2013 IEEE Sixth International Conference on Cloud Computing

Optimization of Electricity and Server Maintenance Costs in Hybrid Cooling Data Centers

Shaoming Chen, Yue Hu, and Lu Peng
Division of Electrical & Computer Engineering, School of Electrical Engineering and Computer Science
Louisiana State University, LA, US
{schen26, yhu14, lpeng}@lsu.edu

***Abstract*—The electricity cost of data centers dominated by server power and cooling power is growing rapidly. To tackle this problem, inlet air with moderate temperature and server consolidation are widely adopted. However, the benefit of these two methods is limited due to conventional air cooling systems ineffectiveness caused by re-circulation and low heat capacity. To address this problem, hybrid air and liquid cooling, as a practical and inexpensive approach, has been introduced. In this paper, we quantitatively analyze the impact of server consolidation and temperature of cooling water on the total electricity and server maintenance costs in hybrid cooling data centers. To minimize the total costs, we proposed to maintain sweet temperature and ASTT (available sleeping time threshold) by which a joint cost optimization can be satisfied. By using real world traces, the potential savings of sweet temperature and ASTT are estimated to be average 18% of the total cost while 99% requests are satisfied compared to a strategy which only reduces electricity cost.**

Keywords-Data centers, electricity cost, server maintenance cost, joint optimization.

I. INTRODUCTION

The total cost of ownership (TCO) in data centers consists of onetime capital costs incurring only at the beginning or upgrade stage of data centers and monthly recurring operational costs including electricity cost, maintenance cost and salaries [4]. According to recent reports, the TCO is dominated by the operational costs [3] [4], among which salaries are largely not a technical but an economic factor. Therefore, we focus on optimization of electricity and maintenance costs in this work.

The growth of the cost of electricity consisting of server power and cooling power outpaces expectations. In 2011, U.S. data centers spent about $7.4 billion in electric power among which server power and cooling power contribute significantly to the total [26]. Several studies try to throttle this increase, though few of them consider the cost of server maintenance.

Prior works employ two methods to reduce energy cost: increasing server consolidation and increasing inlet air temperature. Server consolidation is a powerful tool which has been widely adopted to gain high energy efficiency of server, which results from keeping active servers in high utilization by turning off overprovisioned servers [3][28]. As an alternative approach to save server power, Dynamic Voltage and Frequency Scaling (DVFS) is also used [13]. However, the benefit of DVFS is shrinking because the leakage power is increasing and the voltage of processors is getting very close to its threshold voltage [22]. In addition, DVFS only affects CPU power which amounts to 30% of server power [26]. Server consolidation remains as an effective and practical method to save server power.

To reduce cooling power, increasing inlet air temperature is a common method since increasing inlet air temperature by just one degree can reduce cooling energy consumption by 2 to 5 percent [7]. However, the room of inlet air temperature can be raised is very limited due to the constraint of server temperature below the critical temperature. To keep the constraint with a low cost, there are several prior works advocating thermal-aware workloads placement which distributes workloads according to the thermal map of data centers [3][24]. Unfortunately, these methods cannot maintain energy efficiency of traditional air cooling by keeping high inlet air temperature when data centers are in high utilization [28]. Therefore, a novel cooling system is demanded.

As a practical and inexpensive solution of liquid cooling [14], a hybrid cooling system which combines air and liquid cooling has been proposed and deployed in data centers such as Aquasar, the first hot water cooled supercomputer prototype [36]. The hybrid cooling system uses water to cool down high power density components such as processors and memory which dominate total heat dissipated in servers, while other auxiliary components show low power density are still cooled down by air cooling. In this way, hybrid cooling can remove a mass of heat from datacenter with less power than conventional air cooling.

In addition to the electricity cost coming from servers and cooling systems, hardware maintenance cost is also considerable. According to a typical new multi-megawatt datacenter in the United States, the cost of server repair and maintenance is approximately 50% of the costs of server power and cooling power [4]. Based on the empirical data of a HPC datacenter [28], disks are the most frequently replaced components. Resulting from server consolidation, frequent turning off servers or transition between active state and sleeping state incurs the cost of disk maintenance due to the limited start-stop cycles for disks [8]. Additionally, higher inlet water temperature increases the cost of CPU and memory maintenance, since every 10°C increase over 21°C decreases the lifetime reliability of electronics by 50% [25]. Therefore, rather than restricting chip temperature below a certain threshold, we can balance the saving of the electricity costs and the increase of the costs of hardware maintenance through manipulating inlet water temperature and smoothing the variation of the number of active servers.

978-0-7695-5028-2/13 $26.00
DOI 10.1109/CLOUD.2013.104

The contributions of our work are shown in the following.

• We set up analytical models for server power, cooling power and hardware maintenance model in hybrid cooling data center for the quantitative evaluation. To our best knowledge, we first build a comprehensive framework which covers the evaluation of these costs. This framework provides foundations to optimize the total cost in hybrid cooling data centers.

• We propose a tradeoff between electricity cost and maintenance cost. In this work, we show that the typical optimizations (high inlet water temperature and aggressive server consolidation) which reduce only electricity cost could hurt the maintenance costs.

• To minimize the electricity cost and the maintenance cost, we develop a joint optimization scheme based on dynamic optimal water inlet temperature and server consolidation. Our simulation results show that the method can gain considerable savings of these costs.

