Dynamic Thermal Management in Mobile Devices Considering the Thermal Coupling between Battery and Application Processor

Qing Xie[1], Jaemin Kim[2], Yanzhi Wang[1], Donghwa Shin[3], Naehyuck Chang[2], and Massoud Pedram[1]

[1]University of Southern California, CA, USA, [2]Seoul National University, Korea, [3]Politecnico di Torino, Italy

Email: [1]{xqing, yanzhiwa, pedram}@usc.edu, [2]{jmkim, naehyuck}@elpl.snu.ac.kr, [3]donghwa.shin@polito.it

Abstract—The thermal management is a crucial design problem for mobile devices because it greatly affects not only the device reliability, but also the leakage energy consumption. Conventional dynamic thermal management (DTM) techniques work well for the computer systems. However, due to the limitation of the physical space in mobile devices, the thermal coupling effect between the major heat generation components, such as the application processor (AP) and the battery, plays an important role in determining the temperature inside the mobile device package. Due to this effect, the thermal behavior of one part is no longer independent of the other, but is affected by the temperature of other parts. This is the first work that quantitatively characterizes the thermal coupling between the battery and AP and presents a predictive DTM for mobile devices considering this effect. Simulation results show that the proposed DTM method significantly reduces the thermal violations for the target mobile devices.

Keywords—*dynamic thermal management, smartphones, thermal coupling effect*

I. Introduction

Mobile devices, such as smartphones and tablets, continue to grow in popularity thanks to their computational ability and mobility. Some of the state-of-art mobile processors are considered to be as functional as typical processors in many existing computer systems. Meanwhile, the high-end mobile processor design is concerned with many of the same problems that conventional processors face, including performance, energy efficiency, thermal issues, and reliability concerns, although generally performance takes the back seat to energy efficiency and thermal considerations.

It is well known that elevated chip temperature is one major contributor to lower device reliability (i.e., accelerated aging) and higher leakage power consumption [1]. The negative bias temperature instability (NBTI) and hot carrier injection (HCI) cause the performance degradation [2] as circuits operating at high temperature. In addition, failure mechanisms such as electromigration, stress migration, and dielectric breakdown are exponentially dependent on temperature as well [3]. More importantly, the leakage power consumption is known to be exponentially dependent on the temperature of the application processor (AP). Higher temperature results in greater power consumption, and this effect in turn increases the AP temperature even more. Therefore, developing an effective thermal management policy, which is concerned about energy savings and reliability, is crucial for various mobile device platforms.

Dynamic thermal management (DTM) has been proposed as an effective technique to control the over-heating of the chip and maintain the chip temperature below a *critical temperature, $T_{critical}$*, above which the microprocessor chip could be damaged. A *trigger temperature, $T_{trigger}$ ($T_{trigger} \leq T_{critical}$)*, is defined so that when the chip temperature reaches $T_{trigger}$, DTM is evoked to cool down the chip. The typical DTM response mechanisms (control
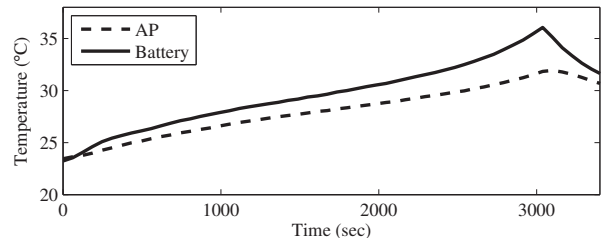
Fig. 1. Temperature profile of the battery and application processor during the 1C battery discharging in Nexus S. The phone is turned off.

knobs) include fetch-toggling, dynamic thread migration, frequency throttling and dynamic voltage and frequency scaling (DVFS) [4]. Practical run-time DTM methods require either distributed temperature sensors or a compact thermal model of the chip (such as HotSpot [5]) that would relate the chip temperatures to the chip-package-heat sink interface characteristics and power density values. In this paper we rely on a combination of sensor readouts and compact thermal modeling of the chip in order to accurately capture the temperature. In addition, we set $T_{trigger}$ to be close to $T_{critical}$.

Conventional DTM methods work well for computer systems where the CPU is assumed to be thermally independent of other components. However, due to the sharing of the rather small physical space, thermal behaviors of different components in mobile devices may interact with each other. Therefore, conventional DTM approaches, which only focus on processors, should be modified to consider the thermal coupling effect between the major heat generation components in enclosures of mobile devices, e.g., AP and battery. Although typical system-on-chip designs avoid direct abutment of the battery and AP, thermal coupling between these two components exists. For example, we clearly observe that the AP temperature rises, when we turn off the AP and discharge the smartphone battery separately. Figure 1 shows how the AP temperature increases as the battery temperature.

