Hot Spot Cooling using Embedded Thermoelectric Coolers

G. Jeffrey Snyder, Marco Soto, Randy Alley, David Koester, Bob Conner
Nextreme Thermal Solutions
Research Triangle Park, NC 27709

Abstract

Localized areas of high heat flux on microprocessors produce hot spots that limit their reliability and performance. With increasingly dense circuits and the integration of high power processors with low power memory, non-uniform thermal profiles will become more dramatic and difficult to manage. Chip scale thermal solutions designed to keep hot spots below a critical temperature unnecessarily overcool the rest of the CPU and add to heat-sink load. Localized hot spot cooling solutions, even active systems that contribute some additional heat, can do a better job controlling hot spot temperatures when efficiently integrated with a heat spreader. Embedded thermoelectric cooling (eTEC) is a promising approach to reduce the temperature of highly localized, high heat flux hot spots generated by today's advanced processors.

Keywords

Localized hot spot cooling. Thermoelectric, Peltier Cooling. Embedded thermoelectric cooler, eTEC

Nomenclature

A	area, m^2
COP	coefficient of performance
E_a	activation energy, eV
f	packing fraction
h	heat transfer coefficient, W/m^2K
I	electric current, A
J	electric current density, A/m^2
k_b	Boltzmann's constant, eV/K
l	thermoelectric element length, m
n	number of thermoelectric couples
Q	heat absorbed, W
S	device Seebeck coefficient, V/K
T	temperature, K
ΔT	temperature difference, K
V_{th}	transistor threshold voltage, V
w	well width, cm
Z	thermoelectric figure of merit, K^{-1}

Greek symbols

α	Seebeck coefficient, V/K
ρ	electrical resistivity, Ω·m
σ	electrical conductivity, $\Omega^{-1}\cdot m^{-1}$
κ	thermal conductivity, W/m K
Θ	thermal resistance, K/W
τ	time constant, s

Subscripts

a	ambient
c	cold side
j	junction
h	hot side
max	maximum performance condition

1. Introduction

With the decreasing feature size and increasing clock speed, in accordance with Moore's law, the design of microelectronics and optoelectronics is facing ever increasing challenges for thermal management. With more power produced, and more densely integrated circuits, higher heat fluxes must be dissipated. This results in higher die temperatures and demand more complex and expensive thermal management systems. Management of overall power consumption is one of the near-term "grand challenges" for enhancing performance identified by the International Technology Roadmap for Semiconductors. The increased power consumption is driven by higher chip operating frequencies, higher interconnect resistances and the increasing leakage from exponentially larger numbers of progressively smaller on-chip transistors. High-performance processor power consumption is forecast to significantly exceed perceived single-chip power limits. Increasing power fluxes adversely affect reliability and performance. Decreasing supply voltages worsen switching currents and noise. Next generation design technology must address this growing power management gap [1]. Site-specific eTEC-assisted heat conduction provides a new design element for power management.

2. Temperature Dependent Issues

An increase in temperature creates lifetime and reliability problems for integrated circuits.

2.1. Reliability

There are many possible temperature dependent failure mechanisms for integrated circuits. One example is failure due to electromigration. In this case the lifetime of an integrated circuit depends exponentially on its temperature, according to the model of Black [2]:

$$\tau = \frac{C}{J^2}\exp\left(\frac{E_a}{k_b T}\right) \quad (1)$$

where τ, is the mean time to failure, C, is a proportionality constant, J, is the current density, k_b is Boltzmann's constant, T is the absolute temperature. E_a is an activation energy, where the typical value for silicon is about 0.68 eV. Accordingly, device reliability depends on the maintaining T below acceptable levels for all the circuits in the chip.

2.2. Subthreshold Leakage

Higher temperatures also compound the thermal management problem by increasing overall power

consumption due to increased subthreshold leakage. The off-state subthreshold leakage current in transistors is becoming a major contributor to total power dissipation as feature size drops. Even today, this static power dissipation is becoming a limiting factor in low power design. The projected power consumption from subthreshold leakage exceeds the total dynamic power consumption as the technology drops below the 65 nm feature size [3].

Leakage current in a transistor occurs across the gate to drain (mitigated by high dielectric constant materials) and from source to the drain (mitigated by lowering the supply voltage or raising the threshold voltage, V_{th}). Designers have reduced supply voltages to below 1 V and are using multiple V_{th} as well. Areas of the chip built with high V_{th} have reduced leakage but sacrifice operating frequency (as it takes the transistors a longer time to turn on with a reduced voltage difference between the supply and V_{th}). Areas of the chip with lower V_{th} have higher speed, but higher leakage as well.

Since subthreshold leakage current is activated by temperature (and therefore increases rapidly with rising temperature [4]) cooling hot regions of the chip may actually result in *reduced* total power consumption.

3. Chip Scale Cooling

Clearly, one solution to the thermal problem is to provide improved cooling of the entire chip, preferably through passive methods (adding no or very little heat to the system being cooled) such as improved heat sink air cooling, use of phase change heat pipes and spreaders, microchannel liquid heat removal, etc. Active cooling techniques (which add a significant amount of heat to the system being cooled), such as using a thermoelectric (Peltier) cooler (TEC) shown in Figure 1, can result in reduced effective thermal resistances [5, 6], particularly if they are properly optimized [7].

However, even well-designed chip scale heat removal systems have size or other physical constraints which eventually limit the dissipation of additional heat. For active cooling of the entire chip with a TEC, it essential to optimize the design and operation of the TEC, considering both the additional heat produced and the temperature difference across the TEC, as well as the characteristics of the heat rejection path, so as to achieve an improved heat sink performance [7]. In short, a high thermoelectric efficiency, or coefficient of performance (*COP*) essential. *COP* is discussed further in Section 6.3.

