

Running Servers around Zero Degrees

Mikko Pervilä
Department of Computer Science
P.O. box 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki, Finland
pervila@cs.helsinki.fi

Jussi Kangasharju
Helsinki Institute for Information Technology
Department of Computer Science
P.O. box 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki, Finland
jakangas@cs.helsinki.fi

ABSTRACT

Data centers are a major consumer of electricity and a significant fraction of their energy use is devoted to cooling the data center. Recent prototype deployments have investigated the possibility of using outside air for cooling and have shown large potential savings in energy consumption. In this paper, we push this idea to the extreme, by running servers outside in Finnish winter. Our results show that commercial, off-the-shelf computer equipment can tolerate extreme conditions such as outside air temperatures below -20°C and still function correctly over extended periods of time. Our experiment improves upon the other recent results by confirming their findings and extending them to cover a wider range of intake air temperatures and humidity. This paper presents our experimentation methodology and setup, and our main findings and observations.

Categories and Subject Descriptors

B.8.1 [Performance and Reliability]: Reliability, Testing, and Fault-Tolerance

General Terms

Experimentation, Reliability

Keywords

Sustainable computing, cooling, empirical system reliability

1. INTRODUCTION

According to an analysis published by HP in February 2009 [3], data centers would be the sixth-largest consumer of electricity if they were classified as a separate industry. By this analysis, research concentrating on reducing data center power consumption should show major benefits from both the green computing and financial viewpoints. In difference to home equipment, whose heat emissions are beneficial to indoor heating in cold environments, the heat generated by

data centers can be collected only in the most recently designed architectures. As collecting the heat seems difficult, we turn our focus into preventing it.

Using outside air to cool the data center can yield energy savings from 40% to 67%, according to HP and Intel [1] respectively. We have begun a small scale experiment to verify the claims of Intel and HP, and also to extend their results to our environment.

Using the naturally cold winter in Finland, we seek to understand in how extreme conditions COTS and server equipment can be operated. During the winter of 2009-2010, outside temperatures of -22°C were measured by the Finnish Meteorological Institute. While these measurements were taken in Southern Finland, much more extreme conditions occur in the Northern parts.

If we can bring the server equipment to tolerate North European conditions, we have shown that Intel's results from New Mexico and HP's from North East England can be extended to most parts of the globe. In addition to extending the area of feasibility, we are also interested in the effects of allowing the intake air a much wider range of variation. If the equipment can tolerate both long- and short-term fluctuations, we could eschew any conditioning of the intake air, including temperature and humidity stabilization.

This paper is structured as follows. In Section 2 we review related work. Section 3 presents our research questions and methodology. In Section 4 we present our main results and findings from our experiment. Section 5 discusses the implications of our results in relation to existing work and presents directions for future work. Finally, Section 6 concludes the paper.

2. RELATED WORK

As far as we have been able to ascertain, the closest works stem from industry white papers explaining the current state of the art of data center cooling. HP has analyzed the magnitude of the industry and reveal some details of their Wynyard data center [3]. A good entry point into current data center cooling solutions is provided by the summary article from Intel's Digital Enterprise Group [5]. It is further supplemented by the proof-of-concept air economizer cited [1] above. Interestingly, Intel's previous report [2] has argued convincingly against air economizers.

In addition to white papers, a more distant relative lies in the field of computer overclocking. A number of competitions have focused on driving COTS motherboards and CPUs well below their normal operational parameters by

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GreenNetworking 2010, New Delhi, India
Copyright 2010 ACM 978-1-4503-0196-1/10/08 ...\$10.00.

employing extreme cooling solutions ranging from liquid nitrogen to geothermal cooling [6].

Our work differs from the white papers in the direct use of intake air with very dynamic temperatures and humidities. Most of the cited cooling solutions assume stable or near-stable input temperatures by conditioning the cooling medium with an intermediary step. Intel’s air economizer article is the closest related work. We seek to extend their previous results by letting the intake air conditions vary in a significantly wider range.

3. FEASIBILITY, RESEARCH QUESTIONS, AND METHODOLOGY

First, the major research question of this work remains whether unconditioned outside air is a feasible cooling solution. If Intel’s proof of concept can be extended to our Northern climate, this would indicate that newly built or *Greenfield* data centers can do without air conditioning units.

