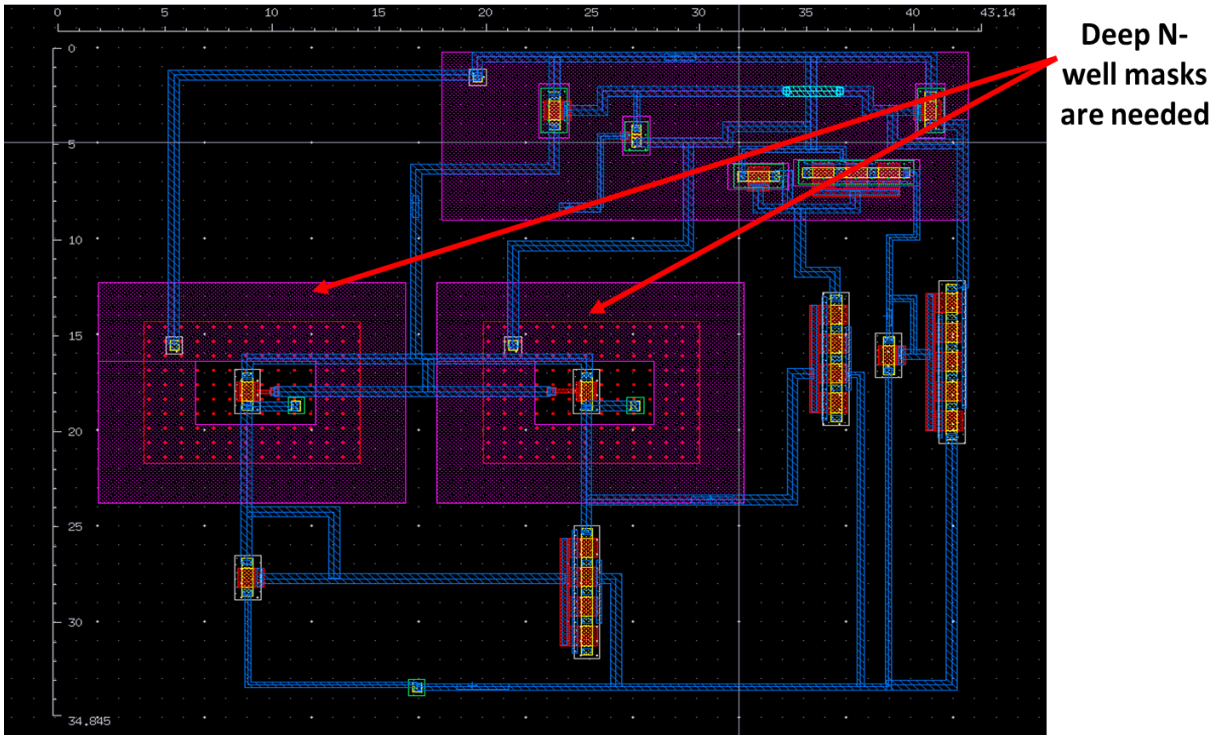


Figure 10: circuit layout

In order to test the circuit “robustness” to different technologies the circuit was also implemented using CMOS-SOI 1µm process. For supply voltage of 5V, an optimal performance was achieved for transistors with body contact shortened to the source (VBS =0 V) and W/L= 4/9 µm/ µm. The results of the different circuit simulation are presented if Figures 11-14.

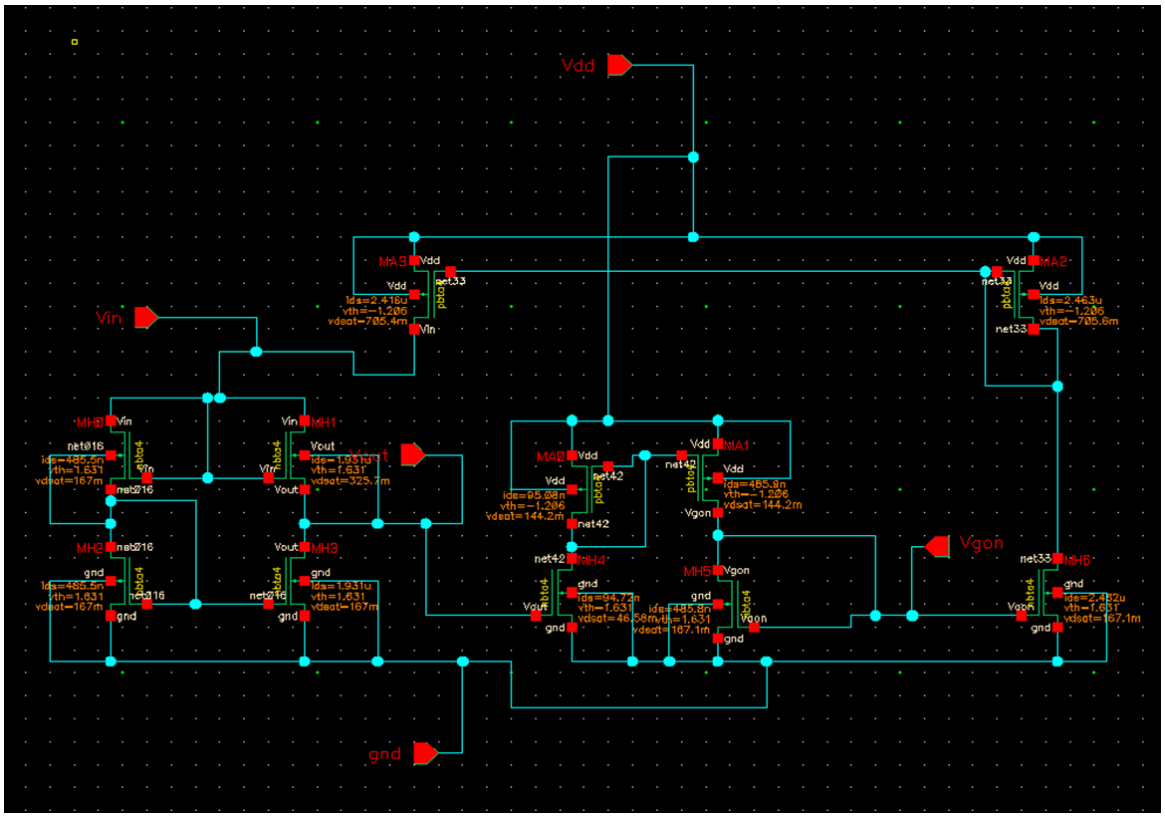
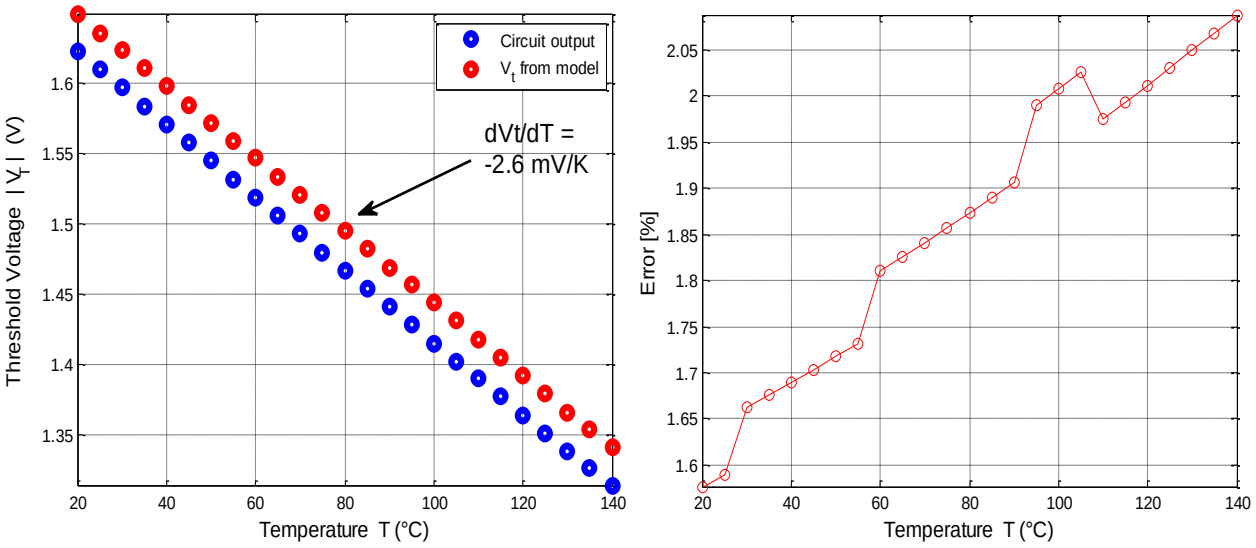


