Workload Management for Air-cooled Data Centers: An Energy and Exergy based Approach

Rohit Gupta [1], *Hosein Moazamigoodarzi* [3], *SeyedMorteza MirhoseiniNejad* [2,3], *Douglas G. Down* [2,3], *Ishwar K. Puri* [1,3*]

[1]Department of Mechanical Engineering, McMaster University, Hamilton, Ontario, Canada.

[2]Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada.

[3]Computing Infrastructure Research Centre, McMaster University, Hamilton, Ontario, Canada.

Note: This is a preprint version of the published article. Publisher full-text is available at doi.org/10.1016/j.energy.2020.118485

AUTHOR INFORMATION

***Corresponding author**

McMaster University

1280 Main St. W.

Hamilton, ON L8S 4L7, Canada

Email: ikpuri@mcmaster.ca

Abstract

The energy required to cool an air-cooled data center (DC) contributes significantly to the cost of operation, which is further exacerbated due to a poor choice of cooling architecture and ineffective IT workload management. Although existing algorithms reduce energy consumption, they do not minimize thermodynamic irreversibility by design. We provide a tradeoff approach that simultaneously minimizes power usage effectiveness PUE and maximizes the exergy efficiency η_{2nd}. The temperature field is predicted inside a contained single-rack DC that is equipped with a rack-mountable cooling unit (RMCU) based on a mechanical resistance model for the fluid flow. This thermal model informs a multi-objective optimization framework based on a genetic algorithm to determine the optimal decision variables and tradeoffs for PUE and η_{2nd}. We investigate the interrelated effects of (1) guidelines that ensure the reliability of the IT equipment, (2) overall network traffic load, (3) spatial IT load distribution, (4) changes in cooling system variables, and (5) multi-objective optimization. Results for the single rack system are presented in a scalable dimensionless form that is applicable for a multi-rack DC containing RMCUs. By considering the first and second laws of thermodynamics, this novel approach improves workload scheduling from both energy and exergy perspectives.

Key words: Data center – Workload management – Energy analysis – Exergy Analysis – Irreversibility – Multi-objective optimization

Nomenclature

	Uppercase letters		***Greek letters***
A	Heat transfer area (m^2)	α	Thermal mass correction factor
C_p	Specific heat capacity (J kg^{-1} K^{-1})	β	Dimensionless airflow of RMCU
D	IT load offered by the data center	β_a	Coefficient of thermal expansion for air (K^{-1})
H	Height of the rack (m)	η	Efficiency
N_F	Number of fans inside RMCU	θ	Dimensionless chilled water temperature
P	Pressure (Pa)	ρ	Density (kg m^{-3})
$\dot{P}$	Power consumption (kW)	ψ	Exergy destruction (kW)
Q	Heat load (kW)		
$\dot{Q}$	Volume flowrate (m^3 s^{-1})		***Subscripts and superscripts (lowercase)***
R	Mechanical resistance (Pa m^{-3} s)	a	Air
T	Absolute temperature (K)	amb	Ambient
T_c	Cold side temperature (K)	b	Back chamber
T_g	ASHRAE guideline temperature, 27 °C	br	Brushes
T_h	Hot side temperature (K)	cf	Condenser fans
U	Overall heat transfer coefficient (W m^{-2} K^{-1})	cr	Critical temperature in the cold aisle
V	Volume (m^3)	ch	Chiller
V	Airflow velocity of RMCU (m s^{-1})	$cool$	Overall cooling system
X	Thermal mass (J K^{-1})	cw	Chilled water
		e	Exhaust of a server
	Lowercase letters	f	Front chamber
g	Gravitational acceleration (m s^{-2})	hx	Heat exchanger
u	Utilization of a server	hw	Hot water
t	Time (s)	i	Vertical server or zone index
		in	Input to the cooling system
	Subscripts and superscripts (uppercase)	m	Exponent of the pressure-flow equation
F	Fans	s	Server
H	Horizontal Direction	w	Water
V	Vertical direction	z	Zone
IT	IT rack	$2nd$	Second law of thermodynamics

1. Introduction

Rapidly increasing computing demand has resulted in data centers (DCs) that are hyper-scale cloud facilities running at several petaFLOPS (floating point operations per second), which now account for nearly 2-2.5% of the world's electricity demand [1], making them environmentally unsustainable and fiscally expensive. From chip to the chiller, a variety of thermal management and energy minimization strategies are employed to more effectively use the cooling energy consumed by a DC [2]. At the chip level, the thermal stress that degrades the chip and its energy consumption are managed through approaches such as dynamic voltage and frequency scaling (DVFS) [3], thread mitigation [4], frequency capping [5], and dynamic power management (DPM) [6]. At the server level the most widely used strategy is to dynamically turn off unutilized servers e.g., through server consolidation [7]. At the rack level, strategies to decrease cooling energy consumption include shortened air paths [8], intelligent workload assignment to servers [9] and switching to liquid-cooled servers [10]. At the facility level, overall energy reduction is typically based on renewable energy sources and harvesting waste DC heat [11].

Workload scheduling is a key tool for reducing DC operating costs while assuring IT equipment reliability. Algorithms assign the IT load to specific servers based on predefined criteria while adhering to thermal reliability guidelines. A strategy should consider (1) cost, (2) computing efficiency, and (3) resilience. Workload scheduling and cooling system control can also use chip temperature-aware approaches [12, 13], where incoming compute loads are assigned to servers with the lowest chip temperatures, maintaining CPU temperatures below 75 °C, which also somewhat reduces the energy budget of a DC.

Another approach is to maintain a uniform temperature at the server exhaust through workload scheduling [14, 15], but this does not consider the cooling energy consumption *a priori*.

Workload can be assigned based on the server inlet temperature and the IT load of neighboring servers [9], or to minimize hot air recirculation inside the DC [16]. These approaches consider a static heat-recirculation matrix based on the pressure-flow characteristics obtained from a computational fluid dynamics (CFD) simulation, but often lead to erroneous dynamic temperature predictions. Since thermal-aware methods attempt to optimize energy through indirect strategies that reduce overcooling, they do not directly consider energy consumption in their objective functions.

Energy consumption can also be placed within the objective function of an algorithm. Energy-aware algorithms that depend on the first law of thermodynamics to minimize the overall DC energy budget [17-20] are effective in reducing cost and carbon footprint. An example is the joint cooling and workload management that simultaneously regulates the workload distribution and setpoint of the cooling equipment [21]. While the approach minimizes energy optimization, it does not address spatial overcooling since it does not account for the loss of available input cooling energy or irreversibility. Lowering the irreversibility in a DC increases its exergetic efficiency and reduces overprovisioning of the cooling capacity.

The literature on irreversibility-based assessments of DCs is sparse. Component level exergy destruction has been considered for an open aisle room-based DC with perforated tiles [22-25]. The role of the number of active racks and hot and cold air mixing on energy and exergy efficiency has been identified [26], as well as the influence of hot air recirculation and cold air bypass [27, 28]. Computationally inexpensive reduced-order exergy models are available [29], for example through flow network modeling (FNM) [30]. Although these exergy analyses consider the DC design, variations in operating parameters and architecture-based thermal management [31], it is unclear how they can be used to explicitly assign server workload.

Therefore, we propose a novel approach for allocating workload that reduces both energy consumption and thermodynamic irreversibility inside air-cooled DCs. The proof of concept is demonstrated for a single rack DC with a rack-mountable cooling unit (RMCU). Workload distribution and cooling parameters are varied simultaneously, based on which a multi-objective genetic algorithm (MOGA) minimizes both energy and irreversibility, and diminishes spatial overcooling. While existing workload scheduling algorithms use CFD and data-driven models, we employ an experimentally-validated physics-based transient zonal model to characterize the flow and temperature field [32]. This improves the predictive extrapolation accuracy to determine a coarse-grained temperature distribution, which is the basis for simultaneously optimizing workload distribution and cooling.

Here, we (1) present a methodology to couple temperature predictions with a MOGA framework that optimizes energy and exergy efficiency, (2) examine the tradeoffs among energy consumption, irreversibility, cooling characteristics, and workload distribution, (3) compare decision variables from two single-objective optimization problems that either minimize energy consumption or irreversibility, (4) identify the benefits of solving a multi-objective optimization problem for different DC IT loads, (5) show the effect of overall IT load on the scheduling algorithm, and (6) generalize the results for the single rack DC architecture to multi-rack homogeneous DCs with RMCUs.

2. Methodology

2.1. Thermal model of the rack-based cooling architecture

We utilize a low complexity, spatiotemporal zonal model [32] to predict the temperature inside an enclosed single-rack micro DC with an RMCU [33]. The model provides spatial temperature predictions of intermediate resolution, lying between a lumped-parameter thermal model and a

more complete CFD simulation. An enclosed rack with separated hot and cold aisles is considered, where the RMCU installed at the bottom of the rack supplies cold air to the IT equipment as shown in Fig. 1.

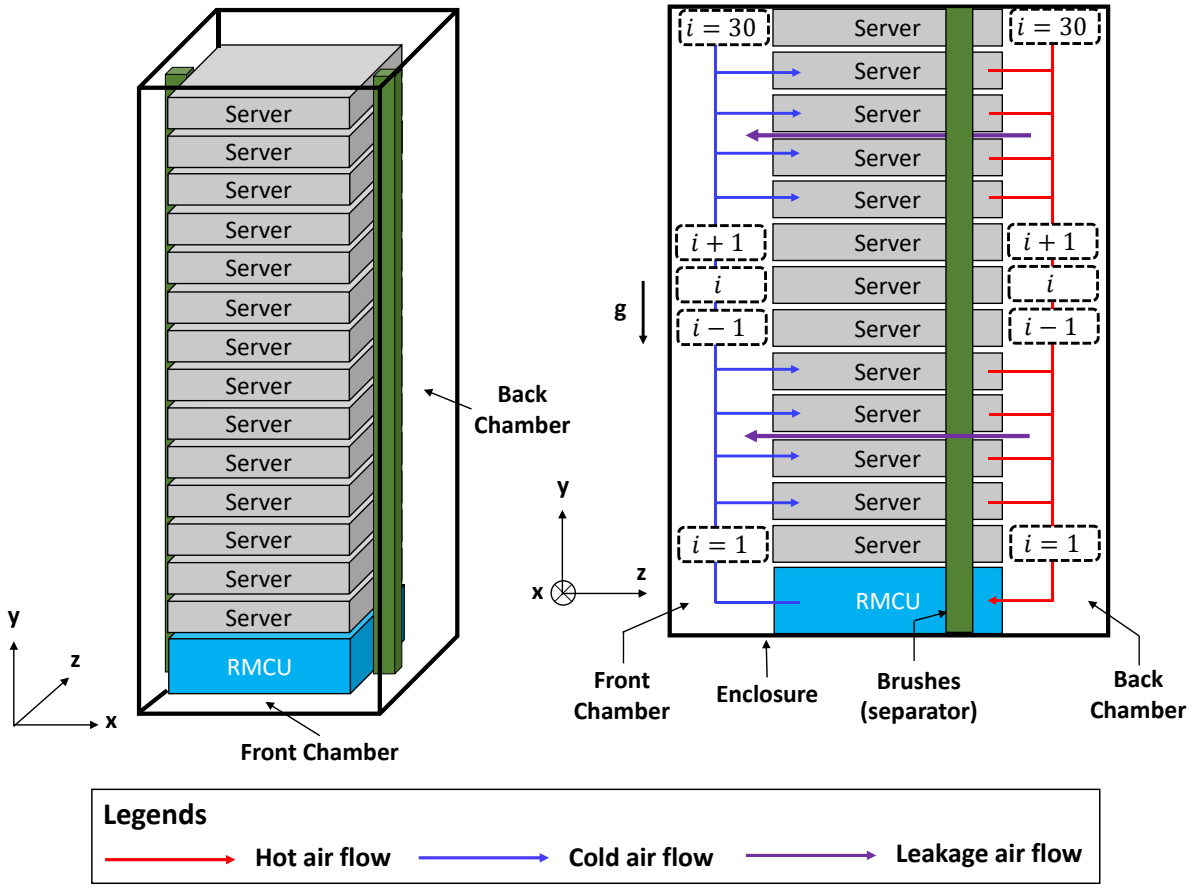

Figure 1: Schematic of the DC rack considered for the case study. The figure on the left depicts a three-dimensional view of the rack, while that on the right describes the salient airflows inside the enclosure. The IT rack consists of 30 1U (1U = 4.4 cm) servers and an RMCU.

