On the Cooling of Energy Efficient Storage

Jian Zhou[1 2], Jun Wang[2], Fei Wu[1 *], Changsheng Xie[1], and Dezhi Han[3]

[1]Huazhong University of Science & Technology
[1]Wuhan National Laboratory for Optoelectronics
[2]University of Central Florida
[3]Shanghai Maritime University
[*]Corresponding Author: wufei@hust.edu.cn

***Abstract*—Energy consumption has become an important issue in storage systems. Existing energy control solutions emphasize power consumption without considering re- liability degradation that results from overburden of those long standing disks. In this paper, we develop a novel multiple criteria optimization scheme based on Fuzzy Decision Making theory, for the Cool Energy Efficient Storage System called CEES. CEES aims to enforce a temperature constraint as well as performance requirements while also keeping energy consumption to a minimum. This is achieved by aggregating all the decision criteria, such as I/O performance, power consumption, temperature and frequency of disk-status transition. We first calculate the satisfaction degree of each criteria. Then, we use the weighted averaging satisfaction degree to determine the system control sequence. The experimental results show that CEES is able to reduce disk temperature by 20-30% as compared with existing control methods, while obtaining comparable performance and power consumption.**

Keywords : Temperature, Fuzzy control, Energy-efficient storage.

I. INTRODUCTION

Due to staggering electricity bills and high disk failure rates, an increasingly large number of energy and temperature control requirements have been introduced on modern storage systems. In order to meet these requirements, data centers have been considering increasing the setpoint temperature at which to run the cooling system[1], [2]. On the other hand, various energy conservation schemes in storage systems have been developed to aggregate heavy workloads on a few disks with a method called Energy Efficient Storage. While both of these methods reduce energy consumption, they could elevate temperature levels on long standing disks and ultimately cause them to overheat. S. Yin [3] studied the reliability model of Energy Efficient Storage, which indicated that the temperature between disks, the age of the disks, and the distribution of workload between disks can elevate their failure rate. N. El-Sayed [2] found that of the three factors listed above, the temperature of the disks had the greatest influence on their failure rate.

Researchers have developed several well recognized temperature constrained schedule schemes for CPUs[4], clusters[5], and connection intensive applications[6]. J. Moore [7] and C. Bash [8] developed open-loop search and optimization methods based on the assumption that the power consumption of chip multiprocessors (CMPs) at each Dynamic Voltage and Frequency Scaling (DVFS) level can be estimated accurately. While this method is effective when the workload pattern is predictable, dramatic workload changes may lead to severe performance degradation or even power constraint violations.

Recently, several closed-loop control solutions have been developed for CMPs using heuristics [4], [5], that employ basic control theory such as model predictive control (MPC). However, MPC requires a pre-defined fixed temperature constraints for the controlled system[9]. The fixed temperature constraint is used to initiate the process of cooling down the disks if their temperature surpass a certain limit. This rigid control has several drawbacks. First, the optimal temperature at which to set constrains can often vary in real world applications. Also, if the workload rises to an extreme level, the MPC will be unable to find a low temperature working groups at which the desired constraint can be achieved with out using mixed constraints or multi-objective functions. If the constraint is set at a high temperature, the disks will be in danger of failure. If we set multiple levels of constraints, the system may still be exposed in frequently disk status switching when performing the transfer between each level. The Fuzzy Decision Making implemented by CEES can allow for a smooth transition between a range of constraints, thus eliminating the possibility of a sudden spike in the storage system. Second, multiple constraints such as warrantable service time, power-state transition frequency of given discs, etc. [3], [10] need to be formalized in the predefined model. This makes the design of a MPC controller very complex. CEES can simplify multidimensional optimization into one dimensional optimization by aggregating all the criteria.

In this paper, we attempt to apply fuzzy control theory to thermal management in order to reduce disk temperature in energy-efficient storage systems. Compared to the existing work, the following contributions has been made.

- We made the first attempt to address the temperature overhead in existing Energy Efficient Storage.
- We employed Fuzzy Decision Making to achieve complexity multi-constraint optimization.

