# TEStore: Exploiting Thermal and Energy Storage to Cut the Electricity Bill for Datacenter Cooling

Yanwei Zhang[†], Yefu Wang[†], and Xiaorui Wang[†,‡]
[†]University of Tennessee, Knoxville, TN [‡]The Ohio State University, Columbus, OH
[†]{yzhang82, ywang38}@utk.edu [‡]xwang@ece.osu.edu

***Abstract*—The electricity cost of cooling systems can account for 30% of the total electricity bill of operating a data center. While many prior studies have tried to reduce the cooling energy in data centers, they cannot effectively utilize the time-varying power prices in the power market to cut the electricity bill for data center cooling. This is in contrast to the fact that various thermal and energy storage techniques available in today's data centers, such as ice or water-based thermal tanks and UPS batteries, can be utilized to store energy when the power price is relatively low. The stored energy can then be used to cool the data center when the power price is high.**

**In this paper, we design and evaluate TEStore, a cooling strategy that exploits thermal and energy storage techniques to cut the electricity bill for data center cooling, without causing servers in a data center to overheat. The proposed TEStore system checks the low prices in the hour-ahead power market and precools the thermal masses in the data center, which can then absorb heat when the power price increases later. Meanwhile, TEStore also checks the energy level in UPS batteries and exploits it as a complementary method in shifting energy demand for data center cooling. On a longer time scale, TEStore is integrated with auxiliary thermal tanks, which are recently adopted by some data centers to store energy in the form of ice. We model the impacts of TEStore on server temperatures based on Computational Fluid Dynamics (CFD) to consider the realistic thermal dynamics in a data center with 1,120 servers. We then evaluate TEStore with workload traces from real-world data centers and power price traces from a real power market. Our results show that TEStore can achieve the desired cooling performance with a much lower electricity bill than the current practice.**

## I. Introduction

With the increasing high server density, cooling has been consuming a major portion of the total operational cost of data centers (*e.g.*, 30% [1]). Reducing cooling cost, *i.e.*, the electricity bill for the energy consumed by the Computer Room Air Conditioning (CRAC) system, is an important concern for data center operators. Many existing solutions (*e.g.*, [2, 3]) concentrate on reducing the cooling energy. However, the energy consumed at different time contributes differently to the cooling cost because the electricity price in the market may fluctuate significantly. For example, in a power market that provides the price in the next hour based on a bidding mechanism (referred to as hour-ahead market), such as the CAPITAL area of New York Independent System Operator (NYISO) [4], the electricity price is below \$40/MWH between 0:00AM to 6:00AM and above \$50/MWh between 9:00AM and 12:00AM, on the day of April 1, 2011. At 9:45AM, the price rises sharply from \$62.91/MWh to \$147.11/MWh.

The fluctuation in power price thus suggests a method to reduce the cooling cost by exploiting some existing energy storage techniques available in data centers. First, various thermal masses in a data center, such as cold air, recirculation air handler coils, supply air ducts, and raised metal floor can be pre-cooled by the CRAC system to a low temperature, such that they can absorb heat later as thermal reserve space [5]. Consequently, when the power price becomes higher later, data centers can choose to use less electricity from the power grid for cooling, resulting in a significantly reduced electricity bill. Second, some data centers [6, 7] have started to adopt auxiliary tanks to store energy in the form of ice, chilled water, etc, for long-term (e.g., daily) thermal storage. Finally, battery-based energy storage techniques have also been recently utilized to buffer energy in today's data centers. For example, the energy stored within the batteries of the uninterrupted power supply (UPS) units has been exploited to optimize server power cost [8, 9]. However, the impacts of energy storage techniques on server temperatures in the data center and their potential on reducing electricity costs have not yet been analytically evaluated.

In order to most effectively utilize the time-varying power prices by exploiting various storage techniques available in a data center, there are some important observations regarding different storage techniques. First, the auxiliary thermal tanks have a high capacity which can buffer a considerable amount of energy [10, 11]. This thus indicates a good opportunity in reserving energy on a long time scale. Second, in contrast to thermal tanks, UPS batteries can only hold a limited amount of energy, because it is designed as a short-time energy buffer before the diesel generator is warmed up to keep the data center powered upon a local utility failure. UPS battery storage can thus serve as a complementary method on a relatively shorter time scale with respect to thermal tanks in buffering energy for data center cooling. Finally, thermal masses is also better to be considered as a short-term storage solution. This is due to the fact that (i) it has a low capacity in storing heat, and (ii) in order to keep the stored energy, the data center needs to reach a temperature lower than the level required by servers, which thus may incur energy waste. Consequently, in order to optimally exploit thermal masses in cutting the electricity bill for data center cooling, thermal masses should be exploited along with UPS batteries on a relatively short time scale, compared to thermal tanks.

In this work, we evaluate the potential cooling cost savings by exploiting thermal masses, auxiliary thermal tanks and UPS batteries available in a data center. To this end, we design TEStore, a smart cooling framework that dynamically

978-3-901882-48-7 

leverages thermal and energy storage techniques on different time scales to cut the electricity bill for data center cooling, without causing servers in a data center to overheat. In particular, TEStore leverages auxiliary thermal storage tanks in buffering energy on the long time scale to utilize the time-varying power prices in the power market. On a shorter time scale, TEStore checks the energy level in UPS batteries and exploits it as a complementary method to thermal tanks in buffering energy for data center cooling. At the same time, TEStore also checks the electricity price in the following invocation period and precools the thermal masses in the data center on a short time scale, which can then absorb heat when the power price increases later. Note that we assume that the electricity price in the next invocation period is known ahead of time. This is a valid assumption in many power markets such as the hour-ahead market [4]. In other types of power markets, the power price of the next invocation period can be predicted with a high accuracy [12]. TEStore then deploys a dynamic control algorithm that optimizes the time-average cooling bill for data center operators, based on Lyapunov optimization [13, 14], which requires no knowledge of future system statistics such as the power prices over time. To handle workload dynamics, TEStore dynamically decides how much energy should be charged to or discharged from the thermal tanks, UPS batteries, and thermal masses in every invocation period. Furthermore, TEStore quantifies the time and energy consumption to precool a data center based on the analytical results from Computational Fluid Dynamics (CFD), a powerful mechanical fluid dynamic analysis approach. Specifically, this paper makes the following contributions:

- We propose a novel precooling strategy that pre-cools a data center when the power price is low. We use CFD to systematically model the thermal dynamics of a data center to derive the energy losses when using the thermal masses inside the data center as a thermal storage system.
- We design a dynamic control algorithm based on Lyapunov drift and Lyapunov optimization to exploit auxiliary storage devices, such as thermal tanks and UPS batteries, in cutting the time-average cooling bill.
- We present the design of TEStore, a smart cooling framework that dynamically coordinates three kinds of storage techniques available in a data center, *i.e.*, thermal tanks, UPS batteries, and thermal masses, to run on different time scale for significantly reduced electricity bill for cooling, despite time-varying workloads and electricity prices.
- We evaluate TEStore in a data center with 1,120 servers with workload traces from real-world data centers and power price traces from a real power market.

