ORIGINAL ARTICLE

Sim2Win: How simulation can help data centers to benefit from controlling their power profile

Sonja Klingert · Nils Wilken · Christian Becker

Abstract To support the grid and integrate renewables, demand response schemes reward the power flexibility of energy consumers. Data centers can profit from this by using power management techniques on all levels of data center architecture: infrastructure, hardware, workload, applications. Even though lately, demand response with data centers has been well researched, most works focus on just one or two techniques and one or two valorization options. This leaves data centers stranded that are not represented by the specific combinations of assumptions and techniques presented in research, and thus a huge potential remains barely touched. To address this challenge, the goal of the presented work is to provide data centers with a framework that can be flexibly instantiated by each data center to assess its individual demand response potential. To achieve this goal, this work presents Sim2Win, a data center simulation framework that can replay any set of different power management strategies in the face of any set of markets for power flexibility. A part of the framework is then instantiated and applied to the workload of a real high-performance data center. It uses workload shifting and frequency scaling in order to market their flexibility on the EPEX spot market and the secondary reserve market in Germany. The results show that by using the inherent flexibility of their power profile on the EPEX spot market the considered data center in 2014 could have earned savings of 7.3% of their power bill.

N. Wilken (✉) · S. Klingert · C. Becker
University of Mannheim, Schloss, 68131, Mannheim, Germany
e-mail: wilken@es.uni-mannheim.de

S. Klingert
e-mail: klingert@uni-mannheim.de

C. Becker
e-mail: christian.becker@uni-mannheim.de

Introduction

Data center energy consumption continues to be a major challenge of the digitization of lifestyle and economy. Even though news coverage is slightly reduced compared to the beginning of the decade, a google scholar research of "data center" and "energy consumption" reveals that the number of research works in the area remains rather stable. In the early years, research focused exclusively on *saving* energy and power in data centers by setting up more efficient equipment. First, results of these research activities

can already be observed in practice. A recent analysis (Avgerinou et al. 2017) showed that the energy efficiency of many data centers in Europe constantly increased since 2008. Since then, attention has also been directed at reducing the detrimental impact of data center power consumption by *adapting* their power profile to the requirements of the grid and the intermittent supply of renewables.

The market approach to this idea is "demand response," which is the temporary adaptation of power demand in response to a market signal. Today, in Europe, the demand response mechanism is institutionalized on the level of the transmission grid and aimed at maintaining the frequency in a small band around 50 Hz. This frequency is influenced by any power fed into or extracted from the grid; therefore, indirectly, answering to demand response requests enables the system to reduce the curtailment of intermittent renewable power and thus increases their share at power consumption. There are many programs that enable demand response with big industry consumers like data centers both in the USA and the EU. These can be grouped into explicit and implicit demand response programs, also called incentive-based versus price-based. According to Coalition (2014), explicit demand response is contract-based programs where users get directly paid to adapt their power profiles upon specific requests; in Europe, these are issued on primary, secondary, and tertiary reserve markets and capacity markets. Implicit demand response is a price-based reaction to sourcing energy at a higher or lower price, e.g., at a wholesale market like the EPEX stock market. Data centers are good candidates for both explicit and implicit demand response schemes as they are highly automated and technically can adapt their power demand in a fine-grained way. Usually, they can do this using a variety of power management techniques on all levels of data center architecture: infrastructure, hardware, workload, applications.

This area has been well researched with most works focusing on single power management strategies like powering off unused servers or geographically migrating virtual machines. Despite hundreds of works showing a high potential of utilizing power flexibility in data centers, in reality, there is hardly any track of data center demand response, especially in Europe. In Germany, for instance, where the here presented simulation and experiments are located, data centers have been monitored for the last 10 years by the Borderstep Institute for Innovation and Sustainability. They show that between 2007 and 2017 not only the overall number of data centers in Germany has increased from nearly 2000 to about 3000, but that also within this time frame the number of big data centers has doubled (Hintemann 2017). In 2017, the overall energy demand from German data centers was at 13.2 TWh (Hintemann 2018). Just assuming that this was consumed using constant power during the year, this would amount to 1.5 GW. This is a huge load, considering that the peak load in whole Germany was at 80.6 GW [1]. Considering further, that in 2014, Gils identified a theoretical load reduction demand response potential in Germany of around 10% of the country's peak load; this figure is even more impressive (Gils 2014). It has to be noted that Gils' work is based on identifying suitable processes in all economic sectors; processes in data centers had not been included.

There is a bundle of reasons for the gap between theoretical demand response potential and referenced implementation thereof, which ranges from a lack of power flexibility market maturity in Europe (Coalition 2017) over business model obstacles, e.g., for colocation providers, to risk aversion in the data center community (Fernández-Montes et al. 2015; Whitney et al. 2014). Among these is also a lack of awareness of data center management about inherent flexibility inside the data center.

The authors are currently also working on a theoretical framework for using an integrated approach that optimizes a set of power management strategies inside the data center, selling the identified flexibility to a set of power flexibility markets ("power flex markets") (Klingert and Becker 2018). However, the challenge remains familiarizing the management of data centers with the general approach to demand response by helping them to identify flexibility options for their power profile and by presenting them opportunities to offer these on power flex markets. One option to do this and control the implementation risk is to simulate several power management techniques and marketing options before implementing them in reality. This approach is selected for the work presented here. There are other simulators for data centers, but as will be shown in "Related work," none of the generic

---

[1] https://www.agora-energiewende.de/fileadmin2/Projekte/2018/Jahresauswertung_2017/Die_Energiewende_im_Stromsektor_2017.pdf

simulators is targeted at simulating demand response with data centers; simulation approaches, on the other hand, that aim at evaluating demand response schemes are regularly not built generically.

Contribution This work is the first to present Sim2Win, a simulation framework that is targeted at replaying any set of power management strategies for any type of data center in the face of any set of markets for power flexibility ("The Sim2Win simulation framework"). A part of Sim2Win is then instantiated and used to simulate workload shifting and frequency scaling in a German high-performance computing (HPC) environment in order to market their flexibility on the EPEX spot market and the secondary reserve market in Germany ("Implemented simulator based on Sim2Win").

The results show that by using the inherent flexibility of their power profile on the EPEX spot market the considered data center in 2014 could have earned savings of 7.3% of their power bill ("Validation and evaluation").

The paper is organized in the following way: it starts with explaining the merits against related work in "Related work" and then introduces the general architecture of the Sim2Win framework in "The Sim2Win simulation framework." The implemented instance of Sim2Win2 and the considered experimental scenario are presented in "Implemented simulator based on Sim2Win." "Validation and evaluation" deals with validation and evaluation on two German power flexibility markets, and "Conclusion and outlook" finally concludes with a short discussion and outlook.

Related work

The work presented spans two interrelated fields of research. On the one hand, it relates to works dealing with how to make data centers more energy efficient in general and how the techniques developed in that context can be used to participate in demand response markets. On the other hand, it also relates to data center simulation research.

