The improvement study for global and local cooling of servers

Wenshuai Zhang¹, Zongliang Bai², Qiang Hu¹, Limei Shen¹*, Yupeng Wang¹ and Huanxin Chen¹

¹Department of Refrigeration and Cryogenics, Huazhong University of Science and Technology, Wuhan, China
²China Ship Development and Design Center, Wuhan, China

*Corresponding author, E-mail: ep_shenlimei@hust.edu.cn

Abstract. Cooling of server is getting more attention these years, which directly influence the efficiency of data processing, storage, and telecommunication. In order to solve the hot issues of sever, a thermal simulation model was built. Finding that the hotspots were mainly concentrated in CPU and Northbridge chips. With regard to CPU, the structure optimization analysis of CPU heat pipe radiator was carried out without changing the volume of heat pipe radiator. The fin thickness and fins number of heat pipe radiator were optimized. The optimized radiator could reduce the CPU hotspot temperature by 5.3%-9.09%. With regard to North Bridge chip, a chip-level cooling method based on Peltier effect was proposed. The cooling performance of thermoelectric cooler was investigated by experiments, and the best working current under different wind speed and cooling load was explored. Compared with the original heat dissipation method (non-thermoelectric cooler), the thermoelectric cooler could improve the starting characteristics and response characteristics of heat dissipation radiator, and the added thermoelectric cooler could reduce the temperature rise of North Bridge chip by 4.45%-31.12%.

1. Introduction

With rapid improvement of the information level, data center has been fully integrated into various fields like economic and social development, which changes the daily life of the citizens [1]. A data center is used to store IT equipment such as computer equipment, server equipment, storage equipment, network equipment, and communication equipment. So the data center can satisfy the special requirement like achieving centralized processing, storage, exchange, and management of data information [2]. To work without rest, data center needs to be considered more economically and safely while there are so many cooling devices and large heat dissipation.

From 1970s to the present, the heat flux density of chips in common circuits has increased from about 10 W·cm⁻² to around 500 W·cm⁻² [3]. With the increase of heat flux density, the temperature and thermal stress of electronic components will increase, making electronic components more susceptible to failure. Practice has proved that with the increase of temperature itself, the failure rate of electronic components and equipment increases exponentially. The lifetime of 50 °C is only 1/6 of that when the temperature is 25 °C. The reliability of system is reduced by 50% when the temperature of a single semiconductor component rises 10 °C [4]. The server in data center is integrated typically by high-density electronic products. Most of the electric power is converted to heat, which is dissipated to the internal space of the server, thereby raising the internal temperature of the server. Moreover, the server is generally
uninterrupted, and the heat of server is generated continuously. An efficient thermal management for the server must be considered in security.

The global data center’s electricity consumption increased from 152 billion kWh in 2005 to 238 billion kWh in 2010, accounting for about 1.3% of the total global electricity consumption [5]. In the United States, the data center’s electricity consumption in 2010 is 91 billion kWh, and it will rise to 140 billion kWh by 2020, which will cost $13 billion annually and generate 150 million tons of carbon emissions. In a data center, the amount of power used for thermal management of server accounts for 33% of total power consumption [6]. Therefore, efficient thermal management for servers are important for economics.

Based on economic and security considerations, good thermal management in the server requires a reliable cooling method and hot spot elimination capability. At present, domestic and international research on server-level and chip-level thermal management mainly includes air-cooled heat dissipation, water-cooled heat dissipation, heat pipe cooling and thermoelectric cooler (TEC) [7-9]. With the increasing heat flux density of electronic devices in the server, the heat dissipation is becoming more and more complicated. In reality, there will be more heat dissipation designs combined with various heat dissipation methods.