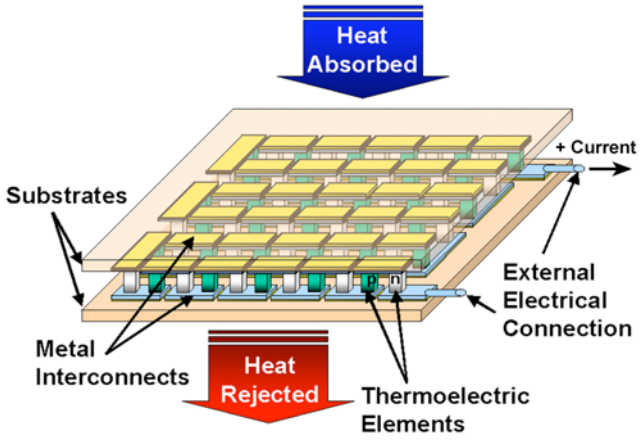

Figure 1. Schematic of a commercially available thermoelectric cooler using bulk materials fabricated on ceramic substrates.

4. Hot Spots

Localized regions of higher temperature, commonly known as hot spots, are becoming more severe as local power density and overall die power consumption increase. The temperature differential across a microprocessor die can vary from 5 to 30 K due to large variations in heat flux density. Similarly, optoelectronic devices can experience temperature variations in excess of 100 K. Today's microprocessors have an average heat flux of about 10-50 W/cm^2 and peak fluxes five times that of the surrounding areas [8]. These large variations in heat flux arise from increasingly higher levels of device integration. For example, the combination of high power logic and low power memory functions on the same chip produces non-uniform power and temperature distributions.

Hot spots occur where there are high speed circuits which typically use low V_{th} for maximum speed, so they have both high dynamic power and high source-drain leakage power dissipation. Some circuits (such as clock drivers) don't form significant hot spots because they occupy such a small area that their heat spreads effectively into the silicon wafer. Conversely, hot spots are a critical problem in larger-area, speed critical circuits, such as fixed point and floating point units in CPUs, phased lock loops / clock generators, and multiple functions in the graphics / video / 3D pipeline, etc. The coolest areas of the microprocessor are in the large cache regions. Between the hottest and coolest areas, chip-wide temperature differences of over 10 K can develop. [9].

Because thermal solutions must ensure that all junction temperatures in the processor (die temperature) do not exceed 90 to 110 ºC, the requirement pertains to the hottest part of the chip: the hot spots. Therefore chip scale thermal solutions, which cool all regions of the chip equally, are over designed to meet target hot spot temperatures. This results in larger, heavier, and more costly thermal solutions. *Thus hot spots, not the whole chip temperature, drive the thermal design*.

An approach to achieve high device performance and reliability is to design the chip with islands of high-speed (low V_{th}) logic that are aggressively cooled, surrounded by larger regions of low power, low-speed, high V_{th} circuits (particularly memory) which are passively managed. Such a

method, even if employing small regions of active cooling, can reduce the overall cooling requirement, allowing a more compact thermal management solution as well [3].

5. Hot Spot Solutions

The mitigation of hot spots by spot cooling or heat spreading is a topic of active research, with a number of passive or active cooling methods under investigation.

5.1. Passive Hot Spot Solutions

Passive heat spreading from the localized hot spots to areas of lower heat flux might be achieved with high thermal conductivity materials, such diamond, or with micro-channel two-phase cooling techniques such as jet/spray cooling or heat pipes. Unfortunately, diamond heat-spreaders are expensive and challenging to integrate with silicon while two-phase cooling systems are difficult to fabricate and complex to model [10].

5.2. Active Hot Spot Solutions

Active cooling requires the input of electrical energy, producing additional heat that must be removed. The requirements for active hot spot cooling to be effective are quite stringent and include the following:

First, sufficient cooling performance is required for the device to produce a temperature difference of at least several degrees under a heat load of a few watts. The solution must have a wide range of design flexibility because different applications have differing requirements, ranging from cooling a few degrees Celsius to tens of degrees Celsius with heat loads ranging from a few hundred milliwatts to several watts.

Second, the cooling solution should be site-specific to concentrate the cooling at the hot spot and minimize the additional heat removal requirement. This significant heat load and small size requirement translates to exceptionally high heat flux requirement, which can exceed 1000 W/cm^2.

Third, efficient cooling is required to minimize the additional heat load that must be removed at all the subsequent stages of the thermal management system.

Finally, a viable hot spot cooling solution requires reasonably simple methods of mechanical, thermal and electrical integration, as well as high reliability.

6. Embedded Thermoelectric Cooling

To date, the requirements delineated above have worked against the adoption of active hot spot chip cooling solutions. However, the recent development of microfabricated thermoelectric devices now provides a promising approach that addresses these critical issues, thus enabling the use of active hot spot thermal management.

Thermoelectric coolers operate by the following principle: When an electric current is driven through a circuit containing two dissimilar materials, heat is absorbed at one junction and released at the other junction. This Peltier effect is particularly strong when one material is an n-type semiconductor and the other a p-type semiconductor. The movement of heat provides a solid-state cooling capability that is ideally suited for applications where temperature stabilization, temperature cycling, or directed cooling is required. By adjusting or reversing the polarity of the applied DC current, heat flow can be proportionally modified allowing precise control of the junction temperature.

Thermoelectric coolers have traditionally been fabricated using bulk materials processing techniques. Commercial bulk thermoelectric coolers are made from thermoelectric pellets with typical dimensions on the order of 1 × 1 mm footprint and 1 mm thickness. These pellets are then assembled into an array of m × n elements between two ceramic substrates. State-of-the art bulk thermoelectrics can be made with pellets as small as 200 μm and provide ΔT_{max} of 70 K [11]. The ceramic substrates make the whole module about 1 mm thick, and greater than 2 – 3 mm in diameter.

