Joint IT-facility optimization for green data centers via deep reinforcement learning

Zhou, Xin; Wang, Ruihang; Wen, Yonggang; Tan, Rui

2021

Zhou, X., Wang, R., Wen, Y. & Tan, R. (2021). Joint IT-facility optimization for green data centers via deep reinforcement learning. IEEE Network, 35(6), 255-262. https://dx.doi.org/10.1109/MNET.011.2100101

https://hdl.handle.net/10356/158611

https://doi.org/10.1109/MNET.011.2100101

Joint IT-Facility Optimization for Green Data Centers via Deep Reinforcement Learning

Xin Zhou, Ruihang Wang, Yonggang Wen, Rui Tan

***Abstract*—The data center market grows rapidly with the increase of data and its corresponding applications (e.g., machine learning, cloud storage, internet of things, etc.). The growth is boosted recently due to the shift of activities online in the COVID-19 pandemic. Reducing the energy consumption of data centers faces various challenges that are further aggravated by the tropical conditions with high temperature and humidity in the tropics like Singapore. The prevailing siloed approach of operating the information technology (IT) and the facility systems separately has resulted in wasteful over-provisioning. The recently proposed approaches for energy usage minimization under various constraints including thermal safety scale poorly with the data center size and often result in non-optimal solutions. To advance the state of the art, we apply deep reinforcement learning (DRL) to address the scalability problem and achieve optimality over a long time horizon in reducing data center energy usage. In particular, we deploy the data-driven deep model and physical rule based model in lieu of the physical data center during the training and validation phases to manage the thermal safety risks caused by DRL's strategy of learning from errors.**

***Index Terms*—Data center, energy efficiency, deep reinforcement learning, optimization**

I. Introduction

The data center market grows rapidly with the endless increase of data. Singapore is a data center hub of the Southeast Asia, accounting for more than 60% of the region's data center market with an annual growth rate of 10% [1]. However, dense data centers operating in the tropical condition aggravates the city state's energy demand. In 2015, data centers accounted for 9% of Singapore's electricity sales [2]. According to the Green Data Center Technology Roadmap [3], Singapore has the potential to significantly increase the energy efficiency of its data centers, projecting to a cumulative saving of about US$5 million in energy costs by 2030.

A data center provides various services for the users via many servers that are usually called nodes. These networked nodes are mutually coupled in thermodynamics, rendering the data center a cyber-physical system (CPS). The aim of optimizing the data center's energy efficiency is a sophisticated CPS problem that needs to respect to various constraints and requirements from both the cyber and physical aspects, including the overall computing throughput, serviceability, and hotspot prevention. As highlighted in the Roadmap [3], the prevailing siloed approach of operating information technology (IT) and facility subsystems has resulted in wasteful over-provisioning in data center design and operation. Heat is the crucial factor that couples the two subsystems in a data center as illustrated in Figure 1. Specifically, the IT subsystem in a data center generates heat in its computation efforts and the cooling subsystem moves the heat generated by the IT subsystem out of the data center building. However, both subsystems have complicated system dynamics in their heat generation and dissipation. Moreover, the regulation of both subsystems faces a key challenge caused by their distinct time constants. The IT subsystem tends to respond within seconds, while the cooling subsystem's response often takes effect in minutes. The time constant mismatch renders the control of the two subsystems challenging, particularly given the current practice that they are managed by two independent departments in the same organization for enterprise data centers or two entities for co-location data centers. Should they both perform the optimization in their own metrics, their optimization measures could even counter-affect each other. For example, when noticing a hot spot in the data center, the IT manager might decide to shift the load to another rack and the thermal manager might decide to send more cold air to that row. Without proper IT-facility coordination, the load may be migrated to another rack, whereas the added cooling is directed to the original rack that now has lower utilization. As a result, the same hot spot remains unsolved and simply moves to another location in the data center.

To advance the data center operation beyond the siloed scheme, prior research has investigated joint IT-facility optimization to increase energy efficiency for enterprise data

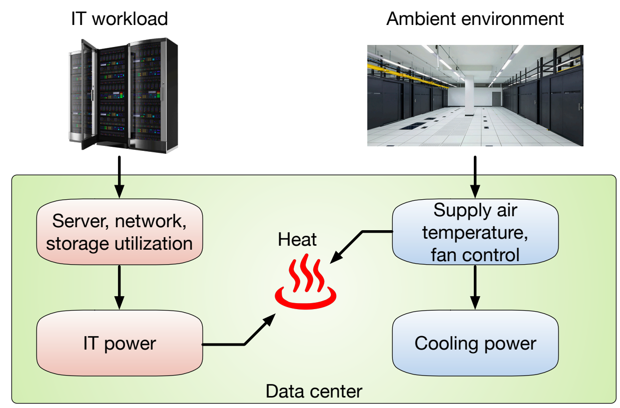


Fig. 1. Heat as a coupling factor between IT and facility subsystems. The IT subsystem in a data center generates heat in the computation tasks and the cooling subsystem dissipates the generated heat.

Xin. Zhou is with the School of Communications and Electronics, Jiangxi Science and Technology Normal University, China. Email: 1020161201@jxstnu.edu.cn. Ruihang. Wang, Yonggang. Wen, Rui. Tan are with the School of Computer Engineering, Nanyang Technological University, Singapore. Email: ruihang001@ntu.edu.sg, {ygwen, tanrui}@ntu.edu.sg.

