

Differentiating the Roles of IR Measurement and Simulation for Power and Temperature-Aware Design

Wei Huang[‡], Kevin Skadron[‡], Sudhanva Gurumurthi[‡], Robert J. Ribando^{*}, and Mircea R. Stan[†]
Departments of Computer Science[‡], Electrical and Computer Engineering[†], Mechanical and Aerospace Engineering^{*}
University of Virginia, Charlottesville, VA 22904

Abstract

In temperature-aware design, the presence or absence of a heatsink fundamentally changes the thermal behavior with important design implications. In recent years, chip-level infrared (IR) thermal imaging has been gaining popularity in studying thermal phenomena and thermal management, as well as reverse-engineering chip power consumption. Unfortunately, IR thermal imaging needs a peculiar cooling solution, which removes the heatsink and applies an IR-transparent liquid flow over the exposed bare die to carry away the dissipated heat. Because this cooling solution is drastically different from a normal thermal package, its thermal characteristics need to be closely examined.

In this paper, we characterize the differences between two cooling configurations—forced air flow over a copper heatsink (AIR-SINK) and laminar oil flow over bare silicon (OIL-SILICON). For the comparison, we modify the HotSpot thermal model by adding the IR-transparent oil flow and the secondary heat transfer path through the package pins, hence modeling what the IR camera actually sees at runtime. We show that OIL-SILICON and AIR-SINK are significantly different in both transient and steady-state thermal responses. OIL-SILICON has a much slower short-term transient response, which makes dynamic thermal management less efficient. In addition, for OIL-SILICON, the direction of oil flow plays an important role by changing hot spot location, thus impacting hot spot identification and thermal sensor placement. These results imply that the power- and temperature-aware design process cannot just rely on IR measurements. Simulation and IR measurement are both needed and are complementary techniques.

1 Introduction

Non-ideal CMOS technology scaling results in ever-increasing power density, especially local power density. With the shift towards a multi-core and even many-core design paradigm, it is likely that the total power will also rise. This is caused by the increased number of cores, the improved circuit delay and higher frequency, the relatively constant die area due to manufacturing yield, and the slow supply voltage scaling. Since all the power consumed is eventually dissipated in the form of heat, the increase of both power density and total power present severe thermal challenges to chip designers. The dissipated heat has to be efficiently removed by a thermal package to prevent circuit malfunction and thermal runaway.

Thermal packages are expensive, especially for high-performance processors, making package design for the worst-case thermal scenario prohibitive. Instead, designers usually address the typical case. As a result, to prevent thermal emergencies, dynamic thermal management (DTM) [1, 16] has become

indispensable for modern thermally constrained chips. On the other hand, engaging DTM introduces performance penalties, so there is a tradeoff between the cost of the package and the DTM-induced performance degradation.

Obviously, accurate thermal characterization of the silicon die is crucial for the development of efficient thermal solutions. Compared to temperature estimations using design-time thermal models, runtime thermal characterization is preferable because it offers direct measurements with the chip running actual workloads in realtime. Recently, IR thermal imaging [7, 12] has been gaining popularity in both industry and academia for characterizing thermal behavior. It has been used to design DTM and temperature sensor placement and sampling rates. It provides unprecedented runtime real thermal measuring resolutions both in time and space. Unfortunately, IR thermal imaging also needs peculiar cooling configurations, because metal is not transparent to IR. For the infrared light emitted from the silicon to reach the IR camera, one technique is to expose the IR-transparent silicon die by removing package components, such as metal heatsink and spreader, and adding an IR-transparent liquid flow over the bare die to carry away the dissipated heat. This leads to drastically different transient behavior from a heatsink and its thermal characteristics need to be closely examined.

In this paper, we investigate the thermal characteristics of two different cooling configurations—forced air flow over a copper heatsink (AIR-SINK) vs. laminar oil flow over bare silicon die (OIL-SILICON). The former is now the prevailing cooling solution for high-performance processors, while the latter is generally used for real-time infrared (IR) thermal imaging of the silicon since silicon and certain oils are both transparent to IR. We modify the HotSpot thermal tool [16] to model OIL-SILICON and add the secondary heat transfer path that spans C4 bumps in flip-chip packages, package substrates, solder ball and printed-circuit board (PCB) for better accuracy.

This paper makes the following contributions:

1. We show that the presence or absence of a heatsink fundamentally changes the thermal behavior of the die, leading to significantly different temperature-aware design implications.
2. We modify HotSpot to model OIL-SILICON and the secondary heat transfer path.
3. We compare AIR-SINK and OIL-SILICON, and show that even with the same overall equivalent convective thermal resistance, the steady-state temperature distribution can be drastically different, in terms of both maximum temperature and temperature gradient.
4. We find OIL-SILICON has a much slower *short-term* transient response. As a result, both transient hot spot loca-

tions and the rate of transient temperature rise can be different between *AIR-SINK* and *OIL-SILICON*. On the other hand, *OIL-SILICON* has a faster *long-term* transient response, making the warmup phase from ambient to steady-state temperature shorter than *AIR-SINK*.

5. Additionally, for *OIL-SILICON*, we show that the flow direction has a significant impact on across-chip temperature distribution, especially when there is no heat spreader: steady-state hot spot locations can change with the direction of oil flow.
6. Due to different steady-state and transient responses, *AIR-SINK* and *OIL-SILICON* can lead to different choices of DTM parameters, such as engagement duration, thermal sampling frequency, thermal threshold, on-chip sensor placement scheme, and number of sensors needed, etc.

The purpose of this paper is to differentiate the roles of IR and simulation. The thermal design process cannot just rely on IR measurements. Simulation and IR measurement are both needed and are complementary techniques. Even after a chip is manufactured, the design of the software interface to the hardware thermal facilities must take into account the spatial and temporal differences created in IR measurement.

This paper is organized as follows. Section 2 provide some background and the related work. Section 3 shows the modifications of HotSpot to model the oil flow and the secondary heat transfer path. Section 4 characterizes and compares transient and steady-state thermal responses of *AIR-SINK* and *OIL-SILICON*. Section 5 presents architectural implications. Finally, Section 6 summarizes the paper and points out future work.

2 Background and Related Work

This section provides a taxonomy of different cooling mechanisms as well as a taxonomy of thermal characterization methods.

2.1 Chip Cooling Mechanisms

Depending on the cooling mechanism, popular thermal packages fall into one of the following categories:

- *Convective Cooling*: Either air or liquid can be used as coolant. The coolant flow can be forced or natural. Examples are forced air cooling over heatsink as in most desktop systems; natural convection for low-cost chips without a fan and a heatsink; forced water cooling in overclocked systems and some server systems; forced liquid cooling over bare silicon die for infrared runtime thermal characterization [7, 12]; microchannel cooling in more thermally-constrained chips [8], etc.
- *Phase-Change Cooling*: Here, heat is first absorbed to change the phase of the coolant (the liquid-to-gas phase change) and later removed by a heat exchanger to the ambient, and the coolant phase is changed back as a result of either the drop in temperature or the presence of a compressor (the gas-to-liquid phase change). There has to be a circulation mechanism to repeat the phase-change process. Examples are heatpipes in most laptop computers; droplet impingement spot cooling on hot chip surface [5]; refrigeration in data centers; phase-change heat spreader to achieve more uniform temperature across the die [15], etc.