The paper is organized as follows. Section II describes the modeling, design and analysis of Fuzzy Decision Making functions. Section III provides the implementation details and Section IV presents extensive experiments and results. Section V introduces related work. Finally, Section VI concludes the paper.

II. Design of CEES

In order to perform temperature-constrained control, we designed a feedback control loop. The key components in the control loop include a *temperature sensor* on each disk, a *power monitor* connected to the power supply circuit of each disk and an *online model estimator*. Let SP denote the sampling period, and k denote the sampling point. S_k denotes current time active disk set at time k, L denotes the data layout, $Perf_k$ denotes the required performance by the user at time k, P_k denotes the measured disk power consumption, T_k denotes the measured disk temperature and O_k denotes the operation time since the last spin down or spin up a disk. S_k, L, $Perf_k$, P_k, T_k and O_k are all input signals for our fuzzy controller to calculate an appropriate active disk set in the next time window SP. We express our feedback control loop in the following function:

$$S_{k+1} = F(S_k, L, Perf_k, P_k, T_k, O_k) \tag{1}$$

where, F denotes the feedback algorithm.

The control loop is invoked periodically and its period is chosen based on a trade-off between actuation overhead and system settling time. The following steps are invoked at the end of every control period. 1) The monitors collect the temperature and power consumption of each disk as well as the overall performance status of the storage system. 2) The online model estimator updates its parameter based on the collected data. 3) Based on the collected data, the controller computes future temperatures of the storage system and selects the working disks accordingly.

We propose Fuzzy Decision Making (FDM) as the feedback algorithm F [11] for the temperature-constrained scheduler for Energy Efficient Storage. This method defines a range of temperature constraints and quantifies the constraints with a satisfaction degree from 0 to 1; thus, a smooth control can be achieved.

The controller uses the system model to predict the control behavior over sampling periods, H_p. This is called the *prediction horizon*. The control objective is to select an input trajectory that minimizes the *objective function*. The objective function, is formulated to represent the satisfaction of the decision criteria and control goals after applying the control actions in the entire prediction horizon, H_p. An input trajectory includes the control inputs in the following H_m sampling periods. S_{k+1}, ..., S_{k+H_m}, where H_m is called the *control horizon*. The notation S_{k+i} means that the disk state vector of the storage system at time $k+i$ depends on the conditions at time k. Once the input trajectory is computed, only the first element S_{k+1} is applied as the control input to the system. At the end of the next sampling period, the prediction horizon slides on sampling period and the input is computed again based on the feedback from the performance, power and temperature monitors.

