

Sensor Communication Model Using Cyber-Physical System Approach for Green Data Center

<https://doi.org/10.3991/ijim.v13i10.11310>

Masnida Hussin ^(✉), Raja Azlina Raja Mahmood, Mas Rina Mustaffa
University of Putra Malaysia, Selangor, Malaysia
masnidah09@gmail.com

Abstract—Energy consumption in distributed computing system gains a lot of attention recently after its processing capacity becomes significant for better business and economic operations. Comprehensive analysis of energy efficiency in high-performance data center for distributed processing requires ability to monitor a proportion of resource utilization versus energy consumption. In order to gain green data center while sustaining computational performance, a model of energy efficient cyber-physical communication is proposed. A real-time sensor communication is used to monitor heat emitted by processors and room temperature. Specifically, our cyber-physical communication model dynamically identifies processing states in data center while implying a suitable air-conditioning temperature level. The information is then used by administration to fine-tune the room temperature according to the current processing activities. Our automated triggering approach aims to improve edge computing performance with cost-effective energy consumption. Simulation experiments show that our cyber-physical communication achieves better energy consumption and resource utilization compared with other cooling model.

Keywords—Cyber-physical system; green data center; energy consumption; wireless communication

1 Introduction

Data centers for centralized and distributed computing are fast-growing area in the IT community and continue to expand in both size and number (Hussin, Muhammed, & Raja Mahmood, "An adaptive energy allocation for high-performance computing systems using a cyber-physical approach.", 2017; Abbasi, Jonas, Banerjee, Gupta, & Varsamopoulos, 2013; Zomaya & Rizvandi, 2012). The complete infrastructure of such data center includes server stacks, sensors and cables are purposely to make resources available for high-computational requirements. Such scenarios consumed massive amounts of electrical energy and have high cooling costs in running the data center. It is reported that high-performance computing data centers use more electricity for their cooling systems than small town sin each day (Zomaya & Rizvandi, 2012).

Furthermore, computer systems not only consuming vast amount of electrical power also emit excessive heat; this often results in system unreliability and performance degradation (e.g., (2014, 2015; Save Energy and Reduce Emissions to Achieve Sustainable Development and Improve Corporate Competitiveness, 2012; Macana, Quijano, & Mojica-Nava, 2011)). The data centres have set their internal ambient temperatures with respect the standard and guidelines of the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) (Inc, 2015). The guideline from ASHRAE is important because the processing elements in data centres must always be kept cool; otherwise, they can be damaged owing to overheating. Some data centres (Grange, Da Costa, & Stolf, 2018; Gu, Liu, Zhang, Huang, & Jia, 2015) have reduced cooling costs by taking advantage of the external environment to cool computational components in data centres. This maximized the number of days per year that external air could be used for cooling either through direct air-side or water-side economizers. In order to sustain a reliable computing environment, some companies (Save Energy and Reduce Emissions to Achieve Sustainable Development and Improve Corporate Competitiveness, 2012) have invested in cooling technologies wherein the heat generated from data centres is drawn back by a cooling generator.

In recent years, cyber-physical systems (CPSs) have attracted significant interest (Tham & Luo, 2013; Abbasi, Jonas, Banerjee, Gupta, & Varsamopoulos, 2013; Rajkumar, 2010)[14-16]. CPSs are computing systems that interact with the physical environment based on information from both the physical and cyber spaces. Interactions between computational equipment and the physical environment optimistically can be leveraged to improve the energy efficiency of data centers. In this work, we study adaptive interaction between system performance (i.e., resource utilization) and server room condition (i.e., temperature), aiming for energy efficiency. Our CPS communication model is constructed using heuristic that incorporate resource utilization analysis and physical sensor communication to achieve the (most) suitable temperature for server cooling. The communication model then helps administration to fine-tune energy consumption in data center. We utilized the sensor network architecture for bringing new optimization opportunities into the energy management. The results obtained show that our automated triggering temperature controlling approach is able to achieve energy efficiency without significantly affecting resource utilization. Hereafter, the terms ‘server room’ and ‘data center’ are used interchangeably.

The remainder of this paper is organized as follows. Section II describes related work on energy efficiency using CPSs. Section III details the models used in the paper. Our adaptive cooling system is presented in Section IV. Experimental settings and results are presented in Section V. Finally, Section VI concludes the paper.