Figure 11: DC simulation results at 300K



(a) (b)

Figure 12: (a) The output of the Vt extractor circuit and Vt according to Spice model as a function of temperature (b) the circuit "sampling" error

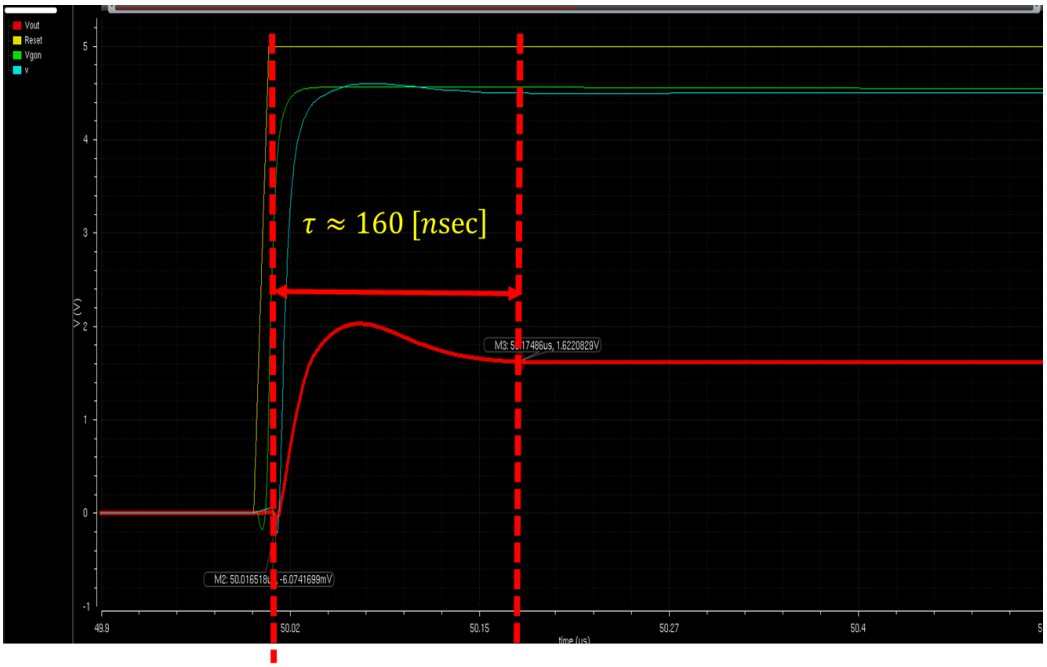


Figure 13: transient simulation results at 300K

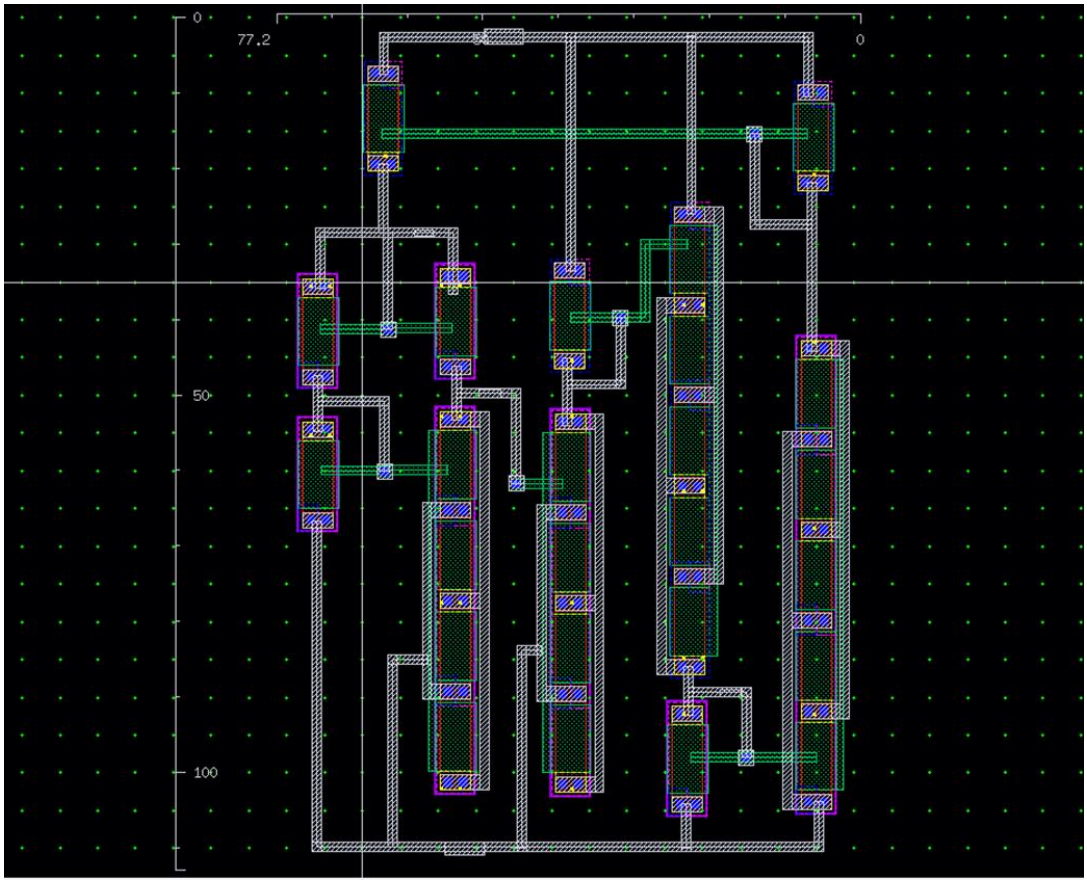


Figure 14: circuit layout

In this technology the expected power consumption is higher in compare to the CMOS based circuit (~24µWatt). However, the sensitivity to temperature changes (dVt/dT) is higher resulting in slightly smaller error in reading temperature changes (~5K). The total circuit area is ~ 80x100 [µm x µm] and the circuit response time is longer (~160 [nsec]).

3. Step 3 - Design of a CMOS SPAD imager with integrated temperature sensors

In order to test the success of the new thermal management paradigm several single temperature sensors as well as Vt extractor circuits were integrated into a matrix of CMOS SPAD sensors fabricated using TowerJazz 0.18µm CMOS process. We have chosen challenging 3D CMOS SPAD imager/optical radar to demonstrate how the TOP-M thermal management approach improves performance. Temperature control improve the imager performance due to

- Heating during avalanche impacts several of the sensor's parameters, such as time uncertainty (time jitter), thermal generation rate (DCR false count) and photon detection efficiency (PDE). Temperature control diminishes the correlation between successive avalanche pulses, hence ensuring stable performance over time .

- The on chip circuitry (such as digital counters) is highly susceptible to temperature due its dependence on charge carrier mobility, which decreases as temperature rises. Temperature control will improve the digital circuit resolution by preventing the mobility degradation.

- Preventing the temperature rise will also decrease thermal noise.

We have designed a 64 x 64 matrix of SPAD pixels with integrated readout. Each pixel contains:

- A CMOS diode biased above breakdown voltage to sense each photon causing avalanche in the device (~20V)
- A quenching resistor to recognizes avalanche and applies negative feedback upon SPAD biasing,
- A counter to measuring signal intensity
- Buffers to store the data

There are also integrated readout circuits on chip enabling simultaneous readout of 8 pixels. We have scattered single temperature sensors as well as several Vt extractors near the matrix to enable so measure the chip temperature during the imager operation. Fig. 15 and 16 present a top view of the entire chip and of the integrated Vt extractor circuit. This chip was designs for long wavelength range ~769 [nm].