The model predicts temperatures based on a zonal approach, where temperature, pressure and flow rates are considered uniform within each zone. The zonal approach simplifies the spatial dependence of field variables and decreases the time required to execute the algorithm as compared to that required for full-field CFD simulations. Servers are considered to be variable heat sources with magnitudes that depend on their utilizations. Neglecting momentum and energy transfer to

the ambient, six types of control volumes are identified, i.e., (1) zones lying in the fronts of servers, 2) zones at the backs of servers, 3) zones adjacent to the RMCU air supply, 4) zones adjacent to the RMCU air return, 5) each server by itself, and 6) the RMCU. The cooling system, which consists of fans and a plate-fin air-water heat exchanger, utilizes chilled water from an external circulation loop. A vapor compression refrigeration (VCR) chiller with an ambient air-cooled condenser produces the required chilled water for the RMCU. Fig. 2 is a schematic of the process flow. Temperature prediction in the architecture occur through a two-step process, where an FNM based on the mechanical resistance provides the pressure and airflow in each zone inside the enclosure for use in the energy balance equations to obtain the zonal temperatures.

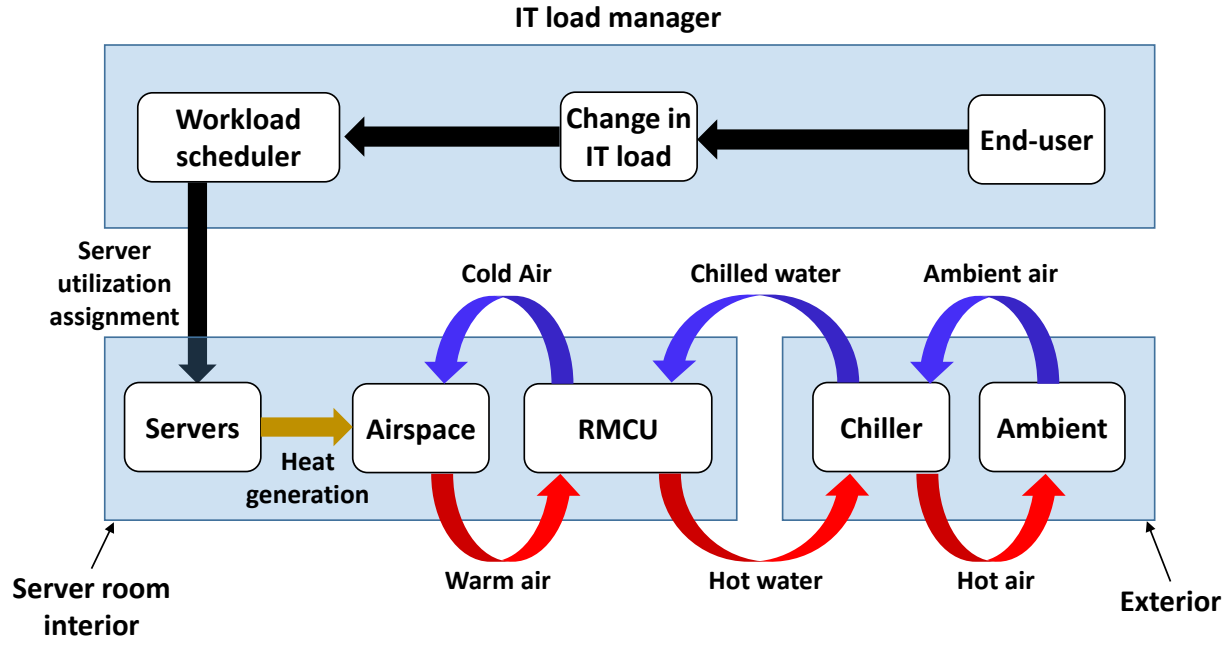

Figure 2: The process flow across different components for the DC case study. The IT rack shown in Fig. 1 is situated within the server room interior block.

2.1.1. Step 1: Determination of the flowrates using mechanical resistances

Since the flow inside an enclosed DC is pressure-driven [32], the mechanical resistance circuit illustrated in Fig. 3 is employed to calculate the flowrates and pressures. The RMCU is equivalent

to a current source that is an airflow of magnitude $\dot{Q}_a$. Each active server, either idle or utilized, is represented as a current source for an airflow $\dot{Q}_i^s$. Unpowered passive servers behave as porous separating media connecting the hot and cold chambers [32, 34] and are represented using a flow resistance R_s. The separators or brushes that prevent leakage airflow and energy transport across the chambers are represented by another resistance R_{br}. Since the principal direction of air travel inside the rack is the vertical direction, the flow resistance in the front chamber is denoted as $R_{f,V}$ and the corresponding resistance in the back chamber by $R_{b,V}$. The magnitudes of these resistances are obtained from our previous work [35]. The model assumes a large flow resistance across the IT rack and the ambient so that the air infiltration from the rack to ambient and in the reverse are minimized [32]. This is readily ensured in practice by hindering leakages through the doors that enclose IT rack by using sealing gaskets.

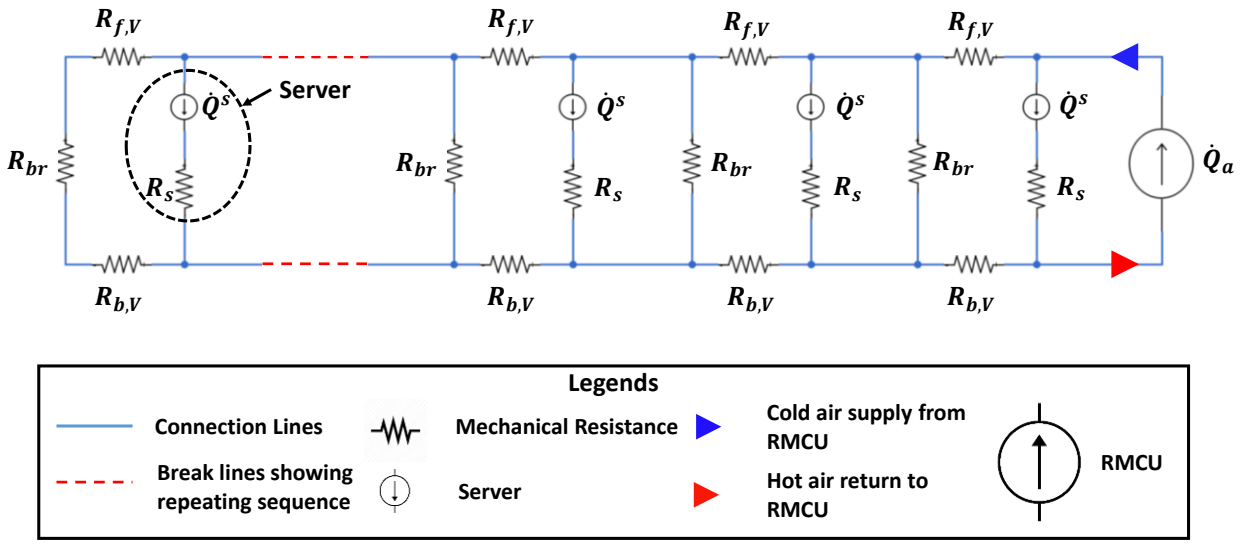

Figure 3: A flow resistance network representation for a server enclosure equipped with thirty servers and an RMCU. For simplification, the dotted red lines represent identical sequences of servers and mechanical resistances inside the enclosure.

The server flowrate is a function of the inlet air temperature of i^{th} server T_i^f [32],

$$\dot{Q}_i^s = \begin{cases} 0.01415 & \text{if } T_i^f < 25°C \\ 0.01415 + \left(T_i^f - 25\right) \times 0.00142 & \text{if } 25 < T_i^f < 35°C \end{cases}. \quad (1)$$

Once the resistances and server flow rates are known, the mass balance equation for each server and its interacting zones can be constructed. Considering the mass balance in the front of an active server,

$$\left[\frac{P_{i+1}^f - P_i^f}{R_{f,V}}\right]^m + \left[\frac{P_{i-1}^f - P_i^f}{R_{f,V}}\right]^m + \left[\frac{P_i^b - P_i^f}{R_{br}}\right]^m - \dot{Q}_i^s = 0, \quad (2)$$

where m denotes the relationship between the flow rate and pressure, P pressure, i the vertical index of a server, f the front of a server and b the back of that server.

The mass balance for the front zones of passive servers,

$$\left[\frac{P_{i+1}^f - P_i^f}{R_{f,V}}\right]^m + \left[\frac{P_{i-1}^f - P_i^f}{R_{f,V}}\right]^m + \left[\frac{P_i^b - P_i^f}{R_{br}}\right]^m + \left[\frac{P_i^b - P_i^f}{R_s}\right]^m = 0, \quad (3)$$

and for the zones at the backs of active servers,

$$\left[\frac{P_{i+1}^b - P_i^b}{R_{b,V}}\right]^m + \left[\frac{P_{i-1}^b - P_i^b}{R_{b,V}}\right]^m + \left[\frac{P_i^f - P_i^b}{R_{br}}\right]^m + \dot{Q}_i^s = 0. \quad (4)$$

The mass balance for the back zone of passive server provides the relation,

$$\left[\frac{P_{i+1}^b - P_i^b}{R_{b,V}}\right]^m + \left[\frac{P_{i-1}^b - P_i^b}{R_{b,V}}\right]^m + \left[\frac{P_i^f - P_i^b}{R_{br}}\right]^m + \left[\frac{P_i^f - P_i^b}{R_s}\right]^m = 0, \quad (5)$$

and for zones in the front of the RMCU in the cold chamber,

$$\left[\frac{P_{i+1}^f - P_i^f}{R_{f,V}}\right]^m = \dot{Q}_a. \quad (6)$$

Utilizing the mass balance for zones in front of the RMCU in the hot chamber,

$$\left[\frac{P_{i+1}^{b} - P_{i}^{b}}{R_{b,V}}\right]^{m} = \dot{Q}_{a}. \quad (7)$$

Appling Eqns. $(1) - (7)$ with $m = 1$ [35], leads to a linear relationship between the flow and pressure for each zone and results in a system of linear algebraic equations that provide the pressure corresponding to each zone.

2.1.2. Step 2: Zonal energy balance

Once the airflows and pressure for each zone are determined by solving the above set of linear equations, the energy balance formulations provide the zonal temperatures. The transient zonal energy balance equation for an active server is [32, 36],

$$\frac{X_s}{2}\left(\frac{dT_i^e}{dt} + \frac{dT_i^f}{dt}\right) = \rho_a C_{p,a} \dot{Q}_i^s \left(T_i^f - T_i^e\right) + \dot{P}_i^s, \quad (8)$$

where T_i^e denotes the exhaust temperature of the i^{th} server, T_i^f the front zone temperature of the i^{th} server, ρ_a the density of air, $C_{p,a}$ the specific heat of air, X_s the thermal mass of a server obtained from the literature [37], and $\dot{P}_i^s$ the power consumption of the i^{th} server, which depends on its utilization [38-40].

An energy balance for the waterside within the RMCU results in [32],

$$\rho_w C_{p,w} V_w \left(\frac{dT_{c,w}}{dt} + \frac{dT_{h,w}}{dt}\right) = \rho_w \dot{Q}_w C_{p,w}\left(T_{c,w} - T_{h,w}\right) + \frac{UA}{2}\left(T_{h,a} + T_{c,a} - T_{c,w} - T_{h,w}\right), \quad (9)$$

and for the airside within the RMCU,

$$\rho_a C_{p,a} V_a \left(\frac{dT_{c,a}}{dt} + \frac{dT_{h,a}}{dt}\right) = \rho_a \dot{Q}_a C_{p,a}\left(T_{h,a} - T_{c,a}\right) - \frac{UA}{2}\left(T_{h,a} + T_{c,a} - T_{c,w} - T_{h,w}\right), \quad (10)$$

where $T_{c,a}$ and $T_{h,a}$ denote the cold side and hot side air temperatures of the heat exchanger inside the RMCU, $\dot{Q}_a$ and $\dot{Q}_w$ the air and water flow respectively, $T_{h,w}$ and $T_{c,w}$ the hot water outlet and chilled water inlet temperature respectively, ρ_w the density of water, $C_{p,w}$ the specific heat of

water, U the overall heat transfer coefficient of the heat exchanger as a function of $\dot{Q}_w$ and $\dot{Q}_a$ [32], A the contact surface area responsible for heat transfer, and V_w and V_a the volumes of the water and air inside the heat exchanger respectively.