The rest of our paper is organized as follows: we describe the structure of hybrid cooling in section 3. In section 4, we build models related to electricity cost and the cost of server maintenance. We propose cost optimization methods in section 5. In section 6, we setup a datacenter model with server performance model and response analysis. In section 7, we analyze the result of these two methods and show their potential savings. Finally, we conclude the paper in section 8.

II. RELATED WORK

Prior works of the cost optimization of data centers fall into two categories: the optimization of electricity cost [30] and the optimization of hardware maintenance cost. To decrease energy consumption of datacenters, many studies were addressed from server level [22][23] , rack level [29] and data center level [8] [11][20][33] . These studies focused on increasing server energy efficiency and reducing server idle power. On the other hand, Moore et al. [24] introduced thermal-aware workloads placement to reduce cooling power in traditional air cooling data centers. On the contrary, other researchers employed advanced infrastructures of cooling systems to solve energy inefficiency of traditional air cooling [4] [16] [31]. However, all these works just aimed at the reduction of either cooling power or server power.

To capture an abroad scope of energy savings, several architects proposed approaches [3] [15] [26] [28] for optimization of cooling power and server power. For an example, Pelley et al. [26] set up a comprehensive framework of total data center power in data centers to optimize server power and cooling power. F.Ahma et al. [3] proposed a joint optimization of server power and cooling power with guaranteeing response time. However, all of these works did not consider the increment of the costs of hardware maintenance.

On the other hand, several papers discussed the issue related to hardware maintenance in data centers [18] [32] [34]. Schroeder et al. [32] analyzed disk replacement rate based on the empirical data, which inspired researchers to study the reliability of hardware in servers.

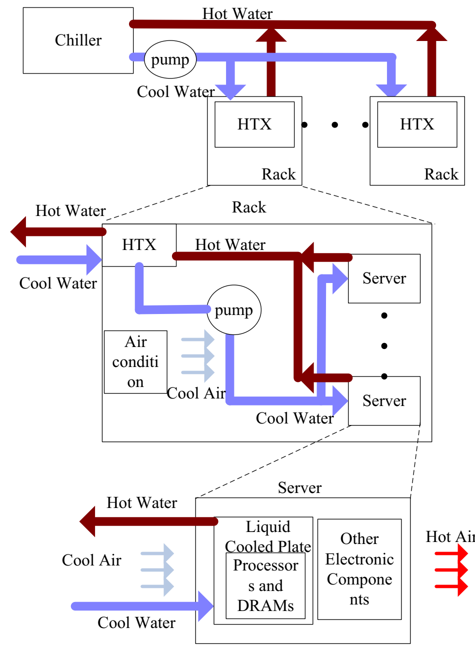

Figure 1: The Structure of Hybrid Cooling

Unlike previous works which focused on the optimization of electricity cost or hardware maintenance in data centers, our approach connects them together. Though Y. Chen et al. [8] minimized the cost of energy and disk maintenance by combing DVFS and server consolidation, the author did not discuss cooling cost and maintenance cost of other components such as processors and memory in servers. Thus, the comprehensive framework is initially addressed in our paper.

III. HYBRID COOLING

The structure of hybrid cooling in modern data centers is shown in Figure 1. The closed liquid loop between the chiller and racks is designed to remove heat dissipation from the racks. The cool water absorbs heat dissipation from the racks and returns back to the chiller with heat. In the closed liquid loop of a rack, the water cooled in the intermediate Heat Exchanger (HTX) is pumped into servers. In a server, the water flowing through a liquid cooled plate takes away power dissipated by processors and memory. Other auxiliary components such as disks, power supply, and chipsets on motherboard are still cooled by the air condition as traditional data centers since these components dissipate less power and, more importantly, exhibit lower power density compared with processors and DRAMs.

IV. COST MODELS

To optimize the electricity cost and the hardware maintenance cost, we setup the cost models which quantitatively estimate the impact of server consolidation and inlet water temperature on the costs when hybrid cooling is used.

A. Electricity costs

The power of a typical data center includes server power, cooling power and power distribution loss. For power distribution loss, PDU and UPS draw 10% of load power [28]. In the following context, the models related to server power and cooling power is addressed.

$$P_{total} = P_{servers} + P_{cooling} + P_{power\ distribution\ loss} \quad (1)$$

For server power, $P_{servers}$ consists of the sum of all active server power and the sum of sleeping server power. The power for all servers is written as:

$$P_{servers} = \sum_{i=1}^{NAS} P_{Servers}(i) + \sum_{j=1}^{NIS} P_{sleep}(j) \quad (2)$$

Here, NAS and NIS denotes the number of active servers and sleeping servers consuming 6 Watts per server [3]. For an active server, the total power consists of the power of processors, the power of memory and the power of other components. The equation is listed as follows:

$$P_{Server} = \sum_{i=1}^{NS} P_{Processor}(i) + \sum_{j=1}^{NM} P_{Memory}(j) + P_{Other} \quad (3)$$

Where NS and NM are denoted as the number of sockets and the number of DIMMs in a server. To simplify the equation, we assume that all servers in data centers have the same number of sockets and the number of DIMMs.

For the power model of components in a server ($P_{Processor}$, P_{Memory} and P_{Other}), we adopt the linear power model, which is shown as follows:

$$P = (P_{TDP} - P_{idle}) * U + P_{idle} \quad (4)$$

where P_{TDP} and P_{idle} indicate the maximum power and idle power of components while U denotes server utilization.