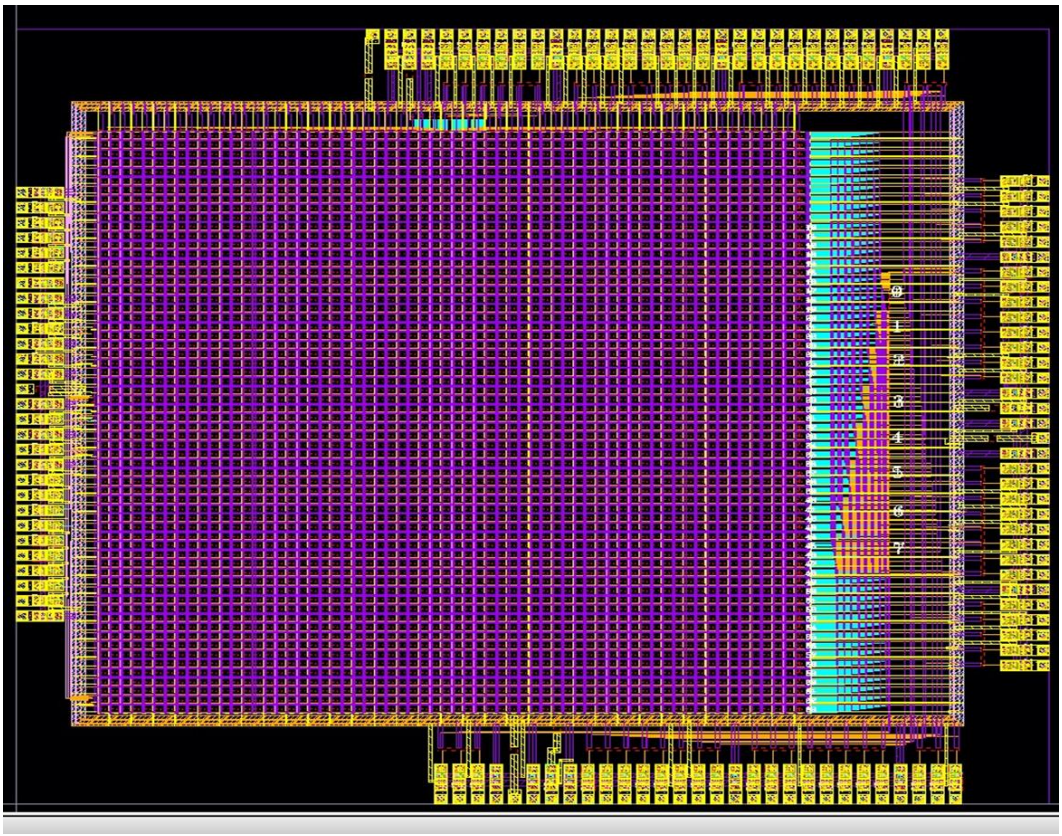


Figure 15: layout of 64 x 64 SPAD Array (top view)

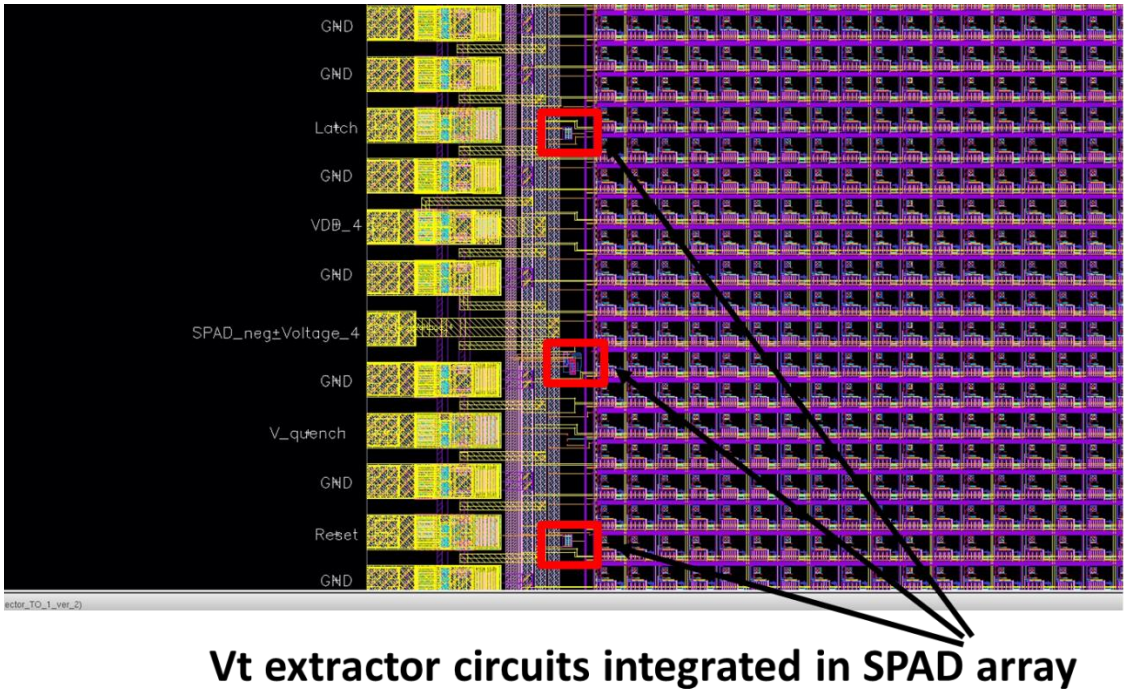


Figure 16: layout of 64 x 64 SPAD Array with Integrated Temperature Sensors

In order to optimize the SPAD performance a second chip was design for wavelength of ~589 [nm]. The first design was based on an array of N+P diodes and the second design is based on an array of P+N diodes. Fig. 17 presents the SPAD and pixel structure of the first design. Fig. 18 presents the SPAD structure, pixel structure and layout of the second design. In the second design we also scattered single temperature sensors as well as several Vt extractors near the matrix to enable temperature measurement during the chip operation.