Applying the energy balance to the cold and hot chamber zones inside the enclosure,

$$\rho_a C_{p,a} V_z \alpha \left(\frac{dT_i}{dt}\right) = \varphi_1 + \varphi_2 + \varphi_3 + \varphi_4 + \varphi_5, \tag{11}$$

where α denotes the thermal mass correction factor, V_z the volume of a zone, and T_i the temperature of a zone. The expressions for φ_1 through φ_5 are reported in Table 1.

Table 1: Expressions for φ_1 through φ_5 in Eq. (11)

Zones in front of the servers		Zones at the back of the servers	
Φ_1 (Flow transport in front chamber)		Φ_1 (Flow transport in back chamber)	
$\left[P_{i+1}^f - P_i^f\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_{i+1}^f - P_i^f}{R_{f,V}}\right] T_{i+1}^f$	$\left[P_{i+1}^b - P_i^b\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_{i+1}^b - P_i^b}{R_{b,V}}\right] T_{i+1}^b$
$\left[P_{i+1}^f - P_i^f\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_{i+1}^f - P_i^f}{R_{f,V}}\right] T_i^f$	$\left[P_{i+1}^b - P_i^b\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_{i+1}^b - P_i^b}{R_{b,V}}\right] T_i^b$
Φ_2 (Flow transport in front chamber)		Φ_2 (Flow transport in back chamber)	
$\left[P_{i-1}^f - P_i^f\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_{i-1}^f - P_i^f}{R_{f,V}}\right] T_{i-1}^f$	$\left[P_{i-1}^b - P_i^b\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_{i-1}^b - P_i^b}{R_{b,V}}\right] T_{i-1}^b$
$\left[P_{i-1}^f - P_i^f\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_{i-1}^f - P_i^f}{R_{f,V}}\right] T_i^f$	$\left[P_{i-1}^b - P_i^b\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_{i-1}^b - P_i^b}{R_{b,V}}\right] T_i^b$
Φ_3 (Flow transport through brushes in front chamber)		Φ_3 (Flow transport through brushes in back chamber)	
$\left[P_i^b - P_i^f\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_i^b - P_i^f}{R_{br}}\right] T_i^b$	$\left[P_i^f - P_i^b\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_i^f - P_i^b}{R_{br}}\right] T_i^f$
$\left[P_i^b - P_i^f\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_i^b - P_i^f}{R_{br}}\right] T_i^f$	$\left[P_i^f - P_i^b\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_i^f - P_i^b}{R_{br}}\right] T_i^b$
Φ_4 (Flow through passive servers)		Φ_4 (Flow through passive servers)	
$\left[P_i^b - P_i^f\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_i^b - P_i^f}{R_s}\right] T_i^b$	$\left[P_i^f - P_i^b\right] \geq 0$	$\rho_a C_{p,a}\left[\frac{P_i^f - P_i^b}{R_s}\right] T_i^f$
$\left[P_i^b - P_i^f\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_i^b - P_i^f}{R_s}\right] T_i^f$	$\left[P_i^f - P_i^b\right] < 0$	$\rho_a C_{p,a}\left[\frac{P_i^f - P_i^b}{R_s}\right] T_i^b$
Φ_5 (Flow through active servers)		Φ_5 (Flow through active servers)	
$-\rho_a C_{p,a} \dot{Q}_i^s T_i^f$		$\rho_a C_{p,a} \dot{Q}_i^s T_i^e$	

2.2. Energy consumption and power usage effectiveness PUE

The server power consumption is a major contributor towards the overall energy consumed by a DC. It consists of the energy consumed by the (1) CPUs and (2) server fans, where the second makes a small 0.95% contribution towards the overall power consumed by a server. Changing the server inlet temperature from 15 °C to 27 °C has been shown to increase the energy consumption of a server by only 0.48% [41]. Therefore, the power consumed by a server is solely a function of its utilization, i.e., the magnitude to which it is stressed [38-40]. The total IT power consumption,

$$\dot{P}_{IT} = \sum_{i=1}^{N} \dot{P}_i^s = \sum_{j=1}^{N} [C_1 + C_2 u_i^s], \tag{12}$$

where $\dot{P}_{IT}$ represents the overall power consumption of the servers, u_i^s and $\dot{P}_i^s$ the utilization level and power consumption of a server with the vertical index i, and C_1 and C_2 are constants that depend on the type and model of server. We assume homogeneous server types within a DC rack, so that C_1 and C_2 are constant throughout. For an HP ProLiant DL360 G5 server, their values are obtained from a datasheet $C_1 = 223.41$ and $C_2 = 154.5$ [42]. For an idle server, $u_i^s = 0$, $\dot{P}_i^s = C_1$ is the static power. For a passive server, $\dot{P}_i^s = 0$.

The overall power consumption in the cooling system,

$$\dot{P}_{cool} = \dot{P}_{ch} + \dot{P}_F, \tag{13}$$

where, $\dot{P}_{ch}$ and $\dot{P}_F$ denote the power consumed by the VCR chiller and the fans inside the RMCU, respectively. These two components of power consumption are calculated using Eqns. (16) and (17) below.

$$y = A_1 x_1 + A_2 x_2 + A_3 x_3, \tag{14}$$

$$y = \left[\left(1 + \frac{1}{COP}\right)\frac{T_{c,w}}{T_{amb}}\right], \qquad x_1 = \frac{T_{c,w}}{Q_{ch}}, \qquad x_2 = \frac{T_{amb} - T_{c,w}}{Q_{ch}T_{amb}}, \qquad x_3 = \frac{\left(1 + \frac{1}{COP}\right)Q_{ch}}{T_{amb}}. \tag{15}$$

$$\dot{P}_{ch} = \frac{\left[Q_{ch} + A_1 T_{c,w} T_{amb} + A_2\left(T_{amb} - T_{c,w}\right)\right]}{T_{c,w} - (A_3 Q_{ch})} - Q_{ch}. \tag{16}$$

Equations (14) and (15) are from the Ng-Gordon universal chiller model [43], where $T_{c,w}$ represents the chilled water temperature leaving the evaporator, T_{amb} the ambient air temperature entering the condenser, Q_{ch} the heat load on the chiller, i.e., the IT load, and COP the coefficient of performance of the chiller based on the cooling load per unit power consumption. In this model, all temperatures are in K and Q_{ch} is in kW. The constants A_1 through A_3 are determined from the performance of a commercially available chiller [44], where multivariate linear regression results in $A_1 = 0.0092, A_2 = 8.953$, and $A_3 = 0.649$. The value of T_{amb} is set to 297.039 K as per the chiller performance data. Eq. (16) describes the power consumption of the chiller by combining Eqns. (14) and (15).

The total power consumption of the fans inside the RMCU,

$$\dot{P}_F = N_F\left[B_1 + B_2\left(\frac{\dot{Q}_a}{N_F}\right)\right], \tag{17}$$

where N_F denotes the number of fans and $\dot{Q}_a$ the total airflow through RMCU. For our case, $N_F = 5$. The constants B_1 and B_2 are determined through experiments for the counter rotating axial fan San Ace 80 9CRB [45], resulting in $B_1 = -0.268$ and $B_2 = 4.12 \times 10^{-3}$. While fan models can be based on linear, parabolic and cubic relations between the flowrate and power consumption, the linear relationship between airflow and energy consumption holds for our operational range $350 \leq \dot{Q}_a \leq 850$ CFM.

The IT and cooling power consumption are combined together using the widely accepted energy-based metric, power usage effectiveness PUE [46-49],

$$PUE = 1 + \frac{\dot{P}_{cool}}{\dot{P}_{IT}}, \tag{18}$$

where $1 \leq PUE < \infty$. The PUE must be minimized, where the ideal, but unrealistic, value $PUE = 1$ implies that the DC consumes none of the total energy consumption for cooling purposes.

2.3. Exergy destruction and the 2nd law of efficiency η_{2nd}

The components of overall exergy destruction are the exergy losses (1) $\dot{\psi}_F$ in fans due to work transfer, (2) $\dot{\psi}_{hx}$ due to heat transfer inside the RMCU, (3) $\dot{\psi}_{ch}$ in the chiller, and (4) $\dot{\psi}_s$ in the servers due to heat dissipation. Here, $\dot{\psi}_s$ depends on the IT load imposed on the rack and end user computing demand, which the DC administrator has no control over. Hence, overall exergy destruction in the optimization framework considers only the first three components that represent cooling cycle inefficiency [31],

$$\dot{\psi}_{cool} = \dot{\psi}_F + \dot{\psi}_{ch} + \dot{\psi}_{hx}. \tag{19}$$

The exergy loss associated with the fan is solely a function of airflow [30], i.e.,

$$\dot{\psi}_F = (1 - \eta_F) N_F \left[B_1 + B_2 \left(\frac{\dot{Q}_a}{N_F} \right) \right], \tag{20}$$

where the efficiency of the fan, $\eta_F = 0.7$, is obtained from a datasheet [45].

The exergy loss due to heat transfer inside the RMCU is determined using the relation [31],

$$\dot{\psi}_{hx} = \rho_a \dot{Q}_a C_{p,a} \left[(T_{h,a} - T_{c,a}) - T_{amb} \ln \left(\frac{T_{h,a}}{T_{c,a}} \right) \right] + \rho_w \dot{Q}_w C_{pw} \left[(T_{c,w} - T_{h,w}) - T_{amb} \ln \left(\frac{T_{c,w}}{T_{h,w}} \right) \right], \tag{21}$$

where $T_{h,a}$ and $T_{c,a}$ denote the average hot air return temperature and cold air supply temperature of the RMCU, $T_{c,w}$ and $T_{h,w}$ are the chilled water supply and hot water return temperatures to the RMCU.

The exergy destruction in the chiller [31],

$$\dot{\psi}_{ch} = \rho_w \dot{Q}_w C_{p,w} T_{amb} \ln\left(\frac{T_{c,w}}{T_{h,w}}\right) + \rho_a \dot{Q}_{cf} C_{p,a} T_{amb} \ln\left(\frac{T_{h,amb}}{T_{amb}}\right). \quad (22)$$

where $T_{h,amb}$ denotes the warm air temperature at the exhaust of the VCR chiller condenser, T_{amb} the ambient air temperature supplied to the condenser, $\dot{Q}_w$ the total volume flow rate through the chilled water loop, and $\dot{Q}_{cf}$ the airflow of the condenser fans. The value of $\dot{Q}_{cf}$ is obtained from the datasheet for the chiller [44].

The second law efficiency η_{2nd} is the exergy efficiency, which is the extent of irreversibility in the system as a fraction of total input exergy, i.e.,

$$\eta_{2nd} = \left(1 - \frac{\dot{\psi}_{cool}}{\dot{\psi}_{in}}\right) \times 100. \quad (23)$$

The exergy input $\dot{\psi}_{in}$ to the cooling system [31],

$$\dot{\psi}_{in} = \dot{P}_{ch} + \dot{P}_F + \rho_a \dot{Q}_a C_{p,a} \left[\left(T_{h,a} - T_{amb}\right) - T_{amb} \ln\left(\frac{T_{h,a}}{T_{amb}}\right)\right]. \quad (24)$$

From a thermodynamic perspective, system irreversibilities must be reduced so that as much of the available input energy can be used as is possible. This can be monitored by comparing η_{2nd} for different scenarios.

2.4. Optimization problem formulation

Using the models in Sections 2.2 and 2.3, our objectives are to (1) minimize PUE and (2) maximize η_{2nd} for the specific DC configuration considered. To examine the interplay between these two objectives, we consider three different optimization problems that are described in Table 2, where (1) the first problem minimizes PUE, (2) the second maximizes η_{2nd}, and (3) the third simultaneously minimizes PUE and maximizes η_{2nd}. The primary decision factors that regulate PUE and η_{2nd} are (1) $T_{c,w}$, (2) $\dot{Q}_a$, and (3) the server utilizations u_i^s, values for which are

optimized. Two classes of servers are considered here, (1) active servers with $0 < u_i^s \leq 1$ and (2) idle servers with $u_i^s = 0$. Turning OFF an idle server, i.e., making it a passive server to save energy changes the thermal and optimization framework significantly [21, 32, 35]. We maintain the maximum server inlet temperature below the ASHRAE guideline temperature $T_g = 27$ ℃ [50].