- *Thermoelectric Cooling (TEC)*: Thermoelectric cooling uses the Peltier effect to pump heat from on-chip hot spots to cooler regions by consuming electric energy. Some PC overclockers use TEC to provide additional cooling for a CPU with another power supply unit.

In this paper we focus on convective cooling, leaving other modes as future work. In particular, we compare forced air cooling over a copper heatsink and forced oil cooling over a bare silicon die. These two cooling methods are used in either design-time or run-time thermal characterizations and the development of thermal management techniques.

2.2 Chip Thermal Characterization

Thermal characterization falls into two categories: design-time and run-time.

- *Design-time silicon thermal characterization* is usually achieved by simulations from full-chip thermal models. These models are especially useful in early design stages where the final design is not yet available. They also let the designer freely explore hypothetical designs without expensive prototypes. There has been abundant work on design-time full-chip thermal models, such as [11, 16, 17, 18]. However, existing studies do not provide flexibility on thermal package modeling, if thermal packaging is modeled at all. Usually, the only configuration is forced air convection with a heatsink. But there are now many other cooling configurations that are increasingly used in the high-performance computing community.
- *Run-time thermal characterization* is preferable because it offers direct measurements with chips running actual workloads, whereas design-time thermal models' accuracy is affected by many other factors such as power estimation, etc. Ideally, design-time models should be validated by real measurements. There are also existing studies on run-time thermal characterization methods. For example [10, 14] provide insights into using on-chip temperature sensors. Most modern high-performance processors, e.g. the IBM POWER series processors [6], have multiple integrated temperature sensors deployed across the entire die. However, due to the discrete nature of sensors, actual hot spots may be missed. Another runtime thermal characterization method is IR thermal imaging [7, 12]. While this technique can capture the detailed thermal map in real time, the cooling configuration is significantly different from the chip's thermal package in normal operation. For example, in [12], an AMD Athlon64 chip was cooled by an IR-transparent oil flow, whereas its actual package will be a heatsink with forced air convection. While the real-time IR thermal characterizations can be used to derive power consumptions of microarchitecture blocks as discussed in [12], this approach can be problematic if IR measurements are directly used to predict what happens when the die is put in its actual thermal package during normal operation. The limited sampling rate of the IR camera may also filter out high-frequency transient thermal fluctuations and miss thermal violations. Therefore, it is important to closely examine the runtime thermal characterizations. This also suggests that runtime thermal characterizations have limits in developing DTM policies for normal operation. Care must be taken when interpreting these measurements.

In this paper, we model more thermal package configurations such as liquid or air forced convection with different flow directions. Among them, one is of particular interest—direct oil cooling of the die without heatsink [12]. This is the only cooling configuration we can find in the literature that has the detailed information needed to make our analysis possible. By using an IR transparent oil that flows directly over the back of the die, frames of silicon transient temperature maps can be recorded by an IR thermal camera.

2.3 Dynamic Thermal Management

There also has been extensive work investigating microarchitecture thermal management techniques, for example [1, 13, 16], to just name a few. However, all these previous papers assume a fixed thermal package and do not consider how a different package choice would affect their techniques.

There are also a few studies that closely relate to our work. For example, IR thermal measurements have been used for the evaluation of general tradeoffs of different dynamic thermal management techniques [4]; and [9] uses IR thermal measurements to guide the thermal sensor placement and calibration. Readings from the thermal sensors are then used to extract useful information about on-chip variations. It would be interesting to see how the observations in this paper would affect the results in [4] and [9].

In this paper, with the improved thermal model, we show that the choice of thermal packages produces drastically different thermal response of different parts of the die. The thermal package alone can change the choice of key thermal management parameters such as DTM engagement duration, thermal sensing frequency and granularity, and on-chip thermal sensor placement. Therefore, being able to explore the tradeoffs among different package choices at design time is desirable. The research presented in this paper suggests another interesting dimension in the design space that chip architects can explore—the thermal package choice. Because of the distinct thermal behavior of the die caused by different packages as we show in the following sections, chip architects have yet another design knob to tune for temperature-aware microarchitecture design.

3 Modeling the Oil Flow

In order to investigate the impacts of the two different cooling mechanisms, we modify the HotSpot thermal modeling tool. One major extension is to model the IR-transparent oil flow over a bare silicon die, which is similar to the cooling setup in [12].

3.1 OIL-SILICON Model Details

The default thermal package configuration in HotSpot is forced convection over a copper heatsink, the heatsink is then attached to a copper heat spreader, which is attached to silicon die via a thin layer of thermal interface material. The secondary heat transfer path is not modeled, because negligible heat is removed through that path. However, if the spreader and heatsink are removed, and a laminar viscous IR-transparent oil flow with low thermal conductivity is directly applied to the back of the silicon die, as in the case of IR thermal measurements, the heat transferred through the secondary path is no longer negligible. Therefore, we added both the oil flow and the secondary heat transfer path, as shown in Fig. 1.

The oil flow is modeled as follows. The overall convection thermal resistance at the oil-silicon boundary can be found

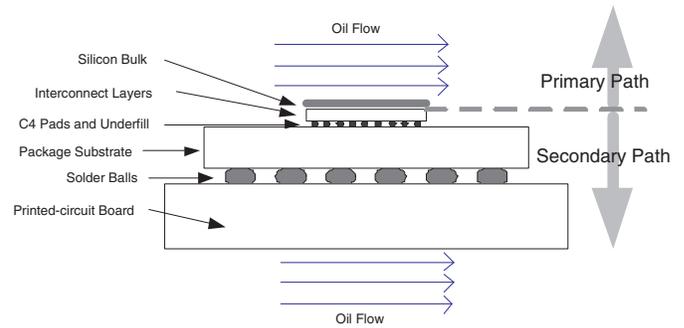


Figure 1. Modified model for oil flow and secondary heat transfer path.

by [3]:

$$R_{conv} = \frac{1}{h_L A_{chip}} \quad (1)$$

where h_L is the equivalent overall heat transfer coefficient, and A_{chip} is the entire silicon area. Furthermore, h_L can be calculated for laminar flow over a smooth flat surface:

$$h_L = 0.664 \frac{k}{L} \text{Re}_L^{0.5} \text{Pr}^{1/3} \quad (2)$$

Here L is the length of silicon along the flow direction, k is the thermal conductivity of the oil, Pr is the Prandtl Number of the fluid, and Re_L is the overall Reynolds Number of the flow.

The transient effect of the oil flow also needs to be modeled. The overall effective thermal capacitance of the oil can be calculated by:

$$C_{conv} = \rho c_p A_{chip} \delta_t \quad (3)$$

where ρ is the density of the oil, c_p is the specific heat of the oil, δ_t is the equivalent thermal thickness of the oil that contributes to heat transfer.