2 Literature Review

The dynamic interaction among computers, networking, and physical systems in the cyber-physical concept leads to fundamentally new designs for energy

management. This has led to many energy management approaches for using CPSs in realizing the best trade-off between satisfying user requests and energy consumption in data centers. The authors in (Rajkumar, 2010) proposed the application of Cyber-Physical-System (CPS) through real-time data monitoring and prediction systems that integrated with cyber aspects. The proposed framework and algorithm have optimized the energy generation and distribution; however, it does not include system controller design. Meanwhile, three energy control strategies are used in (Parolini, 2012) to manage heat generation while improving the quality of computing services, i.e., baseline, uncoordinated, and coordinated strategies. The interesting element in their work is the cyber-physical index (CPI) that used for measuring the trade-off between distribution of cyber and physical effects in data center. An optimization-based decision maker is proposed in (Tham & Luo, 2013) to predict user demand and electricity supply from renewable sources. This prediction approach for calculating demand and supply is based on real-time sensor information and historical data coupled with optimization methods, and able to handle uncertainty in demand, supply, and electricity prices. The authors in (Qi, 2017) proposed the Model Predictive Control (MPC) based scheduling for micro grid. They used two control variables (i.e., turning on or turning off the compressor) to fine-tune the air-cooling temperature. The model constraints are set based on the Heating, Ventilation, and Air Conditioning (HVAC) standard. The model has been implemented utilizing the Simu link/MATLAB platform. Their work similar with our energy management approach in term of system scope that is distributed processing. In our work, we focus on adaptive interaction between wireless sensors and central controllers to assess amount of processing power while adjusting the room temperature appropriately. It aims cost-effective computation while considering energy savings.

3 System Structures

3.1 Server room structure

The model data center or server room used in this work contains several heterogeneous server racks. It is assumed that the server racks are for standard-size servers used in typical server room and mounted inside server cabinets or in rack frame infrastructure. The server room is equipped with air-conditioning (A/C) systems, which are self-contained to fit the temperature for a ‘server room’ that equivalent of a ‘small data center’ common to many types of office. The servers’ cabinets are arranged in rows so that cold air enters at the front and is expelled out the back. The doors to the room and windows are kept closed to ensure that the flow of cold air reaches the server cabinet. For each server rack, there are several number of temperature sensors i.e., 3 to 6 sensors as per suggested by Gartner [7] that allocated at each corner of the rack. The input from each sensor is sent to a computer terminal that acts as a central processing controller.

3.2 Energy model

The energy model in this work attempts to optimize the processing capacity of the servers in order to execute user workloads while maintaining a consistent (suitable) room temperature. For a given server, its power consumption during a particular time period can be measured using information about the power consumed during busy and idle states.

Hence, the power consumption of the computing server in the room is defined at each processor in the server rack as follows.

$$pow_{CPU} = (p_{max} \cdot Et) + (p_{min} \cdot t_{idle}) \quad (1)$$

Where p_{max} is the power usage at 100% utilization (busy), p_{min} is the power usage when servers are idle, s is number of processors, Et is execution time of the processor and t_{idle} is the total idle time of the processor, respectively. Note that the pow_{CPU} in each server rack are different according to the number of CPU resides on it. For this work, we use our previous research (Hussin, Lee, & Zomaya, Efficient Energy Management using Adaptive Reinforcement Learning-based Scheduling in Large-Scale Distributed System, 2011) to collect information on power consumption. The central controller then calculates a function of computing power according to the processing power state (which is mentioned in Section IV) while adaptively increasing, reducing, or maintaining the room temperature.

4 Automated Triggering Approach

4.1 Cyber-physical control components

Regarding the concept of CPSs, there are three major components in our energy management strategy: sensor, computer-generated controller, and actuator; in Fig.1. We designed a centralized sensing strategy to collect room temperature information from available sensors in the room. It means that the racks' sensors sent their information to the central sensor. This central sensor is connected to the central processing controller. The information from each sensor to the central sensor is updated for every 10 time units. Note that the sensing communication merely happens within the central and rack's sensors where there is no communication among other sensors. Hence, it implicitly reduced communication overhead between sensors.

In regards three components in our cyber-physical communication model, it is involved 3 main phases in communication process. Initially, the racks' sensors detected the heat emits from the processing component (e.g., system bus, CPU, memory, storage etc.) in the server rack and passed to the central sensor. In our communication model, the sensor can be wired or wireless category. Then, the central temperature sensor passes the information to the computer-generated controller or central processing controller. After the results been produced, the processing controller then provides alerts in two different ways:

- By triggering actuator (or LED) for changing the color according to the condition of defined processing states and
- By sending notification through mobile phone to the system administrator or any person in-charge of the room. The alert handling system is merely updated when the temperature reaches an unacceptable level (referring to the temperature standard for data center).