Table 2: Mathematical structure of three different optimization problems

Optimization 1	**Optimization 2**	**Optimization 3**
$\underset{u_i^s, T_{cw}, \dot{Q}_a}{\text{minimize}}\ PUE = 1 + \frac{\dot{P}_{cool}}{\dot{P}_{IT}}$	$\underset{u_i^s, T_{cw}, \dot{Q}_a}{\text{maximize}}\ \eta_{2nd} = \left(1 - \frac{\dot{\psi}_{cool}}{\dot{\psi}_{in}}\right)$	$\underset{u_i^s, T_{cw}, \dot{Q}_a}{\text{minimize}}\ PUE$ and $\underset{u_i^s, T_{cw}, \dot{Q}_{cu}}{\text{maximize}}\ \eta_{2nd}$
s.t. $\sum_{i=1}^{n} u_i^s = D$	s.t. $\sum_{i=1}^{n} u_i^s = D$	s.t. $\sum_{i=1}^{n} u_i^s = D$
$max(T_i^f) \leq T_g$	$max(T_i^f) \leq T_g$	$max(T_i^f) \leq T_g$
$10 \leq T_{c,w}(℃) \leq 22℃$	$10 \leq T_{c,w}(℃) \leq 22℃$	$10 \leq T_{c,w}(℃) \leq 22℃$
$350 \leq \dot{Q}_a(CFM) \leq 850$	$350 \leq \dot{Q}_a(CFM) \leq 850$	$350 \leq \dot{Q}_a(CFM) \leq 850$
$0 \leq u_i^s \leq 1,\ i = 1, \ldots, n$	$0 \leq u_i^s \leq 1,\ i = 1, \ldots, n$	$0 \leq u_i^s \leq 1,\ i = 1, \ldots, n$

Optimization problems 1 and 2 are single objective and multi-dimensional with both linear and nonlinear constraints, whereas the third optimization problem is multi-objective. The primary constraint is a linear constraint that signifies that the sum of all u_i^s in the DC should equal the total offered workload D. The secondary constraint is non-linear, which maintains the ASHRAE thermal reliability guideline temperature of 27°C within the cold chamber. Evaluating the maximum temperature in the cold aisle requires use of the coupled FNM and thermal model of Sections 2.1.1 and 2.1.2. The first two problems are solved using the MATLAB *fmincon* function with a *sequential quadratic programming* solver, whereas a genetic algorithm-based multi-objective solver *gamultiobj* is used to solve the third problem. Both functions support linear and nonlinear constraints. The lower and upper bounds for each decision variable (see Table 2) are obtained from the component datasheets of the RMCU and the VCR chiller [33, 44]. The convergence criterion for the objective functions is set to 10^{-6} for all three optimization problems.

3. Results and discussion

3.1. Temperature nonuniformity in the cold aisle

In an air-cooled DC, altering the cooling parameters such as $T_{c,w}$, $\dot{Q}_a$, and $\dot{Q}_w$ significantly changes the (1) flow-field, (2) temperature distribution, (3) energy consumption, and (4) system irreversibility, which lead to tradeoffs across different decision variables. Therefore, we examine the influence of these parameters on the cold aisle temperature distribution. Table 3 lists different cases for a specific workload distribution inside the contained single-rack DC. The two primary decision variables $T_{c,w}$ and $\dot{Q}_a$ influence the energy consumptions of the chiller and fans, respectively. The effect of altering the water flow is neglected because the pumping power is only 2-3% of the overall cooling power [8, 47, 51, 52]. For instance, changing the water flow by ~10% alters the air temperature by 3% [32, 35]. The model considers a fixed speed pump with an adequate valve mechanism-based flowrate control that delivers $\dot{Q}_w = 20$ GPM. As an alternative, a variable speed pump with a variable frequency drive can be used to control the energy consumption of the pump by regulating $\dot{Q}_w$.

Table 3: Case study parameters for the results presented in Fig. 4, where variables $T_{c,w}$ and $\dot{Q}_a$ are changed to constrain the maximum temperature in the cold aisle below 27 ℃.

Case	Chilled water temperature $T_{c,w}$ (℃)	Airflow of RMCU $\dot{Q}_a$ (CFM)	Workload assignment with $D = 15$	Maximum allowable temperature in the cold chamber (℃)	Variance in temperature distribution	Water flowrate (GPM)
1	10	578	Each server has 50% utilization	27	11.6	20
2	12	609			8.6	
3	14	643			6.1	
4	16	682			3.9	
5	18	726			2.2	
6	20	776			1.0	
7	22	834			0.2	

Fig. 4 presents the cold aisle temperature distributions corresponding to the cases listed in Table 3, where $T_{c,w}$ and $\dot{Q}_a$ are varied to constrain the maximum temperature in the cold aisle below 27 ℃. For a fixed workload distribution with all servers at 50% utilization, increasing both $T_{c,w}$ and $\dot{Q}_a$, increases the temperature uniformity in the cold aisle as reflected by the temperature variance reported in Table 3.

The relative magnitudes of natural and forced convection in the DC are characterized through the Richardson number $Ri = g\beta_a\left(T^f_{max} - T^f_{min}\right)H/V_c^2$ [53]. Natural convection is negligible compared to forced convection when $Ri \leq 0.1$. For the DC, $0.004 \leq Ri \leq 0.07$, implying that buoyancy can be neglected for the pressure-driven flow with forced convection.

The maximum air temperature in the cold chamber is observed in front of the server located furthest from the RMCU. This observation is consistent with the FNM depicted in Fig. 3. The flow resistance in the front and back chambers at the i^{th} zone in the chamber is a function of its distance from the RMCU. An increase in the airflow path produces a higher flow resistance and increases the pressure drop [8], which decreases the transport of cold air that proceeds towards the top of the enclosed rack. Therefore, the server located furthest from the RMCU now has the highest intake air temperature. Comparison of cases 1 through 7, shows that the extent of overcooling decreases when $T_{c,w}$ and $\dot{Q}_a$ are simultaneously increased.

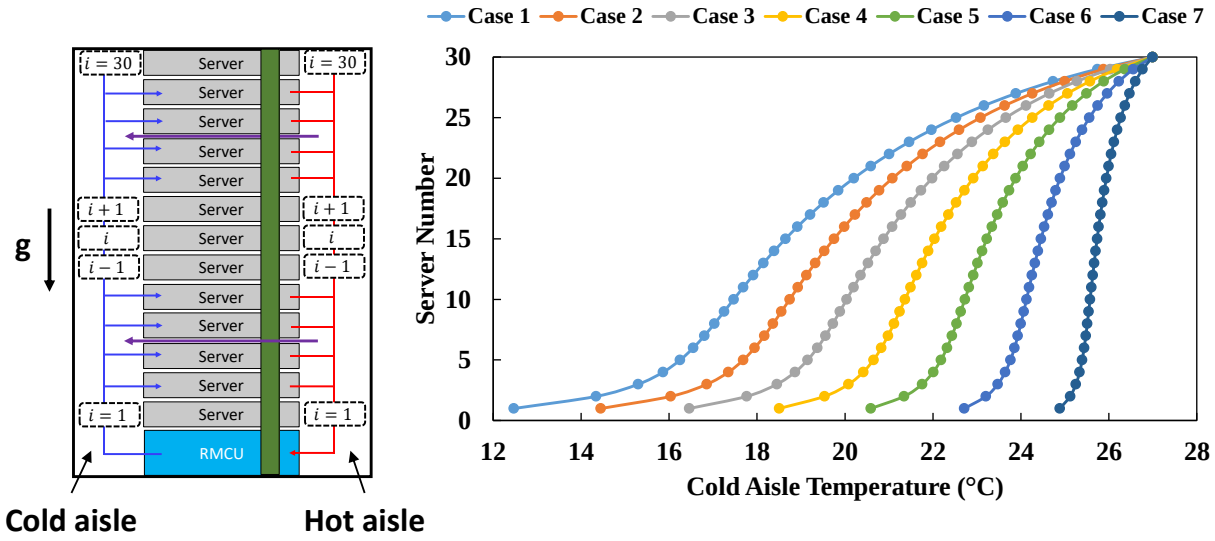

Figure 4: Nonuniformity in the cold aisle temperature due to changes in T_{cw} and $\dot{Q}_a$ for a fixed workload distribution with all servers at 50% utilization for the cases listed in Table 3.

The impact of redistributing IT load is also investigated, where Table 4 lists three cases that are considered. The overall offered workload, D is set to 15 which can be achieved through numerous possible combinations of server utilizations. A fixed speed pump delivers $\dot{Q}_w =$ 20 GPM and with $T_{c,w}$ fixed at 16 ℃.

Table 4: Case study parameters for the results presented in Fig. 5, where $\dot{Q}_a$ and workload distribution are changed to constrain the maximum temperature in the cold aisle below 27 ℃.

Case	Chilled water temperature $T_{c,w}$ (℃)	Airflow of RMCU $\dot{Q}_a$ (CFM)	Workload distribution with $D = 15$	Maximum allowable temperature in the cold chamber (℃)	Variance in temperature distribution	Water flowrate (GPM)
1	16	605	Workload assigned towards RMCU ($u^s_{i=1\to15} = 1$ and $u^s_{i=16\to30} = 0$)	27	3.0	20
2	16	682	Uniform workload assignment ($u^s_{i=1\to30} = 0.5$)		3.9	
3	16	727	Workload assigned far from RMCU ($u^s_{i=1\to15} = 0$ and $u^s_{i=16\to30} = 1$)		4.4	

For the three IT load distributions, $\dot{Q}_a$ is regulated to always maintain the maximum temperature in the cold aisle below 27 ℃ to ensure safe IT equipment operation. Figure 5 presents cold temperature distributions for three workload assignments. Server utilization for these cases are (1) $u^s_{i=1\to15} = 1$ and $u^s_{i=16\to30} = 0$, (2) $u^s_{i=1\to30} = 0.5$, and (3) $u^s_{i=1\to15} = 0$ and $u^s_{i=16\to30} = 1$. In practice, the workload distribution (or server utilization) is controlled through workload manager software. Assigning the IT workload towards the RMCU (Case 1) results in a more uniform temperature distribution for even lower $\dot{Q}_a$. Displacing the workload to servers that lie further away from the RMCU (Case 3) results in a higher variance (Table 4) in the temperature distribution across the rack and increases $\dot{Q}_a$. Consequently, the power consumption of the fans for Case 3 also increases. Therefore, displacing the workload towards the RMCU offers the potential to reduce overcooling and lower fan power consumption.

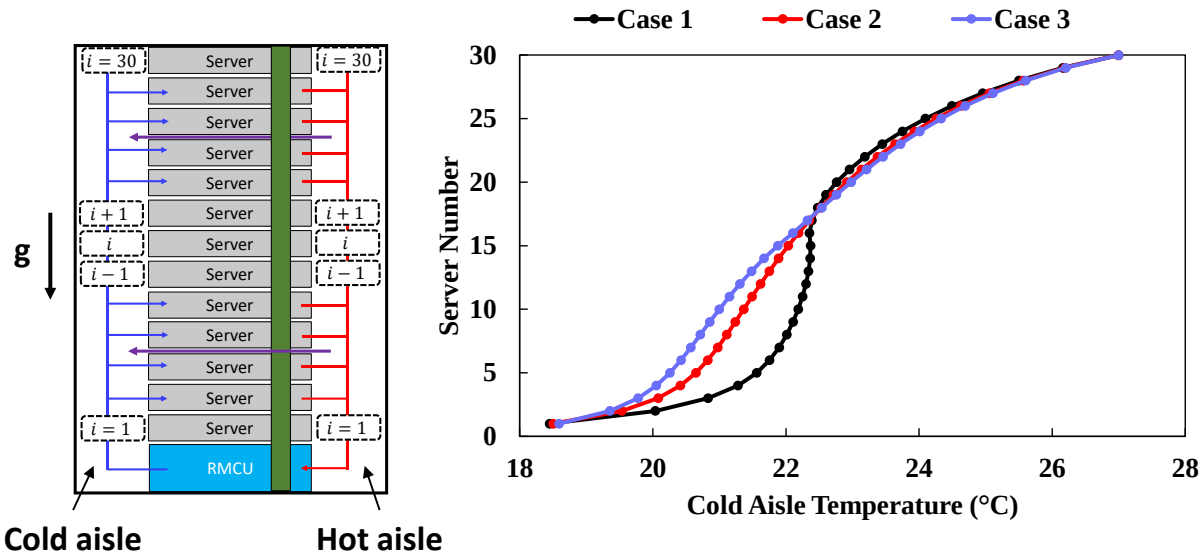

Figure 5: Non-uniformity in the cold aisle temperature due to simultaneous change in cooling parameter and workload distribution. The figure is to be interpreted in accordance with Table 4.

