

REAL TIME THERMAL MANAGEMENT FOR HIGH PERFORMANCE VLSI CHIPS USING DYNAMIC FREQUENCY SCALING AND COOLING TOPOLOGIES

Sachin Magadum^{1,*} and Hansraj Guhilot²

Received: March 12, 2020; Revised: May 07, 2020; Accepted: May 26, 2020

Abstract

The device miniaturization technology made possible to integrate large density of transistors in a chip to produce high computation facility. At the same time, it demands an efficient thermal management technique to handle enormous power dissipation during high performance of the chip. An active thermal management technique proposed here includes run time fully digital onchip temperature sensors, Dynamic Frequency Scaling (DFS) and active cooling topologies to enhance cooling effect of VLSI. The high resolution on chip temperature detector exhibit the temperature resolution of 0.0312°C, time resolution of 106 ps and measurement range from 0°C to 85°C with 21 sensing elements. The proposed thermal management techniques reduce the power dissipation by 30%. To enhance the thermal management capability further, there are numerous cooling methodologies developed and verified at 40 W of die power. With these topologies, the temperature of die at the power of 0 to 40 W lies between 0.5°C and 69°C. This is the worst case range amongst three cooling methodologies. The proposed work also nurtures the features of XILINX plan ahead tool's constrained effective area utilization, time constraints like TIMESPEC, TMRGRP, TPSync, TNM, Operating condition, hierarchy-based synthesis and implementation for thermal management. The performance of Microblaze processors is boosted with effective area utilization, DFS and Thermo Electric Coolers (TEC).

Keywords: Area utilization, DFS, dual TEC, stacked TEC, TDC, thermal management

¹ Electronics and Telecommunication, Rajarambapu Institute of Technology, Sangli, 415414, Maharashtra, India. E-mail: sachimgd@gmail.com

² Electronics and Communication Engineering, Visvesvaraya Technological University, Belagavi, 590018, Karnataka, India.

* Corresponding author

Introduction

It is a known fact that a high density of transistors in a chip suffers from thermal issues. The main performance hindrance parameters of VLSI chips are carelessly considered parameters viz. area, speed and power dissipation. To achieve better performance of the chip, there should be an optimum tradeoff between these three factors. Among these, the power dissipation is always in terms of heat. Hence monitoring the temperature is circuitously monitoring the power dissipation that directly decides the performance of the chip. The major parameters involved in power dissipation are V-Applied voltage, f- Transition frequency and C- Node capacitance with relation $P = CV^2f$. The process of controlling the power dissipation is controlling the effects V, f, C, in terms of temperature. The implementation of autonomous thermal management techniques to maintain health and performance of chip is an interesting research domain. In a multicore scenario, multiple processors run at a different speed with a different workload. Due to non-uniform workload, processors exhibit non-uniform temperature distribution inside the die. Therefore, there is a need for an intelligent, reliable temperature monitoring and control system to sense power dissipation in terms of heat and take necessary action to increase the performance. Performance of processor is also dependent on placement, congestion in placement of processor elements, operating frequency and adoption of cooling techniques.

The implementation of efficient thermal management is a demand for complex MultiProcessor Chips (MPCs) to improve run time performance. The technique of thermal management includes sensing, monitoring and controlling of on-chip temperature. Many researchers addressed the thermal management of chip with the prediction of temperature and workload based on thermal models, and used the techniques like, workload scheduling, sharing and throttling (Zhu *et al.*, 2008; Ayoub *et al.*, 2010; Liu *et al.*, 2011; Sheikh *et al.*, 2012). But the need of

high computing facility increased the chips power dissipation in terms of tens to hundreds watt (Belady and Malone, 2006). Hence the thermal management with passive method (prediction and workload sharing) is not sufficient to maintain the health and performance of computing chips. This requires runtime thermal sensing, monitoring and controlling. The thermal management of onchip multiple processors is reported in (Cochran and Reda, 2010, 2012; Christoforakis *et al.*, 2015) used diode sensor to measure power dissipation in terms of heat. Based on the temperature value the processors are throttled with workload or clock. Here the temperature sensor plays an important role. The inaccurate detection of temperature misleads the thermal management unit and reduces the throughput of processor due to improper scheduling, sharing and throttling.

The temperature detection technique proposed in (Shen and Qiu, 2011) gives the sensed values in analog voltage form. Used temperature transducers (sensors), transducer interface units, operational amplifiers, multiplexer, Analog to Digital Converter (ADC) and clock sources. Here the sensed temperature is in the form of differential voltage. The temperature sensing techniques proposed in (Cochran and Reda, 2010, 2012; Christoforakis *et al.*, 2015) uses diode as temperature sensor that sense temperature in analog form. Hence, ADC is the main part of such temperature sensors to extract temperature. Due to the analog nature of temperature information, the circuit noise may easily get superimposed with temperature and at a long-range, ADC may exhibit nonlinearity problems. Also ADCs in low voltage chips have limitations on voltage range, resolution, sensitivity. For the thermal management of MPCs the digital temperature sensor is proposed in (Shen and Qiu, 2011). Here the author revealed more about the implementation of multicore processors and DFS based on the onchip temperature. To implement DFS technique, the onchip temperature detected with 200 logic elements,

4,096 repetition cycles and 12-bit counter with an external reference clock. With this setup the author claimed that the proposed DFS technique has reduced the chip temperature by 1.8°C.

