Noise Aware Scheduling in Data Centers

Hameedah Sultan
School of Information Technology
Indian Institute of Technology
New Delhi, 110016
hameedah.sultan@gmail.com

Arpit Katiyar
Samsung Research Lab
Noida, Uttar Pradesh 201303
akatiyar.cse.iitd@gmail.com

Smruti R. Sarangi
Computer Science and Engg.
Indian Institute of Technology
New Delhi, 110016
srsarangi@cse.iitd.ac.in

ABSTRACT

As the demand for large scale computing is rapidly increasing to serve billions of users across the world, more powerful and densely packed server configurations are being used. Often in developing countries, and in small and medium enterprises, it is hard to place such servers in sound-proof server rooms. Hence, servers are typically placed in close proximity to employees. The noise from the cooling fans in servers adversely affects employees' health, and reduces their productivity. In this paper, we provide a framework for computer architects to measure the acoustic profile in a data center along with the temperature profile, and estimate the sound power levels at points of interest. Additionally, we studied the noise levels obtained upon using algorithms targeted at homogenizing the temperature profile. We found that these algorithms result in high noise levels, sometimes above the permissible levels. So, we propose two heuristics to redistribute workloads in a data center such that noise can be reduced at certain target locations. We obtain a noise reduction of 2-13 dB when compared with uniform workload distribution, and upto 16 dB as compared to temperature aware workload placement, with a reduction of at least 5-6 dB in 75% of the cases. The performance overhead is limited to 1%.

1. INTRODUCTION

In the last few years, there has been a significant rise in the number of data centers due to increased reliance on internet based services, cloud computing, analytics, and social networking. Data centers have become ubiquitous. They are being set up and used by software vendors, universities, startup companies, internet service providers, and financial firms. Unfortunately, hundreds of high performance servers generate a lot of heat, and it is necessary to cool the data center to keep the temperature below specified limits. Designers typically adopt a multi-level approach to cool a data center. In the first level, we remove the heat generated by the electronic components (processors and memory), and dissipate it to the surrounding air. This is achieved by the fans attached to processors, servers, and racks. These fans blow air over the heat generating components, thus removing excess heat through convective heat transfer. Next, it is necessary to remove the heat generated from the entire data center. This is done by typically blowing cold air from the floor, and removing hot air from the ceiling of the data center. This hot air is cooled by a CRAC unit (computer room air conditioning unit), and that air is again recirculated through the bottom of the data center.

Let us now consider the fans in some more detail. Typically, the temperature at the surface of the processor die reaches 80-100°C for a computationally heavy workload, whereas the ambient temperature of a data center typically needs to be maintained between 18-22°C [1]. This means that the fans have to remove a significant amount of heat by rotating at high speed. Unfortunately, this causes a lot of noise in data centers. It has been shown that for a 50% increase in the fan speed the sound pressure can increase by 10X [2]. Since fans account for an overwhelmingly large portion of the noise in data centers [3], we need to effectively control the fans to minimize noise.

Most server manufacturers report the noise produced by the servers. Subsequently, acoustics engineers take this information into account while designing server racks, and data centers. However, to the best of our knowledge, we are not aware of efforts that use noise as an axis in computer architecture and systems research.

Unfortunately, server noise has very pernicious effects, and can lead to long term hearing loss. OSHA (Occupational Safety & Health Administration) [4] recommends noise levels to be always less than 85 dB. In Europe, the regulations are stricter with a threshold of 80 dB. Even with lower noise levels, the constant hum can introduce hearing problems such as tinnitus, and ultimately lead to hearing loss. A study by Kujawa and Liberman [5] shows that hearing loss is cumulative. A long exposure to moderate noise levels can be as harmful as a small exposure to very high noise levels. The damage to the inner ear continues to build up over time. **Apart from hearing loss, moderate noise levels lead to psychological stress, and negatively affect job satisfaction and productivity [6].**

Our goal in this paper is to highlight the problem of noise in servers to the architecture community.

1.1 Motivation

Let us first motivate the problem by being a devil's advocate and claim that modeling and reducing noise is not important. We can argue that since servers are often kept

ICS '16, June 01-03, 2016, Istanbul, Turkey

DOI: http://dx.doi.org/10.1145/2925426.2926268

in dedicated server rooms, they can be acoustically isolated by having thick anechoic walls. If workers need to perform maintenance work, then they can use noise canceling headphones. However, this represents the best case scenario. Often smaller setups such as start-up companies do not have enough space for additional server rooms with CRAC units. Hence, they place servers in open rooms along with employees' cubicles as has been reported in the popular press [7–9]. By putting servers in a large room, it is easier to dissipate the heat. Moreover, the lack of dedicated server rooms is an acute problem in developing countries. It is often difficult to find the space for a separate room and fit air conditioners. It is easier to place an array of racks in an existing lab that has a moderate amount of air conditioning, and has enough power to run the servers. Even in developed countries, servers are often placed in non-data center environments, such as home offices, hospitals and libraries [9]. As businesses grow, servers are deployed to handle the increased amount of transactions, without building a dedicated server room.

1.2 Alternative Solutions and their Drawbacks

Noise canceling headphones are not a panacea for our problems. Workers often need to stay for a prolonged duration inside the server room to perform maintenance work [10]. They need to talk to each other, and to the original server vendors for rectifying problems. They often remove the headphones while having such conversations (as reported in [10, 11]), but the high noise levels seldom allow clear communication. In this case, we need a noise canceling microphone or a noise canceling headphone on the other side, both of which have been shown to have limited applicability [12]. As evidenced by a recent Forbes article [13], even the best noise canceling headphones that are priced at about 500 USD are primarily effective for removing steady, **predictable**, low frequency (100-500 Hz) noise. They are not suitable for scenarios with transient noise (like servers, or disk racks powering up), and scenarios where the noise frequency is more than 500 Hz. Empirical studies in real world data centers [14] indicate that the noise can go up to 3 KHz, and thus noise canceling headphones have limited applicability. Given these observations, researchers are focusing on active noise control using large loud speakers instead. These approaches [15] propose to broadcast a sound wave that will cancel out ambient noise. The report concludes that this method is not successful especially towards the center of the server room, and thus this solution might not be feasible. The basic point here is that server noise has a very random nature, is spread over a large frequency range and thus it is hard to filter it out. Also note that if improperly designed, active noise cancellation devices may add noise instead of removing it. Moreover, headphones are not very comfortable to wear for a long time. In recent *Reddit* threads [16–22], people who regularly work in data centers have discussed the possible means of ear protection, and accepted noise canceling/isolating headphones/ear muffs as the last resort. Moreover, users have reported several problems such as perspiration, pain in the ear and reduction in the ability to concentrate for extended periods of time. The United Kingdom Health and Safety Executive (HSE) as well as the American Occupational Safety and Health Administration (OSHA) thus **recommend the use of hearing protection only where engineering solutions are not present, and as a short-term measure while other technical means are being developed, or to provide additional protection [23, 24]**.

While these represent the efforts being made at the individual level, at the systems level too, measures are needed to reduce noise levels. Consequently, most server manufacturers today make "quiet" servers that are friendly to office environments, such as Dell [25], HP [26], IBM [27], and Fujitsu [28], thus giving credence to the fact that server noise is a well-recognized problem. However, such systems have limitations in performance.

Commercial products that reduce noise by enclosing servers in noise reduction chambers (often with built-in cooling systems) are also available [29–32]. However, this method suffers from several drawbacks, including airflow and heating problems. If openings are created to allow air flow, they let the noise out as well as pointed out in the community [33] and our in-house simulation studies. Also these solutions are very costly. The cost of a single enclosure, for a 12U server rack can be as high as 2000 USD [17]. Having 20 such enclosures adds an additional 40,000 USD in costs.

Noise reduction in data centers has only recently caught the attention of systems engineers. Several articles in the technical press have talked about the problem of noise in data centers, and measures to reduce it [34–40]. **Interestingly, most of these articles have been written over the last year**. As more and more workloads continue to be ported to data centers, the problem of noise becomes even more significant. Some IT managers have used sound-absorbing acoustic tiles to reduce data center noise [41]. While this helps to some extent, experts who have headed the American Society of Heating, Refrigerating and Air-Conditioning Engineers' Technical Committee, have called for changes in the base design for more noise mitigation [41]. A thermal strategist at HP observes that quieter servers would offer a competitive advantage to companies [41]. These statements suggest that IT managers are ready to welcome any method that can lead to reduction in noise, without a serious power penalty.

Sadly, since the power densities are expected to continue rising, data centers would become even louder, and thus it is now necessary to look for solutions in computer systems domain, which can complement solutions proposed by acoustics engineers.

