Berezovskaya, Yulia; Yang, Chen-Wei; Mousavi, Arash; Vyatkin, Valeriy; Minde, Tor Bjorn

Modular model of a data centre as a tool for improving its energy efficiency

Published in:
IEEE Access

DOI:
10.1109/ACCESS.2020.2978065

Published: 01/01/2020

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Berezovskaya, Y., Yang, C.-W., Mousavi, A., Vyatkin, V., & Minde, T. B. (2020). Modular model of a data centre as a tool for improving its energy efficiency. *IEEE Access*, *8*, 46559-46573. Article 9023486. https://doi.org/10.1109/ACCESS.2020.2978065

Received February 14, 2020, accepted February 24, 2020, date of publication March 3, 2020, date of current version March 17, 2020.
Digital Object Identifier 10.1109/ACCESS.2020.2978065

Modular Model of a Data Centre as a Tool for Improving Its Energy Efficiency

YULIA BEREZOVSKAYA[1], (Member, IEEE), CHEN-WEI YANG[1], ARASH MOUSAVI[2], VALERIY VYATKIN[1,3], (Senior Member, IEEE), AND TOR BJÖRN MINDE[4]
[1]Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, 971 87 Luleå, Sweden
[2]IT Department, SCANIA CV AB, Södertälje, Sweden
[3]Department of Electrical Engineering and Automation, Aalto University, 02150 Helsinki, Finland
[4]RISE SICS North, 973 47 Luleå, Sweden
Corresponding author: Yulia Berezovskaya (yulia.berezovskaya@ltu.se)

This work was supported by the Swedish Energy Agency (Energimyndigheten) through the Cloudberry Datacentres Project under Grant 43090-2.

ABSTRACT For most modern data centres, it is of high value to select practical methods for improving energy efficiency and reducing energy waste. IT-equipment and cooling systems are the two most significant energy consumers in data centres, thus the energy efficiency of any data centre mainly relies on the energy efficiency of its computational and cooling systems. Existing techniques of optimising the energy usage of both these systems have to be compared. However, such experiments cannot be conducted in real plants as they may harm the electronic equipment. This paper proposes a modelling toolbox which enables building models of data centres of any scale and configuration with relative ease. The toolbox is implemented as a set of building blocks which model individual components of a typical data centre, such as processors, local fans, servers, units of cooling systems, it provides methods of adjusting the internal parameters of the building blocks, as well as contains constructors utilising the building blocks for building models of data centre systems of different levels from server to the server room. The data centre model is meant to accurate estimating the energy consumption as well as the evolution of the temperature of all computational nodes and the air temperature inside the data centre. The constructed model capable of substitute for the real data centre at examining the performance of different energy-saving strategies in dynamic mode: the model provides information about data centre operating states at each time point (as model outputs) and takes values of adjustable parameters as the control signals from system implementing energy-saving algorithm (as model inputs). For Module 1 of the SICS ICE data centre located in Luleå, Sweden, the model was constructed from the building blocks. After adjusting the internal parameters of the building blocks, the model demonstrated the behaviour quite close to real data from the SICS ICE data centre. Therefore the model is applicable to use as a substitute for the real data centre. Some examples of using the model for testing energy-saving strategies are presented at the end of the paper.

INDEX TERMS Data center, cooling system, computational system, CPU, cooler, energy consumption, energy-saving, modular model, server, server room, simulation, thermal behavior.

I. INTRODUCTION

In recent years, modern data centres formed a new type of industrial plants with a substantial amount of energy consumption. According to [1], the data centre sector consumed around 1.5% of the global electricity consumption in 2010. Authors of [2] estimate the annual growth rate of energy consumption in this sector to be between 7.5% and 20%, and even the optimistic forecast implies a significant annual increase in electricity use. Since fossil fuel is still the main source of electricity production, the increase of energy consumption in this fast-growing sector raises CO_2 emissions, which have a negative impact on nature [3], [4]. Nevertheless, most modern data centres consume energy in an inefficient manner [5]. Therefore, scientific research that can lead to finding practical methods for improving energy efficiency and reducing energy waste in modern data centres is of high value.

With relation to improving energy efficiency in data centres, computational systems and cooling systems are the subsystems most promising to optimise, as they are the two

The associate editor coordinating the review of this manuscript and approving it for publication was Atif Iqbal.

most significant energy consumers in data centres [6]. Generally, a modern data centre embodies plenty of heterogeneous computational resources consuming a significant amount of energy such as servers and storages, thus selecting the efficient IT-equipment and resource management techniques allows optimizing the energy consumption in data centres [7]. On the other hand, according to recent research, the cooling system alone consumes around 40% of total energy consumption in a modern data centre, hence the choice of a cooling method, cooling equipment and control strategy for cooling system determines the energy efficient data centre [3] [8]. Consequently, the energy efficiency of any data centre mainly relies on the energy efficiency of its computational and cooling systems.

The techniques for optimisation of the energy usage of computational systems can be categorised as (a) static power management (SPM) which relates to the hardware level and implements the optimisation methods at the design time [7], [9]; and (b) dynamic power management (DPM) which adapts the system energy use based on resource requirements at run-time [9]. The SPM is considered as getting the hardware ready for using the DPM [7]. The dynamic power management addresses two levels: hardware and software. Hardware-based DPM techniques include dynamic performance scaling methods, such as dynamic voltage frequency scaling (DVFS) for servers [10], or adaptive link rate for networks [11]; as well as partial or complete dynamic deactivation of idle components [9]. The software-based DPM implements resource management techniques such as virtualisation, scheduling, consolidation [12]–[14], for minimising the number of operating servers without performance loss thus energy efficiency can be achieved by switching off idle servers.

At the same time, there is number of techniques for improving the energy efficiency of cooling systems. This paper only addresses energy efficiency in air-cooled systems, as they are the most popular cooling systems for data centres [15]. An airflow optimization is one of such techniques reducing the coolers utilization due to separating the hot air exhausted by IT-equipment and cold air supplied by coolers to servers' inlet, hereby keeping the servers' inlet temperature in a lower level [8]. In addition, usage of raised-floors and containment of hot aisles can improve the efficiency of the airflow optimization [16]. One more way to achieve energy saving is to flexibly adapt the utilization of the coolers to actual IT-load [17], as well, the local fans to the actual temperature of corresponding CPUs [18]–[20]. Finally, the free-cooling supplying the outside air directly to the server rooms is considered as an effective decision improving energy efficiency in regions with cold climate [21]–[23]. Moreover, the use of economizers allows free-cooling at a higher temperature of outside air [24].

The energy-saving techniques mentioned in the previous two paragraphs require testing their effect on power consumption in data centres. Generally, real data centres cannot be used as a test site because such experiments may violate safety and security protocols. In such cases, modelling allows evaluation of the energy-saving strategies [6], [25]. The idea of this work is to develop a toolbox which enables building models of data centres of any scale and configuration with relative ease. The models should accurately estimate the energy consumption in data centres, along with that, for cooling strategies, it is essential to predict the thermal behaviour of the server rooms under various conditions. In this work, under the thermal behaviour, we mean the evolution of the temperature of all computational nodes and the air temperature in the data centre. And finally, the models should be capable of playing the role of testbeds for different energy-saving policies. To ensure the reliability of conclusions of energy-saving, results produced by the models should be valid and verifiable.

The toolbox bases on the modularity principle, i.e. it is implemented as a set of building blocks which model individual components of a typical data centre, such as processors, local fans, servers, units of cooling systems. The set of building blocks exhibits these properties:

- **extensibility**: the set comprises building blocks for base individual components of a typical data centre and it can be extended by new blocks;
- **reusability**: each block is a complete model of the corresponding component and can be used both as an individual model and as an element of a more complex model; the same block can be used several times;
- **encapsulation**: each block encapsulates all its parameters and equations inside itself and interacts with other blocks and models only via its inputs and outputs;
- **parameters adjustability**: for each block, its internal parameters are set before its integration into a model, parameters for blocks can be taken from specifications of respective components, the toolbox also provides a technique to adjust some blocks parameters using the real data;
- **model evolution**: during the simulation, building blocks calculate their outputs at each time point in response to time-dependent inputs.

The toolbox also provides procedures which use building blocks for constructing models of data centres of different configurations in a relatively easy way. Designed in such a way model is suitable for simulation of the data centre's thermal behaviour and its energy consumption at different levels of detail from servers to the whole data centre and supports examining the different energy-saving strategies applied to the data centre.

To test the reliability of the toolbox and to improve the quality of the model predictions, we used data from a real data centre called SICS ICE located in Luleå, Sweden [26]. The building blocks in the toolbox contains internal parameters which cannot be selected using only specifications. However, the accurate selection of values for these parameters drastically improves the quality of both thermal and energy models. To find out realistic values for the internal parameters, firstly, we have utilised the building blocks to construct the model of Module 1 of SICS ICE research data centre. Next, we have collected real data from the Module 1 such as the utilization of CPUs and their temperature, local fans speed, environmental conditions, energy consumption, etc., and compared those data with results provided by its model after simulation. To reduce the gap between real data and the modelling results, we have applied an optimization method to adjust the internal parameters, after that we have repeated the simulation using the model with adjusted parameters. The comparison of new modelling results with real data has shown that the model constructed from the building blocks simulates the SICS ICE Module 1 quite realistically.