The fully digital on-chip temperature measurement technique proposed in (Javier *et al.*, 2001; Chen *et al.*, 2007; Magadum *et al.*, 2012; Korkian *et al.*, 2018) used long delay line sensor and Time to Digital Converter (TDC). Here the coarse TDC used for temperature detection requires a reference clock along with a ring oscillator. Hence its resolution depends on reference clock frequency. The work depicted in (Jang *et al.*, 2013) measures the temperature by adjusting initial delay difference between two delay chains. This technique has an effective resolution of 1°C. But the delay adjustment is a little complex task due to voltage variations and non-uniform temperature at various parts of FPGA. The distribution of sensing elements over the entire chip makes difficult to identify the hotspot on chip. The identification of hotspot on FPGA is revealed in (Mattada *et al.*, 2016), used the tiny temperature sensors and ring oscillator based Vernier TDC (VTDC) to extract the temperature. Here the single pulse directly extracted from delay line sensor, does not carries satisfactory temperature information. Also, the difference in oscillating periods of two ring oscillators decides the resolution of VTDC. With proper placement satisfactory resolution can be achieved but, finding a proper placement on FPGA is time-consuming due to non-uniform temperature distribution and non-uniform activities inside the FPGA. Also, VTDC is sensitive to onchip temperature; hence use of separate FPGA for computation is not feasible. To provide thermal stability and to reduce temperature measurement error the multiphase clock-based Vernier TDC is developed (Sachin Magadum *et al.*, 2017). This TDC uses Phase Locked Loops (PLL) and complex clock select circuitry. Due to small time difference between various phases, the clock select circuitry many times fails to meet minimum timing requirements and turns output into tristate. But here the PLLs provide better stability to temperature.

The thermal stability to conversion circuitry and exposure of sensing elements to temperature can be achieved by manual place and route facility provided in reconfigurable logic. The Xilinx plan ahead tool provides flexibility to place temperature sensing elements and thermally critical blocks on FPGA with assigning necessary constraints (Xilinx User Guide 2012). This helps to accomplish high performance with better sensitivity and accuracy. The timing-driven detailed placement technique for FPGAs to optimize critical paths is addressed in (Dhar *et al.*, 2017). This technique used to optimize performance specific paths. But in proposed work, the said concept is utilized in a reverse way such that placement of sensing elements should provide proportional delay based on temperature and remaining functional blocks must be adhered to minimum possible (optimized) delay.

After setting the road map of thermal detection the monitoring and controlling parts are essential to complete the loop of thermal management of MPCs to boost the performance. In proposed thermal management techniques, the single stage and multistage structure of Thermoelectric Coolers (TECs) are used along with DFS technique. To cool the chip (FPGA) after detection of significant increase in on-chip temperature, these TECs are operated at their maximum Coefficient of Performance (COP) (Patel *et al.*, 2015). An assessment of use of TECs for die cooling is reported in (Simons *et al.*, 2003). The series of papers on modeling of TEC and used for IGBT based die cooling depicted with simulation in (Caroff *et al.*, 2013; Wunderle *et al.*, 2013; Springborn *et al.*, 2014).

The modern computing chips (CPU) come up with a very high density of transistors with high computation power. Hence the effective implementation of thermal management in run time with identification of hotspot and cooling entire chip is the theme of this proposed research work. The proposed work significantly contributes in effective thermal management of chips to improve the performance at high operating power dissipation. This incorporates the configuration of onchip

processors, design and implementation of digital temperature sensors, high resolution temperature to digital converter and cooling methodologies. The proposed paper is organized as follows; the methodology of proposed work is depicted in next section, results are discussed in results and discussion section and finally the conclusion section.

Methodology

The conceptual block diagram of proposed work is as shown in Figure 1. The network of processors is configured with Fast Simplex Link (FSL) communication. The PLL and clk select circuitry are used for DFS of processors. The thermal sensing unit is a combination of on-chip temperature sensor and high resolution temperature to digital converter. The on-chip temperature sensor dedicated to each processor gives the information of load (power dissipation) of the processor. The thermal sensing unit is set with a threshold value, if the processor temperature exceeds above the set value the clock select circuitry selects the clk of PLL and scales the operating frequency of processor. This DFS technique reduces the processor operating speed and hence the power dissipation. As a load of processor increases the operating frequency reduces and once the optimum operating frequency is reached, the cooling unit gets triggered and cools the chip. The various cooling methodologies and their results are discussed in further sections. The VIRTEX 5 FPGA is used for the implementation DFS technique.