A. Multi-criteria Constraints in Power Consumption Optimization

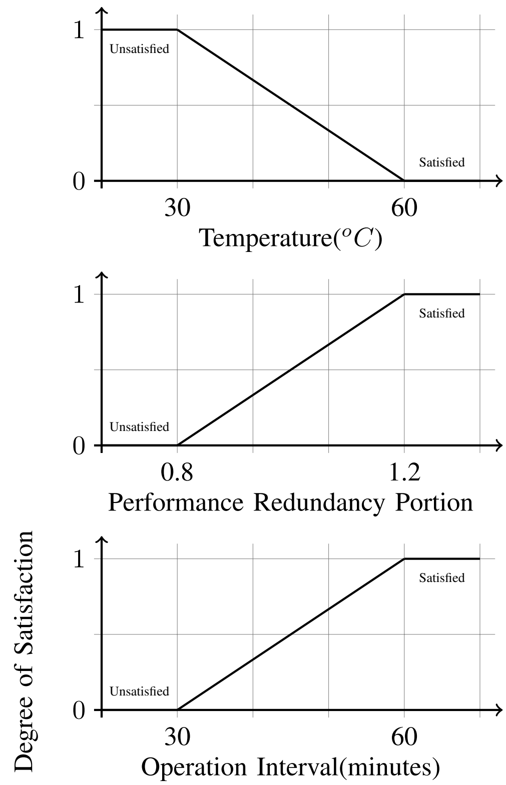

Fig. 1. Liner Satisfaction Membership

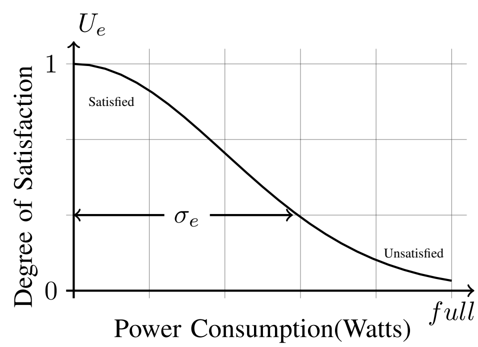

Fig. 2. Exponential Satisfaction Membership

Because I/O performance, Power Consumption, Temperature, and Operation Intervalare measured with separate continuous units, their individual satisfaction levels can be mapped to the interval [0,1]. The control object is subjected to three sets of constraints. First, the performance level should meet the minimal requirements of the workload. Second, the temperature of each disk should be below a given threshold (e.g., 50^oC). Third, the interval between power up and shut down of a certain disk should be above a given time (e.g., 30 minutes). In MPC, these constraints are the “hard” constraints, which could create a scenario in which no optimal solution can be achieved [12]. Therefore, in our design, *fuzzy constraints* are used. For simplicity and reserving the "hard" constraints to indicate the physical limitation, the satisfaction of the *fuzzy constrains* is expressed by the liner membership function shown in Figure 1.

The fuzzy temperature constraint criteria indicate how well the temperature of each disk are satisfied. Basically, there are two types of temperatures regarding the disk failure model. Sankar states that there is an exponential relationship between the temperature and disk failure models [13]. El-Sayed points out that for temperatures below 50^oC, disk failure rates grow as a linear function [2] when the temperature increases. Both of the models can be used as the temperature membership function. For simplicity, the fuzzy temperature constraint is formulated by the following equation:

$$\mu_{t(k+i)} = \begin{cases} 1, T_{k+i} \leq \sigma_{t0} \\ \frac{\sigma_{t1}-T_{k+i}}{\sigma_{t1}-\sigma_{t0}}, \sigma_{t0} < T_{k+i} < \sigma_{t1} \\ 0, T_{k+i} \geq \sigma_{t1} \end{cases} \tag{2}$$

Where, σ_{t0} and σ_{t1} is the temperature at which the constraint is fully satisfied and unsatisfied respectively. For a regular storage system, σ_{t0} could be 30^oC, and σ_{t1} could be 55^oC.

The fuzzy performance constraints criteria indicate how well the performance requirements are satisfied. Sufficient performance redundancy are essential for storage systems, in order to ensure successful and reliable data service. The fuzzy performance constraints can be formulated by the following equation:

$$\mu_{p(k+i)} = \begin{cases} 0, \frac{Perf_{k+i}}{Perf_k} \leq \sigma_{p0} \\ \frac{\frac{Perf_{k+i}}{Perf_k}-\sigma_{p0}}{\sigma_{p1}-\sigma_{p0}}, \sigma_{p0} < \frac{Perf_{k+i}}{Perf_k} < \sigma_{p1} \\ 1, \frac{Perf_{k+i}}{Perf_k} \geq \sigma_{p1} \end{cases} \tag{3}$$

Where, σ_{p0} and σ_{p1} is the portion of performance redundancy at which the performance constraints are fully unsatisfied and satisfied respectively. These constraints could be selected based on specific workloads. For a regular storage system, σ_{p0} could be 0.8, and σ_{p1} could be 1.2.

The fuzzy operation interval constraint criteria indicate the time interval between two disks status switching. It is formulated by the following equation:

$$\mu_{O(k+i)} = \begin{cases} 0, O_{k+i} \leq \sigma_{o0} \\ \frac{O_{k+i}-\sigma_{o0}}{\sigma_{o1}-\sigma_{o0}}, \sigma_{o0} < O_{k+i} < \sigma_{o1} \\ 1, O_{k+i} \geq \sigma_{o1} \end{cases} \tag{4}$$

Where, σ_{o0} and σ_{o1} is the interval at which the constraints fully unsatisfied and satisfied respectively. For a regular storage system, σ_{o0} could be 30 minitus, and σ_{o1} could be 60 minitus.