The paper is structured as follows. Section II provides an overview of related works in the area of modelling of data centres and describes novelty and main features of the toolbox under discussion. Section III represents a description of the toolbox, it considers the toolbox structure, the intention and implementation of building blocks, as well as auxiliary blocks calculating environmental conditions, and procedure of creating data centre models using the building and auxiliary blocks. Section IV describes the toolbox verification, it contains a description of the testbed and its model constructed from building blocks, as well it presents real data collected for verification the model and compares the modelling results with actual data. Section V demonstrates the process of adjustment of the internal parameters of the building blocks and displays comparing the modelling results after parameters adjustment. Section VI considers the case studies and section VII concludes the paper.

II. RELATED WORKS

There are many models to research energy efficiency in data centres. For example, in [6], authors take the task of studying the existing literature covering more than 200 energy consumption models from individual components to whole data centres. Although such models are essential in their own right, it is more important that they can be incorporated into energy-efficient techniques [6]. For example, tools for simulation of airflows and temperature distribution inside a data centre, described in [27]–[30], combined with its energy model, could be used in testing energy-efficient control strategies of cooling systems. Authors of [27], [28] simulate thermal distribution and airflow using computational fluid dynamic (CFD). The CFD method requires massive calculations and its integration into real-time control is not reasonable [29]. In [30] authors suggest a model predicting the temperature at points inside a data centre as the superposition of temperatures at all other outlets of racks and computer room air conditioning (CRAC) units. Based on the model, authors in [29] propose the optimized control of CRAC units in a data centre. Article [31] suggests a modelling framework for data centres and a collection of MATLAB routines simulating the heat exchange within a data centre and its energy consumption based on the framework.

The previously mentioned models can help to improve energy efficiency in data centres. However, they are limited to control of only CRAC units' operation modes as they model primarily airflows and air temperature distribution inside a data centre. These limitations can be improved by creating a data centre model, which supports the testing of various energy-saving strategies applicable to both individual servers and the entire data centre. The model should be capable of simulating a data centre on a server level, namely: (1) thermal evolution of computational nodes (CPUs); (2) the performance of local fans attached to the server for cooling the CPUs; and (3) power consumption of the CPUs and the local fans. Also, the model simulates a server room level namely: (1) evolution of the environmental conditions such as airflows, humidity and temperature distribution inside the data centre; (2) performance of CRAC units and other components of the cooling system such as chiller, heat exchanger, pump and cooling tower; and (3) energy consumption of individual components of data centre and total energy consumption based on dynamically changed users' activity (IT load).

There exist comprehensive models of data centres simulating their thermal behaviour and energy consumption. For instance, in [32] authors propose a coherent model which covers data centre operation from CPU to cooling tower and calculates the static operating states (power consumption, the temperature at inlet and outlet of servers, airflow through the servers and water flow from cooling units) of the data centre and its components under various operating conditions. In [33], the authors suggest a flexible simulation platform integrating models of power usage and thermodynamic properties of IT-equipment and data centre facilities on different levels of detail. SimWare [34] is a holistic simulator of warehouse-scale computers which unites detailed temperature, power, and performance models for servers and cooling units and considers the impact of heat recirculation and air supply timing. The models in [33] and [34] differ in approaches to modelling the thermal distribution in the data centre, the first one uses computational fluid dynamics (CFD), whereas, the second utilises a static distribution matrix.

The models mentioned in the previous paragraph are intended to support designers and administrators of data centres in the process of planning new infrastructure or improving the existing ones. They enable the simulation of energy and thermal behaviour of a data centre to assess its efficiency and optimise the design by adapting its parameters after each simulation not during the simulation. The goal of this work is in developing the model capable of substitute for the real data centre at examining the performance of different energy-saving strategies. Most of such strategies consist of tuning control parameters in response to the conditions change or the appearance of particular events. Examples of such parameters are the utilisation of local fans or cooling units, and the workload of computational nodes. During the simulation, the model provides information about data centre operating states at each time point (as model outputs) and takes values of adjustable parameters as the control signals from system implementing energy-saving algorithm (as model inputs).

In [35] authors propose a simulation tool GDCSim, which implements an automated holistic simulation technique providing online feedback for data centre design and management. GDCSim comprises cyber-physical simulation engine providing fast predictive models for the performance of servers and cooling units, the power consumption of data centre components, the temperatures distribution inside the data centre in response to control decisions made during the simulation. The models can't predict the performance and power consumption of local fans thus they are incapable of assessing the efficiency of strategies aimed at maintaining the higher temperature at servers' inlets. Such strategies enable reducing the energy consumption of room cooling units, whereas the energy consumption of local fans increases to maintain the allowable temperature of CPUs at the cost of increased rotation speed. In fact, due to the huge number of local fans in the data centre, the energy cost of the higher fan rotation speed can eventually surpass the savings of reducing CRAC units' power consumption. In this work, we propose a toolbox applicable for evaluating the strategies related to performance and power consumption of local fans.

Summarising all the previous discussion, this work suggests a novel modelling toolbox which enables building models of data centres of any scale and configuration with relative ease. The models are intended to predict the thermal behaviour and energy consumption for different levels of data centre complexity, from a single server to server room. The modular toolbox is implemented as a kit of Simulink blocks, modelling individual components of a typical data centre and exhibiting the properties of extensibility, reusability, encapsulation, parameters adjustability, and dynamical behaviour. The toolbox also provides a technique to adjust blocks parameters, which cannot be selected using only specifications. For a specific data centre, the model is capable to be used to examine different control and maintenance strategies of the datacentre. Before use, the model needs to be tuned by adjustment of parameters based on real data from the data centre. The model is also applicable for evaluating the control and maintenance strategies with relation to local fans and their performance, fault prediction and power consumption. To the best of the authors' knowledge, no other modelling tool provides all the above-listed capabilities for modelling data centres.

III. TOOLBOX DESCRIPTION

This section presents the description of the proposed toolbox for creating models of data centres of any configuration. The toolbox is based on the modularity principle and incorporates the Simulink blocks divided into two groups. The first group includes building blocks representing individual components of a typical data centre such as CPUs and local fans to build server level models, as well cooling blocks and humidifiers to construct models of data centre cooling systems. The second group includes supporting blocks calculating the environmental conditions inside the data centre such as air temperature, airflow and relative humidity. And finally, the toolbox provides procedures which use building blocks for constructing models of data centres of different configurations in a relatively easy way. The procedures implemented as Matlab scripts build models of data centre components of different levels from server to the server room. They take as input the information about component configuration, for example, type and number of CPUs and local fans building-up a server or type and number of servers in a rack or type and number of racks and cooling units constituting the server room.

A. BUILDING BLOCKS

1) CPU

The CPU block models power consumption and thermal behaviour of a real processor unit. At each simulation time point, the block calculates CPU power consumption, CPU temperature, and amount of heat generated by the CPU. These three values determine the outputs of the CPU block. To motivate its inputs and internal parameters, it is necessary to describe the internal equations estimating the block outputs.

The first internal equation of the CPU block calculates its power consumption. The change in the CPU's power consumption has a direct impact on change its temperature and generated heat. According to [20] [31], the linear function of the CPU utilization shown in (1) calculates the CPU's power consumption close to its real values.

$$P_{CPU} = P_{idle} + (P_{max} - P_{idle})\, Util \tag{1}$$

Equation (1) shows that the block needs input port setting the CPU utilization ($Util$), and two internal parameters one for the peak power consumption (P_{max}), and another for the power consumption in idle mode (P_{idle}).

According to [20], [31], the CPU thermal evolution can be modelled using (2). To calculate the CPU temperature, standard Simulink block Integrator uses derivative presented in (2), and as the second input, the Integrator block takes the value of the CPU temperature at the starting point of the simulation. Therefore, the initial CPU temperature ($T_{CPU}(0)$) is one more internal parameter of the CPU block.

$$\dot{T}_{CPU} = \frac{P_{CPU} - \frac{1}{R}(T_{CPU} - T_{in})}{C_{CPU}} \tag{2}$$

Here, the value of the current power consumption (P_{CPU}) is calculated inside the block using (1), and the CPU temperature (T_{CPU}) is the value of the temperature at previous time point. The thermal resistance of the CPU (R) is calculated inside the block using (3). The heat capacity of the CPU (C_{CPU}) is an internal parameter of the CPU block. And finally, equation (2) shows that the block needs one more input port for the inlet temperature (T_{in}).

The equation (3) adapted from [20] models the thermal resistance of the CPU. Where the constants C_1 and C_2 depend on the properties of the airflow and the CPU package; the parameter k depends on the level of turbulence in the airflow [20]. The equation (3) also shows that the block needs an input port for the volumetric airflow rate through the CPU (CFM).

$$R = \frac{C_1}{CFM^k} + C_2 \tag{3}$$

The CPU block contains an internal block for calculating the thermal resistance using (3). The parameters k, C_1, and C_2 are determined in experiments with this internal block using the data about thermal resistance from [20].

Finally, the third output which is the heat amount generated by the CPU is calculated using (4).

$$\dot{Q} = c_p \cdot f_{kg} \cdot (T_{out} - T_{in}) \tag{4}$$

where

$$T_{out} = \left(1 - \frac{1}{c_p \cdot f_{kg}} \cdot \frac{1}{R}\right) T_{in} + \frac{1}{c_p \cdot f_{kg}} \cdot \frac{1}{R} \cdot T_{CPU} \tag{5}$$

Equations (4) and (5) show that the block needs two values: the air heat capacity (c_p) and the rate of air masses moving (f_{kg}). The second value can be found by converting the volumetric airflow rate (CFM) from cubic feet per minutes into kilograms per second. However, before converting the CFM value, it should be corrected to determine how much of the airflow from the local fans reach the CPU, therefore one more internal parameter of the CPU block is the correction factor ($kCFM$). In real cases, this value is less than 1 and we apply it to CFM value at the block input before converting it into the kilograms per second. The air heat capacity (c_p) is set as the constant parameter of the block.

