Heat to Power: Thermal Energy Harvesting and Recycling for Warm Water-Cooled Datacenters

Xinhui Zhu Weixiang Jiang Fangming Liu* Qixia Zhang Li Pan Qiong Chen Ziyang Jia
National Engineering Research Center for Big Data Technology and System,
Key Laboratory of Services Computing Technology and System, Ministry of Education,
School of Computer Science and Technology, Huazhong University of Science and Technology, China

***Abstract*—Warm water cooling has been regarded as a promising method to improve the energy efficiency of water-cooled datacenters. In warm water-cooling systems, hot spots occur as a common problem where the hybrid cooling architecture integrating thermoelectric coolers (TECs) emerges as a new remedy. Equipped with this architecture, the inlet water temperature can be raised higher, which provides more opportunities for heat recycling. However, currently, the heat absorbed from the server components is ejected directly into the water without being recycled, which leads to energy wasting. In order to further improve the energy efficiency, we propose Heat to Power (H2P), an economical and energy-recycling warm water cooling architecture, where thermoelectric generators (TEGs) harvest thermal energy from the "used" warm water and generate electricity for reusing in datacenters. Specifically, we propose some efficient optimization methods, including an economical water circulation design, fine-grained adjustments of the cooling setting and dynamic workload scheduling for increasing the power generated by TEGs. We evaluate H2P based on a real hardware prototype and cluster traces from Google and Alibaba. Experiment results show that TEGs equipped with our optimization methods can averagely generate 4.349 W, 4.203 W, and 3.979 W (4.177 W averagely) electricity on one CPU under the *drastic*, *irregular* and *common* workload traces, respectively. The power reusing efficiency (PRE) can reach 12.8%∼16.2% (14.23% averagely) and the total cost of ownership (TCO) of datacenters can be reduced by up to 0.57%.**

***Index Terms*—datacenter energy, warm water cooling, thermal energy harvesting, energy recycling, thermoelectric generator**

I. INTRODUCTION

It is no secret that datacenters are consuming enormous energy not only for IT equipment but also for other non-IT equipment (e.g., UPS and cooling system). For instance, electricity consumed by U.S. datacenters was up to 70 billion kWh in 2014, making up 2% of the national total energy cost [43]. However, nearly 48% of the energy is consumed by non-IT equipment, while 79% of the non-IT energy consumption can be attributed to the cooling systems [38]. This trend impels the revolution of datacenters to improve its energy efficiency by means of reducing the electricity consumption associated with the cooling systems.

*Corresponding author: Fangming Liu (fmliu@hust.edu.cn). This work was supported in part by the NSFC under Grant 61722206 and 61761136014 (and 392046569 of NSFC-DFG) and 61520106005, in part by National Key Research & Development (R&D) Plan under grant 2017YFB1001703, in part by the Fundamental Research Funds for the Central Universities under Grant 2017KFKJXX009 and 3004210116, in part by the National Program for Support of Top-notch Young Professionals in National Program for Special Support of Eminent Professionals.

Compared with air cooling, water cooling emerges as a new energy-efficient paradigm for datacenters. By enabling higher rack densities and improved cooling efficiency, it is reported that the water cooling system can reduce the cooling energy consumption by averagely 22% as compared with the air cooling system [3], [24]. In fact, to prevent hardware failures in practice, the facility water should be cold enough to cool down the hottest server, which usually runs at the highest CPU utilization. Hence, chillers need to frequently work in the water cooling system to ensure that the facility water maintains at a low-temperature level (e.g., 7°C∼10°C), which contributes to a large amount of energy consumption. However, the servers with lower CPU utilization usually do not need such "cold" water to take away their heat. In other words, this over-provision strategy can easily lead to energy waste [34], [35].

To further improve the energy efficiency of water-cooled datacenters, the key lies in reducing the energy consumption of the chiller. It is reported that by simply raising the temperature of facility water from 7°C∼10°C to 18°C∼20°C, the cooling energy consumption can be saved by as much as 40% [35]. Based on the fact that servers in datacenters are in low utilization most of the time [31], [32], warm water (e.g., 40°C∼45°C) cooling strategy is proposed to reduce the cooling costs [8]. However, when some servers get overload suddenly, they may exceed the safe operating temperature in a few seconds, while the chiller needs a relatively long time (e.g., several minutes) to cool the water and delivers it to the overheated servers, which can lead to cooling lag/mismatching and the risk of cooling failure [24].

Even though it is reported that the effect of high temperatures on hardware reliability is not so high [15], the hot spot issue in warm water cooling systems still needs a feasible solution. Recently, Jiang et al. [24] propose a hybrid cooling architecture integrating each CPU with a thermoelectric cooler (TEC) to provide extra cooling for individual servers dynamically. By fine-grained and adaptive cooling control, it can economically handle the cooling mismatching according to workload variations. In this way, warm water cooling becomes a promising method which can not only cut down the need for expensive and energy-consuming chillers, but also provide more possibilities for harvesting and recycling the heat in the "used" warm water. For example, ASHRAE Thermal Guidelines [4] "W5" suggests using >45°C water to cool datacenters so that the outlet water is hot enough for heat recovery and heating buildings [7].

However, heat recovery by district heating has its limitations. For one thing, a huge project is needed to transfer the heat to the required district (e.g., complex piping arrangement, urban heating system redesign and planning). For another thing, heat is not always in great demand from season to season, from district to district. In fact, district heating is more common in high-latitude areas suffering from cold winters, like northern Europe [14]; while tropics like Singapore are hot all year round, and accordingly have a low requirement on heating. In addition, most datacenters are located in warm areas, where the peak-hour heat capacity of datacenters exceeds the heat demand of residential homes from spring to autumn (e.g., April to October in Washington D.C., San Francisco, and Huston [7]), leading to a heat mismatching. Last but not least, thermal energy is difficult to store.

Facing these issues, we wonder if there is another way of recycling the thermal energy collected from the "used" warm water. Is it possible to transform the heat into electricity so as to provide part of the power for datacenters themselves? The answer lies in a simple but powerful component, i.e., the thermoelectric generator (TEG), which can generate electricity under temperature difference between its two sides. For example, a previous work has considered deploying small-scale TEGs on the chip in smartphones to harvest heat and turn it into electricity [11]. Motivated by this, in this paper, we advocate that the heat absorbed from the server components can be recycled to further improve the energy efficiency of the warm water-cooled datacenters. Accordingly, we propose Heat to Power (H2P), an economical and energy-recycling warm water cooling architecture, where TEGs harvest thermal energy from the "used" warm water and generate electricity for reusing in warm water-cooled datacenters. In this context, the contributions of this paper are summarized as follows:

- We explore the TEG as a new component to recycle waste heat in warm water-cooled datacenters and present the characteristics of TEGs. We propose Heat to Power (H2P), an economical and energy-recycling warm water cooling architecture, which incorporates TEGs into the existing water cooling system. We consider the placement and scale of TEGs and carefully design a thermal energy harvesting and recycling module.
- We build a hardware prototype to validate the feasibility of H2P. Empirical measurements are conducted on the prototype to demonstrate the generation capacity of TEGs under the influence of different factors, such as the inlet temperature, flow rate, and CPU utilization.
- Since the electricity generation is limited according to the measurement results, we propose some efficient optimization methods, including an economical water circulation design, fine-grained adjustments of the cooling setting (i.e., the inlet water temperature and the flow rate) and dynamic workload scheduling (i.e., workload balancing) for increasing the power generated by TEGs. Experiment results show that TEGs equipped with our optimization methods can averagely generate 4.349 W, 4.203 W, and 3.979 W (4.177 W averagely) electricity on one CPU under three types of workload traces (i.e., *drastic*, *irregular* and *common*) from Google and Alibaba. The power reusing efficiency (PRE) can reach 12.8%∼16.2% (14.23% averagely).
- We provide a comprehensive analysis of the total cost of ownership (TCO) of H2P, which can be reduced by up to 0.57% (i.e., saving around $410,000 per year for a datacenter with 100,000 CPUs).
- We also discuss the limitation of using TEGs to provide electricity for warm water-cooled datacenters. Besides, we give some suggestions on the potential applications of the TEG-enabled water cooling system, and explore the future development of TEGs and warm water-cooled datacenters.

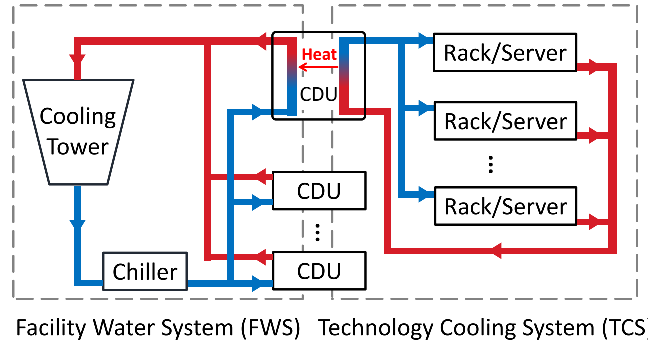


Fig. 1: Water cooling system architecture in a datacenter.

The rest of the paper is organized as follows: Sec. II provides the background and motivation of heat recycling in warm water-cooled datacenters. Sec. III characterizes TEGs, proposes our H2P system architecture, and highlights the key design considerations. Sec. IV presents our hardware prototype and the measurement results and analyses based on the prototype. Sec. V introduces the policies to increase the generated electricity. Sec. VI discusses current potential applications and future development of TEGs. Sec. VII introduces the related works and Sec. VIII concludes the paper.

II. BACKGROUND AND MOTIVATION

In this section, we first introduce the water cooling system architecture for datacenters. Then we illustrate the advantages of adopting warm water cooling in datacenters. Finally, we highlight the motivation for heat recycling by TEGs in warm water-cooled datacenters.