1.3 Proposed Solution

In this paper we propose an algorithm to schedule workloads in order to minimize noise. Since high temperature results in fans rotating at higher speeds, temperature should be correlated with noise. We explored the performance (in terms of noise) of heuristics targeted at reducing temperature hotspots. We concluded that such heuristics do not reduce noise sufficiently and reliably. So we propose 2 new heuristics explicitly targeted at reducing noise levels, without affecting the temperature distribution. Instead of sophisticated solutions that either cause physical discomfort, increase the construction cost and TCO by hundreds of thousands of dollars, scheduling is a cheap option. A lot of modern data centers already have virtualization technologies installed that allow seamless job migration between servers [42]. All that we are proposing is to minimally enhance this framework such that noise reduction can be one of the criteria for job migration. We show that with such minimal overhead (both software or hardware) it is possible to bring the noise down to OSHA and European Union Directives' limits.

Our contributions in this paper are as follows. We first introduce the basics of acoustics and formulate five noise reduction problems in Section 2. Subsequently, we propose a system that can be implemented in a data center to collect temperature and noise information, and effectively schedule workloads to minimize noise at target points of interest (see Section 3). We propose two novel workload scheduling(migration) heuristics for this purpose in Section 4. We evaluate the efficacy of our heuristics by simulating the thermal and noise profiles in a mid-sized data center, using Ansys Icepak and *AcouSTO* [43] respectively. For getting the profiles of real benchmarks we conduct in-vivo measurements with loop ammeters. In Section 5, we describe the design of our simulated data center. and in Section 6, we show that our workload scheduling algorithms can reduce noise by 2-13 dB for different scenarios. Finally, we conclude in Section 7. For additional details and results, interested readers can take a look at the Appendices (available online [44]).

2. BACKGROUND AND RELATED WORK

2.1 Basic Physics

Sound is a longitudinal wave, which travels in a medium involving a succession of compressions and rarefactions. Human hearing responds logarithmically to a sound level, hence a logarithmic scale is more suitable to represent acoustic noise (measured in decibels (dB)). The decibel scale can be used to express the amplitude of an acoustic noise source both in terms of pressure and power.

The sound pressure level, L_p (or SPL), for a sound source with sound pressure (p_{rms}) is:

$$L_p = 20 \times log_{10}\left(\frac{p_{rms}}{p_{ref}}\right)(dB) \tag{1}$$

where, $p_{ref} = 20 \times 10^{-6} Pa$

The sound power level, L_w (or PWL), for a source with power $\mathcal{P}\, Watts$ can be calculated by Equation 2.

$$L_w = 10 \times log_{10}\left(\frac{\mathcal{P}}{\mathcal{P}_{ref}}\right)(dB) \tag{2}$$

where, $\mathcal{P}_{ref} = 10^{-12} Watt$.

A change of 3 dB in the level of sound is just noticeable to the human ear, 5 dB is clearly noticeable, while a change of 10 dB is perceived to be two times louder [46]. Figure 1 gives us an idea about the noise levels of typical scenarios.

2.2 Characterization of Sound Waves

The speed of propagation of sound in a medium depends on the temperature. The speed of sound in air, at 20 °C and 1 atmosphere pressure is 343 m/s. At other temperatures it can be calculated by Equation 3.

$$c = 332 + 0.6T_c \tag{3}$$

where, T_c is the air temperature in °C

When multiple sources are present, they can be correlated or uncorrelated. In acoustic theory, correlated sources are defined to have a precise time and frequency relationship between them whereas uncorrelated sources do not have such a relation. Correlation between sources governs how we calculate the overall impact of these sources at a particular point. In the simplest form, we consider n sources with PWLs L_1, L_2,..., L_n at any point. Then the total sound power level (L_t) at that point can be calculated using Equation 4 if they are perfectly correlated and from Equation 5 if they are uncorrelated.

$$L_t = 20 * log[\sum_{i=1}^{n}(10)^{L_i/10}] \tag{4}$$

$$L_t = 10 * log[\sum_{i=1}^{n}(10)^{L_i/10}] \tag{5}$$

Appendix A (available online) describes the physics of noise propagation in some more detail.

2.3 Characterization of Noise Produced by Cooling Fans

A cooling fan propels a constant volume of cold air towards an electronic component per second. The cold flowing air continuously transfers heat from the region where it is generated to some other region via convective heat transfer. The amount of cooling depends on the flow rate of the fan; a higher flow rate achieves a higher amount of cooling. Also, as the flow rate increases, the noise produced by a fan also increases [2].

The amount of cooling achieved (ΔT) is governed by the mass flow rate $\dot{m}$ of the fan, and the heat transfer rate $\dot{Q}$:

$$\dot{Q} = \dot{m} C_p \Delta T \tag{6}$$

where, C_p = specific heat of air, and ΔT = temperature difference.

The volumetric flow rate, $\dot{V}$ of a fan is related to the mass flow rate $\dot{m}$ by Equation 7.

$$\dot{m} = \rho \dot{V} \tag{7}$$

Equations 6 and 7 can be combined to determine the volumetric flow rate needed to achieve a given amount of heat dissipation, as follows:

$$\dot{V} = \dot{Q}/\rho C_p \Delta T \tag{8}$$

Empirical Method for Estimation of Acoustic Noise.

Lyon and Bergles [2] model the noise generation due to unsteady aerodynamics inside closed electronic packages. The authors establish that acoustic noise produced by cooling air passing through electronic packages has two components. The first is the noise produced by the moving air stream around the blades of the fan and the motor, which is given by Equation 9. The sound level produced by a cooling fan is proportional to the product of the pressure drop across the fan and the flow rate. Although, Equation 9 does not explicitly show the pressure term but its effect is incorporated by expressing it in terms of the other parameters. The second component is due to the turbulent (unsteady) flow of air through the electronic package. As the fast moving air flows through the circuit board, noise is created due to friction between air and the components on the circuit board.

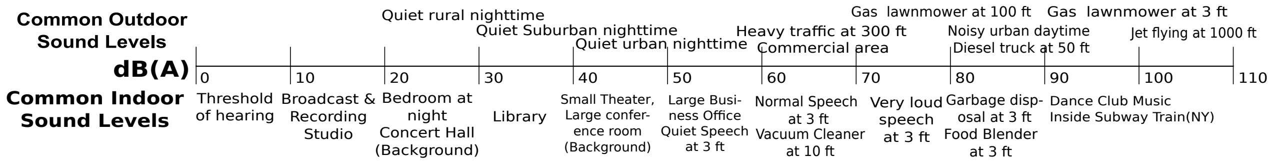

Figure 1: Sound pressure levels in various scenarios [45]

The authors model a relation between the overall PWL and airflow between parallel plates of circuit boards as shown in Figure 2. They propose Equation 10 to estimate the total sound power per unit length. Estimating turbulent sound power involves two steps. First they estimate the noise radiated by one patch as shown in Figure 2, and then they combine all these coherent patches across the length of the board to estimate the overall sound power.

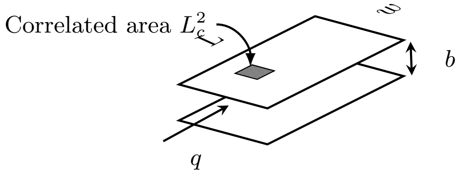

Figure 2: Turbulent flow in a circuit board [2]

$$\mathcal{P}_{fan} = \frac{323 \times \mathcal{P}_{ref} L^2 q^5}{D^{10}} \qquad (9)$$

where, $\mathcal{P}_{ref} = 10^{-12} Watt$, $D = 2b$ and q is the mass flow rate.

$$\mathcal{P}_{turb} = \frac{\pi}{16} \times \frac{C_D^2}{\rho^4 c^2} \times \frac{L_c}{w^4 b^6} \times q^5 \qquad (10)$$

where, ρ is the density of the air, c is the speed of sound in air and C_D is the drag coefficient (resistance of an object to the movement of air). For electronic packages, C_D is taken to be 0.1 and $L_c = b$.

The total sound power generated by a mass flow rate of q can be calculated by adding Equations 9 and 10.

2.4 Objective

Our main objective is to invite the attention of the computer architecture community towards designing data centers with lower noise levels. We consider the following 5 situations for acoustic noise reduction:

- Minimize noise inside an office in a corner. Data centers typically have a small office in the corner for monitoring the health of the data center.
- Minimize noise in a set of offices along a side.
- Minimize noise at a given point where a worker is assumed to be located.
- Minimize noise at multiple points where multiple workers are assumed to be present. These workers are assumed to be mobile workers who are performing repairs at different sites in the data center.
- Minimize noise on one side of the data center. This is often required when there is an adjoining room that shares a common wall with the data center.

These use cases have been inspired by real world problems and personal communications with data center managers.