Demand response with data centers

There has been a lot of research on demand response with data center in the last decade. A general overview can be found in Kong and Liu (2015); the survey of Giacobbe et al. (2015) is limited to cloud computing environments. Also European projects as All4Green[2] and DC4Cities[3] have been dedicated to demand response with data centers. Contrary to the presented work, however, they used simulation only for a specific data center and did not present a generic framework.

In this research, area markets for power flexibility are basically modeled along the characteristics of either certainty versus uncertainty and explicit versus implicit demand response. Whereas the market for implicit demand response is generally modeled as a (set of) price vector(s), models of explicit demand response are more complex. Wang et al. (2012), for instance, is an early work of an optimization framework that uses an HPC data center network with geographical load balancing reacting to signals from the utility. They model the US emergency demand response, where the reward is based on the locational market price on the wholesale market. This is not comparable to the secondary reserve market in Germany which is modeled as an example for European reserve markets in the presented paper, where there is an increased complexity through the combination power and energy rewards. To our knowledge, only the research group of the EU project Geyser as in Arnone et al. (2017) deal with the European version of secondary reserve markets. However, rather than looking into the economics of a data center bidding into the reserve market, they take a pure electro-physical point of view.

Grouping the research according to the power management techniques applied, it can be easily seen that many works focus on just one strategy (Le et al. 2016; Ghamkhari and Mohsenian-Rad 2012; Bhattacharya et al. 2013; Wang et al. 2012; Tran et al. 2016; Ghasemi-Gol et al. 2014; Liu et al. 2013). Examples are, among others, Aksanli and Rosing (2014), who use batteries to store energy for times of high energy prices and through peak shaving estimate energy cost savings of $480,000 per year for an event-based simulation of a 21 MW data center. Ghasemi-Gol et al. (2014) developed an approach that uses load shedding, which is achieved by switching off/on servers to

adjust the power consumption. They found that their data center model with an interactive workload using trace-based inter-arrival times can reduce the energy costs by up to 13% when the load shedding optimization is applied, i.e., the electrical load is reduced while response times are still within limits.

Apart from the here presented approach, only few works look into more than one power management technique of shaping a data center's power demand profile. In Liu et al. (2013), electricity costs are minimized by avoiding the critical peak intervals through the use of workload shifting and local renewable energy generation. Using real data traces for a mixed workload (batch and interactive), they show that the application of their algorithms can provide cost savings of up to 40% due to controlling uncertainty. Tang et al. (2013) examine the potential for a very small data center (300 kW) to participate in a demand response program. For the power demand provision, they consider to change both the temperature setpoint of the cooling equipment and then apply an optimal workload dispatch algorithm using regression power models. Recently, Cioara et al. (2018) conducted a simulation-based experiment combining workload shifting, thermal storage facilities, and battery storage within a data center to provide power demand flexibility to demand response markets. Cupelli et al. (2018) used a model predictive control approach, integrating the thermal characteristics of a specific data center testbed to simulate data center optimization as a response to dynamic prices and simulated workload profiling requests using thermal buffering and workload shifting. Also, Arnone et al. (2017) performed a simulation based on a real data center in order to show demand response participation options. They focus on the physical interaction between the data center and the power grid.

What makes all these works different from our approach is that they deal with but one specific use case in one specific data center, whereas we provide a generic framework to enable data centers to assess their individual demand response potential through simulation.

There is one work of Postema and Haverkort (2018) which also proposes a simulation framework for the evaluation of different power management strategies in data centers. However, in contrast to this work, it does not focus on the aspect of assessing the potential to participate in demand response markets.

Another research area that is interesting in the context of demand response with data centers is the area of energy efficiency in data centers in general. As many approaches in this area also use techniques and strategies to change the power profile of a data center, these techniques could also be used to participate in demand response markets. One example is the work of Wilde et al. (2015), who optimized the control of hot-water cooling circuits to simultaneously ensure stable operating conditions and make the data center more energy efficient. To make the results of these approaches comparable, many performance metrics for data centers were developed. One of these metrics, the power usage effectiveness (PUE), is also used in this work to calculate the cooling power consumption. Capozzoli et al. (2014) provide an exhaustive overview of the existing performance metrics. However, as this research area has a different focus than our work, it is only of minor importance.

Data center simulation

In recent years, energy costs have a share of up to 50% at the total operational cost of a data center (Laganà et al. 2018), and a lot of effort was invested in the development of energy- and workload management approaches to make data centers more efficient (Kliazovich et al. 2012). In order to test these before going live, a variety of simulation environments were developed (Calheiros et al. 2011).

Ostermann et al. (2011) proposed the simulation framework *GroudSim*. It was made for scientific workloads and relies on a discrete-event simulation core. The framework provides some basic analysis features for the evaluation of the simulation runs and allows the user to model computational and network hardware, job submissions, component failure, background load, and data center costs. Meisner et al. (2012) developed the *BigHouse* simulation framework. Instead of a micro-architectural model for each server, *BigHouse* uses a combination of several queuing theory and stochastic models to simulate the data center, thus creating a distributed discrete-event simulation core. Just recently, Ahmed et al. (2017) created a simulation environment to show the results of demand response–based scheduling on the trade-off between the workload's energy consumption and performance in terms of execution time. However, as in many cases, there is no cost or benefits at all associated with this approach.

Additionally, the evaluation is based on data from various different sources that are wildly combined. The *GreenCloud* framework, an extension to the Ns2[4] network simulator, was introduced by Kliazovich et al. (2012). This simulator focuses on recording the power consumption of data center components and thus its energy cost. It can be used to simulate two-tier, three-tier, and three-tier high-speed data centers. Recently, Rahmani et al. (2018) developed a model for a modular simulation of the energy consumption of a data center. They laid the focus especially on the detailed modeling of the energy consumption of each component of the data center.

In contrast to the previously mentioned frameworks, *EMUSIM* is a combination of a simulator and an emulator (Calheiros et al. 2013). The emulator of *EMUSIM* is used to create application profiles that are fed into the simulator. The authors designed this framework to improve the evaluation of application behavior when executed on cloud data centers. The simulation part of *EMUSIM* is based on *CloudSim* (Calheiros et al. 2011). *CloudSim* enables the user to model system resources as well as the behavior of each data center component (e.g., virtual machines (VM) vs. resource provisioning). It also provides the possibility to model inter-networked federations of several data centers and not only the simulation of a single data center. This framework is widely used, e.g., by HP Labs in the USA and has been extended frequently, for example, with network models (*NetworkCloudSim* (Garg and Buyya 2011)) or just lately with a physical cooling model (*CoolCloudSim* (Cristian et al. 2018)). *CloudSim* is also the basis of the simulator DCSim (Schulze et al. 2012), which was used as a basis for this work.

In contrast to these frameworks, the presented Sim2Win framework explicitly focuses on simulating the participation of data centers in demand response markets.