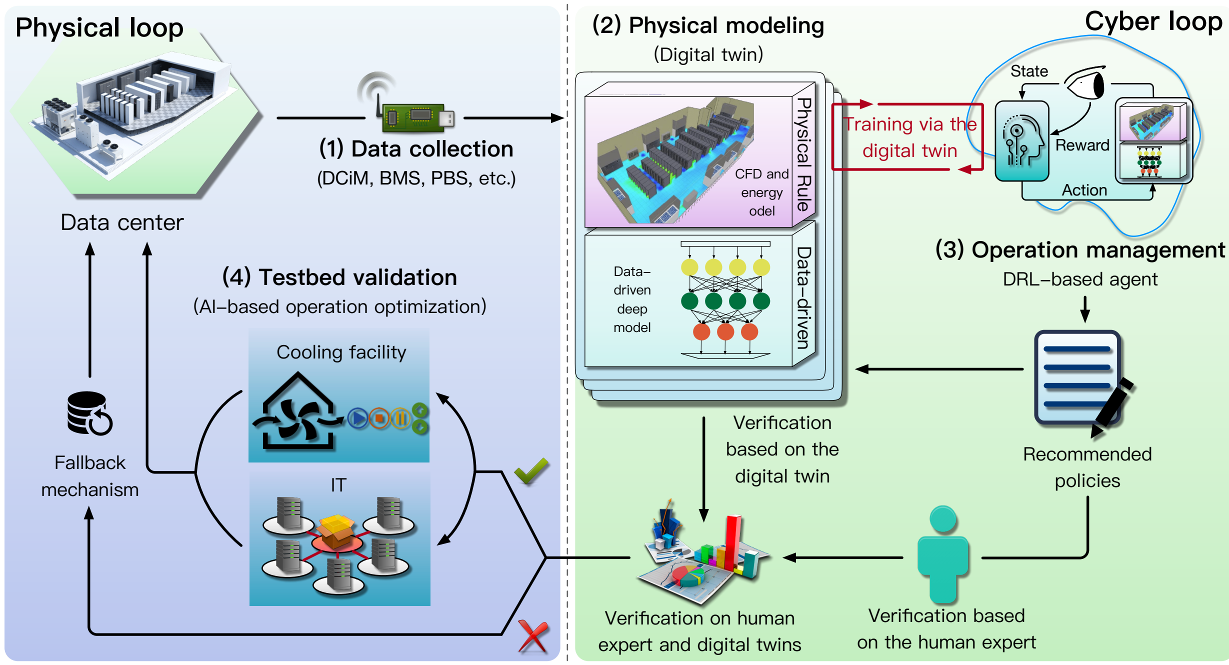


Fig. 2. Workflow of the proposed DRL-based approach. Our DRL-based approach consists of four main modules including data collection, physical modeling, operation management, and testbed validation.

centers [4]. Most existing joint IT-facility optimization approaches often assume some static or dynamic models for the target system. Based on constrained optimization formulations, heuristic algorithms have been proposed to decide the online control actions [5]. Owing to the development of the machine learning techniques, the learning-based techniques are recently used to predict the computing resources, temperature, and power consumption in advance to obtain more proactive control actions via heuristic algorithms. However, due to the complexity of the dynamics of the data center, the constrained optimization problems are hard in general and the heuristic algorithms often result in non-optimal solutions. Furthermore, since the modern data centers are being equipped with massive heterogeneous devices, the online optimization approach generally needs extensive search and scales poorly with the size of the data center and the time horizon for the optimization [6]. Google has announced that their trial with machine learning for data center management has resulted in cooling cost saving [7]. They have also automated machine learning-based management [8]. However, no detail of their approach is mentioned in the two technical posts [7], [8].

In this article, we present our research of applying deep reinforcement learning (DRL) for joint IT-facility optimization to address the aforementioned scalability challenge of modern data centers. Different from the online optimization approaches, our DRL-based approach iteratively interacts with the target system (i.e., the physical data center) by observing its real-time state and exerting the control action on the system. The global optimal control policy is derived from the iterative interaction, which generates the control actions for the operation optimization of the data center, including allocating the tasks to the servers, adjusting the airflow of the air conditioner, etc. Unlike the online optimization approach, trained DRL-based controllers do not need to solve any computation-intensive optimization problem at run time.

However, applying the control policy determined by the DRL may result in various risks including breaches of service requirements and even thermal unsafety. To address this challenge, our approach builds the data-driven deep model, energy model, and computational fluid dynamics (CFD) model based on real data collected from the data center to derive the training of the DRL-based controllers and verify the control policies. Specifically, we apply DRL to develop (1) load-aware target cooling that proactively manipulates cooling capacity and dispatching in response to dynamic IT workload and (2) thermal-aware task scheduling that optimizes the IT workload allocation in the presence of thermal dynamics. To deal with the different time constants of the IT and facility subsystems, we further develop (3) iterative IT-facility control optimization that proceeds iteratively between the IT side and the facility side to achieve a global optimal solution while satisfying various operational requirements. From the preliminary experimental results, our proposed load-aware target cooling approach achieves power consumption saving up to 15% and 30% for air cooling and water cooling systems, respectively.

II. System Overview

This section overviews our DRL-based joint IT-facility control approach. Figure 2 overviews our approach consisting of four main modules such as data collection, physical modeling, operation management, and testbed validation.

A. Data collection

The data collection module collects real-time operational information from the physical data center, including IT-related performance counters (e.g., CPU, I/O, etc.) and ambient condition measurements (e.g., temperature, humidity, and airflow, etc.). Static information including the location, layout, equipment configuration, and structure of the data center is also collected.

B. Physical modeling

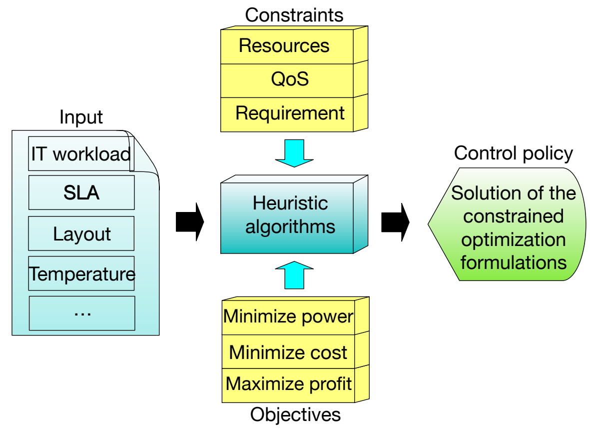
The physical modeling module extracts the underlying behaviors of the real-world system to build the digital twin which presents the digitalized model of the physical system and imitates system dynamics according to certain inputs. Such models are employed as the environment in the DRL. The DRL gradually learns the control policy by iteratively probing and interacting with the environment. The control actions generated by the control policy before its convergence are unpredictable and can be unsafe. Should the unpredictable control actions be applied to the physical data center, chances are that the unsafe control actions could result in large service requirement breaches and even physical damages in the data center. In our approach, the training of the DRL-based controllers is based on three types of computational models. The first type is a deep model that captures the complex interplay between IT and facility subsystems in the data center. The deep model is trained based on massive operational data collected from the physical system. The second type is a 3D CFD model for thermal analysis of the data hall, which is built with only standard information of the infrastructure (e.g., the layout of the data center, the server specification sheet, the air conditioner configuration, etc.). The third type is an energy model that focuses on power analysis of both the data hall and chiller plant. These models are used to (i) generate training data in extreme conditions that are not covered by the real data collected from the physical system, (ii) validate the control policies determined by the trained DRL-based controllers before being applied in the physical data center. Moreover, the physical rule based CFD and energy models can replace the data-driven deep model to conduct the training of the DRL-based controllers in the situation of data shortage.

C. Operation management

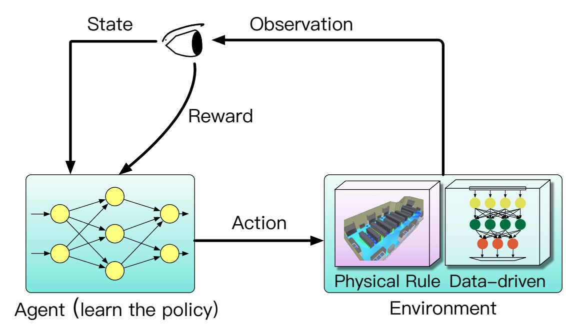
The operation management module applies the DRL-based controllers to solve the joint optimization problem and derive the optimal control policies for the IT and facility subsystems. It reduces the energy cost while ensuring thermal safety and business continuity. In particular, we propose three technologies including the thermal-aware task scheduling, load-aware target cooling, and iterative IT-facility optimization, which work synergistically to optimize the energy efficiency of the data center.