2.5 Related Work

Although workload redistribution is a well researched area, most such strategies are targeted at reducing the total power consumption and minimizing thermal hotspots, while our objective is to reduce noise levels. Since there is no prior work in the computer architecture/systems communities (to the best of our knowledge) that looks at smartly scheduling loads in a data center to reduce noise, the closest related work is in the area of workload distribution to minimize thermal hot spots. We list some of the important contributions in this area. Raid et al. [47] describe a method for altering cooling fan flow rates in multi-socket CPU systems to decrease the overall energy consumption. In comparison, Moore et al. [48] take a more global view by exploring hot-spots and cold-spots inside a data center, and then based on thermodynamics principles devise a scheduling algorithm to minimize temperature. Sharma et al. [49] proposed a simpler algorithm where the workload assigned to a rack is inversely proportional to its temperature. This allows workloads to migrate from hot spots to cold spots and thus the temperature profile gets homogenized. A later approach by Vasic et al. [50] uses a thermodynamic model of a data center, and consists of a scheduling algorithm designed using control theoretic concepts. The authors model a closed loop control system consisting of the data center, a scheduler, and a cooling unit. The temperature aware scheduler accounts for small changes in the workload, while large fluctuations are handled by the cooling unit.

However, the problem of noise reduction is different. We were not able to directly apply temperature reduction algorithms to reduce noise. The effects of high temperature are far more localized than noise that propagates further. Additionally, **there is a complex relationship between temperature, reduction of fan speed, location of the rack, and the workload assigned**. It is hard to capture this relationship analytically, or use solutions targeted to reducing temperature and total power. Heuristics explicitly targeted towards noise reduction performed the best in our evaluation.

Noise reduction problems have been of interest in industrial environments. However, most industrial noise reduction techniques reduce noise at the source (usually by design), in the transmission path (usually by the use of isolators and sound absorbing materials), or at the receiver (by isolating the receiver or providing hearing protection, or workforce rotation). Noise reduction at the source is usually specific to the noise producing machine, and relies on the underlying mechanism by which noise is produced. Isolating the servers is not possible in our case, because of air flow and maintenance issues. We have already discussed the problems associated with hearing protection at the receiver level. None of these methods can provide a solution for our case. Hence our endeavor is to provide a method for noise control, which can complement the noise reduction offered by better design.

3. ARCHITECTURE

There are two main components apart from the regular components present in a data center - a central server (CS) and a rack controller (RC).

3.1 Central Server

A central server (CS) is responsible for distributing a set of jobs in the data center. Each rack has a rack controller (RC) that collects power and temperature information and periodically sends it to the CS. Using this information, the CS redistributes workloads among the servers based on the noise reduction criteria. The target locations have microphones installed, and the sound level information is also passed to the CS. We assume that tasks can freely migrate between servers. This can be achieved by VM migration technologies such as VMware's vMotion [51]. After redistributing the workloads, the individual RCs compute locally optimal configurations. The aim is to minimize the flow rate of fans such that the ambient temperature around a rack is below a certain threshold.

Our contributions in this space are a set of heuristics used by the RC and CS to generate good local and global workload configurations respectively. The RC and CS can either be dedicated servers, software entities, or specialized hardware modules running on FPGAs.

3.2 Rack Controller

The rack controller (RC) has two functions. The first function is to periodically take power and temperature measurements and send it to the CS. We can measure the amount of power by either having ammeters in series with the power supply lines of the rack, or by using loop ammeters that are non-intrusive. Additionally, each rack has a set of temperature sensors that measure the temperature at the inlet of servers. The maximum temperature is sent to the CS.

The second function of the RC is to achieve a stable configuration after the CS redistributes the workloads in the data center. The CS sends a message to the RC that contains a list of tasks and their destination IP addresses. These are forwarded to the servers in the rack. They subsequently initiate process migrations. Our scheme is oblivious to the nature of the process migration mechanism. Once, the new workload runs on a rack the RC modulates the fan speeds such that the ambient temperature is less than a given threshold. The terms used are defined in Table 1.

Symbol	Meaning
n	Number of racks in the data center
θ	Rack temperature limit for safe operation
δ	Minimum amount by which the fan speed can be changed (in cfm)
l	Number of racks that are close to the target point
m	Number of racks that are far away
W	Array containing the power of each rack
F	Array containing the fan flow rates of each rack
A	Array containing the amount of power added to each rack
D	Array containing ranks of the racks sorted on the basis of distance
T	Array containing the temperature of each rack
W_{ex}	Power dissipation of the extracted workloads that have to be migrated

Table 1: Glossary of terms

4. THE WORKLOAD DISTRIBUTION ALGORITHM

4.1 Workload Distribution by CS

Initially the CS distributes workloads uniformly over all the running racks. Subsequently, the RC sets fan speeds such that the temperature constraint is met. The algorithm to set the fan speed is as follows. For the sake of simplicity, we assume that the fan speeds of each server in a rack are the same. This is not a very unrealistic assumption given that rack temperature was found to be relatively immune to the power levels of individual servers. At the outset, we measure the inlet temperature of the servers for each possible combination of rack power and fan speed (flow rate). We store this information in a lookup table.

At runtime, our task is to find the fan flow rates which would satisfy the temperature constraint for the given workload distribution. From the lookup table, we can directly determine this for a given server, if all other servers are idle. But at runtime, the neighboring servers too would be running workloads. However, we found the data center to act as a thermal low pass filter. As a result, the inlet temperature at a server is not very strongly dependent on the workloads running on neighboring racks. It is more correlated with the total power dissipation in the data center, and the nature of the cooling. For a given data center layout, once the workloads are distributed, both of these are fixed. Hence the RC can more or less determine the fan flow rate required based on the values stored in the lookup table. There might be some minor violations of the threshold at some servers at runtime. We fix this by increasing the fan speed by one step in each iteration, till the temperature falls below the threshold. We found that we needed no more than 2-3 iterations for this purpose.

We assume two functions in Algorithm 1 – $getMax-Workload(rackId, rackFlow)$, and $getMinFlowRate(rackId, rackPower)$. The former function returns the maximum amount of workload (measured in terms of power dissipation) that can be placed on a rack given a flow rate, and a temperature threshold. The latter function returns the minimum fan flow rate that is required to run a given workload with power, $rackPower$, and remain within a temperature threshold. Note that in our algorithm we measure the workload in terms of Watts, and assume that it can fluidly be transferred between servers. A mapping to real processes will be described in Section 5.2.

In the next step, we move workloads from the racks close to the target location where we want to minimize noise, to the remote racks based on two different heuristics.

ALGORITHM 1: Workload Redistribution Algorithm

```
Input: $W$, $D$
Output: $W$ after redistribution of workload.
$\theta \leftarrow 26$; $flag1 \leftarrow 0$; $flag2 \leftarrow 0$; $W_{ex} \leftarrow 0$;
$prevSoundLevel \leftarrow 1000$; $soundLevel \leftarrow 0$;
repeat
    StabilizeTemperature() ;
    ExtractWorkload() ;
    HeuristicI() OR HeuristicII() ;
    StabilizeTemperature() ;
    if $(prevSoundLevel < soundLevel)$ ;
    then
        $flag1 \leftarrow 1$;
    end
    $prevSoundLevel \leftarrow soundLevel$;
until $flag1 = 1$;
```

Procedure StabilizeTemperature( )

```
repeat
    $flag2 \leftarrow 1$;
    Measure temperature at observation points($T$) with
    workload $W$, fan speed $F$;
    for $i \leftarrow 1$ to $n$ do
        if $(T[i] > \theta)$ ;
        then
            $flag2 \leftarrow 0$;
            $F[i] \leftarrow F[i] + \delta$;
        end
    end
until $flag2 = 1$;
```

Procedure ExtractWorkload( )

```
for $i \leftarrow 1$ to $l$ do
    $j \leftarrow D[i]$; $F[j] \leftarrow F[j] - \delta$;
    $W_{new} \leftarrow getMaxWorkload(j, F[j])$;
    $W_{ex} \leftarrow W_{ex} + W[j] - W_{new}$;
    $W[j] \leftarrow W_{new}$;
end
```

Procedure HeuristicI( )

```
repeat
    for $i \leftarrow 1$ to $m$ do
        $j \leftarrow D[n - i + 1]$;
        $F[j] \leftarrow F[j] + \delta$;
        $W_{new} \leftarrow getMaxWorkload(j, F[j])$;
        $W_{ex} \leftarrow W_{ex} + W[j] - W_{new}$; $W[j] \leftarrow W_{new}$;
        $checkMaxWorkload()$;
    end
until $W_{ex} \leftarrow 0$;
```

Procedure HeuristicII( )

```
repeat
    for $i \leftarrow 1$ to $m$ do
        $j \leftarrow D[n - i + 1]$;
        $W_{prev} \leftarrow W[j]$;
        $W[j] \leftarrow W[j] + A[j]$;
        $F[j] \leftarrow getMinFlowrate(j, W[j])$;
        $W_{ex} \leftarrow W_{ex} - W[j] + W_{prev}$;
        $checkMaxWorkload()$;
    end
until $W_{ex} \leftarrow 0$;
```

4.2 Workload Extraction from Adjoining Racks

Let us assume there are n racks in our data center. We have to determine the amount of workload that can be shifted to the servers that are located farthest from the point where we want to minimize noise. To do this, the CS selects l racks for workload extraction, physically closest to the location at which we want to minimize noise. The rack controllers of the corresponding racks reduce the fan speeds by a fixed value, δ. As a consequence, the workload on these l racks has to be reduced (to meet the temperature constraint). From the pre-computed data, the CS figures out the new workload that can be placed on these servers for the reduced flow rate. The difference between the old and the new power consumption for each rack is added to the amount of power that has to be migrated, W_{ex}. This task is accomplished by the function $ExtractWorkload()$ in Algorithm 1.