3.2. Impact on PUE and η_{2nd} of changing $T_{c,w}$, $\dot{Q}_a$, and workload distribution

To investigate the extent of overcooling when $T_{c,w}$, $\dot{Q}_a$, and the IT load distribution are changed, we consider two dimensionless metrics. The first PUE, defined in Eq. (18), is the overall power consumption as a fraction of the IT equipment power consumption. The exergy efficiency of the cooling system, based on the second law of thermodynamics, is expressed by Eq. (23) and accounts for the loss of available cooling energy due to inherent system irreversibilities.

We define a non-dimensional number $\theta \times \beta$ where,

$$\theta = \frac{T_{c,w}}{T_{amb}}, \qquad \beta = \frac{\dot{Q}_a}{\dot{Q}_{IT}} = \frac{\dot{Q}_a}{125 \times \dot{P}_{IT}}. \tag{25}$$

Here, θ is the ratio of chilled water temperature to the ambient dead state temperature and β compares the RMCU airflow $\dot{Q}_a$ with the total airflow requirement for all servers $\dot{Q}_{IT}$. A large value of θ lowers the chiller power consumption and decreases heat transfer irreversibility between the cold water and the ambient dead state. Although, the total airflow through the servers as a function of inlet air temperature is provided by Eq. (1), we utilize a standard DC guideline that prescribes 125 CFM airflow to be provided for each kW IT load increment [54], i.e., $\dot{Q}_{IT} = 125 \times \dot{P}_{IT}$, to ensure the thermal reliability of servers. A larger β results in higher fan power consumption and lowers system irreversibility by diminishing the temperature gradients in the airspace. The impact of changing $\theta \times \beta$ on the PUE and η_{2nd} for different maximum allowable air temperatures in the cold aisle, i.e., 26, 27, and 28 ℃ is illustrated in Fig. 6. For all cases, the total workload assignment $D = \sum_{i=1}^{30} u_i$ is set to 15, where each of the 30 servers inside the rack is 50% utilized.

Examination of Fig. 6, shows that increasing $\theta \times \beta$ increases PUE and also improves η_{2nd}. If $T_{c,w}$ is increased, $\dot{Q}_a$ must also be simultaneously increased to satisfy the maximum

allowable cold aisle temperature T_{cr}. As a result, the chiller power consumption decreases and that of the fans increases. Since fan operation is costlier than chiller operation for modular DCs [44, 45], increasing $\theta \times \beta$ increases PUE. On the other hand, from a second law perspective, increasing θ reduces the heat transfer irreversibility in the chilled water stream and higher β reduces the irreversibility due to spatial air temperature gradients. Therefore, with an increase in $\theta \times \beta$ the η_{2nd} improves. Hence, larger values of η_{2nd} are desirable for minimizing the system irreversibility but this increases PUE, increasing the cooling power consumption, which is undesirable.

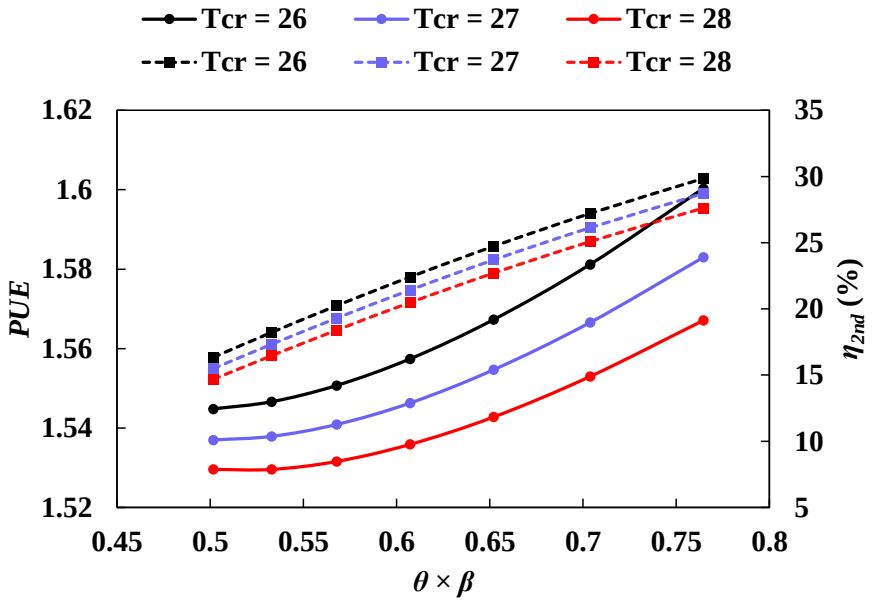

Figure 6: Effect of cooling system operating parameters on PUE and η_{2nd} for different maximum allowable temperatures in the cold aisle. The solid lines with circles represent variations in PUE, whereas the dashed lines with square symbols represent changing η_{2nd}.

Fig. 7 presents the tradeoffs in PUE and η_{2nd} as a function of $\theta \times \beta$ for different workload distributions when the maximum air temperature in the cold aisle is maintained below 27 ℃. The total offered workload $D = \sum_{i=1}^{30} u_i = 15$. Increasing $\theta \times \beta$ increases both η_{2nd} and PUE. This can be addressed by changing the workload distribution, where three cases are considered, i.e., (1) the workload is displaced towards the RMCU ($u^s_{i=1\rightarrow 15} = 1$ and $u^s_{i=16\rightarrow 30} = 0$), (2) there is

uniform workload distribution with 50% utilization ($u^s_{i=1\to30} = 0.5$), and (3) the workload is displaced away from the RMCU ($u^s_{i=1\to15} = 0$ and $u^s_{i=16\to30} = 1$). The lowest PUE is obtained for the first case when the workload is displaced towards the RMCU, but this also leads to the lowest η_{2nd} since system irreversibility increases. When the workload is displaced away from the RMCU for Case 3, PUE increases the by ~5% compared with the Case 1 since the flow path resistance between the heat source and cooling unit increases. This increases the pressure drop which resists the incoming cold air from the RMCU, a resistance that increases as $\theta \times \beta$ is increased. As a result, Case 3 alters the temperature distribution inside the DC slightly, leading to a small increase in PUE. Increasing $\theta \times \beta$ for this case reduces irreversibility within the DC airspace and chilled water stream, improving η_{2nd}. The tradeoff between increasing PUE and decreasing η_{2nd} implies that minimization of power consumption and irreversibility cannot be achieved simultaneously for all operational cases. Therefore, we solve three nonlinear optimization problems in the subsequent sections that (1) minimize PUE, (2) maximize η_{2nd}, and (3) simultaneously minimize PUE and maximize η_{2nd} through a tradeoff.

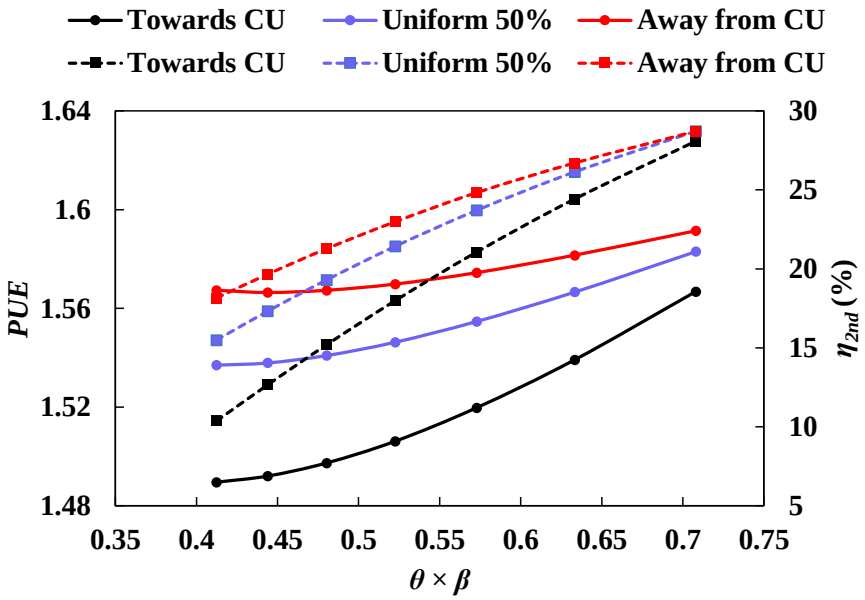

Figure 7: Effect of cooling system operation parameters on PUE and η_{2nd} for different workload distributions. The solid lines with circles represent variations of PUE whereas dashed lines with square symbols represent changes in η_{2nd}.

3.3. Energy optimization by minimizing PUE

The first optimization problem from Table 2 is the PUE minimization problem, which intrinsically minimizes cooling energy consumption. A thermal reliability constraint is imposed to limit the maximum temperature in the cold chamber to below 27 ℃. Table 5 shows results for different D following the methodology presented in Section 2.4. Optimal values of PUE and the cooling decision variables $T_{c,w}$ and $\dot{Q}_a$ are also presented in Table 5. Fig. 8 represents the optimal workload assignments and resulting cold aisle temperature distributions across the rack.

Table 5: Optimal solutions for the different PUE minimization cases presented in Fig. 8.

Case	Total offered workload (D)	Airflow of RMCU $\dot{Q}_a$ (CFM)	Chilled water temperature $T_{c,w}$ (℃)	Minimized PUE	Variance in cold aisle temperature distribution	Corresponding value of η_{2nd}
1	6.0	498	10	1.57	11.5	11.1
2	10.5	493	10	1.52	10.0	10.8
3	15.0	488	10	1.49	8.1	10.4
4	19.5	483	10	1.46	7.0	10.0
5	24.0	550	13.7	1.44	5.0	14.3

As D increases, the additional servers that lie towards the top of the rack become utilized. The maximum achievable computing load $D_{max} = 30$ for this case. For the smallest value investigated, $D = 6$, servers adjacent to the RMCU are utilized since allocating workload closest to the cooling unit results in the lowest flow path resistance. This improves advective transport of the cold air in the front chamber, reducing overcooling of air by simultaneously regulating $T_{c,w}$ and $\dot{Q}_a$. Consequently, PUE is lowered across a wide range of D values. Since fan operation in the RMCU is more expensive than operating the chiller, low values of $T_{c,w}$ and $\dot{Q}_a$ are favored to minimize overall energy consumption. An increase in D improves PUE since the load factor,

$L.F. = D/D_{max}$ increases. As $L.F.$ increases the coefficient of performance (COP) of the chiller reaches its designed nominal capacity and the energy efficiency of the cooling system improves. A higher $L.F.$ also increases the uniformity in the cold aisle, increasing the mean air temperature. This increase in the average temperature results in a lower value for $T_{amb} - T^{f}_{mean}$, reducing the system irreversibility due to heat transfer. Therefore, for the highest D, η_{2nd} improves by ~ 4%, as shown in Table 5. Since the goal is to minimize PUE, the solution does not optimize exergy efficiency. The resulting tradeoffs in PUE and η_{2nd} are illustrated in Figs. 6 and 7.