Close to the oil-silicon interface, there is a thin layer of oil that is relatively hotter than the rest of the oil flow. This is governed by the continuity across the oil-silicon boundary. Oil temperature at the boundary is the same as silicon temperature, and gradually decreases over a thickness δ_t to the ambient temperature of the free flow. According to [3], the thermal boundary layer thickness (δ_t) can be calculated by

$$\delta_t = \frac{4.91L}{\text{Pr}^{1/3} \sqrt{\text{Re}_L}} \quad (4)$$

We also add the secondary heat transfer path, which includes on-chip interconnect layers, C4 bumps, package substrate, solder balls and printed-circuit board (PCB). These layers are added in the same way as the existing layers (i.e. silicon, interface material, heat spreader and heatsink), with the PCB layer connected to another oil layer modeled similar to the oil layer over the silicon, as illustrated in Fig. 1.

3.2 OIL-SILICON Model Validation

We validate the oil flow model with ANSYS¹, a finite-element commercial software package that includes detailed computational fluid dynamics analysis. Notice that ANSYS and HotSpot

¹<http://www.ansys.com/>

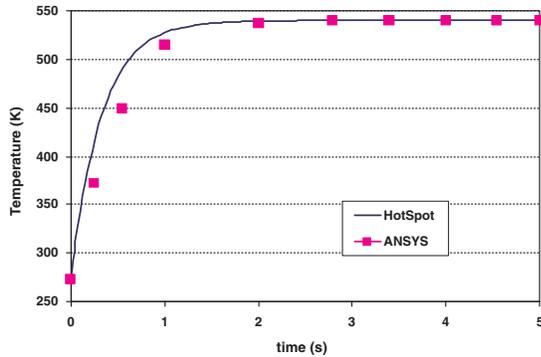


Figure 2. Comparison of transient responses of ANSYS and the modified HotSpot with a 10m/s oil flow over 20mm×20mm×0.5mm silicon dissipating 200W power step uniformly.

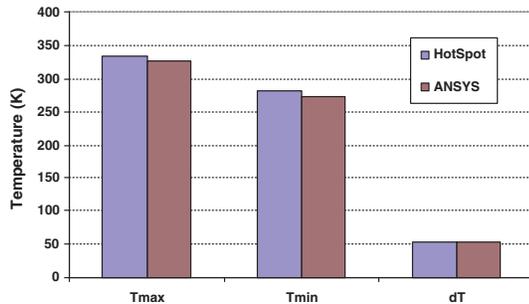


Figure 3. Comparison of steady-state responses of ANSYS and modified HotSpot with a 10m/s oil flow over 20mm×20mm×0.5mm silicon, a center area of 2mm×2mm dissipates 10W.

are independent from each other, and ANSYS is intrinsically more accurate yet less efficient. We put a silicon chip of size 20mm×20mm×0.5mm in an oil flow of 10m/s. We apply a 200W power step at time zero uniformly across the die and probe the temperature at the chip center. Fig. 2 compares the transient thermal response at the center of the heat source. The equivalent convection thermal resistance is about 1.0K/W. As can be seen, the time it takes the silicon to reach steady-state are quite similar in both cases, indicating that our transient model is quite accurate. Notice the thermal time constant is on the order of a second.

To further validate the steady-state response of the oil thermal model, we use the same settings as in Fig. 2, except that we reduce the heat source to 2mm×2mm and 10W at the center of the die. This way, we create greater spatial temperature gradient across the die. Fig. 3 shows the comparison of our oil flow model with ANSYS for on-die maximum temperature (T_{max}), minimum temperature (T_{min}) and temperature difference (dT).

We do not use direct IR measurements to validate the oil thermal model for two reasons: an IR setup is extraordinarily expensive; and our analysis (presented later in this paper) suggests

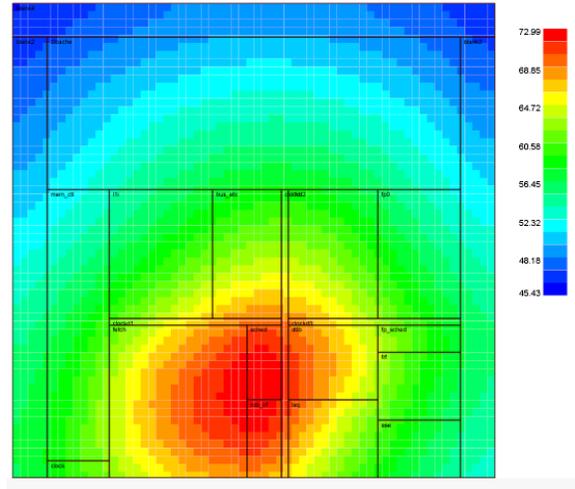


Figure 4. Steady-state responses of an AMD Athlon processor from the modified HotSpot, using average power values derived in [12].

that IR misses some fundamental phenomena. However, we perform a qualitative validation by comparing the steady-state temperature maps predicted by the model and the IR measurements from [12] for an AMD processor. The floorplan is derived from the die photo of the processor, and the power numbers are extracted from [12]. We compare model results with a typical IR thermal snapshot in [12]. The hottest temperature at the “Sched” block is about the same as that in [12] (73°C vs. ~70°C), the coolest temperatures are also close (45°C vs. ~45°C, excluding the blank area on the edges as in [12]). Notice that we assume uniform power per unit as opposed to the detailed IR image in [12]. The secondary heat transfer path model is also validated in this case, since it is included in the IR measurements. Fig. 4 shows the simulated temperature map of the AMD Athlon processor.

To isolate the effect of the secondary heat transfer path with the oil flow, Fig. 5(a) shows that modeling the secondary path is necessary because a significant portion of heat is removed through the secondary path to the oil flow. Without the secondary path, the predicted temperatures are much higher (over 10° for the AMD Athlon processor). On the other hand, for forced air cooling with a copper heatsink, Fig. 5(b) shows that adding the secondary path does not make a noticeable difference (less than 1%). This is because almost all heat is transferred through the primary path, which has a much lower thermal resistance.

4 Distinct Thermal Responses

With the modified thermal model, we investigate the impacts of *AIR-SINK* and *OIL-SILICON* on both transient and steady-state thermal responses of the the same silicon die.

4.1 Transient Response

There are two types of transient responses that are interesting—the warmup phase when everything starts to heat up from ambient temperature; and the fast oscillations around steady-state where everything has reached a fairly stable temperature.

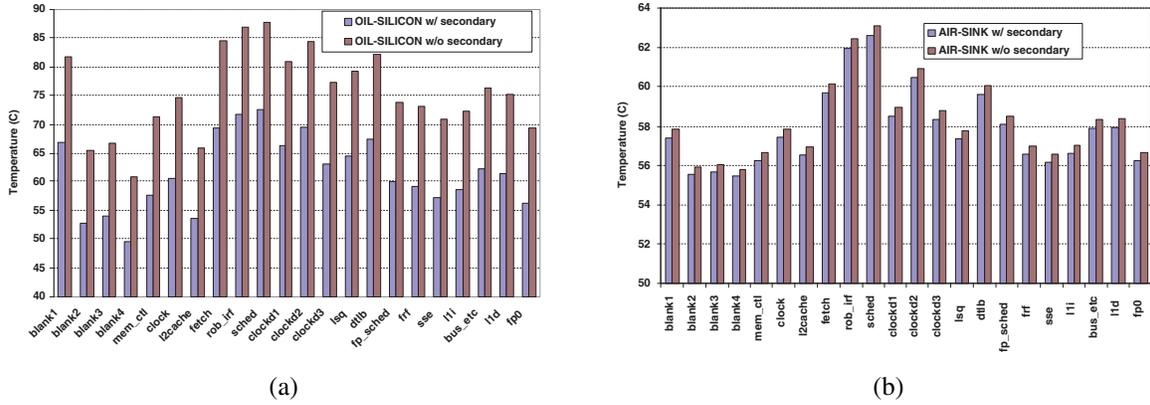


Figure 5. (a) Modeling the secondary heat transfer is necessary in the case of *OIL-SILICON* where heat removed via secondary path is significant, otherwise leading to significant errors. (b) On the other hand, in the case of *AIR-SINK*, the impact of the secondary path is negligible.