A. Water Cooling System Architecture

Fig. 1 shows a widely-used water cooling system architecture in a datacenter. As plotted, it comprises two liquid loops. The facility water system (FWS) is separated from the technology cooling system (TCS) by the coolant distribution unit (CDU) [5]. This water cooling system has one or more CDUs, which distribute TCS coolant to servers in a single rack or several racks and transfer heat from TCS to FWS by using liquid-to-liquid heat exchangers. In FWS, heat is removed mainly by the cooling tower via evaporation. If the ambient air temperature is high, chillers need to further cool the facility water, which is quite energy-consuming. In TCS, CDUs regulate the coolant temperature and the flow rate by using valves and centralized pumps.

B. Warm Water Cooling for Datacenters

Traditional water cooling systems consume a large amount of energy, especially consumed by the chillers. With advanced devices like TECs to deal with hot spots, less or even no use of chillers can be achieved and the temperature of inlet facility water can be raised higher. However, warm water cooling brings about the following concerns.

Is warm water cooling safe? Compared with the hard disk, CPU is more resistant to high temperature and temperature variation [8]. The previous work [24] shows that for Intel Xeon E5-2650 V3, using 40°C∼45°C water will not cause CPU temperature to exceed the maximum operating temperature (i.e., 78.9°C) even under 100% CPU utilization. However, CPU temperature will exceed the maximum operating temperature when the facility water is over 50°C and CPU utilization is over 70%. Due to the averagely low CPU utilization, it is feasible to use 40°C∼50°C water for cooling. In spite of this, warm water cooling still faces the risk of cooling failure when a hot spot emerges with sudden high CPU utilization. In this case, the hot spot servers should be cooled down immediately, while the facility water needs several minutes to be cooled by the chiller. To tackle this problem, a hybrid cooling architecture has been proposed, which integrates TECs to provide extra and timely fine-grained cooling for each server [24]. Hence, equipped with this architecture and method, the inlet water temperature can be raised higher.

What are the benefits of warm water cooling? First, warm water cooling avoids over-cooling some servers running at a low utilization and allows less or even no use of the chiller, which is a kind of active cooling equipment and quite energy-consuming. The main cooling task can be undertaken by the cooling tower via evaporation. Second, the difficulty and challenge of recycling waste heat in the conventional water cooling system using cold water to cool servers are that the heat, though plentiful and continuous, is of low-grade [14]. Higher inlet temperature means higher outlet temperature. After harvesting heat from server components, the "used" warm water is an available and higher-quality heat source. Recently, datacenter heat recovery has been attracting growing attentions. Innovative green datacenters usually cooperate with the district heating system (DHS), providing heating for urban facilities such as office buildings, swimming pools, etc. If the outlet water from datacenters is hot enough, DHSs do not need to further warm it up, so that more energy can be saved.

C. Heat Recycling in Datacenters

In the past, industry and academia improved the Power Usage Effectiveness (PUE) [46] mainly by reducing IT costs and cooling costs. For example, Google has reduced average PUE values close to 1.1 and represented a highly efficient design [17]. Meanwhile, the limitations of existing technologies leave not much room for PUE reduction [12]. Hence, Energy Reuse Effectiveness (ERE) [46] is proposed by Green Grid to measure the benefit of reuse energy from a datacenter, in the form of:

$$ERE = \frac{E_{IT} + E_{Cooling} + E_{Power} + E_{Lighting} - E_{Reuse}}{E_{IT}}.$$

Maximizing energy reuse enables the ratio less than 1, spurring datacenter practitioners to focus more on reusing waste energy for higher energy efficiency. Hence, the massive heat from datacenters will not be viewed as an unwanted byproduct anymore, it can be turned into an opportunity for energy reuse. For example, waste heat can be harvested and recycled in the following three manners.

District heating: One common solution to reuse waste heat is district heating. Datacenters need to cooperate with DHSs, which is a huge project involving both hardware (e.g., complex piping arrangement, urban heating system redesign and planning) and software (e.g., a win-win cooperative mechanism between datacenters and DHSs) design. This remedy is not so applicable to middle-latitude and low-latitude countries in which the heat demand is not all year round, thus they need to bear a significant cost to construct a mature urban heating system. Also, the function of heat is limited. If the energy is in the form of electricity, it can play an even greater role in datacenters requiring electricity more than heat.

Heat to electricity: Previous works have used the thermoelectric generator (TEG) to turn heat into electricity. Small-scale TEGs are placed on the chip in the smartphone or on CPUs of servers to harvest heat [11], [29]. However, TEGs are made of adiabatic materials. Directly attaching TEGs to the CPU may decrease the heat dissipation effect of the CPU (will be shown in Sec. III-B later). Besides, the cold side of TEG is directly exposed to the air, which will gradually reduce the temperature difference between the hot side and cold side of TEG and accordingly degrade its electricity generation.

Combined cooling, heat and power (CCHP): This solution refers to the simultaneous generation of electricity and useful heating and cooling mostly from the combustion of gas [39]. In CCHP, the waste heat is typically collected as steam that can either used for heating, or for generating electricity by delivering steam through a turbine to an absorption chiller, which is used for cooling.

However, since the construction and maintenance costs of CCHP are much higher (e.g., CCHP system design, facilities deployment, and complex piping arrangement), we propose a TEG-integrated system H2P, which is easy to install in the existing water-cooled datacenters at very low cost. Besides, most CCHP systems need gas as energy supply with stricter fire and explosion protection, while H2P does not require additional energy for electricity generation in existing datacenters. Thus, either to integrate TEGs at the building/datacenter scale, or at the server/CPU scale, H2P has its unique advantages that can not be replaced by CCHP. In H2P, we place TEGs at

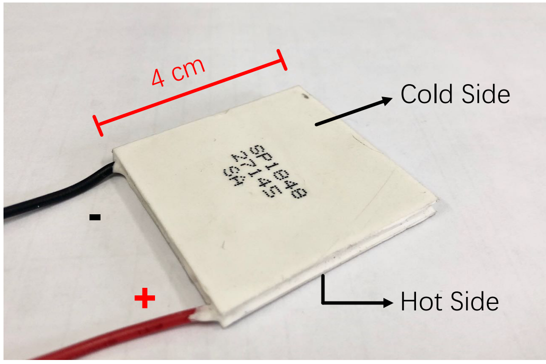


Fig. 2: A thermoelectric generator.

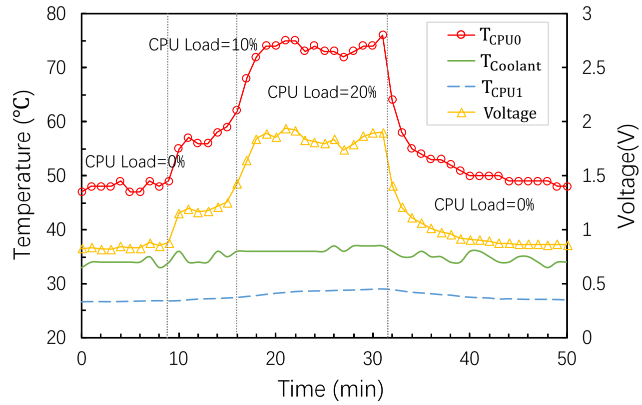


Fig. 3: TEG can hardly conduct heat. Note that a TEG is sandwiched by CPU0 and its cold plate, while CPU1 is directly pressed by the cold plate (CPU Type: Intel Xeon E5-2650 V3, Maximum Operating Temperature: 78.9°C).

the server/CPU scale (which has the highest temperature) to better reuse waste heat. Notably, CCHP and TEG-integrated solutions can be combined with each other to further improve energy efficiency.

III. HEAT TO POWER: KEY CHARACTERIZATION AND SYSTEM ARCHITECTURE

In this section, we first introduce the basic concept of TEGs. Then we analyze the feasibility of integrating TEGs into warm water-cooled datacenters. Finally, we propose Heat to Power (H2P) system for warm water-cooled datacenters.

A. What is TEG?

The thermoelectric generator (TEG) is a device that can convert heat into electricity. It consists of several pairs of n-type and p-type semiconductors sandwiched by two ceramic chips. When there is a temperature difference between the two chips, TEG generates a voltage. This phenomenon is called the Seebeck effect [6].

Physical characteristic: The commercially available TEG (model name: SP 1848-27145) is usually made of Bi_2Te_3 and has a size of $4cm \times 4cm$, as shown in Fig. 2. Such a TEG's resistance is about 2∼2.5 Ω. Its appropriate ambient temperature ranges from -60°C∼120°C, which makes it applicable in the datacenter environment. The open-circuit voltage of a TEG can be calculated as [11]:

$$V_{oc} = n \times \alpha_{TEG} \times \Delta T_{TEG}, \quad (1)$$

where n is the number of n-type and p-type pairs; α_{TEG} is the Seebeck coefficient of TEGs, and ΔT_{TEG} is the temperature difference between the hot side and the cold side of a TEG. As a result, the larger the temperature difference, the higher the voltage.

Cost analysis: We bought TEGs from *Alibaba.com* [2] at a cost of \$1 per piece. There are no moving parts and no working fluids inside the TEG, which means no maintenance and no extra costs [6]. Moreover, it has a long lifespan of no less than 28∼34 years when working with constant heat sources which are exactly available and reliable in datacenters [7], [14], [45].

B. Where to Place TEGs in Datacenters?

To figure out the feasibility of integrating TEGs into warm water-cooled datacenters, we first need to notice that there exists a positive correlation between temperature difference and the generated electricity. To obtain more electricity, we need to find out how to make the temperature difference as large as possible and try to put the TEGs as much close to the hottest place in the whole cooling system as possible. In the traditional water-cooling system, the CPU is pressed tightly by a cold plate, through which the coolant keeps flowing and absorbs heat. The cold plate is made of metal and has a high thermal conductivity. However, TEG is almost adiabatic.