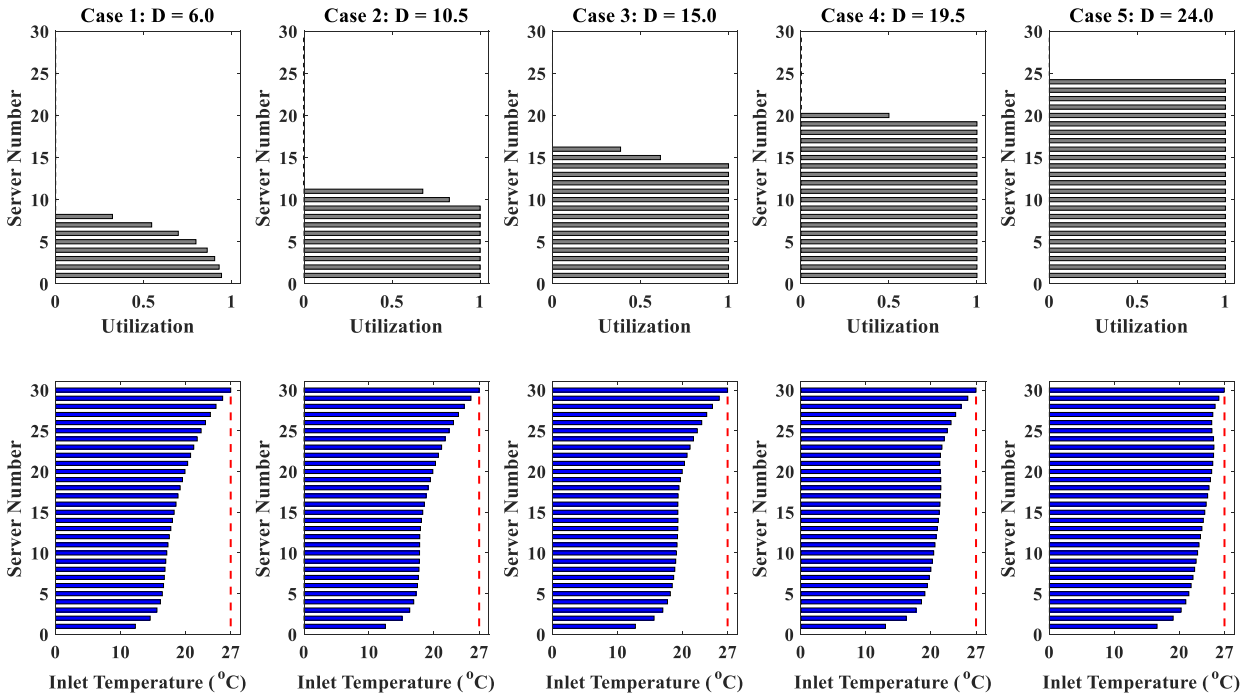

Figure 8: PUE minimized workload assignment and resultant cold aisle temperature distributions across the rack for different offered workloads D. The red dotted line corresponds to the maximum allowable server inlet air temperature, 27 ℃. The figure should be interpreted in accordance with Table 5.

3.4. Exergy optimization by maximizing η_{2nd}

The second optimization problem (see Table 2) is the η_{2nd} maximization problem, which reduces the system irreversibility. A nonlinear thermal reliability constraint is imposed that limits the

maximum temperature in the cold chamber to below 27 ℃. The optimization problem is solved for the different D listed in Table 6 using the methodology in Section 2.4. The optimal values of η_{2nd} and cooling decision variables $T_{c,w}$ and $\dot{Q}_a$, are also provided in Table 6. Fig. 9 represents the optimal workload assignment obtained by solving the maximization problem and resulting cold aisle temperature distribution across the rack.

Table 6: Optimized solutions for different η_{2nd} maximization cases presented in Fig. 9.

Case	Total offered workload D	Airflow of RMCU $\dot{Q}_a$ (CFM)	Chilled water temperature $T_{c,w}$ (℃)	Maximized η_{2nd}	Variance in cold aisle temperature distribution	Corresponding value of PUE
1	6.0	850	22	30.05	0.066	1.68
2	10.5	850	22	29.92	0.057	1.63
3	15.0	850	22	29.75	0.048	1.59
4	19.5	850	22	29.50	0.043	1.55
5	24.0	850	22	29.13	0.052	1.52

Similar to the PUE minimization problem, η_{2nd} maximization also assigns workload closest to the cooling unit for the wide range of D values. However, optimal values of the cooling decision variables $T_{c,w}$ and $\dot{Q}_a$ are different from those obtained from the PUE minimization problem. Maximizing η_{2nd} increases $T_{c,w}$ and $\dot{Q}_a$ which reduces the (1) exergy destruction for the chilled water stream and (2) results in a uniform temperature distribution, as shown in Fig. 9. Therefore, two major heat transfer irreversibilities in the system i.e., $T_{amb} - T_{mean}^{f}$ and $T_{amb} - T_{c,w}$ are minimized. By minimizing the $T_{amb} - T_{mean}^{f}$, the irreversibility due to thermal gradients in the airspace is lowered. On the other hand, minimizing $T_{amb} - T_{c,w}$ reduces the irreversibility associated with the chilled water stream.

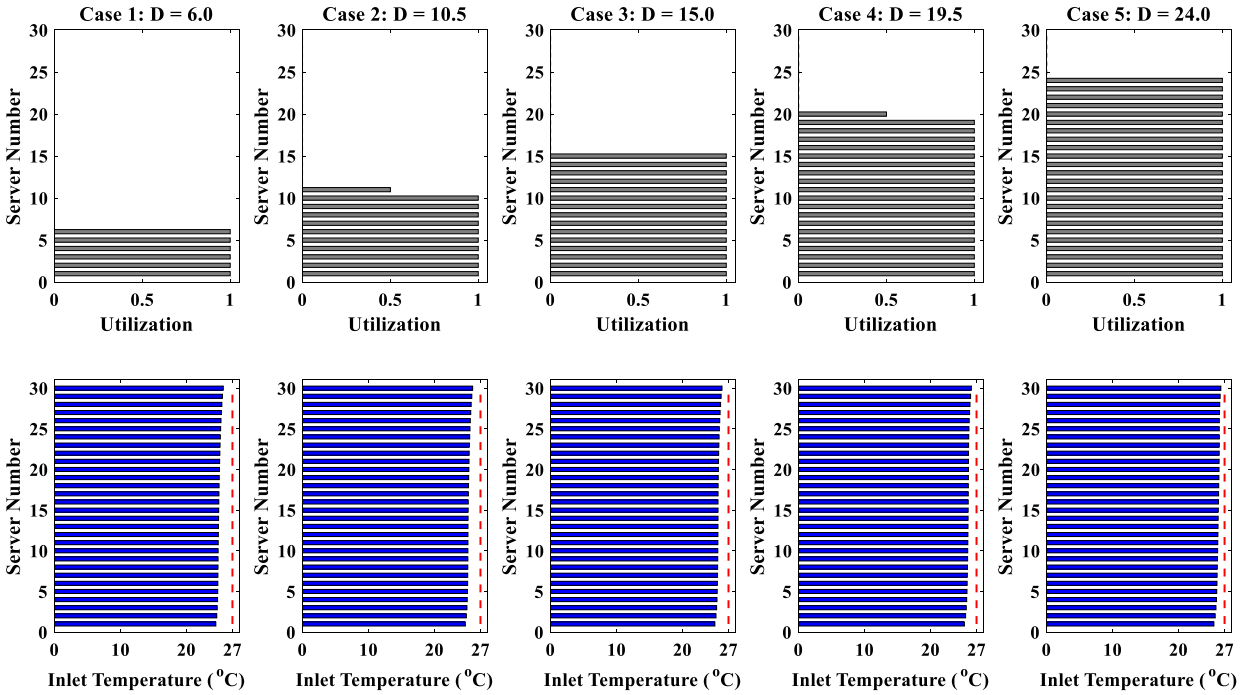

Figure 9: η_{2nd} maximized workload assignment and resulting cold aisle temperature distributions across the rack for different offered workloads D. The red dotted line corresponds to the maximum allowable server inlet air temperature, 27 ℃. The figure should be interpreted in accordance with Table 6.

3.5. Multi-objective optimization

Fig. 10 presents results from two different single objective optimization problem as a function of $L.F.$ for the DC. By minimizing PUE, the exergy efficiency of the system decreases by approximately 19%, which leads to a significant loss of available input cooling energy. To achieve maximum exergy efficiency, PUE increases by 7%, leading to higher energy consumption. The value of η_{2nd} for the PUE minimization problem increases by ≈ 4% for $L.F.$ = 0.8, as shown by the black dashed line in Fig. 10. This improvement is attributed to the irreversibility decrease in the chilled water stream and a smaller spatial air temperature gradient shown in Table 5.

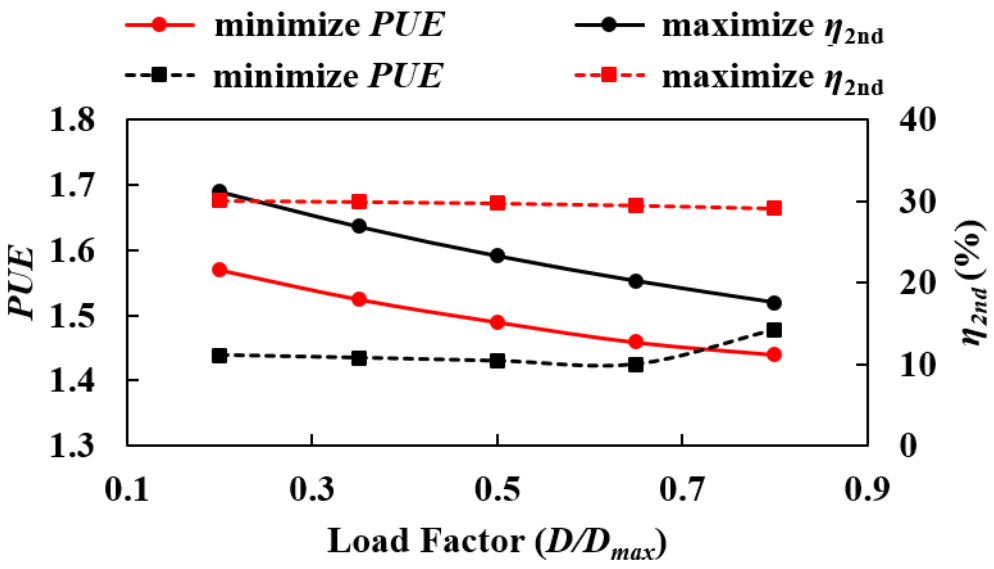

Figure 10: PUE and η_{2nd} as a function of $L.F.$ for PUE minimization and η_{2nd} maximization problems. Solid lines with circles represent variations of PUE, whereas dashed lines with square symbols represent those for η_{2nd}. The optimization details for the figure are provided in Sections 3.3 and 3.4.

To address the tradeoff between PUE and η_{2nd}, a multi-objective optimization problem subject to the same set of constraints is considered. The structure of the third optimization problem is provided in Table 2. The solution is obtained using a genetic algorithm-based multi-objective solver where PUE is minimized and η_{2nd} is maximized simultaneously. The multi-objective problem is solved for the $L.F.$ values specified in Tables 5 and 6. Fig. 11 shows the Pareto optimal front of PUE *vs.* η_{2nd} obtained by solving the multi-objective optimization problem for five different $L.F.$ values, 0.2, 0.35, 0.5, 0.65, and 0.8. Generally, the trends suggest that for all $L.F.$ η_{2nd} improves at the cost of PUE. For a higher $L.F.$ (marked as desired region in Fig. 11), PUE decreases while η_{2nd} improves so that increasing the network traffic load is a promising method to reduce the overcooling of a DC rack.

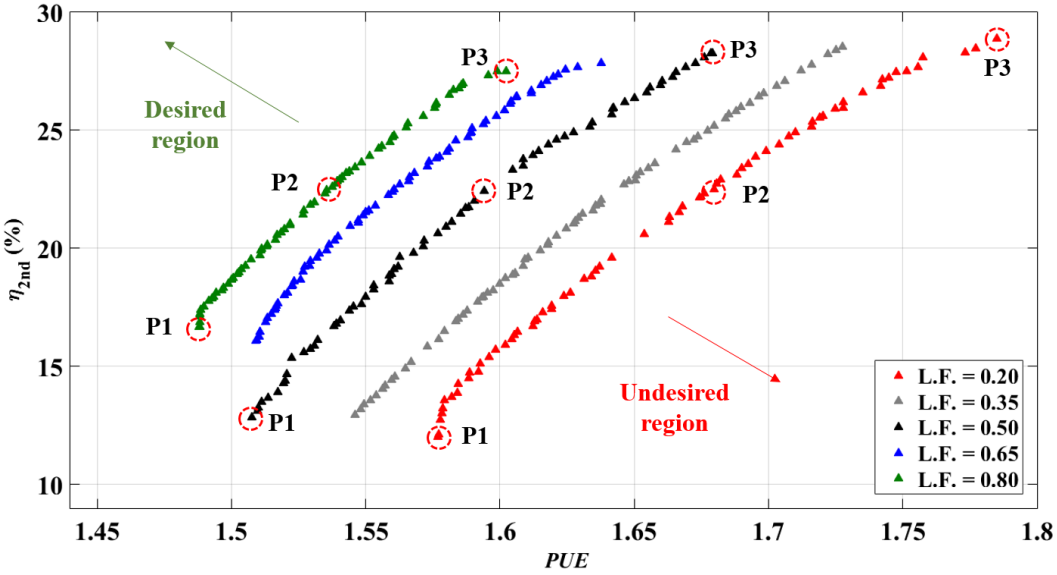

Figure 11: Pareto front of PUE and η_{2nd} obtained by solving the multi-objective optimization problem for five different $L.F.$ imposed on the DC. The symbols in the figure should be interpreted in accordance with Table 7 and Fig. 12.