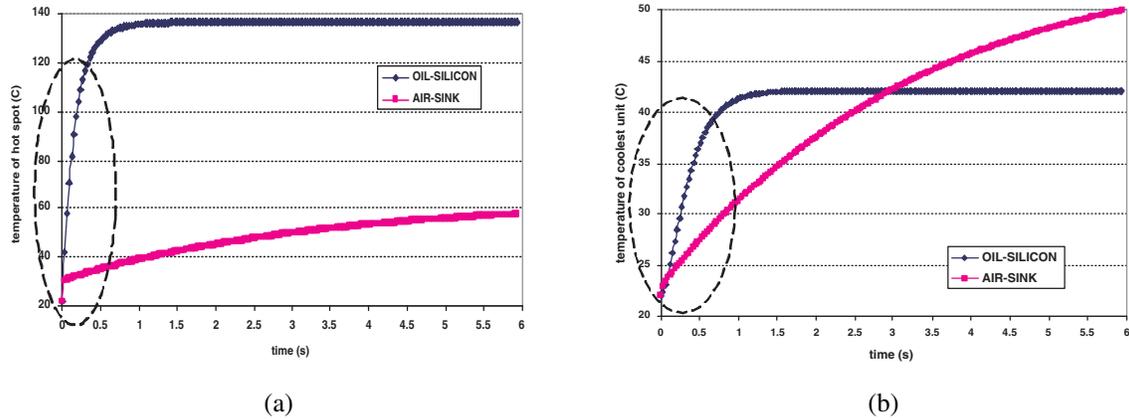


Figure 6. Warm-up transient silicon temperatures of (a) hot spot and (b) cool spot with *OIL-SILICON* and *AIR-SINK*. *OIL-SILICON* shows faster transient thermal time constant during warmup. Notice that for *AIR-SINK*, the steady-state temperature at the cool block is actually higher than *OIL-SILICON*. The cross-die average temperatures of the two cases are not far away (56°C for *AIR-SINK* and 62°C for *OIL-SILICON*), because the two cases have the same R_{conv} . Also notice there is an instant “jump” at the beginning for *AIR-SINK*, indicating two distinct time constants at different time scales.

4.1.1 Long-term transient response during warmup

OIL-SILICON and *AIR-SINK* have distinct transient responses relative to each other. Fig. 6 shows the warmup phase of both cooling configurations. Both of them have the same equivalent overall convection thermal resistance ($R_{conv} = 1.0\text{K/W}$). For illustrative purposes, we apply power for about 6 seconds duration to one hot block that occupies a small area of the die, whereas the other blocks in the die have no power consumption. The power density is 2.0W/mm^2 . We look at the temperatures of the hottest block and the coolest block. For the long-term response, *OIL-SILICON* reaches steady-state temperature much faster than *AIR-SINK*. This can be easily explained by the fact that the copper heatsink and copper heat spreader have a much larger thermal capacitance than the thin layer of IR-transparent oil, hence a longer thermal RC time constant during warmup. Another observation from Fig. 6 is that, although the two configurations have the same effective convection thermal resistance (R_{conv})

and hence about the same average chip temperature (62°C vs. 56°C), the steady-state hot spot temperature of *OIL-SILICON* is much hotter than that of *AIR-SINK* (137° vs. 63° in steady state). *OIL-SILICON* also has a lower steady-state temperature at the coolest unit than *AIR-SINK* does (42° vs. 55°). This is because the lateral heat spreading is much better in copper. The low thermal conductivity of the oil leads directly to higher on-chip temperature gradient.

4.1.2 Short-term transient response

The short-term transient thermal response is of more interest, because for high performance systems that do not shut down frequently, the systems spend most of their time oscillating with relatively high frequencies around some stable operating points. For this experiment, we use the same floorplan as in the experiment of Fig. 6. In the power trace, we apply the power to the hot block for 15ms and then turn the power off for 85ms. Assuming

this power on and off phases repeats periodically, we can use the average power derived from this power trace to solve for steady-state temperatures, and use them as the initial temperatures for the transient thermal simulation. Fig. 8 shows the short-term transient responses of *OIL-SILICON* and *AIR-SINK*. As can be seen from the figure, the short-term transient responses of the two cooling configurations are significantly different. The heat-up and cool-down phases of *OIL-SILICON* is more linear than those of *AIR-SINK*. A more significant difference is that it takes much longer for *OIL-SILICON* to cool down.

The reason for this phenomenon is again the difference in thermal capacitances between the two cases. For *AIR-SINK*, since the thermal capacitance of heatsink is orders of magnitude higher ($\sim 250\times$) than that of silicon due to the heatsink size, the short-term heat pulse (i.e. with high frequency) can only heat up the silicon, leaving the temperature of the heatsink almost unchanged. These lead to the fact that local short-term thermal time constant is mainly determined by the silicon thermal time constant (see Fig 7(a)):

$$\tau_{short,sink} = R_{th,Si}C_{th,Si} \quad (5)$$

On the other hand, in *OIL-SILICON*, the thin thermal layer of oil is very tiny in volume (about $100\mu\text{m}$ thick for a 10m/s oil flow), so its thermal capacitance is much smaller even compared to that of silicon! So for very short power pulses (i.e. very high frequency), the silicon temperature remains almost constant. For all other time scales, thermal capacitance of the oil can be neglected due to its small value, and the thermal resistance of the silicon can also be neglected because it is usually much smaller than the convection thermal resistance. Therefore, as shown in Fig. 7 (b), the only dominant thermal time constant is

$$\tau_{all,oil} = R_{conv}(C_{th,Si} + C_{th,oil}) \approx R_{conv}C_{th,Si} \quad (6)$$

From this analysis, clearly, there are two phases for the case of *AIR-SINK*—the short-term (milliseconds or less) thermal response is determined by $R_{th,Si}C_{th,Si}$; the long-term (seconds to minutes) thermal response is determined by $R_{conv}C_{sink}$. In the case of *OIL-SILICON*, both short-term and long-term thermal responses are determined by $R_{conv}C_{th,Si}$. Comparing Eqn. (5) and Eqn. (6), because R_{conv} is two orders of magnitude greater than $R_{th,Si}$ (e.g. 1.042K/W vs. 0.0125K/W in our setup), the *short-term* thermal time constant of *OIL-SILICON* is much longer than that of *AIR-SINK*.