The Sim2Win simulation framework

Section "Related work" shows that although there are many general data center simulators they do not extend to modeling demand response with data center. And although there is a plethora of works that deals with demand response with data center, most of

these only focus on one or two particular power management strategies and one or two particular power flex markets. Tapping on the real potential for demand response with data center requires the integration of power management in general with power flex markets in general, i.e., on both direct, contract-based, and indirect, price-based demand response markets. We show how this can be achieved by creating a simulation framework with the following requirements:

– R1: In order to represent various types of data centers, the simulation framework must be enabled to model both batch and interactive workload, being provided via physical servers or VMs.
– R2: In order to provide the user with the possibility to test several different power management strategies, the simulation framework needs to be sufficiently flexible to integrate more than one strategy and to allow for the later addition of new, yet un-identified power management strategies at all levels of the data center architecture (infrastructure, hardware, workload).
– R3: Closely linked to R2, in order to allow for various types of strategies, the simulator must offer starting points for manipulating power at all levels of the architecture: infrastructure, hardware, workload, applications.
– R4: In order to represent the value of a power management strategy for the data center (e.g., frequency scaling), the simulation framework has to provide the possibility to model the impact of such techniques on data center cost, e.g., via run-time models and power consumption models.
– R5: In order to simulate the reaction of the data center to explicit demand response requests, the simulation framework has to provide an event-based component that handles such requests, initiating power management accordingly.
– R6: In order to simulate the reaction of the data center to an implicit demand response, the simulation framework has to be able to continuously adapt to dynamic energy prices.

Architecture

The architecture of the Sim2Win framework builds on the data center simulation framework DCSim (Basmadjian et al. 2013). DCSim was aimed at optimizing the interplay between the smart grid and a virtualized data center for interactive workloads. In order

---

to become a generic solution for supporting demand response with data center via simulation, DCSim was extended to Sim2Win. To address this overall goal and meet the whole set of requirements, the architecture follows a modular design principle. Figure 1 illustrates this approach.

The design follows a tree-like structure that has its root at the *SimulationController* component. Sim2Win's design is basically structured into two parts: Facade and Simulation Core. The *SimulationController* component, the only component in the Facade part, is designed to provide functionality for the control of a simulation. Through its connection to a database, the *SimulationController* is able to conveniently monitor and store important simulation data. All the other design components are located in the Simulation Core part. They provide the features and functionalities which are required to form a working simulation core. The *data center* component represents a complete data center within the simulation framework. As shown in Fig. 1, all other design components in the Simulation Core part are subcomponents of the *data centere* component. The physical hardware components *Server*, *Other Power*

Consumers (OPC), and *Heat, Ventilation, Air Condition (HVAC)* each includes a specific instantiation of a power model for servers, OPC, and HVAC respectively. They are required to fulfill R3 and R4. Obviously, these power models might be interdependent.

The Sim2Win framework uses an event-based internal communication mechanism. Thus, it requires a component that handles the events occurring during a simulation. This task is taken care of by the *EventHandler* component which is also responsible for allocating the data center's hardware resources to the current workload. The *DRRequestHandler* component deals with the communication in the case of explicit demand response (see R5), whereas the *EnergyPrice* component simply weighs the total energy consumed with—static or dynamic—energy prices (required by R6). The workload running on the data center is represented in the lower part of the architecture: batch workload is handled by *BatchJob* components, one for each batch job, parenting *SLA* (where SLA stands for service-level agreement) and *RuntimeModel* components (feeding into requirements R1, R3, and R4). For interactive workload consisting of services (e.g., web services), the design of Sim2Win contains a *Ser-*
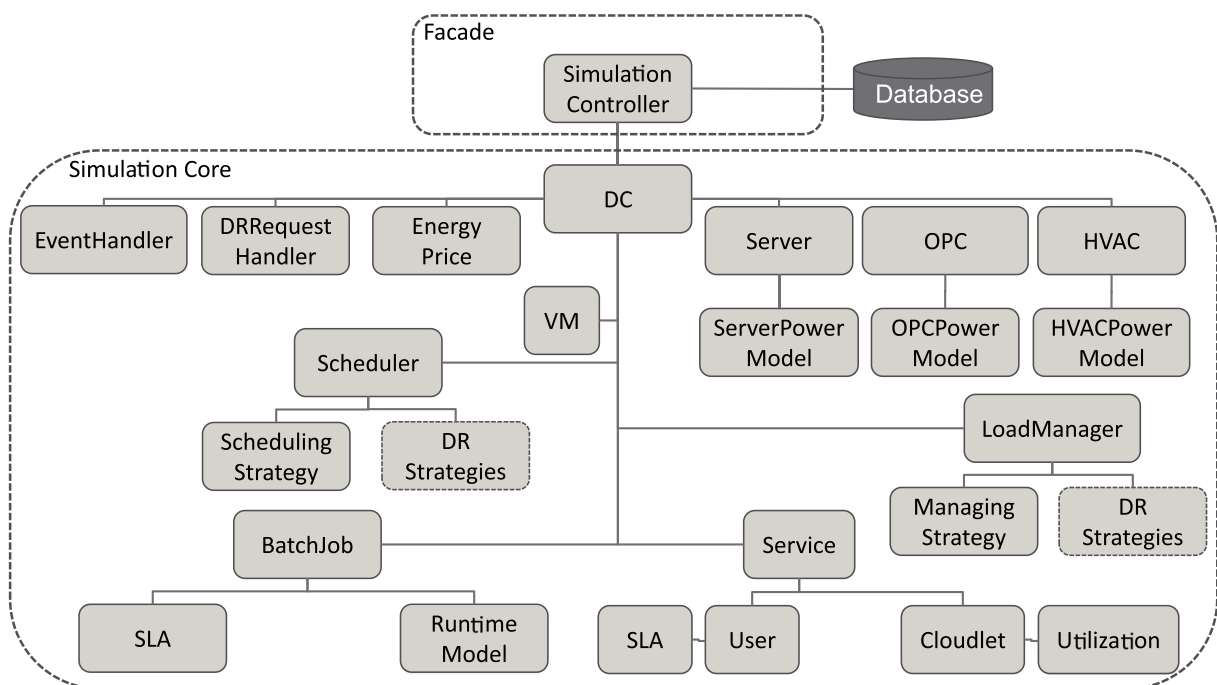

Fig. 1 Overview of Sim2Win design

vice component. It has two subcomponents, namely the *User*, which models one distinct user (and SLA contract) of a service, and *Cloudlet* components representing parts of complete services. Thus, a service can be split into several parts that run on different hardware devices causing different levels of *Utilization*.