2. Thermal physical model

The server studied in this paper is DELL R930, which is a 4U server with wide application in the market. The three-dimensional size is 716mm×423mm×178.8mm. The main components inside the server are CPU, mainboard, Northbridge chip, power source, hard disk, RAM and fan. The parameters of the main components are showed in Table 1. Combining the parameters of each component, a lumped parameter method (using a single device as an unknown block, giving the unknown block a specific property such as material properties, thermal load, etc.) is used to establish a thermal simulation model. The temperature distribution of server is showed in figure 1.

Table 1. Parameters of main component in server.

| Devices          | Size(mm)   | Quantity | Single power (W) | Physical property                                      |
|------------------|------------|----------|------------------|-------------------------------------------------------|
| CPU              | 45×52×5    | 4        | 150              | Thermal conductivity: 12 W/(M·K); Rjc=0.298 K/Wa, Rjb=5 K/Wb |
| Northbridge chips| 25×25×5    | 1        | 30               | Thermal conductivity: 12 W/(M·K)                       |
| Mainboard        | 558×419×2.5| 1        | 0                | Surface thermal conductivity: 19.9 W/(M·K)             |
| RAM              | 70×30×1.5  | 96       | 3                | Orthogonal thermal conductivity: 0.37 W/(M·K)          |
| Hard disk        | 95×70×10   | 24       | 10               | Surface thermal conductivity: 19.9 W/(M·K)             |
| Power source     | 195×40×85  | 4        | 60               | Orthogonal thermal conductivity: 0.37 W/(M·K)          |
| Fan              | 126×126×42 | 6        | 0                | XYZ three-direction damping coefficient:110, 1000 and 1000 |
| Fan for power source | 40×40×28  | 4        | 0                | Maximum wind volume:185 CFM; Maximum wind pressure:746 Pa |
|                  |            |          |                  | Maximum wind volume:24 CFM; Maximum wind pressure:270 Pa |
It can be seen that the maximum temperature in the server is 74.11 °C at the Northbridge chip. The maximum temperature of CPU is 65.45 °C, the maximum temperature of the RAM is 36.53 °C, the maximum temperature of the hard disk is 40.09 °C, and the maximum temperature of the power source is 42.99 °C. The maximum temperature of the RAM, hard disk and power source are lower than 45 °C, which meets the thermal design goal. The hotspots are mainly concentrated in the CPU and Northbridge chips. These two areas are in need of improved thermal design to enhance heat dissipation.

3. Optimize study of CPU heat pipe radiator

The temperature distribution of CPU and heat pipe radiator are showed in figure 2. The maximum temperature is 65.45 °C at the CPU, while the lowest temperature is 32.25 °C at the fins with a difference of 33.2 °C. The average CPU temperature is 55.7 °C, while the average temperature of fins is 40.8 °C with a difference of 14.9 °C. There is a large temperature difference between CPU and fins. The heat is transferred from CPU to fins through four heat pipes. Because the thermal conductivity of heat pipe is difficult to increase, it is feasible to improve the heat exchange between fins and heat pipe.
The main effects of heat transfer between fins and heat pipe are fin thickness and fins number. As the internal space of server is compact, the structure of heat sink needs to be optimized by selecting the optimal solution among fin thickness and fins number. Combined with actual processing factors, the fin thickness in the simulation varies from 0.3mm to 0.8mm at the interval of 0.1mm, and the fins number varies from 20 to 60 at the interval of 5. The simulation results are showed in figure 3.

![Figure 3. Temperature of CPU hotspot varies with fin thickness and fins number.](image)

It can be seen from figure 3 that the lowest temperature of the CPU hot spot exist when the fins number are 30-40 pieces and the fin thickness is about 1.2mm. The data of low temperature region is selected through the fitting equation analysis, which is given by

\[ Z = \frac{p_1 + p_3x + p_5y + p_7x^2 + p_9y^2 + p_{11}xy}{1 + p_2x + p_4y + p_6x^2 + p_8y^2 + p_{10}xy} \]  

(1)

Where Z is the hot spot temperature (°C). Where x and y represent the number and thickness of fins. The value of \( p_1 - p_{11} \) are listed in Table 2.