It is well known that the step transient response of an RC ladder is exponential. Due to the greater short-term time constant for *OIL-SILICON*, for the time duration of milliseconds, the slow exponential curve looks linear locally. The asymmetry between the heat-up and cool-down phases in *OIL-SILICON* can be explained by the location of the initial temperature on the exponential curve. In Fig. 8, the initial temperature happens to be closer to the origin of the exponential curve, therefore, the heat-up phase is steeper whereas the cool-down phase is at the tail of the exponential curve and thus slower and more flat.

Based on the above analysis, another interesting experiment is to observe scenarios where the transient hot spot location changes. For example, if we run the transient simulation for a processor similar to Alpha EV6 from the steady state, and apply 2W first to the IntReg for 10ms and no power on FPMaP, then after 10ms, IntReg is turned off and FPMaP starts dissipating 2W. Fig. 9 shows that at 14ms, FPMaP becomes the new hot spot in *AIR-SINK*, whereas in *OIL-SILICON*, IntReg still remains as hottest spot. This can be explained by the fact that *AIR-SINK* has

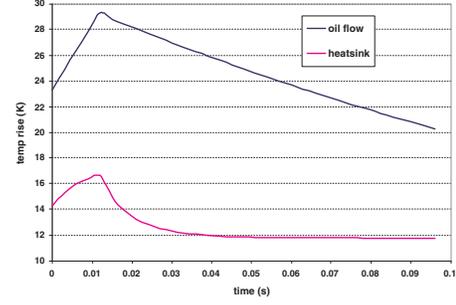


Figure 8. Short-term transient silicon temperatures of the hot spot with *OIL-SILICON* and *AIR-SINK* after the warmup phase. Also notice the asymmetry between heat-up and cool-down for *OIL-SILICON*.

a faster short-term transient response, hence the cooling down of IntReg and heating up of FPMaP happen faster.

To summarize the transient thermal responses—*OIL-SILICON* has faster long-term response but slower short-term response than *AIR-SINK*. These differences have a big impact on the DTM decisions and on-die thermal sensing and sensor placement, as we will see in Section 5.

4.2 Steady-State Response

The major difference on steady-state thermal responses between *OIL-SILICON* and *AIR-SINK* is the absolute temperatures and across-die temperature gradient. Fig. 10 shows the steady-state temperature maps of a processor similar to Alpha EV6 running *gcc* benchmark with *OIL-SILICON* and *AIR-SINK*, respectively. As can be seen, *OIL-SILICON* has about 30 degrees hotter maximum temperature and about 55 degrees higher across-die temperature difference. This is caused by the much less lateral heat spreading in *OIL-SILICON* as a result of the absence of copper spreader and heatsink.

Another factor we have not considered in *OIL-SILICON* so far is the oil flow direction. Oil flow direction is important because the local convection thermal resistance along the flow over silicon die is a strong function of location, i.e.,

$$R_{conv,local} = \frac{1}{h(x)A_{local}} \quad (7)$$

Here $h(x)$ is the local heat transfer coefficient, where x is the distance along the oil flow from the chip edge, and can be expressed as:

$$h(x) = 0.332 \frac{k}{x} \text{Re}_x^{0.5} \text{Pr}^{1/3} \quad (8)$$

Basically, $h(x)$ is highest at the leading edge of the die, and decreases along the flow direction. This translates into lower convection thermal resistance at the leading edge of the die, and higher convection thermal resistance along the way. Therefore, heat generated from units that are close to the leading edge is removed more efficiently. The significance of considering oil flow direction is demonstrated in Fig. 11. Because the hottest unit is usually IntReg, which is on the top edge of the chip, an oil flow from top to bottom (last column in Fig. 11) is the most effective way to cool down this hot unit. In fact, IntReg is cooled so well that it is no longer the hottest unit. The new hottest unit becomes

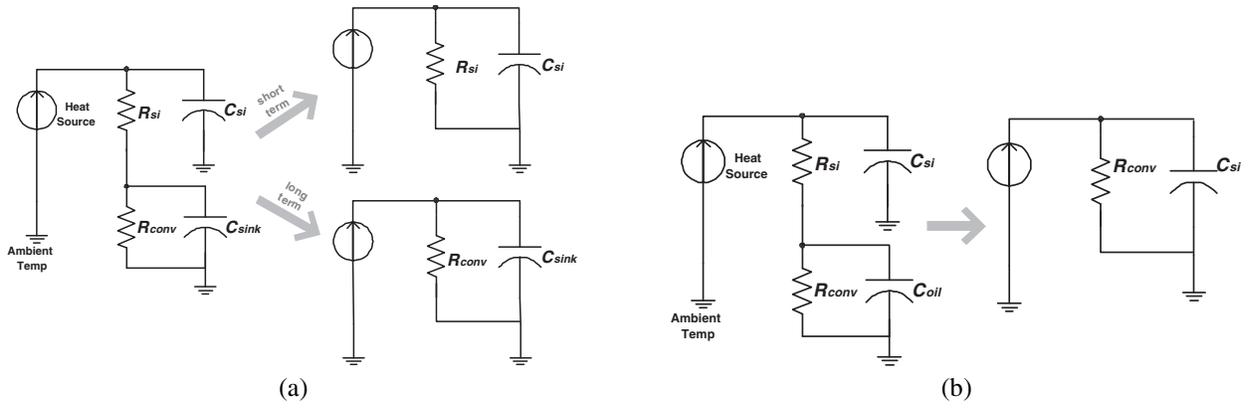


Figure 7. (a) Equivalent thermal circuits of *AIR-SINK* and its simplified versions for short-term and long-term thermal responses; (b) Equivalent thermal circuits of *OIL-SILICON* and its simplified version.

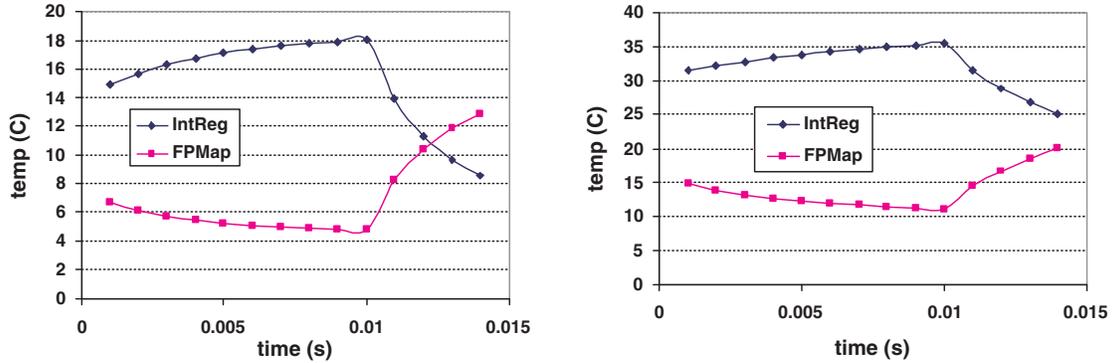


Figure 9. Left: (*AIR-SINK*) after the power switching, FPMaP becomes hottest spot at 14ms; Right: (*OIL-SILICON*) IntReg remains the hottest spot at 14ms after the power switching.