The configuration of power model in a server is shown in Table 1. For processors, its idle power amounts to 10% of the TDP [9], while 4 HDD hard disks are assumed to be installed in the server to fit memory intensive applications. The specification is derived from a typical server [9].

According to the hybrid cooling structure, the cooling power can be divided into two parts: the liquid power and air cooling power:

$$P_{cooling} = P_{liquid\_cooling} + P_{air\_cooling} \quad (5)$$

To estimate cooling power, $E = Q/COP$ is employed where E denotes the energy to remove the heat dissipation Q from data centers and COP (Coefficient of Performance) is defined as a metric to evaluate the efficiency of cooling system [24]. According to prior studies [3][24], COP_{air} (coefficient of performance) can be derived in the following equation: $COP_{air} = (0.0068 \times T^2 + 0.0008 \times T + 0.458)$ where T is the inlet air temperature.

The power of liquid cooling consists of the power of chiller and the pump power [16]. The chiller efficiency for a typical chilled water system is also written as: $COP_{liquid} = E/Q$ [5]. COP_{cooled} is written in terms of inlet water temperature: $COP_{liquid} = T * 0.18 - 0.4836$ based on the specification of water-cooled screw compressor chiller [2]. The water pump power is calculated by this equation [16]:

$$P_{pump} = N \times \frac{V_w \times \Delta P_w}{\eta_{pump}} \quad (6)$$

Table 1 Configurations

Server Configurations			
Part	**#**	**TDP(w)**	**Idle power(w)**
Processor	2	150W	15W
Memory	8	10W	5W
Others	-	124W	73.6W
Hybrid Cooling Configurations			
Parameter		**Value**	
T_{inlet_water} (°C)		25	
T_{inlet_air} (°C)		25	
V_w (GPM)		1	
η_{pump}		70%	
ΔP_w(psi)		4.2	
Thermal Reliability Configurations			
θ_{CP} (°C/W)		0.3	
θ_{MP} (°C/W)		4.75	
θ_p (°C/W)		0.03	
Maintenance Cost Configurations			
Start-stop cycles for disks		40000	
CPU maintenance price ($)		300	
Disk maintenance price ($)		200	
Memory maintenance price ($)		150	

where N is the number of servers and V_w is the water volume flow rate. ΔP_w denotes the water side pressure drop based on the flow resistance. Finally, η_{pump} indicates the pump efficiency.

Overall, the cooling power of the data center is calculated as follows:

$$P_{cooling} = \frac{Q_{liquid\ cooled}}{COP_{liquid}(T_{inlet\_water}) * t} + \frac{Q_{air\ cooled}}{COP_{air}(T_{inlet\_air}) * t} + P_{pump} \quad (7)$$

where t is a time interval during which server components dissipate the heat $Q_{liquid\ cooled}$ and $Q_{air\ cooled}$. The heat $Q_{liquid\ cooled}$ removed by liquid cooling, while the heat $Q_{air\ cooled}$ consisting of the heat dissipated other components in active servers and inactive servers. Shown in the Table 1 is the configuration of hybrid cooling derived from [16]. the pump power of a server is 0.6 watt and is negligible compared to the chilling power.

Overall, the electricity cost of the data center is written as:

$$EC = K_{\$} P_{total} \quad (8)$$

Here, P_{total} and $K_{\$}$ respectively denote the power consumed the data center and commercial KWH Billing Rate which comes to 9 cents/KWH.

B. The costs of hardware maintenance

As we have addressed in the introduction, arising temperature and frequent consolidation could accelerate server aging processes. Due to high power density of DRAM and CPU, we focus on their maintenance cost. In addition, even though hard disks have a low power density, their limited number of lifetime start-stop cycles is heavily impacted by frequent server consolidations. Therefore, we also take the cost of disks maintenance into account.

1) Thermal model

To investigate the costs of processor and memory maintenance, we have setup up thermal models.

The CPU temperature T_C is calculated as follows [17]:

$$T_C = T_{inlet} + (\theta_{CP} + \theta_p) * Q_C \tag{9}$$

Here, T_{inlet} is the inlet water temperature and Q_C is the power dissipated by the CPU. Thermal resistance of the processor package and TIM (Thermal Interface Material) layer is denoted by θ_{CP}. The value of θ_{CP} is derived from [16]. The thermal resistance of cold plate which varies with water flow is denoted by θ_p, according to the specification of Lytron CP20 cold plates [16]. For the reliability issue of CPU, there is a threshold temperature for processor chips as 90°C [25].

For DRAM, the temperature T_M is given as follows:

$$T_M = T_{inlet} + (\theta_{MP} + \theta_P) * Q_{MP} \tag{10}$$

where Q_{MP} is the power dissipated by memory. Thermal resistance of chip package of DRAM is denoted by θ_{MP} derived from [1]. There is a threshold temperature for DRAM as 85°C [19]. The characteristics of thermal package of DRAM, CPU and cold plates are listed in the Table 1.