To the best of our knowledge, this is the first papers that points out the strong thermal coupling between the AP and battery in a mobile device enclosure, and provides a quantitative characterization of this effect. In particular, we build the thermal RC-circuit model to account for this effect and extract corresponding parameters through practical experiments. We also present a DTM method combining the thermal sensor readouts, look-up tables (LUTs), and the RC-circuit thermal modeling. For each task, we calculate the minimum frequency that guarantees to meet deadline constraints and lookup the maximum frequency that avoids thermal violations from pre-characterized tables. The DTM policy is based on the relationship between these two frequencies. In case of a thermal violation is predicted, a DTM response mechanism is activated, e.g., reducing the operating frequency or dropping the task, to avoid the potential violation of $T_{critical}$. Simulations targeting real smartphone platforms show significant reduction of the number of thermal violations by considering the thermal coupling effect.

II. BACKGROUND

A. Related Work

Many research works have been conducted on the dynamic thermal management for conventional computer systems. In the architecture level, Brooks et al. [4] described a couple of three micro-architectural and scaling techniques to cool-down the temperature of the microprocessors. Skadron et al. [5] adopted a feedback control theory to evoke the DTM and used the fetch-toggling as the response mechanism. The work aforementioned are mainly reactive DTM techniques, which have respond time limitation that prevents the usage of dynamic voltage scaling (DVS) due to its high invocation overhead [6]. The authors in [7] proposed a predictive DTM method and combined the architectural adaption and DVS for multimedia applications. The authors in [8] proposed proactive thread migration, DVFS and temperature balancing to minimize impact of thermal hotspots and temperature gradients. For real-time systems, Lee et al [9] proposed DTM for MPEG-2 decoding, which processes tasks at the minimum frequency meeting deadline constraints and sacrifices the video quality if there is a thermal violation. The authors in [10] derived the DTM policy using reinforcement learning technique: defining the processor temperatures and workload levels as the environment states and the available operating frequencies as the action set.

As the feature size scales down, the proportion of the leakage energy consumption in the total energy keeps increasing due to the lower threshold voltage, smaller channel length and gate oxide thickness. The authors in [11] improved the model for leakage current and showed that it has a exponentially relationship with the die temperature. In [12], Shin et al. considered the cooling mechanism for the CPU, accounted for the cooling energy consumption, and solved for the energy-optimal DTM policy. They jointly optimized the cooling fan usage and operating frequency selection to minimize the total energy consumption, including both of computing and cooling. Authors in [13] proposed the thermal management policy to extend the service life of the batteries in portable devices. None of the aforementioned work targets the mobile devices and considers the thermal coupling between the AP and battery.

B. Thermal and Power Model of Batteries

The general thermal model of a lithium battery is very complex due to the non-uniform temperature distribution and non-linearity of the chemical reactions. The temperature variation for polymer lithium batteries is negligible, and thereby we measure the temperature at the battery surface as the battery temperature. The charge and discharge reactions for lithium-ion secondary battery are endothermic and exothermic, respectively. At the system level, we assume the chemical reaction is reversible so that the total heat generation has two components: entropy change heat and ohmic heat [14],

$$Q_{bat} = Q_{bat,r} + Q_{bat,s} = I_{bat}(t)^2 \cdot r_{int} - T_{bat} I_{bat}(t) \frac{\partial V^{OC}}{\partial T}, \quad (1)$$

where $I_{bat}(t)$ is the charge or discharge current, V^{OC} is the open circuit voltage (OCV). The value of the $\partial V^{OC} / \partial T$ varies with the battery materials. The internal resistance of the battery can be empirically modeled as a function of the state-of-charge (SoC) of the battery [15],

$$r_{int} = b_1 e^{b_2 \cdot SoC} + b_3, \quad (2)$$

where $b_{1,2,3}$ are coefficients obtained through experiments that discharge batteries at different rates [16].

The power loss at the at the battery side includes the power loss on the internal resistance and the conversion loss in the DC-DC converter. The output power at the downstream of the DC-DC
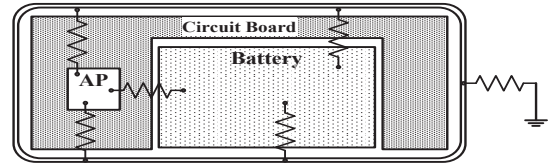

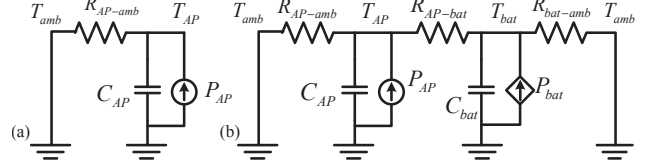Fig. 2. Conceptual diagram of thermal resistors in mobile devices.