Second, the equipment failure rate affects both financial and green endeavors. Financial endeavors are by definition mainly interested about the price. If the outside air technique is feasible but causes a higher equipment failure rate than by using familiar air conditioning, the projected costs must be carefully considered. If the failure rate rises only a little or not at all, replacement costs must be balanced with the purchase and energy costs of air conditioning. For green endeavors, this equation becomes trickier, as the comparison would need to factor in the amount of resources consumed by the manufacture and logistics of new components.

Third, a minor research question concerns which components will fail first. In what is called *industry tribal knowledge*, subjective viewpoints about the humidity or cold breaking component X run rampant. If the extreme temperature and humidity shifts indeed cause certain components to regularly fail, we should be able to detect this as a common-cause failure on multiple hosts nearly simultaneously.

Finally, we deliberately included some hosts from a series of workstations that we already knew to be unreliable. Their problems have to do with the hardware temperatures elevating due to bad air flow circulation. We were interested to see in how far the cooler outside conditions would alleviate the known problems, if at all.

The test setup was taken in two consecutive phases. To begin with, a prototype test was undertaken to ascertain that a real measurement was worth the trouble. After the prototype test completed successfully, a normal phase of testing was setup and started. Both phases are described in the following sections.

3.1 Prototype and Normal Phases

For the weekend from Friday, Feb. 12th to Mon 15th, we ran a generic PC sandwiched between two hard plastic boxes. The boxes did not really impede air flow or contain any heat, but served to protect against snow reaching the computer internals and melting into water. During the test, we monitored both hard drive S.M.A.R.T. readings and the internal temperature sensors through Linux’ *lm-sensors* package. The local meteorological measurement unit located in the building next to ours recorded temperatures as low as -10.2°C for the weekend, with an average of -9.2°C .

The prototype survived the test, remaining operational for the whole weekend. Readings recorded by *lm-sensors* showed

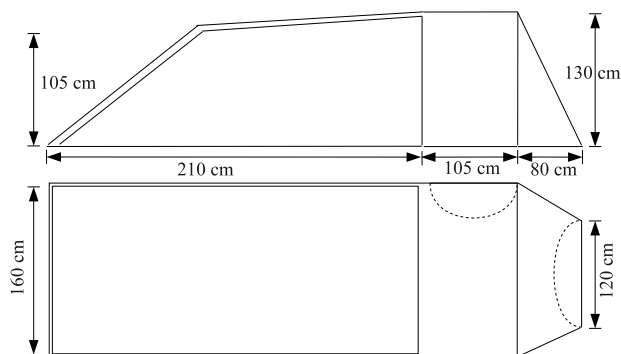


Figure 1: Schematic for tent shielding the computer hardware from rain and snow.

that the CPU had been operating in temperatures as low as -4°C . While this result is surprising, similar readings have been noted by the overclocking communities.

We were forced to stop the prototype test the following week due to external constraints: the two plastic boxes that we had borrowed for this test had to be returned. Nevertheless, we deemed the test a success and scheduled a more extended test to begin the following Friday (Feb. 19th).

Operating on a shoestring budget, we asked for permission from the department’s IT staff to reuse old workstations destined for recycling. Due to local tax regulations, hardware removed from usage can not be given out to employees or donated. As things were, most of the PCs were still fully operational. In addition, we procured some workstations which were considered unreliable, mentioned above, and also a batch of seven rack-mountable servers. The computer equipment is more thoroughly described in Section 3.4.

The main problem to overcome was how to shield the computers from water or, in our case, snow. Many solutions were considered, but in the end we opted for a lightweight tent aimed for three-person camping trips.

3.2 The Tent

In order to maximize air flow, the ideal protective construction would be something resembling an outside storage shed with only minimal cover, e.g., of the kind that hardware stores use for construction materials. Due to time, location, and resource constraints, we were forced to compromise with the protective solution.

We located the tent on the roof terrace of the Department of Computer Science¹. The location is very good, since a power outlet designed for outdoors use is positioned just next to the site, and access to the roof terrace is monitored by video surveillance.