The toolbox also provides the auxiliary file containing sets of parameters for different CPU types. At data centre model design, each CPU block added to the model either leave the default values of parameters or set up them from the file.

2) LOCAL FAN

The Local Fan block allows modelling the behaviour of the fans which are attached to the servers and responsible for cooling the processors. The block calculates two outputs: the local fan power consumption, and its airflow rate measured in cubic feet per minute (CFM). The rate of airflow of the local fan enters the input port of the corresponding CPU block and helps to calculate the CPU thermal behaviour. To motivate inputs, and internal parameters of the block, it is necessary to describe its internal equations.

The estimation of the local fan power consumption and its airflow rate are based on the well-known fan affinity laws depicted in (6) and (7):

$$P = P_{max} \cdot \left(\frac{RPM}{RPM_{max}}\right)^3 \tag{6}$$

$$CFM = CFM_{max} \cdot \left(\frac{RPM}{RPM_{max}}\right) \tag{7}$$

Equations (6) and (7) show that the block needs the current local fan speed measured in rotation per minutes (RPM) as input. Also, it needs three internal parameters: the maximum local fan power consumption (P_{max}), the second one is the maximum local fan speed (RPM_{max}), and the maximum rate of airflow (CFM_{max}).

The toolbox also provides the auxiliary file containing sets of parameters for different local fan types. At data centre model design, each local fan block added to the model either leave the default values of parameters or set up them from the file.

3) COOLING

The blocks described in this subsection represent models of cooling units utilised in server rooms of data centres. As in this work, we deal with air-cooling in data centres, the cooling blocks simulate devices which ensure the supply of cold air into the data centre. Most such devices use fans to supply cold air and vary in sources of the cold air, for example, free-cooling systems utilise ambient air, whereas, some other systems use air chilled by the liquid refrigerant. The cooling block identifies the number and parameters of fans to calculate the power consumption of the respective cooling device, as well uses temperature of cold air to estimate the chill amount supplied into the server room. Also, the cooling block calculates the airflow rate measured in cubic meters per second.

The toolbox contains two cooling blocks, namely Global Fan block and SEE Cooler block. The first one models a big fan ensuring the supply of outside air into the server room. It calculates the power consumption of the global fan using the well-known fan affinity law (6) and the amount of chill supplied by a global fan into the server room using (4). In this equation, T_{in} is the temperature of outside air, and T_{out} is the air temperature in the server room. The second block models the behaviour of real coolers: SEE Cooler HDZ using in the SICS ICE data centre. It calculates the power consumption using (6) because the main energy consumers in the SEE Cooler are fans. The block uses (4) to estimate the amount of chill supplied by coolers into the room. In this equation, T_{out} is the air temperature in the server room, and T_{in} is the coolant temperature. Both blocks utilise equation (7) to estimate the airflow rate measured in cubic feet per minute (CFM), and then they convert CFM in cubic meters per second.

The toolbox also provides the auxiliary file containing sets of parameters for different global fan types. At data centre model design, each global fan block added to the model either leave the default values of parameters or set up them from the file.

4) HUMIDIFIER

Maintaining the humidity inside data centres within reasonable limits is vital because too low values can lead to the accumulation of static electricity, whereas too high humidity can cause corrosion or even short circuit [36]. Therefore, data centres utilise humidifiers to control inside air moisture. The toolbox provides the Humidifier block, which models the air washer system. Such systems use a veil of water droplets or wetted surface to cool and humidify the air passing through them as a result of heat and mass exchange between air and water [37], [38]. The Humidifier block models the power consumption of humidifier, the temperature of the air at the air washer outlet and its moisture content which is the ratio of the mass of water containing in the air to the mass of dry air, wherein the water is evaporated [37].

At calculating the power consumption, we assume that the main power consumer in the humidifier is a pump, and it runs at a constant power which is set as the Humidifier block parameter based on the technical specifications.

To calculate the temperature of the humidifier outlet air and its moisture content, we have used the equations (8) and (9) respectively, adopted from [38].

$$T_{out} = T_{in} + \eta \cdot (T_w - T_{in}) \tag{8}$$

$$w_{out} = w_{in} + \eta \cdot (w_w - w_{in}) \tag{9}$$

Here humidifying efficiency of the air washer (η) and the water droplets temperature (T_w) are set as the block parameters, their values can be set in accordance with specification of the selected device. The moisture content of water droplets (w_w) corresponds to 100% relative humidity. Also, equations (8) and (9) demonstrate that the Humidifier block needs following inputs: the temperature of the humidifier inlet air (T_{in}), and its moisture content (w_{in}). The humidifier takes the outside air thus its conditions are inputs to the block. As far as moisture content is concerned, it is not often possible to find information about it, and it is much easier to obtain information about the relative humidity of the air ($RH\%$). Equations (10)-(12) adopted from [37] enable estimating the moisture content (w) in the air based on its relative humidity.

$$RH = \frac{P_V}{P_{SAT}} \cdot 100 \tag{10}$$

$$P_{SAT} = e^{\frac{1500.3+23.5 \cdot T}{234+T}} \tag{11}$$

$$w = 0.622 \cdot \frac{P_V}{P_{atm} + P_V} \tag{12}$$

The equation (12) shows that block needs value of the atmospheric pressure as a parameter.

The toolbox also provides the auxiliary file containing sets of parameters for different air washer types. At data centre model design, each Humidifier block added to the model either leave the default values of parameters or set up them from the file.

5) TOOLBOX EXTENSION

The toolbox can be extended with new blocks. That means if it is necessary to model a new component of a data centre such as RAM card, hard drive or some specific type of servers, the Simulink block can be created and added to the toolbox. To be an adequate representative of the modelled component, the block should calculate the energy consumption, the temperature and the amount of heat generated by the component. Thus, at creation of the block, the first thing is to determine its equations and inner parameters. For this purpose, it is necessary to investigate the technical specification of modelled components, the parameters also can be adjusted using the real data. The second thing is to construct those equations using the standard Simulink blocks and then encapsulate them into the block. When the block is ready, it can be added to the toolbox.

B. BLOCKS MODELLING CONDITIONS INSIDE DATA CENTRES

1) TEMPERATURE

This subsection describes the Temperature block evaluating the air temperature inside a server room. At each simulation time point, the block calculates the air temperature in a specific area inside the server room, for example, at rack inlet or outlet, as well as, at cooler inlet and outlet, and provides those temperature values to control or maintenance system. To describe thermal evolution, we use (13).

$$T_{air} = T_0 + \frac{1}{c_p \cdot m}\left(\sum_{i=1}^{n} k_{R,i} \int_{t_0}^{t_0+\Delta t} \dot{Q}_{Rack,i}(t)\,dt - \sum_{i=1}^{c} k_{C,i} \int_{t_0}^{t_0+\Delta t} \dot{Q}_{Cooler,i}(t)\,dt\right) \tag{13}$$

This equation demonstrates that the block requires input ports establishing the amount of heat supplied by racks (where n is the number of racks) and the amount of chill supplied by coolers (where c is the number of coolers). The heat capacity (c_p) and air mass (m) are the constant parameters of the block. The initial air temperature (T_0) is the internal parameter which depends on initial conditions in the server room.

In equation (13), the coefficients at integrals in the first and the second sums ($k_{R,i}$, $i = 1, \ldots, n$) and ($k_{C,i}$, $i = 1, \ldots, c$) determine the degrees of influence of heat supplied by racks and chill provided by coolers, respectively. These coefficients are internal parameters of the block and they depend on the location of the area where temperature is calculated. For example, the temperature at the inlet of a specific rack is more susceptible to the influence of heat provided by the next racks than by more distant racks. It is possible to assign the values of the coefficients from theoretical reasoning or adjust them based on real data. The toolbox provides the procedure for such coefficients' adjustment when the data of air temperature in the real server room are available.

2) HUMIDITY

The block described in this subsection evaluates the relative humidity and dew point temperature inside a server room. Dew point temperature is the temperature at which the water vapour changes into liquid (condensation). Relative humidity and dew point temperature are important properties of air inside the server room. ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) takes them in account specifying recommended and allowable operational ranges of air conditions [39]. At each simulation time point, the block calculates the relative humidity and dew point temperature in a specific area inside the server room and provides those values to control or maintenance system.

To estimate the relative humidity in a specific area of the server room, the block uses equation (10). This equation together with equations (11) and (12) show that the block requires input ports for the moisture content (w) and temperature (T) in the corresponding area. The block uses calculated relative humidity to estimate dew point temperature utilizing approximate equation (14) adopted from [37].

$$T_{DP} = \frac{4030\,(T + 235)}{4030 - (T + 235) \cdot ln\,(RH/100)} - 235 \quad (14)$$

3) AIRFLOW

This subsection describes the block evaluating the airflow in a specific area of the server room. It utilises equation (15) which takes airflow rates ($f_{m3/s,i}$) from all cooling blocks supplying air into the analysed area and summates them.

$$f_{m/s} = \frac{\sum_{i=1}^{C} k_i \cdot f_{m^3/s,i}}{S} \quad (15)$$

In the equation, C is the number of coolers affecting the airflow, coefficients k_i ($i = 1, \ldots, C$) determine the degrees of influence of the coolers, and S is the square area through which the airflow passes. The coefficients k_i are internal parameters of the block and they depend on the location of the area where the airflow is calculated. It is possible to assign the values of the coefficients from theoretical reasoning or adjust them based on real data. The toolbox provides the procedure for such coefficients' adjustment when the data of the airflow in the real server room are available.

