Control Engineering Practice 79 (2018) 154–169

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

An optimized active disturbance rejection approach to fan control in server

Qinling Zheng [a,*], Zhan Ping [a], Simone Soares [b], Yu Hu [c], Zhiqiang Gao [b]

[a] Memory Platform Lab in Samsung Semiconductor Inc., San Jose, CA 95134, USA
[b] Cleveland State University, Cleveland, OH 44115, USA
[c] Cleveland State University, Cleveland, OH 4411, USA

ABSTRACT

As more and more massive data storage drives are used in super high density, the power used to cool the servers has become an increasingly large component of the total power consumption. Therefore, improving server cooling efficiency has become an essential requirement in data centers. However, because the thermal dynamics of the server system has characteristics such as nonlinearity, significant inter-loop coupling, and continuously fast-changing/unknown workload disturbances, these pose huge challenges to control engineers and data center architect engineers. To address the above concerns, this paper presents an active disturbance rejection control based decoupling control algorithm for flash storage systems and CPUs in a one-unit (1U) server to simultaneously improve fan power consumption efficiency and regulate the server components' temperature to avoid downgraded performance caused by overheating. In the study, a benchmark system is established based on the Samsung Mission Peak (MP) server where the thermal characteristics and existing solutions are both systematically evaluated. Performance of the design concept is proven in both simulation and a hardware testbed. With the proposed control solution, experiment results show that the temperature overshoot is greatly eliminated, temperatures are more tightly controlled, and the drive throttling rate are greatly decreased. Furthermore, the proposed method is shown to be able to save up to 45% energy versus a PID controller, save about 77% energy versus dynamic fan speed control method, and save about 98% energy versus native fan speed control.

1. Introduction

Reducing the energy consumption for cooling in servers is a major challenge considering data center energy costs today, as the power used to cool the servers has become an increasingly large component of the total power consumption. There have been many studies on server power management and several of the studies (Greenberg, Mills, Tschudi, Rumsey, & Myatt, 0000; Huang et al., 2011; Mittal, 0000; Patel, Bash, Sharma, Beitelman, & Friedrich, 2003) have shown that every 1 W of power used to operate a server often requires an additional 0.5–1 W of power, needed by the cooling equipment to extract the heat. In particular, with increasingly dense computing infrastructures, such as 1U blade servers, more powerful processors and higher density flash storage system, the server fans can often consume a significant amount of power.

In recent years as the capacities of SSD have been quickly improving, there has been rapid growth in the use of 1U blade servers. Commercially available flash storage system density is also continuously getting higher, where 20–60 SSD drives are routinely packaged in one enclosure. The heat dissipation has become a design challenge that can no longer be relegated to a simple control technology. The problem calls for systematic investigations and better solutions because if it is not handled properly, the over temperature caused by the excessive heat generated by the SSDs could easily lead to degraded system performance or result in device shutdown. As a result, thermal management inside the server has become a critical challenge for the sake of system performance and total cost of ownership (TCO).

Most servers regulate temperature using a simple default native fan speed control algorithm, which linearly increases the PWM duty cycle based on the server inlet temperature (Chen, Tan, Xing, & Wang, 2014). Dynamic fan speed control is a method proposed very recently which is an improvement of the native fan speed control algorithm where fan speed is regulated based on CPU utilization (Fancontrol, 0000). Both methods are a type of feedforward control with the control signal generated based on mapping logic. The disadvantage of these approaches is that the control signal cannot continuously adjust which will cause wasted energy consumed by the fans. In contrast, feedback control is a type of approach where the control signal is able to continuously change based on temperature sensor readings. Currently, the field of server thermal control is dominated by the traditional PID technology, the most popular and classic feedback control method, where the temperature deviation is used to activate the control action in a reactive manner. In particular, in the temperature control loop the temperature is first measured by a temperature sensor and then used in the PID controller to calculate the corresponding control action. To

* Corresponding author.
E-mail addresses: jazzy.zheng@samsung.com (Q. Zheng), john.ping@samsung.com (Z. Ping), simone.cnas@yahoo.com.br (S. Soares), yuhu.leo@gmail.com (Y. Hu), z.gao@csuohio.edu (Z. Gao).

https://doi.org/10.1016/j.conengprac.2018.07.003
Received 31 January 2018; Received in revised form 29 June 2018; Accepted 3 July 2018
0967-0661/

make matters worse, when a large number of SSDs are squeezed into one small package, thermal coupling can no longer be ignored. Each SSD will have significant thermal impact to the SSDs adjacent to it and this strong coupling is what makes effective thermal management an industry-wide challenge. Existing PID technology was conceived more than a century ago for single-input, single-output (SISO) processes with tedious but otherwise manageable means of parameter tuning. Compared with the data flow, the thermal behavior of SSD is slow and difficult to maneuver, in the sense that once the heat is built up in a massive data storage device, it cannot be easily or quickly removed. The challenge posed by lumping together a large number of devices into a single package calls for better and more systematic methods of thermal control, which is critical in the future development of high-density storage devices.

These issues provide the motivation for a systematic study of server thermal problems and a comprehensive evaluation of the existing solutions, followed by the development of advanced control algorithms to be implemented for the server as well for the other high-density storage systems in the future. To address the issues above, this paper proposes an active thermal regulating/decoupling methodology for the thermal variance and interference.