There has been significant progress in recent years in making thermoelectric coolers using microfabrication techniques. Bi_2Te_3 and Sb_2Te_3 based materials have been microfabricated into micron scale thermoelectric devices by thick film electroplating [12], vacuum deposition [13, 14], and chemical vapor or MOCVD deposition [15] techniques. These microfabrication techniques are ideal for producing "Embedded Thermoelectric Components" (eTEC's), wherein the eTEC is unobtrusively integrated within the IC package.

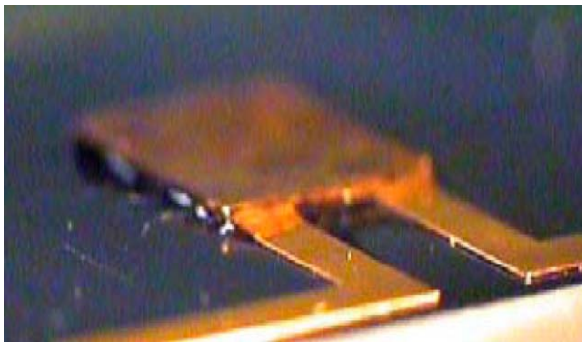

Figure 2. Embedded thermoelectric cooler, with total height of 100 microns and 2.5 mm x 2.5 mm footprint

6.1. Thermoelectric Cooling Performance

The heat (Q) absorbed at the cold side of a thermoelectric cooler is the result of Peltier cooling offset by thermal conduction and Joule heating losses. The following equation describes basic thermoelectric cooling:

$$Q = ST_c I - K\Delta T - \tfrac{1}{2} I^2 R \qquad (2)$$

Where S is the device Seebeck coefficient, T_c is the cold side temperature, I is the electric current, K is the thermal conductance, ΔT is the temperature difference across the device and R is the electrical resistance of the device.

Figure 3 shows the predicted performance curves for a 49-couple, 2.5 × 2.5 mm eTEC. As current increases, Peltier cooling (ST_cI) increases along with Joule heating ($I^2R/2$). As ΔT increases, the heat pumped Q decreases due to thermal conduction ($K\Delta T$) loss.

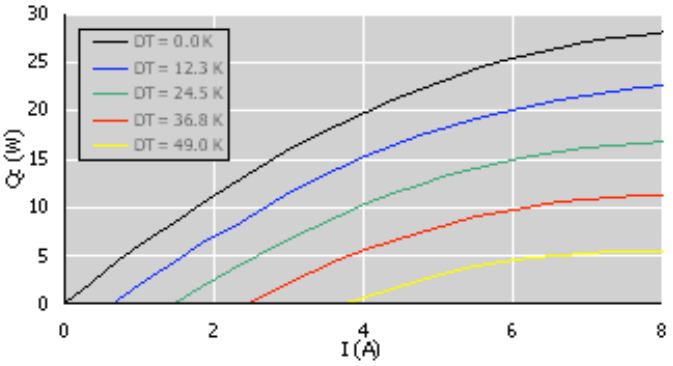

Figure 3. Predicted performance curves (Q for various applied currents, I) of a 49-couple eTEC at T_h = 85 ºC.

Figure 4 shows the same predicted performance as Figure 3 but for constant heat pumped Q. As current increases, the temperature difference ΔT increases to a maximum (then decreases) for a given heat load Q.

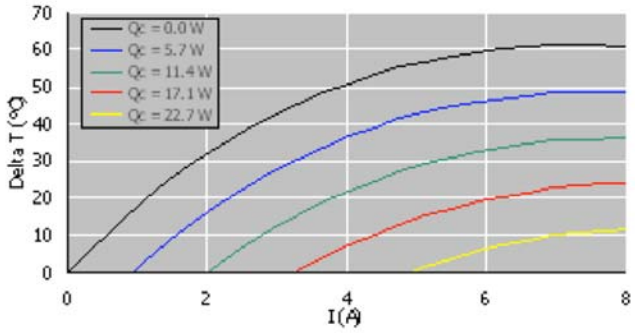

Figure 4. Predicted performance curves (ΔT for various applied currents, I) of a 49-couple eTEC at T_h = 85 ºC.

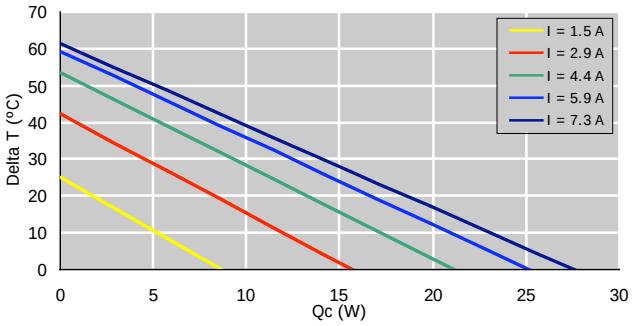

Figure 5. Predicted performance curves (Q and ΔT for various applied currents, I) of a 49-couple eTEC at T_h = 85 ºC.

Figure 5 combines Figures 3 and 4 to show ΔT versus Q for various currents, I. Above a current of about 7.3 A, ΔT and Q decrease because of Joule heating.