D. Testbed validation

The testbed validation step tests the control policies in Singapore's National Supercomputing Center to validate our proposed approaches while all the generated control actions will be verified by the human experts and physical models in advance. As illustrated in Figure 2, we collect massive data from the physical data center to build computational models in the forms of a deep model, an energy model, and a CFD model and then apply the DRL to seek a global optimal control policy. Consequently, the generated control policies are applied in the data center to gain energy savings. This cycle can proceed iteratively to improve the energy efficiency of data centers with a managed risk exposure.



(a)



(b)

Fig. 3. The workflows of a) existing approaches and b) DRL-based approach. (SLA: Service Level Agreement; QoS: Quality of Service.) The existing approaches assume some static models for the system but often cannot solve the extremely complicated constrained optimization problems due to the complexity of the models. Meanwhile, Our proposed DRL-based approach iteratively interacts with the target system and automatically learns the optimal control policies without solving the complicated constrained optimization formulations explicitly.

III. Data Analytics and Modeling

A. Existing approaches versus our DRL-based approach

The IT and facility subsystems in the data center have complex system dynamics in heat generation and dissipation. Moreover, the increasing demands of computation power have resulted in scaling up (i.e., upgrading existing servers) and scaling out (i.e., adding more servers) of the data center. The existing approaches often formulate constrained optimization problems that aim at maximizing the profits or minimizing the energy cost of the data center. Figure 3(a) illustrates the workflow of these existing approaches. However, due to the complexity of the system dynamics, the existing approaches often cannot even find objective function or solve the extremely complicated constrained optimization problems. Consequently, most of the existing approaches achieve local optimum only and cannot scale well with the size of the data center if the constrained optimization problem is solved in an online manner. To address these challenges, we apply DRL for joint IT-facility optimization [9]–[11]. In the DRL-based approach, the constraints are introduced into the reward function which guides the training process. The violation of constraints results in the penalty decreasing the reward. As a result, the violation will eventually reduce the probability of violating again to gain more reward. The control policy is optimized automatically in a trial-and-error manner under the guidance of these constraints without solving the complicated constrained optimization formulations explicitly. The workflow of our DRL-based approach is shown in Figure 3(b). Since DRL-based approach needs to explore the action space holistically, the control policies are unpredictable before the training completes. As a result, applying the control policies determined by the DRL before its convergence incurs various risks including breaches of service requirements and even physical damages. To avoid this risk, our approach builds the data-driven deep model, energy model, and the CFD model based on real data collected from the data center to drive the training of the DRL-based controllers and to verify the control policies determined by the controllers at run time.

B. Data-driven deep model

Given the data collected from the physical data center, we build a deep model to capture the sophisticated relationship between IT and facility subsystems. In our development, we adopt the long short-term memory (LSTM) model [12], which is a neural network architecture designed for short-term memory that can last for a long period of time. Thus, the LSTM network is well-suited to address the time series prediction problems. In our approach, we train the LSTM network based on the collected data to construct a model that can predict the future state of the data center (e.g., CPU utilization, ambient temperature, IT and facility power consumption, etc.) with the historical states. Owing to a large amount of operational data from the physical data center, the model is iteratively calibrated until it delivers accurate results. Driven by the calibrated model, the DRL agent explores the action space to find a global optimal control policy. Compared with the first-principle and analytical models adopted by existing approaches, our data-driven deep model can well scale with the size of the data center.

C. Physical rule based model

In addition to the data-driven deep model, the CFD model and energy model are also built for the data hall and chiller plant based on the basic information of the infrastructure (e.g., the layout, the server specification sheet, cooling facility configuration, etc.). With a commercial CFD software and the opensource platform EnergyPlus, we build and calibrate these two models that can be used to simulate the thermal dynamics and power consumption of the data center. As an example, Figure 4 shows the simulation result based on a constructed CFD model. The physical rule based models can simulate the operations of the data center over a long period with fine time granularity. Several detailed examples are as follows.

- We use the physical models to generate simulated training data that mainly covers the extreme conditions not included by the real data collected from the physical data center. The generated data is then used to calibrate the deep model to cover the extreme conditions during the DRL process. With these simulated training data, we can achieve higher prediction accuracy and better risk management.
- We use the physical models to validate the control policies determined by the trained DRL-based controllers before being applied in the physical data center. Sometimes even the trained DRL-based controllers cannot cover all the possible situations due to the complicated dynamics of the data center or absence of real operational data. This leads to the safety concern when applying the control policy directly to the data center. To address this concern, we apply the control policies on the physical models to validate the generated control actions by observing the simulation results of these models.
- We use the physical models to generate training data for uninstrumented variables. Specifically, it is often difficult to collect some operational data (e.g., workload trace, power consumption, etc.) from the data center due to resource constraints and confidential requirements. Applying the physical models to supplement the collected real data traces can improve the training of the DRL-based controllers.

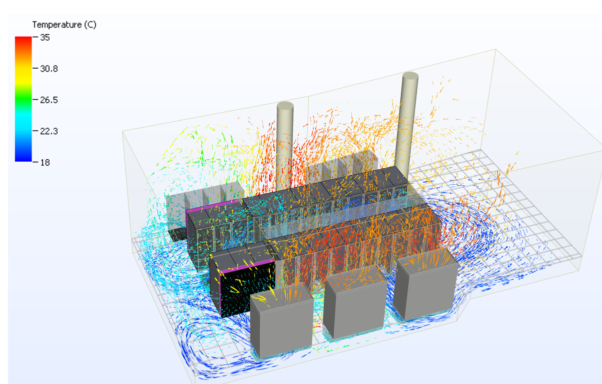


Fig. 4. The CFD model and simulation results of National Supercomputing Center Singapore (plotted with 6Sigma). Since the control policies are unpredictable before the training completes, our approach uses the CFD model to drive the training of the DRL-based controllers and to verify the generated control actions at run time.

IV. DRL-based Operation Optimization

This section presents the design of the DRL-based approach for data center and its application to solve the joint

IT-facility optimization problem. Our DRL-based approach iteratively interacts with the environment (i.e., data-driven deep model, energy model and CFD model) in lieu of the physical data center to avoid the potential thermal safety risk if the DRL agent interacts directly with the physical system. Once the DRL training completes, we apply the trained DRL-based controllers to address (1) load-aware target cooling, (2) thermal-aware task scheduling, and (3) iterative IT-facility optimization.

In practice, the operation optimization is not only critical for designing data centers, but also for operating them. For the former, our iterative IT-facility optimization approach that jointly controls both IT and facility subsystems can be easily applied to the data centers in the planning phase. The physical sensors and Data Center Infrastructure Management (DCIM) software package can be easily installed on the data center in the planning phase, resulting in convenient data collection. Moreover, most of the latest facility subsystems provide automatic control mechanisms compared with the manual control mechanisms of the older ones. These features render our iterative IT-facility optimization an available approach for manipulating both IT and facility subsystems in a unified manner. For the latter, the shortage of physical sensor in IT subsystem and automatic control mechanisms of the facility subsystem makes it difficult to apply the iterative IT-facility optimization approach to the operating data center. Data center upgrades are essential, but challenging and costly for data center managers. However, our load-aware target cooling and thermal-aware task scheduling are still readily viable for the operating data centers. The thermal-aware task scheduling that only controls the IT subsystem is suitable for the operating data centers equipped with facility subsystems with no automatic control mechanisms. On the contrary, the load-aware target cooling that only controls the facility subsystem can be used in the data centers with strong confidentiality requirements for the IT subsystems (e.g., data centers supporting banks).