In recent years, Active Disturbance Rejection Control (ADRC) has been proposed and continuously developed to address this very issue, to the degree that replacing PID is very promising. The principal of ADRC is to actively seek and estimate disturbances and cancel them out with control effort. The idea was first proposed by Han in 1995 (Han, 1955, 1998), formally presented in the English literature in 2001 (Gao, 2002; Gao, Huang, & Han, 0000), and fully designed philosophy in 2006 (Han, 2006). In particular, the transition of ADRC from an academic idea to an emerging technology was recently documented (Gao, 2015). Crucial in the process is the simplification and parameterization (Gao, 2003) that allows seamless integration of the time domain formulation and the frequency domain insights (Li, Qi, Xia, Pu, & Chang, 2015; Zheng & Gao, 2016). This led to the initial attempts at a rigorous stability analysis (Zheng, Chen, & Gao, 2009a; Zheng & Gao, 2012; Zheng, Gao, & Gao, 2007), where not only the stability characteristics but also the bandwidth-error bound relationship were established. Following the initial analysis of the ADRC, various groups have begun to use more advanced mathematical techniques, the scope of which has been well documented recently (Chen, Yang, Guo, & Li, 2016; Guo & Zhao, 2015; Huang & Xue, 2014; Madoński & Herman, 2015; Xue & Huang, 2015, 2017). It turns out that ADRC is a particularly suitable solution for thermal regulation problems and numerous papers have shown its effectiveness in both process control systems and coupled multi-input, multi-output (MIMO) systems (Miklosovic & Gao, 2005; Sun et al., 2017; Vincent et al., 2011; Wang et al., 2012; Zheng, Chen, & Gao, 2009b; Zheng, Gao, & Tan, 2011).

Reformulating the thermal problem in terms of disturbance rejection, the new proposed advanced active control strategies are designed for the MIMO fan control system which is based on thermal dynamics of the drive-zones and CPU-zones within the 1U Samsung Mission Peak blade server to manipulate the operation of the fans. This approach simultaneously addresses the device performance (such as SSD throttling rate) and energy cost, and is practical, easily implementable, and tunable for customers. Performance of the design concept is proven in simulation where the new and existing solutions are compared in terms of thermal characteristics and power efficiency. Furthermore, the control parameters are optimized with a genetic algorithm where the performance and energy are both considered while the cost function is constructed. It is shown that the proposed control design is timely and effective in heat dissipation with greatly improved power efficiency.

The paper is organized as follows. Section 2 systematically derives the mathematic model of the MP server thermal system and introduces the drive-fan thermal model, the CPU-fan thermal model, and the fan power model. In Section 3, an active disturbance rejection based thermal decoupling approach is proposed to address the coupling problem for systems with large uncertainties of the internal dynamics and significant unknown external workload disturbances. The control algorithm is optimized in the same section. To further prove the feasibility of the proposed method, simulation and hardware evaluation test results where the server performance and the energy cost are both addressed in Section 4, followed by concluding remarks in Section 5.

Table 1
Drive-fan thermal zones.

FAN	Drive zone	SSD	No. of SSD
FAN 1 (right first)	ZONE 1 (right first)	Dr. 1 ~ Dr. 8	8
FAN 2	ZONE 2	Dr. 9 ~ Dr. 12	4
FAN 3	ZONE 3	Dr. 13 ~ Dr. 16	4
FAN 4	ZONE 4	Dr. 17 ~ Dr. 21	5
FAN 5	ZONE 5	Dr. 22 ~ Dr. 25	4
FAN 6	ZONE 6	Dr. 26 ~ Dr. 30	5
FAN 7	ZONE 7	Dr. 31 ~ Dr. 36	6

2. Server thermal system

Thermal control has become an increasingly challenging industry-wide problem in high-density storage systems, especially with multiple devices in one system (or plant) that are thermally coupled to each other. However, since there is a lack of understanding of the devices' (e.g. SSDs and CPUs) thermal characteristics, state of the art server fan control often results in over provisioning of airflow leading to higher power consumption. In this section, the server thermal system model is developed. Two individual thermal zones, drive-fan and CPU-fan zones, are introduced in Part 2.1 and 2.2, respectively. The models that present the total power consumption of the fans are described in Part 2.3.

2.1. Drive-fan thermal zone model

The Mission Peak reference platform features up to 36 front-serviceable next-generation small form factor (NGSFF) SSD bays contained in only 1U as shown in Fig. 1. NGSFF devices are specifically designed for all-flash servers and optimized for 1U designs. Furthermore, with dual-socket Intel Skylake-SP CPUs and up to 24 DDR4 DIMMs of memory built-in, the system is also great for I/O intensive local workloads, such as real-time analytics and database servers.

For simplicity of analysis, fully loaded front SSDs are divided into 7 thermal zones, each of them is controlled by a fan accordingly as shown in the right part of Fig. 1. Table 1 explains the SSDs and fan selection for each thermal zone.

The mathematic thermal dynamics of a single drive-fan system is introduced in 2.1.1, followed by the whole system with 7 thermal zones as one lumped thermal system in Section 2.1.2.

2.1.1. Single drive-fan system model

The model relating the drive temperature to the fan speed, the drive utilization, and the ambient temperature in MP can be obtained using a similar method as proposed in Wang et al. (0000). Within one drive-fan thermal zone, the thermal resistance R_{Dr} between all the drives and the ambient air flowing through the blade is defined as:

$$R_{Dr} = \frac{T_{Dr} - T_{amb}}{Q_{Dr}}, \tag{1}$$

where Q_{Dr} is the heat transferred per unit of time between the drive and the ambient air; T_{Dr} and T_{amb} are the drive temperature and ambient temperature.

The power of all the drives in the ith zone is modeled as a function of their utilization as follows:

$$P_{Dr\_ZONE_i} = \sum_n \left(g_{Dr} \times L_{oad_n} + P_{Dr,idle_n} \right), \quad i = 1, \dots, 7, \tag{2}$$

where L_{oad_n} and $P_{Dr,idle_n}$ represent the workload and idle power of the nth drive in the server. Note that Eq. (2) shows that the working load and drive-zone power are linearly related and g_{Dr} is the slope of