A Testbed and Data Yields for Studying Data Center Energy Efficiency and Reliability

Duc Van Le[1*]   Yingbo Liu[1*]   Rongrong Wang[1]   Rui Tan[1]   Lek Heng Ngoh[2]

[1]School of Computer Science and Engineering, Nanyang Technological University, Singapore
[2]Info-communications Media Development Authority, Singapore
{vdle,liuyb,rrwang,tanrui}@ntu.edu.sg, NGOH_Lek_Heng@imda.gov.sg

ABSTRACT

Many data centers still adopt low temperature setpoints, resulting in high energy consumption of their cooling systems. In our ongoing research project, we have built a data center testbed consisting of three test rooms hosting servers and network equipment. In particular, the testbed is equipped with cooling, heating, and ventilation systems that are under our full control. Thus, we can operate the testbed in a wide range of server room conditions. The test rooms and servers are instrumented with hardware and software sensors monitoring more than 200 measurement points. An extensive test plan is being executed to collect a large volume of data to understand the impact of environmental conditions on server's computing performance and reliability as well as the energy efficiency of the testbed. The results will provide important guidelines for building and operating energy-efficient data centers. We will work with relevant authorities towards providing the data collected from the testbed to the research communities.

CCS CONCEPTS

• **Hardware → Reliability and Performance - Temperatures**;

KEYWORDS

Data center, environmental condition, reliability, energy efficiency

1 INTRODUCTION

To improve the energy efficiency of data centers (DCs), a common approach is to raise the temperature setpoint of the Computer Room Air Conditioning (CRAC) system. A previous study has shown that increasing the temperature setpoint by 1°C can save energy by 2-5% [1]. However, raising the temperature setpoint may decrease the server's performance in terms of hardware reliability and computing throughput. El-Sayed *et al.* [2] have observed the increased

---

rates of hardware errors and failures due to the increased server room temperatures. Therefore, a key question in the design and operation of DCs is how to determine the temperature setpoint to well balance the CRAC energy consumption and the potential negative impact of high temperatures on the server performance and reliability. To answer this question, it is essential to understand the details of how the server room temperature will affect power consumption, computing throughput, server hardware reliability.

Several existing studies [3, 4] have applied wireless sensor networks to achieve better visibility of a DC's operating conditions. For example, the Microsoft researchers [4] implemented a data center Genome system using wireless sensors called *Genomotes* to monitor the environmental conditions in DCs. However, in the Genome system, the sensors are deployed in Microsoft's production data centers that operates at relatively low temperature setpoints. Therefore, their sensing data and observations do not cover a large temperature range for us to understanding the system performance and reliability under extreme temperature conditions. Moreover, the collected dataset is not publicly available.

To address the limitations of the above studies, we have built a DC testbed that hosts a number of servers with system failures and hardware damage allowed. In our testbed, various sensors are deployed to monitor the environmental conditions in real time, including temperature, relative humility, air pressure, air flow rate, etc, in the server room and on the computing and networking equipment. In particular, our DC testbed is equipped with a CRAC system capable of maintaining cooling parameters including temperature and air flow rate in large ranges. Therefore, we can conduct experiments under a variety of environmental conditions.

Our main objective is to observe the DC's states under realistic workloads under various environmental conditions. We have conducted a number of experiments in which the cooling conditions (e.g., inlet temperature and air flow rate setpoints) and the server operating parameters (e.g., CPU utilization, memory read/write speed, and workload types) are controlled in specified ranges. During the tests, four different types of measurements are collected, which are environmental, energy, performance, and reliability measurements. The collected data are stored in a centralized database for further analysis on modeling DC energy consumption, servers' performance and reliability under different environment conditions.

Through experimentation and analysis of collected data, the ultimate objective of this project is to stipulate good DC operation practices such that the DC energy efficiency is improved while maintaining the desired server performance and reliability. To the best of our knowledge, this is the first work that has full access to a real DC system that allows the failures of hardware components under controllable and challenging environmental conditions. We

Table 1: Experiment Setting

Parameters	Minimum	Maximum	Step Size
Inlet air temperature	25°C	37°C	1°C
Air flow rate	2500 m^3/h	12500 m^3/h	2500 m^3/h
Servers' CPU utilization	10%	100%	20%
Hard disk read/write speed	10 MB/sec	100 MB/sec	20 MB/sec

will work with relevant authorities towards providing the generated data to both academia and industry for studying DC thermal and energy management.