The rest of the paper is organized as follows. Section II introduces the overall architecture of the proposed TEStore system. Section III presents the modeling and the algorithm for TEStore. Section IV discusses the simulation strategy and evaluates TEStore with real-world traces. Section V reviews the related work and Section VI concludes the paper.

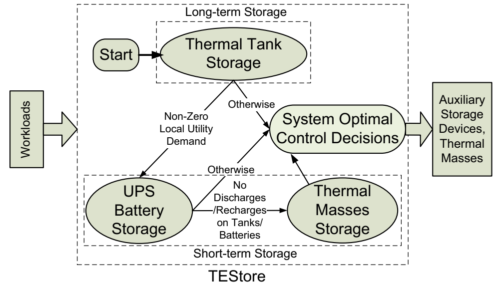



Fig. 1. Control flow of the proposed TEStore system.

## II. Design Methodology of TEStore

In this section, we present a high-level description of TEStore. In particular, in response to time-varying workloads and electricity prices, TEStore periodically determines the optimal control strategy in exploiting different storage techniques to cut the electricity bill for data center cooling, such as (i) when to leverage which type of storage techniques, and (ii) how much to discharge or recharge thermal tanks and UPS batteries, and how much to precool a data center, without causing servers in the data center to overheat.

### A. TEStore Coordination Strategy

In this part, we discuss in detail how TEStore integrates thermal masses, UPS batteries, and thermal tanks, on different time scales. Figure 1 demonstrates the way how the proposed TEStore system coordinates different storage strategies in every invocation period. In particular, TEStore begins with exploiting the long-term storage, and coordinates different storage strategies as follows.

1) Based on the incoming workload from customers and the power price in the local power market, as well as the current status of the long-term storage system (*e.g.*, the energy level in the thermal tanks), TEStore determines the optimal control decision on how much to recharge or discharge the thermal tanks, and thus the amount of energy that should be drawn from the local utility, according to the proposed algorithm on exploiting auxiliary storage devices to reduce cooling bill for data center operators as discussed in Section III-B.
2) If TEStore determines a non-zero amount of energy demand from the local utility after Step One (indicating that the long-term storage itself cannot fully service the incoming workload of the current invocation period), TEStore starts to exploit the short-term UPS battery storage as a complementary method. In particular, a similar algorithm as in Step One is employed in leveraging UPS battery storage based on the non-zero amount of local utility demand from Step One.
3) If TEStore determines no expected discharge and recharge operations on either thermal tanks or UPS batteries after Steps One and Two, that is, neither the long-term thermal tank storage nor the short-term UPS battery storage is used in the current invocation period, TEStore checks the short-term precooling strategy for opportunities in cutting the electricity bill for data center cooling. An optimal precooling decision is then determined according to the precooling strategy discussed in

Section III-A.

4) At the end, TEStore enforces the system control decision from Step One and Step Two, or from Step Three. That is, a consequent optimal decision on how much to discharge or recharge thermal tanks and UPS batteries, or how much to precool the thermal masses is performed for the current invocation period.

### B. Short-term Storage System

On a short time scale, TEStore checks to precool the thermal masses in a data center, which can then absorb heat when the power price in the local power market increases later.

In particular, the thermal masses, *i.e.*, masses inside the data center such as the cold air, the recirculation air handler coils, can be used for thermal storage. The data center can be precooled by adjusting the CRAC output temperature, *i.e.*, the temperature of the air blowing out from the CRAC units. The CRAC output temperature is adjustable via tuning the temperature of the water pumping into the CRAC units. The temperature of the water can in turn be tuned either by using a variable capacity compressor (for systems based on direct expansion units), or by modulating a chilled water supply valve (for systems based on water cooled units) [5], or by adjusting the amount of cold water pumped out from the auxiliary thermal storage. Compared with putting chillers on battery-based energy storage systems which can also be used to store energy on the short time scale, the thermal masses has almost negligible initial cost and lower operating cost, because it does not require purchasing additional equipment and replacing batteries frequently. However, the major limitation for the thermal masses as a thermal storage system is that they can only store a limited amount of heat. Also, to keep the stored energy in the storage system, we need to keep the data center at a temperature lower the level required by the servers, which may incur energy waste. Thus, to minimize the energy losses, in our TEStore system, we only precool the data center in a minimum length of time (*e.g.*, 5 minutes) before the energy price increases and use all the stored thermal energy in the thermal masses immediately after the price increases.

At the same time, TEStore also checks the energy level in UPS batteries and exploits it on the short time scale as a complementary method in shifting energy for data center cooling. In particular, UPS batteries are typically designed to hold enough capacity to power the entire data center at its maximum power needs for anywhere between 5-30 minutes, which thus indicates a good opportunity to shift data center power demand from time to time [8].

### C. Long-term Storage System

The auxiliary thermal storage tanks can be used to buffer energy on a much longer time scale, compared to the thermal masses and the UPS batteries existing in today's data center.