4.3 Workload Addition to Far-away Servers

Next, we have to figure out the amount of workload that has to be added to each of the m racks that are far away. In the beginning, additional workloads are added only to racks that are already running. If all the running servers reach the maximum power limit, then we start adding workloads to the previously switched off servers. We have used two different heuristics to determine the workloads that should be added to each rack.

4.3.1 Heuristic-I

In this method, the CS maintains a sorted list of the racks in the data center. They are sorted in the order of proximity to the points of interest. The CS sends out a signal to the RCs of the m remotest racks to increase their fan speeds. As a consequence, the amount of power that can be placed on those servers increases. From the pre-computed data set, the CS figures out the new workload that can be accommodated for the increased flow rate, and operates these servers at the increased power levels. The difference between the old and the new workloads (workload added) is reduced from W_{ex}. The CS continues migrating tasks until W_{ex} becomes zero. This method is implemented using the procedure $Heuristic1()$ in Algorithm 1.

4.3.2 Heuristic-II

In this approach, the workload to be added is calculated by a weighted proportion based method. Let W_{min} ($\neq 0$) be the lowest possible power consumption of a job and let $\mathcal{D}[i]$ be the distance of the i^{th} rack from the point of interest. In the case where multiple workers are present, $\mathcal{D}[i]$ is based on the geometric mean of the distances of the workers from the i^{th} rack. $D[i]$ denotes the rack numbers in increasing order of distance from the target location. The amount of power $A[j]$ added to the j^{th} rack is calculated using Equations 11 and 12.

$$\mathcal{W}[j] = \frac{\mathcal{D}[j]}{\sum_{k=n-m+1}^{n} \mathcal{D}[k]} \tag{11}$$

$$A[j] = (\left\lceil \frac{\mathcal{W}[j] * W_{ex}}{W_{min}} \right\rceil + 1.0) * W_{min} \tag{12}$$

The farther a rack is, the larger is the weight $\mathcal{W}[j]$ assigned to it. The weight is normalized by the total distance of the far-off racks, so that the sum of the added workload is equal to the total extracted workload[1]. The CS selects the m farthest servers, and adds $A[j]$ amount of power to the j^{th} rack in steps of W_{min}, where $A[j]$ is obtained by weighting the extracted power by the factor $\mathcal{W}[j]$. If this additional power results in a violation of the maximum power limit of a rack, then only the permissible amount of power is added and the corresponding amount is subtracted from W_{ex}. This process is coordinated by the CS. The RC checks using the pre-computed rack data, whether fan speeds for this rack need to be updated since adding extra workloads may cause a temperature violation. The flow rates are updated accordingly, if required. In this step, the CS also checks if it has to start some new servers to accommodate the extracted power. This process is continued until W_{ex} becomes zero. In Algorithm 1, the procedure $HeuristicII()$ is used to implement this method.

In Algorithm 1, the function $checkMaxWorkload()$ checks whether all the currently running servers have reached their maximum power limit. If this is the case, then it turns on a server (the server that is closest to the one being considered in the current iteration). Lists W, F, A and D are each of size n, where $W[i]$ denotes the power consumed by the i^{th} rack and $F[i]$ is the flow rate of its fans.

4.4 Temperature Stabilization and SPL Distribution

Next, the RC measures the temperature profile for the new workload distribution. To ensure that the target temperature requirements are met, the CS again invokes the function $StabilizeTemperature()$. Finally the SPL is measured at the target locations (using microphones) to determine the reduction in sound level obtained, and sent to the CS. If a reduction is obtained with respect to the previous distribution, then the CS re-iterates over the whole process, by reducing the fan speeds once again by δ. After a few iterations we will hit a local minima. Most of the workload at this point would have been moved to remote servers. The fans at those servers would be running at close to maximum speed. This might lead to an increase in the overall sound level at the target locations. The CS saves the workload distribution for the previous settings, until the sound level for the current distribution is obtained. When the obtained sound level for the current distribution is greater than the previous iteration, the algorithm stops and we obtain a local minima of the sound profile at the target location. The complete algorithm is represented pictorially in Figure 3.

Let us now describe what happens if jobs exit the system, or if new jobs enter the system. If jobs exit the system, then we do not do anything immediately. The CS runs the redistribution algorithm on a periodic basis. The next time that the redistribution algorithm runs, it will find jobs that can be migrated from the nearest l servers to the farthest m servers. However, if new jobs are added to the system, the CS is consulted before assigning a new job to a server. In this case the CS either uses Heuristic - I or II to map the jobs to the m farthest servers if possible (subject to power and temperature constraints). If not, it tries to start new servers, and then assigns jobs to the l nearest servers as the last resort. To turn off entire racks, the CS runs a periodic algorithm that tries to find servers that have a very little load (< 150 W). Before shutting them off, the CS tries to move their jobs using either heuristic I or II.

The rescheduling period of the algorithm is three data center thermal time constants, which represents an equitable trade-off between the time it takes for temperatures to stabilize and the responsiveness of the algorithms. We simulated the time required to obtain a stable temperature profile in several different scenarios, and observed it to be of the order of minutes, with an average of 3 minutes. Hence we run our rescheduling algorithm every 10 minutes. In [52], the same rescheduling time was used. The overhead due to the algorithm arises primarily from the workload migration time, plus pre and post migration overheads. Sherif et al.[53] studied workload migration overheads in detail, and showed that if live migration is performed, the downtime has a lower bound of 0.314 s and an upper bound of 1.519 s for a typical 1 GB VM on a 10 $Gbps$ link. Hence, the overhead because of our algorithm would be bounded between 0.05 % and 0.25 %. For a slower link speed of 1 $Gbps$, the lower and upper bounds on the downtime are 0.314 s and 9.498 s respectively. Hence the overhead is between 0.05% and 1.6%. In any case, the overhead is less than 2%, and is therefore affordable. References [54, 55] also report similar or lower overheads.

5. EVALUATION

5.1 Experimental Setup

Our data center ($12.19m \times 3.66m \times 9.14m$) model consists of 2 CRAC units, 2 power distribution units (PDUs), a raised floor air distribution plenum ($0.46m$ above the data center floor), ceiling plenum ($0.61m$ below the ceiling), cable trays, perforated floor tiles and 44 racks organized in 4 parallel rows as shown in Figure 8. Each rack contains 15 Dell PowerEdge R710 servers. Each such server has 2 Intel Core i7 CPUs, 64 GB main memory, and two 300 GB hard disks connected in a RAID 1 configuration.

5.1.1 Data Center Model for Thermal Simulation

For thermal simulation, we model the data center using Ansys Icepak [56], which is based on the Fluent [57] CFD package. Icepak supports two different methods for modeling a data center. The first approach is *server-based*, where we model a data center at the granularity of individual servers (including all its components such as the PCB, CPUs, heat sinks, and fans). These servers are then stacked to create racks, which are then arranged in rows to create a data center. The second approach is *macro-based*, where built-in macros are used for the various data center components such as racks, CRACs, perforated floor tiles, and power distribution units (PDUs) (Figure 8). Each rack is characterized by its power dissipation, and air flow. Scaling a server-based model to the level of a medium sized data center results in very complex and slow simulations. Additionally, we found issues with convergence. Hence, we used a macro-based model that does not have these problems, and simultaneously preserves accuracy. For meshing also, we have two approaches: conformal meshing and non-conformal meshing. The simulation time of a given model depends on the complexity of the mesh, and the nature of meshing. The conformal meshing scheme produces a mesh

---

[1] Note that we make appropriate adjustments to ensure that the workload added is equal to the workload removed