A fuzzy set in the appropriate domain characterizes both the *fuzzy goals*, such as reducing the power consumption, and the *fuzzy constraints*, such as temperature constraints. Applying FDM in the temperature-constrained power controller allows the combination of goals and constraints to be achieved. The estimated power consumption is mapped to the interval [0, 1] which indicates how well the power consumption satisfies the goal to minimize it by utilizing a membership function. The *fuzzy goals* are defined by a Gaussian membership function that is never been zero, as shown in Figure 2. That indicates that high power consumption is an allowable but not a desirable state. The *fuzzy goal* can be formulated by the following equation:

$$\mu_{e(k+i)} = exp(-\frac{(S_{k+i}P_k)^2}{2{\sigma_e}^2}) \tag{5}$$

Where S_{k+i} is the disk state vector at time $k+i$, P_k is the power consumption vector of each disk, σ_e is used to determine how fast the *fuzzy goal* approach as the power consumption increase. For a regular storage system, the σ_e could be half of the max energy consumption in a sampling period.

B. Aggregation of Criteria for Energy Efficient Storage

Because all of performance criteria can effect each other, the satisfaction levels of all of these factors must be aggregated into an equation and maximized. The fuzzy criteria aggregation is the process that computes the joint satisfaction of all the criteria [12]. The confluence of goals and constraints can be done by aggregating the membership values. The membership value $\mu_\pi(k)$ for the control sequence π is obtained using the aggregation operators $\otimes$, $\otimes_g$ and $\otimes_c$ to combine the decision criteria in Equation 6:

$$\begin{aligned} \mu_\pi(k) = &(\mu_{e(k+1)} \otimes_g ... \otimes_g \mu_{e(k+H_p)}) \\ &\otimes (\mu_{p(k+1)} \otimes_c ... \otimes_c \mu_{p(k+H_p)}) \\ &\otimes (\mu_{t(k+1)} \otimes_c ... \otimes_c \mu_{t(k+H_p)}) \\ &\otimes (\mu_{o(k+1)} \otimes_c ... \otimes_c \mu_{o(k+H_p)}) \end{aligned} \tag{6}$$

In Equation 6, $\otimes_g$ denotes an aggregation operator for combining the goals, $\otimes_c$ denotes an aggregation operator for combining the constraints, $\otimes$ denotes an aggregation operator for combining the aggregated goals and constraints. The operator $\otimes_g$ is the average operator. It computes the average satisfaction of the energy consumption. So, the accumulated energy consumption, rather than the energy consumed in certain sampling period, is taken in to account. The operator $\otimes_c$ is the minimum operator. The smallest satisfaction of the constraints are chosen as the decision criteria. The operator $\otimes$ is the weighted averaging operator.

The control object is to find a solution that best satisfies the joint constraint of all the fuzzy criteria. The ideal results in which we desire is a situation that all the criteria are satisfied. In this case, the storage system will get 1) sufficient performance redundancy, 2) low temperature of disks, and 3) few disks spin down spin up actions, and 4) minimal energy consumption in the prediction horizon. The requirement manifested by an "and" operator of the criteria values. The ordered weighted averaging aggregation operators [14] are used in the aggregation of goals and constraints. Assume the ordered weight vector W is $[w_e, w_p, w_t, w_o]$, the μ_π can be expressed as:

$$\mu_\pi = w_e\mu_e + w_p\mu_p + w_t\mu_t + w_o\mu_o \tag{7}$$

Where $w_e + w_p + w_t + w_o = 1$.

The translation of control goals and constraints to a membership value avoids the specification of the criteria in a large multidimensional space. The decision criteria 7 should be satisfied as much as possible, which corresponds to the maximal value of the overall satisfaction. Thus, the optimal sequence of control actions π^* is found by the maximization of μ_π.

$$\pi^* = argmax(\mu_\pi) \tag{8}$$

III. Prototype and Implementation

In this section, we describe our physical experimental testbed and benchmarks, as well as the implementation details of each component in CEES control loop.

We develop CEES prototype on a machine with an Intel Dual-Core E5200 2.5GHz CPU, 2G Bytes DDR2 memory running an Ubuntu 11.04 operating system. We implement a trace replayer on our power-proportional data layout based storage system testbed and measure the performance and power consumption of disks for each run. Due to the hardware limit, the prototype contains 16 disks, which are connected by an SAS cable. To measure the power consumption of disks, we adopt a plug-in real-time multi-meter, called ZH102.