C. CONSTRUCTORS

1) THE SERVER BLOCK CONSTRUCTOR

A typical server comprises processor units, hard drives, RAM cards, and local fans for cooling. This work neglects hard drives and RAM cards hence the Server block is constructed from CPU blocks and Local Fan blocks. Servers often differ from one data centre to another thus we suggest a constructor block instead of a particular server block. The Server constructor builds the Server block based on information about the number of CPUs and local fans comprising the server and their types. To design the Server block, the constructor utilises parameter files specifying the parameters of CPU and Local Fan blocks, according to the selected types. The constructed Server block can be added to the Toolbox for further use. For example, the Toolbox already contains a Server block modelling the server consisting of two CPUs and six local fans, such kind of servers are utilised in Module 1 of the SICS ICE data centre.

The Server block calculates the total power consumption of all local fans and CPUs forming the modelled server, the temperature of each CPU, and the heat amount generated by the server. The block calculates output values at each simulation time point under their input values. The Server block inputs are the workload of all its CPUs, rotational speeds of all server local fans, and the air temperature at the inlet to the server.

2) THE RACK CONSTRUCTOR

In data centres, servers are mounted in racks. Therefore, in our model, rack is a common modelling element. We introduce the Rack block as a container for Server blocks. Its main aim is to organise all the Server blocks forming the rack in the same model location and configure the input data for the Server blocks. Such Server block inputs as CPUs workload and local fans rotational speed are specific for each server in the rack and depend on operating modes of the servers determined by the model user. On the other hand, the temperature at the server inlet is calculated during the modelling and depending on the level of detail some groups of the servers can obtain the same inlet temperature value.

The Rack block allows separating the servers on different groups depending on their location inside the rack such that all servers in one group obtain the same inlet temperature value. Thus, the Rack block constructor needs a number of such groups, named zones, as well as a number and types of servers in each zone as input values. The constructed Rack block can be added to the Toolbox for further use. For example, the Toolbox already contains Rack block modelling the rack comprising three zones with 6 servers in each zone. The similar zoning is used in Module 1 of the SICS ICE data centre, its sensors in front of racks measure the temperature at the inlet to servers in three different levels: top, middle, bottom. The system of sensors collects the real-time data from the module to verify the model.

3) THE SERVER ROOM CONSTRUCTOR

This work considers a typical server room as a group of server racks and a set of cooling units. In this work, we deal also with the cold/hot aisles layout [16]. In such a layout, the racks form rows which localise the cold and hot aisles. The coolers supply cold air to the front of racks in each row (a cold aisle), and the racks exhaust the hot air from the backside into the space between rows (a hot aisle), which isolates the hot air from other spaces of the server room and supplies heat to coolers' return ducts. Understanding the airflow management model is very important because when modelling such a system, it is necessary to consider the distribution of hot and cold air masses within the server room and calculate the air temperature separately for cold and hot aisles.

The intention of the Server Room model is simulation of thermal behaviour inside the modelled server room and power consumption of the server room and its separate components. Thus, the Server Room model constructor needs the number of rack rows, the number and types of racks in each row, as well as, the number and types of coolers corresponding to each rack row or group of rack rows as input values. Besides rack rows and sets of coolers, the Server Room model comprises environmental blocks calculating the air temperature and humidity, as well airflow at inlet and outlet of each rack. For example, this work contains the model for server room consisting of ten server racks organised into two rows, each row provided with two cooling devices. Such a configuration corresponds to Module 1 of the SICS ICE facility. Later, section III describes the Module 1 configuration and its model in detail.

IV. TOOLBOX VERIFICATION

To test and evaluate the reliability of described earlier simulation toolbox, we use it to create a model of the data centre and then compared model behaviour with the real system. The verification process has passed through the following steps. The first one is selecting the testbed and development of its model using blocks and constructors of the toolbox. The second step is collecting data from the testbed and theirpreparation to be used by the model, and configuring the model with consistency to the data namely setting necessary data from the testbed as inputs for the model, setting the period of simulation in accordance with the period in data, and defining the data for comparing with model results. The last step is running the model and comparing the model results with real data. This section describes each one of the abovementioned steps in detail.

A. MODELLING THE TESTBED

The testbed used in this work is a testing and experimentation facility named SICS ICE (Infrastructure and Cloud data centre test Environment). The SICS ICE is a data centre envisioned for different research projects and established by SICS Swedish ICT in partnership with Luleå University of Technology and it is located in Luleå, Sweden [26]. In this study, we focus on the simulation of Module 1 of the SICS ICE facility [40] which is demonstrated in FIGURE 1.

As shown in FIGURE 1 (a), the module is a 5 m×6.5 m room, consists of 10 server racks (Rack 1-10), and 4 cooling devices (SEE Cooler 1-4). The racks assemble the computational capabilities of Module 1, each rack contains 18 servers each of them equipped with four 4TB 7200rpm hard drives, 256GB RAM, two Intel Xeon E5-2620 v3 2.4GHz CPUs and six Delta PFR 60 × 60 local fans [40]. In terms of modelling of thermal behaviour and energy consumption, we use only models of CPUs and local fans as components of server model. In Module 1, racks form two rows of five racks in each, a view of one row is shown in FIGURE 1 (b). Four SEE Coolers HDZ-2 are distributed to supply cold air so that each row of racks is cooled by two coolers as shown in FIGURE 1 (a).

Module 1 provides a system of sensors and meters measuring air conditions and collecting the real-time data applicable for verification of data centre models. First, the temperature and humidity sensor kits Raritan DX2-T3H1 contain three temperature sensors, and the middle sensor is also measuring humidity [40]. They are placed at three different heights in front and rear of all rack cabinets measuring the temperature at their inlets and outlets in three different levels: top,

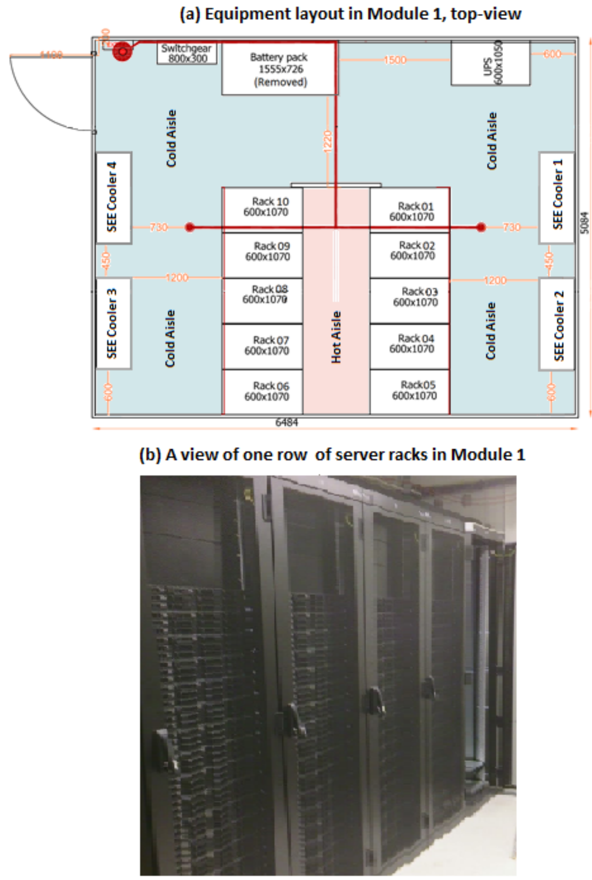


FIGURE 1. A view of Module 1 of the SICS ICE facility.

middle and bottom. Humidity sensors are installed at the middle height, both in front and rear of the rack, and measure relative and absolute humidity at racks inlets and outlets. From datasheet the sensors can measure temperature in the range 20°C to 70°C with accuracy +/– 0.5°C and humidity in the range 0% RH to 100% RH with accuracy +/– 2.5%. Besides that, each rack provides Raritan DPX-AF1 sensor measuring the airflow rate at the middle height of each rack inlet [40]. Data from all the environmental sensors are collected twice a minute (i.e. sampled at 30 s). The energy consumption of racks is measured by vertical Power Distribution Units (Schleifenbauer hPDU) [40], at 300 seconds sampling interval.

To construct the model of Module 1 based on its real layout shown in FIGURE 1 (a), the Server Room constructor embeds 2 rows of 5 Rack blocks into the model. For each row, the constructor also adds to the model 2 SEE Cooler blocks supplying chill to the row. For each rack, the constructor embeds into the model two environment blocks, the first one estimates air temperature, humidity and airflow at the inlet to the rack, and the second one evaluates air temperature, humidity and airflow at the outlet of the rack. These environmental blocks model the system of sensors in Module 1. Each Rack block in the model comprises 18 Server blocks organised into three zones (bottom, middle, and top), of 6 Server blocks in each zone. So we have built the model of Module 1.

B. DATA COLLECTION AND PREPARATION

To find out what data are required to verify the server room model, it is worthwhile to examine the input and output ports of the main blocks forming the model. To simulate Module1, the corresponding data from the real plant must be available in input ports of its main blocks by the time of the simulation. Using these input data, the model generates output data, which have to be compared with the corresponding data from the real plant in the same period. So, the model requires input data for simulation and output data for comparison purposes.