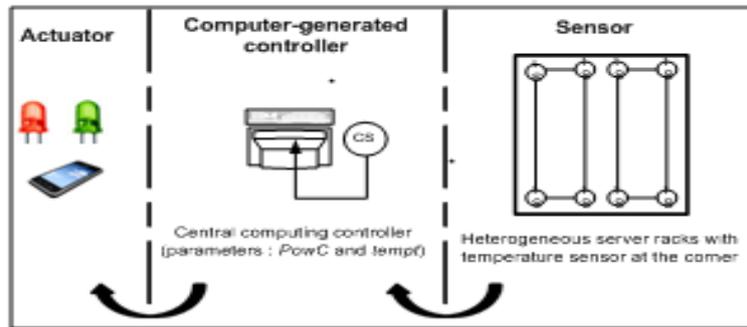


Fig. 1. Physical components and its sensor communication

4.2 Autonomous decision-making system

Our decision-making system consists of three stages are data collection, filtering and matching. All stages are controlled and occurred at the central processing controller, where information on heat of processing racks, room temperature level and processing power are assembled. In the first stage, the collection of information on processing power is realized for prioritizing it according to the level of importance in various processing states. In order to reach accurate data collection, we pre-processed our simulation environment for 30 minutes. The processing state is monitored and determined by the power consumption within the duration. If it is more than that, the reduction of energy consumption is hard to control and sustain [xx]. We then filter the data and categorizing it into respective processing states. If the average power consumption of the server rack is more than 80% of the highest power consumption of the CPU in the rack, its state is considered as busy. If the lowest power consumption is 20% lower than the average power consumption—i.e., there are only several servers performing processing activity—then the state is set to normal. Otherwise, the processing state is in the idle mode.

In the third stage, the matching strategy is applied between the processing state and the average room temperature, *tempt*. In order to minimize the temperature to the most suitable level, we introduce a *tempt threshold* and a *fair factor*. The *tempt threshold* represents the reference temperature of the server room and is assumed to be controllable. In this work, we use a *tempt threshold* between 18°C and 27°C, where this constraint reflects the environmental guidelines of the ASHRAE [7]. Meanwhile,

the *fair factor* is associated with increasing and decreasing percentages of the average room temperature. The servers in data centers typically operate at 50% of their maximum utilization [18], hence the room temperature is increased or reduced with an increment of 0.5 with respect to the processing state (i.e., for the busy mode, new $tempt = (tempt - (80\% + 0.5))$, and for the idle mode, new $tempt = (tempt + (20\% + 0.5))$). There is no change made to the room temperature if the processing state is in the normal mode.

The communication between CPS components is then occurred once there is a new value of temperature that needed to be adjusted, the processing-controller produced two different outputs (i.e., LED light-on and mobile message). For the color-changing LED there are 2 different colors where a red LED is on for busy mode and a green LED is on for idle mode. Meanwhile, the mobile messaging merely sent to the recipient for taking an action when the LED color is on for 30 seconds (either in busy or idle states). It is because we try to reach suitable temperature level for consistent duration of time. The recipient or server administrator is necessary to response to such alert especially when the current temperature reading is outside of the *tempt threshold* range.

5 Results and Discussion

We conducted simulations using different sensor capacities with workload runs at background to evaluate the performance of our cyber-physical communication model. The performance studied using two main metrics reutilization rate and energy consumption. The utilization rate, RU is defined as the percentage of time servers are busy servicing user workloads and is given by total number of workloads divided by total execution time of the processor. Meanwhile, the average energy consumption denotes the energy consumed to fulfill users' processing requirements and includes room temperature levels:

$$\beta = \left| \left| average\ PowCPU + \left(\frac{\sum diff}{\sum s} \right) \right| / \alpha \right| \quad (2)$$

where $diff$ is the total difference of temperature value changes (in °C) within time α (in seconds), and s is the number of available sensors in the room.

In our simulations, there are four racks of servers that each contained 4 to 12 processing cores/CPU's. The speed of a processing core is uniformly distributed within the range of 10 to 100. We designed workloads for the cores that followed a Poisson distribution with the mean inter-arrive time being 2-time units. The processing cores in this work are considered to be 'elastic' where it can go from being *capable* status (available to perform computation in a timely manner) to *incapable* status (unavailable to perform computation by deadline) at any time without further communication. For each server rack, there is temperature sensors where their capabilities varied based on sensor detection delay. We chose inter sensor detection time delay between 5 and 10 units. Each sensor updates the temperature value to the central sensor for every 10-time units.

a) Experimental results

Experiment 1: Impact of cooling model on energy consumption

We compared our CPS based communication model against the Model Predictive Control (MPC) model in (Li & Shi, 2014). We do not re-simulate their work for our data center. In fact, we merely used the MPC model that process and manage energy according to variable targets and limits to calculate expected changes in the cooling states. For fair comparison, the variables and thresholds are used same as that applied in our CPS model.