In this paper, we assume an ice-based thermal storage tank [10] for the data center cooling usage. The energy loss rate of exploiting a typical ice-based energy storage tank can be analyzed as follows. First, a tank has an energy loss of about 1% to 5% per day caused by heat loss. Second, the energy loss for a chiller becomes more significant with a lower chiller output temperature. In order to turn water into ice for an ice-based tank, a chiller needs to cool the water to the phase change temperature of $0°C$, which is significantly lower than the temperature of the water used in water-based CRAC systems ($10°C$ to $17°C$) and thus incurs more energy waste. On the other hand, because the ice is made at night, the low temperature in the environment reduces the energy consumption of the chillers. The transmission loss of the electric power becomes lower at night. Overall, considering all the energy losses and gains, a typical ice-based thermal storage system has an overall energy loss of about 13% [11]. Furthermore, the energy loss rate for a thermal storage tank becomes lower when better thermal materials are used in the tank. For example, if using phase change materials (PCM) such as hydrated salts with a phase change temperature of about $8.3°C$ [11] instead of water, the energy loss of the chiller becomes smaller. Note that because the energy loss is mainly induced in the recharging operations to cool the water to the phase change temperature of $0^oC$, we consider that the energy loss mainly happens in the recharging operations when TEStore exploits thermal tanks on the longer time scale.

## III. Algorithm Design

In this section, we discuss the algorithms we design to exploit different storage techniques on different time scales in cutting the electricity bill for data center cooling.

### A. Precooling Strategy

On a short time scale, to reduce cooling cost by precooling the data center before the energy price increases, it is important to quantify the increased cooling energy consumption when precooling the data center and the decreased energy consumption after the precooling process. Therefore, we first establish a model between the cooling energy and the CRAC output temperature. We then design our precooling algorithm.

*1) Cooling Energy:* The cooling power of a data center can be modeled as [3, 2] :

$$CoolingPower = HeatRemoved/COP \tag{1}$$

where $COP$ is the coefficient of performance, a variable that depends on the output temperature of the air conditioning facility. We use a COP model from a chilled-water CRAC unit at the HP Labs Utility Data Center [2]:

$$COP(T_{out}) = (0.0068T_{out}^2 + 0.0008T_{out} + 0.458) \tag{2}$$

where $T_{out}$ is the outlet temperature of the CRAC system.

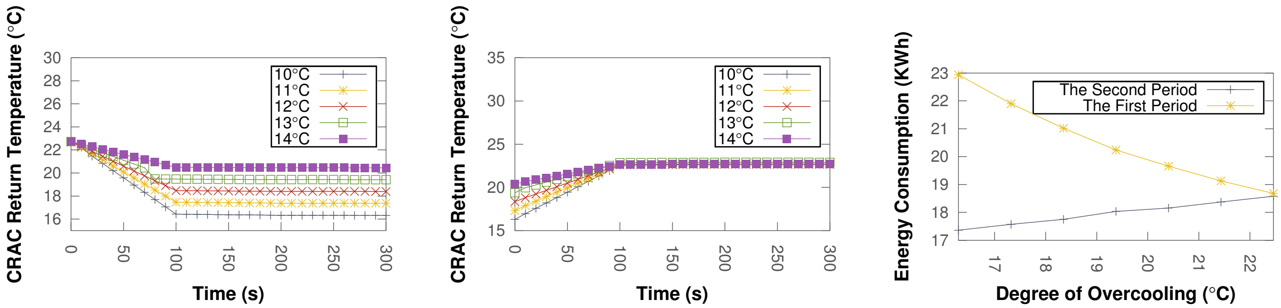
The heat removed in (1) is calculated as [15]:

$$HeatRemoved = m * C_p * (T_{in} - T_{out}) \tag{3}$$

where $m$ is the mass flow rate. $C_p$ is the specific heat, a constant number. $T_{in}$ is the inlet temperature of the CRAC system. The constant parameter $m * C_p$ can be obtained by comparing the removed heat and input/output temperature of the CRAC system.

Therefore, the cooling energy in each invocation period is:

$$\begin{aligned} Energy(k) &= \int_{kT}^{(k+1)T} CoolingPower(k)dt \\ &= \int_{kT}^{(k+1)T} \frac{m * C_p * (T_{in} - T_{out})}{COP(T_{out})} \end{aligned} \tag{4}$$




(a) Average temperature of the air returning to the CRAC unit when the CRAC output temperature is turned down to precool the data center at time 0.

(b) Average temperature of the air returning to the CRAC unit after precooling.

(c) Cooling energy when we precool the data center in the first period and return to the traditional cooling strategy in the second period. Precooling degree is plotted as the CRAC input temperature.

Fig. 2. Energy consumption modeling under different cooling strategies.

The impacts of TEStore on server temperatures, *i.e.*, how the CRAC input temperature $T_{in}$ in (4) changes when adjusting the CRAC output temperature $T_{out}$, can be modeled based on CFD analysis. CFD is a fluid mechanics approach that analyzes problems of fluid flows based on numerical methods and algorithms. Several popular software such as AirPAK and Fluent can be used for CFD modeling. These software packages allow modeling the thermal behavior of today's data centers and quantify the temperature changes over time.

We derive the energy consumption to precool the data center and the energy consumption after the data center returns to a commonly used traditional cooling strategy that does not consider cooling energy storage. We first derive how $T_{in}$ changes as a function of time. Therefore, we perform a series of experiments in the CFD simulation environment. We fix the utilization of the data center to a level (*e.g.*, 25%). For the data center cooled to a steady state with a traditional cooling strategy, we turn down the CRAC output temperature to different levels lower than the one required by the traditional cooling strategy and plot the temperature changes in Figure 2(a). Similarly, we plot the temperature changes when the CRAC output temperature is turned from several low levels back to the level required by the traditional cooling strategy in Figure 2(b). Then, as demonstrated in Figure 2(c), for different depth of precooling in terms of the CRAC input temperature, we derive the energy consumption to precool the data center and the energy consumption when the data center returns from an precooled state using the temperature curves in Figure 2(a), 2(b) and Equation (4). We then repeat the process for all levels of data center utilizations.

*2) Algorithm Design:* Our algorithm is invoked periodically with the same period that the price changes, *e.g.*, 15 minutes in the hour-ahead market in New York, or an hour in the day-ahead market in the same region. Our algorithm only optimizes the cooling cost in two adjacent periods. This is because the settling time of the cooling system (*e.g.*, 3-4 minutes [3]) is typically shorter than the period that the price changes (*e.g.*, 15 minutes or 60 minutes). It can be proved that, in this case, it is sufficient to consider only two adjacent periods. In an offline profiling, we derive the energy consumption curve in two adjacent periods if we precool the data center in the first period and return to the traditional cooling strategy in the second one, *i.e.*, a curve similar to Figure 2(c), for every utilization level. In every period, our algorithm checks the electricity price in the current period and in the next period in the power market. Then based on the current level of utilization in the data center, the algorithm checks the energy consumption curve (Figure 2(c)) and finds a degree of precooling $T_D$ with the least electricity cost in the two periods if we precool the data center in the first period and return to the traditional cooling strategy in the second one.