A. DRL for data center operation optimization

Traditional reinforcement learning algorithms such as Q-learning evaluates the value of the state in which an action is taken. This value can be assessed by the effectiveness of taking action at this state. Given values of sufficient state-action pairs, taking the action with the largest effectiveness often leads to optimal policy and high efficiency. Therefore, a look-up table to store the values of state-action pairs plays a critical role in reinforcement learning. However, in the context of data center operations, the massive possible state-action pairs will result in a huge Q-table that is difficult to store and query. Thus, the tabular Q-learning approach is not suitable for a complicated environment.

To tackle the above challenge, the Deep Q-Network (DQN) approach employs a neural network to replace the Q-table [9], where this network is trained to approximate the Q-table and can deal with the large volume of state-action pairs. In DQN, the states are further mapped to Q values of different actions via a linear transform. With this feature, DQN is suitable for the control problem with a discrete action space. Therefore, we use DQN to address the problem of computing task scheduling that assigns the tasks to a finite number of servers, where the assignment, i.e., the action, is discrete.

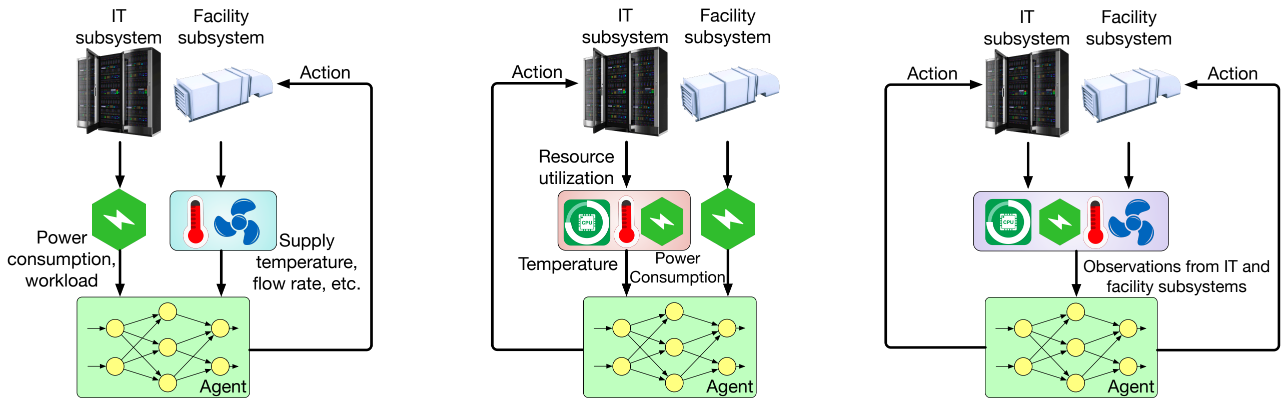
However, the original DQN cannot deal with cooling control, because the cooling control inputs are continuous values. To address this issue, we apply the Deep Deterministic Policy Gradient (DDPG) [11], which extends the DQN algorithm. Specifically, we employ two neural networks, that is, the actor-network computes the action predictions (e.g., continuous values that represent the control factors of the cooling facility) for the current state, whereas the critic-network evaluates the value of the current state and the action given by the actor. This feature renders DDPG a suitable solution for load-aware target cooling since DDPG computes and evaluates the continuous action directly.

B. Applications of DRL to specific tasks

In most data centers, the IT managers focus on task scheduling to satisfy the service-level agreement and guarantee the quality of experience with less or no concern on the temperature and the power consumption of the facilities, which may lead to hot spots and wastage of energy. On the contrary, the facility managers adapt actions to the cooling facilities according to the temperature with less or no concern on the workload, which may cause cooling over-provisioning. To jointly address these problems, we propose three advanced operation optimization technologies.

1) Load-aware target cooling: Developed mainly for facility managers, this technology dispatches cooling capacity in response to IT workload. We adopt the DDPG algorithm and take the workload and the power consumption of IT subsystem into account for the cooling optimization as shown in Figure 5(a). Note that the power consumption of the IT system closely relates to the workload. Moreover, the changes in workload incur the fluctuations of power consumption within a few seconds and then the temperature changes with respect to the power consumption within a few minutes. Our method proactively dispatches the cooling capacity according to the workload, which differs from most existing approaches that adjust the cooling facility passively with respect to temperature [13].

The target system of this research (i.e., Singapore's National Supercomputing Center) has both air cooling and water cooling that run independently. The air cooling system blows the cold air through the corridor and racks to remove the heat, whereas the water cooling system pumps the water directly into the heat sinks of CPUs to reduce the core temperature. We focus on optimizing the adjustments of the airflow rate and the pump flow rate, which are the actions in our proposed DRL-based approach. On the other hand, we use the workload and the power consumption of IT subsystem, the power consumption of facility subsystems and the ambient temperature to represent the state since they all affect the cooling load. By setting a proper reward which is a function of the power consumption of the target cooling facilities and the ambient temperature, our DRL-based approach starts from a certain state and iteratively adjusts the action iteratively



(a) (b) (c)

Fig. 5. Three applications of our proposed DRL-based approach: a) load-aware target cooling; b) thermal-aware task scheduling; c) iterative IT-facility optimization. The load-aware target cooling regulates cooling subsystem in response to IT workload. The thermal-aware task scheduling allocates the computing tasks to the IT subsystem in response to the thermal dynamics in the data center. The iterative IT-facility optimization is to jointly optimize the IT and facility subsystems simultaneously.

according to the reward. This iterative training automatically learns a global optimal control policy that achieves the desired trade-off among the workload of IT subsystem, the ambient temperature, and the energy cost of facility subsystem.

2) Thermal-aware task scheduling: The DRL can be also used to schedule the computing tasks in the IT subsystem in response to the thermal dynamics in the data center, subject to various constraints including the temperatures of the IT subsystem, and the power consumption of IT and facility subsystems. As one scenario of our joint optimization, we introduce the thermal-aware task scheduling [14] in this subsection. Figure 5(b) overviews our approach.

The traditional task scheduling approaches focus on optimizing the resource utilization and reducing the waiting time of the tasks. They rarely take the temperature or power consumption into account. Thus, they often result in over-provisioning and wastage of cooling energy. The thermal-unaware scheduling may assign the computing tasks to the servers in the already hot areas, creating hot spots. Consequently, under the prevailing cooling scheme, once a certain server exceeds a specified temperature upper bound, the facility manager can only increase the overall power of the cooling facilities to remove the hot spot, resulting in energy wastage. Our approach employs a DQN-based task scheduling approach to allocate computing tasks in a thermal-aware manner, which aims to reduce the overall power consumption of IT and facility subsystems while maintaining the temperature of the servers without significantly reducing the computing throughput.