2) Thermal Reliability model of electronic devices

After we have obtained the chip temperature of electronic devices, we can predict the lifetime of electronic devices based on the thermal reliability model of electronic devices. The main factors to determine the lifetime of electronic devices are power and chip temperature [10]. For memory, the lifetime prediction model [18] is adopted. MTTF (mean time to failure) is widely used to represents the predicted lifetime of electronic components for processors: $\mathrm{MTTF} = 1/\lambda$. For the prediction of the lifetime of processor and memory, λ is the number of failures per million hours and calculated according to Military Handbook MIL-HDBK-217F [38].

$$\lambda = (C_1\pi_T + C_2\pi_E)\pi_Q\pi_L \tag{11}$$

$$\pi_T = 0.1\exp\left(\frac{-E_a}{8.617 \times 10^{-5}}\left(\frac{1}{T_p + 273} - \frac{1}{298}\right)\right) \tag{12}$$

Here, E_a is the effective activation energy (Ev) and T_P is the temperature of electronic devices. The parameters (C_1,C_2,π_E,π_L π_Q) are derived from [38]. We have scaled the lifetime of CPU and memory according to recent studies [18][34]. The lifetime of CPU is expected to be 7 years when chip temperature is 70 ºC [34], while the expected lifetime of 2GB DRAM is 5 years when its temperature is 65 ºC [18].

C. The costs of hardware maintenance

After the thermal reliability of electronic devices has been introduced, we evaluate the costs of processors and DRAM maintenance based on their thermal reliability is given as follows:

$\mathrm{RC} =$ the cost of hardware maintenances $/\mathrm{MTTF}$.

For a time interval, MTTF is calculated based on their thermal reliability model and current chip temperature. The cost of a CPU, a disk and a memory maintenance are \$300, \$200 and \$150 respectively as shown in Table 1, according to the maintenance ranging from \$300 to \$150 [4]. Based on the thermal reliability model, the cost of CPU and memory maintenance in an active server is specified as follows:

$$RC_{Server} = \sum_{i=1}^{NS} RC_{Processor}(i) + \sum_{j=1}^{DM} RC_{Memory}(j) \tag{13}$$

Here, the costs of DRAM and CPU maintenance are increased by higher inlet water temperature, though the auxiliary components are still cooled down by air cooling. Their little heat dissipation, much lower power density and fixed inlet air temperature result in their little cooling power and their stable maintenance cost. Additionally, the lifetime of hard disks is heavily impacted by server consolidations due to hard disk limited number of lifetime start-stop cycles [12], while the impact of utilization and temperature is still unclear [27]. On the other hand, switching on/off servers incurs relatively little maintenance cost of other components such as processors and memory compared with that of hard disks. The cost of disk maintenance is computed by the following equation:

$$RC_{Disk} = \frac{Price}{start - stop\ cycles} \tag{14}$$

As we know, the number of lifetime start-stop cycles for hard disks is 40000 [8].

Overall, the cost of hardware maintenance of data center is listed as follows:

$$RC = \sum_{n=1}^{ND} RC_{Disk}\,[NAS(t-1) - NAS(t)]^+ + \sum_{k=1}^{NAS} R_{Server}(k) \tag{15}$$

$$[A]^+ = A\ if\ A > 0\ or\ [A]^+ = 0\ if\ A \le 0$$

where ND and $\mathrm{NAS(t)}$ respectively denotes the number of disks in a server and the number of active servers in the data center at the time t. $[\mathrm{NAS(t)} - \mathrm{NAS(t-1)}]^+$ represents the number of servers which have been turned off.

Consequently, we have set up models for electricity cost and the cost of hardware maintenance to evaluate our approach which optimizes the total cost. The models have been validated with the costs of our campus datacenters.

V. COST OPTIMIZATION IN DATA CENTERS

We formulate the total cost in equation (16) based on the equations (8) (15) with constraints. Since we only focus on the operational cost of data centers, we pick up a typical specification for our heuristic data center shown in Table 1.There are two important decision variables T_{inlet_water} and NAS, while other variables are determined by available servers, server performance and characteristics of traces, which are also treated as parameters. For example, NS denotes the total number of servers, while MINS denotes the minimal required number of active servers which is determined by traces. Our objective is to minimize the total cost with the constraints:

$$\min\{\mathrm{TC} = \sum_{n=1}^{ND} RC_{Disk} * [\mathrm{NAS(t-1)} - \mathrm{NAS(t)}]^+ + \sum_{i=1}^{\mathrm{NAS}} \mathrm{RC_{Server}(i)} + \mathrm{K_\$} * (\mathrm{P_{IT}} + \mathrm{P_{cooling}})\} \tag{16}$$

Subject to

$T_C \le 90$ °C and $T_M \le 85$ °C $\mathrm{MINS} \le \mathrm{NAS} \le \mathrm{NS}$

The space of feasible solutions of this discrete optimization is too large, resulting in that exhaustively searching the

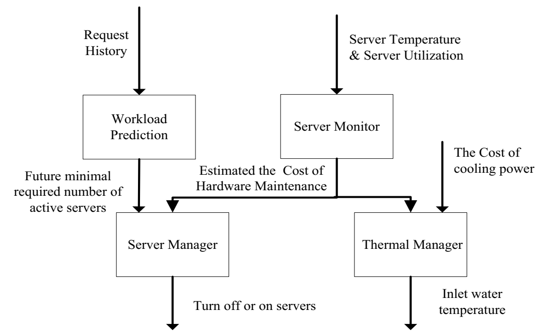

Figure 2: The overview of costs optimization system

global optimal solution is impossible. To optimize the total cost of electricity and hardware maintenance, we proposed to trace local optimal solution by dynamically manipulating T_{inlet_water} and NAS corresponding to the fluctuation of workloads.