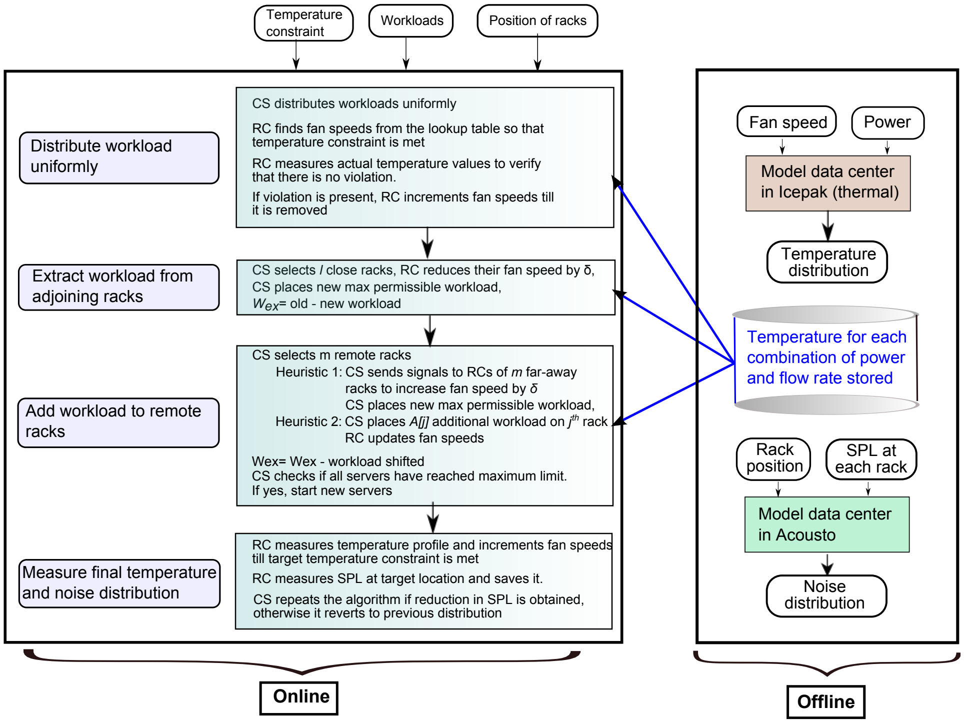

Figure 3: The complete workload redistribution algorithm

of the same size for each object. Hence, small objects unnecessarily increase the mesh complexity of larger objects around them. Non-conformal meshing allows every object to be meshed separately, allowing larger objects to be meshed with a coarser mesh. Hence we use the non-conformal meshing scheme, which reduces the mesh complexity, and has a reasonable simulation time and accuracy, along with a stable temperature profile.

Figure 4 shows the non-conformal mesh generated for the data center.

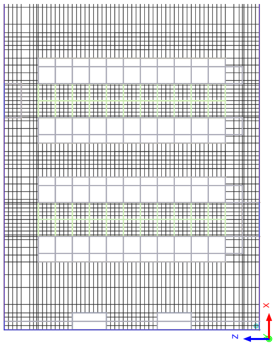

Figure 4: Data center mesh created using non-conformal meshing

We use our thermal simulation model to estimate the temperature of a data center given a distribution of workloads across the servers and a set of fan speeds.

The typical power consumption in a data center has been reported to range between $4-8\ kW$ per rack ([58]). The latest survey by the Data Center User's Group ([59]) cited the average power density per rack in a data center to be 5.94 kW ([60]). To qualitatively verify our model, we used synthetic data and randomly assigned an intermediate workload (4000 $Watt$) and low flow rate (500 cubic foot per minute(cfm)) to half the racks (in the first two rows closest to the CRAC units). To the remaining racks, we assigned a larger workload (6000 $Watt$) and a flow rate of 1000 cfm. We performed a CFD analysis of the model and found out that the solution converges quickly and the flow pattern obtained (Figure 6) is similar to what we intuitively expect (Figure 5). Figure 7 shows the temperature contour on the faces of each block in the data center model. We can observe that the temperature contours are consistent with the air flow since areas in which cold air from the CRAC units reaches have a lower temperature, and areas in which warm air flows have a higher temperature. Also, areas around the racks with a larger workload have a greater average temperature.

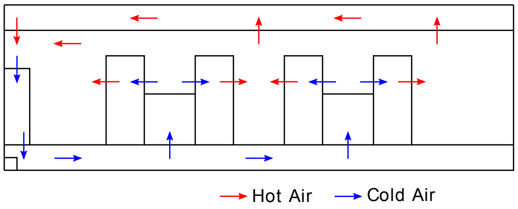

Figure 5: Expected air flow pattern inside the data center

5.1.2 Data Center Model for Acoustic Simulation

Once thermal simulation is done, we need to model the

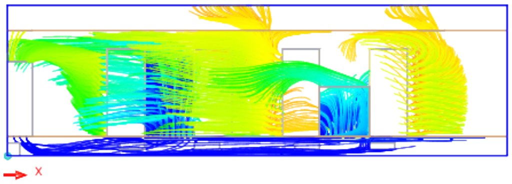

Figure 6: Simulated air flow pattern in the data center

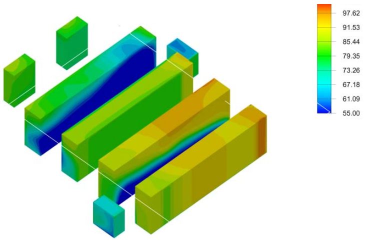

Figure 7: The temperature contours for different rows of racks in the data center

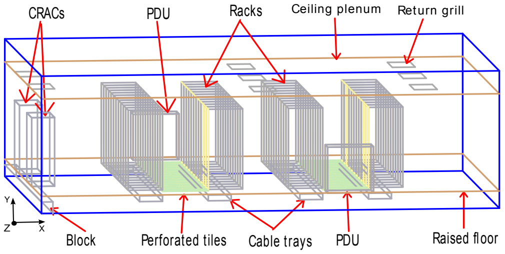

Figure 8: Model of a data center

acoustic noise distribution with the cooling fans acting as sources of noise. The sound power generated by the fans depends on their flow rate as given by Equations 9 and 10. To find the acoustic profile, we used the open source tool *AcouSTO* [43]. *AcouSTO* is a popular acoustic simulator, which uses the boundary element method to solve the Kirchhoff-Helmholtz equation. To verify the accuracy of *AcouSTO*, we conducted experiments with real sound sources and measured the acoustic field. We placed a sound source producing a single frequency ($250\,Hz$) noise in a room and measured the SPL at various points in the room. A sample *Gmsh* model of the room is shown in Figure 9. The open area in the model consisted of windows and doors. The room had painted brick walls (absorptivity, $\alpha = 0.01$), glass windows and doors (absorptivity, $\alpha = 0.25$). Figure 10 shows the top view of the room with the location of the source (2.09, 0.0, 2.7), and the points where we measured the SPL. For the acoustic simulation, we calculated the source power from the measured SPL at a point close to the source using Equation 13 and appropriately specified the boundary conditions.

$$L_w = L_p + \left| 10 \times log_{10}\left(\frac{Q}{4\pi r^2}\right)\right| \tag{13}$$

where, Q is the directivity factor.

Table 2 shows the values of observed and simulated SPL at various observation points. The maximum error was found to be around 1.2 dB, or 2.2%, which may be due to our point source model assumption and errors in specifying exact boundary conditions. We also verified the SPL produced by a single point source of 1 $Watt$ placed in an open field at 20 °C. The theoretical SPL value at 1 $metre$ from the source is 109.17 dB, which matches the result obtained by *AcouSTO* up to the fourth decimal place.

Distance from source (m)	Observed SPL(dB)	Simulated SPL (dB)	Absolute difference
0	83	81.8	1.2
1.2	75	76.1	1.1
2.4	69	69.9	0.9
3.6	72	72.6	0.6
4.8	68	67.5	0.5
6.0	73	72.1	0.9

Table 2: Observed and simulated SPL (dB) for verifying the setup

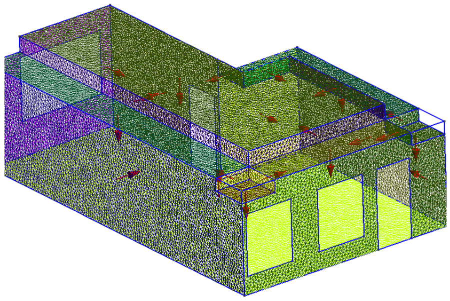

Figure 9: Geometry of the room with the structure of the mesh

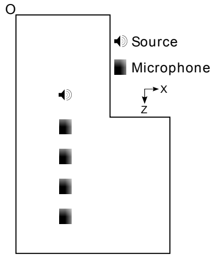

Figure 10: Locations of the source and microphones

We start out by manually creating a mesh with the publicly available *Gmsh* [61] mesh generator. We specify the absorption and reflection coefficients of the boundaries. Subsequently, we set the orientation of the normal vector of each boundary surface, which decides its reaction to an incident sound wave. For the data center boundaries, we choose inward oriented normal vectors, making the walls behave as a cuboidal cavity, whereas for the racks, we create two surfaces per face - one with the normal vector facing inwards and the other with the normal vector facing outwards. Since sound sources are present inside the racks, an inward oriented normal vector is required, and the outward normal vector represents sound waves that propagate through the rack, and are reflected off its outer surface. Lastly, we set the appropriate temperature at different points since the speed of sound is dependent on temperature. For simplicity, each rack is taken to be a single acoustic noise source by lumping together the noise produced by each server. To measure the SPL at points of interest, we specify microphones distributed on a plane in a mesh format generated by using *Gmsh*.