Dcache, which is further away from the leading edge of the oil flow.

In the case of *AIR-SINK*, the impact of the direction of air flow is negligible. This is because the temperature distribution inside the heatsink is quite uniform thanks to the high thermal conductivity and good lateral spreading of copper. Additionally, in most forced air cooling packages, the fan is placed on top of the heatsink, providing an impinging flow that is very uniform on the heatsink surface, resulting in more uniform heat transfer coefficient over a surface that is already close to isothermal.

5 Implications on Architectural DTM

The impacts of *OIL-SILICON* and *AIR-SINK* on silicon thermal responses have been demonstrated in Section 4. For the same chip running the same workload, they show drastic difference in both transient and steady-state temperatures. If IR or other runtime measurements are being used in the design process, being aware of the differences during the development of DTM techniques and on-chip sensor placement is thus crucial for the design to be efficient and free of thermal hazards. Based on the observations in Section 4, we list various aspects of the temperature-aware microarchitecture design that should be carefully considered.

5.1 DTM engagement duration

Depending on how fast a functional unit or a processing core responds to a DTM policy, the duration of the DTM engagement is chosen to minimize performance penalties. For example, *AIR-SINK* has a faster transient response than *OIL-SILICON*. Therefore, DTM should be engaged for a shorter duration for a chip cooled by *AIR-SINK* than the same chip cooled by *OIL-SILICON*. Figs. 12(a)-(b) are the simulated temperature trace snippets from an Alpha EV6 running *gcc*. We use SimpleScalar² with integrated Watch [2] and the modified HotSpot for the simulations.

Fig. 12(a) shows the case of *AIR-SINK*, where the overall thermal convection resistance (R_{conv}) is 0.3K/W; Fig. 12(b) shows the case of *OIL-SILICON*, where the overall thermal convection resistance at the oil-silicon interface is artificially set, for comparison purposes, to 0.3K/W as well³; Ambient temperature is

²<http://www.simplescalar.com/>

³We think a fair way of comparison is to use the same case-to-ambient thermal resistance for the two systems. Bringing the peak die temperatures of air-sink and oil-silicon to the same level would be a proper approach if *OIL-SILICON* is used in actual products. But since it is mainly used in an IR measurement setup that emulates *AIR-SINK*, it is more convincing to use the same case-to-ambient thermal resistance (R_{conv}) for a fair comparison of the two systems.

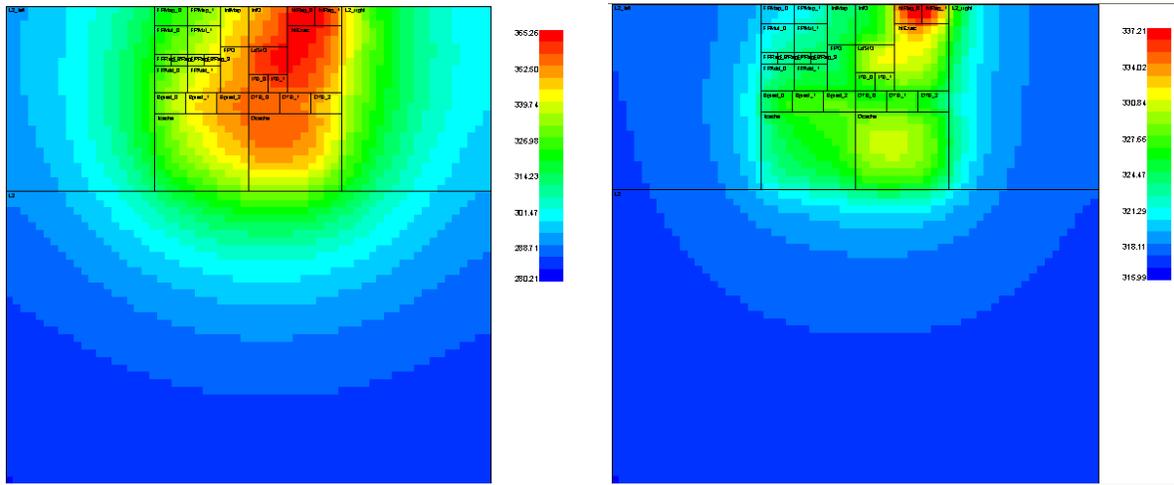


Figure 10. Left: steady-state thermal map of EV6 for *OIL-SILICON*; Right: steady-state thermal map of EV6 for *AIR-SINK*. Notice the color scales are not the same.

units	Oil Flow Directions			
	left to right	right to left	bottom to top	top to bottom
L2_left	41.93	44.39	44.61	41.84
L2	42.05	42.04	41.68	42.29
L2_right	46.32	42.7	46.65	42.38
Icache	71.61	75.93	79.67	66.6
Dcache	96.02	88.81	100.48	82.37
Bpred	73.11	79.03	82.88	66.42
DTB	76.36	72.79	83.81	60.63
FPAdd	63.53	68.77	72.8	56.06
FPReg	60.04	64.64	68.42	52.51
FPMul	56.45	60.21	63.99	48.41
FPMap	50.16	52.81	56.02	42.94
IntMap	65.3	65.46	73.34	49.29
IntQ	72.02	68.98	80.21	49.92
IntReg	104.91	97.85	112.37	67.9
IntExec	95.34	86.24	102.69	67.85
FPQ	67.45	67.92	76.37	53.31
LdStQ	90.79	87.28	99.83	68.84
ITB	77.35	74.55	85.76	59.68

Figure 11. Steady-state temperatures of EV6 with four oil flow directions. Notice the differences in temperatures and the change of hottest spot from IntReg to Dcache in the last column.

set to a typical 45°C in all cases. Notice that only the top five hottest blocks' temperature traces are plotted. The temperatures of the hot units in *OIL-SILICON* are much higher than *AIR-SINK* because of the very high local power densities, the absence of copper spreader and heatsink, and the low thermal conductivity of the oil. However, the overall average chip temperatures of the two cases are still about the same. This is because the L2Caches that occupy most of the die area is cooler in *OIL-SILICON*, which balances the impact of the smaller hotter units in the core.

There are several observations from Fig. 12: (1) As shown before, the heat-up and cool-down phases of *AIR-SINK* are significantly shorter than *OIL-SILICON* ($\sim 3\text{ms}$ vs. much more than 15ms). This confirms our earlier derivation that *OIL-SILICON*

has a much slower transient response. Also notice that 3ms is typically shorter than IR camera's sampling interval, therefore, IR thermal measurements could miss thermal emergencies within that time scale. In Fig. 12(a), the processor with *AIR-SINK* spends more times in phases where temperature is almost constant, whereas for *OIL-SILICON*, the processor spends most of time in the transient phases, meaning it takes longer time to bring the processor out of potential thermal emergencies in *OIL-SILICON*. Clearly, DTM techniques are more efficient in *AIR-SINK* than *OIL-SILICON*, and shorter engagement duration is preferred. A quantitative DTM performance comparison is almost impossible because the transient temperature scales of *AIR-SINK* and *OIL-SILICON* are so different. Reducing R_{conv} in *OIL-SILICON* might be a viable option to bring the peak temperatures to the same scale as that of *AIR-SINK*, but that only works for this particular simulated phase of *gcc*, so it is still not a fair comparison. Therefore, we only provide qualitative comparisons. (2) The hottest spot in *AIR-SINK* is more distinct (IntReg) than in *OIL-SILICON* with the same R_{conv} . It is hard to identify which unit is hotter in the latter case. This is because in *OIL-SILICON*, the vertical heat transfer from silicon to oil is not efficient due to the low thermal conductivity of the oil, making the lateral heat transfer within silicon a comparable heat transfer path to the vertical transfer path, heating up the closer neighboring units and making them equally hot. On the other hand, in *AIR-SINK*, most of the lateral spreading happens inside the copper spreader and heatsink, making the temperatures of units in silicon more distinct. (3) In *OIL-SILICON* the BPred is cooler than other shown units most of the time, whereas in *AIR-SINK*, due to the fast transient change, the BPred can sometimes be hotter than other units, which confirms the results shown in Fig. 9.