| Factor | Value       | Factor | Value       |
|--------|-------------|--------|-------------|
| p1     | 2074188.37577837 | p7     | 8447.04536584058 |
| p2     | 6808.31467794571  | p8     | 12686.5560100466  |
| P3     | 391459.082943302  | p9     | -220305.32066387 |
| p4     | 43713.9054894401  | p10    | 1224.84621009029  |
| p5     | 16682.5774025351  | p11    | 34466.0709908196  |
| p6     | -133.185514924707 |        |              |

The genetic algorithm is used to find optimal solution of Eq.(1). The optimal fin thickness and fins number are 1.5 mm and 30 respectively. Then the server is simulated under the optimal conditions. The result shows that the hot spot temperature of CPU is 62.00 °C, which is 5.3% lower than that of original structure (65.45 °C).

The same optimization method is proceed on optimizing CPU structure under different powers. The optimized and original CPU hotspot temperature are showed in figure 4. It is obvious that when the CPU power increases, the performance of optimized heat sink is superior. For example, when the power is 165W, the hot spot temperature of optimized CPU heat pipe radiator is 66.57 °C, which is 9.09% lower than that of original structure (73.33 °C).
4. Optimize study of Northbridge chip

Figure 1 shows that the highest temperature of chip is 74.11 °C, while the lowest temperature of heat sink is 59.3 °C with a difference of 14.81 °C. The average temperature of the north bridge chip is 69.1 °C, while the average temperature of the heat sink is 60.9 °C with a difference of 8.2 °C. There is a large temperature difference between chip and heat sink, where the heat dissipation performance can be further improved. A thermoelectric cooler (TEC) could be added for optimization according to the compact space. The size of the Northbridge chip is 25mm×25mm×5mm, and the selected TEC model is TEC12706.

4.1 Experimental procedure

By comparing the cooling method with or without TEC120706, the influence of TEC current, heating power, and wind speed were analyzed. The starting and response characteristics of TEC were also explored. The test rig is showed in figure 5. The inner wall of the air duct is 800mm×120mm×120mm. The wind speed is provided by a fan (24V DC) with a dimension of 120mm×120mm×40mm. A heating module with raised portion is used to simulate the North Bridge chip. The raised portion is 25mm×25mm×5mm and heated by three heating rods. The maximum power for heating rod was 50 W.

Figure 5. Test rig.

The internal connection of heat sink and heat source in the experiment is showed in Figure 6. When there was no TEC, the heat source was directly connected to the heat sink, and the contact interfacial was coated with thermal grease to reduce the contact thermal resistance. When there was TEC, the heat source was firstly connected with the lower side of copper piece, and the upper side of copper was connected to the cold end of TEC. And the hot side of TEC was connected to the heat sink. The copper piece was used to promote the heat transfer from heat source to various points of TEC’s cold side. In this experiment, two thermocouples were placed at the heat source. One at the center of the heat source, the other at the end of the diagonal.
4.2 Experimental result

The hot spot temperature rise is the highest temperature of the heating module minus the inlet air temperature. As the inlet air temperature is constantly changing during the experiment, the performance of heat sink is evaluated by the rise of hot spot. When the wind speed is fixed at 0.5 m/s and the heating power changes from 10 W to 35 W at the interval of 5 W, the rise of hot spot temperature changes with electrical current, as shown in figure 7(a). The optimal current increases with the increase of heating power, and the optimal current of TEC is in the range of 2A-2.5A. When the experiment heating power is fixed at 30W and the experiment wind speed changes from 0.3m/s to 1.8m/s at the interval of 0.2m/s, the optimal current increases with the increase of wind speed, as showed in figure 7(b). The optimum current of TEC is in the range of 2.5A-3.2A.