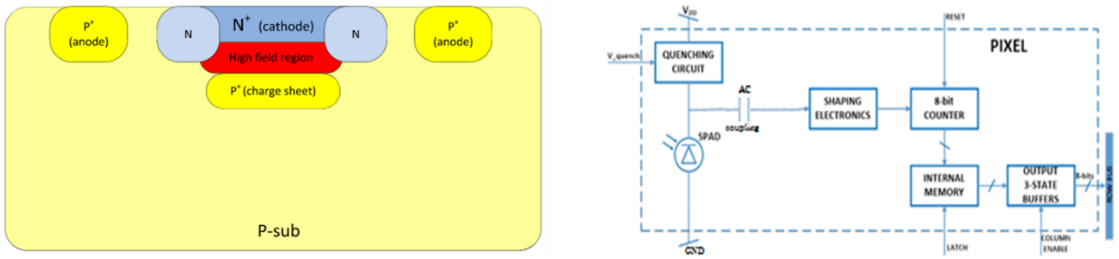


Figure 17: The SPAD and pixel structure of the first design

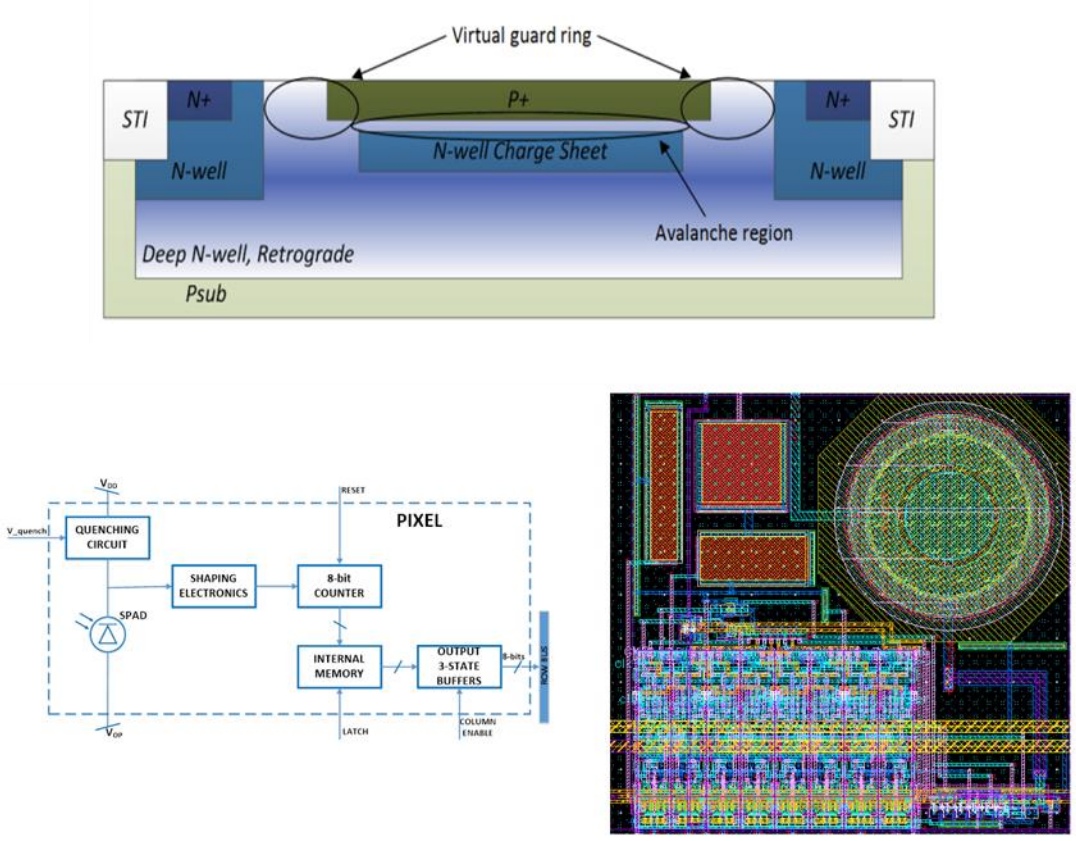


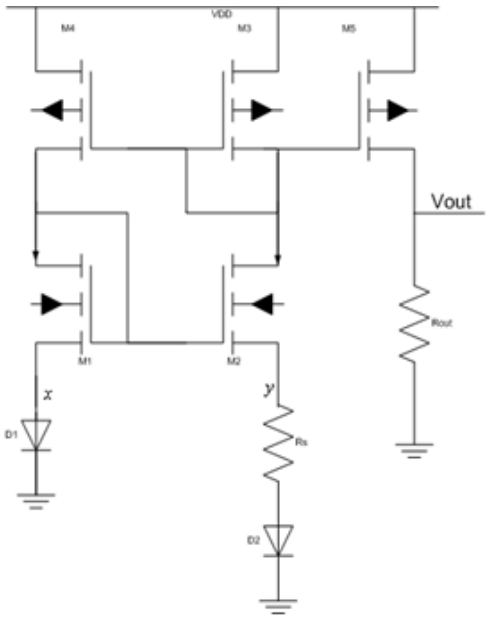
Figure 18: The SPAD structure, pixel structure and layout of the second design

The first design tape-out was in September 2016; however, due to problems in the fabrication process the first design will come back for characterization in August 2017. The second design tape-out was in April 2017 and will come back from fabrication in August 2017. Results of the characterization and performance of both design will be presented in THermPED meeting t in COMCAS 2017.

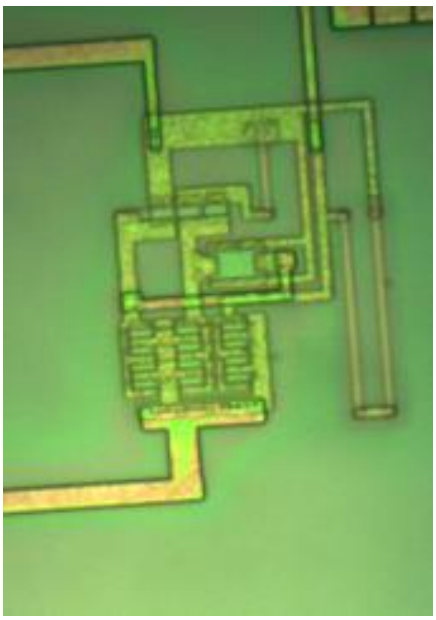
4. Step 4 – Real time measurements of the integrated temperature sensors

The optimal Vt extractor circuit design (presented in section 2) was manufactured using CMOS 0.18µm process and CMOS-SOI 1µm process and characterized in the measurement setup presented in section 1.1. The circuit thermal performance was evaluated by comparing it to standard temperature sensing circuit used in the semiconductor industry, i.e. *Proportional to Absolute Temperature* (PTAT) circuit manufactured in the same technologies. These circuits are compared in terms of sensitivity, linearity, accuracy, calibration needs, area and possibility of measure temperature during the operation of the chip (on-line).

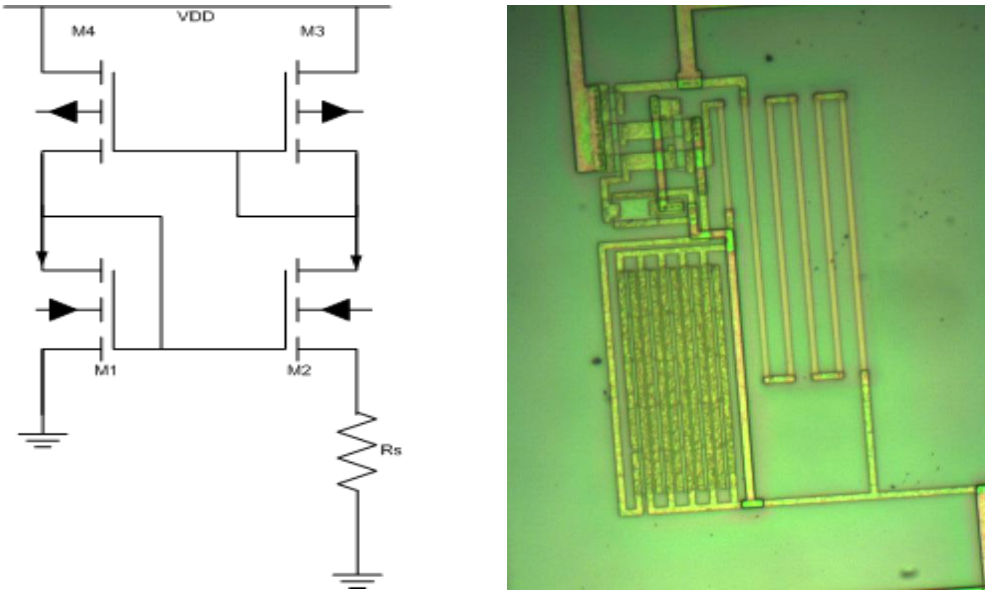
PTAT circuits are widely used to generate temperature independent bias currents and as temperature sensor in temperature measurement systems. They are also employed in-bandgap reference circuits which are commonly used to generate temperature independent bias voltages and as reference in measurement systems. A PTAT circuit which uses the exponential dependence of diodes upon temperature to generate output voltage proportional to the chip local temperature is shown in Fig. 19(a)-(b). A PTAT circuit which is based on resistors to generate the output voltage is shown in Fig. 19(c)-(d). Fig. 20(a)-(b) shows photographs of the Vt extractor circuits manufactured in both technologies.