The middle layer of the Sim2Win architecture deals with managing the workload of the data center. In the case of batch workload, this is done through the *Scheduler* component, and for the case of interactive workload, this is done through the *LoadManager* component. Both of them are operated differently dependent on "normal" vs. "DR event-based" operation. For example, in the case of batch workload, addressing R2, the *SchedulingStrategy* component defines the regular scheduling of a data center's workload. However, when a demand response event is activated, the *DRStrategy* components define strategies on how to use the pre-defined power management techniques to provide power demand flexibility. The more different types of demand response strategies are available, the more starting points for manipulation of power are available. Thus, the *DRStrategy* components address R4.

The components that are responsible for a data center's reaction to adaptation requests are the scheduler component and the load management component for workload-related power management strategies, the VM component for power management of virtual machines and the physical hardware components for hardware-based adaptation. The way that these power management techniques are implemented depends on the underlying power models in a concrete instantiation of the framework architecture. Such an instantiation will be described in the next section.

Implemented simulator based on Sim2Win

Parts of the framework introduced in "The Sim2Win simulation framework" are instantiated[5] in Java for a German high-performance data center (HPC) with a heterogeneous scientific batch workload running on a cluster of almost identical compute nodes. The characteristics of the concrete simulation instance are highly dependent on the available data: as the data source provided relates to the year 2014, also the evaluation data of the German power flex markets are 2014 data.

The power market structure has not changed since then; the share of renewable-based electricity generation has since then increased to 46% on average in 2019[6]. For the database component, a SQLite database is used. In the following, we present the experimental settings and data traces followed by the models and demand response strategies implemented in this instance.

Experimental settings

In this work, an experimental setting consists of a data center and a demand response market in which the data center participates.

Considered data center

The data traces used for the data center simulation are provided by a large-scale HPC system in Germany and cover the whole year 2014. Therefore, the results are meaningful in the context of a real data center in Germany; being an operating environment, the origin of the data cannot be disclosed.[7] In order to derive more detailed models of some components, e.g., cooling power, it would have been possible to add other data sources and adapt them to the current system; however, for reasons of consistency, this approach was not chosen.

The traces are derived from a homogeneous (in terms of the installed system software stack and system hardware) HPC system with more than 9000 compute nodes, each of which features 2×8 core Intel Sandy Bridge processors having a thermal design power of 130 W[8] and a maximum CPU frequency of 2.7 GHz. The default operating frequency is set to 2.3 GHz. The system uses the IBM LoadLeveler[9] for the management of resources. The workload schedule that is produced by the LoadLeveler system is reused in the simulated baseline scenario. Workload is mainly batch processing with complex algorithmic and computational background. In 2014, the total energy consumption of the considered data center was

---

[5]Code is available upon request.

[6]www.energy-charts.de/ren_share.htm?source=ren-share&period=annual&year=all

[7]Upon personal request data access might be mediated.

[8]The thermal design power quantifies a power value that an electric component should not exceed during operation.

[9]https://www.ibm.com/support/knowledgecentre/en/SSFJTW_5.1.0/com.ibm.cluster.loadl.v5r1.load100.doc/am2ug_ch1.htm

roughly 20,000 MWh; its theoretical peak power is near 4 MW. Cooling technology is hot-water cooling, classified into the ASHRAE W4 (ASHRAE 2011).

Data provided were acquired via a real-time monitoring toolset for the year 2014.

Job data The job data trace contains information for every job that was executed in the operating environment of the considered data center in 2014. It includes among others JobID, submission-, start-, and end-time, allowed maximum frequency, energy (EN), and average power consumption (AP). After data cleaning, the job data trace contained almost 400,000 job records. Job runtimes are very heterogeneous with an average of 3.5 h, a maximum of 52 h and a median of 0.104 h. The same holds true for the EN and AP values, as well as for the occupation of nodes: On average, jobs run on 32 nodes, the mean, however, is 2 nodes only. As expected, the average frequency was 2.38 GHz, very close to the default frequency of 2.3 GHz. Through a simple script, the job data was turned into time series data (see Fig. 2).

PUE and IT power time series The two other data traces available are time series data: The "IT power trace" is the power in KW measured at the main power lines that supply the room which contains servers, storage, network, internal cooling pumps, and PDUs. In 2014, the average IT power consumption was 1892 kW with a standard deviation of 312 kW. The "PUE data trace" contains hourly values for the PUE for the complete year of 2014 so that cooling power could be calculated (Fig. 2). It has a range between 1.06 and 1.35, and regression analyses showed that cooling reacts only little to changes in jobs and IT power and follows rather a seasonal than a diurnal pattern. Some missing values were estimated by linear interpolation.

Considered markets for power flexibility

For the evaluation of the Sim2Win framework, two German power markets were chosen: the EPEX Day Ahead market representing implicit demand response, and the secondary reserve market for explicit demand response.

EPEX Day Ahead market. The EPEX Day Ahead market is a European Exchange market where by trading at 12 pm hourly prices are determined for each hour of the consecutive day.

Secondary reserve market. The European system of reserve power is made up of generally three reserve
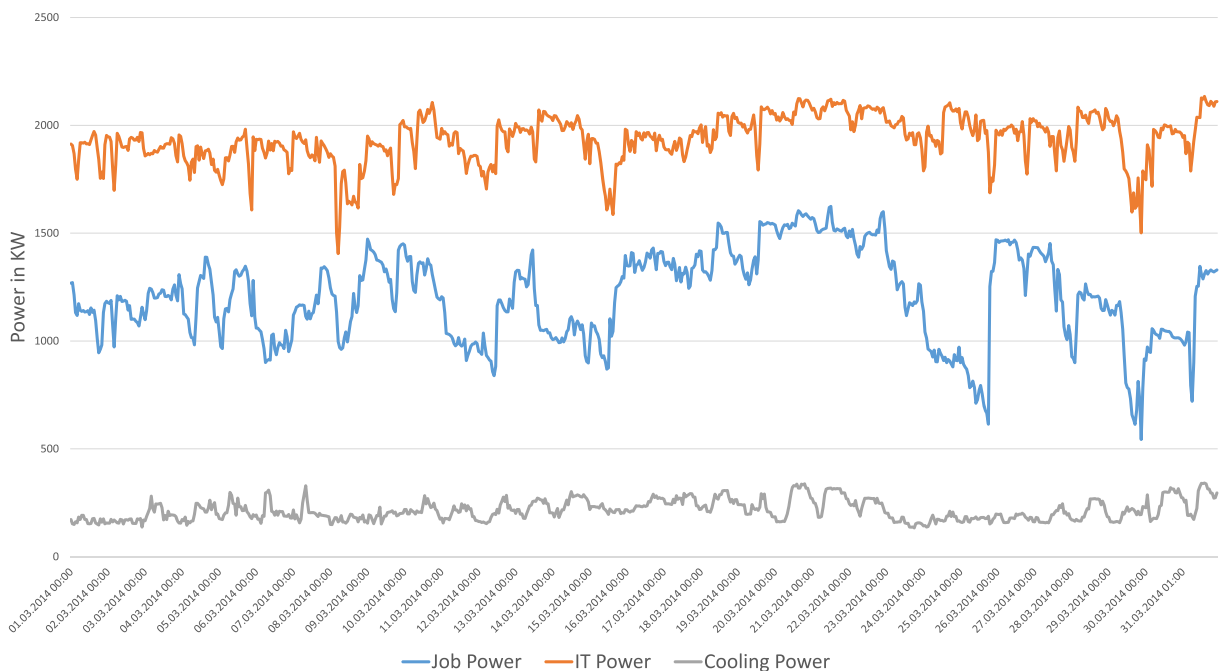