Network of Processors

The provision of Xilinx Platform Studio (XPS) is used to configure identical Microblaze processors with equal resources and peripherals. The FSL established the communication between processors. The processors are made to execute complex codes with proper combination of functions that offers high switching activities. The configured processors also operates timers in complex modes that exhibit enough activities to produce enough power dissipation inside the

processor for testing. The provision of manual place and route provided by XILINX plan ahead tool is effectively utilized to place the processors at a particular place to form a network on FPGA. The options like *create Pblock*, *Auto create Pblock*, percentage of area utilization of processor and hierarchical synthesis models are explored for testing of thermal management technique.

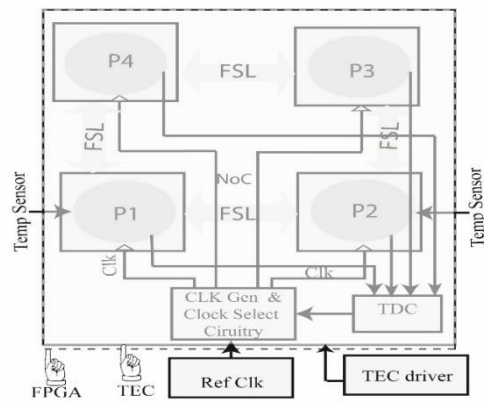


Figure 1. Proposed concept of thermal management technique

Temperature Sensing Unit

The temperature sensing unit is composed of digital temperature sensor and TDC. The proposed temperature sensor made up of 21 inverters and pulse accumulation unit to improve the resolution. This setup handles temperature information accurately with an error of less than $\pm 0.1^\circ\text{C}$. The relation between temperature and time delay of each sensing element used here is shown in Equation (1).

$$T_D = \frac{\left(\frac{L}{W}\right)C_L}{\mu_{\text{CoX}}(V_{\text{DD}} - V_T)} \ln \left(\frac{1.5V_{\text{DD}} - 2V_T}{0.5V_{\text{DD}}} \right) \quad (1)$$

where, L- MOSFET's channel length, W – MOSFET's channel width, CL- Load capacitance, μ - Electron mobility, V_T – Switching threshold voltage, V_{DD} - Supply voltage.

The overall pulse time provided by the delay line after cycled through itself is as per the Equation (2).

$$T_{\text{Total}} = T_D * n * C_y \quad (2)$$

where, n -no of inverters, C_y - No of cycles.

The placements of these sensing elements are done on each processor to identify the thermal activities of processors. The thermal values of each processor are sampled and read after every second by temperature sensitive Time to Digital Converter (TDC).

The TDC is a basic building block of temperature sensing unit, which detects the temperature by converting time into digital form. The 16 true phase TDC is used to detect the temperature with the time resolution of 106 ps.

Dynamic Frequency Scaling Unit (DFS)

The DFS unit composed of PLL and Clock Select Circuitry (CSC). Initially, the processors are configured to operate at a particular frequency of 50MHz. The processors are insisted to run with complex functions. The bare metal flow method is used to run the processors with complex arithmetic, data transfer and switching functions to test the temperature behavior of processors through temperature sensing unit allocated to each processor. The temperature sensing unit continuously monitors the temperature and when temperature reaches above the set threshold value it gives a signal to DFS unit. The DFS unit changes the clock based on temperature value. The DFS unit operates the processors on 50 MHz, 66 MHz, 75 MHz, and 99 MHz frequencies. The CSC has a multiplexer that selects the clock and assigns it to the processor based on temperature value from TDC. The results of this technique are discussed in results and discussion section. Once the optimum frequency is reached (in this case 50 MHz), further downscaling of frequency malfunctions the operation of processors. Hence the proposed cooling methodologies significantly contribute to

improving operating speed and performance of processors inside the FPGA.

Cooling Topologies

The main components of cooling section are Thermoelectric Cooler (TEC) and heat sink. Based on cooling strategy and arrangement of cooling devices, there are three topologies presented in this paper. If the temperature rises at any processor, cooling section cools the entire chip and brings down to optimum temperature level. The following are the three cooling topologies proposed in this paper.

i) One side one TEC cooling

Figure 2(a) shows one side cooling with single TEC and heat sink. The cold surface of TEC is placed on die. The silicon thermal interface material is used to fill the gap between die surface and TEC cold surface. The hot surface is aligned on heat sink which is made up of aluminum 6061 that has good heat conductivity. The thermal conductivity of heatsink can be further improved by using pure copper with little cost overhead. The Marlow DT12-6 TEC is used for single side cooling which has an operating current of 0.1 Amps to 5.6 Amps. The satisfactory performance range of TEC current to cool the die is from 2.5 Amps to 3.5 Amps. Beyond this upper range of operational current, the effective resistance

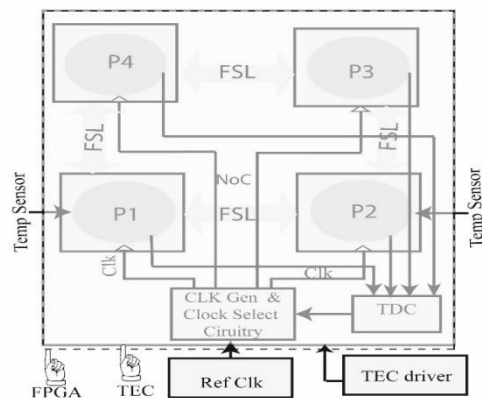


Figure 2. Proposed concept of thermal management technique

varies and heat flux from hot side dominates on cold side and the cold side temperature starts increasing. Hence, it increases the die temperature. At lower range of current the TEC cool flux is not able to penetrate into the die and cool the chip. Hence, die heat dominates on cool side of the TEC and increases the heat. The effective resistance offered by TEC at different temperature is shown with Equation (3).