The input ports of the Server block determine the data required for simulating the whole IT-system. As mentioned earlier, the Server block has three input ports, namely, CPUs utilization, local fans speed, and the temperature at the server inlet. Therefore, to simulate the IT-system of Module 1, we have collected data about utilisation reported by all CPUs in the server room, the rotation speed of all local fans reported by servers, and temperature at all racks inlets measured by sensors described in the previous subsection. All these data are collected at 30 seconds of the sampling interval. The gathered data were converted into separate Matlab time series objects saved as MAT files, and each server block obtained as inputs the corresponding MAT files: 2 files with CPU utilisations as each server comprises two CPUs; 6 files with rotation speeds of local fans as each server contains six local fans; 1 file with air temperature at the server inlet, this file is the same for all servers in a specific zone of the rack. These files set the simulation of the IT system.

Using the input data, the Server block generates outputs which determine what real data are required for comparison purpose. The data are the CPUs temperature and server power consumption. Therefore, for comparison purpose, we have collected data about the temperature reported by all CPUs in the server room at each 30 seconds, and energy consumption of all racks measured by PDUs described in previous subsection. After the simulation, the model results can be compared with real data. For each rack, real data about the energy consumption of only the whole rack are available, therefore it is necessary to sum modelled energy consumption of all servers comprising the rack. Thus, for the IT system, the temperature of all CPUs and the energy consumption of all racks can be compared with real data to estimate the accuracy of modelling results.

Similar to the IT-system, model of the cooling system requires input data for simulation and output data for comparison purposes. Since the main component of the cooling system is the SEE Cooler block, its input ports determine the required input data. As mentioned earlier, the SEE Cooler block has three input ports the percentage of the maximum fans' speed, the coolant temperature, and the air temperature at the cooler inlet. To simulate the cooling system of Module 1, we have collected data about utilisation (percentage of maximum rotational speed) reported by all coolers in the server room. The gathered data were converted into separate Matlab time series objects saved as MAT files, and each Cooler block obtained as its first input the corresponding MAT file. The coolant temperature is a constant value, it is equal to 18 °C for the testbed. The third input of Cooler block is calculated by the model for each time point.

The output ports of the Cooler block determine what real data we need for comparison purpose. So, we have collected data about the energy consumption of all coolers measured by the SEE Cooling units with a granularity of 10W sampled at 60 seconds, so that after the simulation, the model results can be compared with real data to estimate the accuracy of modelling. The second output which is the chill amount ejected by cooler into the server room is verified indirectly by applying it to the calculation of air temperature inside the server room in environmental blocks.

The environmental blocks use as inputs the values calculated by other blocks of the model and estimate the air temperature, its related humidity and airflow at inlets and outlets of racks. For comparison purpose, we have collected data about the air temperature, relative humidity and airflow at inlets and outlets of all racks in Module 1. Data about temperature were gathered for three rack zones: bottom, middle, and top, as the temperature sensors are located in such a manner. Humidity and airflow were measured only at middle rack zones. After the simulation, the model results can be compared with real data to estimate the model accuracy.

C. COMPARING SIMULATION RESULTS AND REAL DATA

The final step in model verification is to compare the results of the simulation and the data measured in the real server room. The outputs of the simulation fall into two categories: thermal and energy outputs. The thermal outputs represent the values for the temperature of CPUs as well the air temperature in cold and hot aisles. The energy outputs represent the values of energy consumption of racks and coolers. The modelling results for later comparing are formed by the running two models. The first one simulates the IT-system of Module 1 separately so that all its input ports such as the utilisation of all CPUs, rotation speed of all local fans, and air temperature at inlets of all racks are set to the corresponding real data. The second model simulates the entire Module 1 with a focus on thermal behaviour inside the server room namely air temperature in cold and hot aisles. This model comprises models of all racks in the server room and models of all coolers. It takes real data of all CPUs utilisation, all local fans rotational speed, and all coolers utilisation as inputs. The model calculates air temperature at inlets and outlets of all racks. Later paragraphs in this subsection discuss results of models runs at default values of all blocks parameters.

The run of the first model has generated data of temperature evolution of all CPUs and energy consumption of all racks in the server room. FIGURE 2 shows a comparison of the modelled CPU temperature to corresponding real data. As the model of Module 1 includes 360 CPU blocks, it is impractical to demonstrate results for all CPUs. To simplify the results representation, the mean of absolute values of differences between model and real temperature values

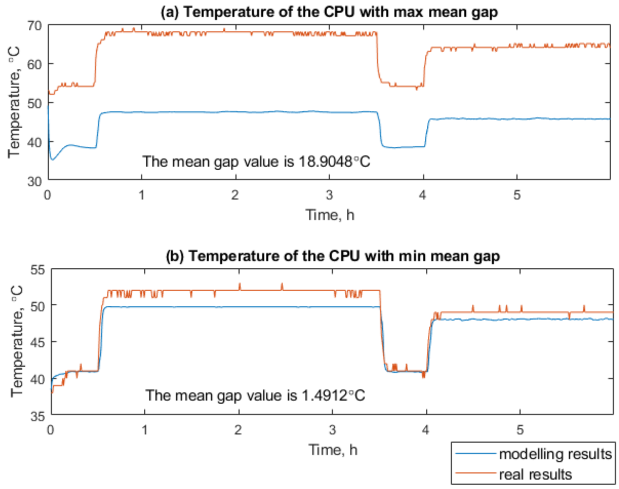


FIGURE 2. **Comparison of modelled and real CPUs temperature.**

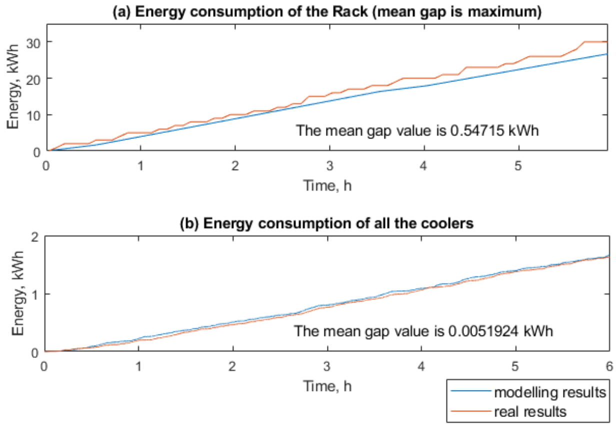


FIGURE 3. **Comparison of modelled and real energy consumption.**

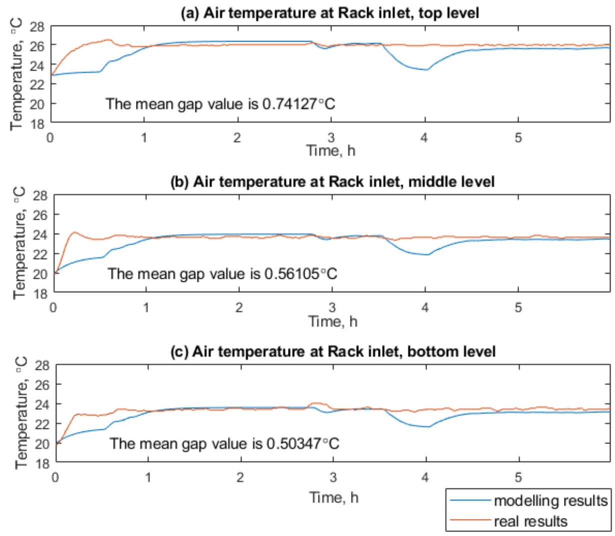


FIGURE 4. **Comparison of modelled and real air temperature in the cold aisle of the server room.**

is calculated for each CPU. FIGURE 2 demonstrates two extreme cases: FIGURE 2 (a) shows a comparison of temperature evolution for CPU with a maximum mean the gap between model results and real data; FIGURE 2 (b) shows a comparison of temperature evolution for CPU with a minimum mean the gap.

Both the graphs demonstrate the similar trends in modelled CPU temperature and real values, however, FIGURE 2 (a) has a significant gap between two curves, whereas in FIGURE 2 (b) the gap is quite small. It can be explained by differences in the airflow organisation over the different CPUs. As mentioned earlier, the CPU block has a parameter ($kCFM$) determining how much of the airflow from the corresponding local fans reaches the CPU. This parameter depends on the server arrangement, its configuration, as well, on some other factors including also the back pressure of the server. In the performed simulation, the $kCFM = 1$ for all CPU blocks, that is, the whole airflow achieves the CPUs. The analysis shows that all CPUs in the model tend to be colder than real CPUs at the same input values, hence, the values of $kCFM$ are overestimated and it is necessary to adjust it for each CPU individually. Next section discusses the adjustment of the parameter $kCFM$ for all CPU blocks integrated into the model.

Concerning rack energy modelling, after comparing the energy outputs of the first run of the model with real data, it is possible to conclude that the model of a rack dependably estimates its energy consumption. FIGURE 3 (a) shows a comparison of the modelled Rack energy consumption to corresponding real data. The figure presents the results of Rack which has a maximum mean gap between modelled and real data of all racks. The graph demonstrates that modelled energy consumption less than real, and it is true for all racks, that because of the model considers only the power consumption of CPUs and local fans, ignoring the rest equipment of servers.

The run of the second model has generated data of air temperature evolution at inlets and outlets of all racks and energy consumption of all coolers in the server room. FIGURE 3 (b) demonstrates a comparison of modelled total energy consumption of all coolers in the server room to real data. The graph shows that the two curves are almost the same and the mean gap is negligibly minor. Such accuracy of modelling can be explained by constructing the Cooler block as a model of particular cooling device: SEE Coolers HDZ-2.