A. Trace Replay Framework

In our prototype system, we implement the trace replay framework as a C program running under Ubuntu Server 12.04. Because of the heavy integration of our control algorithm, a new trace replay code are developed rather than using an existing one. It has $8.5k$ code lines in all. The main idea is to use the "libaio" programming library to asynchronously access I/O of the storage subsystem (sending I/O requests to the storage subsystems). The trace replay framework consists of 6 modules: performance requirements monitor, power & temperature monitor, model predictor, fuzzy controller, data layout manager, and virtual disk layer.

B. Power Proportional Layout

We implemented power proportional layout [15] in our storage system. In the experiment, we made three replicas of the whole data set in the virtual disk. The first replica is spread on three disks, which are called disk *Group 1* in the experiment. The second replica is stored on five disks, which are called disk *Group 2* in the experiment. The last replica is saved on eight disks, which are called disk *Group 3* in the experiment. We use a total of 16 disks in the experiments.

C. Write Off-loading

CEES use an alternated version of Write Off-loading [16] technique to maintain the power proportional feature. Write Off-loading is an energy saving technique for storage system, which prolong the idle time of standby disks by redirecting writes to active disks temporarily. It redirects the writes to the currently active disks and updates the writes to the other replicas when a specific trigger occurs. While the trigger used in the original Write Off-loading is the off-loaded time span or data amount, the only trigger used in our experiment, including CEES and other baselines, is that the controller changes the active disk sets based on the constraints.

IV. Experimental Results and Analyses

We present our experimental results and associate in-depth analyses here.

A. Workload

We use a mixture of real-world and synthetic traces to comprehensibly study the impact of different storage architectures on a wide spectrum of enterprise-scale workloads. Table I presents salient features of our workloads. We employ a write-dominant I/O trace from an OLTP application running at a financial institution and a popular Internet web search machine [17] made available by the Storage Performance Council (SPC), henceforth referred to as the Financial1, Financial2 and Websearch traces. These traces are collaboratively collected by HP and Storage Performance Council. Exchange [18] was collected at the Microsoft Exchange 2007 SP1 server, which is a mail server for 5000 corporate users. MSN [18] was collected at the Microsoft's several Live file servers. Develop [18] was obtained from a file server accessed by more than 3000 users to download various daily builds of Microsoft Visual Studio. Radius [18] was obtained from a RADIUS authentication server that is responsible for worldwide corporate remote access and wireless authentication.

Workloads	Avg.Req.Size read/write(KB)	Read (%)	Avg.Req.Arrv. Time (ms)
Financial1	2.25/3.75	23.2	8.19
Financial2	2.3/2.9	82.3	11.08
Websearch	15.15/8.6	99.9	2.99
Exchange	15.15/14.5	30.8	1179
MSN	9.6/11.1	67.2	513
Develop	18.45/10.95	88.6	1985
Radius	124.25/12.45	17.1	9475

TABLE I
Enterprise-Scale Workload Characteristics.

B. Baselines

The first baseline system, referred to Rabbit[15], is a power proportional storage. It ensures ideal power-proportionality, by providing multiple gears of storage to work under different performance requirements. All the baselines including CEES are use the same data layout in order for a fair comparison.

The second baseline, referred to Simple Feedback Controller (SFC), is a simple temperature feedback control loop. SFC

	Temp.(^oC) Min/Avg/Max	Power(W) Avg	Actions	Satisfy Avg
CEES	33.5/38.6/39.8	48	4	0.79
MPC	36.0/38.0/38.5	46	9	0.46
SFC	30.2/40.3/42.0	47	6	0.38

TABLE II
MULTI-CONSTRAINTS OPTIMIZATION.

Replay Speed	Temp.(^oC) Min/Avg/Max	Power(W) Avg	Actions	Satisfy Avg
x4	33.5/38.6/39.8	88	4	0.72
x8	36.0/39.2/38.5	130	5	0.63
x16	37.9/45.3/42.0	156	8	0.48
x32	41.2/48.3/52.0	162	10	0.35

TABLE III
MULTI-CONSTRAINTS OPTIMIZATION ACCELERATED TRACE REPLAY.

represents a typical feedback solution which use the real-time per-disk temperature to control the behavior of disks without relying on an online model estimator. We compare our CEES against SFC to show that a well-designed simple temperature feedback controller may still fail to enforce accurate temperature control and thus degrade performance.