$$R_{\text{eff}} = R_{\text{ref}} [1 + \alpha(T_{\text{actual}} - T_{\text{ref}})] \quad (3)$$

where, R_{ref} - Resistance offered at reference temperature, T_{ref} - Reference temperature, α - thermal coefficient, and T - actual temperature

The operating of TEC without heat sink is not a good design of cooling system. To avoid the risk of damage to TEC from overheating, the heat sink is used. The heat sink attached to TEC hot side has 16 internal plate fins of extruded/cast type to dissipate maximum heat. Based on cooling requirement and amount of heat to be dissipated in an external environment, the heat sink with fan is a good option for better die cooling approach. The results of this topology are discussed in results and discussion section.

ii) Stacked TEC cooling with heat pipes

To improve the Delta T max (i.e. the temperature difference between hot side and cold side of TEC) and cooling effect, the stacked TEC topology is used. Figure 2(b) shows the arrangement of stacked TEC with heat pipes to dissipate maximum heat from the die. The two TECs are stacked based on their size in a pyramid pattern. The Die_TEC (TEC1) is directly in contact with FPGA has same size as die i.e., 40 mm*40 mm with an operational current range of 2Amps to 2.5 Amps. The Top_TEC (TEC2) which is stacked on Die_TEC has 30 mm*30 mm size which is less than Die_TEC to avoid blocking of entire hot side of Die_TEC. This helps Die_TEC's hot side temperature to dissipate in the external environment and the central part gets cooled

by Top_TEC. The Top_TEC has capacity to overcome the heat dissipated by some part of Die_TEC. The Top_TEC's hot side is attached to heatsink with silicon thermal interface material providing proper thermal conductivity between TEC and heatsink. The heat pipes further contribute to dissipating the Die_TEC's hot side heat into heatsink as per the Equation (4). Hence, this topology allows die to operate at maximum power or processors at maximum frequency to increase the speed and performance. The results of this topology are discussed in results and discussion section.

$$Q = hA(T_{\text{surr}} - T_{\text{sur}}) \quad (4)$$

where, A - Surface area where heat transfer takes place, T_{surr} - Temperature of surrounding area T_{surf} - Temperature of surface (Hot side of TEC), h - Heat transfer coefficient = $q/\Delta T$, q -Heat flux, ΔT -Difference in Temperature.

iii) Dual side dual TEC cooling

In the scenario where many components connected on a single PCB, each component experiences the mutual heat due to individual components' power dissipation. When chip exhibits high power dissipation, the surrounding components also experience the heat of die. Figure 2(c) shows the dual side cooling topology where two TECs are connected on both side of Die or PCB. This arrangement is more effective to the onboard die having Dual or Quad inline Packages. This is also applicable to area IO Ball Grid Array (BGA) package with proper interface between Die/PCB and TEC cold side. The Top_TEC is controlled by thermal sensing unit. Based on maximum power dissipation or heat, the bottom TEC (Lo_TEC) can be operated with constant current to maintain the die and its surrounding components at normal temperature range. The bottom TEC can also be used as a heater to avoid moisture saturated as an effect of condensation due to overcooling at low die power dissipation. The results of this cooling topology are discussed in results and discussion section.

Results and Discussion

The proposed temperature sensor designed with 21 inverters carries temperature information in the form of pulses. These pulses are cycled through the same delay line and accumulated no. of pulses to construct a single pulse that carries sufficient temperature information. The single pulse generated with pulse accumulation technique provides the time variation of $\sim 3\text{ns}$ at every 1°C rise in temperature. With pulse accumulation sensor and 16 phase TDC offered the temperature resolution of 0.03125°C with temperature measurement error less than $\pm 0.1^\circ\text{C}$.

Figure 3 shows the time required to reach set threshold temperature value due to percentage of FPGA's Pblock area utilization by processor, running at full load with set operating frequency. The XILINX plan ahead tool provides the flexibility to place modules on FPGA with required percentage of area among selected Pblock boundary. The boundary of Pblock is larger than the area required for processor then the processor resources spread over the Pblock with less congestion, hence less impact of mutual temperature produced by surrounding modules. Hence takes more time to reach set threshold. This time also varies with operating frequency of processor. As percentage of area utilization increases (Pblock boundary is reduced) the processor resources densely placed within the Pblock. Due to this, the impact of surrounding components is considerable when processor running at high speed with full load. This may spread heat to other processors that are on same die. The identification of hotspot on each Pblock is achieved by small temperature sensors (Mattada *et al.*, 2016). To achieve high performance with appropriate thermal management, there must be a tradeoff between area utilization and operating speed. Once optimum % of area utilization and operating speed of a processor are reached then external cooling techniques contribute further for effective thermal management.