Concerning air temperature modelling, after comparing the rack inlet temperature outputs of the second run of the model with real data, it is possible to conclude that the model estimates the temperature at the rack inlets quite realistic. FIGURE 4 demonstrates a comparison of the modelled inlet temperature at all three levels of Rack 1 to corresponding real data. Average gaps in all three pairs of graphs in FIGURE 4 are less than one degree, however, each pair of graphs has two areas with gaps of about or more two degrees. Modelling results in those areas are lower than the actual data, because of the model calculates the air temperature based on (13), which considers only heat emitted by racks and chill

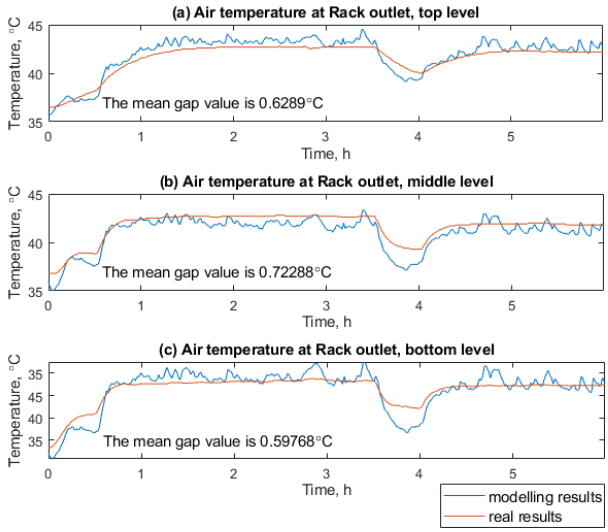


FIGURE 5. Comparison of modelled and real air temperature in the hot aisle of the server room.

supplied by coolers. Analysis the real data used as inputs to the model indicates that coolers operate at approximately constant performance during the whole simulation time while all the racks have close to zero workloads during the time corresponding areas of big gaps between curves so they emit less heat into the air and the model shows lower air temperatures in these periods. The model uses a simplified formula for calculating indoor temperature and does not consider all the influencing factors, but even so, it shows fairly good results.

After comparing the rack outlet temperature outputs of the second run of the model with real data, it is possible to conclude that the model estimates the temperature at the rack outlets quite realistic. FIGURE 5 demonstrates a comparison of the modelled inlet temperature at all three levels of Rack 1 to corresponding real data. Average gaps in all three pairs of graphs in FIGURE 5 are less than one degree, so the model shows fairly good results.

At the end of the section, consider the simulation speed. All models were run on CPU Intel(R) Core(TM) i7-7500U, 2.70GHz, 2 Core(s). The first model comprises ten Rack blocks and calculates only IT system characteristics namely: the energy consumption and the temperature of all 360 CPUs, the energy consumption of all 1080 local fans, and the amount of heat rejected by all 180 servers. The simulation of 6 hours with a fixed-step size equal to 1 second took around 300 seconds at normal simulation mode, and approximately 170 seconds at simulation mode with an accelerator in Simulink. The second model is an extension of the first one, in addition to 10 Rack blocks, it consists of 4 SEE Cooler blocks and 10 Temperature blocks. Besides all characteristics of racks, the model calculates the amount of chill produced by all 4 coolers and their energy consumption. Also, each Temperature block calculates the values of air temperature near the corresponding rack: three values for inlet and three for outlet temperature. The simulation of 6 hours with a fixed-step size equal to 1 second took around 410 seconds at normal simulation mode, and approximately 190 seconds at simulation mode using an accelerator.

V. TOOLBOX TUNING

A. ADJUSTING THE CPU BLOCK

As described in the previous section, the modelled temperature of some CPUs shows a significant gap with the corresponding real data, and this gap can be reduced by the adjustment of the CPU block parameter determining how much of the airflow from the corresponding local fans reaches the CPU. To adjust the parameter value, an optimization method is used. This method utilises an existing MATLAB function called **fmincon**, which finds the global minimum for a constrained nonlinear multivariable function. The **fmincon** function requires the so-called Cost Function, which estimates the mean deviation among modelling results and real values [41], [42]. To optimize the model, this work uses the cost function (J) represented by (16).

$$J\left(k_{CFM}\right) = \frac{1}{2m}\sum\nolimits_{t=1}^{m}\left(T_{CPU,real}\left(t\right) - T_{CPU,model}\left(t\right)\right)^2 \tag{16}$$

In (16), $T_{CPU,real}(t)$ is the real CPU temperature at time t, $T_{CPU,model}(t)$ is the CPU temperature calculated by the model at the same time t, and m is the number of timestamps. Thus, the optimisation method minimises the mean of the squares of the differences between the real temperature values and their values generated by the model.

The **fmincon** function also requires initial values for parameter, and its lower and upper bounds. The initial value of the parameter is set to 1, and $0 < k_{CFM} \leq 1$.

This work uses an auxiliary model to calculate the cost function which is constructed from one CPU block and corresponding Local Fan blocks. The model inputs are MAT-files with real data about the CPU usage, the rotational speed of corresponding local fans, and the temperature at the inlet of the server. The total time of the simulation of the auxiliary model is 6 hours (or 21600 seconds). At each time point, the model calculates the CPU temperature and the square of the difference between the calculated CPU temperature and its real value received from corresponding MAT-file. At the end of the simulation, the vector of the calculated squares of differences is available to calculate the cost function.

The parameter adjustment process goes throughout the following steps. The first step is in the auxiliary model preparation: (1) the initial value of *kCFM* parameter and its lower and upper bounds are determined; (2) the MAT-files with data about the CPU usage, the rotational speed of corresponding local fans, and the temperature at the inlet of the server are set as input values of the auxiliary model; (3) the MAT-file with data about the real CPU temperature is set as a sample for comparison with modelling results. The second step is in running the **fmincon** function which runs the auxiliary model with the current parameter to calculate the cost function value, then generate new parameter value and

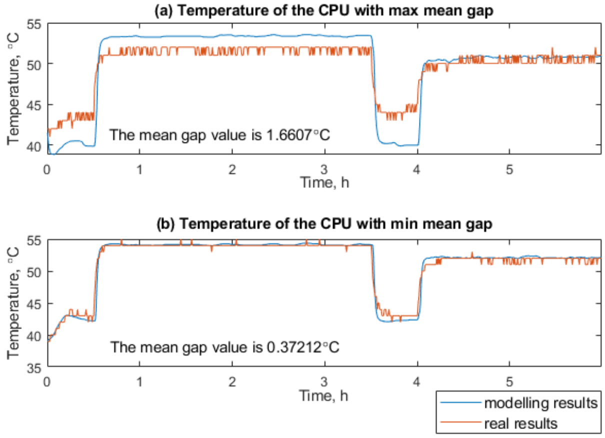


FIGURE 6. **Comparison of modelled and real CPU temperature after optimisation.**

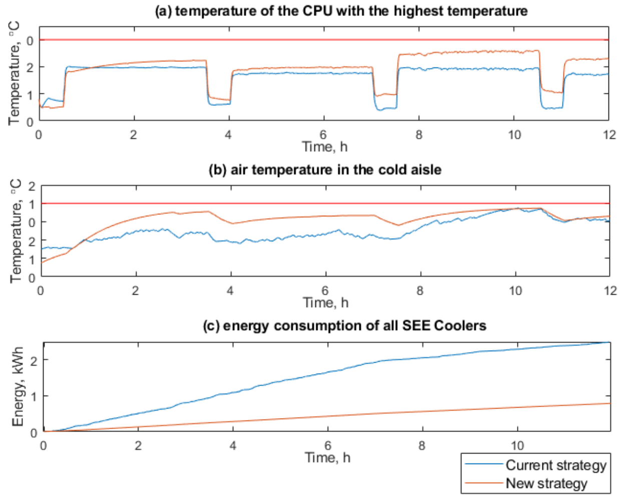


FIGURE 7. **Comparison of modelled and real CPU temperature after optimisation.**

reruns the auxiliary model until the global minimum for cost function is found. The parameter value corresponding the found minimum is considered the adjusted value.

B. RESULTS OF THE CPU BLOCKS' ADJUSTMENT

We applied the procedure of parameters adjustment, described in the previous subsection, to all CPU blocks in the model of Module 1 of the SICS ICE data centre, and after that we ran the model. FIGURE 6 demonstrates the comparison of modelling data generated after the adjustment of the CPU blocks parameters with real values of the temperatures obtained from SICS ICE data centre. The graphs clearly show that modelling results and real data are much closer comparing to the results, which have been generated before the adjustment. Since the maximum mean value of the gap between CPUs temperatures calculated by the model and obtained from the SICS ICE data centre is equal to 1.66 °C it can be concluded that the model calculates the CPUs temperatures fairly reliable.

VI. CASE STUDIES

The previous sections describe the toolbox for the modelling of data centres and present an example of the developed model. Since the model generates data which are fairly similar to the real data, they can be used to examining different cooling strategies on their energy-efficiency. This section demonstrates the usage of the toolbox to construct models for examining different cooling strategies.