(a)

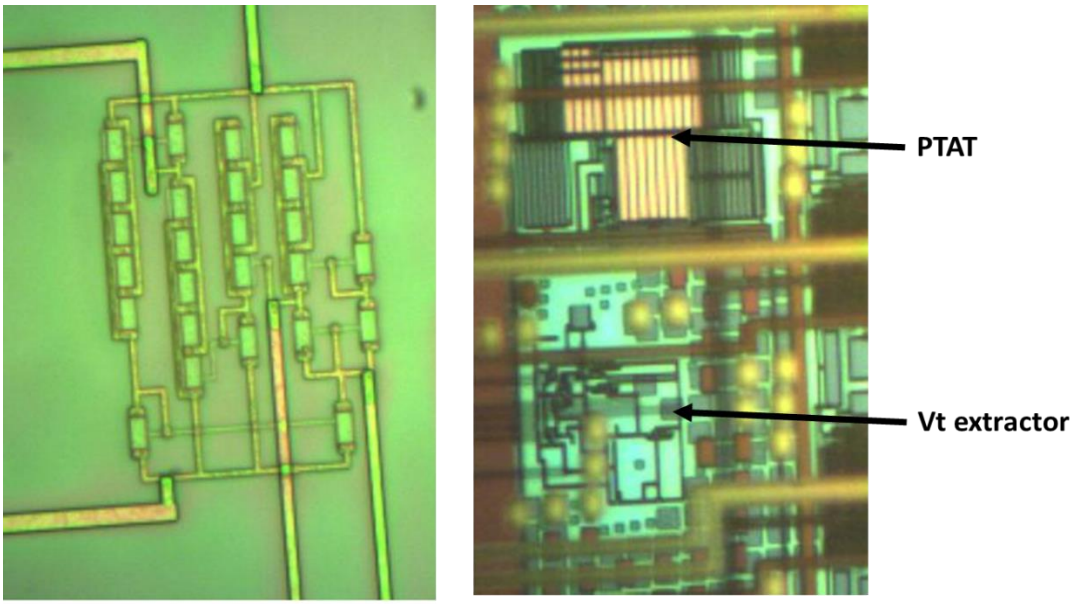


(b)



(c) (d)

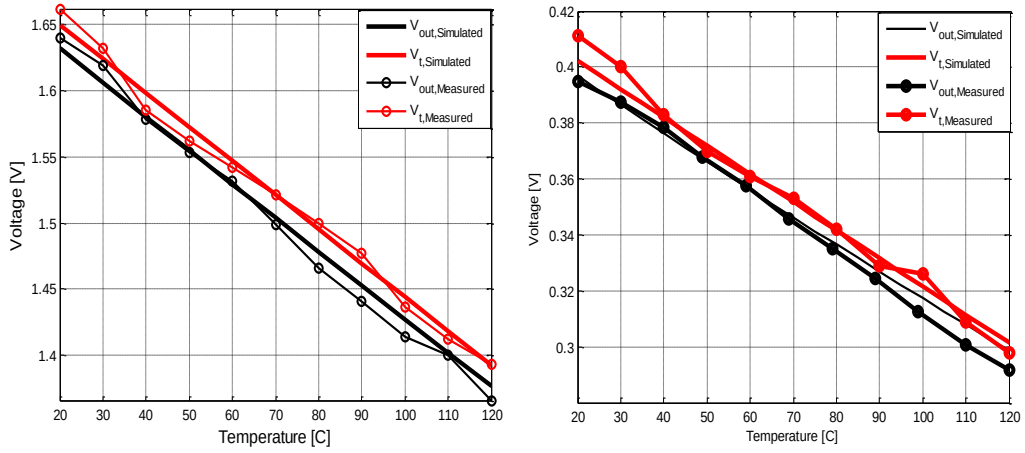
Figure 19: Scheme and photograph of a PTAT circuit implemented in CMOS technology based on diodes (a)-(b) and resistors (c)-(d).



(a) (b)

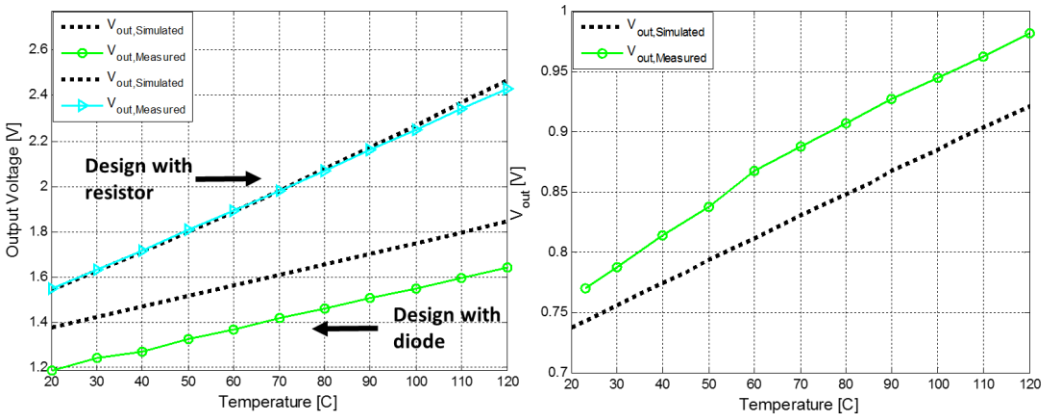
Figure 20: Photograph of a Vt extractor circuit implemented in (a) CMOS-SOI 1µm process and in (b) CMOS 0.18µm process.

The output voltage of all circuits was measured as a function of temperature at the range of 20C to 120C and is presented in Fig. 21 and Fig. 22. Fig. 22(a) presents the output voltage of two PTAT designs (resistors based and diode based) manufactured using CMOS-SOI 1µm process. Fig. 22(b) presents the output voltage of PTAT circuit which is based on resistors manufactured using CMOS 0.18µm process. In order to evaluate each circuit accuracy and performance the measured output voltage was compared to the simulated voltage extracted from spice electrical simulations and the nominal threshold voltage calculated by SPECTRE MOSFET models, both presented in Fig.21 and Fig.22.



(a) (b)

Figure 21: Measured and simulated output voltage of the Vt extractor circuit implemented in (a) CMOS-SOI 1µm process and in (b) CMOS 0.18µm process. The threshold voltage extracted during the thermal characterization process and the threshold voltage extracted from electrical Spice simulations are presented as well for both technologies.



(a) (b)

Figure 22: Measured and simulated output voltage of the PTAT circuits implemented in (a) CMOS-SOI 1µm process and in (b) CMOS 0.18µm process.

As seen from Fig. 21(a)-(b) the Vt extractor circuit output is linear with temperature, highly accurate and produces a maximum error of 1% and 3% from the simulated and measured Vt values for the CMOS-SOI and CMOS processes respectively. The output voltage sensitivities to temperature variations are -2.7 mV/C and -1.1 mV/C for the CMOS-SOI and CMOS processes respectively. These sensitives are similar to the sensitivities calculated during the processes' thermal characterization; hence the accuracy of the temperature reading is determined by the precision of the threshold voltage extraction technique.

The error in the temperature extraction (ΔT) has been obtained by calculating the difference, ΔV, between the measured output voltage of the extractor and measured Vt vs. T, which is then converted in temperature using the sensitivity (dVt/dT). Accordingly, the accuracy is estimated by

$$\Delta T = \frac{\left| V_{out\ extractor,measured} - V_{t,measured} \right|}{dV_t / dT}$$

Errors of ±3C and ±4C have been calculated for the CMOS-SOI and CMOS processes respectively.

As Fig. 22(a)-(b) shows all the PTAT circuit exhibit linear dependence in temperature in the entire temperature range; however, there is an offset between the measured and simulated results in circuits manufactured by both technologies which decreases the accuracy of the sensor and requires circuit calibration. The output voltage sensitivities to temperature variations are 4.4mV/C, 9mV/C and 2.2mV/C for the diode /resistor based PTAT circuits manufactured in CMOS-SOI process and PTAT manufactured in CMOS processes respectively. The accuracy of the PTAT circuit in CMOS –SOI process is ±7C when using the exponential dependence of diodes and ±2C when using the temperature dependence of polysilicon resistors. The accuracy of the PTAT circuit in CMOS process is ±7C.

Both circuits' exhibit linearity over the entire temperature range. Although the PTAT circuits have higher sensitivity their accuracy is highly

dependent on process variations; hence in order to improve the sensor's precision carful calibration is needed for each process. In addition, the power consumption during operation and the needed chip area for the PTAT circuits are higher than those need for the Vt extractor circuit. The implemented Vt extractor circuit has a small error (1%) under nominal conditions, and reasonable error under extreme corners. A low dependence on process variation and supply voltage variation makes the design robust, process-independent and no calibration is needed; hence, the Vt extractor circuit is a better solution for local temperature sensing when using CMOS and CMOS-SOI technologies. In Table 3, the main properties of both circuits are summarized and compared.