A diagram of the tent is presented in Fig. 1. When erected, the tent consists of a roughly tube-shaped, double-layered structure of polyester fabric. Soon after installing the equipment, we were forced to make repeated modifications to the structure, as the tent proved surprisingly good at retaining heat. Later changes include removing the inner layer.

There are four main factors affecting the inside temperature of the tent. These are, in order of importance, the

¹An hourly webcam image of the terrace (with the tent) is available at <http://www.cs.helsinki.fi/Exactum-kamera/>

outside air temperature, sunlight and wind speeds, power draw of equipment, and which tent flaps are open.

We tried to reduce direct sunlight hitting the tent fabric by installing a partial, reflective foil cover of the same material used in first-aid rescue sheets. The purpose of these sheets is to keep incapacitated patients warm in cold environments. For our purposes, the reflective cover measurably decreases the internal temperatures, as we later show in Section 4.1.

Wind speed remains somewhat of a problem. As the tent is designed to actually block out the wind chill effect, we have tried to modify the structure by cutting open the internal fabric and removing the protective tarpaulin from the bottom. As our terrace is elevated higher than the roof, some cool air is able to circulate through the floor and into the tent. This way, the electronic equipment is still protected by the outer fabric, but the heat dissipation factor is as high as possible.

The last modification to normal operation was to let the outer front door remain in a half-open position. This seems to improve air flow from the back and through the bottom of the tent on days with even a moderate amount of wind.

3.3 Measurements Taken

Following data center best practices, our analysis concentrates mainly on the temperatures and relative humidities (RHs) surrounding the electronic equipment. These measurements are separated into data gathered from inside and outside of the tent.

For outside data gathering, we were fortunate enough to receive access to the Department of Physics' weather station located just outside our building. The station is known as SMEAR III and is co-operated with the Finnish Meteorological Institute, who provide data gathering services for nearly all interested parties.

Inside the tent, we used a Lascar EL-USB-2-LCD data logger as the sensor device. Measurement error for the unit is $\pm 0.5^\circ\text{C}$, $\pm 3.0\%$ RH typically and $\pm 2^\circ$, $\pm 6.0\%$ RH maximum. Data loggers of this type are used by companies transporting edibles, for example. The advantage of the data logger is that it is machine readable, although only by manually inserting the device into a USB port. Due to this, we have been forced to remove a number of outliers in the measurements caused by removing the data logger and carrying it indoors. These outliers have been removed from the graphs.

Finally, in order to gauge the amount of heat generated by the hardware we used a Technoline Cost Control unit. It has recently been tested [4] by local colleagues and found to perform very admirably given its price. The unit was used to measure normal and maximum power draw of the server hardware. The total load of the tent remained below 1100 W during normal operation, including the mechanical fan installed later.

3.4 Hardware

In total, we operate 19 computers in three form factors. The first set is from a small vendor ("A") using COTS hardware to build "cloned" desktop machines. These machines are built in medium tower cases and contain two hard drives formed into a Linux multiple devices software mirror. The second is a large vendor ("B") producing mass-manufactured small form factor PCs as workstations. Only a single hard drive can fit in the case due to the form factor. The third

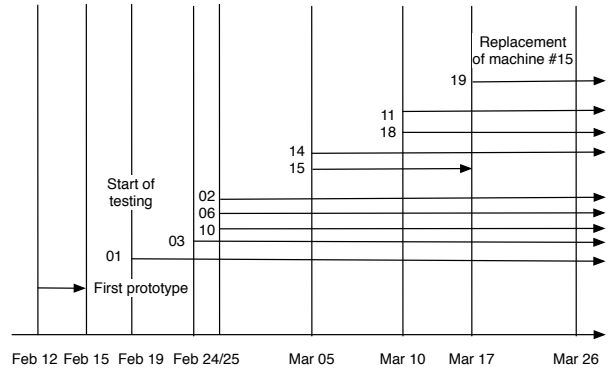


Figure 2: Dates of when servers were installed.

is a large vendor ("C") offering rack mounted heavy duty servers in the 2U form factor. There are five hard drives in each, two of which compose a hardware mirror, and the remaining three a stripe set with parity.