Fig. 3. Comparison of RC-thermal circuit models: a) conventioal thermal model; b) thermal model considering the thermal coupling effect.

converter supports the application processor and other components in the mobile devices. We use a factor η to denote the converter efficiency, which has typical value around 85% in mobile devices. Thus we have,

$$V^{CC}(t) = V^{OC}(t) - I_{bat}(t) \cdot r_{int}. \quad (3)$$

$$P_{down}(t) = \eta \cdot V^{CC}(t) I_{bat}(t) = P_{AP}(t) + P_{other}(t). \quad (4)$$

Since r_{int} of the Li-ion battery is typically small, thus we approximately let V^{OC} equals to V^{CC} in (3). Therefore, insert (4) to (1), we have,

$$Q_{bat}(t) = \left(\frac{P_{down}(t)}{\eta V^{OC}(t)}\right)^2 \cdot r_{int} - T_{bat}(t) \left(\frac{P_{down}(t)}{\eta V^{OC}(t)}\right) \frac{\partial V^{OC}}{\partial T}. \quad (5)$$

C. Thermal and Power Model of the Application Processor

The power consumption of the application processor has two components: dynamic part and static (leakage) part. It is known that the leakage power has a strong dependence on the die temperature. Accurately modeling the power consumption of the AP is quite complicated since it contains many parts working at different frequency and voltage levels. We derive the approximated total power consumption versus voltage, frequency, and temperature based on our own measurements.

The thermal behaviors of the application processor is more straightforward than batteries. According to [5], thermal resistance is proportional to the package thickness and inverse of the interface area between the AP package and ambient environment. The thermal capacitance is proportional to the thickness and area of the AP package. The heat generation is proportional to the power consumption of the AP.

III. PROBLEM STATEMENT

A. RC-circuit Thermal Model of Thermal Coupling

Figure 2 shows a conceptual diagram of the mobile devices. Differ from previous work, we include a thermal resistor between the AP and the battery to reflect the thermal coupling effect. Figure 3(b) models Figure 2 using a RC-circuit manner. The thermal model includes elements as follows: T_{AP}, T_{bat} and T_{amb} denote temperatures of the AP, battery, and ambient environment; C_{AP} and C_{bat} denote thermal capacitances of the AP and battery; P_{AP} and P_{bat} are the heat generated by the AP and battery; and R_{AP-amb} and $R_{bat-amb}$ are thermal resistors between the ambient environment and the AP and battery, respectively. Compare to conventional RC-circuit thermal models like Figure 3(a) in [5], [12], a thermal resistor that corresponds to the thermal coupling between the AP and battery
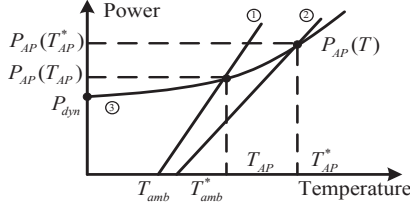
Fig. 4. The effect of the elevated ambient temperature due to the thermal coupling n the mobile devices.

is considered. Note that in Figure 3(b), the heat generation of the battery is described as a dependent source since it is determined by the total current demand of the AP and other components, as shown in (3)∼(5).

We write the Kirchhoff equations for the proposed RC-circuit thermal model shown in Figure 3(b) as follows,

$$C_{AP}\frac{dT_{AP}}{dt} = P_{AP} - \frac{T_{AP} - T_{amb}}{R_{AP-amb}} - \frac{T_{AP} - T_{bat}}{R_{AP-bat}},$$
$$C_{bat}\frac{dT_{bat}}{dt} = P_{bat} - \frac{T_{bat} - T_{amb}}{R_{bat-amb}} - \frac{T_{bat} - T_{AP}}{R_{AP-bat}}. \quad (6)$$

It is worthy to point out that the conventional RC-circuit thermal model in 3(a) is a special case of the proposed model: one can obtain the conventional model by disconnecting the battery heat source and short the T_{bat} to T_{amb}. In that case, the AP has two parallel-connected thermal resistors between T_{AP} and T_{amb}.

B. Effect of the Thermal Coupling

At the steady-state, equations set (6) has closed-form solutions as follows,

$$T_{AP} = T_{amb} + \frac{R_{AP-amb}R_{bat-amb} + R_{AP-bat}R_{AP-amb}}{R_{bat-amb} + R_{AP-bat} + R_{AP-amb}} \cdot P_{AP}$$
$$+ \frac{R_{AP-amb}R_{bat-amb}}{R_{bat-amb} + R_{AP-bat} + R_{AP-amb}} \cdot P_{bat}. \quad (7)$$

$$T_{bat} = T_{amb} + \frac{R_{AP-amb}R_{bat-amb}}{R_{bat-amb} + R_{AP-bat} + R_{AP-amb}} \cdot P_{AP}$$
$$+ \frac{R_{AP-amb}R_{bat-amb} + R_{AP-bat}R_{bat-amb}}{R_{bat-amb} + R_{AP-bat} + R_{AP-amb}} \cdot P_{bat}. \quad (8)$$

We solve the T_{AP} by plotting both left and right hand sides in (7) and finding the cross-point, as shown in Figure 4. In Figure 4, the line ③ denotes the temperature-dependent P_{AP} at a fixed operating frequency, i.e., fixed dynamic power consumption. The leakage power consumption increases super-linearly with respect to the AP temperature. Line ① stands for solutions without considering the thermal coupling and line ② stands for closed-form solutions of (7). In general, the intercept of ②, denoted by $T_{amb}^{\star}$, does not equal to T_{amb} because the last term in (7) is generally non-zero. The slope of ② is smaller than ① because high P_{AP} causes elevated T_{bat}, which plays a negative role in heat dissipation of the AP. Figure 4 shows that ignoring the thermal coupling between the battery and AP could result in underestimation of T_{AP} and P_{AP}.