One use case is in using the model of Module 1 of the SICS ICE data centre to analyse the current cooling strategy and search for a more energy-saving strategy. Just as a reminder, Module 1 contains two rows of racks and each row is cooled by two SEE Coolers. To reduce the energy consumption of the SEE Coolers, the possible cooling strategy suggests equalising loads of both coolers chilling the specific row of racks. Currently, both SEE Coolers for each row use different fan speeds. In this new strategy, we equalise a load of coolers by setting their speeds to the minimum speed amongst them. With this energy-saving assumption, we simulate the new strategy. Finally, we compared thermal behaviours and energy consumption patterns of current and new strategies.

FIGURE 7 (a) shows the comparison of the temperature of the CPU with the highest temperature, which in this case is CPU2-Server9-Rack, in the aforementioned strategies. The red line in this diagram shows the maximum allowable CPU temperature which is 70 ^{0}C. Also, FIGURE 7 (b) shows the air temperature in the cold aisle for both strategies, and the red line shows the maximum allowable air temperature, which is 27 ^{0}C. FIGURE 7 (c) shows the energy consumption of all SEE Coolers of Module 1 for both strategies. For the 12 hours of simulation time, the current strategy consumes 2.5 kWh, while the new strategy consumes 0.7857 kWh. Therefore, the new strategy saves (2.5034–0.7857=) 1.7177 kWh. Thus, the percentage of energy-saving by the new in 12 hours is 68.6% which is a very considerable amount. And the developed model shows the possibility to examine energy-saving strategies.

More use cases using only particular model blocks from our toolbox are provided in our previous works. For example, in [18] and [19] a simple model of a server room was built and several operational modes of local and global fans were investigated. In [23] a model of a server room with two cooling types was created. The modelled types of cooling were the regular air-cooling and free-cooling with adiabatic humidification. The model was used for exploring the role of flexible humidity control in data centres for energy saving. In [43], the model of Module 1 helps to examine the impact of the distribution of IT load between servers on energy consumption of data centre cooling systems. In [44], the model was used as a substitute for the real system to generate dataset comprising records of both normal and fault cases for further analysis and construction of a fault detection system.

VII. CONCLUSION

In this paper, we presented a modular toolbox for modeling of arbitrary data centres. The toolbox comprises of Simulink blocks which model individual components of a regular data centre. Each block is a complete model of the corresponding component encapsulating all parameters and equations describing its behaviour. The system provides constructors which have access to parameters of the blocks and can set them to values corresponding to the type of modelled component. Therefore, each block can model different types of components. Also, each block interacts with other models only via its inputs and outputs, so it is possible to modify the inner implementation of the block without changing its functionality. Thus, the system is extendable by adding the new modifications of existing blocks as well as by creating new blocks.

The modular modelling system contains constructors which can build models of data centre systems of different levels from server to the server room. These constructors take as input the information about the system configuration, for example, type and number of CPUs and local fans building-up a server or type and number of servers in a rack or type and number of racks and cooling units constituting the server room. Using the server room constructor, we built the model of Module 1 of SICS ICE data centre and compared modelling results with actual data from the data centre. Comparison has shown similar trends in modelling results and real data, however, the building blocks in the toolbox contain internal parameters which cannot be selected using only specifications. The system provides methods to adjust the internal parameters of the building blocks. The comparison of modelling results after adjusting with real data has shown that the model constructed from the building blocks simulates the SICS ICE Module 1 quite realistically. Once configured, the model is ready to use as a substitute for the real data centre for examining different control and maintenance strategies.

Thus, the article describes the modelling toolbox, presents the created with its help model of the specific data centre, and also shows the model adequacy. However, the presented toolbox is more general and enables building models of data centres of different configurations. If some components of the data centres are not represented in the toolbox, their models can be created and added to it. To provide relatively accurate results, the models of new components should be verified. The process of verification requires the data from the component and comparing the modelling results with these data. In some cases, the inner parameters of the model block are not available from technical specifications or other sources, therefore, the block requires the adjustment technique similar to one described in Section V. Thus, the process of preparing the new block adequately representing the data centre component takes the following steps: determining the main purposes of the block, creating the Simulink model which implements the purposes, initial selection of parameters, collecting data from a real device, comparing modelling results and real data, performing adjustment of parameters, if necessary. And time consumed by this process depends on the model equations complexity, number of inner parameters, availability of technical documentation and data from the modelled component, and number of parameters needing the adjustment.

REFERENCES

[1] J. G. Koomey, *Growth in Data Centre Electricity Use 2005-2010*. Analytics Press, 2011.

[2] P. A. Corcoran och Andrae, "Emerginng trends in electricity consumption for consumer ICT," Nat. Univ. Ireland, Galway, Ireland, Tech. Rep., 2013.

[3] M. Avgerinou, P. Bertoldi, and L. Castellazzi, "Trends in data centre energy consumption under the European code of conduct for data centre energy efficiency," *Energies*, vol. 10, no. 10, p. 1470, Sep. 2017.

[4] B. Whitehead, D. Andrews, A. Shah, and G. Maidment, "Assessing the environmental impact of data centres part 1: Background, energy use and metrics," *Building Environ.*, vol. 82, pp. 151–159, Dec. 2014.

[5] P. Bertoldi, "A market transformation programme for improving energy efficiency in data centres," in *Proc. Proc. MELS, Taming Beast, ACEEE Summer Study Energy Efficiency Buildings*, Pacific Grove, CA, USA, 2014, pp. 1–13.

[6] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 732–794, 1st Quart., 2016.

[7] M. Zakarya, "Energy, performance and cost efficient datacenters: A survey," *Renew. Sustain. Energy Rev.*, vol. 94, pp. 363–385, Oct. 2018.

[8] J. Ni and X. Bai, "A review of air conditioning energy performance in data centers," *Renew. Sustain. Energy Rev.*, vol. 67, pp. 625–640, Jan. 2017.

[9] A. Beloglazov, R. Buyya, Y. C. Lee, and A. Zomaya, "A taxonomy and survey of energy-efficient data centers and cloud computing systems," *Adv. Comput.*, vol. 82, no. 2, pp. 47–111, 2011.

[10] T. Guérout, T. Monteil, G. Da Costa, R. N. Calheiros, R. Buyya, and M. Alexandru, "Energy-aware simulation with DVFS," *Simul. Model. Pract. Theory*, vol. 39, pp. 76–91, Dec. 2013.

[11] C. Gunaratne, K. Christensen, B. Nordman, and S. Suen, "Reducing the energy consumption of Ethernet with adaptive link rate (ALR)," *IEEE Trans. Comput.*, vol. 57, no. 4, pp. 448–461, 2008.

[12] T. C. Ferreto, M. A. S. Netto, R. N. Calheiros, and C. A. F. De Rose, "Server consolidation with migration control for virtualized data centers," *Future Gener. Comput. Syst.*, vol. 27, no. 8, pp. 1027–1034, Oct. 2011.

[13] K. T. Rao and P. S. L. S. Kiran och Reddy, "Energy efficiency in datacenters through virtualization: A case study," *Comput. Sci. Technol.*, vol. 10, pp. 2–6, 2010.

[14] L. Liu, O. Masfary, and N. Antonopoulos, "Energy performance assessment of virtualization technologies using small environmental monitoring sensors," *Sensors*, vol. 12, no. 5, pp. 6610–6628, May 2012.

[15] A. H. Khalaj and S. K. Halgamuge, "A Review on efficient thermal management of air- and liquid-cooled data centers: From chip to the cooling system," *Appl. Energy*, vol. 205, pp. 1165–1188, Nov. 2017.

[16] I.-N. Wang, Y.-Y. Tsui, and C.-C. Wang, "Improvements of airflow distribution in a container data center," *Energy Procedia*, vol. 75, pp. 1819–1824, Aug. 2015.

[17] K. Zhu, Z. Cui, Y. Wang, H. Li, X. Zhang, and C. Franke, "Estimating the maximum energy-saving potential based on IT load and IT load shifting," *Energy*, vol. 138, pp. 902–909, Nov. 2017.

[18] A. Mousavi, V. Vyatkin, Y. Berezovskaya, and X. Zhang, "Towards energy smart data centers: Simulation of server room cooling system," in *Proc. IEEE 20th Conf. Emerg. Technol. Factory Autom. (ETFA)*, Luxembourg, U.K., Sep. 2015, pp. 1–6.

[19] A. Mousavi, V. Vyatkin, Y. Berezovskaya, and X. Zhang, "Cyber-physical design of data centers cooling systems automation," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Helsinki, Finland, Aug. 2015, pp. 254–260.

[20] Z. Wang, C. Bash, N. Tolia, M. Marvah, X. Zhu, and P. Ranganathan, "Optimal fan speed control for thermal management of servers," in *Proc. ASME/Pacific Rim Tech. Conf. Exhib. Packag. Integr. Electron. Photonic Syst., MEMS, NEMS (InterPACK)*, 2009, pp. 709–719.

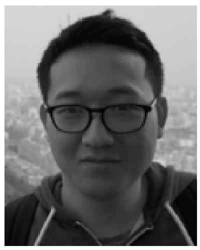
[21] H. Zhang, S. Shao, H. Xu, H. Zou, and C. Tian, "Free cooling of data centers: A review," *Renew. Sustain. Energy Rev.*, vol. 35, pp. 171–182, Jul. 2014.

[22] H. M. Daraghmeh and C.-C. Wang, "A review of current status of free cooling in datacenters," *Appl. Thermal Eng.*, vol. 114, pp. 1224–1239, Mar. 2017.

