Integrated Control of Multiple Data Centre Cooling Units*

Shirin Mozaffari
Dept. of Computing and Software
McMaster University
Hamilton, Ontario, Canada
mozafs1@mcmaster.ca

Ghada Badawy
Computing Infrastructure Research Centre
McMaster University
Hamilton, Ontario, Canada
badawyg@mcmaster.ca

Douglas Down
Dept. of Computing and Software
McMaster University
Hamilton, Ontario, Canada
downd@mcmaster.ca

***Abstract*—Energy management is one of the biggest challenges in today's fast growing world of technology. While data centres play a critical role as a facility for storage and IT operations, they also consume a significant amount of energy to operate, especially for cooling. Efficient cooling of data centres involves meeting temperature constraints while minimizing power consumption. We first demonstrate the applicability of a zonal model for a modular data centre. We then leverage this zonal model to demonstrate that a joint workload assignment and cooling control problem has a simple solution: equally load the servers and maximize the setpoints of the cooling units to meet redline temperature constraints.**

***Index Terms*—data centres, cooling control, zonal model**

I. Introduction

Data centres are an integral part of today's technology. With the growing demand for data centers to meet computational needs, there is pressure to decrease data centre-related costs. As of 2018, data centres use on the order of 200 terawatt hours of electricity per year, and this amount is increasing at a significant rate [1]. Cooling accounts for a considerable fraction of the overall power consumption in data centres. In a data centre as large as 5000 square feet about 40% of power consumption is accounted for by cooling [2]. By reducing the amount of power needed to cool servers, the overall power consumption can be decreased. Efficient cooling of data centres involves meeting temperature constraints while minimizing power consumption. In this paper, we are interested in the opportunities that may be available through integrated control of multiple cooling units. Currently, in data centres with more than one cooling unit, each of the cooling units is controlled independently. In other words, cooling units do not collaborate with each other. This mode of operation results in each cooling unit needing to be set for the worst case, which results in over cooling and is not energy efficient. Coordinating cooling units has the potential to decrease the power consumption of a data centre by eliminating this over cooling. Furthermore, coordinating with workload management may help mitigate cooling unit power consumption.

There are some techniques currently used to reduce power consumption in data centres. For example, to control power consumption of servers, processors use dynamic voltage and frequency scaling (DVFS), letting them change voltage or frequency dynamically during run time [3], [4]. This is a server-level optimization which is not in the scope of this research. *Server consolidation* is another technique that assigns workload to a minimum number of servers and shuts down the rest [5]–[7]. On the other hand, *load balancing* is the technique which distributes workload uniformly among all of the servers to achieve evenly distributed power to both reduce hot spots and decrease latency [8]. A number of previous works have examined the problem of controlling multiple cooling units in a data centre, the focus of this work. Most of these works – and our work – focus on different zones within a data centre. A zone or a region of influence of a cooling unit is an area within which the temperatures of different points are the same, within some tolerance. While operating, locations closer to a cooling unit have lower temperatures than distant locations, creating a region or a zone. The shape and location of a zone are not static and vary between different architectures and cooling unit placements within a data centre [9]. Physical barriers such as wires and walls also affect the shape of zones. It is worth noting that the concept of zones in data centres and using zonal models for the aim of controlling or monitoring temperature has always been a matter of interest. This comes from the fact that using a zonal model can decompose a larger problem into smaller and potentially more tractable problems. Some research has solely focused on methodologies to define zones in a data centre. For example, Hamann et al. [10] use thermal equations to find zones in a raised floor data centre. However, in most works the manner of defining zones or the parameters that affect the zones are not discussed. Song et al. [11] is one example of such work. They focus on a zonal model to describe airflow and temperature patterns in a data center. In their work, they regard a zonal model as a thermodynamics-based intermediate approach and do not talk about the process of finding zones and defining their boundaries. Bash et al. in [12] outline a control scheme for dynamic thermal management of air-cooled data centres. In this research, each CRAC's supply air temperature is perturbed and the subsequent change at each rack's inlet is recorded. Regions of influence or zones are defined by this process, an example of the shapes of zones is shown in Figure 1. Their

This research was supported by an NSERC Collaborative Research and Development grant CRDPI506142-16 and by Cinnos Mission Critical Incorporated.

data centre also uses a raised floor structure to cool the servers.