The third baseline, referred to Model Predict Control (MPC), is a recent power management solution in CPUs[4]. As we discuss before, the MPC control of CPUs cannot be directly used in storage system. We implement the MPC controller under the framework of SFC control. It shares most part of the SFC controller. A fundamental difference between SFC and MPC controller is that SFC simply uses the moving average temperature of a selected disk by one step, depending on whether the measured temperature is lower or higher than the set point. In contrast, MPC computes a predicted temperature level for each disk according to the temperature prediction model.

C. Experiment

To demonstrate the superiority of the Fuzzy Constraint Satisfaction design in CEES, we conducted a series of real word experiment.

1) Multi-Constraints Optimization: One of the main contribution of CEES is that it provides a better way to perform multi-constraints optimization between performance requirement, temperature, power consumption and number of disk status transfer actions.

In this experiment, we carefully choose the parameters for CEES, MPC and SFC, which guarantee that the system can run smoothly in a manipulated environment without any interference on performance. The experiment runs for a fixed 3 hours under WebSearch traces. We get the different running state in Table II. We can see that CEES provides a better way to control aggregated satisfaction of multiple constraints compared to MPC and SFC, since they do not do multiple constraints' optimization and are not able to adjust the constraints automatically. In this table, CEES has fewer total control actions and achieves better aggregated satisfaction degree. For MPC and SFC, aggregated satisfaction degree can be designed to be higher if we configure the system in the right way at each run. However, this is not practical.

In order to study the system behavior under heavy workloads, we accelerated the replay speed while replaying WebSearch traces. In Table III, we can find that the controller relatively consumes more power when the workload is low. This is because of the exponential power consumption satisfaction model has smaller slope at the lower power consumption points. This design can help reserve more system performance and reduce the temperature when the workload is low. When the workload is high, the controller works harder. In this experiments, the system reduce the system I/O performance to maximize the satisfaction degree.

2) Temperature: In this experiment, we evaluate the reduction of overall disk temperature of CEES at different I/O patterns, compared to the power proportional storage — Rabbit [15]. In order to do a fair comparison, we choose looser constrictions for SFC and MPC, so that they are able to run stably at various conditions, because they do not have a well designed constrict change policy. Figure 3 plots the normalized average temperatures of disks with different control methods under different workloads. The result shows that CEES can significantly reduce the temperature adaptively and accurately based on the online model prediction and fuzzy decision making.

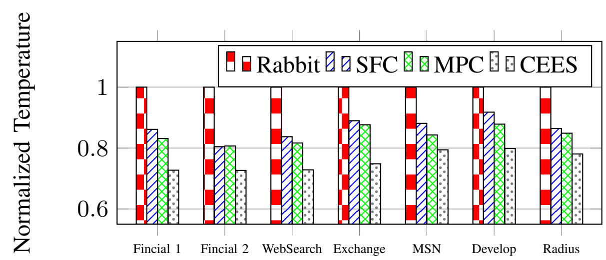

Fig. 3. Normalized Temperature under Different Traces.

3) Energy Consumption: In this experiment, we evaluate the energy consumption under different workloads. Figure 4 shows that CEES has a minimal impact on the energy consumption compared with normal power proportional storage.

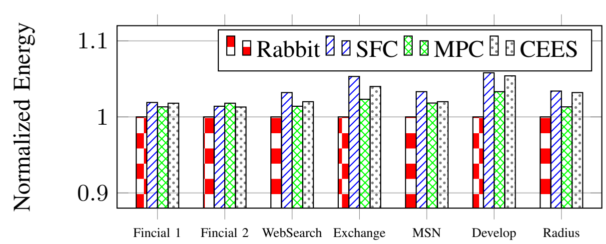

Fig. 4. Normalized Energy Consumption under Different Traces.