3) Iterative IT-facility optimization: The proposed load-aware target cooling and thermal-aware task scheduling approaches control the facility and IT subsystems, respectively. To jointly optimize the IT and facility subsystems, there are several challenges. First, two subsystems have two time constants as the IT subsystem responds within seconds whereas the facility subsystem responds within minutes. The optimization measures of them may counter-affect each other when they optimize their own metrics. Second, to jointly control both subsystems, it is essential to observe the state of both subsystems and generate two kinds of control actions simultaneously. Consequently, the regulation of both subsystems based on DRL leads to larger state and action spaces compared with the regulation of one subsystem. Third, it is difficult to apply the traditional DRL algorithms such as DQN and DDPG to learn a global optimal control policy that generates two control actions laying on discrete space and continuous space, respectively.

To address these challenges, we have developed a two-time scale IT-facility optimization approach based on a Parameterized action space based Deep Q-Network (PADQN) algorithm [15]. The iterative IT-facility optimization approach derives the optimized policy to jointly control the IT and facility subsystems to bring them to an ideal balance that results in the reduction of energy consumption and improvement of energy efficiency as shown in Figure 5(c).

V. Evaluation

We have evaluated the proposed DRL-based operation optimization approaches on Singapore's National Supercomputing Center. We collected a wide range of operational data from the target data center, including temperature and power consumptions of IT and cooling facilities, server specifications, cooling configurations, etc. Based on these data, we applied the deep learning based approach, CFD software, and energy platform to model the sophisticated relationship between the IT and cooling facilities to build the data-driven model, CFD model, and energy model, respectively. Subsequently, the load-aware target cooling, thermal-aware task scheduling, and iterative IT-facility optimization algorithms are trained on the physical rule based and data-driven digital twins. The evaluation results are shown in Figure 6.

Compared with the baseline approach with manually designed control settings based on expert domain knowledge, the load-aware target cooling dynamically and proactively regulates cooling facilities according to changes on the IT load, which saves about 15% cooling power of the data center while guaranteeing the rack inlet temperatures not exceeding a predefined threshold [13]. In contrast, the thermal-aware task scheduling controls job allocations of the IT subsystem based on observations of ambient temperatures inside the data hall, which results in saving more than 9% power of the IT subsystem in comparison with the baseline using heuristic algorithms. Moreover, this approach reduces the processor temperature significantly [14]. Alternatively, different from the siloed baseline approaches controlling the IT and cooling facility separately, the iterative IT-facility optimization approach jointly operates the IT and cooling subsystems and saves up to 15% total power consumption of the data center [15]. In the real-world data center, the control optimization often has a specific target or is limited by special conditions. For instance, a co-location data center that rents its space and cooling capacity to the customer and ensures the system running within the range of SLA (Service-Level Agreement). The co-location provider can only optimize the control of the cooling facilities to save power since he does not have access to the IT system. On the contrary, the tenants want to optimize the scheduling of intensive tasks and computing resources to improve efficiency and reduce the task waiting time while they do not care about how the cooling system works. Joint control optimization can be carried out only when access to both IT and cooling systems is available. Therefore, we proposed these three DRL-based approaches because they can be flexibly employed in different scenarios to improve system operation and management of the data center for better business continuity and energy efficiency.

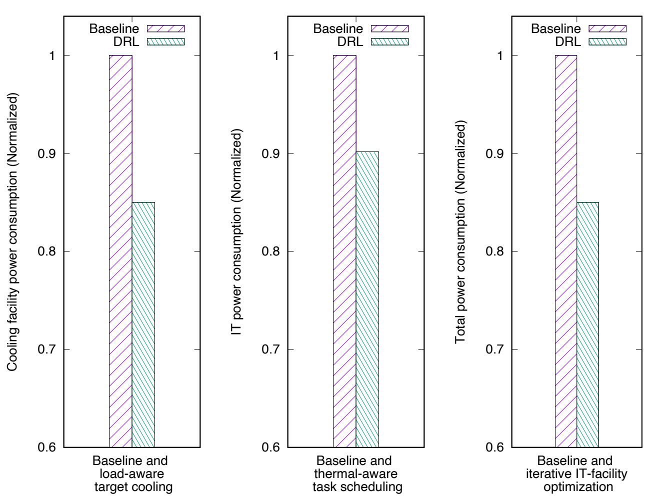


Fig. 6. Experimental results of the DRL-based optimization approaches. The load-aware target cooling optimizes the cooling facility control for improving energy efficiency of the data center. The thermal-aware task scheduling focuses on IT subsystem control optimization by proactively exploring thermal dynamics of the data hall to improve the computing energy efficiency of servers. The iterative IT-facility jointly regulates both IT and cooling facilities to improve the operation of the holistic data center system.

VI. CONCLUSIONS

We have proposed a joint IT-facility solution of digital twins and AI technologies for green data centers. Specifically, the proposed DRL-based approaches including the load-aware target cooling, thermal-aware task scheduling, and iterative IT-facility optimization aim to improve the energy efficiency while ensuring the thermal safety of the data center. In the proposed solution, the digital twins are used to provide massive training data and validate the control policy from the DRL-based approaches. Compared with the existing online optimization approaches, our solution applies a black-box approach as well as the digital twin to address the system dynamics in a complicated and evolving data center, relieving us from tedious, detailed modeling of system dynamics. Furthermore, it is a self-learning system that minimizes human intervention, reducing the potential of human errors in judgment and execution. As the data center becomes a mission-critical infrastructure, our DRL-based approach that interacts with the trained computational model can be effectively used to learn a global optimal control policy for improving the energy efficiency of the data center without affecting the operations of the physical data center. Currently, we are validating the effectiveness of the joint IT-facility optimization. In the future, real-world applications of our proposed solution will be conducted for technology validation and commercialization.

VII. ACKNOWLEDGEMENTS

This research is in part supported by the Nation Research Foundation, Prime Minister's Office, Singapore under its Energy Programme (EP Award No. NRF2017EWT-EP003-023) administrated by the Energy Market Authority of Singapore, its Green Data Centre Research (GDCR Award No. NRF2015ENC-GDCR01001-003) administrated by the Infocommunications Media Development Authority, its Central Gap Fund ("Central Gap" Award No. NRF2020NRF-CG001-027), and its Sustainable Tropical Data Centre Test bed programme (STDCT).