At 99 MHz operating frequency with 100% area utilization by processor takes 2 min

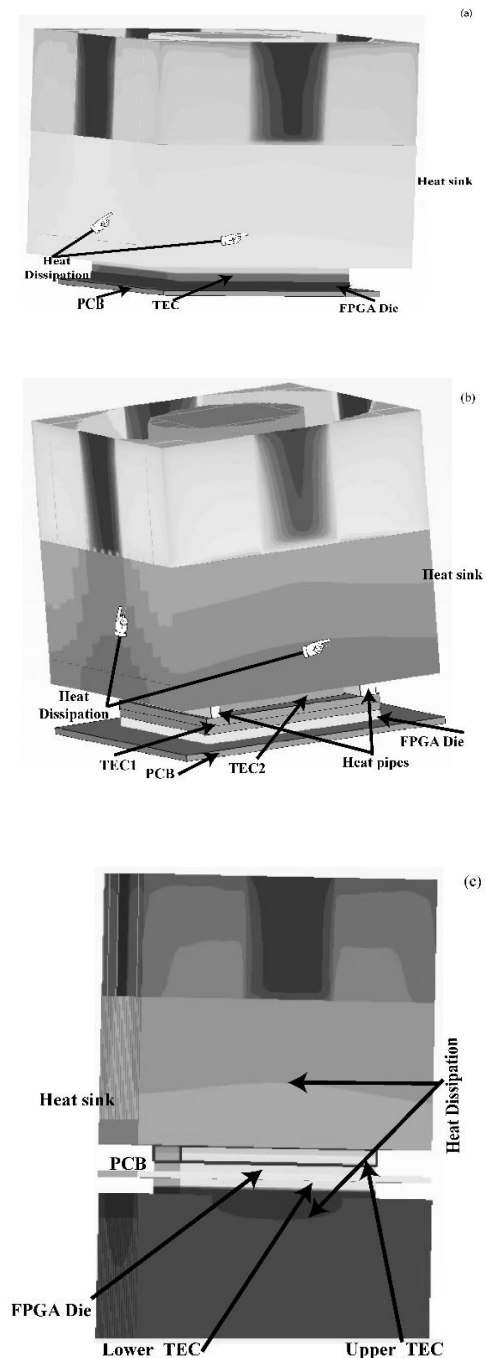


Figure 3. Cooling Topologies a) Single Side Cooling; b) Stacked TEC cooling; and c) Dual side cooling management technique

of time to reach set threshold temperature (48°C) with fixed hardware and functions running inside the processors. As temperature threshold is reached, the operating frequency is switched to 50 MHz and then the processor has taken 4.55 min of time to reach same threshold. Through this, the maximum 9% of internal power dissipation is reduced by switching frequency from 99 MHz to its half. This value varies with 75 MHz and 66MHz. But reducing the frequency to reduce the power dissipation hampers the speed of the processor. Hence keeping processors operating frequency at 99 MHz maximum the external TEC is used to cool the chip. The DFS with TEC cooling reduced the temperature of 30% in 3 min of maximum TEC current (i.e. 3 Amps). Further to operate processors at higher speed with low power dissipation the different cooling technologies been introduced and analyzed the behavior at different die power dissipation.

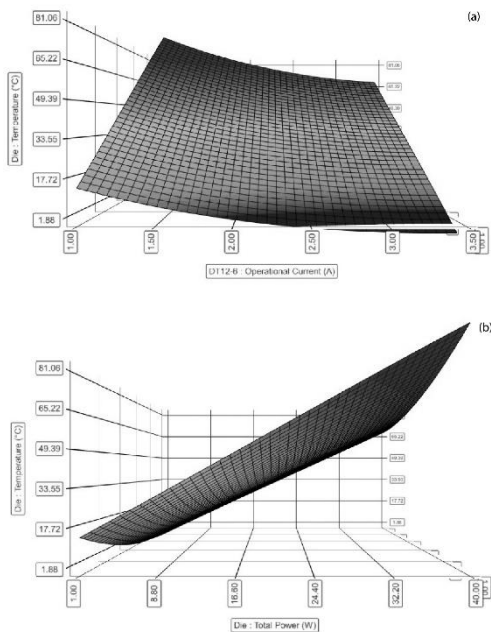


Figure 4. Variation of temperature due to die power and TEC current a) TEC current vs Die temperature; and b) Die power dissipation vs Die temperature

One side one TEC cooling technique cools the die based on the current flowing through the chip. Figure 4(a) shows the temperature variation due to die power and TEC current. Figure 4(b) shows the die temperature variation due to TEC current and die power. The maximum die temperature reached here at maximum tested die power (40 W) and maximum COP current of TEC is up to 69°C . As current increases through TEC at low internal die power i.e. below 5W the die temperature reduces below the lower range of commercial grade FPGA. With this die may get frozen and increases the chance of breaking. Hence at low power dissipation the TEC can be turned off. This can be achieved by thermal management unit to keep temperature well within the commercial grade range.