[23] Y. Berezovskaya, A. Mousavi, V. Vyatkin, X. Zhang, and T. B. Minde, "Improvement of energy efficiency in data centers via flexible humidity control," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Florence, Italy, Oct. 2016, pp. 5585–5590.
[24] J. Cho, T. Lim, and B. S. Kim, "Viability of datacenter cooling systems for energy efficiency in temperate or subtropical regions: Case study," *Energy Buildings*, vol. 55, pp. 189–197, Dec. 2012.
[25] S. M. Rivoire, "Models and metrics for energy-efficient computer systems," Ph.D. dissertation Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, 2008.
[26] *URL of SICS ICE Data Centre Located in Luleå, Sweden*. Accessed: Sep. 13, 2019. [Online]. Available: https://www.sics.se/projects/sics-ice-data-center-in-lulea
[27] J. Ni, B. Jin, B. Zhang, and X. Wang, "Simulation of thermal distribution and airflow for efficient energy consumption in a small data centers," *Sustainability*, vol. 9, no. 4, p. 664, Apr. 2017.
[28] E. Wibron, A.-L. Ljung, and T. Lundström, "Computational fluid dynamics modeling and validating experiments of airflow in a data center," *Energies*, vol. 11, no. 3, p. 644, Mar. 2018.
[29] M. Song, K. Chen, and J. Wang, "Numerical study on the optimized control of CRACs in a data center based on a fast temperature-predicting model," *J. Energy Eng.*, vol. 143, no. 5, Oct. 2017, Art. no. 04017041.
[30] H. Bhagwat, U. Singh, A. Deodhar, A. Singh, and A. Sivasubramaniam, "Fast and accurate evaluation of cooling in data centers," *J. Electron. Packag.*, vol. 137, no. 1, Oct. 2014, Art. no. 011003.
[31] L. Parolini, "Models and control strategies for data center energy efficiency," Ph.D. dissertation, Carnegie Mellon Univ., Pittsburgh, PA, USA, 2012.
[32] G. Zhabelova, M. Vesterlund, S. Eschmann, Y. Berezovskaya, V. Vyatkin, and D. Flieller, "A comprehensive model of data center: From CPU to cooling tower," *IEEE Access*, vol. 6, pp. 61254–61266, 2018.
[33] M. vor dem Berge, J. Buchholz, L. Cupertino, G. Da Costa, A. Donoghue, G. Gallizo, M. Jarus, L. Lopez, A. Oleksiak, E. Pages, W. Piatek, J.-M. Pierson, T. Piontek, D. Rathgeb, J. Salom, L. Sisó, E. Volk, U. Wössner, and T. Zilio, "CoolEmAll: Models and tools for planning and operating energy efficient data centres," in *Handbook on Data Centers*, S. U. Khan and A. Y. Zomaya, Eds. New York, NY, USA: Springer, 2015, pp. 191–245.
[34] S. Yeo and H.-H.-S. Lee, "SimWare: A holistic warehouse-scale computer simulator," *Computer*, vol. 45, no. 9, pp. 48–55, Sep. 2012.
[35] S. K. Gupta, A. Banerjee, Z. Abbasi, G. Varsamopoulos, M. Jonas, J. Ferguson, R. R. Gilbert, and T. Mukherjee, "GDCSim: A simulator for green data center design and analysis," *ACM Trans. Model. Comput. Simul.*, vol. 24, no. 1, 2014, Art. no. 3.
[36] T. Evans, "Humidification strategies for data centers and network rooms," White Paper 58 Rev.1," APC Schneider Electr., West Kingston, RI, USA, 2004.
[37] *ASHRAE Handbook: Fundamentals, SI Edition*, ASHRAE, Atlanta, GA, USA, 2001.
[38] N. Fonseca and C. Cuevas, "Experimental and theoretical study of adiabatic humidification in HVAC&R," *Applicaions Ingeniare. Revista Chilena De Ingeniería*, vol. 18, no. 2, pp. 244–253, 2010.
[39] *2011 Thermal Guidelines for Data Processing Environments—Expanded Data Centre Classes and Usage Guidance*, document ASHRAE TC 9.9, ASHRAE Technical Committee (TC) 9.9 Mission Critical Facilities, Technology Spaces, and Electronic Equipme, 2011.
[40] J. Gustafsson, S. Fredriksson, M. Nilsson-Mäki, D. Olsson, J. Sarkinen, H. Niska, N. Seyvet, T. B. Minde, and J. Summers, "A demonstration of monitoring and measuring data centers for energy efficiency using opensource tools," in *Proc. 9th Int. Conf. Future Energy Syst. (e-Energy)*, Karlsruhe, Germany, 2018, pp. 506–512.
[41] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[42] J. J. Schneider and S. Kirpatrick, *Stochastic Optimisation*, Berlin, Germany: Springer-Verlag, 2006.
[43] Y. Berezovskaya, A. Mousavi, V. Vyatkin, and X. Zhang, "Smart distribution of IT load in energy efficient data centers with focus on cooling systems," in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2018, pp. 4907–4912.
[44] Y. Berezovskaya, C.-W. Yang, A. Mousavi, X. Zhang, and V. Vyatkin, "A hybrid fault detection and diagnosis method in server rooms' cooling systems," in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, Helsinki, Finland, Jul. 2019, pp. 1405–1410.

YULIA BEREZOVSKAYA (Member, IEEE) received the degree in mathematics and computer science from Pomor State University, Arkhangelsk, Russia, in 1999. She is currently pursuing the Ph.D. degree in electrical engineering with the Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Sweden.

From 2001 to 2004, she was a Post-Graduate Student in computational machines and systems with the Department of Applied Mathematics, Pomor State University, Arkhangelsk, Russia. From 2004 to 2015, she was a Senior Lecturer with the Department of Programming and High-Performance Computing, Institute of Mathematics, Information and Space Technologies, Northern (Arctic) Federal University, Arkhangelsk. From 2015 to 2018, she was a Research Assistant with the Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology. Her research interests include the modeling of complex distributed systems and agent-based architectures, novel methods based on cognitive algorithms, machine learning, and bio-inspired heuristics into the agent-based intelligent control systems.



CHEN-WEI YANG received the B.E. degree (Hons.) in computer systems engineering, the M.E. degree (Hons.) from The University of Auckland, New Zealand, in 2011 and 2012, respectively, and the Ph.D. degree from the Luleå University of Technology, Sweden, in 2018.

He is a Domain Expert in the field of distributed systems engineering, with particular interests in energy systems, datacenters, and manufacturing systems. His research expertise includes the development of SOA systems with application in requirements design, data modeling, systems engineering, requirements validation, model transformation, and substation automation design with an emphasis in IEC 61850 and IEC 61499.

ARASH MOUSAVI received the B.Sc. and M.Sc. degrees in computer science from the Department of Computer Science Electrical and Space Engineering, Luleå University of Technology, Sweden, and the Ph.D. degree in computer science from the National University of Malaysia, in 2011.

He has completed Postdoctoral Fellowship with the Department of Computer Science Electrical and Space Engineering, Luleå University of Technology, from 2013 to 2015. He has been working in the field of computer science as a Senior Researcher, a Senior Lecturer, a Software Engineer, and a System Analyst-Programmer, since 1995. He is currently working as a System Architect and a Product Owner at the IT Department, SCANIA CV AB, Södertälje, Sweden. His research interests include distributed computing, multiagent technology, semantic technology, and intelligent industrial automation.

VALERIY VYATKIN (Senior Member, IEEE) received the Ph.D. degree in applied computer science from the Taganrog State University of Radio Engineering (TSURE), Russia, in 1992, the Dr.Eng. degree in electrical engineering from the Nagoya Institute of Technology, Japan, in 1999, the Dr.Sc. (Eng.) degree in information and control systems from TSURE, in 1999, and the Habilitation degree from the Ministry of Science and Technology of Sachsen-Anhalt, Germany, in 2002. He was a Visiting Scholar with Cambridge University, U.K., and had permanent academic appointments with the University of Auckland, Auckland, New Zealand; and the Martin Luther University of Halle-Wittenberg, Halle, Germany, Japan, and Russia. He is on joint appointment as a Chaired Professor (Ämnesföreträdare) of dependable computation and communication systems with the Luleå University of Technology, Luleå, Sweden, and a Professor of information and computer engineering in automation with Aalto University, Helsinki, Finland. His research interests include dependable distributed automation and industrial informatics; software engineering for industrial automation systems; artificial intelligence, distributed architectures, and multiagent systems applied in various industry sectors, including smart grid, material handling, building management systems, data centers, and reconfigurable manufacturing.

He received the Andrew P. Sage award for the Best IEEE Transactions paper, in 2012. He has been the Chair of the IEEE IES Technical Committee on Industrial Informatics, since 2016.

TOR BJÖRN MINDE received the master's degree in electrical engineering from the Royal Institute of Technology, Stockholm, 1986. He was a General Manager building up Ericsson Labs and a number of other new research labs within Ericsson. He was the Head of research strategies with Ericsson Corporate Research, involved in defining 4G and 5G. He has led the team that has received a number of awards in the latest years; BDVA i-Space Gold Level 2018, DCD Global Awards 2017 and 2019 for the datacenter initiative of the year, Datacloud Global Award 2019 for the edge and smart city test-bed. He has 35 years of experience at Ericsson in mobile communication, cloud, machine learning, and multimedia and applications. He has been an Adjunct Professor of system technology with the Luleå University of Technology, Luleå, since 2001. He is currently the Director and the Head of labs for RISE Space Data Science and Engineering and Datacenter Technology Research Lab. His research interests are sensor, context aware, datacenter, and machine learning technologies. His main focus has been on large-scale cloud and big data infrastructures for energy efficient, autonomous, and smart grid integrated operations.