REFERENCES

[1] Enterprise Innovation. (2013) Singapore is Top Data Center Hub in SE Asia: Report. (2021, March 15). [Online]. Available: www.enterpriseinnovation.net/article/singapore-top-data-center-hub-se-asia-report-1301859302.
[2] IMDA, Huaiwei, and Keppel. (2017) IMDA, Huaiwei and Keppel Data Centres Sign Memorandum of Intent to Conduct a Joint-Feasibility Study for a First-of-its-kind High-Rise Green Data Centre Building. (2021, March 15). [Online]. Available: https://www.imda.gov.sg/-/media/imda/files/about/media-releases/2017/annex-a---fact-sheet-for-high-rise-green-dc.pdf?la=en.
[3] IDA, NSCC, and NRF. (2012) Green Data Centre Technology Roadmap. (2021, March 15). [Online]. Available: www.nccs.gov.sg/docs/default-source/default-document-library/green-data-centre-technology-roadmap.pdf.
[4] W. Zhang, Y. Wen, Y. W. Wong, K. C. Toh, and C. H. Chen, "Towards Joint Optimization Over ICT and Cooling Systems in Data Centre: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1596–1616, 2016.
[5] L. Cupertino, G. D. Costa, A. Oleksiak, W. Piatek, J. M. Pierson, J. Salom, L. Siso, P. Stolf, H. Sun, and T. Zilio, "Energy-Efficient, Thermal-Aware Modeling and Simulation of Data Centers: The CoolEmAll Approach and Evaluation Results," *Ad Hoc Networks*, vol. 25, pp. 535–553, 2015.
[6] M. Polverini, A. Cianfrani, S. Ren, and A. V. Vasilakos, "Thermal-Aware Scheduling of Batch Jobs in Geographically Distributed Data Centers," *IEEE Transactions on Cloud Computing*, vol. 2, no. 1, pp. 71–84, 2014.
[7] Google. (2016) DeepMind AI Reduces Energy Used for Cooling Google Data Centers by 40%. (2021, March 15). [Online]. Available: https://blog.google/outreach-initiatives/environment/deepmind-ai-reduces-energy-used-for/.

[8] Google. (2018) Safety-First AI for Autonomous Data Centre Cooling and Industrial Control. (2021, March 15). [Online]. Available: https://deepmind.com/blog/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control/.
[9] V. Mnih, K. Kavukcuoglu, D. Silver, and et al., "Playing Atari with Deep Reinforcement Learning," *arXiv preprint arXiv:1312.5602*, 2013.
[10] L. Huang, X. Feng, C. Zhang, L. Qian, and Y. Wu, "Deep Reinforcement Learning-based Joint Task Offloading and Bandwidth Allocation for Multi-user Mobile Edge Computing," *Digital Communications and Networks*, vol. 5, no. 1, pp. 10–17, 2019.
[11] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous Control with Deep Reinforcement Learning," *arXiv preprint arXiv:1509.02971*, 2015.
[12] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
[13] Y. Li, Y. Wen, D. Tao, and K. Guan, "Transforming Cooling Optimization for Green Data Center via Deep Reinforcement Learning," *IEEE transactions on cybernetics*, vol. 50, no. 5, pp. 2002–2013, 2019.
[14] D. Yi, X. Zhou, Y. Wen, and R. Tan, "Toward Efficient Compute-Intensive Job Allocation for Green Data Centers: A Deep Reinforcement Learning Approach," *IEEE International Conference on Distributed Computing Systems*, pp. 634–644, 2019.
[15] Y. Ran, H. Hu, X. Zhou, and Y. Wen, "Deepee: Joint Optimization of Job Scheduling and Cooling Control for Data Center Energy Efficiency using Deep Reinforcement Learning," *IEEE International Conference on Distributed Computing Systems*, pp. 645–655, 2019.

Xin Zhou is an Assistant Professor at School of Communications and Electronics, Jiangxi Science and Technology Normal University. He received his M.E. and Ph.D. degrees from the Department of Information Engineering, Hiroshima University, Japan in 2013 and 2016 respectively. His research interests include reconfigurable architectures, parallel computing, parallel architecture, FPGA computing, deep reinforcement learning, and green data center. He is expert in learning-based ICT and cooling facilities control optimization and impact analytics.

Ruihang Wang received his M.E. and B.E degrees from the School of Precision Instrument and Opto-electronics Engineering, Tianjin University (TJU), China. He is currently a PhD student in Nanyang Technological University (NTU), Singapore. His research interests include cyber-physical system, machine learning and digital twin technology.

Yonggang Wen is the Professor of Computer Science and Engineering at Nanyang Technological University (NTU), Singapore. He has also served as the Associate Dean (Research) at College of Engineering at NTU Singapore since 2018. He received his PhD degree in Electrical Engineering and Computer Science from Massachusetts Institute of Technology (MIT), Cambridge, USA, in 2008. He serves on editorial boards for IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia, etc. His research interests include cloud computing, green data center, distributed machine learning, blockchain, multimedia network and mobile computing. He is a Fellow of IEEE.

Rui Tan is an Associate Professor at School of Computer Science and Engineering, Nanyang Technological University, Singapore. He was a Research Scientist (2012-2015) and a Senior Research Scientist (2015) at Advanced Digital Sciences Center. He received the Ph.D. (2010) degree in computer science from City University of Hong Kong, the B.S. (2004) and M.S. (2007) degrees from Shanghai Jiao Tong University. His research interests include cyberphysical systems, sensor networks, and ubiquitous computing systems. He received the Best Paper Awards from IPSN'17, CPSRSG' 17, Best Paper Runner-Ups from IEEE PerCom'13 and IPSN'14.

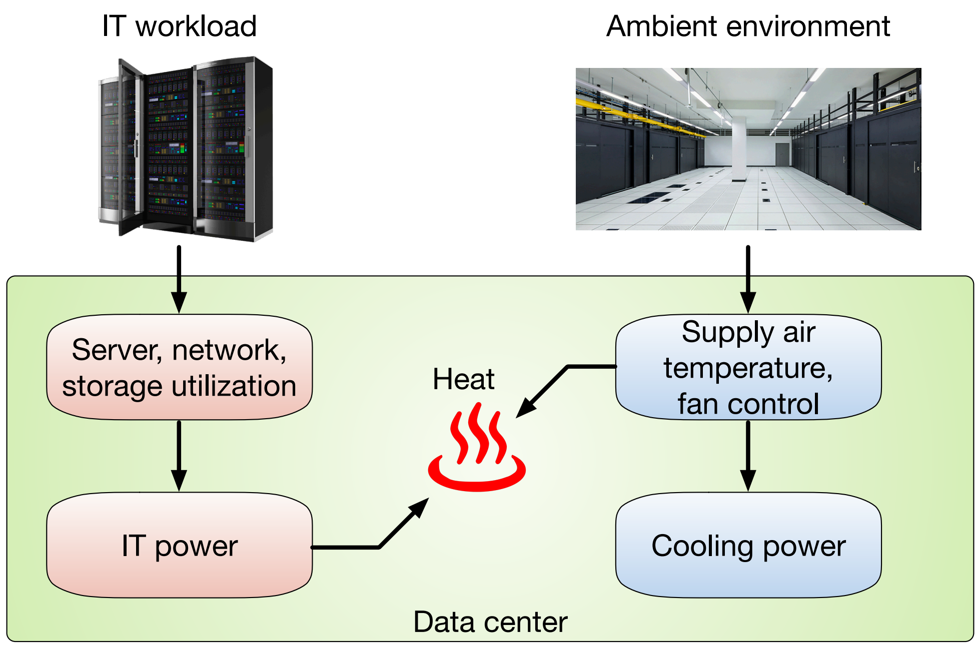


Fig. 1. Heat as a coupling factor between IT and facility subsystems. The IT subsystem in a data center generates heat in the computation tasks and the cooling subsystem dissipates the generated heat.