The stacked TECs with heat pipes shown in Figure 2(b) dissipate more heat from the die as compared to topology 1. Figure 5(a) shows the change in die temperature due to die power and TECs current. Figure 5(b) shows the die temperature due to top TEC current and die power, with keeping lower TEC current at its maximum COP. The maximum possible die temperature at tested max power (40 W) is 63.3°C . In this topology, both TECs are operated at their maximum COP currents. The heat flux generated at hot side of lower TEC is brought to heat sink by heat pipes. The heat sink embedded with heat pipes absorbs

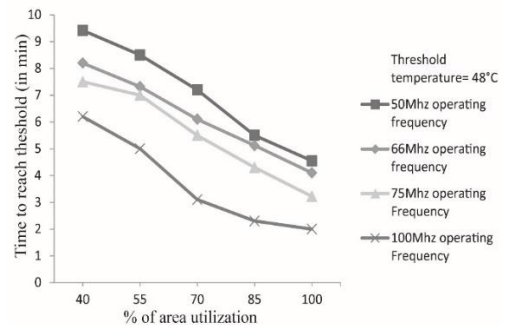


Figure 5. Time to reach threshold temperature with area utilization and operating frequency

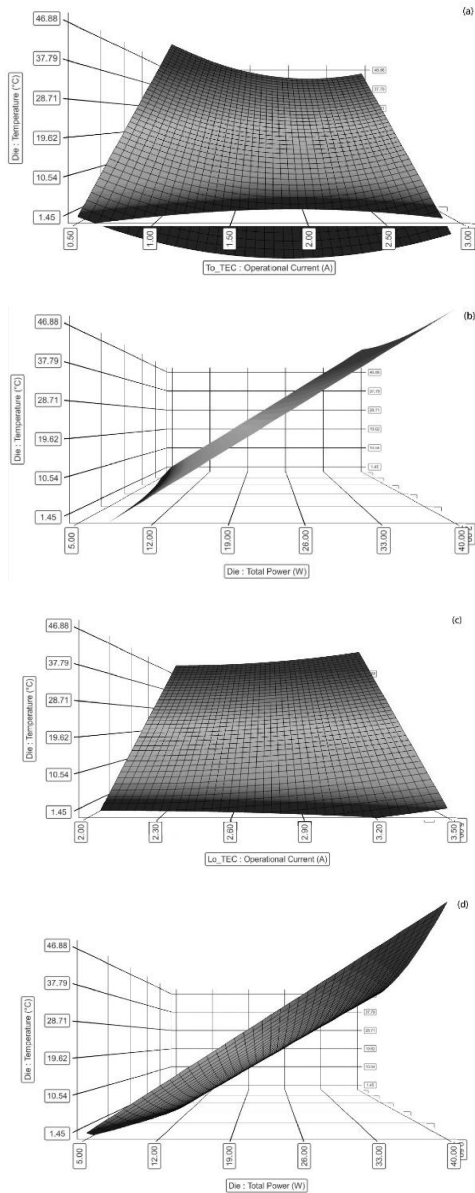


Figure 6. Die temperature due to die power and dual side cooling a) Variation of die temperature due to cooling through top TEC; b) Variation of die temperature due to die power dissipation and bottom TEC current; c) Variation of die temperature due to cooling through bottom TEC; and d) Variation of die temperature due to die power dissipation and cooling through top TEC

maximum heat through heat pipes. The cold surface of top TEC cools the hot side of bottom TEC. Its cold side is in contact with die surface. The combination of heat pipes and TEC avoids hot flux to dominate on cold side and hence the die experience more cold flux and cools the chip. The heat carried by heat pipes is as shown in Figure 7. As heat pipes carry most of the temperature, the onboard components which are near to die are also experience the heat of die/heat pipes. This problem is tackled in topology 3 i.e. Dual side cooling.

The contribution of dual side cooling thermal management technique is as same as topology 2 at low power dissipation. But as power dissipation inside the die increases, this topology plays a significant role in cooling the chip as shown in Figure 6. The maximum temperature of die at maximum tested power is 45.1°C. Also the heat experienced by surrounding components is less as compared to other topologies as shown in Figure 7. Based on power dissipation, the current through lower TEC is varied to avoid the overcooling

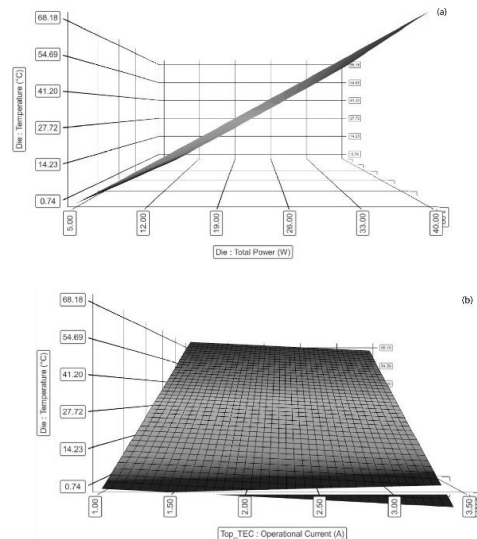


Figure 7. Variation of temperature due to die power and stacked TEC current a) Variation of die temperature due to Die power dissipation; and b) Variation of die temperature due to cooling through Top TEC