To investigate the thermal conductance of TEG, we make a measurement in a Dell T7910 server equipped with two CPUs of the same type, i.e., Intel Xeon E5-2650 V3, whose maximum operating temperature is 78.9°C [10]. The two CPUs are connected in parallel in the water circulation, hence the flow rate and the inlet temperature in the two branches are almost the same. The only difference is that a TEG is sandwiched by CPU0 and its cold plate, while CPU1 is directly pressed by the cold plate. We divide fifty minutes into four phases, in which CPU load is equal to 0%, 10%, 20%, 0% in sequence. We stress the same load on CPU0 and CPU1 each phase and then record their temperatures, the coolant temperature and the open-circuit voltage of TEG. As can be seen in Fig. 3, CPU0's temperature experiences twists and turns in the whole process, and is very close to the maximum operating temperature at a load of 20%, while the temperatures of coolant and CPU1 are relatively stable which means the heat from CPU0 is hardly absorbed by coolant. The variation of voltage accords with CPU0's temperature.

Hence, considering the CPU's safety and the electricity generation capacity, we place the TEG module at the outlet of each CPU, which is the hottest place in the whole circulation.

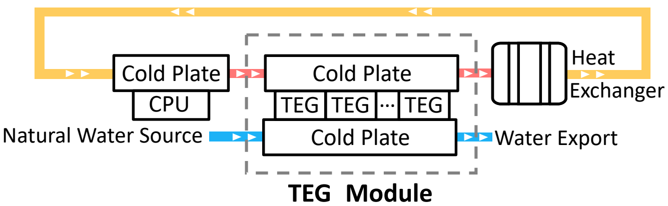


Fig. 4: Heat to Power (H2P) system architecture.

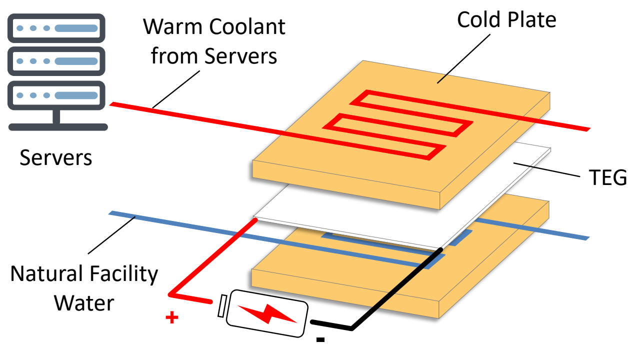


Fig. 5: A thermoelectric generation module.

C. System Architecture of H2P Integrating TEGs

Fig. 4 plots the system architecture of H2P. The TEG module (details in Fig. 5) consists of several TEGs connected electrically in series at the outlet of each server. To utilize the electricity generated by TEGs, the output voltage must reach a certain value. However, according to our measurement results (will be shown in Sec. IV-B later), the output voltage of one TEG is not enough, so we choose a simple and effective collecting-in-series method to raise the voltage, which is also the most commonly used scheme in current research [22], [23]. Under this circumstance, the maximum output power occurs when the load resistance equals the whole TEG module's resistance [22].

The TEG module is sandwiched by two cold plates, which are connected into two separate liquid loops. In particular, on the hot side of TEGs, the water in the cold plate is the warm coolant from the servers. This flow of water has been used to cool down the CPU of servers and reaches over 40°C after taking away the heat of CPUs. After flowing through the TEG module, it flows to the heat exchanger and transfers heat to FWS. On the cold side of TEGs, the natural water source is the domestic water or the running water from nature, which is around 20°C.

In fact, the emerging water-cooled datacenters tend to be located by the sea or the lake, for accessing to natural water resources. AliCloud Qiandao Lake datacenter is designed to use deep water from Qiandao Lake, the temperature in which stabilizes perennially at 15°C$\sim$20°C. This source of water can be a perfect natural water source for TEGs.

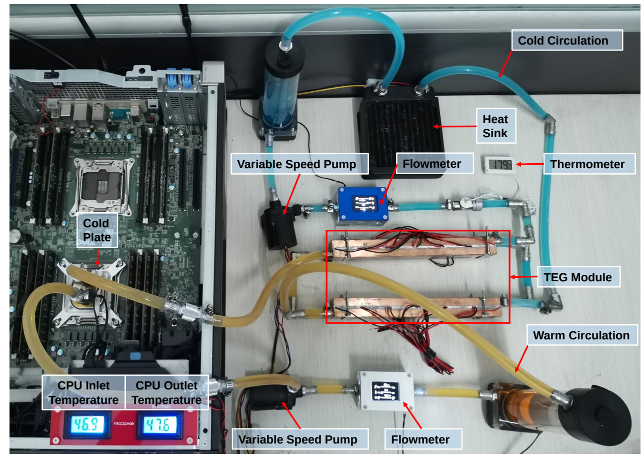


Fig. 6: Hardware prototype of Heat to Power (H2P) system (TEG Model: SP 1848-27145, CPU Type: Intel Xeon E5-2650 V3).

IV. HARDWARE PROTOTYPE AND EMPIRICAL MEASUREMENT

To evaluate the performance of the H2P system, we build a hardware prototype on a commonly-used server. Then, to figure out the power generation capacity of TEGs under the influence of different factors, such as the inlet temperature, flow rate, and CPU utilization, we conduct a series of empirical measurements on this prototype.

A. Hardware Prototype of H2P

To investigate the power generation capacity of TEGs, we build a hardware prototype of H2P in a Dell T7910 server equipped with an Intel Xeon E5-2650 V3 CPU. The TEG module consists of 12 TEGs whose model is SP 1848-27145 as shown in Fig. 6. Six TEGs as a group are sandwiched by two $4cm \times 24cm$ cold plates. There are two water circulations in this test-bed: the warm circulation filled with yellow coolant and the cold circulation with blue coolant. The warm circulation, acting as the TCS loop, includes three cold plates, a variable speed pump, a flowmeter and temperature sensors. One $4cm \times 4cm$ cold plate presses the CPU tightly to transfer heat from CPU to the yellow coolant. The other two $4cm \times 24cm$ cold plates are respectively attached to two groups of TEGs' one side, in order to create a hot surface for TEGs. Temperature sensors are used to monitor the inlet and outlet water temperatures of CPU. Note that CPU outlet temperature can be viewed as the TEG module inlet temperature of the warm circulation. The cold circulation comprises two $4cm \times 24cm$ cold plates both to create the cold surface for TEGs, a variable speed pump, a heat sink to simulate heat dissipation of the natural water source, a flowmeter and a temperature sensor to monitor the water temperature of the cold circulation. We use *Fluke 2638A* data acquisition system to record the temperatures and voltages simultaneously.

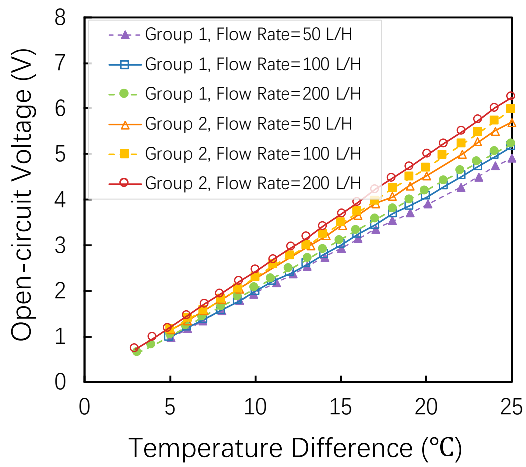


Fig. 7: Open-circuit voltage variation of 6 TEGs connected in series with temperature difference between warm coolant and cold coolant at different flow rates (the warm coolant and the cold coolant are set to the same flow rate).

B. Empirical Measurement on Prototype

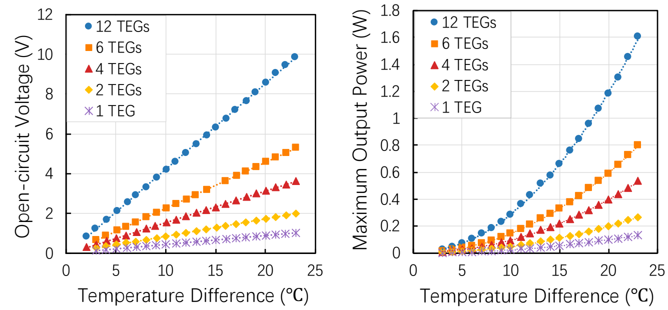
We assume that the temperature of natural water in cold circulation stabilizes at 20°C. Warmer TCS coolant brings larger temperature difference, accordingly more electricity generation, but meanwhile a higher risk of cooling failure in CPU, leading to a trade-off between *TEGs' electricity generation capacity* and *CPU's safety*, which are two critical factors in this system that need to be carefully considered.

1) TEGs' Electricity Generation Capacity: To evaluate the performance of TEGs, we need to adjust the temperature difference between the warm circulation and the cold circulation and measure the open-circuit voltage of the TEG module. We fix the temperature of the cold circulation T_{cold} at 20°C, and then adjust the outlet water temperature of CPU T_{warm_out} which is also the TEG module's inlet temperature of the warm circulation. The temperature difference between warm circulation and cold circulation ΔT is:

$$\Delta T = T_{warm\_out} - T_{cold} = T_{warm\_out} - 20. \tag{2}$$

First, we test the effect of flow rate on thermoelectric generation. We divide the twelve TEGs into two groups, and in each group six TEGs are connected in series. Then we measure the open-circuit voltage variation of two groups with temperature difference between warm coolant and cold coolant at different flow rates, shown in Fig. 7. The warm coolant and the cold coolant are set to the same flow rate each time. It is observed that the voltage increases linearly with the temperature difference between warm coolant and cold coolant. The larger the flow rate, the higher the voltage. Yet this improvement may be too little to be worth making, because a larger flow rate means more power consumption of the pump and more heat ejection, which will increase the workload of the cooling equipment.