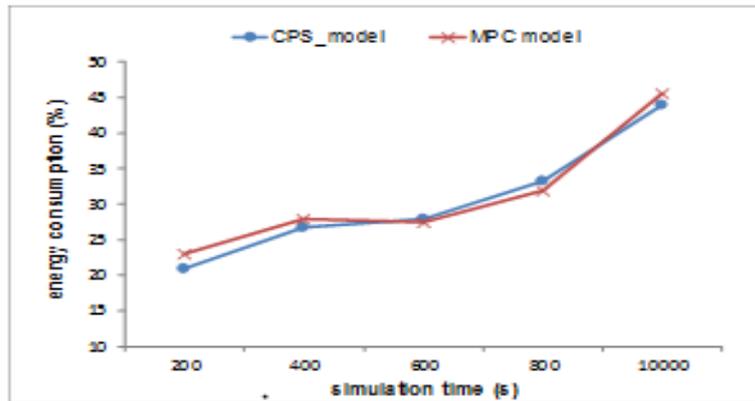


Fig. 2. Comparison of energy consumption

The effectiveness of CPS communication model is comparative with MPC model regarding energy consumption, as presented in Fig. 2. Interestingly, the energy consumption shows similar patterns as the load increases. It does not significantly differ and is observed to be about 5% on average. It is also demonstrated that our adaptive temperature setting strategy succeeds in heavy loads.

Experiment 2: Impact of sensor information-gathering approach on energy management

In this experiment, we investigated on how *CPS_model* deals with energy consumption when there is different pattern in distribution of sensor information. We compared the centralized information gathering-*sensor*, where each sensor passed information to the central sensor. Meanwhile, in the decentralized information gathering *d-sensor* each sensor is directly sent to the central processing controller. The utilization rates *RU* for *c-sensor* and *d-sensor* were not significantly different; in Fig. 3. This indicates that the method of information gathering (e.g., centralized or decentralized) had little influence on energy consumption in the server room.

Results shown in Fig. 3 clearly confirm the superior performance of *c-sensor* in terms of energy consumption. Specifically, *c-sensor* is able to ensure that a well-adapted level of energy with respect to different workloads in the system. This was mainly due to the fact that *c-sensor* effectively minimized communication overhead during data fusion, leading to consumed low energy over time.

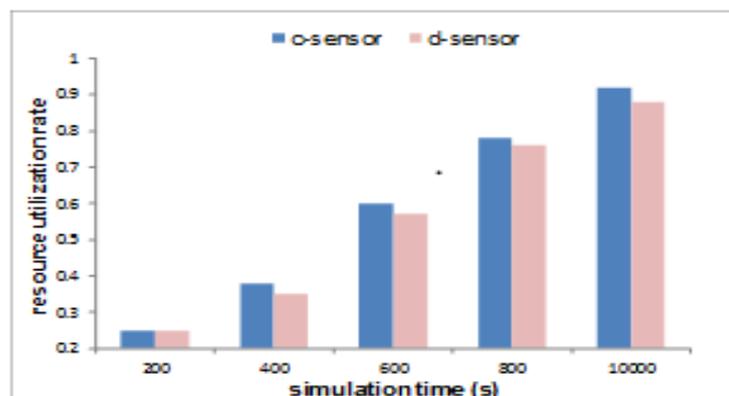


Fig. 3. Comparison of information gathering methods in terms of resource utilization

6 Conclusion

Information technology (IT), energy management, and cooling control systems have to work together in order to effectively achieved green data centers. Server rooms are experiencing rapid growth in both size and number, and aim to support the processing requirements of HPC applications. Such server room must equipped appropriate cooling system in order to sustain capability of its processing devices. In this work, we proposed an adaptive temperature-setting strategy using CPS sensing for monitoring the cooling system in HPC server room. Our cyber-physical communication model offers a fine-tune temperature strategy through multi-sensor communication and able to produce real-time notification system. Based on results from extensive experiments, this communication model demonstrated a substantial positive effect with better energy saving under different workload conditions, and optimistically to achieve green data center.

7 Acknowledgement

This work is supported by University Putra Malaysia under Grant GP/2018/9621200.