In order to distinguish faults caused by our chosen conditions, a control group was installed into the department's basement. Computers are thus installed pairwise so that identical units are placed into the control group in the basement and the test group in the tent on the terrace. The department's basement doubles as a protection shelter for staff, thus, the control group operates in a sparsely furnished environment with stable, office-type air conditioning. The operating conditions are therefore well within specifications.

For our tests we installed ten hosts from vendor A, four from B, and four from C, yielding a symmetric nine hosts in the basement and nine in the tent. A timeline of when servers were added is depicted in Fig. 2. The numbering refers to the servers on the terrace. (The 19th server was used to replace one server that partially failed during the test; see Section 4.2.)

3.5 Load

All servers execute a synthetic workload, which consists of packing a Linux kernel source directory with the standard tar and bzip2 archive programs. After packing, each compressed tarball is verified by calculating its md5sum hash function and comparing the result with an initial value calculated before installation. If the results differ, the packed tarball is stored. If not, the tarball is overwritten in the next cycle.

Each host executes its synthetic load every 20 minutes. In order to avoid synchronization, some fuzz is added to the starting phase: each host sleeps for 0 to 119 seconds before commencing the archival process.

At the time of writing, we have collected results from a total of 119516 executed runs from the 19 hosts executing the synthetic load. Of these results, six have been found faulty and examined more thoroughly in Section 4.2.2.

Some load is additionally generated by the monitoring host, which recovers all calculated md5sums and data gathered from the local sensors every 20 minutes. The transfer is done using public-key authentication through an OpenSSH tunnel, and new files are transferred by the rsync program.

4. CURRENT KNOWLEDGE

At the time of writing, the first server installed has been operating for three months with only very minor glitches, which are described in Section 4.2. The last of the hosts was installed March 13th, meaning two and a half months of operation at the time of writing. Of the eighteen hosts installed initially, one has encountered two transient system failures, and after having been taken indoors, has remained in stable operation. A failure rate of 5.6% may seem harsh initially, but Intel has reported a comparable rate of 4.46% during their experiment [1].

Despite the relatively small number of transient system failures we have learned a number of lessons. In the following section, we take a look into the development of temperatures and relative humidities inside the tent and review the faults encountered more thoroughly.

4.1 Temperature and Humidity

Figure 3 shows the evolution of the temperatures both inside and outside the tent during the experiment. The outside temperature is from the SMEAR III station and the inside temperature is from the Lascar data logger. The figure also shows a few key events, marked R, I, B, and F, which are explained below. Because the Lascar data logger arrived late, tent-internal temperature and humidity data from the early parts of the experiment are missing.

Figure 3 shows a number of changes in the tent’s internal temperature. Major operations undertaken to limit the heat retained by the tent fabric have been marked with letters beneath the figure. In order of appearance, the coding is R for installation of the reflective foil cover, I for removal of the inner tent, B for partial removal of the bottom tarpaulin, and F for installation of a common tabletop motorized fan.

What is clearly visible towards the end of the graph is the week-long heat period encountered in Helsinki during the third week of May. Outside temperatures rose quickly to relatively high temperatures of 20 – 25° C, causing a subsequent elevation to 25 – 30° C in the tent’s internal temperature. After that one week of unusually warm weather, Helsinki has now normalized to much more usual temperature levels.

Relative humidities are shown in 4. Because RH values are defined by their ambient temperatures, the figure is somewhat difficult to analyze. What is visible, however, is that the tent has been able to retain more stable relative humidities than outside air, although sharp temperature drops are still visible. As we increase air flow to lower the inside temperatures, the humidity also begins to vary more intensely.

4.2 Faults Encountered

During the full test run thus far, we have encountered four cases of system failures and six cases of miscalculated synthetic loads. Two of the system failures can be written off as being caused by hardware faults present even before the test. The other two are more difficult to explain. We will describe the system failures first, and then move on to the synthetic loads.