Fig. 8a and Fig. 8b respectively show the measurement results of the open-circuit voltage and maximum output power



(a) (b)

Fig. 8: The impact of coolant temperature difference and the number of TEGs connected in series on (a) open-circuit voltage and (b) maximum output power.

when different numbers of TEGs are connected in series. Because the flow rate has little effect on thermoelectric generation of TEGs, we fix the flow rate at 200 L/H. We define v as the open-circuit voltage of one TEG and V_{oc_n} as the open-circuit voltage of n TEGs are connected in series electrically. We observe that V_{oc_n} is nearly n times of v. Then according to our measurement, we can model V_{oc_n} as a linear function of coolant temperature difference ΔT:

$$v = 0.0448\Delta T - 0.0051, \tag{3}$$

$$V_{oc\_n} = n \times v = n \times (0.0448\Delta T - 0.0051). \tag{4}$$

The maximum output power occurs when the load resistance equals the TEG module's resistance, and the load resistance of one TEG R_{TEG} is measured as 2 Ω. We define P_{max_n} as the maximum output power of a TEG module consisting of n TEGs and the curve can be modeled as a quadratic function of coolant temperature difference ΔT:

$$P_{max\_1} = \frac{(\frac{1}{2}v)^2}{R_{TEG}} = \frac{1}{8}v^2, \tag{5}$$

$$P_{max\_1} = 0.0003 \times \Delta T^2 - 0.0003 \times \Delta T + 0.0011, \tag{6}$$

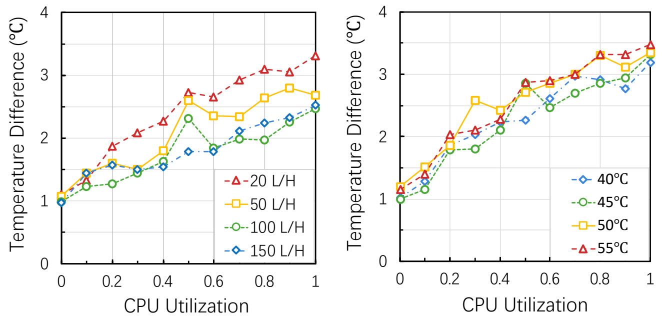
$$P_{max\_n} = n \times (0.0003 \times \Delta T^2 - 0.0003 \times \Delta T + 0.0011). \tag{7}$$

Note that the temperature difference in Fig. 8 ranges from 0°C∼25°C, this is because in our prototype, there is only one CPU releasing heat. In a real datacenter with thousands of CPUs, the temperature difference can be higher than 25°C, and the maximum output power of 12 TEGs can be higher than 1.8 W.

2) The Outlet Water Temperature of CPU: We define T_{warm_in} as the inlet water temperature of CPU (i.e., the temperature of water to cool CPU), and ΔT_{out-in} as the temperature difference between the outlet water and the inlet water of CPU. Hence, the outlet water temperature of CPU T_{warm_out} is

$$T_{warm\_out} = \Delta T_{out-in} + T_{warm\_in}. \tag{8}$$

As shown in Fig. 9a and Fig. 9b, ΔT_{out-in} fluctuates within 1°C∼3.5°C, and is mainly affected by CPU utilization, while the inlet temperature and flow rate have little effect on it.



(a) (b)

Fig. 9: Temperature difference between outlet water and inlet water with (a) varying CPU utilization and flow rate (this is an average result of four inlet water temperatures), (b) varying CPU utilization and inlet water temperature (the flow rate is fixed at 20 L/H).

3) CPU's Safety: CPU temperature is mainly affected by three factors: CPU utilization u, the inlet water temperature of CPU T_{warm_in}, and the flow rate f. We conducted several measurements on a Dell server equipped with an Intel Xeon E5 2650 V3 CPU, whose maximum operating temperature is 78.9°C [10].

First, we fix the flow rate at 20 L/H, and investigate the impact of coolant temperature and CPU utilization on CPU temperature. As can be seen in Fig. 10 , when CPU utilization exceeds 50%, the frequency starts to increase slower and finally settles down at about 2.5 GHz, because the CPU frequency governor is set to the "powersave" mode. The variation trend of CPU temperature roughly matches that of frequency.

Then, we conduct a measurement of CPU temperature with different coolant temperatures and flow rates when CPU utilization is 100%. As shown in Fig. 11, at each flow rate, CPU temperature grows linearly with coolant temperature. Given a fixed coolant temperature supply, CPU temperature lessens with increasing flow rate. This is because the faster the coolant flows, the more heat can be taken away from CPU per unit time. However, the improvement of the cooling efficiency is not unlimited, when the flow rate reaches a certain degree (e.g., above 250 L/H), it has little effect on CPU temperature. We observe that the flow rate also has an impact on the line's slope, which increases as the flow rate decreases.

C. Implications from Measurement Results

Based on the measurement results above, we can see that the electricity generation capacity of the TEG module in our proposed H2P is mainly determined by the outlet water temperature of CPU T_{warm_out}. Furthermore, T_{warm_out} is mainly determined by the inlet water temperature of CPU T_{warm_in}, since T_{warm_out} is only 1°C∼3.5°C higher than T_{warm_in}. Hence, the key to raising the generation capacity lies in raising T_{warm_in}. However, the upper limit of T_{warm_in} is determined by the hottest server in this circulation. Therefore, we should identify the optimal number of water circulations in a datacenter, dynamically adjust the cooling setting and

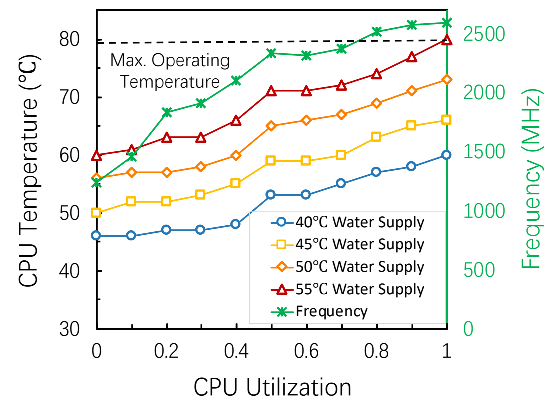


Fig. 10: CPU temperature variation with different coolant temperatures and CPU utilization (CPU Type: Intel Xeon E5-2650 V3, Frequency Governor: powersave, Maximum Operating Temperature: 78.9°C, Flow Rate: 20 L/H).

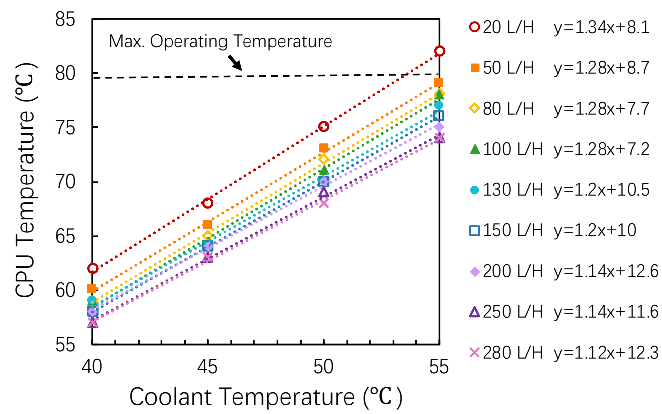


Fig. 11: CPU temperature variation with different coolant temperatures and flow rates (CPU Type: Intel Xeon E5-2650 V3, Frequency Governor: powersave, Maximum Operating Temperature: 78.9°C, CPU utilization: 100%).

schedule CPU load so as to optimize the electricity generation capacity of the TEG module.

V. ELECTRICITY GENERATION OPTIMIZATION AND TRACE-DRIVEN EVALUATION

Since the electricity generation is limited as shown by the measurement results, we investigate how to increase the generated electricity from both hardware-level and software-level. In particular, we first explore how to optimize the water circulation design for the construction of datacenters. Then, we investigate how to optimize the cooling control and the workload scheduling of all water-cooled servers to generate more electricity. We evaluate the performance of H2P by comparing two schemes and using different types of traces. At last, we provide a comprehensive analysis of the total cost of ownership (TCO).

A. Water Circulation Design

We consider H2P in a homogeneous datecenter equipped with a cluster of 1,000 servers. Generally, servers in one or several racks are controlled by one CDU and share the same water circulation, equipped with a chiller and a centralized pump. In this circulation, we assume that the inlet water temperature and the flow rate of each server are the same. As we mentioned in Sec. IV-C, the key to raising the generation capacity lies in raising T_{warm_in}, while the upper limit of T_{warm_in} is determined by the hottest server in this circulation. Hence, how many servers are arranged in a water circulation should be taken into consideration. Obviously, each server monopolizing one circulation is the most energy-efficient and can generate most electricity, because in this way we can adjust T_{warm_in} of each CPU. However, it is not practical to equip each server with a chiller and a pump, considering the technical limit and the capital expenses.