Fig. 2 Data traces of IT, job, and cooling power, March 2014

markets that amend each other: the primary reserve market is servicing unexpected shortfalls of supply and demand until the resources of the secondary reserve market are up and running. They bridge the time until the tertiary reserve market takes over. The German secondary reserve market has an activation period of maximum 15 min with full provision after maximum 5 min. It is auction based; in 2014, auctions were carried out weekly. There are four separate auctions, one for each combination of the provision times (main vs. secondary) and reserve types (positive vs. negative) (Consentec 2014). A bid includes the maximum amount of provided reserve power (in MW), a power compensation price (PP) (€/MW), and an energy compensation price (EP) (€/MWh).

Bids are chosen after the offers are sorted according to their prices and accumulated until the necessary adaptation size is reached (Consentec 2014). As the minimum bid size (5 MW) is much larger than what the considered data center can offer, it is assumed that the data center participates in the secondary reserve market via an aggregator who in return is estimated to keep 30% of the returns. [10]

Market data traces. The data traces used for the secondary reserve market are accessible via the transparency pages of the German transmission operators [11]. For the Day Ahead market, the data traces are sourced from the EPEX Day Ahead website.[12]

Implemented models

There are two sets of implemented models: one refers to power and cost of data center operation and the other to the handling of the power market side.

Server power model

Dayarathna et al. (2016) provide a well-structured overview of existing power models in the data center environment, among these server and server cluster power models. In order to be enabled to use CPU frequency as a starting point for demand response strategies, the server power model of Elnozahy et al.

(2002) was slightly adapted. Server power is here defined as $P_{\mathrm{serv}}(f) = Af^3 + P_{\mathrm{idle}}$, where A is a server and application-specific constant that represents server capacitance and the activity of the server gates, P_{idle} is the server's idle power, and f is the CPU frequency of the server. This implies that we assume servers to be either idling or fully utilized so that server power can be described frequency based. As for this model, data for both server characteristics and the nature of applications are required which our data traces do not contain, finally, a modified version validated on benchmark application data (Shoukourian et al. 2015) was chosen:

$$P_{\mathrm{serv}}(f) = k_1 f^3 + k_2 \qquad (1)$$

where k_1 and k_2 are application- and server-specific fitting parameters (Shoukourian et al. 2015). The structure of the model in Eq. 1 is similar to a linear regression model that uses f^3 as its only variable. This model has the advantages of being closely linked to the causal frequency-based power model on the one hand and being easy to fit to any data center on the other hand.

To fit the model to the available data traces (see "Considered data center") in this work, the WEKA data mining framework is utilized (Hall et al. 2009). However, the model could not be applied directly, as the provided data contains only the maximum allowed values of frequency for each application, not the implemented frequencies. In order to do a model fit, jobs with similar characteristics with regards to the En, AP, and AP/node values are clustered into 30 pseudo-job classes, by the k-means implementation of the WEKA framework. Thus, the job records that end up in one cluster have similar power consumption characteristics, which is beneficial to jointly model the APC/node values of these records, but they cannot be considered to be actually records of the same application type.

Cooling power model

As with all models, obviously, also for cooling models there is a trade-off between explanatory power and data availability. Many data center simulators use thermodynamics-based power models, linking workload, wet-bulb temperature, airflows, and/or server inlet temperatures in order to create a full picture of the causalities involved. However, this also leads to

---

huge requirements for data monitoring which most of today's data centers do not fulfill. The same applies to the data center that provides us with cooling data: they monitor PUE (power utilization effectivenes (Tipley et al. 2009)) traces on an hourly basis; apart from this only outside and wet-bulb temperature values are available. Data analysis shows that the PUE depends on both IT power (P_{IT}) and temperature (more or less equally on wetbulb and outside temperature), which is accounted for by using hourly and not average values to calculate cooling power, so that the following formula can be used:

$$Cooling\,Power = (PUE * P_{IT}) - P_{IT} \qquad (2)$$

Unfortunately, this model does not allow to apply cooling-based demand response strategies and thus reduces the demand response potential in this simulation. Even though this is a drawback of the current simulator, it reflects reality as cooling infrastructure in legacy data centers often consists of heterogeneous technologies; and in these cases, according to many conversations with data center operators, they are just content to see that the cooling environment is working well with all kinds of workload profiles. The modular architecture of the presented data center simulating framework of course allows to easily exchange this model for a more complex one should more data be available.

Other power consumers

The same issue applies to data on the power consumption of other consumers like PDUs, storage equipment, lighting, and network equipment. Server and cooling power generally account only for about 80% of data center power (Rahmani et al. 2018; Hintemann et al. 2017). Most of the components of other power consumers (OPC) depend to a considerable degree on server power; in the currently considered data center, the correlation coefficient is at 0.85. Therefore, due to a lack of fine-grained data, OPC are modeled simply as the fraction of the server power consumption that is equal to the difference between IT power consumption and server power consumption:

$$OPC = \frac{P_{IT} - P_{\text{serv}}}{P_{\text{serv}}} \qquad (3)$$

where P_{serv} is the server power consumption and P_{IT} the IT power consumption, i.e., data that were taken from the monitoring system in the server room.

Its median value is 0.4, meaning that around 71% of the IT power consumption originates in the servers. In the simulation model, this fraction is used to calculate an estimated total IT power consumption from the server power consumption. This calculated IT power consumption is then multiplied with the real PUE to derive cooling and thereby total facility power.

Data center cost

The cost of running the data center is modeled in terms of energy and SLA cost. The energy baseline cost is computed merely by multiplying the consumed energy (kWh) with the baseline energy price (€/kWh) which also contains a fraction of averaged power cost. In order to keep track of peak power changes due to power management strategies, the overall peak power is monitored in the simulation tool and taken into account in cost where necessary.

SLA cost is considered in so far as they are affected by power management strategies. As the SLA model of the considered data center is not to be published, SLA cost was modeled based on Garg et al. (2014): Constructing artificial deadlines based on Garg et al. (2014), a relative delay D is computed that is penalized with a default usage price, so that

$$D = \frac{(AFT - SLA_{DL})}{default\,Runtime} \qquad (4)$$

where $default\,Runtime$ is the runtime of a job as originally specified in the workload trace. The AFT of a job is calculated using the necessary adaptation time for shifting workload and the run-time model explained in "Implemented demand response strategies" for power management through frequency scaling. And SLA_{DL}, finally, is the contractual SLA deadline.