5.2 Simulation Methodology

We simulated the data center using a C++ program that invokes Ansys Icepak and $AcouSTO$. The C++ program simulates the functionality of the CS and the RCs. For the workloads, we consider a bag-of-tasks model, and assume that different servers in the data center run independent tasks. We created 40 combinations of 8 benchmarks with applications in financial analytics, computer vision, and data mining (from the Parsec suite [62] - *blackscholes*, *bodytrack*, $ferret$, $freqmine$, *swaptions*, *canneal*, *streamcluster*, and $swaptionsLarge$) and measured the server power at steady state using a loop ammeter. We randomly distribute the combinations across the servers.

The rated maximum power consumption of the Dell Power Edge R710 server is 415.9 W [63], which results in a maximum rack power consumption of around 6000 W. We discretize this range at intervals of 1000 W and use the algorithm in [64] to find a subset of workloads whose power values add up to 1000 W. More than 99% of the time, we were able to find a subset with an error of $\pm 25W$. Similarly, we discretize the fan flow values in the range $[500, 1500]$ cfm at intervals of 250 cfm. Note that the exact nature of the benchmarks is not important for demonstrating the efficacy of our algorithm. We primarily need to be able to find a set of workloads whose power values add up to multiples of approximately 1 kW. The benchmarks used demonstrate that this is indeed possible in practice.

5.2.1 Utilization Scenarios

We consider 3 scenarios for evaluation, with 50%, 75% and 90% of the servers running, as mentioned in Table 3. For each server utilization scenario, we compare the noise levels at target locations using a uniform workload distribution, a distribution based on the algorithm in [49] (proposed to remove thermal hotspots), and after applying heuristic-I and heuristic-II.

Some thermally efficient workload placement algorithms reduce hotspots by scheduling workloads at the colder servers. A highly cited paper [49] proposed a technique for workload redistribution to maintain uniform temperatures across the data center. We have implemented this technique and evaluated the noise levels obtained using this algorithm. We compare the results obtained using our heuristics with this method.

The CRAC units supply cold air at a temperature of 12.78 °C and with a mass flow rate of 7.21 kg/s. The threshold temperature in the data center is 30 °C. A maximum workload of 8000 $Watt$ can be accommodated by a rack.

The locations of the target points considered for noise reduction are given in Table 4. We assume the origin of our data center to be located in the middle (the floor is on the x-z plane). The corner office is of dimensions $4ft \times 12ft \times 6ft$ ($= 1.2192m \times 3.6576m \times 1.8288m$), with the center at $(5.49, 0, 3.66)$. In the multiple offices case, we consider a set of five offices of equal dimensions located along one wall. The workers are considered to be confined to within 2 feet of the locations mentioned in Table 4.

5.2.2 Estimation of Acoustic Power

We assume that the rack flow rate is divided uniformly among all the 15 servers in a rack. For each server, we estimate the fan and turbulent noise power using Equations 9 and 10, by taking $w = 0.44m$, and $b = 0.086m$ as per the server dimensions [65]. Finally, we calculate the acoustic power level for each server and rack. We consider the noise generated by server fans only, and neglect the noise from the CPU fans, because the CPU fans typically generate much lower noise ($30 - 40$ dB [66]) as compared to server fans ($55 - 59$ dB [65]). Table 6 shows the calculated PWL for various flow rates.

Data center boundaries are made of brick walls (absorptivity, $\alpha = 0.01$) and rack cabinets are made of steel (absorptivity, $\alpha = 0.39$). These values of absorptivity correspond to a sound frequency of $250\,Hz$. We evaluate sound levels at this frequency because cooling fans generate the largest amount of noise around this frequency [67].

	Operational racks (%)	Total workload in the data center (kW)
Scenario-I	50	132
Scenario-II	75	165
Scenario-III	90	200

Table 3: Utilization scenarios

6. RESULTS

Table 5 shows the SPL at various points considered for noise reduction in different scenarios. Table 7 summarizes the noise reduction obtained in the various scenarios. The optimal (l, m) values obtained empirically are (22,22) for all settings, other than the corner office (18,26).

We observe that temperature based distribution is better than uniform distribution only when the data center utilization is low. Also, the heuristics directed at explicitly reducing noise outperform the temperature based distribution method in nearly all the evaluated cases. Our heuristics give upto 16 dB lower noise levels, with at least 5 dB lower noise in 75% of the cases. These results clearly show that the heuristics developed for reducing noise are really needed and temperature based scheduling algorithms are not very effective for this purpose.

In Table 5 (50% utilization scenario: 132 kW total power) we find that the noise reduction obtained using the two heuristics is roughly 8 dB for the corner office, 6 dB for multiple offices and the mobile worker, 9-10 dB for multiple (2) workers, and 1-2 dB for the side wall. When we consider multiple workers, we report noise values for each point. For all other cases, we consider the mean noise value

Name	Center (x, y, z)
Corner Office	(5.49, 0, 3.66)
Multiple Offices	(5.49, 0, 0)
Mobile Worker	(0, 0, 0)
Multiple Workers	(4.6, 0, -2) & (2.1, 0, 2.6)
Side Wall	Plane z = -4.57

Table 4: Various target points and their location

Algorithm	Corner Office	Multiple Offices	Mobile Worker	Multiple Workers I	Multiple Workers II	Wall
50% Utilization Scenario						
Uniform Distribution	83.12	80.72	93.60	86.19	95.71	79.66
Temperature based Distribution	79.96	79.68	95.15	81.84	91.82	78.08
Heuristic-I	74.73	74.60	87.37	76.56	86.81	77.80
Heuristic-II	75.47	79.42	89.40	77.08	85.97	77.92
75% Utilization Scenario						
Uniform Distribution	85.10	91.82	93.93	84.19	96.45	77.19
Temperature based Distribution	81.59	84.13	98.27	89.27	91.36	82.34
Heuristic-I	80.38	83.33	90.33	78.60	87.57	77.96
Heuristic-II	80.50	88.19	90.33	85.18	90.66	76.02
90% Utilization Scenario						
Uniform Distribution	85.30	88.47	102.48	91.65	100.08	82.10
Temperature based Distribution	89.04	89.23	105.39	86.75	103.05	89.04
Heuristic-I	84.80	82.63	98.40	90.20	94.72	80.89
Heuristic-II	79.64	82.24	98.65	90.50	87.02	80.23

Table 5: Sound pressure level (dB) at various points of interest

Flow rate (cfm)	PWL (dB)
500	77.5
750	86.32
1000	92.5
1250	97.5
1500	101.38

Table 6: Sound power level for various flow rates

	Algorithm	Corner Office	Multiple Offices	Mobile Worker	Multiple Workers I	Multiple Workers II	Wall
50% Utilization	Heuristic-I	8.39	6.12	6.23	9.63	8.90	1.86
	Heuristic-II	7.65	1.30	4.20	9.11	9.74	1.74
75% Utilization	Heuristic-I	4.72	8.49	3.60	5.59	8.88	-0.77
	Heuristic-II	4.60	3.63	3.60	-0.99	5.79	1.17
90% Utilization	Heuristic-I	0.50	5.84	4.08	1.45	5.36	1.21
	Heuristic-II	5.66	6.23	3.83	1.15	13.06	1.87

Table 7: Acoustic noise reduction (dB) with heuristics for various points of interest

at all the grid points. It is difficult to reduce noise for the side wall because we need to move a lot of jobs away from it, and this starts violating temperature constraints at distant nodes. Heuristic-I is marginally better than Heuristic-II for this case. This is because when the data center utilization is lower, the extracted workload can be accommodated in a small number of servers. Heuristic-I steps up fan speeds of the farthest servers one by one until the extracted workload is accommodated. However, with Heuristic-II, workloads are added to all the distant servers in inverse order of their distances. Thus the noise is lower at the target location with Heuristic-I.

Let us now consider the 75% utilization scenario (165 kW total power) Here, the gains are lower (but still significant). Since we have a larger workload, and lesser room to move tasks, we reach maximum power and temperature thresholds easily. Here also Heuristic I is better.

However, we see a reverse effect with 90% utilization (see Table 5). The gains are between 2-13 dB, and Heuristic-II is better. In specific, for the case of 2 workers, the noise reduces from 100.1 dB with the uniform distribution, and 103.1 dB with the temperature based distribution, to 87 dB, which is significant. The reason for the increased variance is that the overall noise levels are higher, and any redistribution of the workloads can significantly reduce noise if it is effective. For high power levels (200 kW), most of the servers have already reached the power threshold. Hence in Heuristic-I the fan speed of almost all the remote racks are increased. Heuristic-II however does a more intelligent redistribution. Thus we may not need to increase the fan speed of all the remote racks, which results in lower noise levels than Heuristic-I. The online appendices [44] contain a detailed set of results: SPL profiles for each scenario, temperature distribution, and power consumption of different combinations of benchmarks.