C. Problem Formulation

Application processors typically have accurate temperature sensors embedded inside the package or on the printed circuit board [17], and thereby we can utilize these thermal sensors in the DTM methods. Although thermal sensors may suffer from noise issue, it is beyond the scope of this work. Although the temperature distribution inside the AP package are generally uneven among the junctions, we approximately focus on the temperature of the entire AP package at the system-level. We propose the DTM method combining the

accurate temperature predictions using the proposed RC-circuit thermal model with the thermal sensors and pre-characterized LUTs.

We consider stationary periodic soft real-time system and describe a task T_i by (a_i, W_i, d_i), where a_i, W_i and d_i denote the arrival time, workload and deadline for i-th task, respectively. The typical magnitude d_i are set to be in the order of tens to hundreds of milliseconds. We adopt a common assumption in the real-time system that the next task arrives no sooner than the deadline of the current task, $a_i + d_i \leq a_{i+1}$. We consider discrete operating frequency levels in the AP. The supply voltage, denoted by $V_{dd,k}$, is set as the minimum voltage that can support the AP running at operating frequency f_k. We consider K number of discrete operating frequency levels $(f_1, f_2, ..., f_K)$ for practical usage and thereby K number of voltage levels $(V_{dd,1}, V_{dd,2}, ..., V_{dd,K})$.

We aim to maximize the performance of the application processor subject to thermal constraints. The AP performance is defined as the percentage of tasks that are finished by their deadlines. Previous work in [18] revealed that keeping the die temperature at the maximal allowable value results in the maximum throughput. Therefore, we let the AP running at the temperature level close to the $T_{critical}$. We first qualitatively characterize the thermal coupling between the AP and battery through a series of experiments. Then we utilize the characterized thermal model to provide accurate predictions of the AP temperature, given the current temperature readings. In case that a thermal violation is predicted, we first try to avoid it by reducing the operating frequency which may violate the deadline constraint but still finish the task before the arrival time of the next task. It causes a relatively small amount of *quality of service* (QoS) degradation. However, we have to drop the task if no thermally safe operating frequency can be selected to finish the task by next arrival time, which causes a large amount of QoS degradation.

IV. METHODOLOGY

A. Parameter Extraction of the RC-Circuit Thermal Model

We setup the following experiments to extract the parameters of the proposed thermal model in Figure 3(b).

1) Characterizing the Battery: We first setup an experiment to characterize r_{int}, $\partial V^{OC}/\partial T$, $R_{bat-amb}$, and C_{bat}. We take the battery out from the smartphone, connect the battery terminals to a programmable active load device and place it in a constant-temperature chamber which maintains constant ambient temperature. We discharge the battery with various constant C-rating currents and derive r_{int} based on the voltage differences at different current ratings. We calculate the SoC of the battery by Coulomb counting and determine parameters b's according to (2). After separating thermal contributions made by ohmic heat and entropy change heat, we extract other parameters using curve fittings.

2) Characterizing the Battery-AP Thermal Coupling: We derive R_{AP-bat} and R_{AP-amb} in this experiment. The idea is to heat up only the battery and then measure the AP temperature change, while the AP is turned off. However, direct discharging battery using an active load has two issues: 1) it is impossible to precisely control the heat generation rate of the battery, given the complex battery thermal model in (1); and 2) the heating-up process is limited by the battery capacity. Therefore, we heat up the battery with a heater that wraps around the battery. The battery heater enables us to emulate a controllable constant-power heating such that both of the battery and AP reach their thermal equilibriums, i.e., the steady-states. We measure the steady-state temperatures and solve for R_{AP-bat} and R_{AP-amb}, by letting $P_{AP} = 0$ and P_{bat} to be the heat generation rate set in the experiment in (7).
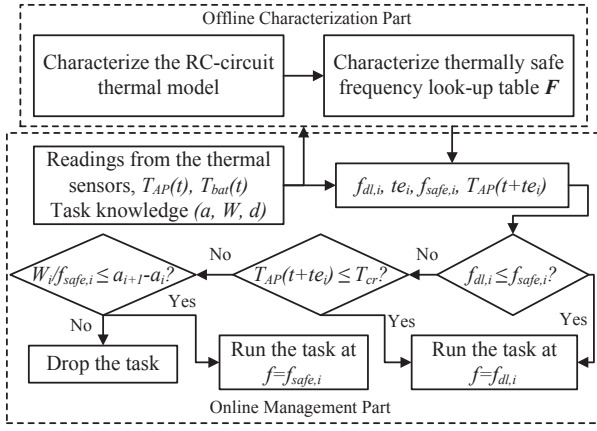
Fig. 5. The flowchart of the proposed DTM method.