A. The Overview of Cost Optimization System

For the manipulation of T_{inlet} and NAS, we proposed a structure shown in Figure 2. In this structure, there are four modules, Workload Prediction, Server Monitor, Server Manager and Temperature Manger, working together to reduce the total cost. The workload prediction collects request history and predicts future request trend based the history. The module also can predict the future minimal required number of active servers. The server monitor collects the temperature and utilization information of servers and estimates the cost of hardware maintenance. Acquiring the average server utilization from the server monitor, the temperature manager adjusts inlet water temperature. The Server manager dynamic allocates servers according to the predicted future minimal required number of active servers.

B. The Impact of Inlet Water Temperature

To investigate the impact of the inlet water temperature on the total cost, we divide the total cost into two parts: the cost of cooling power and CPU and memory maintenance which are affected by the inlet water temperature, and the other costs which are unaffected denoted by C.

$$TC = K_{\$} * P_{cooling} + \sum_{i=1}^{NAS} RC_{Server}(i) + C \qquad (17)$$

As the inlet water temperature increases, $P_{cooling}$ decreases based on the function of COP, while RC_{Server} increases at the same time according to equations (9)-(13). There should be an optimal temperature to balance the cost of cooling power and the costs of CPU and memory maintenance. The optimal temperature (or sweet temperature) is adjusted according to workloads since the two costs also vary with the change of workloads.

C. The Impact of Server Consolidation

The other substantial variable NAS is facilitated by server consolidation which lively migrate jobs cross servers, with the upper bound of available servers and the low bound of service level agreement. Under these constraints, its cost and benefit are investigated in the following.

```
//NAS : the Current Number of Active Servers
If NAS < FMRNAS [T]
        NAS = FMRNAS [T]
Else
        // Turn off servers
        If NAS > Max(FMRNAS [T,T+ ASTT])
        // Turn off(NAS - Max(FMRNAS [T,T+ ASTT])) servers
                NAS = Max(FMRNAS [T,T+ ASTT])
        Else
                // Do nothing
                pass
```

Figure 3: The algorithm based on ASTT

1) The cost of Server Consolidation

It is well known that server consolidation could save the electricity cost. Unfortunately, it increases the cost of disk maintenance, according to equation (15). Furthermore, the transition between the active state and the sleeping state, servers wastes energy. We formulate the cost for server consolidation denoted by C_{cs}. The cost C_{cs} per a server is calculated as follows:

$$C_{cs} = \sum_{j=1}^{ND} RC_{Disk} + P_{max} * T_T * K_{\$} \qquad (18)$$

where T_T is the time of the two transitions including two job migrations (20 seconds for one[8]) and two transitions between the active state and sleeping state (5 seconds for ACPI S3 state [21]). Therefore, T_T is estimated to be 50 seconds. P_{max} and $K_{\$}$ respectively represent the maximum power for a server and denotes commercial KWH Billing Rate.

2) The benefit of Server Consolidation

The reward of server consolidation depends on the length of server sleeping time for once turning off. In other word, the benefit is determined by the length of the period of turning off servers without violation of user level agreement. The length of this period is referred as available sleeping time (AST) which indicates the maximal server sleeping time. Thus, the benefit of turning off N servers is denoted by $B_{sleeping} \times AST \times N$. Here, $B_{sleeping}$ denotes the benefit of turning off a server for a minute.

To optimize server consolidation, we define available sleeping time threshold (ASTT) as follow:

$$C_{cs} = \sum_{j=1}^{ND} RC_{Disk} + P_{max} * T_T * K_{\$} \qquad (19)$$

When the available sleep time of servers is longer than ASTT, the servers should be turned off. Otherwise, the server should keep running. We design an algorithm shown in Figure 3 based on the concept. Generally, the algorithm conservatively turns off servers to mitigate the cost of server consolidation.

In this algorithm, the decision of turning off servers requires the knowledge of Future Minimal Required Number of Active Servers (FMRNAS) which is bound by the constraint of service level agreement (SLA).The performance of this algorithm depends on how accurately FMRNAS is predicted. Therefore, we will introduce two different predictions combined with the algorithm in the following sections.

3) ASTT-P Available sleeping time threshold based on a perfect prediction

Firstly, we assume that we have a perfect predictor which indicates FMRNAS accurately. Given this knowledge, ASTT-P is designed to minimize the total cost by selecting an available sleeping threshold without the impact of inaccurate predictions. The exact value of optimal available sleeping threshold is impossibly obtained by solving equation (18) since $B_{sleeping}$ is slightly affected by other factors such as inlet water temperature.

4) ASTT-AR: Available Sleeping time threshold based on the autoregressive model (AR model).

The adopted prediction based on the autoregressive model [35] which is widely used for pattern prediction is listed in the following equation to estimate FMRNAS:

$$\dot{SN}(T) = (K+1)(C + \sum_{i=1}^{a} A_i * SN(T-i)) \quad i = 1 \cdots a \qquad (20)$$

where $\dot{SN}(T)$ denotes predicted FMRNAS at time T while $SN(T-i)$ denotes PMRNAS at time $(T-i)$. C and A_i are tuned to reduce overprovision servers and guarantee the response time in offline. K is updated according to the percentage of requests whose response time is satisfied. When the percentage is below the requirement, K increases to reserve more servers to handle spike requests. Otherwise, K is decreased. The goal in this paper is to satisfy more than 99% requests. In our paper, we focus on the benefit of ASTT by utilizing the mature pattern prediction, though it might be replaced by advanced tools.

In the following section, the model of a datacenter is built up to quantitatively evaluate the benefit of sweet temperature and ASTT.