In practice, we suppose that a water circulation has n servers, so that 1,000 servers are divided into 1,000/n water circulations. In the i-th water circulation, the j-th CPU utilization is $u_{i,j}$ and its temperature is $T_{CPUi,j}$. The safe temperature of CPU is T_{Safe}, which can be pre-defined and adjusted (e.g., 80% of CPU's maximum operating temperature). Note that pro-longed operation at close to the maximum temperatures may cause CPU performance degradation and shorten the CPU lifespan [9]. Hence, $T_{CPUi,j}$ should be reduced by $\Delta T_{i,j} = T_{CPUi,j} - T_{Safe}$. According to the measurements in Fig. 11, T_{CPU} has a linear function of coolant temperature $T_{Coolant}$: $T_{CPU} = k \times T_{Coolant} + b, (k \in [1, 1.3])$. T_{warm_in} of the i-th circulation should be reduced by ΔT_i, k times of which is the maximum value in the set $\{\Delta T_{i,j}\}$, defined by:

$$k \times \Delta T_i = max\{\Delta T_{i,1}, ..., \Delta T_{i,n}\}. \quad (9)$$

The energy cost of the chiller in the i-th water circulation can be represented by:

$$E_{chiller\_i} = C_{water} * \Delta T_i * n * f * t * \rho / COP_{chiller}. \quad (10)$$

The total energy cost of all chillers can be represented by:

$$E_{chiller} = \sum_{i=1}^{1000/n} E_{chiller\_i}. \quad (11)$$

$C_{water} = 4.2 \times 10^3$ J/(kg·°C) is the heat capacity of water. It means 4.2×10^3 Joule heat needs to be added to (or removed from) 1-kilogram water to change the water temperature by 1 °C. $n * f * t$ is the total water volume that flows into heat exchanger during time t, where n is the number of servers and f is the flow rate assumed as a constant (e.g., 50 L/H) in that water circulation. ρ is the density of water. $COP_{chiller}$ (i.e., coefficient of performance) describes the energy efficiency of a chiller [49], which is determined by the ratio between the heat that chiller removes from water and the energy it consumes (i.e., COP = Heat removed by chiller / Energy consumed by chiller). We can assume $COP_{chiller}$ as 3.6 [24].

The optimization problem is to minimize the total cost containing electricity charge of chillers and the cost of chillers, represented by:

$$\min\{E_{chiller} * Charge_{electric} + (1000/n) * Cost_{chiller}\}. \quad (12)$$

Because the workload of a certain CPU is subject to the users' random actions, the CPU temperature that can be computed from the CPU utilization is a random variable subjected to normal distribution. We define n random variables $T_1, T_2, ..., T_n$, representing the temperatures of n CPUs and satisfying the relationship $T_i \sim N(\mu, \sigma^2)$. So we have the probability density function:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}, \quad (13)$$

and the probability distribution function:

$$F(x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\} dx. \quad (14)$$

$T_{(1)} \leq T_{(2)} \leq ... \leq T_{(n)}$ are order statistics of $T_1, T_2, ..., T_n$. For the maximum value $T_{(n)}$, its probability distribution function can be represented by:

$$F_{T_{(n)}}(x) = F^n(x), \quad (15)$$

and then we calculate the derivative to get its probability density function:

$$f_{T_{(n)}}(x) = n(F(x))^{n-1} f(x). \quad (16)$$

Hence, we can calculate the expectation of the highest temperature of n CPUs expressed as $E(T_{max})$:

$$E(T_{max}) = E(T_{(n)}) = \int_{-\infty}^{+\infty} x f_{T_{(n)}}(x) dx, \quad (17)$$

and the expectation of ΔT_i:

$$E(\Delta T_i) = (E(T_{max}) - T_{safe})/k. \quad (18)$$

In this way, the total cost in Eq. 12 can be expressed as a function of n. This exploration can give some suggestions on the design and construction of the future warm water-cooled datacenters.

B. Software-based Optimization Strategies

According to the measurement results in Sec. IV, the optimization problem is to optimize the maximum output power of the TEG module P_{TEG} by maximizing T_{warm_out}, while limited by CPU's safe operating temperature T_{safe}. The measurement results of CPU temperature T_{CPU} are plotted in Fig. 12. In this 3-D space formed by these discrete points, the coordinate of each point is (u, f, T_{warm_in}) (i.e., CPU utilization, flow rate, inlet water temperature), and the color of each point represents T_{CPU}. The darker the color, the higher the temperature. T_{warm_out} can also be plotted in such a 3-D space, because each $(T_{CPU}, f, T_{warm\_in})$ maps a T_{warm_out}. Hence, if we know the CPU utilization, and determine a cooling setting $\{f, T_{warm\_in}\}$, we can easily get T_{CPU} and T_{warm_out} from the measurement results, and then

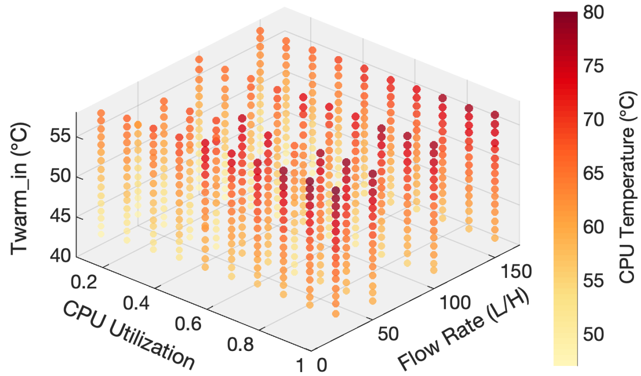


Fig. 12: The measurement results of CPU temperature can form a 3-D discrete space.

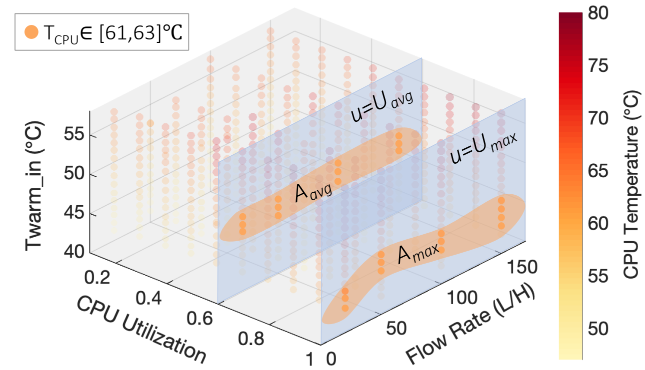


Fig. 13: Selecting the points whose CPU temperatures are between 61°C and 63°C (CPU utilization = U_{max} or U_{avg}, T_{safe} = 62°C).

calculate P_{TEG}. Note that although our measurements cannot cover all the points in the space, T_{CPU} changes continuously and linearly with its variables (see Fig. 10 and Fig. 11). Therefore, we can extend our limited measurements to a general relationship, and the discrete points can be fitted to a continuous space, which can function as a look-up space in practical use.

1) Adjusting Water Temperature and Flow Rate: In a water circulation with n servers, during each time interval t (e.g., 5 minutes), we determine T_{warm_in} and f as below:

- *Step1:* At the beginning of t, CPU utilization of n servers are $\{u_1, u_2, ..., u_n\}$, we find the maximum value U_{max} in this set and draw a plane $U : u = U_{max}$ which is perpendicular to the CPU Utilization axis.
- *Step2:* Set the CPU safe operating temperature T_{safe}, and we search in the space for the points whose T_{CPU} are in the proximity of T_{safe} (e.g.,within $[T_{safe} - 1, T_{safe} + 1]$ °C) to form a new space X.
- *Step3:* Let the plane U_{max} and the space X intersect and then get $\{f, T_{warm\_in}\}$ of each point in the area of this intersection: $A_{max} = U_{max} \bigcap X$. Compute P_{TEG} under each cooling setting $\{f, T_{warm\_in}\}$ (Eq. 2 and Eq. 7), and choose the maximum P_{TEG} to determine the optimal cooling setting.

For example, in Fig. 13, we set T_{safe} as 62°C. T_{CPU} of the points in the area A_{max} are between 61°C and 63°C.

2) Balancing Workload: As we mentioned in Sec. IV-C, the upper limit of T_{warm_in} is determined by the hottest server in this circulation. If we can balance the workload on all the servers, their CPU temperature can be close to each other. In other words, this strategy can flatten the cooling demands of servers in a loop, so that chillers can work less and T_{warm_in} can be set higher, which means T_{warm_out} and the power generated by TEGs will be higher.

Under this workload balancing strategy, when we aim to determine the cooling setting $\{f, T_{warm\_in}\}$, the plane we draw in *Step1* will be $U : u = U_{avg}$, where U_{avg} is the average value of $\{u_1, u_2, ..., u_n\}$. The intersection of U_{avg} and X will be A_{avg}. In Fig. 13, we can see that T_{warm_in} of the points in A_{avg} are generally higher than those in A_{max}. Hence, power generation can be optimized in this way.

C. Trace-Driven Evaluation

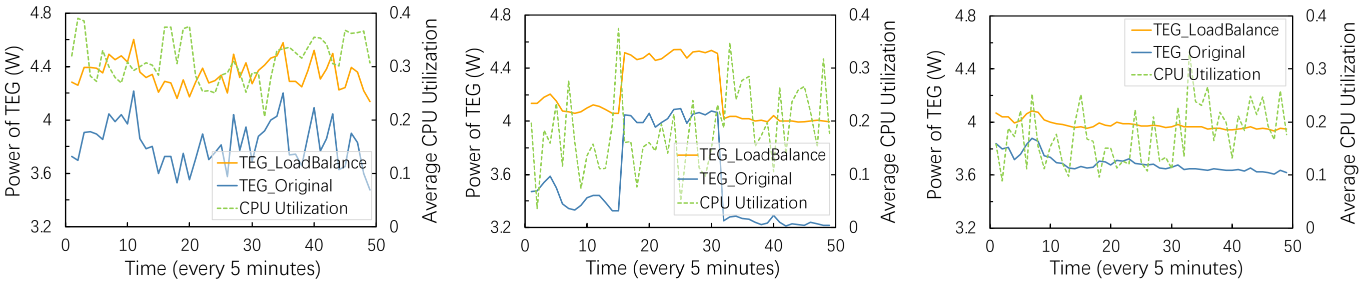
To evaluate the performance of H2P, we set three types of traces:

- **Drastic**: This trace is from Alibaba cluster [1], which contains the CPU usage of 1,313 servers in 12 hours. The variation of CPU utilization has drastic and frequent fluctuations.
- **Irregular**: Google trace [18] provides the CPU usage of 12.5k servers for a month. We select CPU utilization of 1,000 servers for 24 hours to compose an abnormal workload trace. The variation of CPU utilization is relatively common, but with occasional high peaks.
- **Common**: Similarly, we select another 1,000 servers for 24 hours from Google trace. The variation of CPU utilization is relatively common and has very little fluctuations.