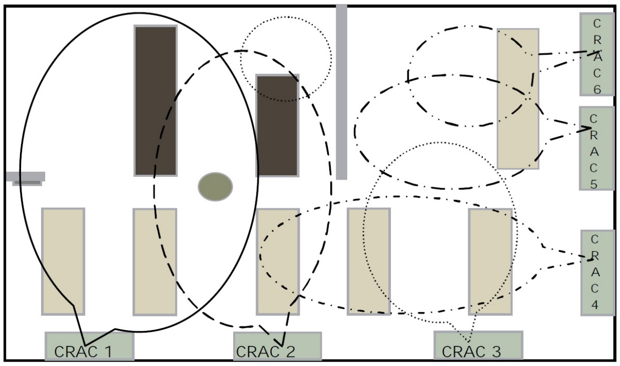

Fig. 1. Temperature at front and side of racks for an experiment [12]

In most of the previous research in this area, cooling units and servers are physically separate, for which it seems plausible to have different zones. They make use of separate cooling units located far from servers and raised floor cooling. In contrast, our data centre is a modular system in which servers and cooling units are in an enclosure and thus are not separate from each other. This in row cooling architecture is known to be more efficient than conventional cooling [13]. In the next section, we show that a zonal model is still applicable and zones can be clearly defined. In addition we explore which parameters affect the zones and to what degree. In the subsequent section we explore the implications of a zonal model for the joint control of workload and cooling.

II. Modular Data Centre Zonal Model

In this section, we study the plausibility of a zonal model for a data centre with in row cooling units. The architecture of our data centre and details of our experiments can be found in [14]. We use a data centre with five racks and two in row cooling units, one at each end. In a typical deployment the set points of the cooling units are set to equal values at the lowest possible magnitude, which results in overcooling. To investigate the effect of individual cooling units on the inlet temperature distribution within the racks, we conducted experiments where we incrementally changed the difference between setpoints of cooling units. We altered cooling unit setpoints, workload, and fan speed and monitored the changes in the resulting temperature distribution. We observed the changes in the shapes of the zones as we varied the different parameters and determined that the only factor that changes the shape of the zones is the relative values of the setpoints. We now give a short summary of our experiments. Complete details are in [14].

We chose different setpoints for the cooling units to be able to observe the potential zones. As a result of the setpoint difference, regions of influence were observed for each cooling unit. We determined that when the cooling units have different set points, based on the inlet temperatures of the servers there exist two zones, corresponding to the region of influence of each cooling unit.

The experiments were designed to consolidate this zonal hypothesis. We wanted to explore the effect, if any, of the operational parameters including workload, setpoint of each cooling unit, and fan speed of each cooling unit. We altered each of these parameters one by one and recorded the temperature reported by temperature sensors installed at the front of servers. We observed the changes in the shapes of the zones and determined that the only factor that changes the shapes of the zones is the relative value of the setpoints. One set of experiments is shown in Figure 2 (the reader is referred to [14] for additional experiments). Here all servers are equally loaded with the given utilizations. The value in each cell gives the temperatures at particular locations in front of the racks. Cells directly below a rack name are temperature values at the top, middle, and bottom of the centre of the corresponding rack. Values between rack names are for the top, middle, and bottom between the corresponding racks. Cells in blue have temperature values closer to the cooling unit with lower set point (16°C on the right), while cells in red have temperature values closer to the higher set point (24°C on the left). As can be seen, the changing utilization has only a minor effect on the shapes of the zones.

Utilization 80%

C.U 2 - 24 °C		Rack 5		Rack 4		Rack 3		Rack 2		Rack 1		C.U 1 - 16 °C
	25.13	NR	24.81	24.81	25	NR	25.19	24.06	18.38	17.75	16.81	
	24.94	24.56	23.5	23.31	24	23.5	21.31	19.69	17.63	17.06	17.25	
	22.63	23.5	26	22.69	21.25	20.31	17.5	16.63	16.56	16.44	16.63	

Utilization 50%

C.U 2 - 24 °C		Rack 5		Rack 4		Rack 3		Rack 2		Rack 1		C.U 1 - 16 °C
	25.31	24.63	25.13	25.06	25	24.5	24.63	22.63	17.75	17.63	16.81	
	25.19	25.06	23.88	23.44	23.19	23.44	21.44	19.5	17.19	17.06	17.06	
	23	22.44	25.38	21.94	19.56	18.81	17.25	16.69	16.63	16.44	16.81	

Utilization 20%

C.U 2 - 24 °C		Rack 5		Rack 4		Rack 3		Rack 2		Rack 1		C.U 1 - 16 °C
	26.06	24.38	24.88	24.81	24.63	24.31	24.19	21.88	16.94	16.75	16.06	
	24.88	24.69	23.69	23.13	22.88	22.94	20.88	18.81	16.38	16.31	16.31	
	22.8	21.88	24.38	21.19	19	18	16.5	16	15.88	15.75	16.06	