VI. EXPERIMENT SETUP

A. Datacenter

Recalling the models related to the costs of electricity and hardware maintenance, we combined them with server performance model and real traces to simulate our prototype data center which consists of 1024 servers cooled by hybrid cooling.

1) Server performance model & response time analysis

We assume a server in our datacenter provides 2.6 Gbytes/sec service rate and the mean of response time should be bound by 6 ms for SLA [8]. To calculate the FMRNAS at a time interval, we use GI/G/m model [6] to determine how many servers can satisfy a demand based on the following equation:

$$\widetilde{W} = \frac{1}{\mu} + \frac{P_m}{\mu(1-\rho)} * \left(\frac{C_A^2 + C_B^2}{2m}\right) \qquad (21)$$

$$P_m = \rho^{\frac{m+1}{2}} \text{ if } \rho \le 0.7 \quad P_m = \frac{\rho^m + \rho}{2} \text{ if } \rho > 0.7$$

where $\widetilde{W}$ is the mean response time. $1/\mu$ is the mean service time of a server. $\rho = \frac{\lambda\varphi}{mf}$ is the average utilization of servers. $\lambda, \varphi,$. C_A and C_B are derived from trace characteristics[3]. We use this performance server and response time model to acquire the minimal required number of active servers at every time slot. For a time interval, we choose 5 minutes as the minus unit [3].

B. Traces

We use five traces downloaded from the Internet traffic Archive [37]: Clarknet-HTTP, NASA-HTTP, Saskatchewan-HTTP, UC Berkeley IP and WorldCup. The lengths of them range from 14 days to 30 days and all of trace files cover several peak requests. We have scaled the traces to meet our datacenter performance.

VII. RESULTS

A. The impact of the optimization based on Sweet Temperature

As illustrated in equation (17), when the server power is fixed, the total cost is only related to cooling and hardware maintenance. Figure 4 illustrates the impact of the inlet water temperature changing from 15°C to 35°C on the cooling cost and the cost of hardware maintenance of a datacenter with 30% utilization. These costs are normalized against the total costs when inlet water temperature is 15°C. Increasing inlet water temperature reduces cooling power especially when the temperature is below 25°C. However, high inlet water temperature also increases the cost of hardware maintenance of CPU and memory. Observed from Figure 4, we can find an optimal inlet water temperature (25 °C in this case) which minimizes the total cost when utilization is fixed at 30%. In the following context, we will refer the sweet temperature to the optimal inlet water temperature. This observation justifies that high inlet water temperature

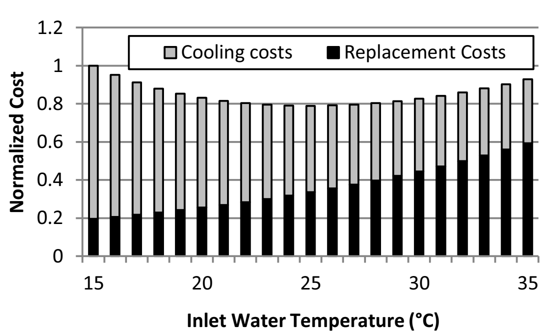

Figure 4: The impact of inlet water temperature on the costs of cooling power and hardware maintenance

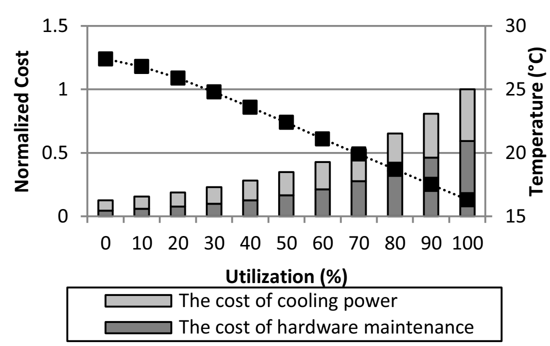

Figure 5: The variation of Sweet Temperature and these costs corresponding to the utilization of the data center

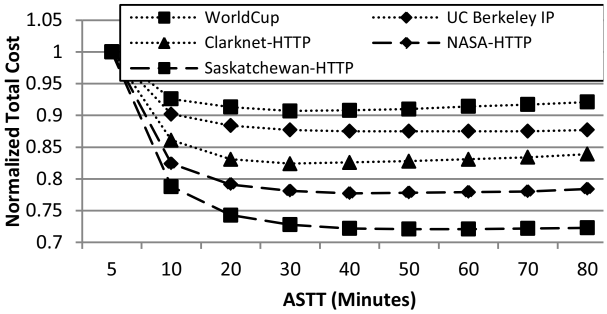

Figure 6: The Normalized total cost reduced by ASTT-P when ASTT from 5 to 80 minutes in five traces

is reasonable in datacenters when the current average server utilization is low (below 30%). Otherwise, high inlet water temperature could hurt the cost of hardware maintenance during the high utilization.

Figure 5 shows the cooling and hardware maintenance costs of our datacenter when its utilization varies from 0% to 100%. The right vertical axis of the figure illustrates sweet temperatures for different utilizations. In the figure, the total costs for all utilizations are the lowest for the datacenter cooled by water at corresponding sweet temperatures. When the utilization of the datacenter is low, warm inlet water temperature offers more benefit since the cost of cooling power is larger than the cost of hardware maintenance (e.g. in our simulation result, the cost of cooling power is 1.65 times of the cost of hardware maintenance when the utilization is 10%). On the other hand, as the datacenter utilization increases, we must keep a cold chilling water to cool down the heating hardware and slow the growth of hardware maintenance especially when their temperatures are close to the critical temperatures. Consequently, to minimize the total costs, inlet water temperature should be dynamically adjusted according to the data center utilization.