We compare the power generation capacity of H2P under two schemes:

- ***TEG_Original***: a baseline that datacenters are equipped with TEGs and adjust the cooling setting but do not consider any workload scheduling methods.
- ***TEG_LoadBalance***: an optimization of *TEG_Original* by adding workload balancing strategy.

Fig. 14 shows the workload traces (*Drastic*, *Irregular*, *Common*) and the corresponding power generation capacity of H2P under three different workloads and two different workload scheduling methods (*TEG_Original* and *TEG_LoadBalance*). As plotted in Fig. 14a, when the CPU utilization is high, the corresponding power generation capacity of H2P is low. This is because when the CPU utilization of a server is high, its CPU temperature is also high, while the inlet water temperature should be low enough to cool down the sever. Under this low inlet water temperature, the power generation capacity is consequently low. The generated electricity of H2P shows the same trend in both Fig. 14b and Fig. 14c. In particular, the average generated power of *TEG_Original* under the three workload traces (*Drastic*, *Irregular*, *Common*)



(a) Drastic (b) Irregular (c) Common

Fig. 14: The electricity generation under three types of CPU utilization.

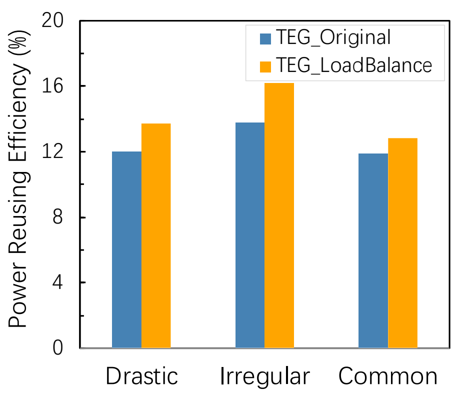


Fig. 15: Power reusing efficiency of TEG/CPU under three types of CPU utilization.

are 3.725 W, 3.772 W and 3.586 W, respectively (3.694 W averagely). The peak values of the generated power are 4.210 W, 3.935 W and 4.035 W respectively. After workload balancing optimization, the average generated power of *TEG_LoadBalance* are 4.349 W, 4.203 W, and 3.979 W for the three traces, respectively (4.177 W averagely). The peak values are 4.595 W, 4.554 W, and 4.082 W. Thus, we can conclude that the workload balancing strategy can effectively improve the average generated power by approximately 13.08%, (i.e., from 3.694 W to 4.177 W).

Power Reusing Efficiency: Because we only focus on recycling heat from CPU, not target on optimizing energy consumption in water circulation (e.g., chillers and pumps), which already exists in water cooling system even without H2P, we define the power reusing efficiency (PRE) as:

$$PRE = \frac{TEGs'\ Power\ Generation}{CPUs'\ Power\ Consumption}. \tag{19}$$

By our measurement of an Intel Xeon E5-2650 V3 CPU, CPU's power consumption can be expressed as a function of CPU utilization u with root mean square error less than 5 W, in the form of:

$$P_{CPU} = 109.71 \times log(u + 1.17) - 7.83. \tag{20}$$

By Eq. 20, we can estimate the power consumption of three different traces and compare CPU power consumption with TEGs' power generation. As shown in Fig. 15, PRE of *TEG_Original* are 12.0%, 13.8%, 11.9% for *Drastic*, *Irregular*, *Common* workloads, respectively. After workload balancing optimization, PRE of *TEG_LoadBalance* are 13.7%, 16.2%, 12.8% for the three trace, and the average PRE is 14.23%.

TABLE I: Parameters used to model TCO [27].

Description	Value	Unit
DCInfraCapEx	21.26	\$ / (server $\times$ month)
ServCapEx	31.25	\$ / (server $\times$ month)
DCInfraOpEx	7.63	\$ / (server $\times$ month)
ServOpEx	1.56	\$ / (server $\times$ month)
TEGCapEx	0.04	\$ / (server $\times$ month)
TEGRev (*TEG_Original*)	0.34	\$ / (server $\times$ month)
TEGRev (*TEG_LoadBalance*)	0.39	\$ / (server $\times$ month)

D. TCO Analysis

To calculate the total savings from TEGs, we analyse the total cost of ownership (TCO) of datacenters with and without H2P, respectively. The capital expenses, operating expenses and the revenue are list in Table. I.

Capital Expenses (CapEx): We have shown in Sec. III-A that the purchase price of an SP 1848-27145 TEG is \$1 per piece with a long lifespan of 28$\sim$34 years. We assume that a TEG can be used for at least 25 years. In H2P, a server is equipped with 12 TEGs. Hence, TEGCapEx can be calculated as 0.04 \$/(server$\times$month). Datacenter infrastructure capital expense (DCInfraCapEx) is the cost of building a datacenter amortized to each month, including the cost of land, uninterruptible power supply (UPS) and power infrastructure, cooling infrastructure and so on [27]. Note that server capital expense (ServCapEx) is not included in DCInfraCapEx.

Operating Expenses (OpEx): Since TEGs mostly do not need long-term maintenance, we neglect its operating expense. The operating expenses of servers and datacenter infrastructures are defined as ServOpEx and DCInfraOpEx.

Revenue (Rev): We define the revenue from the generated electricity by TEGs is TEGRev. According to the simulation result in Sec. V-C, average generated electricity under the two schemes (*TEG_Original* and *TEG_LoadBalance*) are 3.694 W and 4.177 W respectively. Considering the price of electricity is 13 cents/kwh [16], TEGRev (*TEG_Original*) is 0.34

\$/(server×month) and TEGRev (*TEG_LoadBalance*) is 0.39 \$/(server×month).

TCO: Suppose that there is a cluster with 100,000 CPUs equipped with 1,200,000 TEGs. The TEGRev (*TEG_LoadBalance*) in 24 hours can be evaluated as 10,024.8 kwh, (i.e., \$1,303.2/day). The purchase price for this system can be evaluated as the total price for TEGs, \$1×1,200,000. As a result, the break-even point of this system will be 920 days.

In order to better illuminate the benefit of integrating H2P to a datacenter, we calculate the TCO in a datacenter with and without H2P system respectively, which can be represented as follows:

$$TCO_{noTEG} = (DCInfraCapEx + ServCapEx) + (DCInfraOpEx + ServOpEx), \quad (21)$$

$$TCO_{H2P} = TCO_{noTEG} + TEGCapEx - TEGRev. \quad (22)$$

Consequently, *TEG_Original* and *TEG_LoadBalance* can respectively reduce TCO by 0.49% and 0.57%. For a datacenter with 100,000 CPUs, TCO_{H2P} can save \$350,000 ∼ \$410,000 for a year.

VI. DISCUSSION

In this section, we first analyze the limitation of TEGs, which limits the power generation capacity. Next, we explore how to store the generated electricity more effectively. Finally, we present some potential applications of TEG-enabled H2P in warm water-cooled datacenters and discuss the future development of TEGs and datacenters.

A. Limitation of TEG

Limited by the material and technology, the power generation efficiency of TEG is low at present. Besides, waste heat in datacenters is low-grade, making it hard to fully utilize TEG's advantages. The storage of power generated by TEGs is also a difficult problem. Hence, exploiting TEGs in datacenters is not meant to replace other heat recovery solutions (e.g., combining TEGs with CCHP). In this paper, we only focus on optimizing the maximum output electricity generated by TEGs cost-efficiently.

B. Electricity Storage for TEGs

Since the generated power of the TEG module is fluctuant and time-varying with the temperature difference, it may exceed or be lower than power demands if the TEG module is connected with electrical appliances directly. Usually, during the peak hours (e.g., midday to the evening) the CPU load is generally high [42], which requires lower inlet water temperature to cool CPUs, resulting in lower power generated by TEGs. While during the off hours (e.g., late night and early morning) when most users are asleep and CPU load is low, the inlet water temperature can be raised to a relatively high level, leading to higher power generated by TEGs. The power demands of different electrical appliances are also different. Hence, to handle the irregular power mismatches, energy storage devices (e.g., batteries, super-capacitors) are needed. Compared with batteries, super-capacitors (SCs) can achieve a higher energy efficiency of 90%∼95% while have a higher capital cost. To exploit the merits of both, a hybrid energy buffering system [31] comprising batteries and SCs has been proposed to shave peaking power mismatching between power demands of servers and renewable power sources. Similarly, a small-scale energy buffering system can be exploited for TEGs' electricity storage.

C. Potential Applications of TEGs

1) TEGs for powering TECs: As we mentioned above in Sec. II-B, in warm water-cooling systems, hot spots occur as a common problem where a hybrid cooling architecture integrating thermoelectric coolers (TECs) emerges as a new remedy [24]. Although this hybrid cooling system reduces the usage of chillers, its energy consumption still has room for further reduction. Because TECs bring extra energy consumption into the system; besides, the TEC removes heat from the CPU to the water quickly, which helps not only cool the CPU but also heat the water. Hence, compared with not working, the outlet water temperature of CPU is higher when TEC is working. It is more advantageous to integrate our H2P system into this hybrid cooling system than conventional water cooling systems.