3) Characterizing the AP: We derive the thermal capacitance of the AP, C_{AP}, in this experiment. Note that we cannot determine C_{AP} in Section IV-A2 because the steady-state temperature is independent of the thermal capacitance. There are generally two ways to derive the RC thermal time constant of the AP: heating-up or cooling-down. We use the latter method because accurately measuring the AP power dissipation is rather complicated. Therefore, we heat up the entire smartphone using the chamber, then place it to room environment and record the cooling-down temperature profile of the AP. The C_{AP} is derived using the measured RC time constant and previously determined R_{AP-amb}.

B. Dynamic Thermal Management

In this section, we present our DTM method, which consists of an offline characterization part and an online management part. Figure 5 shows the flowchart of the proposed DTM method. We first characterize the proposed RC-circuit thermal model, which allows us to accurately predict the AP thermal behaviors. Then we define *thermally safe frequency* f_{safe} as the maximum frequency that we can keep the AP running without causing the thermal violation after a time interval Δt, which is given by,

$$f_{safe}(T_{AP}, T_{bat}, \Delta t) = max\{f_k | T_{AP}(t+\Delta t) \leq T_{critical} \\ \text{and } f_k \in (f_1, f_2, ..., f_K)\}. \quad (9)$$

Since the thermally safe frequency in (9) is independent of the task, we pre-characterize all f_{safe} at different conditions and store them into a LUT, denoted by $F_{safe}(T_{AP}, T_{bat}, \Delta t)$, using the proposed thermal model and the relationship between the power consumption and operating frequency.

The online parts focus on maximizing the performance subject to the thermal constraint. We consider the stationary periodic soft real-time task set in this work. Thus, for a task (a_i, W_i, d_i), we define the *deadline frequency* $f_{dl,i}$ as the minimum operating frequency that can finish the task while meeting the deadline constraint. The corresponding *execution time* te_i is defined as the time to finish the task using $f_{dl,i}$.

$$f_{dl,i} = min\{f_k | \frac{W_i}{f_k} \leq d_i \text{ and } f_k \in (f_1, f_2, ..., f_K)\}, \\ te_i = W_i/f_{dl,i}. \quad (10)$$

Operating the AP at the frequency lower than $f_{dl,i}$ results in deadline violations. However, running the AP at frequency higher than $f_{dl,i}$ is wasteful from the energy perspective, which also causes unnecessary higher AP temperature. We access the thermal sensors and determine the thermally safe frequency for Task i using the knowledge of the task set and the pre-characterized LUT,

$$f_{safe,i} = F_{safe}(T_{AP}(t), T_{bat}(t), a_{i+1} - a_i). \quad (11)$$
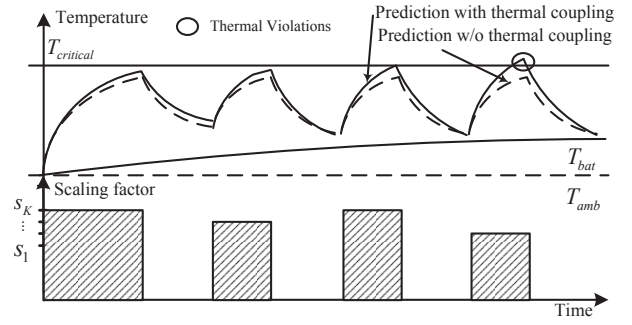

Fig. 6. Comparison of temperature prediction with and without considering the thermal coupling effect. The scaling factor is the ratio between the actual operating frequency and the maximal frequency.

At the time instance of a_i, we apply the following DTM policy to determine whether to execute this task the operating frequency.

1) $f_{dl,i} \leq f_{safe,i}$: we commit to run Task i at $f_{dl,i}$ as it is thermally safe and guarantees to finish by the deadline.
2) $f_{dl,i} > f_{safe,i}$: it is not necessary to cause thermal violations since in general $a_i + d_i \leq a_{i+1}$. Thus we solve the differential equation (6) for $T_{AP}(t+te_i)$ and if,
   a) $T_{AP}(t+te_i) \leq T_{critical}$: we commit to run Task i at $f_{dl,i}$ as it is thermally safe.
   b) $T_{AP}(t+te_i) > T_{critical}$: there is a thermal violation. We try to resolve it with QoS degradation and if,
      i) $W_i/f_{safe,i} \leq a_{i+1} - a_i$: we commit to run Task i at $f_{safe,i}$. It is thermally safe but causes a small QoS degradation.
      ii) $W_i/f_{safe,i} > a_{i+1} - a_i$: we have to drop Task i as there is no way to finish it before the arrival time of next task and meet the thermal constraint. It causes a large QoS degradation.

Figure 6 compares the AP temperature predictions with and without considering the thermal coupling effect conceptually. In the presence of thermal coupling, since the battery temperature is typically higher than the ambient environment, ignoring the thermal coupling effect underestimates the AP temperature. As the battery discharges, T_{bat} increases and the underestimation of the AP temperature increases as well. Since we keep the AP temperature close to the maximum allowable value for the sake of performance, underestimations potentially cause thermal violations. Thus, considering the thermal coupling in temperature prediction is crucial for the DTM method.