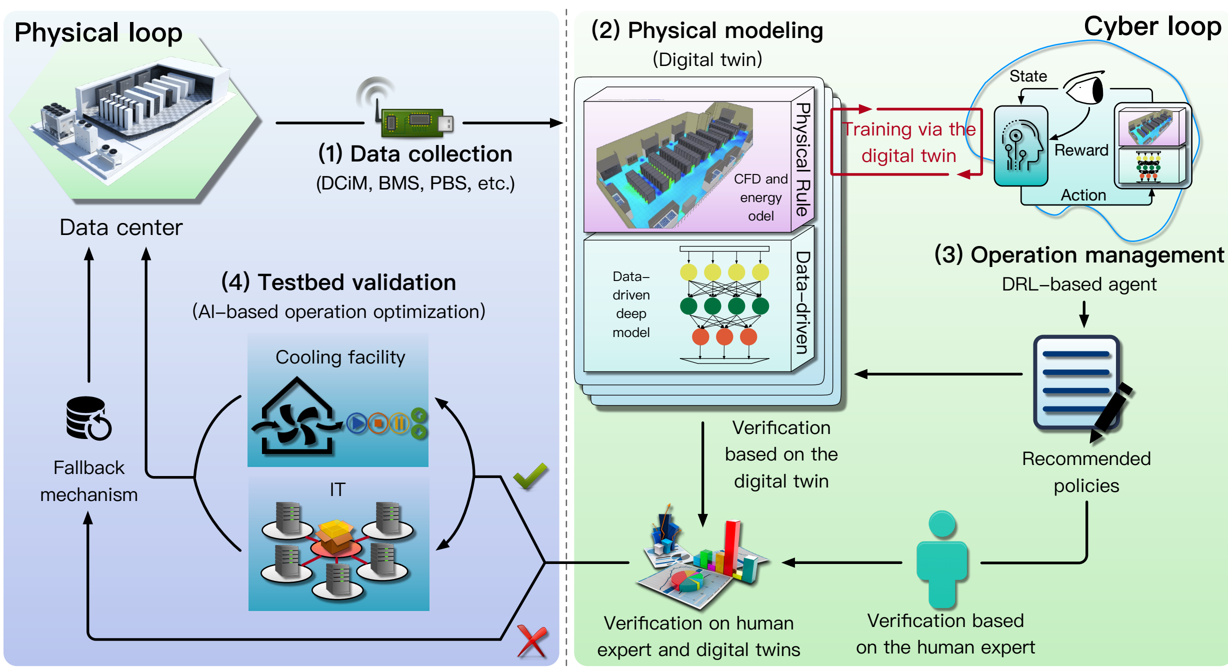


Fig. 2. Workflow of the proposed DRL-based approach. Our DRL-based approach consists of four main modules including data collection, physical modeling, operation management, and testbed validation.

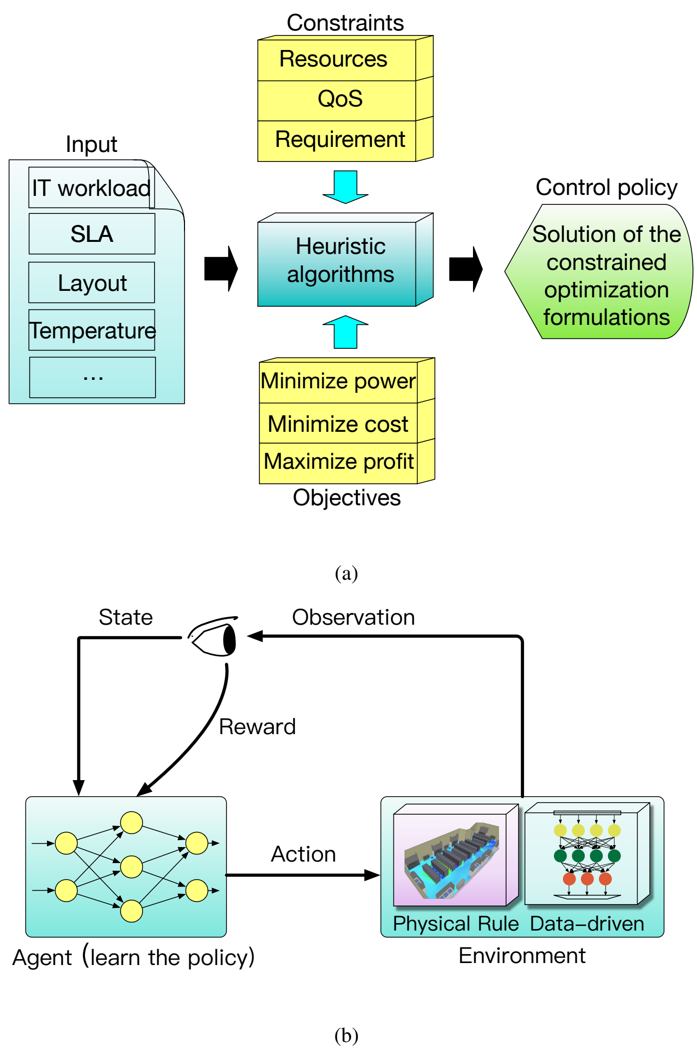


Fig. 3. The workflows of a) existing approaches and b) DRL-based approach. (SLA: Service Level Agreement; QoS: Quality of Service.) The existing approaches assume some static models for the system but often cannot solve the extremely complicated constrained optimization problems due to the complexity of the models. Meanwhile, Our proposed DRL-based approach iteratively interacts with the target system and automatically learns the optimal control policies without solving the complicated constrained optimization formulations explicitly.