4.2.1 System Failures

The first problem was discovered in the host that has been in continuous operation for the longest span of time. This host has encountered outside temperatures of –22° C. After the initial period in the most extreme cold, the host’s

lm-sensors started to malfunction. Before the failure, the motherboard’s sensor chip had reported CPU temperatures of below –4° C, followed by clearly erroneous readings of –111° C. After detecting the anomaly, we tried to redetect the sensor chip with hopes of resetting its internal readings. Instead, the opposite resulted, and the sensor chip ceased to be detected at all. After a week, we risked a warm system reboot, which caused the sensor chip to work again. It is difficult to say if the sensor hardware or its accompanying kernel modules were the root of the fault. However, no further problems have been detected on this host.

Host #15 from vendor B encountered a system failure on Saturday, March 7th at 04:40 (a.m.). The host in question was running in the tent. After an inspection and reset on the following Monday, no cause for the failure could be determined. The failure was initially marked as transient and the host resumed normal operations in the tent.

Unfortunately, the same host encountered another failure on Wednesday, March 17th at 12:20 (p.m.). The host was reset in outside conditions but could not resume normal operations. It was again taken inside for an inspection. A standard Memtest86+ run caused another system failure within a few hours. After this, the host was left to operate in an indoors environment. No further failures have been detected on the host. Note that this host was from the series we already knew to have frequent defects. We must thus concur that during this test, the system series known to be defective operated no better in outside conditions. None of the hosts in the control group have failed yet, and neither has the new host that replaced host#15 in the tent.

Finally, the two problems that we can explain relate to the network infrastructure. In order to share the network connectivity we employed two 8-port network switches known to contain cosmetic errors, i.e., an annoying whining sound during normal operation. Both of the switches encountered a failure after a week or so of tent operation. After some testing, the remaining switch that had never been used for this test manifested an identical failure state. We can therefore conclude that the problem is inherent in these individual switches and existed even before we began our test.

4.2.2 Wrong hashes

Our synthetic load has encountered problems in 6 out of a total of 119516 test runs. The ratio of tent/basement errors is as follows: two hosts placed outside reported one wrong md5sum hash each, and one host placed inside reported four wrong hashes. Of the problematic archive files, we were able to recover the two most recent ones.

While inspecting the tarballs with the bzip2recover utility, it became clear that only a single one of the 396 bzip2 compression blocks had been corrupted. No errors have been reported by the file system or the kernel, and the hard drives have passed their S.M.A.R.T. long test runs. The current conjecture of a failure cause is therefore a memory error. All three hosts that have reported faulty hashes contain memory chips without error-correcting parities.

By calculating the size of the source directory to be compressed, the average block size of the compressed tarball, and the amount of cycles we have estimated the amount of memory pages read and written to lie in the ballpark of 14 billion. If the estimate is correct, and the six faulty archives are caused by a single memory page fault each, the failure ratio is around one in 2,5 billion.