Power flexibility markets

As mentioned in "Considered markets for power flexibility" the Sim2Win framework is evaluated on the market side through the EPEX spot market and the secondary reserve market in Germany.

Secondary reserve market. As explained in "Considered markets for power flexibility," a demand response event on the reserve market is described in terms of provision type (positive or negative),

adjustment height, starting time, duration, and compensation. The product implemented here is positive reserve power. It is assumed that the data center reacts to the request in real time and thus provides the total offered reserve power for the whole timespan of the demand response event. In order to take the decision on how to reply to the request the adjustment height must be below the *maximum technically achievable adjustment height*: This is defined by shifting the entire affected workload out of the demand response event time window and scaling the unshiftable remains to minimum frequency.

When workload is shifted, the new schedule is optimized so as to cause minimal energy and SLA cost, with the energy cost being calculated with the energy price of the timestep at which the demand response request is issued. The optimization procedure evaluates the combination of each possible amount of shifted jobs with all possible scaling frequencies in order to determine the combination that fulfills the requirements of the demand response event and causes minimal costs. Finally, the costs for an adaptation to a demand response event are determined by copying the simulation twice in the starting condition. In one copy, the demand response event is issued, whereas the second copy will not be adjusted and thus executes the original schedule. Subsequently, the simulation is advanced in the two copied instances until the two instances are in the same state again. The additional costs that are caused by the reaction to the

demand response event are then calculated as the difference between the energy and SLA cost of the two simulation instances (see Fig. 3).

EPEX Spot market. In contrast to the secondary reserve market, the EPEX spot market is an implicit market. This means that it is not based on demand response events, but on the implicit signal of dynamic changes in the energy price. As mentioned in "Considered markets for power flexibility," the EPEX Spot market can be used until 12 pm to buy energy for specific hours of the following day. Thus, the data center can optimize its energy costs, by scheduling workload preferably into periods in which the energy price is low. However, in contrast to the secondary reserve market, the data center can decide *voluntarily* to adapt to the dynamic price.

The workload is scheduled in a way that tries to minimize the energy and SLA cost of a job, with the energy cost being calculated with the dynamic EPEX Spot energy price. For each job, the optimization procedure evaluates the combination of all possible start times in the next 24 h in 5 min steps and all possible execution frequencies, in order to determine the combination that induces minimal costs. If a start at the optimal timestep is not possible, it checks less optimal ones where combinations that cause less cost are tried first. If a job cannot be scheduled, the entire scheduling process stops in order to ensure that large jobs will be scheduled eventually.
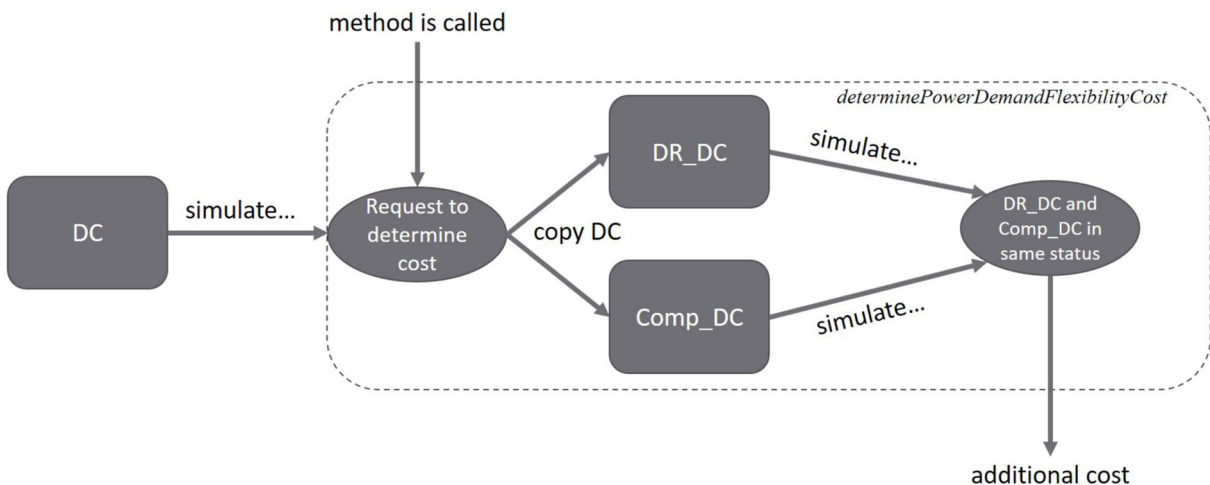

Fig. 3 Procedure to determine additional cost caused by the adaptation to a demand response event

Implemented demand response strategies

The power models implemented in the simulator allow the manipulation of two knobs, thus currently enabling two power management strategies to control the data center's power profile: CPU frequency can be changed, and jobs can be scheduled to another point in time. These two power management strategies are obviously interdependent, as only the CPU frequency of such workload can be manipulated that has not been shifted.

Shifting workload. The shifting strategies in Sim2Win are executed based on the assumption that jobs cannot be halted and resumed later, as the considered data center does not use a virtualization technique. This means that only those jobs that are in the queue but have not started at the considered time slot can be shifted in time. The earliest time they can be resumed is after the necessary adaptation duration. The earliest starting time in the case of preponing is the submission time. The shifting strategy, which is used for explicit demand response, uses a "shortest time to deadline first" heuristic to determine the order in which the jobs should be shifted. This means that in the case of a request of positive power reserve provision, this heuristic is applied by ordering the jobs in a descending order with regards to Θ_{STDF} values, which are calculated as follows:

$$\Theta_{STDF}(x) = \frac{SLA_{DL}(x) - t_{estFinish}(x)}{number Of Nodes(x)} \quad (5)$$

where x is a batch job, $SLA_{DL}(x)$ is the SLA deadline of x, $t_{estFinish}(x)$ is the estimated finish time of x, and $number Of Nodes(x)$ indicates the number of nodes that x utilizes. In the scenario in which no SLA cost is considered, shifted jobs are re-scheduled on a "first come, first serve" basis with backfilling. For implicit demand response, a strategy is used that schedules each job in such a way that the sum of energy cost and SLA cost is minimized.

Frequency scaling. The frequency scaling power management strategy, which is used for explicit demand response, simply scales all jobs to a requested frequency. After a possible demand response event, all jobs are scaled back to their originally specified execution frequency. The strategy, which is used for implicit demand response, determines a frequency for each job such that the SLA and energy costs are minimized. Frequency scaling has an impact on the runtime of a job which influences the SLA cost associated with this strategy (see "Data center cost"). To calculate the "actual finish time," the runtime of a job needs to be assessed. This is done using the concept of computational versus memory boundedness of a job: CPU frequency proportionally increases or decreases computing time—the share of a job that is memory bound, however, is unaffected. This observation leads to a slightly adapted formula of Etinski et al. (2012):

$$\frac{T(f)}{T(f_{max})} = \beta(\frac{f_{max}}{f} - 1) + 1 \quad (6)$$

where $T(f)$ is the job's runtime at frequency f, $T(f_{max})$ is the job's runtime at a nominal frequency f_{max}, and β is a fitting parameter that depends on the degree of memory boundedness of a job. A value of $\beta = 0$ indicates a purely memory-bound application; a value of $\beta = 1$ a purely CPU bound application (Etinski et al. 2012). Fitting the parameter with typical applications from the currently used data center leads to 15 different β values, one for each level of frequency. However, as the spread among those is rather small ($\pm$ 4.5%), an average value is chosen.