Thus, our algorithm precools the data center if the predicted cost is lower than the traditional cooling strategy. In order to use the minimum energy to achieve the desired depth of precooling at the end of the current period, we use the traditional cooling strategy between time $kT$ and $(k+1)T-T_s$, and turns the CRAC output temperature to $T_D$ from time $(k+1)T-T_s$ to time $(K+1)T$ where $T_s$ is the settling time of the temperature of the data center. Because our algorithm either uses the traditional cooling strategy, or uses a CRAC output temperature lower than the value that the traditional cooling strategy uses, our algorithm achieves the same or better cooling performance as the traditional cooling strategy does, *i.e.*, the temperature of the 80 hottest servers will not exceed the desired value of $25°C$. Note that we assume that the electricity price in the near future is known, which is a valid assumption in several types of power markets such as the hour-ahead market. In other types of power markets, our algorithm can be integrated with price prediction algorithms [12] to achieve short-term cost savings.

### B. Cooling Strategy in Exploiting Auxiliary Storage Devices

In this section, we propose a dynamic control algorithm based on Lyapunov optimization [13, 14] to exploit auxiliary storage devices such as thermal tanks and UPS batteries in cutting the time-average cooling bill. The theory of Lyapunov optimization is commonly used in optimizing the time averages of certain quantities subject to time-average constraints on other quantities, which can then be solved with a common mathematical framework that is intimately connected to queueing theory. The resulting optimal control actions are chosen over time solely in reaction to the existing system state, *e.g.*, the energy level of the auxiliary storage devices, the incoming workload and power prices in the current time slot, without the knowledge of future system statistics. The dynamic control algorithm can reach the optimal time-average electricity bill by the distance $O(1/V)$, where $V$ is limited by the capacity of the storage devices.

*1) Workload Model:* To make TEStore a practical solution, the system workload $U(k)$ (in units of data center CPU utilization) in every single invocation period should be fully serviced with the required cooling energy which is denoted as $W(k)$ (in units of energy), as derived in Section III-A1. A combination of the energy drawn from utility and the energy in auxiliary storage devices can be used to cool the data center in order to obtain this goal. Let $P(k)$ be the total energy drawn from the power grid in the invocation period $k$, out of which $R(k)$ is the total amount of energy used to recharge the storage devices deployed in the data center. Without loss of generality, $lf_r$ represents the energy loss rate whenever recharging the storage devices (*e.g.*, 13% for thermal tanks). Thus, $(1-lf_r)\cdot R(k)$ is the actual amount of energy recharged into the devices. $D(k)$ denotes the total energy discharged from the storage devices in the invocation period $k$. With the loss rate of $lf_d$ while discharging the storage devices, $(1-lf_d)\cdot D(k)$ is the actual amount of energy used to service the workload in the invocation period $k$. Thus, the following constraint must be satisfied in every single invocation period in order to guarantee the data center cooling performance

$$W(k) = P(k) - R(k) + (1 - lf_d) \cdot D(k)$$

Furthermore, the average cooling power consumption of a data center should not exceed a power constraint $P_{peak}$ of the invocation period $k$ due to the local utility limitation. Thus, we have $0 \leq \frac{P(k)}{T} \leq P_{peak}$, given the duration $T$ of a single invocation period.

*2) Auxiliary Storage Device Model:* In this part, we discuss the characteristics of the recharge and discharge behaviors on the auxiliary storage devices. We assume that a storage device should not be recharged and discharged simultaneously within the same invocation period of the proposed cooling strategy. We thus have

$$R(k) \cdot D(k) = 0$$

Due to the fact that frequent discharge and recharge operations may have a negative impact on the lifetime of the storage devices, we introduce an amortized lifetime-related cost (in dollar per operation), denoted as $C_{dc}$ and $C_{rc}$ for discharge and recharge, respectively. For example, regarding a new battery with the cost of $M$ dollars and a lifetime of $N$ discharge/recharge cycles, one can expect $C_{dc} = C_{rc} = \frac{M}{N}$, given the assumption of a similar impact on battery lifetime by recharge and discharge operations. This is consistent with a recent study [8].

Second, let $Y(k)$ denote the energy level in the auxiliary devices at the beginning of the invocation period $k$. It is clear that the storage devices should have a finite capacity $Y_{max}$ such that $Y(k) \leq Y_{max}$. Without loss of generality, a lower bound $Y_{min}$ on the energy level is also assumed. To this end, we have another constraint

$$Y_{min} \leq Y(k) \leq Y_{max}$$

The dynamics of the energy stored in the devices then follows $Y(k+1) = Y(k) + (1-lf_r)\cdot R(k) - D(k)$.

Next, we assume that the maximum amount by which we can recharge or discharge the storage devices in a data center in every invocation period is bounded by $R_{max}$ and $D_{max}$, which is consistent with the study in [8]. Thus, two more constraints are

$$0 \leq (1 - lf_r) \cdot R(k) \leq R_{max}, 0 \leq D(k) \leq D_{max}$$

*3) Optimization Formulation:* The objective of TEStore in exploiting the auxiliary storage devices is to dynamically determine the control actions such as $P(k)$, $R(k)$ and $D(k)$ over time, in order to minimize the time-average cooling bill for data center operators while guaranteeing the required cooling energy and complying to the constraints rendered by the operations on storage devices. We now explicitly formulate the problem as follows.