Fig. 2. Temperature zones for varying utilization

III. Integrated Control

One of the most important use cases of a zonal model is to address the problem of distributing workload. Minimizing the power consumption of cooling units while maintaining temperature constraints can be formulated as an optimization problem. The key idea in simplifying this problem is that we are considering one inlet temperature per zone rather than taking each server's inlet temperature into account. (Doing the latter would require a more detailed thermal model, which can be difficult to achieve in practice, both due to the inherent complexity of developing such thermal models and the dynamic nature of data centre environments.) In order to better clarify the notation for the optimization problem, we define a function $Z(j)$, which denotes the zone of server j (the cooling unit that is associated with the server). Because we only have two zones, $Z(j)$ only assumes the values 1 and 2. Note that the remainder of our results are for two zones, but are easily generalized to an arbitrary number of zones.

The total power consumption of a data center can be written in terms of the power consumption of cooling units plus the IT power (the amount of power that servers use), $P_{total} = P_{IT} + P_{cooling}$. If the IT power is a direct function of the entire workload (U), then it is relatively insensitive to where the workload is placed. We assume this to be the case.

The power consumed by a cooling unit is directly proportional to the power consumed by the servers (the power consumed by a server is typically a convex function of its utilization) and inversely proportional to the Coefficient of Performance. The Coefficient of Performance or CoP is the ratio of the heat removed by the cooling unit (the heat generated by the servers) to the amount of work done by the cooling unit. It should be noted that CoP of a cooling system is not constant and is an increasing function of the cooling unit setpoint (see [15] and [16]).

Below is the optimization problem. Decision variables are the utilizations of the servers (u_j) and the cooling unit setpoints ($T_{setpoint,i}$). In addition, $T_{cpu,j}$ is the CPU temperature of server j, $T_{in,j}$ is the inlet temperature of server j, T_{hot} is the maximum allowable server CPU temperature, and N is the total number of servers. We assume that all servers are identical. By solving this optimization problem, we will find the optimal setpoint temperatures and server utilizations in order to minimize the power consumption of cooling units.

$$\underset{u_j, T_{setpoint,i}}{\text{Min}} \sum_{i=1}^{2} P_i$$

subject to

$$T_{cpu,j} = f_1(u_j, T_{in,j}),\ j = 1, \ldots, N \tag{1}$$

$$T_{cpu,j} \leq T_{hot},\ j = 1, \ldots, N \tag{2}$$

$$P_i = \frac{\sum_{j:Z(j)=i} f_3(u_j)}{CoP(T_{setpoint,i})},\ i = 1, 2 \tag{3}$$

$$CoP(T_{setpoint,i}) = f_4(T_{setpoint,i}),\ i = 1, 2 \tag{4}$$

$$T_{in,j} = T_{setpoint,i} + \eta,\ Z(j) = i, \quad j = 1, \ldots, N,\ i = 1, 2 \tag{5}$$

$$n_i = f_5(T_{setpoint,1}, T_{setpoint,2}),\ i = 1, 2 \tag{6}$$

$$\sum_{j=1}^{N} u_j = NU \tag{7}$$

This problem entails keeping CPU temperature below a predefined threshold (2), 70°C is a typical value. The constraint (1) can be found in [17] and is used to calculate $T_{cpu,j}$ as a function of $T_{in,j}$ and u_j. Constraints (3) and (4) give the power consumption of the cooling units (previously discussed). The constraint (5) is the key constraint that simplifies the model, being a zonal estimate for $T_{in,j}$ in relation to $T_{setpoint,i}$. It follows the zonal model and states that a server's inlet temperature is equal to the corresponding cooling unit's setpoint plus a constant value (η), where η can depend on the application. Having said that, we will see later that our structural results for the optimization problem are independent of η. Finally, constraint (6) is a mapping from the cooling unit setpoints to the size of each zone.

With the optimization problem formulated, we can solve it for u_j and $T_{setpoint,i}$ to jointly optimize all the parameters and make a holistic conclusion about workload distribution and cooling unit control.

We will see that under mild conditions, the following simple policy is optimal.

1) Set the cooling unit setpoints to the same temperature and as high as possible.
2) Distribute the workload uniformly over all servers (servers in different zones have the same utilization).