The eTEC voltage as a function of the applied current for different ΔT is shown in Figure 6. The voltage required by the thermoelectric cooler is the sum of the Seebeck ($S\Delta T$) and Ohm's law (IR) terms:

$$V = S\Delta T + IR \tag{3}$$

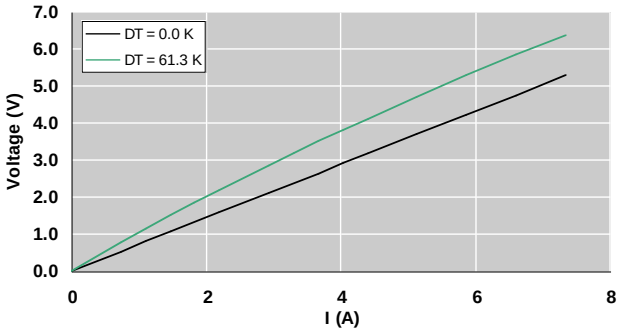

Figure 6. Predicted performance curves (V for various applied currents, I) of a 49-couple eTEC at T_h = 85 ºC.

The performance of a thermoelectric cooler can be predicted based on a few characteristic measurements. These include the maximum temperature difference (ΔT_{max}) when there is no heat (Q = 0) absorbed at the cold side, the electric current in the device (I_{max}) that gives this maximum temperature difference; and the maximum heat flow (Q_{max}) with no temperature difference (ΔT = 0).

For example, it can be shown that for constant S, K, R:

$$\Delta T_{\max} = \frac{S^2 T_c^2}{2KR} = \frac{ZT_c^2}{2} \tag{4}$$

$$I_{\max} = \frac{ST_c}{R} \tag{5}$$

$$Q_{\max} = \frac{S^2 T_c^2}{2R} \tag{6}$$

The dimensionless quantity $ZT = \frac{S^2}{RK}T$ is often referred to as the thermoelectric figure of merit, which is related to the material figure of merit $zT = \frac{\alpha^2}{\rho\kappa}T$, where α is the Seebeck coefficient, ρ, electrical resistivity, κ, thermal conductivity and T, absolute temperature.

In the simplest case (constant S, K, R, and T_c) the performance curves of the type shown in Figures 3, 4, and 5 are described by the following equation:

$$\frac{Q}{Q_{\max}} = \frac{I}{I_{\max}}\left(2 - \frac{I}{I_{\max}}\right) - \frac{\Delta T}{\Delta T_{\max}} \tag{7}$$

The performance curves for the more practical case of constant T_h (Figures 3-6) are only slightly different.

6.2. High Heat Flux

In addition to the thermoelectric material properties (α, ρ, κ), the electrical and thermal contacts as well as the number

and geometry of the thermoelectric elements influence the above performance parameters. While the maximum temperature difference ΔT_{max} depends primarily on *Z of the device* (which includes most of the effects of the electrical and thermal contacts), Q_{max} and I_{max} depend strongly on the size and number of the thermoelectric elements within the TEC module.

In particular, it can be shown that

$$\frac{Q_{max}}{A} = \frac{f}{l}\frac{\alpha^2 T^2}{2\rho} \tag{8}$$

$$I_{max} = \frac{Af}{nl}\frac{\alpha T_c}{2\rho} \tag{9}$$

where A is the area of the thermoelectric device, l is the length (i.e., thickness) of the thermoelectric elements, and f is the fraction of A covered by thermoelectric elements. Thus the cooling heat flux (Q/A) is directly proportional to f/l. In the above equations, α, and ρ are effective, averaged materials properties.

Control of Q/A can best be achieved by adjusting the length l, or filling fraction, f rather than the materials properties. For very high Q/A, l needs to be reduced since the largest f can be is 1. This length also controls the speed of the thermoelectric cooler because the thermal time constant is proportional to l^2 [16]. Accordingly, small l devices, produced by thin-film techniques, have high heat flux capability and are exceedingly fast.

Because the heat is transported only from the cold side to the hot side of the thermoelectric cooler, the distance of heat transfer is also l. Thus, long (i.e., thicker), narrow thermoelectric elements [17, 18], can move heat a further distance, but only with low heat fluxes.

Several chip side TEC spot cooling systems have been investigated using exceedingly short thermoelectric elements made from materials that could be integrated with IC processing, such as thermoelectric micro-coolers, based on SiGe/Si or InGaAsP materials [19, 20]. Such structures have demonstrated high heat fluxes because of their short (< 5µm) thermoelectric elements. However, these devices have small ΔT_{max} (a few K), giving them low efficiency or coefficient of performance.

In contrast, $(Bi,Sb)_2Te_3$-based superlattice thermoelectric materials have shown a high figure of merit at the thin-film element level [21] and have produced devices with ΔT_{max} in the range of 55 K and ~65 K for hot-side temperature of 25° C and 75° C respectively [22].

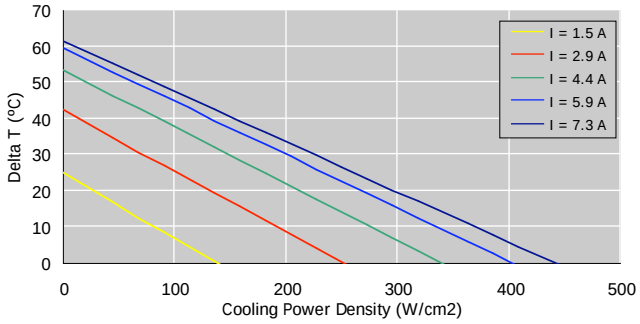

Figure 7. Predicted performance curves (Q" for various applied currents, I) of a 49-couple eTEC at T_h = 85C.

Figure 7 shows the potential for very high heat flux densities using an eTEC fabricated with thin-film $(Bi,Sb)_2Te_3$-based materials.

6.3. High *COP*

Cooling efficiency is characterized by the coefficient of performance (*COP*), defined by the ratio of the heat removed to the power input.

$$COP = \frac{Q}{IV} \tag{10}$$

For example, cooling 3 Watts with a solution having a *COP* of 0.5 requires 6 W of input power and results in a total of 9 W of heat to be rejected to the ambient. This additional heat must be managed not only at the chip level but also adds additional heat load to the system, rack, and data center thermal management systems. If the cooling solution has a *COP* of 2.0 then only 1.5 W of input power would be needed to cool the 3 W.