**Original Optimization Problem:**

$$Minimize: \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\{P(k)C(k) + I_{D(k)}C_{dc} + I_{R(k)}C_{rc}\} \tag{5}$$

subject to:

$$W(k) = P(k) - R(k) + (1 - lf_d) \cdot D(k) \tag{6}$$

$$0 \leq \frac{P(k)}{T} \leq P_{peak} \tag{7}$$

$$R(k) \cdot D(k) = 0 \tag{8}$$

$$Y_{min} \leq Y(k) \leq Y_{max} \tag{9}$$

$$0 \leq (1 - lf_r) \cdot R(k) \leq R_{max} \tag{10}$$

$$0 \leq D(k) \leq D_{max} \tag{11}$$

where the expectation $\mathbb{E}$ in the objective function is used to capture the potential randomness of the control decision in every invocation period. $K$ represents the total number of invocation periods in the simulation of minimizing the time-average cooling bill for data center operators by TEStore. $C(k)$ is the electricity price of the $k^{th}$ invocation period, and $I_{D(k)}$ and $I_{R(k)}$ are indicator functions on discharge and recharge operations of the auxiliary storage devices, respectively.

$$I_{D(k)} = \begin{cases} 1 & D(k) > 0 \\ 0 & else \end{cases} \quad I_{R(k)} = \begin{cases} 1 & R(k) > 0 \\ 0 & else \end{cases}$$

*4) Algorithm Design:* The uncertainty of the system inputs such as the time-varying incoming workloads and power prices over time make it difficult in solving the original problem (5) - (11). Even if the statistics of the system inputs are known ahead of time, the finite capacity of the storage devices and the constraints (9), (10), and (11) still make it a challenging problem to solve. Traditional methods, such as Dynamic Programming [16], may result in a highly complex computation effort, since they need to compute the optimal control decisions out of all possible combinations of the storage device charge level and the system inputs such as $W(t)$. Therefore, we introduce a Lyapunov optimization-based dynamic control algorithm to solve the optimization problem of minimizing the time-average cooling bill [8, 13]. Specifically, we first relax the constraint (9) to be a time-average quantity as shown in Equation (12), because the total discharged energy and recharged energy on the storage devices should be balanced in the long run [8].

$$\bar{R} - \bar{D} = 0 \tag{12}$$

where $\bar{R} = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\{(1-lf_r)R(k)\}$ and $\bar{D} = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\{D(k)\}$. In particular, $\bar{R}$ and $\bar{D}$ are the time-average quantity on the amount of energy recharged into, and discharged from the storage devices, respectively, and the expectation is with respect to the randomness of the control policy. The original problem is then be transferred into an alternative problem with the same objective function as (5), but subject to constraints (6) (7) (8) (10) (11) and (12). Being next, a virtual queueing state $H(k)$ is introduced as Equation (13) to meet the time-average constraint (12). The dynamics of the virtual queues then follows Equation (14), according to the dynamics of the energy that is stored in the auxiliary storage devices, *i.e.*, $Y(k+1)$, as discussed in Section III-B2.

$$H(k) = Y(k) - V \cdot \chi_{min} - D_{max} - Y_{min} \tag{13}$$

$$H(k+1) = H(k) + (1-lf_r) \cdot R(k) - D(k) \tag{14}$$

where $\chi_{min} = max[C(0), C(1), C(2), \cdots]$ is to ensure the performance guarantee of the proposed *CCMA* algorithm. As a result, according to the theory of Lyapunov optimization [13, 14], an alternative problem formulation is generated, based on which our dynamic control algorithm is developed.

**Alternative Optimization Problem:**

$$\begin{aligned} Minimize: & H(k)[(1-lf_r)R(k) - D(k)] \\ & + V[P(k)C(k) + I_{D(k)}C_{dc} + I_{D(k)}C_{rc}] \end{aligned} \tag{15}$$

subject to: (6) (7) (8) (10) and (11).

*Cooling Cost Minimization Algorithm (CCMA)*: At the beginning of every invocation period $k$, this algorithm observes the current virtual queueing state $H(k)$ and the system incoming events ($W(k)$, $C(k)$), and then makes an optimal control decision on ($P(k)$, $R(k)$, $D(k)$), according to the optimization problem (15). The virtual queueing state is then updated for the next invocation period $(k+1)$, according to the dynamics of the virtual queues (14).

*5) Algorithm Performance:* The proposed dynamic control algorithm computes the control decision in every invocation period towards the minimization of the time-average cooling bill, without knowledge of future system statistics. In every invocation period, the algorithm tries to recharge the storage devices with $H(k)$ being negative (which indicates that the current energy level of the storage devices is lower than some threshold as defined in Equation (13)) and a lower power price $C(k)$, and it will discharge the storage devices when the energy level of the storage devices is higher than the threshold, *i.e.*, $H(k)$ being a positive value. Instead of making decisions purely based on the power prices, which may result in a sub-optimal cooling bill [8], the proposed dynamic control algorithm approaches the optimal control decisions by observing the energy level in the storage devices, as well as the power price. More rigorous mathematical proofs regarding the algorithm performance can be found in [8, 13, 14].

## IV. System Evaluation

In this section, we conduct extensive experiments to evaluate the proposed TEStore system.

### A. Simulation Environment

We use AirPAK, a CFD software package which allows modeling the thermal behavior of existing data centers and quantifies the temperature changes over time, to simulate the cooling in a data center. We simulate a data center that consists of 1120 blade servers deployed in four rows, with each row hosting seven 40U racks. In particular, the servers are modeled according to some state-of-the-art products such as IBM Systems x350 M2, with 100W of idle power consumption, 300W at 100% CPU utilization, and 5 W in the standby mode. This is consistent with the data center configuration used in several previous studies [3, 17, 18]. More details on the simulated data center are not presented here due to the space limitations, but are available in [3].

There are two non-trivial factors regarding exploiting UPS batteries [8], which may impact the cooling cost savings. First, UPS batteries have conversion loss: the incoming power is first converted from AC to DC (to store in batteries) and is again converted from DC to AC (to power infrastructures). Thus, a portion of the energy stored in batteries is lost whenever there is a recharge/discharge operation, *e.g.*, about 10-15% for lead-acid batteries. Second, batteries become unreliable as they are recharged/discharged with higher depth-of-discharge (DoD), causing faster degradation in their reliability. This dependence between the useful lifetime of a battery and how it is discharged/charged is expressed via battery lifetime charts as in [19, 8]. In particular, the per-cycle cost (*i.e.*, $C_{rc}$ and $C_{dc}$) is estimated in the range of 30 cents to 100 cents for battery-based energy storage systems with lead-acid batteries under use [20]. To this end, we set $C_{rc}$ and $C_{dc}$ to be a specific value, *e.g.*, 45 cents per cycle in this work. Note that other values can also be easily integrated for evaluation. In contrast to battery-based storage, a key advantage of thermal tank storage is that it does not have the DoD related operating cost. Furthermore, thermal tanks are known to have a long lifetime duration, e.g., 20 to 30 years [21]. Thus, the factors of $C_{rc}$ and $C_{dc}$ are often negligible in leveraging thermal tanks to cut the electricity bill, *i.e.*, $C_{rc} = 0$ and $C_{dc} = 0$. In addition, we assume that the energy loss only happens in the recharge operations for thermal tanks, because the loss is mainly introduced in cooling the water into the phase change temperature of $0^oC$. The value of 13% is set for $lf_r$ and $lf_d = 0$ [11].