Validation and evaluation

In order to first test the correctness of the simulation system, it is *validated* against the real baseline data in "Validation." In order to evaluate the methodology, the following challenge needs to be considered: The data center concerned did not participate in any power flex markets. As a direct comparison of real versus simulated participation of this data center is therefore not possible, we *evaluate* the simulation system against baseline situation (no participation) in "Evaluation" thus testing its usefulness.

Validation

According to Sargent (2004), validation consists of conceptual model validation, model verification, operational validation, and data validation. Data validity was discussed in "Considered data center." Conceptional model validation and verification are partially implied in "Implemented models" as the presented

instance of Sim2Win builds on well-known and validated power and runtime models. Explicitly, validation is carried through by checking the data created by the simulation without power adaptation against all available real data traces for the entire year of 2014. The validation simulation run was executed on a Windows 10 Pro machine with an Intel i7-7600U CPU, which has 2 physical cores at 2.8 GHz, and 16-GB RAM. The validation run, which started at January 1, 2014, and ended at January 6, 2015, needed 558 s (approximately 9.3 min) to complete. As the provided workload trace is on seconds basis, a simulation step length of one real-time second is used. The scheduling interval length is set to one simulation step. This is necessary, because the minimum time difference between the submission time and start time is 0 s in the provided data trace.

Table 1 shows the statistics of this comparison. The high correlation (0.985) and R^2 (0.97) and the low error values between the original data trace and the simulated job data indicate that the simulation reproduces the job power very accurately on the basis of the pseudo-job classes. The accuracy of the simulated IT power consumption, total facility power consumption, and cooling power consumption are not quite as high as for the simulated job power consumption. The reason is that those are based on the IT power consumption of the data center, which contains partially unexplained components, and not on the server data. For the objective of this work, i.e., to demonstrate the value of simulation for a data center to assess the benefit of their inherent flexibility, the accuracy is sufficient.

Evaluation

The current section will show how the simulated data center might have profited from controlling its power profile in March 2014 by valorizing their flexibility on two German power markets. The baseline values against which the demand response simulation runs are compared are given by the real data originating from 2014 provided by the data traces. All evaluation runs were executed on the same machine as the validation run.

The simulation is able to account for two different market side stimuli: event-based adaptation as a form of direct demand response, using the secondary reserve market in Germany as an example. The reason for this choice is that the primary reserve market requires automated adaptation implemented directly by the grid provider—an operational power which a data center will refuse to cede—and the tertiary reserve market rewards amount to merely around half of the rewards on the secondary reserve market (Consentec 2014). And continuous adaptation to hourly changing prices on the European EPEX Day Ahead market as an example for indirect demand response. For the baseline energy price, the static average industrial energy price of 0.1532€/kWh (der Energie-und Wasserwirtschaft 2018) from 2014 is used for the secondary reserve market scenario and the hourly changing price for the EPEX scenario. For the usage price of one compute node hour, the price of 0.36€/node hour, which is the price of a comparable offer (HLRS 2018). This usage price is necessary to calculate the SLA cost on the basis of formula (4).

Secondary reserve market

Consistent with the available data center data trace, 2014 data was used for the secondary reserve market. As explained in "Power flexibility markets," prices on the German reserve markets are determined via bidding for each market participant individually. All successful bids, corresponding PPs (PowerPrice) and EPs

Table 1 Statistics of the comparison between original data traces and simulated data traces

	Correlation	R^2	MAE	MAPE
Job power	0.985	0.97	51.2	4.37%
IT power	0.812	0.659	165.764	10.11%
Cooling power	0.896	0.803	21.508	10.11%
Total facility power	0.808	0.654	187.269	10.11%
Active nodes	0.999	0.999	1.329	0.03%
Running jobs	0.999	0.999	0.007	0.02%

Table 2 Simulation scenarios

Scenarios/power offered	0.2 MW	0.5 MW	0.6 MW
Max bid with SLA	MaxSLA0.2	MaxSLA0.5	MaxSLA0.6
Max bid w/o SLA	Max0.2	Max0.5	Max0.6
Min bid with SLA	MinSLA0.2	MinSLA0.5	MinSLA0.6

(EnergyPrice) are published on *regelleistung.net*[13], which is the internet platform on which the auctions for the different reserve power markets take place. For the evaluation, it is assumed that the data center participated in the auction for positive reserve power provision (main time) from Monday, March 3, 2014, to Sunday, March 9, 2014. This week was chosen, because it is quite representative of the year 2014 in terms of total facility power consumption. The mean total facility power consumption in this week was only 3.6% below the mean total facility power consumption throughout the entire year of 2014. In addition, in this, an interesting volatility pattern of the job power consumption (see Fig. 2) was observed. As the amount of volatility in the job power consumption strongly affects the potential power consumption flexibility of a data center, it is important to use data from a week that includes such volatility patterns to get high-quality evaluation results.

The successful bids in the considered week are used to construct two artificial bids *"as if"* the data center at hand had participated in the secondary reserve market. The historic data state how often secondary reserve power was requested from each participant. Thus, it is possible to reconstruct the distribution of activation events for the artificial bid in terms of number and times.

The artificial bids of the simulated data center are constructed in the following way in order to assess the range of possible financial benefits: The "Max bid" scenario is composed of the maximum accepted PP and the energy price EP that generated the highest income, whereas the "Min bid" scenario combines the minimum accepted PP and the EP that generated the lowest positive income. The maximum accepted PP in this week for positive reserve power was 382€/MW and the EP of the provider that earned the highest revenue through activation compensations

was 63.1€/MWh. The provider that offered this maximum energy price was activated in 90 15-min intervals during the considered week in 2014. The minimum accepted PP in this week for positive reserve power was 271€/MW and the EP of the provider that earned the lowest revenue trough activation compensations was 64.1€/MWh. The provider that offered this minimum energy price was activated in 4 15-min intervals.

Thus, the demand response request trace in the case of the "Max bid" has 90 entries and in the case of the "Min bid" 4 entries, where each entry specifies a demand response event request at the same time at which the real world provider was activated. As the data center has to offer the same product for the entire week, the adjustment height, adjustment length, and provision type values are equal for all entries in the demand response event trace. In order to determine the adjustment height, simulation test runs were carried through. The prognostic power of simulations to assess the adjustment height of power profile adaptations obviously depends on either how well known the workload is by experience or how good the workload forecast of the considered data center is.