Technology [nm]	**Sensor**	**Area [μm²]**	**Maximum Sensitivity [mV/C]**	**Accuracy [C]**	**Power Consumption (T=20C) [μW]**	**Calibration**
CMOS 0.18μm	Vt extractor	44 x 35	-1.1	4	19	Not needed
	PTAT	42x55	2.2	7	72	Needed
CMOS-SOI 1μm	Vt extractor	80x100	-2.7	3	40	Not needed
	PTAT with diodes	178x150	4.4	7	46	Needed
	PTAT with resistors	85x100	9	2	260	Needed

Table 3. The sensors' main properties

5. Step 5 – Integrating micro-channels for cooling

The last step of our project is to integrate our electrical design with micro-channels and use the on-line temperature sensors to close feedback loops. We can apply better liquid cooling by measuring the chip temperature during the cooling process and compensate accordingly, for example by reducing the pumping speed. We based the liquid cooling on **InterCooling** approach were single temperature sensors were attached to

a substrate containing cooling channels which were manufactured separately. The chip temperature rise was controlled by an integrated resistor and the cooling DI water were pumped from front side to backside with external pump at a rate appropriate for reducing the chip temperature. The chip “stackup”, which consists of the electrical design and the cooling channels at the backside is presented in Fig. 23.

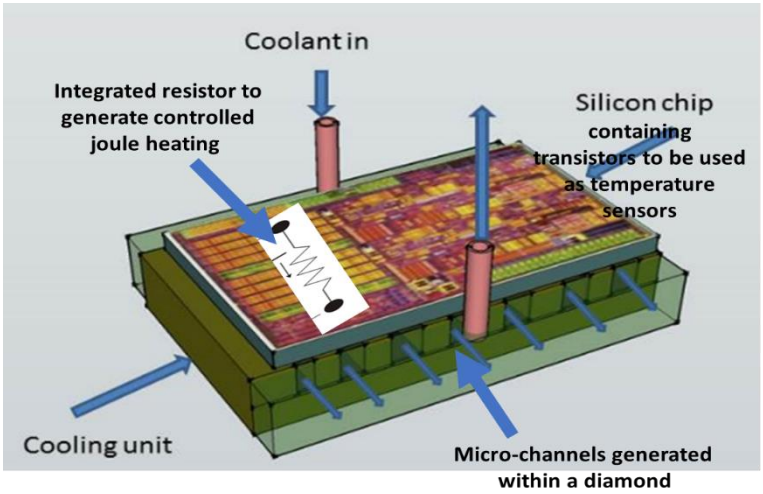
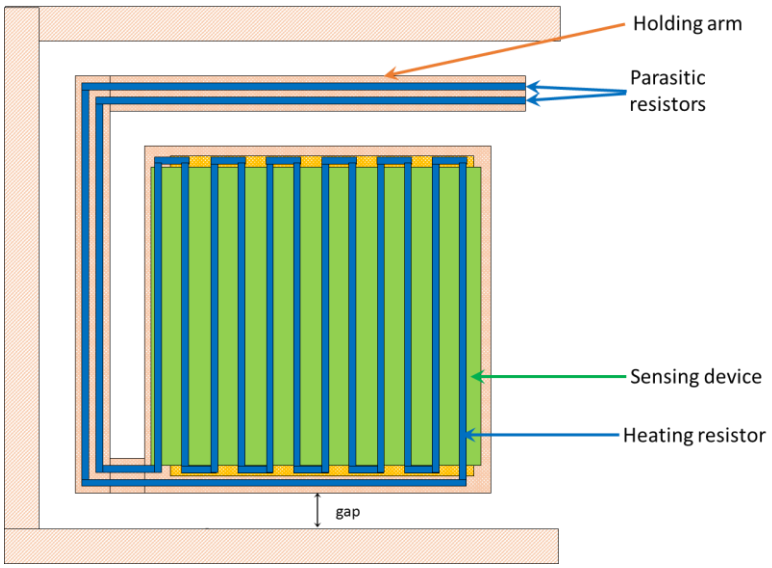


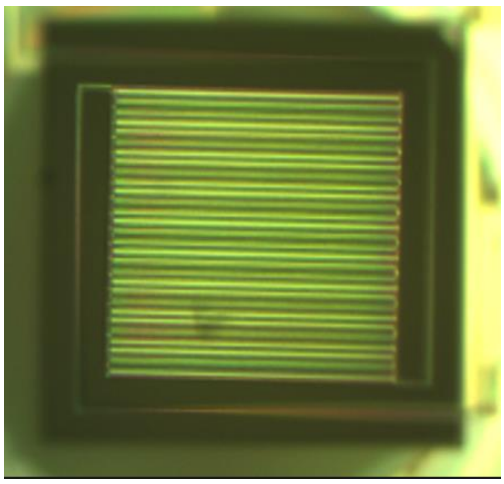
Figure 23: Chip “stackup”

5.1 Chip “stackup" top level (electrical design)

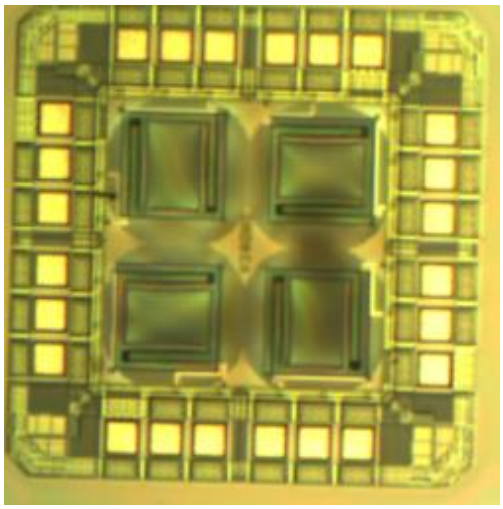
The chip “stackup” top level consists of several single integrated temperature sensors (MOSFET transistors of different sizes) and an integrated polysilicon resistor which was used as a controlled source for joule heating. This chip was manufactured using CMOS-SOI 1µm process. A scheme and photographs of the electrical design is presented in Fig. 24.



(a)



(b)

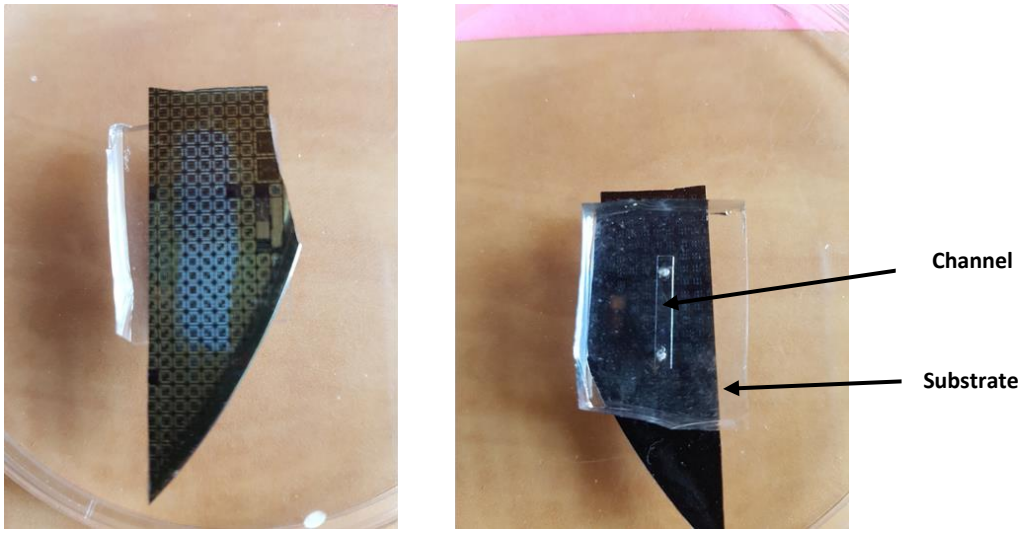


(c)

Figure 24: Scheme and photographs of the electrical design

5.2 Manufacturing of micro-channels ("stackup" backside)

The micro-channels were fabricated in cooperation with Dr. Netanel Korin team from the Biomedical Engineering department at the Technion and are presented in Fig.25 and Fig. 26. Microfluidic devices were fabricated using conventional soft-lithography techniques. A computer-aided design program was used to produce masks for photolithography. Using convectional SU-8 photolithography a master was manufactured for molding. By utilizing the master, Polydimethylsiloxane (PDMS, Sylgard 184 Silicone Elastomer Kit) replica of the channels were produced and irreversibly sealed against a glass slide. The dimensions of the micro-channels used in this study are: 80 µm (height) x 2 mm (width) x 2.2 cm (long). For flow experiments a programmable syringe pump (Harvaerd apparatus) was used to perfuse water at a defined constant flow rate (0.1-1 ml/min).