of die. The effect of overcooling may lower the die temperature below the commercial grade temperature range (0°C) of FPGA as shown in Figure 6(a). At low power and low ambient temperature the onchip temperature of die is to be maintained above the 0°C . This is achieved by lower TEC with reversing the current through it, is as shown in Figures 6(b) and 6(c). Figure 6(d) shows the die temperature due to die power and current variation through top TEC keeping lower TEC current constant. As power dissipation increases, the thermal sensing unit increases the current through lower TEC to cool the die. At high power dissipation, the lower and upper TECs are operated at their maximum COP current. The TECs used in dual side cooling takes less current and provides a better cooling effect as compared to topology 1 and 2. The availability of low voltage high current TECs reduces the power overhead of cooling unit.

The dual side cooling technique also extends its advantage for grade extension of integrated circuits. Since the commercial grade and industrial grade FPGAs are almost same except the operating temperature range. The temperature ranges of commercial and industrial grade chips are 0°C to 85°C and -20° to 125°C respectively. The thermal management topologies 1 and 2 contribute to extending the upper range of temperature by cooling the die at high power dissipation. The topology 3 extends the upper range by cooling the chip with both TECs and lower

temperature range by utilizing lower TEC as a heater. The topology 3 maintains the lower range of temperature of chip well within the commercial grade temperature range just by changing the direction of current through lower TEC even at low ambient temperature.

Table 1 shows the comparison of all three cooling topologies when TECs are operated at their maximum COP currents. At 5 W of die power, the topology 1 has die temperature of 1.5°C and TECs in other topologies are used as heater to maintain proper die temperature or turned off at low die power. At maximum tested die power, the dual side cooling gives a better cooling performance as shown in the last column of Table 1. This topology reduced the die temperature by 34% at 40W of power dissipation.

Table 1. Comparison of proposed cooling topologies

Die Power Dissipation (in Watts)	Topology 1 Die Temperature (in $^{\circ}\text{C}$)	Topology 2 Die Temperature (in $^{\circ}\text{C}$)	Topology 3 Die Temperature (in $^{\circ}\text{C}$)
5	1.5	-	-
10	12.65	1.89	4
15	23.88	12.18	11
20	31.14	22.7	17
25	40.26	31.32	24.3
30	52.22	43.96	34.5
35	59.6	53.64	40.4
40	69	63.3	45.1

Conclusions

The proposed thermal management techniques exhibit better cooling performance. The developed thermal sensing unit showed a resolution of 0.03125°C over the measurement range of 0°C to 85°C with an error of $\pm 0.1^{\circ}\text{C}$. It utilized less than 2% of onchip resources. The DFS with TEC reduced 30% of internal power dissipation without compromise in speed. The provision of manual place and route is utilized efficiently to place thermally critical blocks for their better performance. Further the performance of processor is enhanced by thermal management techniques with three topologies, each tested at 40 W of die power.

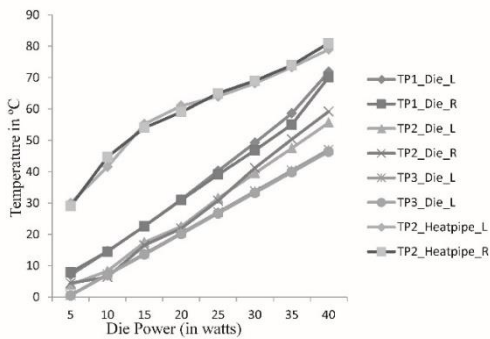


Figure 8. Die surrounding temperature and heat pipe temperature at die power on PCB

The topology 1 exhibit maximum die temperature of 69°C, topology 2 exhibits maximum die temperature of 63.3°C and topology 3 exhibits a maximum temperature of 45.1°C. High rating TECs can be used, if die required to be operated at high power to keep maximum temperature below 85°C. Topology 1 is used where area is concerned and less components mounted on PCB. Topology 2 is used where speed and performance are major concerns along with less no of surrounding