We briefly discuss the hardware and software overhead for the proposed DTM method. Temperature sensors are widely used in the mobile devices for the thermal management purpose. We pre-characterize the RC-circuit thermal model and the lookup table that contains thermally stable frequency of AP for each type of device, thereby the online computation cost is negligible. Since the battery typical has much larger thermal capacitance, we ignore the battery temperature variation during the task processing period and only focus on AP temperature. Compare to reactive DTM method, the proposed DTM method is armed with accurate thermal model to capture the AP thermal behaviors, and thereby accesses thermal sensors in a relatively low rate, which reduces the data-collection overhead. The overhead of DVFS (50~200 μs according to [6]) is also negligible compared to the duration of the task (tens to hundreds of milliseconds.)

V. EXPERIMENTAL RESULTS

We characterize the RC-circuit thermal model for a Google Nexus S smartphone [19], which has 1500 mAh battery and 1.0 GHz single core ARM Cortex-A8 processor. We dissemble the phone and attach
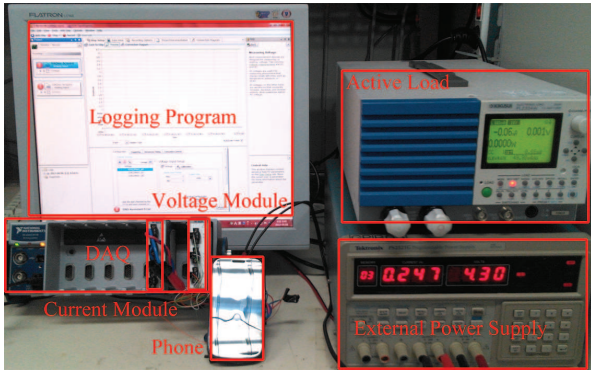
Fig. 7. Experimental setup used in this work.

TABLE I. EXTRACTED PARAMETERS FOR THE PROPOSED RC-CIRCUIT THERMAL MODEL.

Parameter	value	Parameter	value
b_1	3.897 Ω	b_2	-8.765
b_3	0.385 Ω	$\partial V^{OC}/\partial T$	0.62 mW/K
C_{bat}	150.2 J/K	$R_{bat-amb}$	7.58 K/W
R_{AP-amb}	35.8 K/W	R_{AP-amb}	78.8 K/W
C_{AP}	9.0 J/K		

TC1047AVNB thermal sensors [20] onto the center of battery back side, and the center of AP package. We reassemble the smartphone and securely close the battery door so that we keep the original operating condition. the temperature data are logged using NI-DAQ. The experimental setup is shown in Figure 7.

A. Battery Thermal Modeling

We take the battery from the smartphone, connect the battery terminals to the Kikusui PLZ334W programmable active load device. The system is placed in a constant-temperature chamber to keep the ambient temperature constant. We perform the constant-current discharging (0.5C, 1C, and 1.5C) for the battery using the active load and log temperature profiles of the battery. We extract the C_{bat}, thermal resistance, r_{int} and $\partial V^{OC}/\partial T$ using the method discussed in Section IV-A1. The parameters are summarized in Table I. The battery thermal model are validated by comparing simulated results to the measured data, as shown in Figure 8. Simulated thermal behaviors of 1C discharging case show 0.2% error on average.

B. Battery-AP Thermal Interaction Modeling

We wrap Nichrome wire onto the battery to produce controllable constant heating power. We apply 1.5V voltage across the Nichrome wire and 0.770A current, which gives a heating power of 1.125W. We wait for sufficiently long time to let the system reach the thermal steady-state ($T_{AP} = 31°C$, $T_{bat} = 33.5°C$, $T_{amb} = 25.5°C$). We derive R_{AP-amb} and R_{AP-bat} using the method discussed in Section IV-A2. The results are summarized in Table I. Simulated results of T_{bat} and T_{cpu} using extracted parameters match measured data well with an average error of 0.8% and 0.5%, as shown in Figure 9.
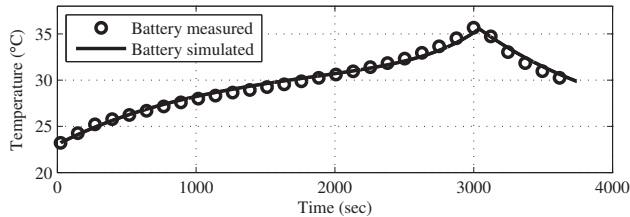

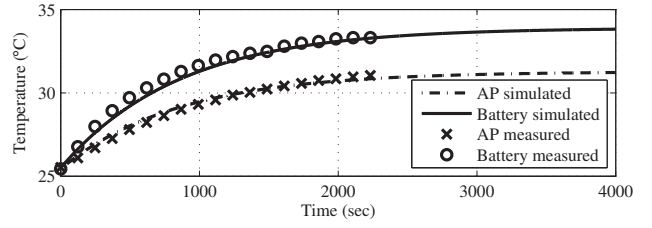Fig. 8. Measured and simulated battery thermal behaviros for a 1C discharging process.