2) TEGs for lighting: Though IT equipment and the cooling part account for the overwhelming majority of datacenter total energy consumption, we still cannot ignore other parts. Lighting costs representing 1%, are still a considerable expense. In fact, lighting in datacenters needs to meet the requirements of illuminance, lighting uniformity, lighting stability, and suppression of glare, while light-emitting diodes (LEDs) can be a perfect lighting solution [33]. The power of an ordinary LED is generally 0.05W and the working current is 20 mA, even high-power LEDs work at 1W and 2W [37], which can save more than 80% of energy than traditional light sources. Based on the experiment results, TEGs in H2P can generate 3W or more electricity, which is enough for supplying power for some of the LEDs used in datacenters.

D. Future Development of TEGs and Datacenters

The TEG used in H2P is SP1848-27145 made of Bi_2Te_3, which is currently the most widely-used TEG in both academia and industry [19], [40]. Its maximum ZT (i.e., the figure of merit, to evaluate whether a given material is a good thermoelectric material) is around 1 at 300-330 K [6], [36] and the conversion efficiency is approximately 5%. In recent years, however, remarkable progress has been made to create materials of a higher ZT, for example, recent laboratory results show that ZT of thin-film Heusler alloys based on $Fe_2V_{0.8}W_{0.2}Al$ can reach 6 at around 360 K [20]. However, this material is still in research, which needs a few years for commercial use. In addition, nanomaterials are also currently being developed for commercialization [36]. Once the new cheap materials of higher ZT are commercially available, a much wider application of these materials in datacenters is possible. Since TEGs only require a temperature difference to generate power, low-grade waste heat can be easily harvested in this way. Equipped with appropriate energy storage devices, the TEG can also be a backup power supply.

The centralized and distributed topologies are two primary energy storage architectures in datacenters currently [31]. To avoid energy losses resulting from double converting (AC-DC-AC) in the centralized UPS system, IT giants such as Google and Facebook have used decentralized batteries (i.e., 12V or 48V DC power) to supply servers. A cabinet of batteries can serve one or several racks. Future DC-based power distribution systems can improve energy utilization and reduce the total system energy use [13]. Our H2P system is appropriate for these DC-supplied datacenters.

VII. RELATED WORK

Cooling and energy management in datacenters: To reduce datacenter energy consumption and costs, previous works have provided energy monitoring and accounting at different levels (e.g., server-level [44], virtual machine-level [25], non-IT equipment-level [26]), cooling and power management at both software-level and hardware-level, primarily leveraging the strategy of under-provision or over-subscription. At software-level, Hsu et al. [21] propose a workload-aware service placement strategy by grouping services with asynchronous peak times under the profile of each power node to reduce power fragmentation and improve power utilization. Manousakis et al. [34] introduce CoolProvision to a free-cooling datacenter and handle the cooling mismatching resulted from under-provisioning by workload throttling (e.g., DVFS), consolidation and deferral. Such software-based solutions usually do not apply to delay-sensitive or performance-sensitive workloads, hardware-based approaches have been exploited to shave power/thermal peak without performance degradation. Liu et al. [31] propose hybrid energy buffers, including batteries and super-capacitors (SCs), and the latter help to overcome the limitations of the former. Skach et al. [41] flatten the cooling demand by the phase-change material (PCM), which absorbs heat by melting during the peak period and releases heat by solidification at the off-peak time. To actively control the melting of PCM and further dig the potential cooling capacity, they create a virtual melting temperature by locating hot jobs together [42]. Jiang et al. [24] propose a hybrid water cooling system, in which the thermoelectric coolers (TECs) and the chiller work together in a fine-grained manner to deal with hot spots under the strategy of warm water cooling. Different from previous works above, we view datacenters' waste heat as a resource that can be reused rather than be ejected into the environment and removed by cooling equipment or materials.

Waste heat reuse: Waste heat reuse is a new direction of energy saving, including heating buildings, converting heat into electricity by thermoelectric generators (TEGs) and so on. CloudHeat [7] is an online reverse auction mechanism for helping datacenters to sell heat to district heating systems. Liu et al. [30] propose "data furnace" by running servers in residential buildings and heating buildings locally. Apart from heating, several promising works on TEGs also have been conducted. First, Yazawa et al. [48] provide a theoretical analysis and proof-of-concept experiment on thermoelectric-powered forced convection cooling of a personal computer microprocessor. Next, Zhou et al. [50] also present a model to accurately estimate the TEG efficiency and they measure the generated power under varying processor workloads. Wu et al. [47] propose a TEG-integrated architecture to harvest waste heat generated by computing devices and transform it into electricity so as to be directly used within the system. They demonstrate that with three TEG modules, the harvested electrical energy is significant enough to power a fan. However, Jayakumar et al. [22] conclude that using thermoelectrics as TECs can provide boosts to performance with additional power consumption (i.e., TEGs increase the processor's temperature and leakage power). By taking processor temperature distribution and reliability into consideration, Lee et al. [28], [29] provide a self-sustaining, fine-grained, microarchitectural-level hot spot cooling mechanism, which attaches several small-sized TECs and TEGs on CPU and does not require additional energy to power the TECs for spot cooling. Dai et al. [11] also use small-scale TEGs to generate energy and further use the generated power to cool down hot spots and recharge micro-supercapacitors in the smartphone.

In summary, most of these works focus on integrating TEGs into PCs or air-cooled datacenters, where the cold side of TEG is exposed to air rather than attached to a stable cooling source. This will gradually reduce the temperature difference between the two sides of TEG and further degrade its electricity generation, which we have discussed in Sec. III-B with empirical experiments. Differently, in H2P, we attach the cold side of TEG to the natural water resource (around 20°C) to ensure its long-term electricity efficiency. Besides, some current TEG-integrated solutions are customized for certain CPU types, while H2P suits all types of CPUs.

VIII. CONCLUSION

In this paper, we propose Heat to Power (H2P), an economical thermal energy harvesting and recycling architecture for warm water-cooled datacenters. In H2P, thermoelectric generators (TEGs) are integrated to harvest thermal energy from the "used" warm water and generate electricity for reusing in datacenters. Considering the electricity generation is limited as shown by the measurement results, we accordingly propose some efficient optimization methods, including an economical water circulation design, fine-grained adjustments of the cooling setting and dynamic workload scheduling for increasing the power generated by TEGs. Based on a hardware prototype and real-world traces from Alibaba and Google, H2P can averagely generate 4.349 W, 4.203 W and 3.979 W (4.177 W averagely) electricity on one CPU under the *drastic*, *irregular* and *common* workload traces, respectively. The power reusing efficiency (PRE) can reach 12.8%∼16.2% (14.23% averagely) and the total cost of ownership (TCO) of datacenters can be reduced by up to 0.57%.

References

[1] Alibaba, "Alibaba cluster workload traces," https://github.com/alibaba/clusterdata[Online Accessed, 6-Apr-2020], 2019.

[2] Alibaba, "Price of thermoelectric generator," https://www.alibaba.com/product-detail/New-40-40mm-Thermoelectric-Power-Generator_62080985453.html?spm=a2700.7724838.2017115.76.78407917cmq3Va[Online Accessed, 6-Apr-2020], 2020.

[3] Asetek, "Newsroom," https://www.asetek.com/newsroom/[Online Accessed, 6-Apr-2020], 2020.

[4] ASHRAE, "Ashrae thermal guidelines," https://datacenters.lbl.gov/sites/all/files/ASHRAE%20Thermal%20Guidelines_%20SVLG%202015.pdf[Online Accessed, 6-Apr-2020], 2015.

[5] ASHRAE, "Water-cooled servers—common designs, components, and processes," https://tc0909.ashraetcs.org/documents/ASHRAE_TC0909_Water_Cooled_Servers_15_Oct_2019.pdf[Online Accessed, 6-Apr-2020], 2019.

[6] D. Champier, "Thermoelectric generators: A review of applications," *Energy Conversion and Management*, vol. 140, pp. 167 – 181, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196890417301851

[7] S. Chen, Z. Zhou, F. Liu, Z. Li, and S. Ren, "Cloudheat: an efficient online market mechanism for datacenter heat harvesting," *ACM Transactions on Modeling and Performance Evaluation of Computing Systems (TOMPECS)*, vol. 3, no. 3, p. 11, 2018.

[8] H. Coles, M. Ellsworth, and D. J. Martinez, ""hot" for warm water cooling," in *State of the Practice Reports*, ser. SC '11. New York, NY, USA: ACM, 2011, pp. 17:1–17:10. [Online]. Available: http://doi.acm.org/10.1145/2063348.2063371

[9] CPUWorld, "Minimum maximum operating temperatures," http://www.cpu-world.com/Glossary/M/Minimum_Maximum_operating_temperatures.html[Online Accessed, 6-Apr-2020], 2018.

[10] CPUWorld, "Intel xeon e5-2650 v3 specifications," http://www.cpu-world.com/CPUs/Xeon/Intel-Xeon%20E5-2650%20v3.html[Online Accessed, 6-Apr-2020], 2020.

[11] Y. Dai, T. Li, B. Liu, M. Song, and H. Chen, "Exploiting dynamic thermal energy harvesting for reusing in smartphone with mobile applications," in *Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems*, ser. ASPLOS '18. New York, NY, USA: ACM, 2018, pp. 243–256. [Online]. Available: http://doi.acm.org/10.1145/3173162.3173188

[12] R. Danilak, "Why energy is a big and rapidly growing problem for data centers," https://www.forbes.com/sites/forbestechcouncil/2017/12/15/why-energy-is-a-big-and-rapidly-growing-problem-for-data-centers/#6f1475d45a30[Online Accessed, 6-Apr-2020], 2017.

[13] W. Deng, F. Liu, H. Jin, B. Li, and D. Li, "Harnessing renewable energy in cloud datacenters: opportunities and challenges," *iEEE Network*, vol. 28, no. 1, pp. 48–55, 2014.