2 DATA CENTER TESTBED

2.1 System Infrastructure

To evaluate the impact of different environmental conditions on DC operating parameters, we design three test rooms with different room temperature control configurations. The first room is equipped with a cooling coil and a space heater, which allow us to control room ambient temperature within a range from 20°C to 40°C. The second room is equipped with cooling coil but no heater. Moreover, these two rooms are also equipped with a fan system which can provide an air flow speed of up to 12500 cubic meter per hour (m^3/h). The third test room is a standard server room in a production data center. It has the same temperature and relative humidity controls as other server rooms in the data center. It is used to generate baseline results on servers' computing performance and energy consumption.

Servers and network equipment from various vendors have been installed on a total of twelve 42U racks in the three test rooms. A smart digital power monitoring system is installed to measure the real-time power consumption of the cooling coils, the heater, the room fans, the server racks and other IT equipment. To monitor the environmental conditions in the test rooms, more than 100 hardware sensors are deployed to measure various parameters such as temperature, relative humility, differential pressure, and chemicals (i.e., SO_2, NO_2 and H_2S) at room and server rack levels. In addition, PID controllers use the sensor measurements as control inputs to operate the cooling coils, heater, and room fans to maintain the room temperatures and air flow rates at the specified setpoints.

2.2 Experiment Setup

We have run various experiments to understand and qualify the complex interactions among the DC operating parameters, including cooling, IT, and power. Table 1 shows the control parameters and their ranges which are used in our experiments. For example, in our experiments, the inlet air temperature is controlled up to 37°C, which is close to the highest temperature (i.e., 36.7 °C) recorded in Singapore in last 10 years. To conduct each experiment, we first setup control parameters to desired values and then run the server system with a specific benchmark workload for a certain time period, e.g., 1 hour. In addition, control programs are developed to continuously run long-term experiments in which control parameters are automatically changed in multiple loops.

2.3 Sensor Data Acquisition

We use a web-based data logging application to manage the acquisition and visualization of the data measurements from the hardware sensors and software sensors deployed in the environment, power supply, cooling system, as well as servers. We also deploy a number of software packages such as *PMI*, *SAR*, *Syslog* and *lm-sensor* to access the servers' built-in sensors and performance counters to collect the servers' real-time operating parameters.

The sensor measurement data can be divided into four categories: environmental, energy, reliability, and performance. Specifically, the environmental data relate to the conditions of the test room, servers and outdoor weather. Example include room/rack temperature and relative humidity, air velocity, servers' inlet/outlet temperatures, processor core and disk temperatures, etc. The energy data include power measurements of cooling devices, server racks and other IT equipment. Moreover, the reliability data relate to server hardware failures which consist of correctable and un-correctable memory errors, latent sector errors and additional hard disk SMART attributes. The performance data include server computing throughput, memory and CPU utilization, disk read/write throughput, and etc.

All sensor measurements are recorded as time-series data, where each data point is indexed by a timestamp corresponding to when the measurement is sampled. The metadata which consists of sensor location of each sensing point and details of cooling system architecture will be included when we release the dataset of collected sensor measurements. Our dataset can be used for other studies to evaluate their algorithm design on thermal modeling for adaptive temperature control and cooling optimization in a data center.

3 CONCLUSIONS

To understand how the data center operating conditions including environmental and IT device conditions affect power consumption, computing throughput and hardware reliability, we have built a DC testbed equipped with a number of servers, network equipment, and air cooling/heating systems. Based on the testbed, we have conducted various experiments to observe the behavior of the testbed under various servers' operating and environmental conditions. A monitoring system consisting of a large number of hardware and software sensors are developed to measure the environmental and IT equipment operating conditions in real time during the experiments. The collected dataset and analytical results will guide the design and operation of energy-efficient data centers. We will work with relevant authorities towards providing the dataset to the research communities.

REFERENCES

[1] J. Brandon. Setp. 2007. Going Green In The Data Center: Practical Steps For Your SME To Become More Environmentally Friendly. *Processor* 29 (Setp. 2007).
[2] Nosayba El-Sayed, Ioan A. Stefanovici, George Amvrosiadis, Andy A. Hwang, and Bianca Schroeder. 2012. Temperature Management in Data Centers: Why Some (Might) Like It Hot. In *Proc. SIGMETRICS*. 163–174.
[3] R. Khanna, H. Liu, and T. Rangarajan. 2014. Wireless Data Center Management: Sensor Network Applications and Challenges. *IEEE Microw. Mag.* 15, 7 (2014), S45–S60.
[4] Jie Liu and Andreas Terzis. 2012. Sensing Data Centers for Energy Efficiency. *Philosophical Transactions of The Royal Society A* 370 (2012), 136–157.