To explore the influence of TEC, the temperature rise of hot spot with or without TEC was studied, as shown in figure 8. Regards to figure 8(a), the temperature rise of hot spot with TEC is 3.3~8.32°C lower than that of without TEC when the heating power changes from 10W to 35W. It shows that TEC can effectively reduce the hot spot temperature. When the heating power is fixed at 30W and the wind speed changes from 0.3m/s-1.8m/s, the temperature rise of hot spot with or without TEC is showed in figure 8(b). The temperature rise of hot spot with TEC is 0.99~5.89°C lower than that of without TEC.
When the wind speed is 0.5m/s and the heating power is 30W, the starting characteristics of two heat dissipation methods with or without TEC are showed in figure 9(a). It can be seen that the temperature rise with TEC is always lower than without TEC.

When the heating power is 30W, we firstly let the hot spot temperature rise to 75 °C, then turn on the fan to make the wind speed at 0.5m/s, or turn on the TEC with current of 2.5A at the same time. The results of thermal response characteristics are showed in figure 9(b). The hot spot with TEC can reach a lower temperatures faster.

4.3 Temperature distribution
The simulation of server was proposed to analyze the influence of TEC on the global temperature distribution. It is found that the lowest temperature of the Northbridge chip exists when the current of TEC is 2.7A, and the temperature distribution is showed in figure 7. The maximum temperature is 68.36 °C, which is 7.76% lower than the 74.11 °C for the original structure.
5. Conclusion

In this paper, a thermal simulation model based on DELL R930 4U server was established, which found that the hotspots in the server were mainly concentrated in CPU and Northbridge chips. For the hotspots in CPU, the optimization of heat pipe radiator is carried out. According to the thermal resistance network of CPU radiator, genetic algorithm was used to find the best solution of fin thickness and the fins number. The optimum fin thickness and fins number were 1.5 mm and 30 respectively. When the CPU power is 165 W, the hotspot temperature of optimized CPU radiator could be reduced by 9.09%. For the hotspots of North Bridge chip, a chip-level cooling method based on thermoelectric cooler was proposed. The optimal working current of TEC under different wind speed and powers were analyzed through experiment. The experimental result showed that the dissipation method with TEC could reach lower temperature faster than original structure. With the verification of experiment, the simulation of server showed that the temperature with TEC was 7.76% lower than that of original structure without TEC.

References

[1] Yuventi, J. and R. Mehdizadeh 2013. A critical analysis of Power Usage Effectiveness and its use in communicating data center energy consumption. Energy & Buildings. 64(64): p 90-94.
[2] Lin, M., et al 2014. Strategies for data center temperature control during a cooling system outage. Energy & Buildings. 73(2): p 146-152.
[3] Mariam, I. and F. Ahamed 2010. Thermal management of outside plant telecommunication cabinets: Design and CFD modeling methodology. Dissertations & Theses - Gradworks.
[4] Belhardj, S., et al 2003. Using microchannels to cool microprocessors: a transmission-line-matrix study. Microelectronics Journal. 34(4): p 247-253.
[5] Kheirabadi, A.C. and D. Groulx 2016. Cooling of server electronics: A design review of existing technology. Applied Thermal Engineering. 105: p 622-638.
[6] Nada, S.A., A.M.A. Attia and K.E. Elfeky 2016. Experimental study of solving thermal heterogeneity problem of data center servers. Applied Thermal Engineering. 109: p 466-474.
[7] Siedel, B., V. Sartre and F. Lefèvre 2015. Literature review: Steady-state modelling of loop heat pipes. Applied Thermal Engineering. 75: p 709-723.
[8] Chein, R. and G. Huang 2004. Thermoelectric cooler application in electronic cooling. Applied Thermal Engineering. 24(14): p 2207-2217.
[9] Yong, J.L., P.K. Singh and P.S. Lee 2015. Fluid flow and heat transfer investigations on enhanced microchannel heat sink using oblique fins with parametric study. International Journal of Heat & Mass Transfer. 81: p 325-336.