4) Performance: In this experiment, we evaluate the impact of our controller on the performance at different I/O intensities. Figure 5 plots the average I/O delay under different workloads.

It shows that CEES makes the least impact on the performance. This is because CEES has a very accurate prediction module and it can dynamically change the temperature threshold when the temperature vibrates.

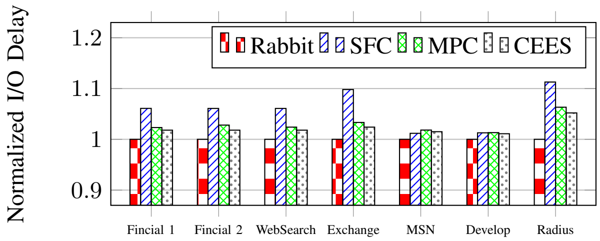

Fig. 5. Normalized I/O Delay under Different Traces.

V. Related Work

Several studies have been conducted for the purpose of studying the relationship between disk temperature and failure rate. A study conducted by Google in 2007 [19] suggested that lower temperatures are actually more detrimental to disk reliability than higher temperatures. However a research conducted by Microsoft in 2011 [13], [20] demonstrated that Annualized Failure Rate (AFR) steadily increases as Hard Disk Drive (HDD) temperature increases. The most recent research conducted by N. El-Sayed in 2012 [2] suggests that the reason why Google's study arrived at the conclusion is because when the study was conducted, many disks of various models were used to aggregate the data and different models of disks failed at different temperatures. N. El-Sayed [2] observed that the increase in failures with respect to temperature tends to be linear, except for very high temperatures (above 50^oC). Because of this relationship between disk temperature and failure [2], [13], [20], [21], a temperature constrained control scheme is essential to the well-being of Energy Efficient Storage systems.

Moorey *et. al.* developed three dynamic thermal managements [7] to deal with the stability of the scheduling arithmetic, avoid the changes in temperature and minimize the cost of thermal management. Heath *et. al.* developed a thermal emulation model called Mercury along with a very simple thermal feedback control called Freon [22]. These arithmetics simply schedule the load to the low temperature node, but may schedule the load to a machine that is hard to cool. In our design, a online temperature estimation model is used to avoid this. Abbasi proposed a two-tier dynamic server provisioning and workload distribution method in developing thermal aware Internet data centers[23]. Weissel *et. al.* used Newton's Law of Cooling to predict dynamic thermal management for distributed systems[24]. Ramos *et. al.* improved the C-Freon in C-Oracle[25] using online thermal prediction. Their basic method is to control the load intensity of a server. Wang *et. al.* introduced the online model to predict the CPU temperature control [4]. However, they do not study at which temperature should we set the constraints. Our design gives a flexible constraint for the temperature while balancing it with multiple other constraints.

VI. Conclusions

In this paper, we present the first study of applying a Fuzzy Decision Making theory in energy-efficient storage systems to manage the energy while adhering to the temperature constraints. Comprehensive experimental results on a physical test-bed show that CEES outperforms two state-of-art algorithms by significantly reducing the temperature in energy-efficient storage systems. More specifically, CEES reduces the temperature by 20% - 30% compared to current methods such as Rabbit and MPC, while maintaining comparable performance and power consumption.

VII. Acknowledgement

This work is supported in part by the US National Science Foundation Grant CNS-1115665, CCF-1337244, the National Science Foundation Early Career Award 0953946, the National Basic Research Program of China (973 Program) under Grant No.2011CB302303, the National Natural Science Foundation of China No. 61300047, No. 61370063, and No. 61472152, the National Natural Science Foundation of Hubei Province No.2015CFB315, and the National Natural Science Foundation of China Grant 61373028 and 61070154.

References

[1] R. Miller, "Google: Raise your data center temperature," http://www.datacenterknowledge.com/archives/2008/10/14/google-raise-your-data-center-temperature/, 2008.

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

[3] S. Yin, K. Li, X. Qin, X. Ruan, H. Chen, J. Xie, and X. Zhu, "Reliability analysis of energy-efficient parallel storage systems," in *High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing (HPCC_EUC), 2013 IEEE 10th International Conference on*. IEEE, 2013, pp. 1448–1455.