The coefficient of performance will depend on the operating conditions I and ΔT as well as S, K, and R. For maximum coefficient of performance, the electrical current is only a fraction of I_{max} [23]:

$$\frac{I_{max\,COP}}{I_{max}} = \frac{\Delta T}{T_c\left(\sqrt{1+ZT_{Avg}}-1\right)} \tag{11}$$

where $T_{Avg} = \frac{T_c + T_h}{2}$. At this optimal current for maximum *COP*, the *COP* (COP_{max}) is given by [23]:

$$COP_{max} = \frac{T_c}{\Delta T}\frac{\sqrt{1+ZT_{Avg}} - \frac{T_h}{T_c}}{\sqrt{1+ZT_{Avg}}+1} \tag{12}$$

The first term, $\frac{T_c}{\Delta T}$, is the thermodynamic maximum coefficient of performance obtainable (the coefficient of performance of a Carnot cycle).

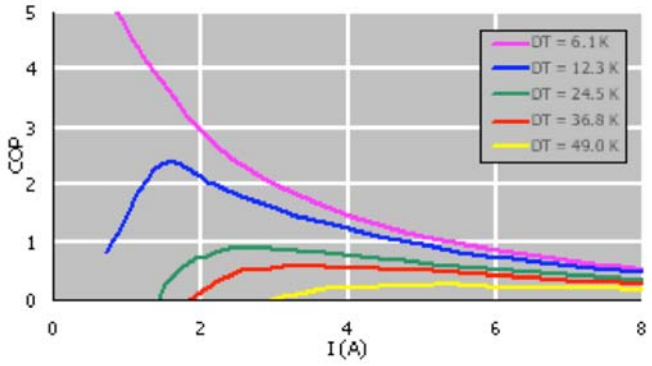

Figure 8. Predicted coefficient of performance (*COP)* curves for various operating currents, *I*, of the eTEC in Figure 2. $I_{max} \approx 8$ A. Plotted for constant T_h conditions.

As shown in Figure 8, the *COP* is largest when the current *I* in the device is substantially less than I_{max}. Thus, to achieve both high *COP* and high heat flux, a very high Q_{max}/A is required that enables the device to pump a high heat flux *Q* with a current well below I_{max}. The solution is to design for an exceptionally large Q_{max} (but operating at a smaller *Q*), by using small *l*, which naturally leads to the use of thin film thermoelectric devices.

Operation in the region of high *COP*, shown schematically in Figure 9, is accomplished by powering the device at a fraction of I_{max}. The inherent high heat flux capability of eTECs enables hot spot cooling by simultaneously meeting the criteria of high *COP*, Q_{max}/A, and ΔT. By adjusting the size (*A*), packing fraction (*f*), and thickness (*l*) the extent of the high *COP* range shown in Figure 9 can be adjusted.

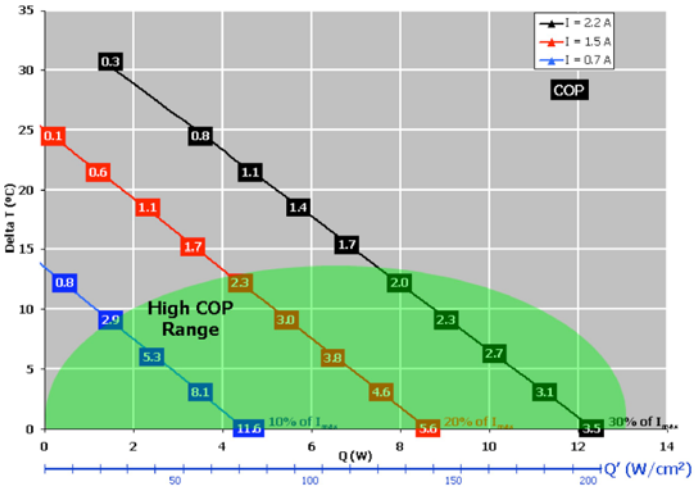

Figure 9. Predicted performance curves (*ΔT, Q and COP* for various applied currents, *I*) of a 49-couple eTEC at T_h = 85C.

6.4. Design Flexibility

The key to optimizing the hot spot cooling solution is to manipulate the size (*A*), packing fraction (*f*), number of couples (*n*) and thickness (*l*) of the thermoelectric cooler. Because each application may have different requirements, design flexibility to carefully tailor module parameters is a must. Since most hot spot cooling applications will require small *l*, microfabrication techniques are essential to achieve the necessary control.

In addition, the system power supply limits the voltage and current available to drive the TEC. Current and voltage requirements drive the design for the number of couples (*n*) in the thermoelectric cooler. As shown above, I_{max} is proportional to $1/n$ (given a fixed thermoelectric area *Af* divided into *n* couples) whereas Q_{max} is not. This allows for independent design of I_{max} and Q_{max}.

The small *l* necessary for hot spot cooling requires that the electrical contact resistance be very low, of the order 10^{-7} $\Omega \cdot cm^2$ [24] (for *l* of about 10 μm). This is because the relative contribution of the contact resistance increases as the resistance of the thermoelectric element (proportional to *l*) decreases. The effective figure of merit Z_{eff} compared to the figure of merit *Z* without contact resistance is given by:

$$Z_{eff} = Z \frac{1}{1 + \frac{2\rho_{contact}}{\rho l}} \qquad (13)$$

where $\rho_{contact}$ is the contact resistivity (Ωcm^2).

Microfabrication of the thermoelectric materials allows unprecedented control of couple geometry and number. In addition, thin film fabrication processes allow the precise surface preparation that is required to achieve these exceptionally low contact resistivities. By combining this capability with the use of engineered, nanostructured materials which promise to increase device efficiency through enhancement of *Z*, dramatic advances in TEC performance are possible. [21, 25].