8 References

- [1] M. Hussin, A. Muhammed and R. A. Raja Mahmood, "An adaptive energy allocation for high-performance computing systems using a cyber-physical approach.," *Advanced Science Letters* 23(6), pp. 5045-5049., 2017. <https://doi.org/10.1166/asl.2017.7306>
- [2] Z. Abbasi, M. Jonas, A. Banerjee, S. Gupta and G. Varsamopoulos, "Evolutionary Green Computing Solutions for Distributed Cyber Physical Systems," in *Evolutionary Based*

- Solutions for Green Computing, Springer, 2013, pp. 1-28. https://doi.org/10.1007/978-3-642-30659-4_1
- [3] A. Zomaya and N. Rizvandi, "A Primarily Survey on Energy Efficiency in Cloud and Distributed Computing Systems.," ArXiv preprint arXiv:1104.5553, 2012.
 - [4] P. A. R. 2014, "(PRACE: Partnership for Advanced Computing in Europe)," 2015.
 - [5] "Save Energy and Reduce Emissions to Achieve Sustainable Development and Improve Corporate Competitiveness," Huawei Whitepaper, 2012.
 - [6] C. Macana, N. Quijano and E. Mojica-Nava, "A survey on Cyber Physical Energy Systems and their applications on smart grids.," in IEEE PES Conference on Innovative Smart Grid Technologies, (ISGT Latin America),, 2011. <https://doi.org/10.1109/isgt-la.2011.6083194>
 - [7] A. S. o. H. R. a. A.-C. E. Inc, "Data Center Storage Equipment- Thermal Guidelines, Issues and Best Practices.," ASHRAE ORG, Atlanta GA, 2015.
 - [8] L. Grange, G. Da Costa and P. Stolf, "Green IT scheduling for data center powered with renewable energy," Future Generation Computer Systems, vol. 86, pp. 99-120, 2018. <https://doi.org/10.1016/j.future.2018.03.049>
 - [9] C. Gu, C. Liu, J. Zhang, H. Huang and X. Jia, "Green scheduling for cloud data centers using renewable resources," in IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs), Hong Kong, 2015. <https://doi.org/10.1109/infcomw.2015.7179410>
 - [10] C. Tham and T. Luo, "Sensing-Driven Energy Purchasing in Smart Grid Cyber-Physical System." IEEE Transaction on Systems, Man and Cybernetics: Systems, vol. 43(4), pp. 773-784, 2013. <https://doi.org/10.1109/tsmca.2012.2224337>
 - [11] R. e. a. Rajkumar, "Cyber-physical systems: the next computing revolution. in 2010.," in Proceedings of the 47th Design Automation Conference. ACM, 2010. <https://doi.org/10.1145/1837274.1837461>
 - [12] L. e. a. Parolini, "A Cyber-Physical Systems Approach to Data Center Modeling and Control for Energy Efficiency.," in Proceedings of the IEEE, 2012.
 - [13] F. & J. X. & M. Y. & J. H. & Y. X. & X. X. & J. T. & L. X. Qi, "Model predictive control based scheduling method for a building microgrid.," in 10.1109/PESGM.2017.8274239., 2017.
 - [14] JM. Hussin, Y. C. Lee and A. Y. Zomaya, "Efficient Energy Management using Adaptive Reinforcement Learning-based Scheduling in Large-Scale Distributed System," in IEEE Computer Society in 40th Int'l Conf. on Parallel Processing (ICPP2011), Taipei, Taiwan., 2011. <https://doi.org/10.1109/icpp.2011.18>
 - [15] H. Li and Y. Shi, "Event-triggered robust model predictive control of continuous-time nonlinear systems." Automatica, vol. 50, no. 5, p. 1507–1513, 2014. <https://doi.org/10.1016/j.automatica.2014.03.015>

9 Authors

Masnida Hussin is the Associate Professor at Universiti Putra Malaysia. She works in the Department of Communication Technology and Networks and is a faculty of Computer Science and Information Technology Her research area includes Computer Networks, Parallel and Distributed Computing and Green Computing. She

has published many articles in reputable Scopus indexed journals. masnidah09@gmail.com

Raja Azlina Raja Mahmood works as a Lecturer at Universiti Putra Malaysia, Department of Multimedia and is a faculty of Computer Science Information Technology. She has done B.Sc. (Comp) (University of Michigan), M.Sc. (UiTM). Her research interest includes Networking, Wireless and Mobile Networks. azlina@fsktm.upm.edu.my

Mas Rina Mustafa is a senior lecturer at Universiti Putra Malaysia, Selangor. She works as a faculty of Computer Science and Information Technology. Her Research interest includes Multimedia Systems and Applications, Multimedia Information Retrieval, Pattern Recognition, Image Processing. masrina@fsktm.upm.edu.my

Article submitted 2019-07-17. Resubmitted 2019-08-24. Final acceptance 2019-08-26. Final version published as submitted by the authors.