[4] Y. Wang, K. Ma, and X. Wang, "Temperature-constrained power control for chip multiprocessors with online model estimation," in *ACM SIGARCH Computer Architecture News*, vol. 37, no. 3. ACM, 2009, pp. 314–324.

[5] X. Wang and M. Chen, "Cluster-level feedback power control for performance optimization," in *High Performance Computer Architecture, 2008. HPCA 2008. IEEE 14th International Symposium on*. IEEE, 2008, pp. 101–110.

[6] G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao, and F. Zhao, "Energy-aware server provisioning and load dispatching for connection-intensive internet services," in *Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation*, vol. 8. USENIX Association, 2008, pp. 337–350.

[7] 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, pp. 5–5.

[8] C. Bash and G. Forman, "Cool job allocation: measuring the power savings of placing jobs at cooling-efficient locations in the data center," in *2007 USENIX Annual Technical Conference on Proceedings of the USENIX Annual Technical Conference*. USENIX Association, 2007, p. 29.

[9] E. F. Camacho, C. Bordons, E. F. Camacho, and C. Bordons, *Model predictive control*. Springer London, 2004, vol. 2.

[10] J. Wang, H. Zhu, and D. Li, "eraid: Conserving energy in conventional disk-based raid system," *Computers, IEEE Transactions on*, vol. 57, no. 3, pp. 359–374, 2008.

[11] J. M. da Costa Sousa and U. Kaymak, "Model predictive control using fuzzy decision functions," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 31, no. 1, pp. 54–65, 2001.

[12] Z. Ruttkay, "Fuzzy constraint satisfaction," in *Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on*. IEEE, 1994, pp. 1263–1268.

[13] S. Sankar, M. Shaw, and K. Vaid, "Impact of temperature on hard disk drive reliability in large datacenters," in *Dependable Systems & Networks (DSN), 2011 IEEE/IFIP 41st International Conference on*. IEEE, 2011, pp. 530–537.

[14] R. R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, no. 1, pp. 183–190, 1988.

[15] H. Amur, J. Cipar, V. Gupta, G. Ganger, M. Kozuch, and K. Schwan, "Robust and flexible power-proportional storage," in *Proceedings of the 1st ACM symposium on Cloud computing*. ACM, 2010, pp. 217–228.

[16] D. Narayanan, A. Donnelly, and A. Rowstron, "Write off-loading: Practical power management for enterprise storage," *ACM Transactions on Storage (TOS)*, vol. 4, no. 3, p. 10, 2008.

[17] "I/o. umass trace repository," http://traces.cs.umass.edu/index.php/Storage/Storage.

[18] "Microsoft enterprise traces," http://iotta.snia.org/tracetypes/3.

[19] E. Pinheiro, W.-D. Weber, and L. A. Barroso, "Failure trends in a large disk drive population." in *FAST*, vol. 7, 2007, pp. 17–23.

[20] S. Sankar, M. Shaw, K. Vaid, and S. Gurumurthi, "Datacenter scale evaluation of the impact of temperature on hard disk drive failures," *Trans. Storage*, vol. 9, no. 2, pp. 6:1–6:24, Jul. 2013. [Online]. Available: http://doi.acm.org/10.1145/2491472.2491475

[21] W. Digital, "Thermal reliability technical paper."

[22] T. Heath, A. Centeno, P. George, L. Ramos, Y. Jaluria, and R. Bianchini, "Mercury and freon: temperature emulation and management for server systems," in *ACM SIGARCH Computer Architecture News*, vol. 34, no. 5. ACM, 2006, pp. 106–116.

[23] Z. Abbasi, G. Varsamopoulos, and S. Gupta, "Thermal aware server provisioning and workload distribution for internet data centers," in *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*. ACM, 2010, pp. 130–141.

[24] A. Weissel and F. Bellosa, "Dynamic thermal management for distributed systems," in *Proceedings of the 1st Workshop on Temperature-Aware Computer Systems*, 2004.

[25] L. Ramos and R. Bianchini, "C-oracle: Predictive thermal management for data centers," in *High Performance Computer Architecture, 2008. HPCA 2008. IEEE 14th International Symposium on*. IEEE, 2008, pp. 111–122.