(a) (b)

Figure 25: Photographs of the chip "stackup" (a) top level – integrated temperature sensors and resistors (b) backside - micro-channel

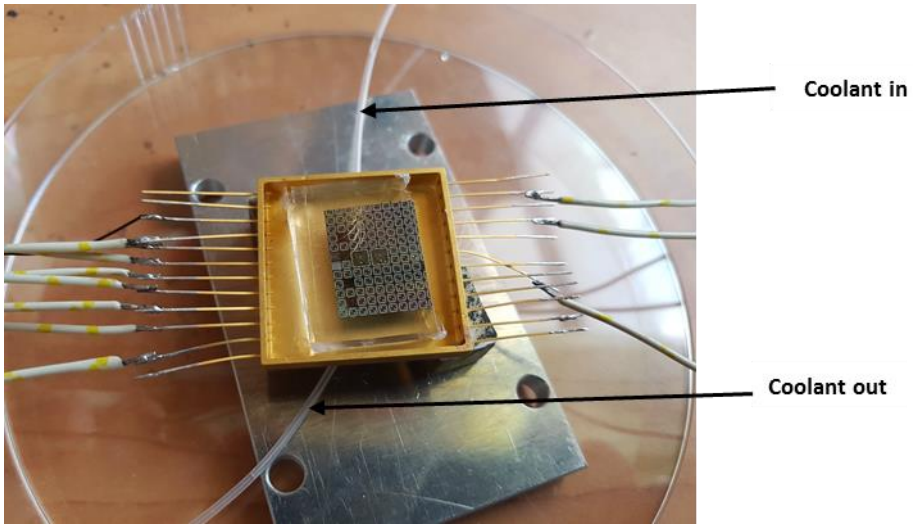


Figure 26: Packaging for electrical and thermal characterization

5.3 Characterization of micro-channels

The cooling of the micro-channels was evaluated in the following process:

1. The chip temperature was increased by applying voltage to the polysilicon resistor causing joule heating. The chip temperature increase (ΔT) as a function of the applied voltage on the resistor (V_R) was calculated using elctro-thermal simulations in COMSOL Multiphysics (see Fig. 27).
2. The simulated temperature rise was compared to the temperature measured by the integrated temperature sensor. The transistor I-V curves were measured and by extracting the threshold voltage and

comparing to the process thermal calibration curves (Vt vs. T in Fig. 12(a)) the chip temperature was evaluated (see Fig. 28 and Fig. 29).

3. Then while applying voltage to the polysilicon resistor the cooling DI water were circulated with an external pump through the micro-channels.
4. The transistor I-V curve was measured again during the continuous cooling and by extracting the threshold voltage the chip new temperature was evaluated (see Fig. 30 and Fig. 31).

This process was repeated for different resistor voltages and the results are shown in Fig.28 till Fig. 31.

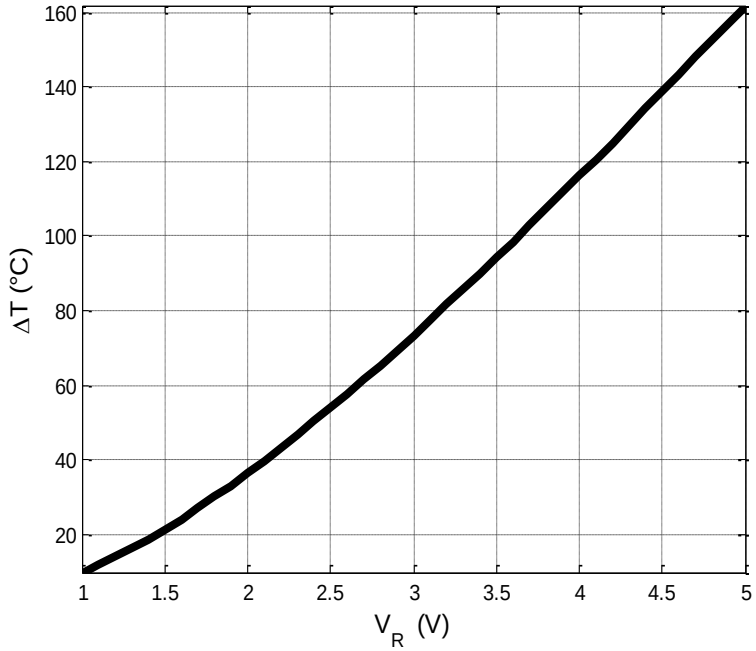
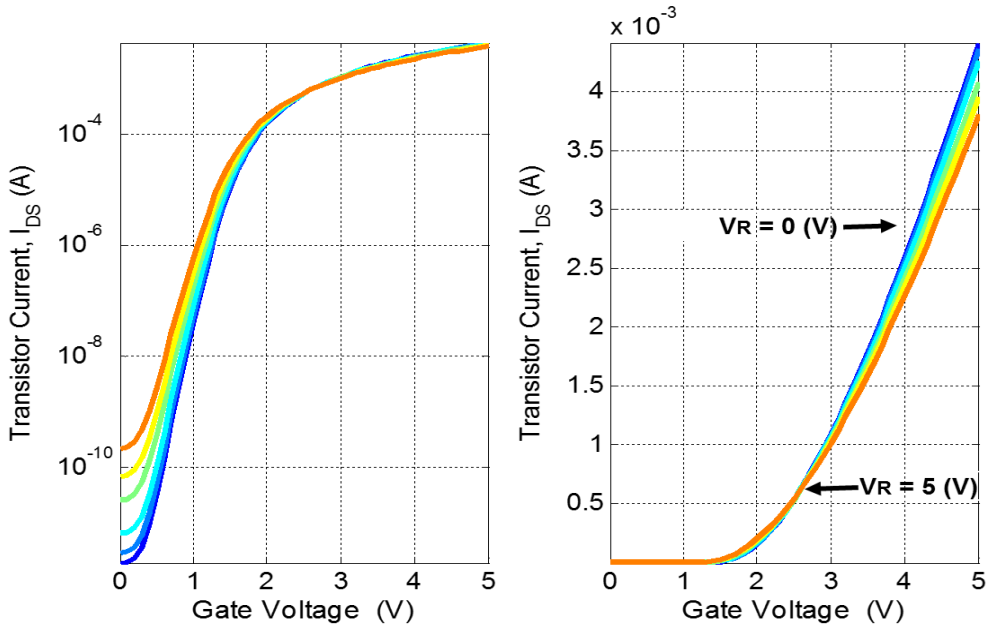


Figure 27: The chip temperature increase (ΔT) as a function of the applied voltage on the resistor (VR) based on simulation in COMSOL Multiphysics



(a) (b)

Figure 28: Example of I-V sweep performed for different resistor voltages for an NMOS transistor with W/L= 178x228µm/10.6 µm: on a semi-log scale (a) and on a linear scale (b).