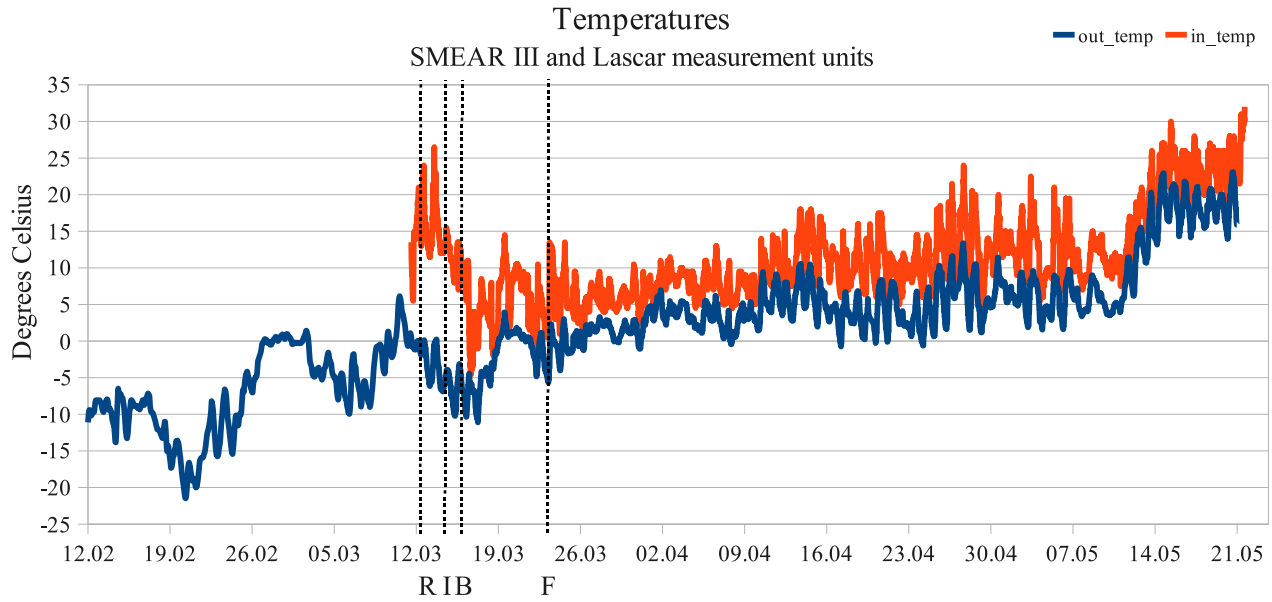


Figure 3: Temperatures outside and inside the tent.

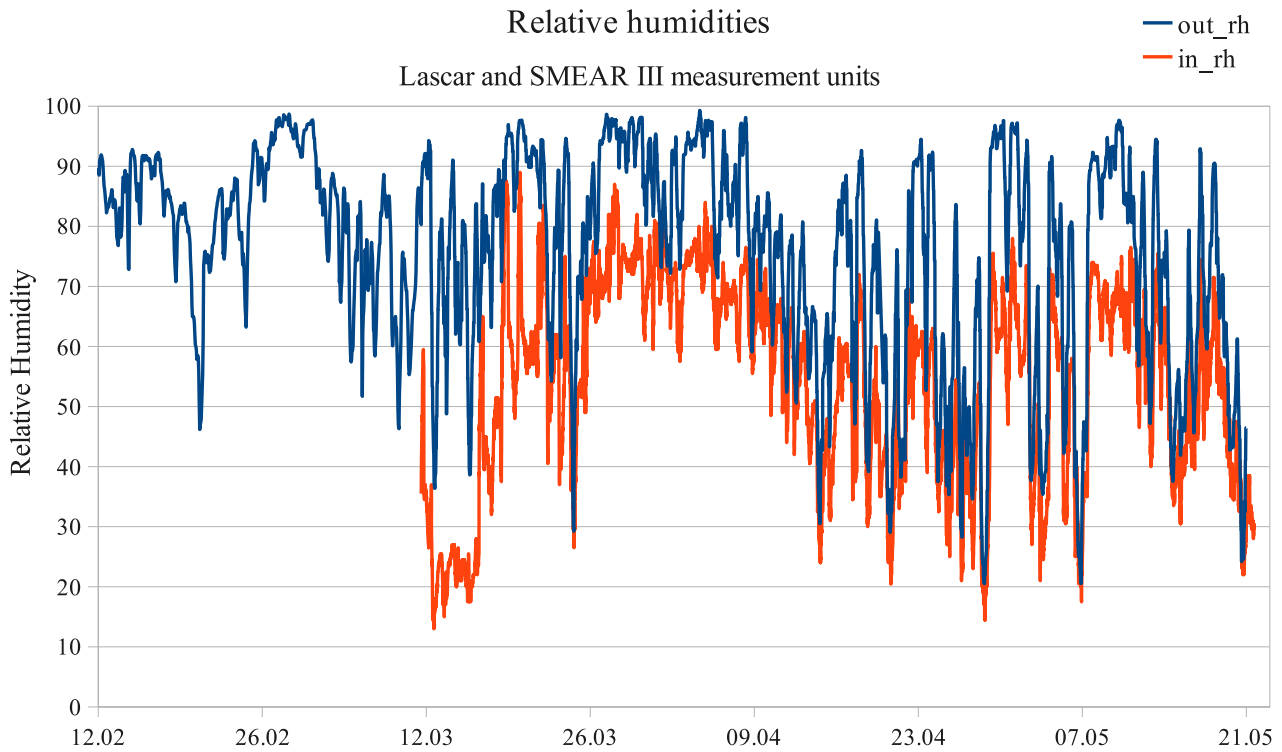


Figure 4: Relative humidities inside and outside the tent. Missing inside measurements are due to the Lascar data logger's delayed arrival.