References

- Ayoub, R., Indukuri, K.R., and Rosing, T.S. (2010). Energy efficient proactive thermal management in memory subsystem. *Proceedings of the International Symposium on Low Power Electronics and Design*: p. 195–200.
- Belady, C. and Malone, C. (2006). Data center power projections to 2014. *Thermomechanical Phenomena in Electronic Systems -Proceedings of the Intersociety Conference*. <https://doi.org/10.1109/ITHERM.2006.1645376>
- Caroff, T., Mitova, R., Wunderle, B., and Simon, J. (2013). Transient cooling of power electronic devices using thermoelectric coolers coupled with phase change materials. *THERMINIC 2013 - 19th International Workshop on Thermal Investigations of ICs and Systems*, *Proceedings*: p. 262–267.
- Chen, P., Shie, M. C., Zheng, Z.Y., Zheng, Z.F., and Chu, C.Y. (2007). A fully digital time-domain smart temperature sensor realized with 140 FPGA logic elements. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 54(12):2,661–2,668.
- Christoforakis, I., Tomoutzoglou, O., Bakoyiannis, D., and Kornaros, G. (2015). Dithering-Based power and thermal management on fpga-based multi-core embedded systems. *Proceedings - IEEE/IFIP 13th International Conference on Embedded and Ubiquitous Computing, EUC 2015*:173-177.
- Cochran, R. and Reda, S. (2010). Consistent runtime thermal prediction and control through workload phase detection. *Proceedings - Design Automation Conference*. <https://doi.org/10.1145/1837274.1837292>
- Cochran, R. and Reda, S. (2012). Thermal prediction and adaptive control through workload phase detection. *ACM Transactions on Design Automation of Electronic Systems*. <https://doi.org/10.1145/2390191.2390198>
- Dhar, S., Iyer, M.A., Adya, S., Singhal, L., Rubanov, N., and Pan, D.Z. (2017). An effective timing-driven detailed placement algorithm for FPGAs. *Proceedings of the International Symposium on Physical Design, Part F1271(ii)*:151–158.
- Jang, J., Kim, J., Oh, R., and Sung, M.Y. (2013). All digital on-chip temperature sensor using dual ring oscillators. *Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems*: p. 181–184.
- Javier, S., Garrido, J., and Boemo, E. (2001). Measurement of FPGA Die Temperature Using Run-time Reconfiguration. In: *Proceedings of the 7th International*, 1304:69-78.
- Korkian, G., Rahmanikia, N., Noori, H., and Clemente, J. A. (2018). Exploration of ring oscillator based temperature sensors network accuracy on FPGA. *2017 19th International Symposium on Computer Architecture and Digital Systems, CADSD 2017, 2018-January*: 1–6. <https://doi.org/10.1109/CADSD.2017.8310728>
- Liu, S., Leung, B., Neckar, A., Memik, S.O., Memik, G., and Hardavellas, N. (2011). Hardware/software techniques for DRAM thermal management. *Proceedings of International Symposium on High-Performance Computer Architecture*: p. 515–525.
- Magadum, S., Naveen, K.V., and Guhilot, H. (2012). Sensor less distributed temperature sensor and control using FPGA. *Int. J. Eng. Innov. Technol.*, 1(5):295–298.
- Magadum, S., Mattada, M.P., and Guhilot, H. (2017). Multiphase clock based vernier TDC on FPGA for on-chip temperature measurement application. *2017 2nd International Conference for Convergence in Technology, I2CT 2017, 2017-January*: 828–831. <https://doi.org/10.1109/I2CT.2017.8226244>
- Mattada, M.P., Magadum, S.M., and Guhilot, H. (2016). Identification of hotspots on FPGA using Time to Digital Converter and distributed tiny sensors. *JISSE*, 5:1–11.
- Patel, J., Patel, M., Patel, J., and Modi, H. (2015). Improvement In The COP Of Thermoelectric Cooler. *Int. J. Sci. Technol. Res.*, 4(8):73–76.
- Sheikh, H.F., Ahmad, I., Wang, Z., and Ranka, S. (2012). An overview and classification of thermal-aware scheduling techniques for multi-core processing systems. In *Sustainable Computing: Informatics and Systems*. <https://doi.org/10.1016/j.suscom.2011.06.005>
- Shen, H. and Qiu, Q. (2011). An FPGA-based distributed computing system with power and thermal management capabilities. *Proceedings of International Conference on Computer Communications and Networks, (ICCCN)*, <https://doi.org/10.1109/ICCCN.2011.6005802>
- Simons, R.E., Ellsworth, M.J., and Chu, R.C. (2003). An assessment of module cooling enhancement with thermoelectric coolers. *American Society of Mechanical Engineers, Electronic and Photonic Packaging, EPP*. <https://doi.org/10.1115/IMECE.2003-42239>
- Springborn, M., Wunderle, B., May, D., Mrossko, R., Manier, C.A., Oppermann, H., Ras, M.A., and Mitova, R. (2014). Thermal management of electrical overload cases using thermo-electric modules and phase change buffer techniques: Simulation, technology and testing. *2014 15th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in*

- Microelectronics and Microsystems, EuroSimE 2014, p. 1–11. <https://doi.org/10.1109/EuroSimE.2014.6813857>
- Wunderle, B., Manier, C.A., Abo Ras, M., Springborn, M., May, D., Oppermann, H., Toepper, M., Mrossko, R., Xhonneux, T., Caroff, T., Maurer, W., and Mitova, R. (2013). Double-sided cooling and thermo-electrical management of power transients for silicon chips on DCB-substrates for converter applications: Design, technology and test. THERMINIC 2013 - 19th International Workshop on Thermal Investigations of ICs and Systems, Proceedings, 2013, p. 253–261. <https://doi.org/10.1109/THERMINIC.2013.6675189>
- Xilinx. (2012). PlanAhead User Guide - UG632, (632):454
- Zhu, C., Member, S., Gu, Z., Member, S., Shang, L., Dick, R.P., and Joseph, R. (2008). Three-Dimensional Chip-Multiprocessor Run-Time Thermal Management. 27(8):1,479–1,492.