First, we see that the zonal model simplifies the optimization problem. If $f_1(u_j, T_{in,j})$ is convex in u_j (a reasonable assumption and satisfied by all proposed models in the literature), it is not difficult to see that the utilizations of servers within a zone should be identical.

Lemma 1. *Assume that $f_1(u_j, T_{in,j})$ is convex in u_j and that the total load assigned to zone i, $\sum_{j:Z(j)=i} u_j$ is fixed. Then it is optimal for all servers in zone i to have the same utilization.*

Proof. By convexity of f_1, $\max_{j:Z(j)=i} T_{cpu,j}$ is minimized if all utilizations are equal. As the COP is increasing in $T_{setpoint,i}$, this allows $T_{setpoint,i}$ to be maximized (such that (2) is tight) and thus P_i is minimized. □

As described in the proof of the previous lemma, the setpoint of each cooling unit should also be maximized to make (2) tight. (Note that by (1) and (5), $T_{cpu,j}$ can be determined directly from $T_{setpoint,i}$). This is the key use of the zonal model.

All that remains is to show that the utilizations of the servers in the two zones should be identical. Before proceeding with the proof of that result, we note that if the power consumed by a server is convex in u_j, then the power consumed by a cooling unit is also convex in u_j (remember all u_j's in a zone are the same) and is directly proportional to n_i.

Theorem 2. *Uniform workload distribution over all servers in all zones minimizes the total power consumption of the cooling units.*

Proof. We define $P_i(u)$ as the power consumption of cooling unit i when the utilization of servers in zone i is u. We begin with equal utilizations of all servers and study deviations from this workload assignment. Without loss of generality, $u + \delta$ is the increased utilization in zone 1 and $u - \delta'$ is the decreased utilization in zone 2.

The deviation made in each zone is a function of the size of that zone. The relation between them can be shown to be $\delta'/\delta = n_1/n_2$, where as a reminder n_i is the number of servers in zone i.

Also, we have $P_i(u) = \bar{P}_i(u)$, where $\bar{P}_i(u)$ is the power consumption of cooling unit i due to a single server in its zone and the total power consumption is equal to the sum of the power consumption of both cooling units:

$$\sum_{i=1}^{2} P_i(u) = n_1\bar{P}_1(u) + n_2\bar{P}_2(u).$$

Our goal is to show the power consumption after deviating from the uniform workload distribution has a larger value than the power consumption of the uniform workload distribution:

$$P_1(u) + P_2(u) \leq P_1(u-\delta) + P_2(u+\delta'). \quad (8)$$

or

$$n_1\bar{P}_1(u) + n_2\bar{P}_2(u) \leq n_1\bar{P}_1(u-\delta) + n_2\bar{P}_2(u+\delta\frac{n_1}{n_2}). \quad (9)$$

Because the power function $P_i(u)$ is convex and increasing, (9) can be rearranged as

$$n_1[\bar{P}_1(u) - \bar{P}_1(u-\delta)] \leq n_2[\bar{P}_2(u+\delta\frac{n_1}{n_2}) - \bar{P}_2(u)]. \quad (10)$$

We can write the difference on the right hand side of (10) using the following relation:

$$\bar{P}_2(u+\frac{n_1}{n_2}\delta) - \bar{P}_2(u) = \bar{P}_2'(u)\frac{n_1}{n_2}\delta + \epsilon(u). \quad (11)$$

where $\epsilon(u)$ is the error in using the derivative at u as an approximation. Because the function $\bar{P}_2(u)$ is increasing and convex, $\bar{P}_2'(u-\delta)\cdot\delta$ is an underestimate of $\bar{P}_2(u+\frac{n_1}{n_2}\delta) - \bar{P}_2(u)$, hence the error is positive.

So, we have

$$n_1\bar{P}_1'(u)\delta - n_1\epsilon'(u) \leq n_1\bar{P}_2'(u)\delta + n_2\epsilon(u). \quad (12)$$

In (12), $\bar{P}_1'(u)\delta$ and $\bar{P}_2'(u)\delta$ have the same value, so (12) is equivalent to

$$-n_1\epsilon'(u) \leq n_2\epsilon(u). \quad (13)$$

In (13), both $\epsilon(u)$ and $\epsilon'(u)$ have positive values, therefore (13) is a tautology and is true for all u.

□

With the result of Theorem 2 in hand, we have shown that our proposed algorithm minimizes the total power consumption of the cooling units.