5. DISCUSSION

Alongside with our measurements, the CS Department is building a new computing cluster. Upon completion, the cluster is estimated to become the third fastest supercomputer in Finland. Our server room is being retrofitted with added cooling capacity to accommodate the new equipment.

Calculating the amount of power consumed by the new devices, we know that the cluster can operate at a peak load of 75 kW. In order to cool this load, we have installed three new computer room air conditioning (CRAC) units, drawing a total of 6.9 kW of power. The unit that provides cool water to cool the CRACs operates in a designated heating, venting, and air conditioning (HVAC) area. That unit's power draw is specified as 44.7 kW. Last, the final piece of the setup is the liquid cooling unit positioned on the department's roof, which has a specified power draw of 3.8 kW. If we could just sum those figures up, the new cluster's power usage effectiveness (PUE) rating would be a rather efficient 1.74. Unfortunately, such is not the case, as our existing CRACs take care of some of the thermal load. This means that for PUE, the situation is worse, and more energy is wasted.

At the beginning of our test nobody really knew if the idea was entirely plausible, and if so, how long would the servers run. Now we know that at least a couple of months is a realistic guess, and that sub-zero temperatures or relative humidities above 80% or 90% are not a certified cause for server failures.

With these percentages, a central question concerns whether water can condense in the hardware, potentially short circuiting the electrical components. Our current knowledge is that water has few possibilities to condense in the equipment, as this would require the outside air to suddenly become warmer than the computer cases. As the cases are heated by their internal power draw and their inside air circulates due to the system fans, this phenomena is not as likely as some initial ideas suggested.

The air cooling tests described herein will continue to provide new data and knowledge about malfunctions encountered and shifts in the operating conditions. So far, neither the extreme colds of our winter or the rapidly changing conditions of spring have not been terminal for the hardware. As higher and higher summer temperatures are becoming common, we will see how temperature peaks affect our control group. It is certainly still possible that within the next months of operation, some components may start to regularly fail.

6. CONCLUSION

In this article, we have shown that current computer equipment is able to withstand very wide-ranging temperatures and humidities for extended periods of time. Further experimentation is necessary to find the limits of feasibility for this type of operation. Our future research will extend the initial results herein with more data over longer periods of time, over varying meteorological conditions, and more diverse hardware.

Through our current results, we have been able to independently verify the previous findings of computer manufacturers. These results promise very significant potential reductions in data center energy use, through the use of outside air for their cooling. As our department is also soon running very power-hungry hardware, we are dedicated into finding new and more efficient cooling solutions.

7. ACKNOWLEDGEMENTS

For their assistance at the setup phases of this project, we would like to thank Pekka Niklander and Pekka Tonteri from the IT staffs of the CS Department and HIIT respectively. Pasi Aalto of the SMEAR III project was invaluable in providing us with meteorological data. Finally, Ilona Riipinen and Sampo Smolander from the Physics Department clarified many concepts related to humidity and condensation.

8. REFERENCES

- [1] D. Atwood and J. G. Miner. Reducing Data Center Cost with an Air Economizer, 2008.
http://www.intel.com/it/pdf/Reducing_Data-Center_Cost_with_an_Air_Economizer.pdf.
- [2] D. Garday. Reducing Data Center Energy Consumption with Wet Side Economizers, 2007.
<http://www.intel.com/it/pdf/reducing-dc-energy-consumption-with-wet-side-economizers.pdf>.
- [3] E. S. HP. Eco-Friendly Data Centre Promises Windfall For HP Clients, 2009.
<http://h10134.www1.hp.com/news/features/5138/>.
- [4] L. A. Liikkanen and T. Nieminen. Comparison of End-User Electric Power Meters for Accuracy, 2009.
- [5] M. K. Patterson and D. Fenwick. The State of Data Center Cooling, 2008.
<http://download.intel.com/technology/eep/data-center-efficiency/state-of-date-center-cooling.pdf>.
- [6] Romir. 12 Feet Under: 1000 Square Feet of Geothermal PC Cooling, 2010.
<http://www.overclock.net/cooling-experiments/671177-12-feet-under-1000-square-feet.html>.