B. The impact of the optimization based on ASTT

The total cost by employing ASTT-P with different ASTT (ASTT from 5 to 80 minutes) is shown in Figure 6. The total costs of five traces with different ASTT are normalized against the total cost of five traces when ASTT is 5 minutes. Observed from this figure, the total cost of five traces can be reduced considerably when we select an optimal ASTT for them, though the best ASTT for five traces are not the same (around 30 minutes to 50 minutes) due to the small variation of the benefit of server tion($B_{sleeping}$). In the following, we select 40 minutes ASTT as the optimal ASTT for ASTT-P in the five traces. For ASTT-AR, we also obtained similar curves for five traces, though the optimal ASTT (around 60 minutes) of ASTT-AR for five trace is longer than that of ASTT-P due to the inaccurate prediction and the relatively slow growth of total cost. 60 minutes ASTT is selected as the optimal ASTT for ASTT-AR in the five traces for the following analysis.

Figure 7 shows the benefit comparison between ASTT-P (ASTT = 40 minutes) and ASTT-AR (ASTT = 60 minutes) for five traces. All the total costs are normalized against the total cost of ASTT-P (ASTT = 5 minutes) respectively. ASTT-P offers the most benefit compared with ASTT-AR but it requires an unreachable perfect prediction. As a practical algorithm, ASTT-AR still saves considerable cost while it guarantees the response time of 99% requests in the datacenter.

C. Joint optimization based on sweet temperature and ASTT-AR

Figure 8 shows the benefit when we combine dynamic optimal inlet water temperature (i.e. sweet temperature) and ASTT-AR for the five traces. The total costs of five traces are normalized against the total costs in five traces with

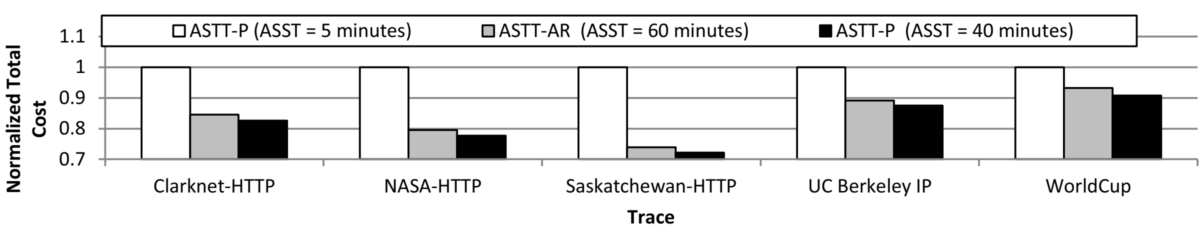

Figure 7: The total cost of ASTT-P with ASST (5 minutes), ASTT-AR with ASST (60 minutes), and ASTT-P with ASST-P with ASST (40 minutes) in five traces

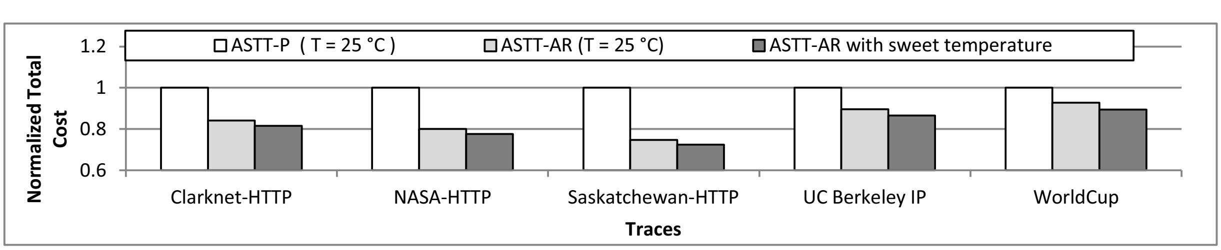

Figure 8: The total cost of ASTT-P with ASST (60 minutes) & fixed inlet water temperature (25 °C), ASTT-AR with ASST (60 minutes) & fixed inlet water temperature (25 °C), and ASTT-P with ASST (60 minutes) & sweet temperature in five traces.

ASTT-P (ASTT = 5 minutes and inlet water temperature fixed at 25 °C) as the baseline which represents a typical scheme. Overall, the total costs of sweet temperature and ASTT-AR offers 18% savings of total cost of five traces compared with the baseline in arithmetic mean based on our simulation results.

VIII. CONCLUSION

The quick growth of electricity bill drives owners of data centers to employ server consolidation and the high temperature of data center. However, the traditional air cooling system offers limited benefit of these two approaches due to its low energy efficiency of cooling power especially. We build a comprehensive framework which covers the costs of server power, cooling power, and hardware maintenance. Based on the models, we introduce a joint optimization of the costs of electricity and server maintenance. The approach gains 18% savings of the total cost and guarantees the response time of more than 99% requests. In the future, our framework will incorporate elaborated reliability models for state of the art servers and power managements which are also important for minimizing costs of data center owners.