### B. Real-World Workload and Electricity Price Traces

To build our workloads in the simulated data center, we use a trace file from real-world data centers to simulate the CPU utilization variations. In particular, it includes CPU utilization data of 5.145 servers, which records the average CPU utilization of each server in every 15 minutes from 00:00 on July 14th through 23:45 on July 20th in 2008 [22]. We then map the utilization trace to our data center by consolidating the utilizations of all the servers and distributing them evenly to our 1,120 servers as plotted in Figure 3. Note that we evenly distribute the utilization of the servers because we focus on cooling in this paper. More sophisticated workload distribution methods (*e.g.*, [3]) can be integrated with TEStore. In addition, to simulate the time-varying electricity prices, we apply the

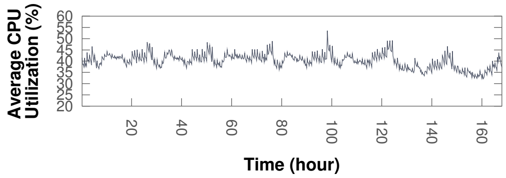



Fig. 3. Data Center CPU Utilization Trace.

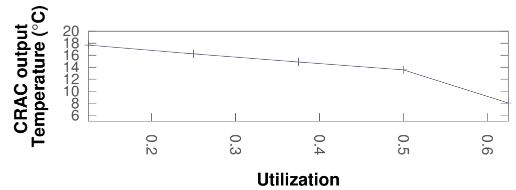



Fig. 4. A commonly used cooling strategy without cooling energy storage: tuning the CRAC output temperature to a level derived from offline profiling.

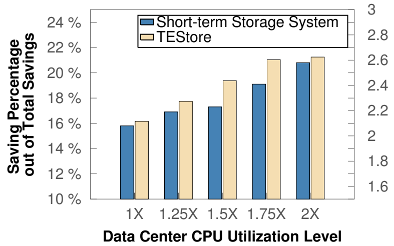



Fig. 5. Cooling cost savings and saving distributions of TEStore under different utilization levels.

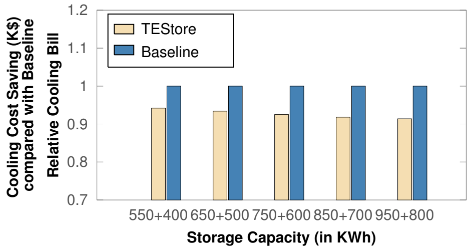



Fig. 6. Cooling cost savings of TEStore under different storage capacities.

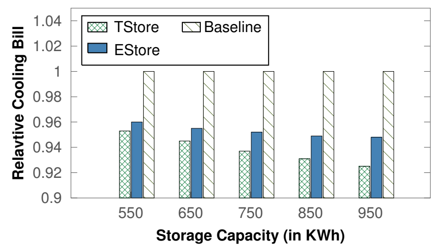



Fig. 7. Cooling cost comparison between different energy storage techniques, under different storage capacities.

hour-ahead price data from the NYISO CAPITAL area in our simulation [4], since it has complete and accurate price data records. It is a 9-month price trace starting from December 1st, 2010 through August 31th, 2011.

### C. Baseline

We assume that the data center has an intelligent cooling strategy which tunes the CRAC output temperature to guarantee the cooling performance. Because this paper focuses on reducing cooling cost using various storage systems, we adopt a commonly assumed intelligent cooling strategy that does not consider cooling energy storage as our baseline. In baseline, we simulate the data center running at different utilization levels. For each utilization level, we manually tune the CRAC output temperature until the data center reaches the desired cooling performance specified in [3], *i.e.*, the average inlet air temperature of the 80 hottest servers reaches $25°C$. We plot the CRAC output temperature in Figure 4. The baseline cooling strategy is to select a fixed CARC output temperature from Figure 4 based on the measured data center utilization.

### D. Effectiveness of TEStore on Cooling Cost Savings

We now perform simulations to demonstrate that TEStore can achieve the desired cooling performance with a lower time-average cooling bill than the baseline cooling strategy.

In this set of experiments, we study the time-average cost savings of TEStore, under the storage capacity of 750 KWh and 600 KWh for thermal tanks and UPS batteries, respectively. We set the parameters $D_{max}$ and $R_{max}$ to 35 KWh. This is consistent with the practice that $R_{max} + D_{max}$ is much smaller than $Y_{max} - Y_{min}$ [8]. As a result, TEStore achieves an overall cooling cost saving of 7.5% with the 9-month electricity power trace from NYISO, compared to the baseline cooling strategy. In particular, out of the total cost saving by TEStore, around 84.2% saving is introduced by exploiting the auxiliary thermal tanks on the long time scale, and the rest is brought in by the short-term storage with respect to UPS batteries and thermal masses. The resultant distribution of the cost savings is due to the coordination strategy deployed by TEStore, as discussed in Section II-A. Note that a higher percentage of cost savings by the short-term storage can be achieved for such workload with higher CPU utilization levels. This is because a heavier workload can lead to a better utilization on the short term storage, due to the limitation on the thermal tank capacity on the long-time scale.