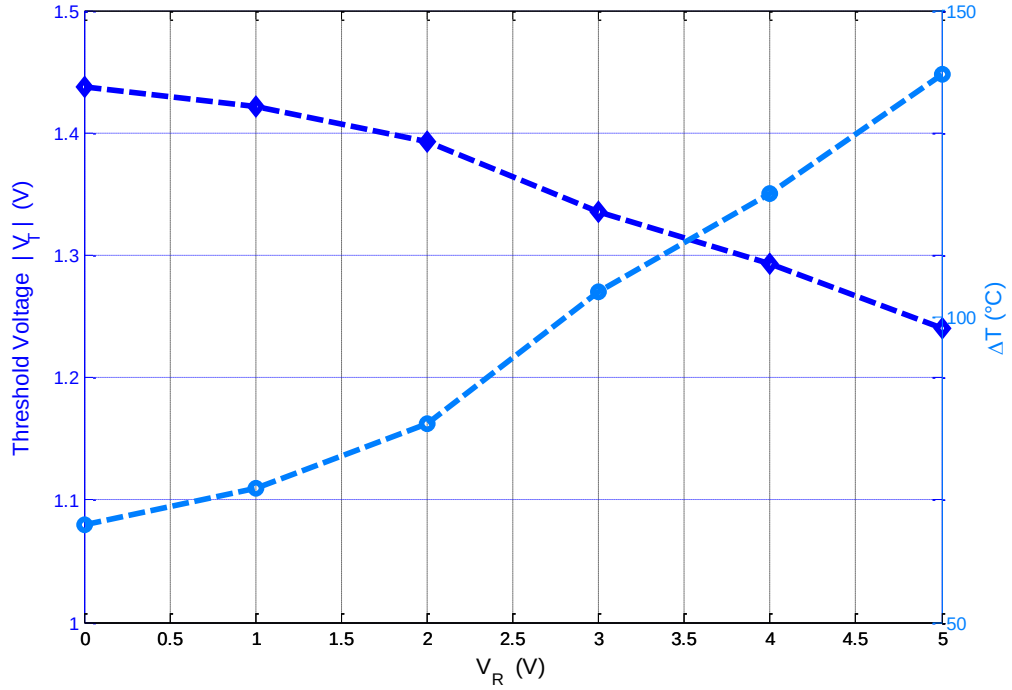
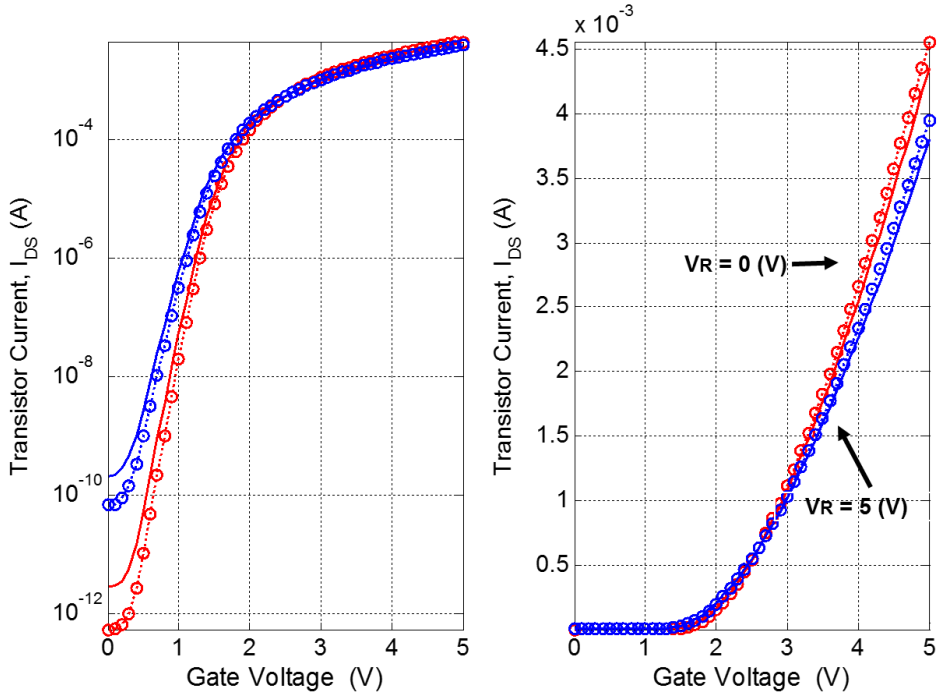


Figure 29: Extracted threshold voltage and the chip temperature rise determined from the process thermal characterization for different resistor voltages.

As can be seen from Fig.28 and Fig. 29 there is substantial increase in the chip temperature, even without applying voltage to the polysilicon resistor (V_R = 0V), due to severe self-heating (ΔT=86C). The temperature rise calculated by extracting V_t and comparing it to the process thermal

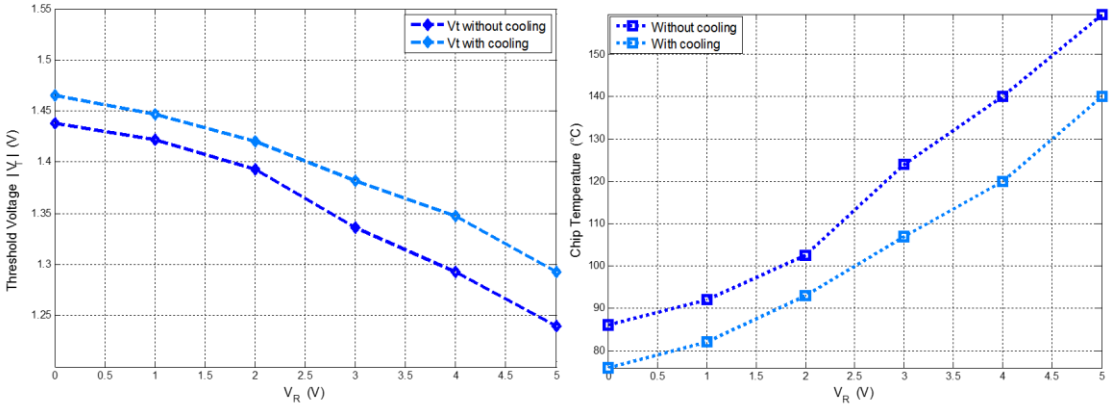
characterization is lower from the one simulated in COMSOL (Fig. 27) due to heat conductance through the substrate.



(a) (b)

Figure 30: Example of I-V sweep performed while **cooling the chip** for an NMOS transistor with W/L= 178x228µm/10.6 µm: on a semi-log scale (a) and on a linear scale (b). The dashed I-V curves were measured with cooling. The solid curves were measured without cooling.

.



(a) (b)

Figure 31: The extracted threshold voltage (a) and the calculated chip temperature (b) with and without cooling.

Fig. 30 present an example of two I-V curves measured while cooling the chip (V_R=0V and V_R=5V) by circulating DI water through the micro-channels. Fig. 31 shows the extracted threshold voltage and consequently

the temperature increase with and without cooling as a function of the applied voltage on the resistor. We can see that by one micro-channel we can decrease the chip temperature by maximum of 20C depending on the pumping speed. The results for all applied voltages are summarized in Table 4.

Cooling		VR = 0 [V]	VR = 1 [V]	VR = 2 [V]	VR = 3 [V]	VR = 4 [V]	VR = 5 [V]
Without Cooling	Vt [V]	1.438	1.422	1.393	1.336	1.293	1.24
	Temperature [C]	86	92	102	124	140	159
With Cooling	Vt [V]	1.465	1.447	1.42	1.382	1.347	1.293
	Temperature [C]	76	82	93	107	120	140

Table 2. Results Summary

6. Summary

In this report we have presented the results achieved in the TOP-M project:

-Thermal characterization of Tower Jazz CMOS 0.18 µm process, mainly the threshold voltage dependence upon temperature (Vt).
-Design, simulation and optimization of Vt extractor circuits for on-line temperature reading and monitoring .
- Design of an array of 64*64 CMOS SPAD sensors .
-Integration of single temperature sensors and Vt-extractor circuits into the SPAD chip.
- Thermal characterization of the first SPAD chip with integrated temperature sensors.
- Study of inter-cooling mechanisms using micro- cooling channels.
Integrating single temperature sensors with the cooling channels and measuring the maximum temperature decrease.

The project's main objectives, completed tasks, significant achievements and suggestions for future cooperation are summarized in the following quad char:

Main Objectives :

- Establishing New Thermal Management Paradigm for high-power mixed signal chips.
- Include thermal effects at the first stage of VLSI design
- Include thermal cooling technology as part of die architecture and packaging.

Significant Achievements:

- 4 Celsius resolution in temperature measurement.
- Accuracy of 3% of the absolute threshold value.
- Maximum temperature decrease due to cooling - 20C.
- Vt extractor chip area – 44x35 [µm x µm].

Completed Tasks:

- Thermal characterization of Tower Jazz CMOS 0.18 µm process.
- Design, simulation, optimization and thermal characterization of Vt extractor circuits and single integrated temperature sensors.
- Integrating of single temperature sensors with the cooling micro-channels.
- Integration and analysis of the micro-channel cooling with COMSOL and cadence simulations.

Future Cooperation:

- **Thermal TMOS IR** uncooled Sensors
- **Thermal TeraMOS THz** uncooled Sensors
- **Thermal GMOS Gas** Sensors
- **CMOS SPAD** camera for detection of gun flash
- **CMOS SiPM** Photomultiplier for **LIDARs**