6.5. Site-Specific Thermal Management

For hot-spot cooling, active, solid-state thermoelectric cooling devices must be positioned to absorb heat at the localized hot spot or after it has spread through the thickness of the silicon substrate and transport it to the heat spreader. Since only the hot spot is being actively cooled, the additional heat input from the cooling device is small compared to the total heat load.

In order to exploit this design principle, microfabrication techniques are essential. For each hot spot, there will be an optimum TEC dimension and heat pumping capacity that maximizes the cooling at the hot spot. These values are dependent on the hot spot size and heat output, as well as the amount of heat spreading that occurs in the substrate before reaching the TEC.

For example, given a 3W hot spot 0.5 x 0.5 mm, and a silicon substrate 0.75 × 0.75 mm, a simple assumption can be made that the heat will spread from the hot spot at an angle of approximately 45°, leading to a 2 × 2 mm hot spot on the back of the substrate with a heat flux of 75 W/cm^2 (ignoring the presence of any non-hot spot related background heat). Bulk TECs typically provide Q_{max}/A of less than 20 W/cm^2, far short of the heat flux necessary. With high thermal

conductivity substrates, higher Q_{max}/A values are possible [11]. However, even these devices do not meet both the high *COP and Q/A* requirements above.

If the TEC size is increased too much in order to increase total heat pumping, it will pump an unnecessarily large proportion of the background heat in addition to the heat from the hot spot, and its effectiveness cooling the hot spot will be reduced. Microfabricated eTECs can easily be fabricated with the optimum geometry, simultaneously meeting the requirement of small size and high heat pumping capacity. An example of an optimized eTEC design in shown in Figure 10.

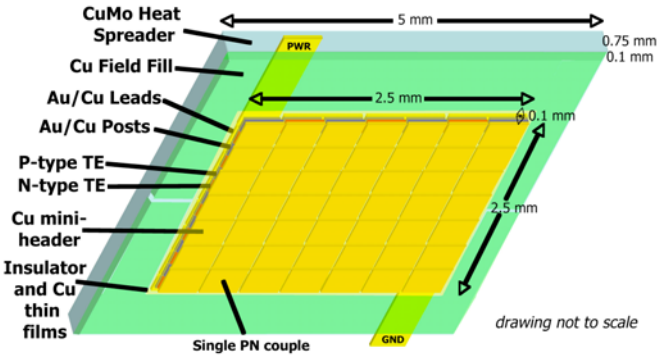

Figure 10. Proposed design of embedded thermoelectric cooler (eTEC) with 49 PN couples integrated into a heat spreader for mating onto an integrated circuit chip (view from below).

6.6. Integration into IC Package

An example of integrating the eTEC into an IC package is shown in Figure 11. The integrated circuit is flip-chip mounted face down to the substrate. A hot spot is shown in red in the middle of the active die area. A thin thermal interface material (TIM1) provides the thermal interface between the backside of the die and heat spreader, which forms the lid of the package. A second, less critical thermal interface material (TIM2) provides a thermal interface to the heat sink. Each of these components has already been optimized (including cost and manufacturing considerations) to minimize the thermal resistance.

The eTEC, about 100 µm thick, is mounted on the heat spreader and effectively recessed (using a metal field fill technique, Figure 10). The ability to fabricate an ultra-thin eTEC is critical to both the performance of cooler and the effectiveness of the existing thermal management solution. When recessed in this fashion, no increase in the TIM1 thickness is required, and the associated increase in thermal resistance is avoided. Further, the fully recessed, planarized surface of the eTEC requires no change in the method of TIM application or its subsequent assembly. Because the eTEC integrates unobtrusively into the existing thermal management solution, its impact on the existing components will be minimal.

In the region around the hot spot, the eTEC produces a small temperature inversion: where the temperature actually *rises* along the direction of heat flow. The eTEC pumps heat from the IC and TIM1 (lowering the temperature of these components) to the heat spreader (raising the temperature of the heat spreader, TIM2 and heat sink). Hence, the eTEC enhances heat conduction from the hot spot to the heat spreader with that tradeoff that somewhat more heat must be removed from the system. The additional heat is minimized by site-specific cooling of the hot spot, and by operating the eTEC in its high COP regime.

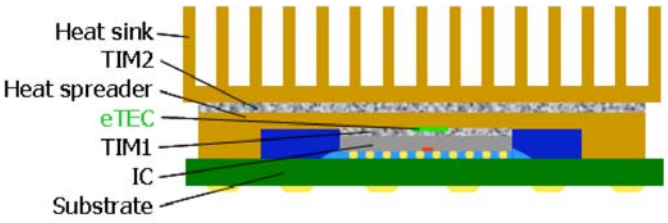

Figure 11. Flip-chip package with eTEC mounted on heat spreader

6.7. Reliability

It is critical that any thermal management solution be highly reliable. While solid-state thermoelectric devices have the advantage of no moving parts, they must be able to withstand large heat flux densities and thermal gradients. Thin-film superlattice films have been stressed with extended power and temperature on/off cycle testing and have shown little change in thermoelectric properties after over 50,000 cycles [26].

Because of the presence of large thermal gradients, careful consideration must be given to thermal expansion mismatches between the materials used in the eTEC integration and packaging. One way to mitigate these effects is to utilize a compliant interface between the eTEC, and the target device. Typically, compliant thermal interface materials are present between the substrate and the heat sink or thermal spreader. These materials will help reduce the thermal stresses on the eTEC.

7. Performance Modeling

For illustration, the following 1-dimensional model can be used to explain the utility of eTEC hot spot cooling.