5.1.1 Cooling configuration for IR measurements of high-performance/high-power chips

As a side note, from Fig. 12 we can see that for a high-performance chip with high power density units, like Alpha EV6, the hot spot temperature is prohibitively high for the oil cooling configuration, even for a low $R_{conv} = 0.3\text{K/W}$ (the oil flow speed would be an unrealistic 100m/s to achieve 0.3K/W con-

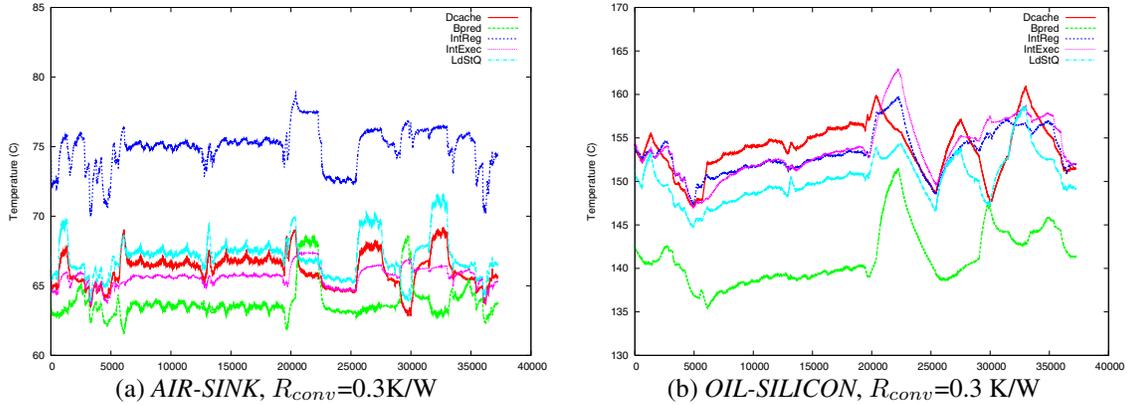


Figure 12. Simulated temperature traces of Alpha EV6 running *gcc*. Only the hottest five blocks are plotted. The cross-die average temperature of the two cases are about the same, due to the cooler L2Cache in *OIL-SILICON* that balances the hottest units plotted here. L2Cache occupies most of the die area. (a) *AIR-SINK* with $R_{conv} = 0.3\text{K/W}$; (b) *OIL-SILICON* with $R_{conv} = 0.3\text{K/W}$; Ambient temperature is 45 degrees Celsius in this case. The high hot spot temperatures for *OIL-SILICON* is just for illustrative purposes. In reality, *OIL-SILICON* has to be cooler and should have a lower R_{conv} , this can be achieved by introducing other cooling mechanisms such as thermoelectric cooling (transient response would be different in that case). The x-axis is number of simulated samples of temperature, samples are taken every 10K clock cycles, corresponding to about $3.3\mu\text{s}$ in our simulation setup.

vective resistance). In reality, for such chips, additional cooling mechanisms other than only the oil flow (e.g. thermoelectric cooling or fast coolant at extremely low temperature) might be necessary to further reduce R_{conv} and hence the hot spot temperatures. In that case, since R_{conv} is lower, the short-term thermal time constant would be also shorter, leading to yet different transient thermal responses from Fig. 12(b). Those cooling configurations also need thorough thermal characterizations, which is beyond the scope of this paper.

5.2 Thermal sensing frequency

In terms of how frequently the on-die sensor should measure the temperature and how fast DTM should respond to a possible thermal emergency, *AIR-SINK* and *OIL-SILICON* are similar. Although the heat-up transient response in *AIR-SINK* is faster, *OIL-SILICON* has a higher base temperature. Therefore, although on a relative scale, *OIL-SILICON*'s transient response is slower, its absolute rate of change is still comparable with *AIR-SINK*. According to Fig. 12(a) and Fig. 12(b), in both cases, IntReg's temperature can increase about 5 degrees in 3ms. If the desired resolution is 0.1 degrees, this leads to a sampling interval of at most $60\mu\text{s}$.

On the other hand, if higher oil flow speed is applied in *OIL-SILICON* to bring down the peak transient temperature to comparable levels as *AIR-SINK*, *OIL-SILICON* needs less frequent temperature sensing because its thermal change rate is slower. Notice that less frequent sensing does not mean less DTM performance penalty, since sensing does not necessarily trigger DTM. Only the DTM engagement duration after the trigger affects performance.

5.3 Thermal sensing granularity

OIL-SILICON has greater cross-die temperature difference (Fig. 10), meaning steeper temperature gradient across the die. In this case, if a thermal sensor is placed off the hot spot location, the sensor error can be significant due to the large gradient from

the actual hot spot to the sensor location. This implies more on-chip temperature sensors are needed. On the other hand, if the same number of sensors are deployed for *AIR-SINK* and *OIL-SILICON*, then the latter needs to put a larger margin on sensor error to account for the bigger accuracy impact with a misplaced sensor. This extra margin in turn will lower the DTM triggering threshold, making more frequent DTM engagements than necessary, hence more performance penalty.

This also implies that if the on-chip thermal sensor placement is determined based on IR thermal measurements, more sensors than necessary may be deployed. Putting more on-chip thermal sensors adds extra hardware complexity, area and power overheads.

5.4 Sensor placement considering oil flow direction

In Section 4.2, we also consider the possibly large impact of oil flow direction on hot spot location in the *OIL-SILICON* scenario. This phenomenon is especially important during on-chip thermal sensor placement. If sensor is placed without considering flow direction, the real hottest spot can be missed. For example, for a top-to-bottom flow over the Alpha EV6 chip with IR-transparent oil, the actual hot spot is the Dcache. If no flow direction is assumed, the steady-state hottest spot would always be IntReg. Therefore, for a chip with only a few on-chip sensors available, placing the sensor at IntReg would miss possible thermal emergencies at Dcache.

On the other hand, if the thermal sensor placement for a chip cooled by *AIR-SINK* is determined by IR thermal measurements, and if the oil flow is from top to bottom during IR measurements, the sensor would be placed at DCache, whereas in normal operation, the hottest spot is usually IntReg. This placement could lead to missing the actual hot spot and thus a thermal emergency.