6.1 Example

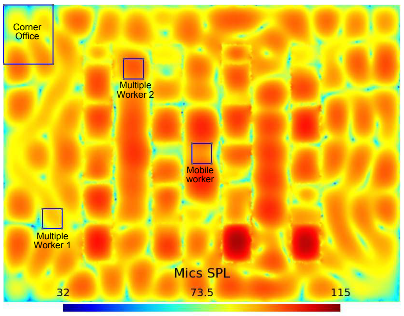

Figure 11: SPL distribution for the uniform workload distribution at 50% utilization

Let us consider the case of the corner office with Heuristic I (50% utilization). Figure 11 shows the SPL distribution at the beginning, and Figure 12 shows the SPL distribution

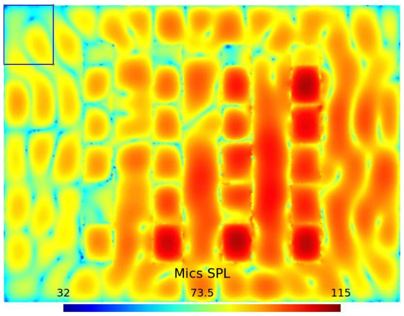

Figure 12: SPL distribution after applying heuristic-I for the corner office at 50% utilization

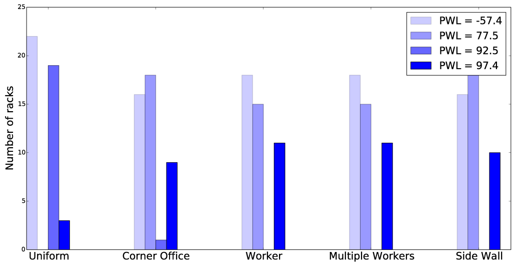

Figure 13: Rack count distribution with PWL for heuristic-I at 50% utilization

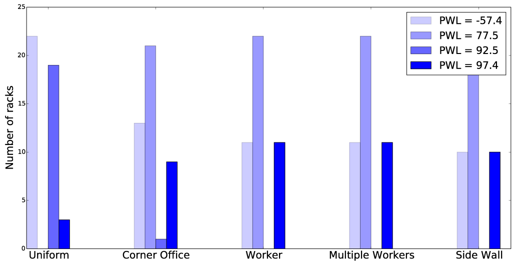

Figure 14: Rack count distribution with PWL for heuristic-II at 50% utilization

after redistribution. We can clearly see that jobs have moved away from the corner office (marked with a rectangle).

6.2 Comparison of the Heuristics

For 50% utilization, Figures 13 and 14 show the rack count distribution versus PWL values for heuristic-I and heuristic-II, respectively. We report the PWL of idle servers as -57 dB[2] (background noise). Initially each running rack was assigned a flow rate of 1000 cfm (using the pre-computed data, which corresponds to 92.5 dB), but after CFD analysis the temperature at some points was found to exceed the threshold, so the flow rate of racks near them was increased by 250 cfm. That is why we get 3 racks with a PWL value of 97.4 dB. After workload redistribution, the acoustic power of the racks close to the points of interest got reduced to 77.5 dB and for racks farther away, it increased to 97.4 dB. Also, in the redistribution process some new racks are assigned workloads, which were initially idle.

We can clearly see why Heuristic I performs better. It has more idle racks. The proportional weighting method tries to distribute jobs in every rack that is farther away, and this is a bad strategy. It is better to have some really noisy servers that are far away (as we can see in the graphs).

6.3 Cost - Benefit Analysis

As discussed in Section 1, a pair of good noise canceling headphones cost 500 USD, dedicated server rooms cost tens of thousands of dollars, and rack enclosures cost thousands of dollars. In comparison, the only additional hardware needed to implement our method is the rack controller. Each data center already has a job manager, which assigns incoming workloads to the different servers. The CS takes the role of the job manager. The RC can be implemented on an Arduino board costing less than 30 USD. We have already discussed the performance overhead in Section 3, which is $< 1\%$ in the general case. If there are additional Quality of Service requirements, we can add a factor in the workload migration algorithm that depends on the memory footprint of the VM to be migrated. This will reduce the number of migrations in the case of very memory intensive applications and ensure that the Service Level Agreement is not violated.

There are two more issues that we would like to address. Most thermally efficient workload placement algorithms try to homogenize the thermal profile such that hotspots are not created. Our algorithm too ensures that the thermal threshold is never exceeded, and hence the cost of cooling would remain the same. Also, the fan power consumption is much lower than the CPU power consumption in a data center. Therefore, for all practical purposes, it is neglected.

7. CONCLUSION

In this paper, we proposed a methodology to model the acoustic profile of data centers using *AcouSTO* and Ansys Icepak. We made several non-trivial design choices along the way that helped us to quickly attain convergence in both Icepak and *AcouSTO*. We verified our setup with real world data.

A temperature based algorithm does not provide any reliable reduction in noise as compared to a uniform distribution. Hence we explored two simple noise aware workload redistribution heuristics that can reduce the noise at points of interest by 6-10 dB in most cases, which is significant. Without applying our technique, the ambient noise was in the range of 85-95 dB, which is not permissible according to OSHA or European regulations. Our approaches were able to reduce the noise level and make the system conformant with OSHA regulations. All of this was achieved with a maximum performance loss limited to 1.6%. Developing more sophisticated noise reduction algorithms is a part of the future work.

8. REFERENCES

[1] T. ASHRAE, “9.9 (2011) thermal guidelines for data processing environments–expanded data center classes and usage guidance,” *Whitepaper prepared by ASHRAE technical committee (TC)*, vol. 9, 2011.

---

[2] Sound levels below the human threshold of hearing ($< 10^{-12}$ W) have negative PWL in dB. Theoretically, if the sound power is zero, the PWL tends to negative infinity.

[2] R. H. Lyon and A. E. Bergles, “Noise and cooling in electronics packages,” *Components and Packaging Technologies, IEEE Transactions on*, vol. 29, no. 3, pp. 535–542, 2006.

[3] C. E. Peterson. (2011) Dell enterprise acoustics. [Online]. Available: http://www.dell.com/downloads/global/products/pedge/en/acoustical-education-dell-enterprise-white-paper.pdf

[4] OSHA. Noise exposure computation occupational safety and health standards 1910.95 app a. [Online]. Available: https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=9736

[5] S. G. Kujawa and M. C. Liberman, “Adding insult to injury: cochlear nerve degeneration after “temporary noise-induced hearing loss,” *The Journal of Neuroscience*, vol. 29, no. 45, pp. 14 077–14 085, 2009.

[6] S. Klitzman and J. M. Stellman, “The impact of the physical environment on the psychological well-being of office workers,” *Social Science & Medicine*, vol. 29, no. 6, pp. 733–742, 1989.

[7] M. Kassner. (2014) Data centers may be hazardous to your hearing. [Online]. Available: http://www.techrepublic.com/article/data-centers-may-be-hazardous-to-your-hearing/

[8] T. S. Dope. (2012) Hearing loss from data center work. [Online]. Available: http://boards.straightdope.com/sdmb/showthread.php?t=672967

[9] Edugeek. (2011) Working environment - server noise. is this affecting my hearing? [Online]. Available: http://www.edugeek.net/forums/general-chat/86747-working-environment-server-noise-affecting-my-hearing.html

[10] P. Thibodeau. (2007) That sound you hear? the next data center problem. [Online]. Available: http://www.infoworld.com/article/2642787/operating-systems/that-sound-you-hear--the-next-data-center-problem.html

[11] (2010) Web hosting talk, data center noise cancellation. [Online]. Available: http://www.webhostingtalk.com/showthread.php?t=958922

[12] Health and S. Executive. Hse - noise: Frequently asked questions. [Online]. Available: http://www.hse.gov.uk/noise/faq.htm#q5

[13] (2014) Are noise cancelling headphones worth it? [Online]. Available: http://www.forbes.com/sites/geoffreymorrison/2014/02/21/are-noise-cancelling-headphones-worth-it/#1c3222b97e26

[14] S. Z. Jalilzadeh, M. Asgari, and M. E. Sadeghi, “A quantitative noise survey in a general room–based data center,” *International Journal of Occupational Hygiene*, vol. 6, no. 1, pp. 16–22, 2014.

[15] S. Sommerfeldt, M. Maughan, J. Daily, J. Esplin, Z. Collins, and M. Shaw, “Acoustical analysis of active control in the server room of a c7 data centers colocation facility,” *Feasibility Report, Brigham Young University*, 2010.