[14] K. Ebrahimi, G. F. Jones, and A. S. Fleischer, "A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 622 – 638, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032113008216

[15] N. El-Sayed, I. A. Stefanovici, G. Amvrosiadis, A. A. Hwang, and B. Schroeder, "Temperature management in data centers: why some (might) like it hot," *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1, pp. 163–174, 2012.

[16] Í. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Parasol and greenswitch: Managing datacenters powered by renewable energy," *ACM SIGPLAN Notices*, vol. 48, no. 4, pp. 51–64, 2013.

[17] Google, "Efficiency: How we do it," https://www.google.com/about/datacenters/efficiency/internal/[Online Accessed, 6-Apr-2020], 2020.

[18] Google, "Google cluster workload traces," https://github.com/google/cluster-data[Online Accessed, 6-Apr-2020], 2020.

[19] R. Goswami and R. Das, "Waste heat recovery from a biomass heat engine for thermoelectric power generation using two-phase thermosyphons," *Renewable Energy*, vol. 148, pp. 1280–1291, 2020.

[20] B. Hinterleitner, I. Knapp, M. Poneder, Y. Shi, H. Müller, G. Eguchi, C. Eisenmenger-Sittner, M. Stöger-Pollach, Y. Kakefuda, N. Kawamoto *et al.*, "Thermoelectric performance of a metastable thin-film heusler alloy," *Nature*, vol. 576, no. 7785, pp. 85–90, 2019.

[21] C.-H. Hsu, Q. Deng, J. Mars, and L. Tang, "Smoothoperator: Reducing power fragmentation and improving power utilization in large-scale datacenters," *SIGPLAN Not.*, vol. 53, no. 2, pp. 535–548, Mar. 2018. [Online]. Available: http://doi.acm.org/10.1145/3296957.3173190

[22] S. Jayakumar and S. Reda, "Making sense of thermoelectrics for processor thermal management and energy harvesting," in *2015 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*. IEEE, 2015, pp. 31–36.

[23] N. Jaziri, A. Boughamoura, J. Müller, B. Mezghani, F. Tounsi, and M. Ismail, "A comprehensive review of thermoelectric generators: Technologies and common applications," *Energy Reports*, 2019.

[24] W. Jiang, Z. Jia, S. Feng, F. Liu, and H. Jin, "Fine-grained warm water cooling for improving datacenter economy," in *Proceedings of the 46th International Symposium on Computer Architecture*, ser. ISCA '19. New York, NY, USA: ACM, 2019, pp. 474–486. [Online]. Available: http://doi.acm.org/10.1145/3307650.3322236

[25] W. Jiang, F. Liu, G. Tang, K. Wu, and H. Jin, "Virtual machine power accounting with shapley value," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, Atlanta, GA, USA, June 2017, pp. 1683–1693.

[26] W. Jiang, S. Ren, F. Liu, and H. Jin, "Non-it energy accounting in virtualized datacenter," in *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, Vienna, Austria, July 2018, pp. 300–310.

[27] V. Kontorinis, L. E. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, D. M. Tullsen, and T. S. Rosing, "Managing distributed ups energy for effective power capping in data centers," in *2012 39th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2012, pp. 488–499.

[28] S. Lee, D. Pandiyan, J.-s. Seo, P. E. Phelan, and C.-J. Wu, "Thermoelectric-based sustainable self-cooling for fine-grained processor hot spots," in *2016 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*. IEEE, 2016, pp. 847–856.

[29] S. Lee, P. E. Phelan, and C.-J. Wu, "Hot Spot Cooling and Harvesting Central Processing Unit Waste Heat Using Thermoelectric Modules," *Journal of Electronic Packaging*, vol. 137, no. 3, 09 2015, 031010. [Online]. Available: https://doi.org/10.1115/1.4030686

[30] J. Liu, M. Goraczko, S. James, C. Belady, J. Lu, and K. Whitehouse, "The data furnace: Heating up with cloud computing," in *Proceedings of the 3rd USENIX Conference on Hot Topics in Cloud Computing*, ser. HotCloud'11. Berkeley, CA, USA: USENIX Association, 2011, pp. 15–15. [Online]. Available: http://dl.acm.org/citation.cfm?id=2170444.2170459

[31] L. Liu, C. Li, H. Sun, Y. Hu, J. Gu, T. Li, J. Xin, and N. Zheng, "Heb: Deploying and managing hybrid energy buffers for improving datacenter efficiency and economy," in *Proceedings of the 42Nd Annual International Symposium on Computer Architecture*, ser. ISCA '15. New York, NY, USA: ACM, 2015, pp. 463–475. [Online]. Available: http://doi.acm.org/10.1145/2749469.2750384

[32] D. Lo, L. Cheng, R. Govindaraju, L. A. Barroso, and C. Kozyrakis, "Towards energy proportionality for large-scale latency-critical workloads," in *Proceeding of the 41st Annual International Symposium on Computer Architecuture*, ser. ISCA '14. Piscataway, NJ, USA: IEEE Press, 2014, pp. 301–312. [Online]. Available: http://dl.acm.org/citation.cfm?id=2665671.2665718

[33] R. Lopez, "Ip-addressable light-emitting diode," Aug. 22 2002, uS Patent App. 09/784,052.

[34] I. Manousakis, Í. Goiri, S. Sankar, T. D. Nguyen, and R. Bianchini, "Coolprovision: Underprovisioning datacenter cooling," in *Proceedings of the Sixth ACM Symposium on Cloud Computing*. ACM, 2015, pp. 356–367.

[35] F. Maurizio, "Rethinking chilled water temperatures can bring big savings in data center cooling," https://blog.se.com/datacenter/2016/08/17/water-temperatures-data-center-cooling/[Online Accessed, 6-Apr-2020], 2016.

[36] A. J. Minnich, M. S. Dresselhaus, Z. Ren, and G. Chen, "Bulk nanostructured thermoelectric materials: current research and future prospects," *Energy and Environmental Science*, vol. 2, no. 5, pp. 466–479, 2009.

[37] H. Murata, "Light emitting diode lamp," Jun. 19 1990, uS Patent 4,935,665.

[38] C. Nadjahi, H. Louahlia, and S. Lemasson, "A review of thermal management and innovative cooling strategies for data center," *Sustainable Computing: Informatics and Systems*, vol. 19, pp. 14–28, 2018.

[39] N. Petchers, *Combined heating, cooling & power handbook: technologies & applications: an integrated approach to energy resource optimization*. Fairmont Press, 2003.

[40] W. Punin, S. Maneewan, and C. Punlek, "Heat transfer characteristics of a thermoelectric power generator system for low-grade waste heat recovery from the sugar industry," *Heat and Mass Transfer*, vol. 55, no. 4, pp. 979–991, 2019.

[41] M. Skach, M. Arora, C.-H. Hsu, Q. Li, D. Tullsen, L. Tang, and J. Mars, "Thermal time shifting: Leveraging phase change materials to reduce cooling costs in warehouse-scale computers," in *Proceedings of the 42Nd Annual International Symposium on Computer Architecture*, ser. ISCA '15. New York, NY, USA: ACM, 2015, pp. 439–449. [Online]. Available: http://doi.acm.org/10.1145/2749469.2749474

[42] M. Skach, M. Arora, D. Tullsen, L. Tang, and J. Mars, "Virtual melting temperature: Managing server load to minimize cooling overhead with phase change materials," in *Proceedings of the 45th Annual International Symposium on Computer Architecture*, ser. ISCA '18. Piscataway, NJ, USA: IEEE Press, 2018, pp. 15–28. [Online]. Available: https://doi.org/10.1109/ISCA.2018.00013

[43] Y. Sverdlik, "Here's how much energy all us data centers consume," https://www.datacenterknowledge.com/archives/2016/06/27/heres-how-much-energy-all-us-data-centers-consume[Online Accessed, 6-Apr-2020], 2016.

[44] G. Tang, W. Jiang, Z. Xu, F. Liu, and K. Wu, "Zero-cost, fine-grained power monitoring of datacenters using non-intrusive power disaggregation," in *Proceedings of the 16th Annual Middleware Conference*, ser. Middleware '15. New York, NY, USA: ACM, 2015, pp. 271–282. [Online]. Available: http://doi.acm.org/10.1145/2814576.2814728

[45] TECMicrosystems, "Te cooler reliability testing and reports," https://www.tec-microsystems.com/faq/reliability-testing-and-reports.html[Online Accessed, 6-Apr-2020], 2019.

[46] B. Tschudi, O. Vangeet, J. Cooley, and D. Azevedo, "Ere: A metric for measuring the benefit of reuse energy from a data center," *White Paper*, vol. 29, 2010.

[47] C.-J. Wu, "Architectural thermal energy harvesting opportunities for sustainable computing," *IEEE Computer Architecture Letters*, vol. 13, no. 2, pp. 65–68, 2013.

[48] K. Yazawa, G. L. Solbrekken, and A. Bar-Cohen, "Thermoelectric-powered convective cooling of microprocessors," *IEEE transactions on advanced packaging*, vol. 28, no. 2, pp. 231–239, 2005.

[49] Z. Zheng, Q. Chen, C. Fan, N. Guan, A. Vishwanath, D. Wang, and F. Liu, "Data driven chiller sequencing for reducing hvac electricity consumption in commercial buildings," in *Proceedings of the Ninth International Conference on Future Energy Systems*. ACM, 2018, pp. 236–248.

[50] Y. Zhou, S. Paul, and S. Bhunia, "Harvesting wasted heat in a microprocessor using thermoelectric generators: modeling, analysis and measurement," in *2008 Design, Automation and Test in Europe*. IEEE, 2008, pp. 98–103.