Fig. 9. Measured and simulated battery and AP thermal behaviors by applying a controllable constant heating power.
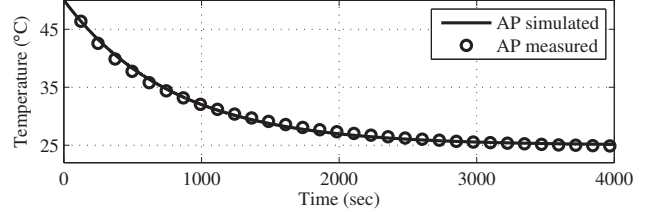

Fig. 10. Measured and simulated AP cooling down process.

C. AP Thermal Modeling

We place the smartphone in a constant-temperature chamber, heat it up to 50°C, and then move it to the room temperature of 25°C. We record the AP temperature decrease curve and obtain C_{AP} using the R_{AP-amb} in Table I. Figure 10 shows that the simulated cooling down curve matches well with the measured curve, with an average error of 0.85%.

D. Simulations of Proposed DTM Method

We compare the simulated the temperature rising of the Nexus S using the proposed RC thermal model and extracted parameters, with the measured temperature profile obtained by running the StabilityTest v2.7 benchmark [21]. Simulation results in Figure 11 show that our system models matches well with measured data at an average error of 1.46%. We carry out following simulations based on our system models.

We setup two periodic task sets: TS1 and TS2. TS2 has relatively shorter interval between task arrival times and the workload of tasks are heavier than the TS1. We assume six available operating frequency levels (389, 503, 655, 760, 950, and 1000) MHz. The power consumption at different frequency levels are derived based on our measurements. According to our measurements by running benchmarks on the Nexus S, the critical temperature that invokes the internal DTM is around 40°C. Considering that modern smartphones typically have multiple processors and thus produce more heat, we set the critical temperature in our simulation slightly higher to be 45°C and 50°C, respectively. While the AP is idle, we assume it is power gated with 95% efficiency. The temperature of the ambient environment is set to be 25°C. We compare the proposed DTM method with a baseline setup: the same DTM method except that the thermal coupling effect is ignored.
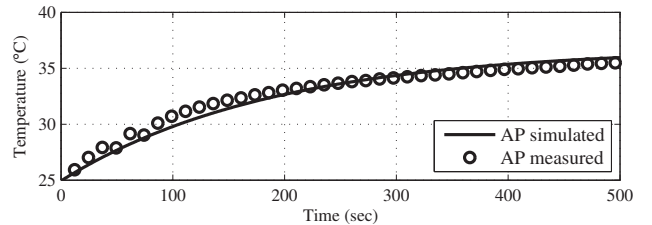

Fig. 11. Verification of AP thermal behaviors on Nexus S during the temperature rising while running the StabilityTest v2.7 benchmark [21].

$T_{critical}$	Method	Thermal Violations % at T_{bat} (°C)				
		28	30	32	34	36
45 °C	Proposed	0.87	0	0.40	0	0
	Baseline	0.87	11.61	17.11	28.32	14.50
50 °C	Proposed	0	0	0	0	0
	Baseline	0	0	1.34	15.57	14.23

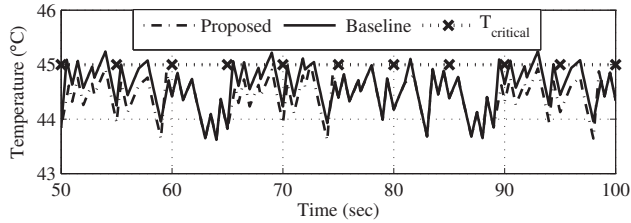| $T_{critical}$	Method	Thermal Violations % at T_{bat} (°C)				
		28	30	32	34	36
45 °C	Proposed	0.25	0	0	0	1.56
	Baseline	7.33	19.10	17.94	41.18	48.54
50 °C	Proposed	0	0	0	0.20	0.05
	Baseline	0	0	5.28	31.26	18.59

Fig. 12.   Thermal violations caused by baseline DTM without considering thermal coupling effect. Simulation for TS1 with $T_{critical} = 45°C$.

Table II and III show the number of tasks causing thermal violation, i.e., $T_{AP} > T_{critical}$, as a percentage of all tasks at given different critical temperatures. Since ignoring the thermal coupling typically underestimates the AP temperature, the baseline setup results in thermal violations much more frequently. The proposed method avoids thermal coupling most of the time, thanks to its accurate temperature prediction. The simulation for TS2 shows more thermal violations than the TS1 because TS2 is heavier. Among each table, less thermal violations are observed at higher $T_{critical}$ because the DTM is more relaxed if the allowable temperature is higher. Figure 12 shows a clip of the simulation result. One can see that the AP temperature of the baseline setup rises above the $T_{critical}$ several times.