We then stress test TEStore by scaling up the original workload (in units of CPU utilization) by a series of different factors, from 1.25 to 2.0. Note that due to the limitation on the CPU utilization trace we use in this work, we cannot verify the TEStore system under workloads with a factor that is greater than 2.0 since the scaled CPU utilization level will exceed 100%. Figure 5 shows the cost saving distribution of the long-term storage and the short-term storage, out of the total cost savings by TEStore system, under different CPU utilization levels from CPU 1X to CPU 2X with the storage capacity of 750 KWh and 600 KWh for thermal tanks and UPS batteries, respectively. As demonstrated in the figure, with the increase of the workload intensity, a higher percentage of cost savings by the short-term storage is achieved. For example, when the CPU utilization level increases from 1X to 2X, the cost saving by the short-term storage increases from 15.8% to 20.8%. Figure 5 also shows that with the increase of the workload intensity, the cooling bill saving (in dollars) by TEStore increases. This is because that the storage system is more effectively exploited in shifting energy to utilize the time-varying power prices with a heavier workload. One may think that TEStore's gain of a lower cooling bill is at the expense of degraded data center cooling performance. Our results show that TEStore achieves the same or better cooling performance. Specifically, TEStore leads to the same or even lower CRAC output temperatures for effective cooling.

To demonstrate the trade-off between the storage capacity and the cooling bill savings, we perform a series of simulations with respect to the different storage capacities. Figure 6 shows that the cost savings increase from 5.8% to 8.6% as the storage capacity increases from 950 KWh to 1750 KWh. In particular, the tank size is set from 550 KWh to 950KWh while the energy storage increases from 400 KWh to 800 KWh. Note that other configurations can be easily verified in this work.

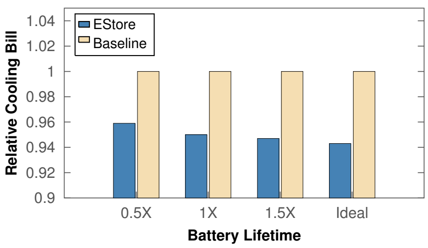



Fig. 8. Cooling bill by EStore under different battery lifetime.

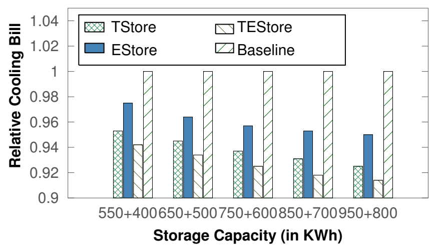



Fig. 9. Cooling cost comparison among different different storage strategies, under a series of different storage capacities.

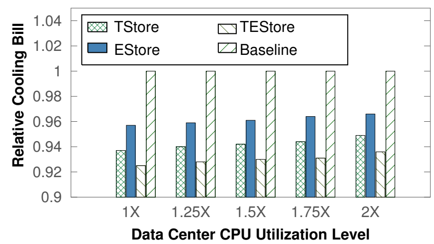



Fig. 10. Cooling cost comparison with respect to different data center CPU utilization levels.

### *E. Comparison of Different Storage Techniques*

In this set of experiments, we study how different auxiliary storage techniques, such as thermal tanks (referred to as TStore) and UPS batteries (referred to as EStore), differentiate in cutting the electricity bill for data center cooling, due to their distinguishing characteristics in discharging/recharging the devices. In particular, we perform precooling strategy along with the auxiliary storage techniques. Note that one can easily compare different auxiliary storage techniques without precooling strategy.

We evaluate the comparison with a series of different storage capacities. Figure 7 shows that TStore can achieve higher cooling cost savings than EStore. In particular, TStore can bring in an up-to-7.4% lower cooling cost for data center operators, compared to a 5.2% cooling cost saving by EStore. This is due to the non-trivial amortized battery lifetime cost in the EStore system. That is, UPS batteries become unreliable and need to be replaced given frequent charge and discharge operations. In this experiment, an average cost of 4.3% is caused by battery replacement, compared to a 1.5% cost by the conversion loss regarding the EStore system.

To demonstrate how the device replacement due to the frequent discharge and recharge operations affects the cooling cost savings, we stress test EStore regarding different UPS battery lifetime. Figure 8 shows the cost savings of running EStore under a series of different battery lifetime from Lifetime 0.5X to Lifetime Ideal with the storage capacity of 750 KWh. Specifically, Lifetime Ideal represents the case that no device replacement is necessary in exploiting UPS battery storage. As shown in the figure, a longer battery lifetime indicates a better cooling cost saving, since a lower per-cycle cost follows a longer battery lifetime. With the ideal battery lifetime, EStore achieves the best cost saving.

### *F. Integration of Thermal and Energy Storage Systems*

In this Section, we conduct experiments to show that a lower data center cooling bill can be achieved with the integration of various storage techniques available in the data center. As shown in Figure 9, data center operators can expect a lower cooling cost by integrating different forms of storage techniques. The reason is that more energy can be buffered by integration, and thus to more wisely utilize the time-varying power prices in reducing the data center cooling cost.

We then stress test the effectiveness of the integration of different storage techniques regarding different data center CPU utilization levels. Figure 10 demonstrates the effectiveness of TEStore with the storage capacity of 750 KWh and 600 KWh for thermal tanks and UPS batteries, respectively. The figure shows that for all different CPU utilization levels, TEStore guarantees the lowest cooling bill for data center operators.

## V. Related Work

Minimizing the electricity cost of data centers has recently received much attention, such as [23, 24, 25, 26, 27, 28, 29, 30]. In contrast to these studies that mainly focus on dynamic workload distribution among geographically distributed data centers, this work introduces a smart cooling strategy by exploiting the thermal and energy storage techniques available in a data center in cutting the cooling bill for data center operators. Data center power and cooling management has also been research extensively (*e.g.*, [3, 2, 5, 31, 8, 9, 18, 32, 33, 34, 35, 36]). However, those studies focus mainly on reducing the power consumption or server temperatures instead of the electricity cost.

Auxiliary thermal tanks are used to buffer energy in order to cool a data center in face of CRAC system power outage [6]. Thermal tanks have also been used to shift energy consumption from day time to night time in buildings (instead of data centers) for cost reduction [11]. A few recent projects have tapped into battery-based storage techniques to cut server power bill [8, 9]. However, to the best of our knowledge, no study has integrated different storage techniques in cutting the electricity bill for data center cooling. In summary, TEStore models the thermal fluid dynamics in a data center, and presents the first study that integrates different storage techniques on different time scales to cut the electricity bill for data center cooling.