IV. Conclusion

We have demonstrated the validity of a thermal zonal model for modular data centres. This, combined with previous insights for enterprise data centres, suggest that using zonal models for operational optimization spans a number of data centre architectures. If one is concerned with cooling power consumption, then we show that if a zonal model is valid, under mild (and practically realistic) conditions, the problem of joint control of cooling and workload management has a simple solution: balance the load on all servers and raise the cooling unit setpoints to the maximum values that respect thermal constraints. There is some discussion of the magnitude of savings possible with this algorithm in [14], but more extensive studies are warranted.

One assumption that would be interesting to address in the future would be that the power consumed by the servers depends only on the overall offered load. Server consolidation and the ability to power down servers (or place servers in deep sleep modes) may mean that balancing the load on all servers may have greater power consumption than if the load were consolidated on fewer servers. It would be interesting to see if a form of these simple policies can be extended to this setting.

References

[1] N. Jones. How to stop data centers from gobbling up the world's electricity. *Nature*, 561(7722):163–167, 2018.

[2] J. Judge, J. Pouchet, A. Ekbote, and S. Dixit. Reducing data center energy consumption. *ASHRAE Journal*, 50(11):14, 2008.

[3] T. Chantem, X.S. Hu, and R.P. Dick. Online work maximization under a peak temperature constraint. In *Proceedings of the 2009 ACM/IEEE International Symposium on Low Power Electronics and Design*, pages 105–110. ACM, 2009.

[4] H. Hanson, S.W. Keckler, S Ghiasi, K Rajamani, F Rawson, and J Rubio. Thermal response to DVFS: Analysis with an Intel Pentium M. In *Proceedings of the 2007 international symposium on Low power electronics and design (ISLPED'07)*, pages 219–224. IEEE, 2007.

[5] D. Meisner, C.M. Sadler, L. André Barroso, W.D. Weber, and T.F. Wenisch. Power management of online data-intensive services. In *ACM SIGARCH Computer Architecture News*, volume 39, pages 319–330. ACM, 2011.

[6] M. Lin, A. Wierman, L.H. Andrew, and E. Thereska. Dynamic right-sizing for power-proportional data centers. *IEEE/ACM Transactions on Networking (TON)*, 21(5):1378–1391, 2013.

[7] A. Krioukov, P. Mohan, S. Alspaugh, L. Keys, D. Culler, and R.H. Katz. NAPSAC: Design and implementation of a power-proportional web cluster. In *Proceedings of the first ACM SIGCOMM workshop on Green networking*, pages 15–22. ACM, 2010.

[8] R.K. Sharma, C.E. Bash, C.D. Patel, R.J. Friedrich, and J.S. Chase. Balance of power: Dynamic thermal management for internet data centers. *IEEE Internet Computing*, 9(1):42–49, 2005.

[9] X. Zhang, T. Lindberg, K. Svensson, V. Vyatkin, and A. Mousavi. Power consumption modeling of data center IT room with distributed air flow. *International Journal of Modeling and Optimization*, 6(1):33–38, 2016.

[10] H.F. Hamann, V. López, and A. Stepanchuk. Thermal zones for more efficient data center energy management. In *2010 12th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pages 1–6. IEEE, 2010.

[11] Z. Song, B. T Murray, and B. Sammakia. A dynamic compact thermal model for data center analysis and control using the zonal method and artificial neural networks. 62(1):48–57, 2014.

[12] C.B. Bash, C.D. Patel, and R.K. Sharma. Dynamic thermal management of air cooled data centers. In *Thermal and Thermomechanical Phenomena in Electronics Systems, 2006. ITHERM'06. The Tenth Intersociety Conference on*, pages 8–pp. IEEE, 2006.

[13] G.C. Bell. Improving Data Center Efficiency with Rack or Row Cooling Devices: Results of "Chill-Off 2" Comparative Testing. U.S. Department of Energy, 2012.

[14] S. Mozaffari. Integrated Control of Multiple Cooling Units. M.Sc. Thesis, McMaster University, 2019.

[15] J.D. Moore, J.S. Chase, P. Ranganathan, and R.K. Sharma. Making scheduling "cool": Temperature-aware workload placement in data centers. In *USENIX annual technical conference, General Track*, pages 61–75, 2005.

[16] E. Pakbaznia and M. Pedram. Minimizing data center cooling and server power costs. In *Proceedings of the 2009 ACM/IEEE international symposium on Low power electronics and design*, pages 145–150. ACM, 2009.

[17] S.M. Mirhoseininejad, G. Badawy, and D.G. Down. EAWA: Energy-aware workload assignment in data centers. *The 2018 International Conference on High Performance Computing and Simulation*, 8(5):811–817, 2018.