Figure 13 shows the relationship between the battery temperature and dropped tasks as a percentage of the all tasks for TS1. High battery temperature brings more difficulties to heat dissipation of the AP. Therefore, we have to drop more tasks to avoid thermal violations. On the other hand, maintaining the AP at a lower $T_{critical}$ temperature results in larger QoS degradation, i.e., more tasks are dropped. Similar trend is observed for tasks having deadline violations.

VI. CONCLUSION

A fundamental difference between the thermal management in mobile devices and conventional computer systems is the significant thermal coupling effect among the major heat generation components, s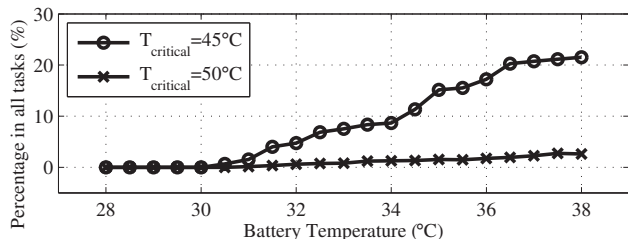uch as battery and application processor (AP), through the thermal conduction. We observed this thermal coupling issue in mobile devices through practical experiments. We proposed an RC-circuit thermal model that reflects the thermal coupling effect and qualitatively characterized the values of the RC-circuit components. We introduced a dynamic thermal management (DTM) method, combining proposed thermal model with the temperature sensors and the pre-characterized LUT, to maximize the smartphone performance subject to the thermal constraint. This paper showed that conventional DTM methods ignoring the thermal coupling effect could incur frequent violations of the thermal constraint and thus may not work correctly. The proposed DTM method accounted for the thermal coupling effect significantly reduced AP thermal violations.

REFERENCES

[1] M. Pedram and S. Nazarian, "Thermal modeling, analysis and management in vlsi circuits: principles and methods," *Proceedings of the IEEE*, 2006.

[2] W. Wang, V. Reddy, R. Vattikonda, S. Krishnan, and Y. Cao, "Compact modeling and simulation of circuit reliability for 65-nm cmos technology," 2007.

[3] JEDEC publication JEP122C, "Failure mechanisms and models for semiconductor devices." http://www.jedec.org.

[4] D. Brooks and M. Martonosi, "Dynamic thermal management for high-performance microprocessors," in *HPCA*, 2001.

[5] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy, and D. Tarjan, "Temperature-aware microarchitecture: Modeling and implementation," 2004.

[6] S. Park, J. Park, D. Shin, Y. Wang, Q. Xie, N. Chang, and M. Pedram, "Accurate modeling of the delay and energy overhead of dynamic voltage and frequency scaling in modern microprocessors," *IEEE T. CAD*, 2013.

[7] J. Srinivasan and S. V. Adve, "Predictive dynamic thermal management for multimedia applications," in *ICS*, 2003.

[8] A. K. Coskun, T. S. Rosing, and K. C. Gross, "Proactive temperature balancing for low cost thermal management in mpsocs," in *ICCAD*, 2008.

[9] W. Lee, K. Patel, and M. Pedram, "Gop-level dynamic thermal management in mpeg-2 decoding," 2008.

[10] Y. Ge and Q. Qiu, "ynamic thermal management for multimedia applications using machine learning," in *DAC*, 2011.

[11] W. Liao, L. He, and K. M. Lepak, "Temperature and supply voltage aware performance and power modeling at microarchitecture level," *IEEE T. on CAD*, 2005.

[12] D. Shin, S. W. Chung, E.-Y. Chung, and N. Chang, "Energy-optimal dynamic thermal management: Computation and cooling power co-optimization," *IEEE T. on Industrial Informatics*, 2010.

[13] Q. Xie, S. Yue, D. Shin, N. Chang, and M. Pedram, "Adaptive thermal management for portable system batteries by forced convection cooling," in *DATE*, 2013.

[14] K. Onda, T. Ohshima, M. Nakayama, K. Fukuda, and T. Araki, "Thermal behavior of small lithium-ion battery during rapid charge and discharge cycles," *J. of Power Sources*, 2010.

[15] M. Chen and G. Rincon-Mora, "Accurate electrical battery model capable of predicting runtime and I-V performance," *IEEE T. on Energy Conversion*, 2006.

[16] D. Shin, Y. Wang, Y. Kim, J. Seo, M. Pedram, and N. Chang, "Battery-supercapacitor hybrid system for high-rate pulsed load applications," in *DATE*, 2011.

[17] Texas Instruments, "OMAP4460 multimedia device data manual." http://www.ti.com/product/OMAP4460.

[18] R. Rao and S. Vrudhula, "Performance optimal processor throttling under thermal constraints," in *CASES*, 2007.

[19] Nexus S specifications. http://www.samsungnexuss.com/nexus-s-specs/.

[20] TC1047AVNB thermal sensor. http://pdf1.alldatasheet.com/datasheet-pdf/view/74999/MICROCHIP/TC1047AVNB.html.

[21] StabilityTest v2.7. https://play.google.com/store/apps/details?id=com.into.stability&feature=search_result.

Fig. 13.   Violated tasks and droped tasks as a percentage of all tasks in Task Set 1 versus different battery temperature.