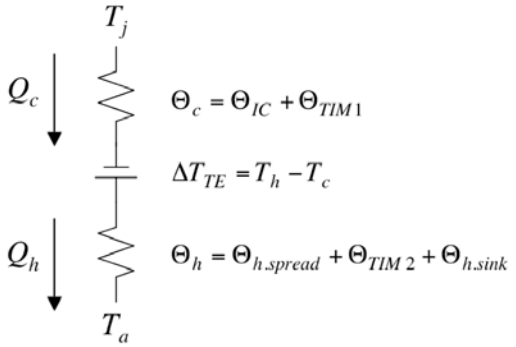

Figure 12. Thermal circuit used for schematic 1-dimensional model of hot-spot (at T_j) cooling with eTEC.

The temperature of the hot spot (junction T_j) is given by

$$T_j = Q_c\Theta_c - \Delta T_{TE} + Q_h\Theta_h + T_a \quad (14)$$

where Θ_c is the total, effective thermal resistance from the hot spot to the TEC, and Θ_h is the total, effective thermal resistance from the TEC to the ambient.

We wish to compare the hot spot temperature T_j without the eTEC ($\Delta T_{TE} = 0$ and $Q_h = Q_c$) to the case with the eTEC ($Q_h = Q_c + Q_c/COP$). The difference in hot spot temperature due to insertion of an eTEC (keeping the thermal resistance of the passive elements the same) is:

$$\Delta T_j = \Delta T_{TE} - \frac{Q_c}{COP}\Theta_h \quad (15)$$

Thus, the condition for hot spot cooling ($\Delta T_j > 0$) becomes

$$\Delta T_{TE} COP > Q_c\Theta_h \quad (16)$$

which again stresses the need for high *COP* and high ΔT coolers. Because there is more heat to reject behind the eTEC ($Q_h > Q_c$), there will necessarily be more ΔT drop needed across Θ_h with the eTEC than without. The above equation then explains that hot spot cooling will occur when ΔT_{TE} is greater than the additional ΔT needed across Θ_h to make up for the added heat.

Figure 13 shows schematically the temperature increase from ambient (T_a) to the junction temperature (T_j) due to the thermal resistance of the various components shown in Figure 11. Conduction moves heat from the hot junction through the IC, TIM1, heat spreader and TIM2, and convection moves the heat from the heat sink to the ambient air.

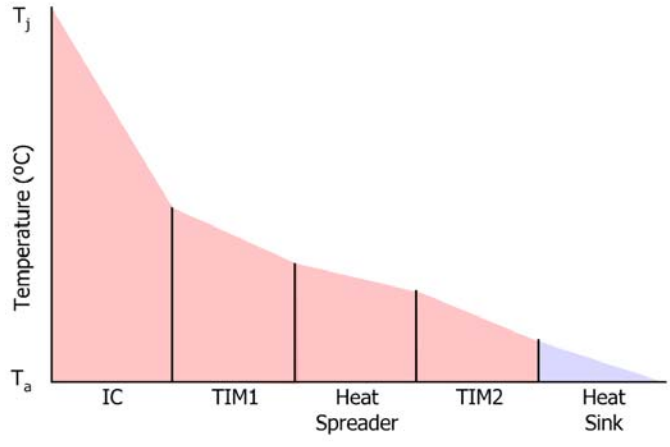

Figure 13. Schematic temperature profile from the hot-spot (T_j) to heat sink for a conventional thermal management solution (not to scale).

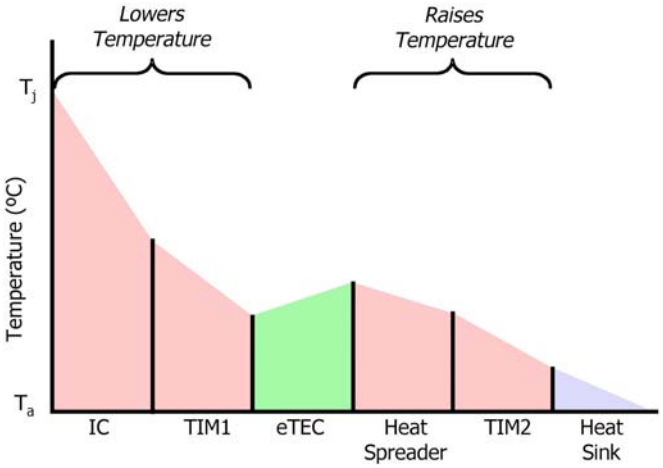

Figure 14. Schematic temperature profile from the hot-spot to heat sink using an eTEC in the thermal management solution (not to scale). Notice that the effective system thermal resistance (proportional to $T_j - T_a$) is reduced compared to Figure 13 even though the individual thermal resistances (slope) of the passive (red) components do not change.

In comparison, Figure 14 shows the effect of mounting the eTEC on the integrated heat spreader as shown in Figure 11. In the region around the hot spot, the eTEC suppresses the temperature by producing a temperature inversion across the eTEC that in turn leads to a lower T_j. This allows the effective system thermal resistance $\frac{T_j - T_a}{Q_c}$ to be lowered without changing the thermal resistance of the passive components (IC, TIM1, Heat Spreader, TIM2).

8. Conclusions

Localized heat flux on microprocessors produces hot spots that drive the design of thermal management solutions. Chip scale thermal solutions designed to keep hot spots below a critical temperature unnecessarily overcool the rest of the CPU and can dramatically increase the amount of the heat that must be rejected by the heat-sink. Active hot spot cooling solutions, provided by embedded thermoelectric cooling (eTEC), will enhance the thermal management, improving reliability and performance.

For site-specific hot spot cooling to work, both exceptionally high heat flux and high *COP* are required in a small thermoelectric device. This can only be achieved in a thin film thermoelectric device that utilizes the large increase in heat flux pumping capability, due to the small size of the thermoelectric elements, while maintaining a high *COP*.