IX. REFERENCES

[1] [Online]. http://www.micron.com/products/dram

[2] [Online]. Catalog of Water-Cooled Screw Compressor Chillers. http://chillertrader.com/manufacturer/McQuay

[3] F. Ahmad and T. N. Vijaykumar, "Joint Optimization of Idle and Cooling Power in Data Centers While Maintaining Response Time," in ASPLOS, 2010, pp. 243-256.

[4] L. A. Barroso and U. Hölzle, The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines. Morgan and Claypool Publishers, 2009.

[5] M. H. Beitelmal and C. D. Patel, "Model-Based Approach for Optimizing a Data Center Centralized Cooling System," Technical Report No. HPL-2006-67, Hewlett Packard Laboratories, April 2006.

[6] G. Bolch, et al. "Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications," John Wiley, New York, 1998.

[7] California Energy Commission, "Summer time energy-saving tips for businesses," consumerenergycenter.org/tips/business summer.html.

[8] Y. Chen, et al., "Managing Server Energy and Operational Cost in Hosting Centers," in SIGMETRICS '05, 2005, pp. 303-314

[9] J. Chang, et al, "Green Server Design: Beyond Operational Energy to Sustainability," in In Proceedings of the 2010 International Conference on Power Aware Computing and Systems, HotPower'10, 2010.

[10] EPSMA, "Guidelines to Understanding Reliability Prediction," 2005.

[11] X. Fan, ,et al, "Power provisioning for a warehouse-sized computer," In ISCA, pages 13–23, 2007

[12] J. G. Elerath, "Specifying reliability in the Disk Drive industry: No More MTBF," In Proceedings of the Annual Reliability and Maintainability Symposium, pp. 194-199,2000

[13] M. Elnozahy, et al, "Energy-Efficient Server Clusters," In Proceedings of the Second Workshop on Power Aware Computing Systems, February 2002.

[14] W. Huang, et al., "PTL-C2D: Power and Thermal Limits from Chips to Datacenters", in Tutorial of ISCA 2012.

[15] W. Huang, et al., "Tapo: Thermal-aware power optimization techniques for servers and data centers," In Proc. of IGCC, 2011.

[16] D. C. Hwang, et al, "Energy savings achievable through liquid cooling: A rack level case study," in ITtherm, June 2010, pp. 1-9.

[17] Intel® Core™2 Duo Processor E8000¹ and E7000¹ Series and Intel® Pentium® Processor E5000¹ Series Thermal and Mechanical Design Guidelines.

[18] S. Li, et al., "System implications of memory reliability in exascale computing," in High Performance Computing, Networking, Storage and Analysis (SC), 12-18 Nov. 2011, pp. 1-12.

[19] J. Lin, et al, "Thermal modeling and management of dram memory systems," in ISCA, pp. 312-322, 2007.

[20] M. Lin, et al, "Dynamic right-sizing for power-proportional data centers," INFOCOM, 2011 Proceedings IEEE , pp.1098-1106, April 2011

[21] Linux Documentation. latest release 2.6, http://www.kernel.org/doc/documentation/power/states.txt.

[22] D. Meisner, et al, "Powernap: Eliminating server idle power," In ASPLOS '09: pages 205–216, 2009.

[23] D. Meisner and T. F. Wenisch, "DreamWeaver: Architectural Support for Deep Sleep," InASPLOS, 2012.

[24] J. Moore, et al, "Making Scheduling "Cool": Temperature-Aware Workload Placement in Data Centers," in Proceedings of USENIX, 2005.

[25] M. K. Patterson, "The Effect of Data Center Temperature on Energy Efficiency," in ITHERM, May 2008, pp. 1167-1174.

[26] S. Pelley, D. Meisner, T. F. Wenisch, and J. W. VanGilder, "Understanding and absracting total data center power," in WEED '09: Workshop on Energy Efficient Design, 2009.

[27] Eduardo Pinheiro, et al. "Failure trends in a large disk drive population," In FAST, 2007

[28] A. Qouneh, C. Li, T. Li, "A quantitative analysis of cooling power in container-based data centers," in IISWC, 2011,pp. 61-71

[29] P. Ranganathan, et al,"Ensemblelevel power management for dense blade servers," in ISCA ,34(2):66–77, 2006.

[30] R. Raghavendra, et al. "No "power" struggles: coordinated multi-level power management for the data center ," in ASPLOS '08, 36(1):48-59, March 2008

[31] B. A. Rubenstein, et al, "Hybrid cooled data center using above ambient liquid cooling," in ITherm, pp.1-10, 2-5 June 2010

[32] B. Schroeder and G. A. Gibson, "Disk failures in the real world: what does an MTTF of 1,000,000 hours mean to you?," in FASTT '07, 2007.

[33] S. Srikantaiah, et al. "Energy aware consolidation for cloud computing," In Proceedings of the 2008 conference on Power aware computing and systems (HotPower'08), 2008

[34] J. Srinivasan, et al, "Lifetime reliability: toward an architectural solution," in Micro, vol. 25, no. 3, pp. 70-80, May 2005.

[35] D. S. Stoffer and R. H. Shumway, "Time Series Analysis and Its Applications," Springer Verlag, New York, 2000.

[36] Severin Zimmermann, et al, "Aquasar: A hot water cooled data center with direct energy reuse, Energy," 43(1): 237-245, July 2012

[37] Traces in the Internet Traffic Archive. [Online]. http://ita.ee.lbl.gov/

[38] USDOD, Reliability Prediction of Electronic Equipment. Military Handbook,MIL-HDBK-217F NOTICE 2, 1995.