[16] (2015) Datacenter hearing protection. [Online]. Available: https://www.reddit.com/r/networking/comments/1wuq9i/datacenter_hearing_protection/

[17] (2015) Hearing protection for data centre work. [Online]. Available: https://www.reddit.com/r/sysadmin/comments/2xckqr/hearing_protection_for_data_centre_work/

[18] (2015) Bluetooth headset for use in a data center. [Online]. Available: https://www.reddit.com/r/networking/comments/3l28dy/bluetooth_headset_for_use_in_a_data_center/

[19] (2015) What type of headset works well in data centers? [Online]. Available: https://www.reddit.com/r/sysadmin/comments/316tve/what_type_of_headset_works_well_in_data_centers/

[20] (2015) Do noise cancelling headphones lessen hearing damage? [Online]. Available: https://www.reddit.com/r/audiophile/comments/2t19ma/do_noise_cancelling_headphones_lessen_hearing/

[21] (2015) Psa: Your beats noise canceling headphones aren’t proper earpro. [Online]. Available: https://www.reddit.com/r/guns/comments/28wn8b/psa_your_beats_noise_canceling_headphones_arent/

[22] (2015) Google data center. [Online]. Available: https://www.reddit.com/r/Cyberpunk/comments/19jvhw/google_data_center/

[23] Health and S. Executive. Hse - noise: Hearing protection. [Online]. Available: http://www.hse.gov.uk/noise/hearingprotection.htm

[24] OSHA. (2014) Safety and health topics | occupational noise exposure. [Online]. Available: https://www.osha.gov/SLTC/noisehearingconservation/

[25] P. Artman and C. E. Peterson. Advanced thermal control: Optimizing across environments and power goals. [Online]. Available: http://www.dell.com/downloads/global/products/pedge/advanced_thermal_control_whitepaper.pdf

[26] HP. Hp proliant microserver gen8. [Online]. Available: http://www8.hp.com/us/en/products/proliant-servers/product-detail.html?oid=5379860

[27] IBM. Ibm system x3100 m4 - features and benefits. [Online]. Available: http://www-03.ibm.com/systems/in/x/hardware/tower/x3100m4/features.html

[28] Fujitsu. Fujitsu server primergy tower systems - fujitsu global. [Online]. Available: http://www.fujitsu.com/global/products/computing/servers/primergy/tower/

[29] APC. Server cabinets - virtually silent, sound proofed server cabinet solutions. [Online]. Available: http://www.kellsystems.com/

[30] GIZMAC. Xrackpro2 noise reducing server racks rackmout enclosure cabinets. [Online]. Available: http://www.xrackpro.com/

[31] Orion. Acoustic racks. [Online]. Available: http://www.rackcabinets.co.uk/cabinets/acoustic-racks.html

[32] A. Products. Silence server noise - quiet acoustic rack enclosures, soundproof materials. [Online]. Available: http://www.acoustiproducts.com/en/index.asp

[33] Serverfault. How can you reduce ambient noise from servers? [Online]. Available: http://serverfault.com/questions/265540/how-can-you-reduce-ambient-noise-from-servers

[34] (2012) What can be done to control noise levels in data centers? [Online]. Available: http://www.idganswers.com/question/3607/what-can-be-done-to-control-noise-levels-in-data-centers

[35] (2015) Making noise about a quiet data center. [Online]. Available: http://blogs.cisco.com/ciscoit/b-dc-03112015-making-noise-about-a-quiet-data-center

[36] (2014) Quieter data center. [Online]. Available: http://www.pducables.com/kens-korners/2014/12/quieter-data-center/

[37] (2014) How to reduce data center noise? [Online]. Available: https://storageservers.wordpress.com/2014/01/10/how-to-reduce-data-center-noise/

[38] (2014) To lower the noise levels in data centers, raise the temperature. [Online]. Available: http://www.itwatchdogs.com/environmental-monitoring-news/data-center/to-lower-the-noise-levels-in-data-centers,-raise-the-temperature-617409

[39] (2015) Noise reduction - asetek. [Online]. Available: http://www.asetek.com/data-center/technology/data-center-benefits/noise-reduction/

[40] S. Wynne. (2011) How are you handling excessive noise in the data center? [Online]. Available: http://blog.rackmountsolutions.net/soundproof-server-rack/how-are-you-handling-excessive-noise-in-the-data-center/

[41] P. Thibodeau. (2007) That sound you hear? the next data center problem | computerworld. [Online]. Available: http://www.computerworld.com/article/2542868/data-center/that-sound-you-hear--the-next-data-center-problem.html

[42] M. F. Bari, R. Boutaba, R. Esteves, L. Z. Granville, M. Podlesny, M. G. Rabbani, Q. Zhang, and M. F. Zhani, “Data center network virtualization: A survey,” *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 2, pp. 909–928, 2013.

[43] U. Iemma and V. Marchese, “*AcouSTO, Acoustics Simulation TOol v1.6*,” http://acousto.sourceforge.net, 2014.

[44] (2016) Appendix of noise aware scheduling in data centers. [Online]. Available: http://www.cse.iitd.ac.in/~srsarangi/tr/noise/appendix.pdf

[45] I. Acentech. (2014) Acoustics & vibration fundamentals. [Online]. Available: http://www.acentech.com/resources/fundamentals.html

[46] A. G. Carol. (2014) The relationship of voltage, loudness, power and decibels. [Online]. Available: http://www.gcaudio.com/resources/howtos/voltageloudness.html

[47] R. Ayoub and T. S. Rosing, “Cool and save: cooling aware dynamic workload scheduling in multi-socket cpu systems,” in *ASP-DAC, 2010*, 2010, pp. 891–896.

[48] J. D. Moore, J. S. Chase, P. Ranganathan, and R. K. Sharma, “Making scheduling” cool”: Temperature-aware workload placement in data centers.” in *USENIX, General Track*, 2005, pp. 61–75.

[49] 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,” *Internet Computing, IEEE*, vol. 9, no. 1, pp. 42–49, 2005.

[50] N. Vasic, T. Scherer, and W. Schott, “Thermal-aware workload scheduling for energy efficient data centers,” in *International Conference on Autonomic Computing*, 2010, pp. 169–174.

[51] VMWare. (2013) vmotion. [Online]. Available: http://www.vmware.com/products/vsphere/features/vmotion

[52] R. Ayoub, S. Sharifi, and T. S. Rosing, “Gentlecool: Cooling aware proactive workload scheduling in multi-machine systems,” in *DATE, 2010*, 2010, pp. 295–298.

[53] S. Akoush, R. Sohan, A. Rice, A. W. Moore, and A. Hopper, “Predicting the performance of virtual machine migration,” in *MASCOTS, 2010*, 2010, pp. 37–46.

[54] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, “Live migration of virtual machines,” in *Symposium on Networked Systems Design & Implementation-Volume 2*, 2005, pp. 273–286.

[55] X. Jiang, F. Yan, and K. Ye, “Performance influence of live migration on multi-tier workloads in virtualization environments,” *IARIA CLOUD*, pp. 72–81, 2012.

[56] A. Icepak, “14.0 documentation,” *ANSYS Inc*, 2012.

[57] A. Fluent, “12.0 documentation,” *Ansys Inc*, 2009.

[58] D. C. Journal. (2013) Raising data center power density. [Online]. Available: http://www.datacenterjournal.com/it/raising-data-center-power-density/

[59] D. C. U. Group. (2013) Data center users’ group. [Online]. Available: http://datacenterug.ning.com/

[60] E. N. Power. (2013) Average power density per rack holds at 5.94 kw. [Online]. Available: http://www.emersonnetworkpower.com/en-US/Solutions/infographics/Pages/graphicoftheweek18.aspx

[61] C. Geuzaine and J.-F. Remacle, “Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities,” *International Journal for Numerical Methods in Engineering*, vol. 79, no. 11, pp. 1309–1331, 2009.

[62] C. Bienia, S. Kumar, J. P. Singh, and K. Li, “The parsec benchmark suite: characterization and architectural implications,” in *PACT*, 2008.

[63] I. Dell. (2011) Dell poweredge r710 power and performance data sheet. [Online]. Available: http://www.dell.com/downloads/global/products/pedge/en/Dell%20PowerEdge_R710_570W_Energy_Star_Data_Sheet_new.pdf

[64] C. E. Leiserson, R. L. Rivest, C. Stein, and T. H. Cormen, *Introduction to algorithms.* MIT press, 2001.

[65] I. Dell. (2011) Poweredge r710 technical guide. [Online]. Available: http://www.dell.com/downloads/global/products/pedge/en/server-poweredge-r710-tech-guidebook.pdf

[66] (2015) Spire sound level measurements: Which heatsink is quietest? [Online]. Available: http://www.frostytech.com/articleview.cfm?articleid=2521&page=3

[67] (2005) Acoustics/noise from cooling fans. [Online]. Available: http://en.wikibooks.org/wiki/File:Noisespectrum.gif