To gain better insight into the thermal dynamics and workload distribution in the Pareto front, nine points are marked (using red dotted circles) on the $L.F.$ curves of 0.2, 0.5, and 0.8. in Fig. 11. The corresponding cooling parameters, workload assignments, and temperature distributions are shown in Fig. 12 and Table 7. For different $L.F.$, optimizations 1 and 2 result in a utilization distribution that places the workloads closest to the RMCU (Figs. 8 and 9). In contrast, the workload distribution obtained from optimization 3 (Fig. 12) distributes the IT load across the rack in a nonuniform manner. Fig. 8 shows that optimization 1 overcools several servers in the vicinity of the RMCU. Therefore, distributing servers with high and low utilizations across the rack and thus regulating the cooling parameters should result in the desired tradeoff across the two objectives. By comparing three points for the same $L.F.$ from Fig. 12, we infer that moving across the Pareto front for a specific $L.F.$ does not influence the spatial workload distribution. The tradeoff in the Pareto front is a result of the synchronized regulation of $\dot{Q}_a$ and $T_{c,w}$ which alters the extent of reduction in overcooling and changes the variance in the cold aisle temperature

distribution (Table 7). Moving from left to right on the Pareto front for each $L.F.$ curve results in a lower variance in temperature, which consequently improves η_{2nd}. However, such an improvement leads to additional expenditure from a PUE perspective.

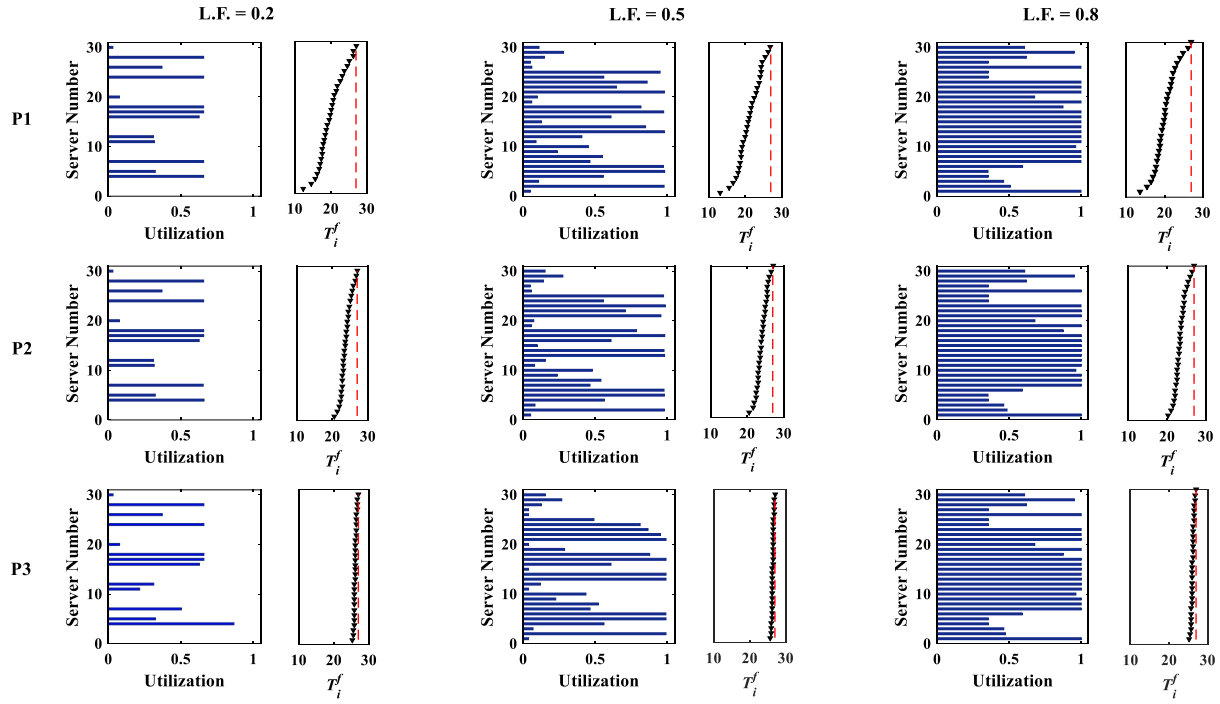

Figure 12: Utilization and temperature distributions for PUE and η_{2nd} aware multi-objective workload assignments. Nine salient points for the different $L.F.$ marked in Fig. 11 are considered for this diagram. The red dotted line corresponds to the maximum allowable server inlet temperature of 27 ℃. The figure should be interpreted in accordance with Table 7.

By comparing point P2 marked on different $L.F.$ curves in Fig. 12 increasing $L.F.$ does not significantly change the temperature distribution but significantly alters the workload assignment. Points P1 and P3 correspond to extrema where PUE is minimized and η_{2nd} is maximized, respectively. Optimizations 1 and 2 concentrate the IT load near the RMCU for all values of $L.F.$ In contrast, the third optimization problem distributes the workload across the rack in a nonuniform manner so that a specific vertical temperature gradient is maintained while accounting for the tradeoff for two different objectives. Obtaining such a Pareto optimal front for a wide range of

$L.F.$ values provides an operational regime diagram (Fig. 11) and helps realize potential savings across the objectives.

Table 7: Salient points for $L.F. = 0.2$, 0.5 and 0.8 obtained from Fig. 11

Load factor	Point	Airflow of RMCU $\dot{Q}_a$ (CFM)	Chilled water temperature $T_{c,w}$ (℃)	η_{2nd}	PUE	Variance in cold aisle temperature distribution
0.2	1	515	10.1	12.0	1.58	12.8
	2	685	18.5	21.8	1.67	2.4
	3	839	21.5	28.5	1.78	0.1
0.5	1	531	10.6	12.9	1.51	10.3
	2	689	18.0	21.5	1.58	2.2
	3	850	22.0	28.3	1.68	0.1
0.8	1	611	10.8	16.9	1.49	8.9
	2	725	17.1	22.4	1.53	2.4
	3	850	22.0	27.5	1.60	0.1

3.6. Normalization of the Pareto front

To generalize the results for a single rack DC and demonstrate their applicability for multi-rack systems, the solutions obtained from multi-objective optimization are normalized. The baseline results obtained by solving single-objective problems at different $L.F.$ are described through Fig. 10. By fitting trend lines to each curve, the minima and maxima for PUE and η_{2nd} are obtained as functions of $0 \leq L.F. \leq 1$. The empirically obtained equations are as follows.

$$[PUE]_{min} = 1.64 - 0.39(L.F.) + 0.18(L.F.)^2, \tag{26}$$

$$[PUE]_{max} = 1.77 - 0.43(L.F.) + 0..15(L.F.)^2, \tag{27}$$

$$[\eta_{2nd}]_{min} = 5.27 + 53.41(L.F.) - 146.66(L.F.)^2 + 117.28(L.F.)^3, \text{ and} \tag{28}$$

$$[\eta_{2nd}]_{max} = 30.43 - 1.52(L.F.). \tag{29}$$

Therefore, the horizontal axis of the dimensionless Pareto front can be represented as,

$$PUE^* = \frac{[PUE]_i - [PUE]_{min}}{[PUE]_{max} - [PUE]_{min}}, \tag{30}$$

and the vertical axis is normalized as,

$$\eta_{2nd}^* = \frac{[\eta_{2nd}]_i - [\eta_{2nd}]_{min}}{[\eta_{2nd}]_{max} - [\eta_{2nd}]_{min}}, \tag{31}$$

where $[PUE]_i$ and $[\eta_{2nd}]_i$ represents the i^{th} point of the Pareto front.

Fig. 13 represents the dimensionless variation of PUE *vs.* η_{2nd} where, the effect of $L.F.$ is normalized using Eqns. $(26) - (31)$. There is a proportionality tradeoff across PUE^* and η_{2nd}^* where maximizing η_{2nd}^* leads to a detrimental effect on PUE^*. Thus, an intermediate value lying on the dimensionless Pareto front should be maintained during operation as per choice of DC administrator. The proposed normalization approach enhances the applicability of the results obtained for a single-rack system for a generalized multi-rack homogeneous DC providing better control over cooling-aware workload scheduling while ensuring the thermal reliability of the system.

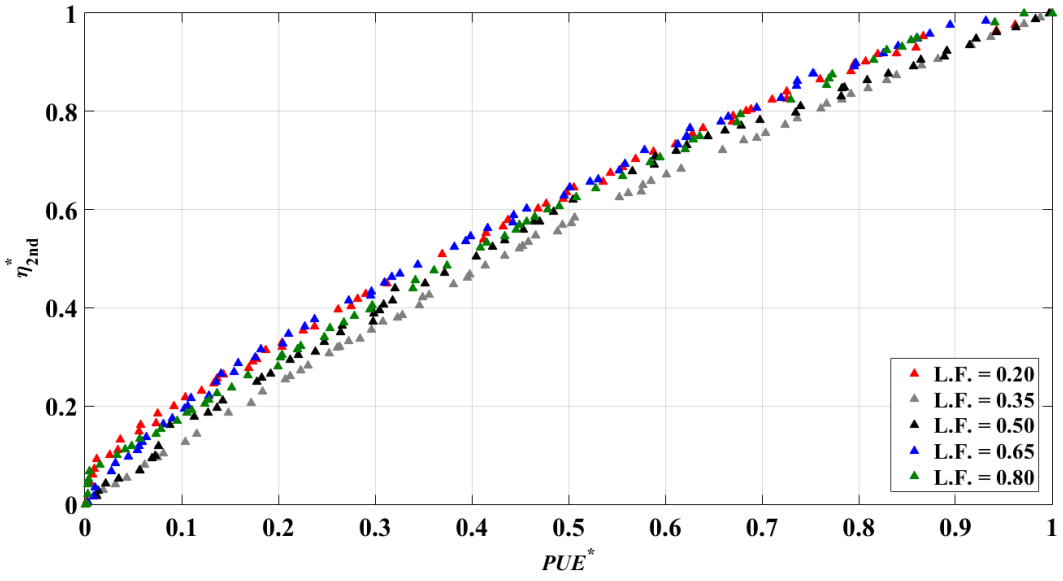

Figure 13: Dimensionless Pareto front of PUE and η_{2nd} obtained by adopting the proposed normalization procedure. The normalization procedure nullifies the effect of $L.F.$

4. Conclusion

We have developed a novel approach to perform energy and exergy aware workload scheduling for air-cooled DCs with RMCUs. The methodology combines (1) a low complexity zonal temperature prediction model, (2) energy and exergy formulation, and (3) an optimization solver resulting in a facile tool for DC workload assignment. In contrast to existing thermal- and energy-aware IT load approaches the irreversibility-based method opens up avenues for future research. The approach simultaneously regulates salient decision variables of the cooling system and the spatial workload distribution.

Important aspects of the investigation include elucidation of the impact on the algorithm of, (1) altering cooling system variables and workload distribution on the energy and exergy efficiencies, (2) additional benefits obtained by considering the exergy efficiency, (3) tradeoffs across the three optimization problems, and (4) the influence of the thermal reliability constraint, IT load factor and workload distribution. Salient findings are:

1. Simultaneously increasing $T_{c,w}$ and $\dot{Q}_a$ while maintaining $T_g \leq 27℃$ increases the temperature uniformity and mean air temperature in the cold aisle.
2. Populating the workload closer to the cooling unit lowers the value of $\dot{Q}_a$ and reduces the power consumption of the fans.
3. An increment in the dimensionless cooling parameter $\theta \times \beta$ improves the exergy efficiency at the cost of increasing the PUE. Therefore, the two desired values of the objectives cannot be achieved simultaneously.
4. The PUE minimization problem does not necessarily minimize overcooling of the servers in the vicinity of the RMCU. On the other hand, the exergy optimization problem lowers the extent of overcooling by increasing both $T_{c,w}$ and $\dot{Q}_a$.