Three key considerations must be taken to ensure sufficient performance of the eTEC. The first is to use high efficiency thermoelectric materials, the second to design and build small, thin thermoelectric devices, and the third is to unobtrusively integrate the eTEC into the existing thermal solution such that there is minimal thermal resistance surrounding the eTEC.

By taking advantage of wafer fabrication process technology, eTECs offer a performance-improvement roadmap resulting from improvements in both materials processing and device geometry. This unique combination makes eTECs ideally suited for the problem of hot spot cooling.

Acknowledgments

The authors would like to acknowledge the support of Matt Cauchy, the Engineering and Operations groups at Nextreme, Rama Venkatasubramanian and the Center for Thermoelectric Research at RTI.

References

1. "International Technology Roadmap for Semiconductors (ITRS)," Semiconductor Industry Association 2005.
2. Black, J. R., "Electromigration - A brief survey and some recent results," IEEE Transactions on Electron Devices, vol. ED-16, pp. 338-347, 1969.
3. Kim, N. S., et al., "Leakage current: Moore's law meets static power," Computer, vol. 36, pp. 68-75, 2003.
4. Mukhopadhyay, S., A. Raychowdhury, K. Roy, "Accurate estimation of total leakage current in scaled CMOS logic circuits based on compact current modeling," Proceedings 2003. Design Automation Conference (IEEE Cat. No.03CH37451), pp. 169-174, 2003.
5. Sauciuc, I., et al., "Thermal Devices Integrated With Thermoelectric Modules With Applications To CPU Cooling," Proceedings of InterPACK 2005, pp. 73243, 2005.
6. Fukutani, K., A. Shakouri, "Optimization Of Bulk Thermoelectric Modules For Chip Cooling Applications," Proceedings of InterPACK 2005, pp. 73410, 2005.
7. Bierschenk, J., D. Johnson, "Extending the limits of air cooling with thermoelectrically enhanced heat sinks," The Ninth Intersociety Conference on Thermal and Thermomechanical Phenomena In Electronic Systems (IEEE Cat. No.04CH37543), pp. 679-84 Vol.1, 2004.
8. Watwe, A., R. Viswanath, "Thermal Implications of Non-Uniform Die Power Map and CPU Performance," Proceedings of InterPACK 2003, pp. 35044, 2003.
9. Copeland, D., "64-bit server cooling requirements," Twenty First Annual IEEE Semiconductor Thermal Measurement and Management Symposium (IEEE Cat. No.05CH37651), pp. 94-98, 2005.
10. Prasher, R. S., et al., "Effect Of Localized Hotspot On The Thermal Performance Of Two-Phase Microchannel Heat Exchanger," Proceedings of InterPACK 2005, pp. 73087, 2005.
11. Semenyuk, V., "Thermoelectric micro modules for spot cooling of high density heat sources," Proceedings ICT2001. 20 International Conference on Thermoelectrics (Cat. No.01TH8589), pp. 391-396, 2001.
12. Snyder, G. J., et al., "Thermoelectric microdevice fabricated by a MEMS-like electrochemical process," Nature Materials, vol. 2, pp. 528, 2003.
13. Stark, I., M. Stordeur, "New micro thermoelectric devices based on bismuth telluride-type thin solid films," Eighteenth International Conference on Thermoelectrics. Proceedings, ICT'99 (Cat. No.99TH8407), pp. 465-472, 1999.
14. Bottner, H., et al., "New thermoelectric components using microsystem technologies," Journal of Microelectromechanical Systems, vol. 13, pp. 414-420, 2004.
15. Venkatasubramanian, R., et al., "Superlattice thin-film thermoelectric materials and devices," Thermoelectric Materials 2003 - Research and Applications Symposium (Mater. Res. Soc. Symposium Proceedings Vol.793), pp. 51-58, 2004.
16. Snyder, G. J., et al., "Supercooling of Peltier cooler using a current pulse," Journal of Applied Physics, vol. 92, pp. 1564-9, 2002.
17. Wijngaards, D. D. L., R. F. Wolffenbuttel, "Characterisation and modelling of planar on-chip integrated Peltier elements for highly localised thermal stabilisation and cooling," Eighteenth Annual IEEE Semiconductor Thermal Measurement and Management Symposium. Proceedings 2002 (Cat.No.02CH37311), pp. 105-112, 2002.
18. Solbrekken, G. L., "Peltier Enhanced Heat Spreading for Localized Hot Spot Thermal Management," Proceedings of InterPACK 2005, pp. 73471, 2005.
19. Wang, P., et al., "Thermoelectric Micro-Cooler For Hot-Spot Thermal Management," Proceedings of InterPACK 2005, pp. 73244, 2005.
20. LaBounty, C., A. Shakouri, J. E. Bowers, "Design and characterization of thin film microcoolers," Journal of Applied Physics, vol. 89, pp. 4059-4064, 2001.
21. Venkatasubramanian, R., et al., "Thin-film thermoelectric devices with high room-temperature figures of merit," Nature, vol. 413, pp. 597-602, 2001.
22. Venkatasubramanian, R., private communication.
23. Heikes, R. R., R. W. Ure, Thermoelectricity: Science and Engineering. New York: Interscience, 1961.
24. Krishnan, S., et al., "Towards a Thermal Moore's Law," Proceedings of InterPACK 2005, pp. 73409, 2005.
25. Harman, T. C., et al., "Quantum dot superlattice thermoelectric materials and devices," Science, vol. 297, pp. 2229-2232, 2003.
26. Alley, R., et al., "Reliability Testing of Thin-Film Superlattice Thermoelectric Device," Proceedings ICT'04. 23rd International Conference on Thermoelectrics, 2004.