## VI. Conclusion

In this paper, we have presented TEStore that leverages thermal and energy storage techniques on different time scales to cut the electricity bill for cooling, without causing servers in a data center to overheat. A dynamic control algorithm is proposed to optimize the time average cooling bill by exploiting different auxiliary storage devices. The impacts of TEStore on server temperature is modeled based on Computational Fluid Dynamics (CFD) to consider the realistic thermal dynamics in a data center. Our trace-driven simulation results show that TEStore achieves the desired cooling performance with a much lower electricity bill than the current practice.

## VII. Acknowledgments

This work was supported, in part, by NSF grants CNS-0720663, CNS-0915959, CNS-1143607 (NSF CAREER Award), and by ONR grant N00014-09-1-0750.

## References


[1] S. Pelley, D. Meisner, T. F. Wenisch, and J. W. VanGilder, "Understanding and abstracting total data center power," in *Workshop on Energy Ecient Design (WEED)*, 2009.

[2] J. Moore, J. Chase, P. Ranganathan, and R. Sharma, "Making scheduling "cool": temperature-aware workload placement in data centers," in *Proceedings of the annual conference on USENIX Annual Technical Conference*, 2005.

[3] F. Ahmad and T. N. Vijaykumar, "Joint optimization of idle and cooling power in data centers while maintaining response time," in *Proceeding of ASPLOS*, 2010.

[4] "NYISO," http://www.nyiso.com/, 1999.

[5] T. D. Boucher, D. M. Auslander, C. E. Bash, and C. C. F. C. D. Patel, "Viability of dynamic cooling control in a data center environment," *Journal of electronic packaging*, vol. 128, p. 137, 2006.

[6] "Thermal storage system provides emergency data center cooling," Intel, Technical report, 2007.

[7] R. L. Mitchell, "Ice balls help data center go green," http://www.computerworld.com, 2010.

[8] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *Proceedings of SIGMETRICS*, 2011.

[9] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and limitations of tapping into stored energy for datacenters," in *Proceedings of ISCA*, 2011.

[10] K. Roth, R. Zogg, and J. Brodrick, "cool thermal storage," *ASHRAE Journal*, vol. 48, p. 94, 2006.

[11] K. W. Roth, D. Westphalen, J. Dieckmann, S. D. Hamilton, and W. Goetzler, "Energy consumption characteristics of commercial building HVAC systems volume III: Energy savings potential," 2002.

[12] F. Nogales, J. Contreras, and A. J. Conejo, "Forecasting next-day electricity prices by time series models," *Power Systems, IEEE Transactions on*, vol. 17, no. 2, pp. 342 –348, may 2002.

[13] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," in *Foundations and Trends in Networking*, 2006, pp. 1–149.

[14] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.

[15] C. D. Patel, R. Sharma, C. E. Bash, and A. Beitelmal, "Thermal considerations in cooling large scale high compute density data centers," in *ITherm*, 2002.

[16] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2007, vol. 1&2.

[17] J. Moore, J.S.Chase, and P.Ranganathan, "Weatherman: Automated, online and predictive thermal mapping and management for data centers," in *Proceeding of ICAC*, 2006.

[18] R. K. Sharma, C. E. Bash, C. D. Patel, and R. J. Friedrich, "Balance of power: Dynamic thermal management for internet data centers," *IEEE Internet Computing*, vol. 9, pp. 42–49, January 2005.

[19] D. Linden and T. B. Reddy, *Handbook of Batteries*. McGraw Hill Handbooks, 2002.

[20] H. Ibrahim, A. Ilinca, and J. Perron, "Energy storage systems characteristics and comparisons," *Renewable and Sustainable Energy Reviews*, vol. 12, pp. 1221–1250, 2008.

[21] "Ice storage systems," http://www.calmac.com.

[22] X. Wang, M. Chen, C. Lefurgy, and T. W. Keller, "SHIP: Scalable hierarchical power control for large-scale data centers," in *Proceedings of PACT*, 2009.

[23] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, "Cutting the electric bill for internet-scale systems," in *Proceedings of SIGCOMM*, 2009.

[24] L. Rao, X. Liu, L. Xie, and W. Liu, "Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment," in *Proceedings of INFOCOM*, 2010.

[25] J. Li, Z. Li, K. Ren, and X. Liu, "Towards optimal electric demand management for internet data centers," *IEEE Transactions on Smart Grid*, 2010.

[26] K. Le, O. Bilgir, R. Bianchini, M. Martonosi, and T. D. Nguyen, "Managing the cost, energy consumption, and carbon footprint of internet services," in *Proceedings of SIGMETRICS*, 2010.

[27] K. Le, R. Bianchini, , T. Nguyen, O. Bilgir, and M. Martonosi, "Capping the brown energy consumption of internet services at low cost," in *Proceedings of IGCC*, 2010.

[28] K. Le, R. Bianchini, M. Martonosi, and T. D. Nguyen, "Cost- and energy-aware load distribution across data centers," in *Proceedings of HotPower*, 2009.

[29] Y. Zhang, Y. Wang, and X. Wang, "Electricity bill capping for cloud-scale data centers that impact the power markets," in *Proceedings of ICPP*, 2012.

[30] Y. Zhang, Y. Wang, and Wang, "Greenware: Greening cloud-scale data centers to maximize the use of renewable energy," in *ACM/IFIP/USENIX 12th International Middleware Conference*, 2011.

[31] Z. Wang, C. Bash, N. Tolia, M. Marwah, X. Zhu, and P. Ranganathan, "Optimal fan speed control for thermal management of servers," in *ASME*, 2009.

[32] C. Lefurgy, X. Wang, and M. Ware, "Server-level power control," in *Proceedings of ICAC*, 2007.

[33] X. Wang and M. Chen, "Cluster-level feedback power control for performance optimization," in *Proceedings of HPCA*, 2008.

[34] Y. Wang, X. Wang, M. Chen, and X. Zhu, "Power-efficient response time guarantees for virtualized enterprise servers," in *Proceedings of RTSS*, 2008.

[35] Y. Wang, R. Deaver, and X. Wang, "Virtual batching: Request batching for energ conservation in virtualized servers," in *Proceedings of the 18th IEEE International Workshop on Quality of Service*, 2010.

[36] X. Wang, Y. Yao, X. Wang, K. Lu, and Q. Cao, "Carpo: Correlation-aware power optimization in data center networks," in *Proceedings of INFOCOM*, 2012.