5. The η_{2nd} maximization problem results in an exergy efficiency of approximately 30% with a ~7% sacrifice on the optimal PUE.
6. The multi-objective optimization problem distributes the workload across the IT rack, whereas the single-objective optimization problem assigns workload closest to the cooling unit.
7. The dimensionless Pareto front enhances the applicability of the results obtained for a single rack modular DC towards a multi-rack system.

5. Acknowledgment

The authors (RG, HM, and SM) gratefully acknowledge research assistantship support provided by McMaster University and financial support from the National Sciences and Engineering Research Council (NSERC). The first author gratefully acknowledges the effort of Ms. Sreya Sarkar of the University of Illinois Chicago for assisting in proofreading.

6. References

[1] W. Van Heddeghem, S. Lambert, B. Lannoo, D. Colle, M. Pickavet, and P. Demeester, "Trends in worldwide ICT electricity consumption from 2007 to 2012," *Computer Communications*, vol. 50, pp. 64-76, 2014.

[2] S. V. Garimella, T. Persoons, J. Weibel, and L.-T. Yeh, "Technological drivers in data centers and telecom systems: Multiscale thermal, electrical, and energy management," *Applied energy*, vol. 107, pp. 66-80, 2013.

[3] S. Herbert and D. Marculescu, "Analysis of dynamic voltage/frequency scaling in chip-multiprocessors," in *Proceedings of the 2007 international symposium on Low power electronics and design (ISLPED'07)*, 2007, pp. 38-43.

[4] J. C. Smolens, R. T. Golla, and M. B. Smittle, "Dynamic mitigation of thread hogs on a threaded processor," ed: Google Patents, 2013.

[5] J.-G. Park, C.-Y. Hsieh, N. Dutt, and S.-S. Lim, "Co-Cap: energy-efficient cooperative CPU-GPU frequency capping for mobile games," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016, pp. 1717-1723.

[6] S.-T. Kung, C.-C. Cheng, C.-C. Liu, and Y.-C. Chen, "Dynamic power saving by monitoring CPU utilization," ed: Google Patents, 2003.

[7] P. Padala, X. Zhu, Z. Wang, S. Singhal, and K. G. Shin, "Performance evaluation of virtualization technologies for server consolidation," *HP Labs Tec. Report*, vol. 137, 2007.

[8] H. Moazamigoodarzi, P. J. Tsai, S. Pal, S. Ghosh, and I. K. Puri, "Influence of cooling architecture on data center power consumption," *Energy*, vol. 183, pp. 525-535, 2019.
[9] 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*, 2005, pp. 61-75.
[10] S. Zimmermann, I. Meijer, M. K. Tiwari, S. Paredes, B. Michel, and D. Poulikakos, "Aquasar: A hot water cooled data center with direct energy reuse," *Energy*, vol. 43, pp. 237-245, 2012.
[11] K. Ebrahimi, G. F. Jones, and A. S. Fleischer, "A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 622-638, 2014.
[12] S. Li, T. Abdelzaher, and M. Yuan, "Tapa: Temperature aware power allocation in data center with map-reduce," in *2011 International Green Computing Conference and Workshops*, 2011, pp. 1-8.
[13] Y. Bai and L. Gu, "Chip temperature-based workload allocation for holistic power minimization in air-cooled data center," *Energies*, vol. 10, p. 2123, 2017.
[14] 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*, vol. 9, pp. 42-49, 2005.
[15] M. T. Chaudhry, T. Ling, S. A. Hussain, and A. Manzoor, "Minimizing thermal stress for data center servers through thermal-aware relocation," *The Scientific World Journal*, vol. 2014, 2014.
[16] Q. Tang, S. K. S. Gupta, and G. Varsamopoulos, "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, pp. 1458-1472, 2008.
[17] S. M. M. Nejad, G. Badawy, and D. G. Down, "Eawa: Energy-aware workload assignment in data centers," in *2018 International Conference on High Performance Computing & Simulation (HPCS)*, 2018, pp. 260-267.
[18] X. Li, X. Jiang, P. Garraghan, and Z. Wu, "Holistic energy and failure aware workload scheduling in Cloud datacenters," *Future Generation Computer Systems*, vol. 78, pp. 887-900, 2018.
[19] A. A. Khan, M. Zakarya, and R. Khan, "Energy-aware dynamic resource management in elastic cloud datacenters," *Simulation Modelling Practice and Theory*, vol. 92, pp. 82-99, 2019.
[20] D. S. Chisca, I. Castineiras, D. Mehta, and B. OSullivan, "On energy-and cooling-aware data centre workload management," in *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, 2015, pp. 1111-1114.
[21] S. MirhoseiniNejad, H. Moazamigoodarzi, G. Badawy, and D. G. Down, "Joint data center cooling and workload management: A thermal-aware approach," *Future Generation Computer Systems*, vol. 104, pp. 174-186, 2020.
[22] A. J. Shah, V. P. Carey, C. E. Bash, and C. D. Patel, "Exergy-based optimization strategies for multi-component data center thermal management: Part I—analysis," in *ASME 2005 Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems collocated with the ASME 2005 Heat Transfer Summer Conference*, 2005, pp. 205-213.

[23] A. J. Shah, V. P. Carey, C. E. Bash, and C. D. Patel, "Exergy-Based Optimization Strategies for Multi-Component Data Center Thermal Management: Part II—Application and Validation," in *ASME 2005 Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems collocated with the ASME 2005 Heat Transfer Summer Conference*, 2005, pp. 215-224.

[24] A. J. Shah, V. P. Carey, C. E. Bash, and C. D. Patel, "Exergy analysis of data center thermal management systems," *Journal of Heat Transfer*, vol. 130, 2008.

[25] A. J. Shah, V. P. Carey, C. E. Bash, C. D. Patel, and R. K. Sharma, "Exergy analysis of data center thermal management systems," in *Energy Efficient Thermal Management of Data Centers*, ed: Springer, 2012, pp. 383-446.

[26] A. J. Díaz, R. Cáceres, J. M. Cardemil, and L. Silva-Llanca, "Energy and exergy assessment in a perimeter cooled data center: The value of second law efficiency," *Applied Thermal Engineering*, vol. 124, pp. 820-830, 2017.

[27] L. Silva-Llanca, A. Ortega, K. Fouladi, M. del Valle, and V. Sundaralingam, "Determining wasted energy in the airside of a perimeter-cooled data center via direct computation of the Exergy Destruction," *Applied energy*, vol. 213, pp. 235-246, 2018.

[28] L. Silva-Llanca, M. del Valle, A. Ortega, and A. J. Díaz, "Cooling effectiveness of a data center room under overhead airflow via entropy generation assessment in transient scenarios," *Entropy*, vol. 21, p. 98, 2019.

[29] A. Bhalerao, K. Fouladi, L. Silva-Llanca, and A. P. Wemhoff, "Rapid prediction of exergy destruction in data centers due to airflow mixing," *Numerical Heat Transfer, Part A: Applications*, vol. 70, pp. 48-63, 2016.

[30] K. Fouladi, A. P. Wemhoff, L. Silva-Llanca, K. Abbasi, and A. Ortega, "Optimization of data center cooling efficiency using reduced order flow modeling within a flow network modeling approach," *Applied Thermal Engineering*, vol. 124, pp. 929-939, 2017.

[31] R. Gupta, S. Asgari, H. Moazamigoodarzi, S. Pal, and I. K. Puri, "Cooling architecture selection for air-cooled Data Centers by minimizing exergy destruction," *Energy*, p. 117625, 2020/04/19/ 2020.

[32] H. Moazamigoodarzi, S. Pal, S. Ghosh, and I. K. Puri, "Real-time temperature predictions in it server enclosures," *International Journal of Heat and Mass Transfer*, vol. 127, pp. 890-900, 2018.

[33] S. Pal, H. Moazami-goodarzi, S. Ghosh, K. Thorn, and I. K. Puri, "Plate-fin heat exchanger suitable for rack-mountable cooling unit," ed: Google Patents, 2019.

[34] H. Moazamigoodarzi, S. Pal, D. Down, M. Esmalifalak, and I. K. Puri, "Performance of a rack mountable cooling unit in an IT server enclosure," *Thermal Science and Engineering Progress*, vol. 17, p. 100395, 2020.

[35] H. Moazamigoodarzi, R. Gupta, S. Pal, P. J. Tsai, S. Ghosh, and I. K. Puri, "Modeling temperature distribution and power consumption in IT server enclosures with row-based cooling architectures," *Applied Energy*, vol. 261, p. 114355, 2020.

[36] M. Kheradmandi, D. G. Down, and H. Moazamigoodarzi, "Energy-Efficient Data-Based Zonal Control of Temperature for Data Centers," in *2019 Tenth International Green and Sustainable Computing Conference (IGSC)*, 2019, pp. 1-7.

[37] Z. M. Pardey, D. W. Demetriou, H. S. Erden, J. W. VanGilder, H. E. Khalifa, and R. R. Schmidt, "Proposal for standard compact server model for transient data center simulations," *ASHRAE Transactions*, vol. 121, pp. 413-422, 2015.

[38] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, pp. 732-794, 2015.
[39] C. Jin, X. Bai, C. Yang, W. Mao, and X. Xu, "A review of power consumption models of servers in data centers," *Applied Energy*, vol. 265, p. 114806, 2020.
[40] F. Chen, J. Grundy, Y. Yang, J.-G. Schneider, and Q. He, "Experimental analysis of task-based energy consumption in cloud computing systems," in *Proceedings of the 4th ACM/SPEC International Conference on Performance Engineering*, 2013, pp. 295-306.
[41] S.-W. Ham, M.-H. Kim, B.-N. Choi, and J.-W. Jeong, "Simplified server model to simulate data center cooling energy consumption," *Energy and Buildings*, vol. 86, pp. 328-339, 2015.
[42] H. Packard, "HP Proliant DL360 G5 Server series-overview," *Consulta: Junio*, 2007.
[43] J. M. Gordon, K. Ng, and H. Chua, "Optimizing chiller operation based on finite-time thermodynamics: universal modeling and experimental confirmation," *International Journal of Refrigeration*, vol. 20, pp. 191-200, 1997.
[44] TRANE, "Air-cooled Liquid Chillers," in *Product catalog,10-60 Tons*, ed, 2004.
[45] S. Denki, "San Ace 80," in *Product Catalog*, ed, 2016.
[46] D. L. Beaty, "Data center energy metric: Power Usage Effectiveness (PUE)," *Ashrae Journal*, vol. 55, pp. 61-63, 2013.
[47] M. P. David, M. Iyengar, P. Parida, R. Simons, M. Schultz, M. Gaynes, *et al.*, "Experimental characterization of an energy efficient chiller-less data center test facility with warm water cooled servers," in *2012 28th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM)*, 2012, pp. 232-237.
[48] M. Ouchi, Y. Abe, M. Fukagaya, H. Ohta, Y. Shinmoto, M. Sato, *et al.*, "Thermal management systems for data centers with liquid cooling technique of CPU," in *13th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, 2012, pp. 790-798.
[49] Y. Pan, R. Yin, and Z. Huang, "Energy modeling of two office buildings with data center for green building design," *Energy and Buildings*, vol. 40, pp. 1145-1152, 2008.
[50] R. Steinbrecher and R. Schmidt, "Data center environments," *ASHRAE Journal*, vol. 53, pp. 42-49, 2011.
[51] M. Iyengar and R. Schmidt, "Analytical modeling for thermodynamic characterization of data center cooling systems," *Journal of Electronic Packaging*, vol. 131, 2009.
[52] D. W. Demetriou, H. E. Khalifa, M. Iyengar, and R. R. Schmidt, "Development and experimental validation of a thermo-hydraulic model for data centers," *HVAC&R Research*, vol. 17, pp. 540-555, 2011.
[53] D. W. Demetriou, "Thermally aware, energy-based techniques for improving data center energy efficiency," 2012.
[54] K. Brown, W. Torell, and V. Avelar, "Choosing the optimal data center power density," *Schneider White, Boston, MA*, p. 2, 2014.