Adding more sensors would be another option, but there are several difficulties: sensors are hard to be inserted into dense array structures such as caches and register files; sensors don't measure the *in-situ* hot spot temperatures and they also dissipate

heat; the speed of the sensor might limit the sampling rate; calibration issues; area overhead, etc. We think a proper way is to combine IR and sensor measurements and thermal modeling to achieve a better thermal design.

Existing studies by Hamann et al. [7] and Renau et al. [12] take IR measurements of a chip and use them to reverse engineer the power map of the chip. In this case, assuming we have a multi-core chip, and each core is dissipating similar amount of power—under an IR camera that captures the thermal map of the chip with an oil flowing left to right across the die, the cores on the right side of the die appears hotter, which results in an artifact of higher reverse-engineered power consumption for those cores. Therefore, during the mapping from IR temperature measurements to detailed power estimations, the factor of oil flow direction still needs to be taken into consideration. As a matter of fact, Hamann et al. have taken the flow direction into consideration for more accurate power extraction in their work [7].

6 Conclusions and Future Work

In this paper, we characterize the different thermal responses of forced air-cooling over an attached heatsink (*AIR-SINK*) and laminar oil-cooling over a bare silicon die (*OIL-SILICON*), and investigate their implications on temperature-aware microarchitecture design. *AIR-SINK* is the dominant cooling solution for the majority of high-performance processors, whereas *OIL-SILICON* is used in infrared thermal imaging of silicon. By modifying HotSpot with oil flow and secondary heat transfer path, we identified different transient and steady-state thermal responses of both *AIR-SINK* and *OIL-SILICON*. The direct convective cooling of silicon via the oil flow and the low thermal conductivity of the oil cause *OIL-SILICON* to have higher hot spot temperature, larger thermal gradients, and slower short-term transient response. In addition, the direction of the oil flow also plays an important role of hot spot location due to the non-uniform heat transfer coefficient along the flow path.

While IR-transparent oil cooling over bare silicon opens the way to direct and detailed runtime thermal measurements, the unique cooling configuration it requires also makes the silicon transient and steady-state thermal responses drastically different from a chip attached to a conventional air-cooled heatsink. DTM design based on *OIL-SILICON* measurements generally results in longer engagement duration of DTM than necessary, increasing performance overhead. It also leads to more on-chip thermal sensors to account for the greater across-die thermal gradient, hence increasing hardware complexity. The flow direction dependency of hot spot location also causes challenges of thermal sensor deployment and the accuracy of temperature-to-power reverse conversion.

Notice that in this paper, we only present analysis and results for the architecture implications of two specific chip cooling configurations. The entire design space of thermal packages and interaction with temperature-aware architecture-level performance needs thorough and quantitative analysis. This will be an interesting future work.

Another interesting future goal is to enhance a design-time thermal model to reconcile the differences among thermal packages. For example, it could be useful to ascertain the thermal response of a chip with air-cooled heatsink based on the IR measurements from an oil-cooled bare silicon die. Certain factors such as the temperature dependency of leakage power and the feedback loop from DTM may make such a derivation more complicated.

Acknowledgement

The authors would like to thank the anonymous reviewers for their insightful comments and suggestions. We thank Hendrik Hamann from IBM Research for interesting discussions and feedback. We also thank Jose Renau from University of California, Santa Cruz, for extensive discussions, feedback and the generous help with the IR setup information and the SESC-THERM model. This work is funded by NSF grants CRI-0551630 and EHS-0509245, as well as a grant from Intel Research.

References

- [1] D. Brooks and M. Martonosi. Dynamic thermal management for high-performance microprocessors. In *Proc. High-Performance Computer Architecture (HPCA)*, pages 171–182, January 2001.
- [2] D. Brooks, V. Tiwari, and M. Martonosi. Watch: A framework for architectural-level power analysis and optimizations. In *Proc. International Symposium on Computer Architecture (ISCA)*, pages 83–94, June 2000.
- [3] Y. A. Cengel. *Heat and Mass Transfer: A Practical Approach*. McGraw-Hill, Inc., New York, New York, 2007.
- [4] J. Choi, C. Y. Cher, H. Franke, H. Hamann, A. Weger, and P. Bose. Thermal-aware task scheduling at the system software level. In *Proc. of International Symposium on Low-Power Electronic Design (ISLPED)*, August 2007.
- [5] E. N. Wang et al. Micromachined jets for liquid impingement cooling of VLSI chips. *Journal of Microelectromechanical Systems*, 13(5):833–842, October 2004.
- [6] J. Clabes et al. Design and implementation of the POWER5 microprocessor. In *Proc. of International Solid State Circuits Conference (ISSCC)*, February 2004.
- [7] H.F. Hamann, A. Weger, J. A. Lacey, Z. Hu, P. Bose, E. Cohen, and J. Wakil. Hotspot-limited microprocessors: Direct temperature and power distribution measurements. *IEEE Journal of Solid-State Circuits*, 42(1):56–65, January 2007.
- [8] J. M. Koo, S. Im, L. Jiang, and K. E. Goodson. Integrated microchannel cooling for three-dimensional electronic circuit architectures. *Journal of Heat Transfer*, 127(1):49–58, January 2005.
- [9] E. Kursun and C. Y. Cher. Variation-aware thermal characterization and management of multicore architectures. In *Proc. of International Conference on Computer Design (ICCD)*, October 2008.
- [10] K. J. Lee, K. Skadron, and W. Huang. Analytical model for sensor placement on microprocessors. In *Proc. of Proceedings of the International Conference on Computer Design (ICCD)*, pages 24–27, October 2005.
- [11] P. Li, L. T. Pilleggi, M. Asheghi, and R. Chandra. IC thermal simulation and modeling via efficient multigrid-based approaches. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(9):1763–1776, September 2006.
- [12] F. J. Mesa-Martinez, J. Nayfach-Battilana, and J. Renau. Power model validation through thermal measurements. In *Proc. of International Symposium on Computer Architecture (ISCA)*, 2007.
- [13] M. Monchiero, R. Canal, and A. Gonzalez. Design space exploration for multicore architectures: A power/performance/thermal view. In *Proc. of ACM International Conference on Supercomputing (ICS)*, June 2006.
- [14] R. Mukherjee and S. O. Memik. Systematic temperature sensor allocation and placement for microprocessors. In *Proc. of Design Automation Conference (DAC)*, pages 542–547, June 2006.
- [15] S. Murthy, Y. Joshi, and W. Nakayama. Orientation independent two-phase heat spreaders for space constrained applications. *Microelectronics Journal*, 34(12):1187–1193, December 2003.
- [16] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan. Temperature-aware microarchitecture. In *Proc. International Symposium on Computer Architecture (ISCA)*, pages 2–13, June 2003.
- [17] T. Y. Wang and C. C. P. Chen. 3-D Thermal-ADI: a linear-time chip level transient thermal simulator. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 21(12):1434–1445, December 2002.
- [18] Y. Yang, Z. P. Gu, C. Zhu, R. P. Dick, and L. Shang. ISAC: Integrated space and time adaptive chip-package thermal analysis. *IEEE Transactions on Computer-Aided Design*, 26(1):86–99, January 2007.