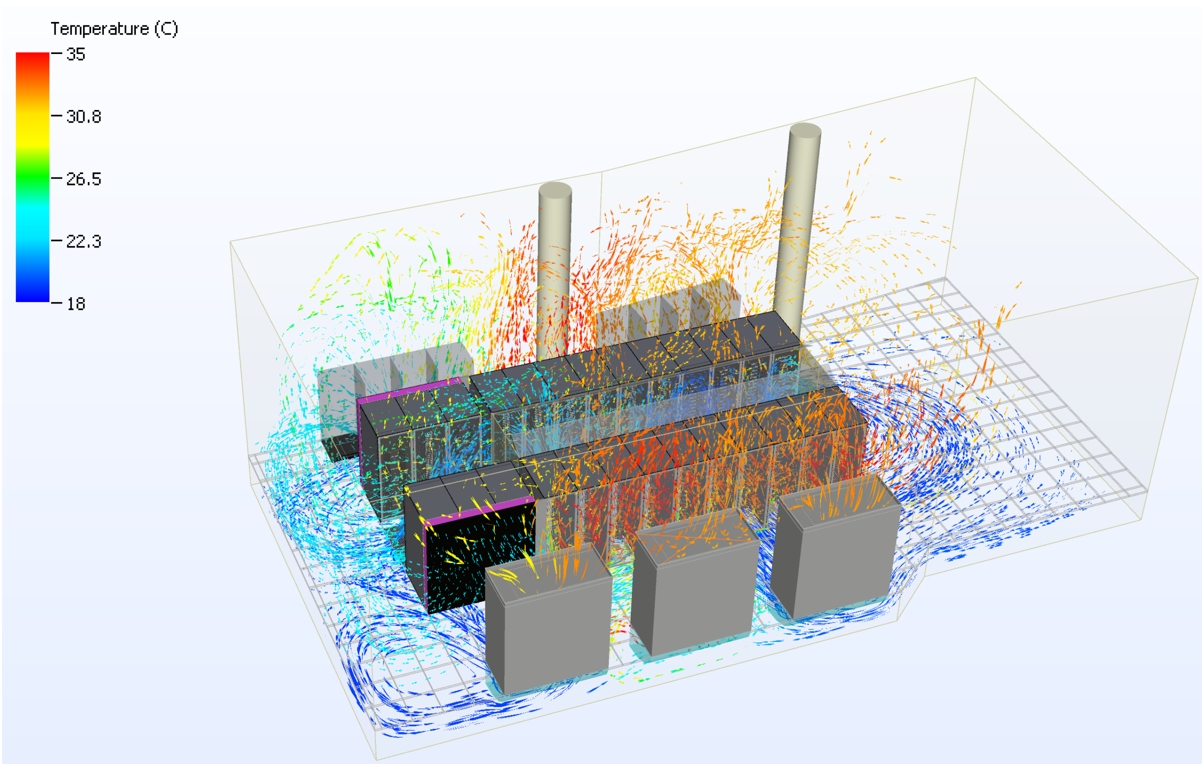


Fig. 4. The CFD model and simulation results of National Supercomputing Center Singapore (plotted with 6Sigma). Since the control policies are unpredictable before the training completes, our approach uses the CFD model to drive the training of the DRL-based controllers and to verify the generated control actions at run time.

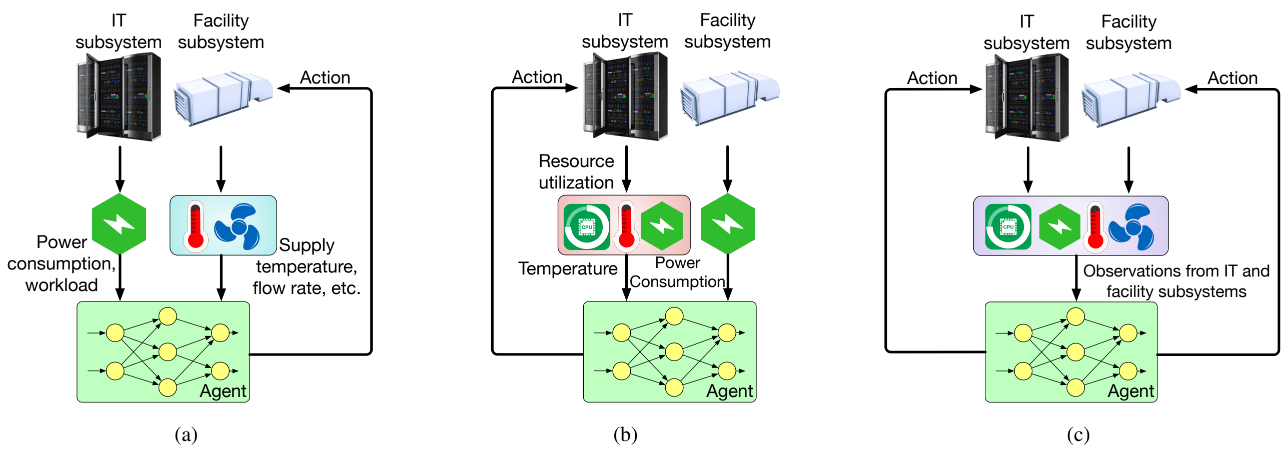


Fig. 5. Three applications of our proposed DRL-based approach: a) load-aware target cooling; b) thermal-aware task scheduling; c) iterative IT-facility optimization. The load-aware target cooling regulates cooling subsystem in response to IT workload. The thermal-aware task scheduling allocates the computing tasks to the IT subsystem in response to the thermal dynamics in the data center. The iterative IT-facility optimization is to jointly optimize the IT and facility subsystems simultaneously.

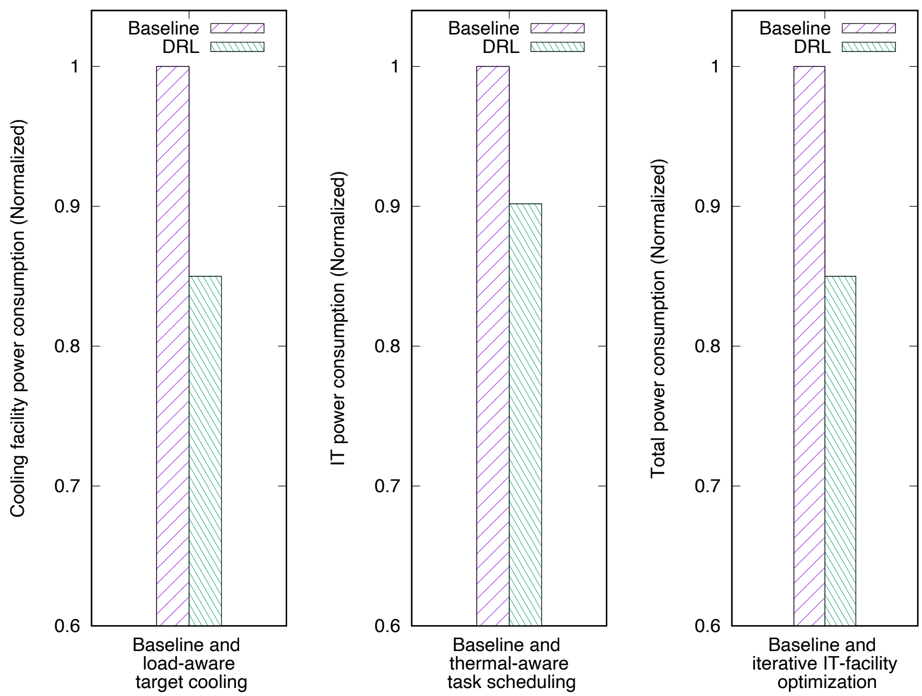


Fig. 6. Experimental results of the DRL-based optimization approaches. The load-aware target cooling optimizes the cooling facility control for improving energy efficiency of the data center. The thermal-aware task scheduling focuses on IT subsystem control optimization by proactively exploring thermal dynamics of the data hall to improve the computing energy efficiency of servers. The iterative IT-facility jointly regulates both IT and cooling facilities to improve the operation of the holistic data center system.