The simulations showed that adjustments beyond 700 KW are infeasible, so for each scenario ("Max bid" and "Min bid") three simulations were executed, offering 0.2 MW, 0.5 MW, and 0.6 MW. Representative for these simulation runs, the execution time of the 0.5MW simulation run was recorded. It started at January 1, 2014, and ended at May 15, 2014, and needed a total of 1373 s (approximately 22.9 min) to complete. For the "Max bid" scenario, also a simulation run without SLA cost was carried out, due to the fact that the real SLA cost of the real data center are not to be published: thus, the whole scope of potential benefit from offering power flexibility on the secondary market is shown. In this run, scheduling was done in "first in, first out" order combined with backfilling. Table 2 summarizes the scenarios.

---

[13]https://www.regelleistung.net/ext/tender/

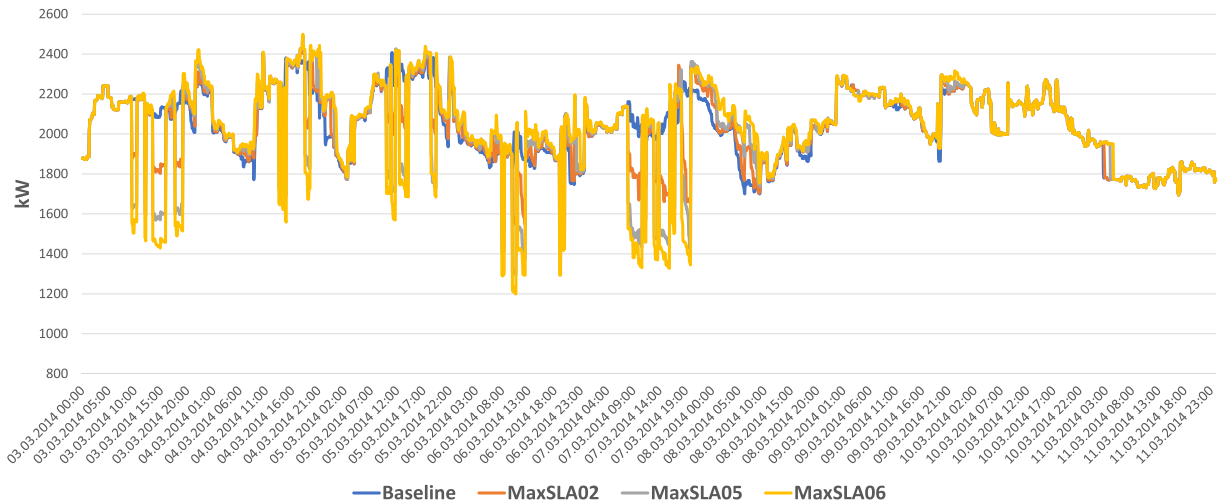$$Benefit_{PP} = Power_{offered} \times PP \qquad (7)$$

Fig. 4 Comparison of facility power between baseline and demand response events

Assuming that for the duration of each event the total offered reserve power is provided, the compensations are calculated by adding the reward for offering the power $Benefit_{PP}$ (7) and the reward for the energy supplied in the events $Benefit_{EP}$ (8).

$$Benefit_{EP} = \#DREvents \times EP \times Power_{offered} \times 0.25 \quad (8)$$

Here, $Power_{offered}$ is the power offered on the reserve market, which in the case of the energy reward must be

turned from power into energy values and multiplied with $\#DREvents$, the number of demand response events.

Results from bidding into the secondary reserve market

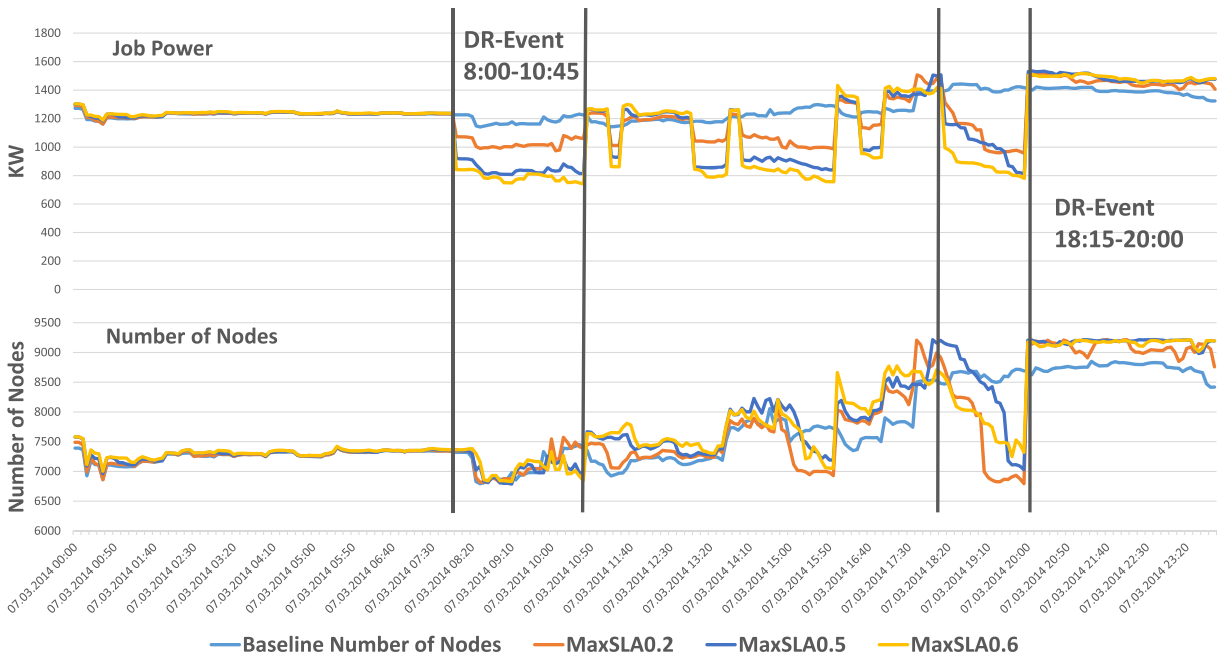First, the three different scenarios described in the first line of Table 2, i.e., the Max bid scenario with SLA

Fig. 5 Comparison between baseline run and demand response simulations, 7th of March

cost for the three reduction levels 0.2 MW, 0.4 MW and 0.6 MW, were carried through. The results are shown in Fig. 4.

It can be seen that some demand response events were directly adjacent to each other, with the highest number of activations on Friday, 7 March, and that adjustments linger a while after the completion of the adjustment week (March 7). This day can be found in detail in Fig. 5. It traces the job power and the number of active nodes for the said scenarios.

There are two noteworthy events depicted in Fig. 5, the first being in the morning (8:00–10:45 hours). Power reduction starts immediately, evoked by a reduction of CPU frequency, which, as assumed, is implemented instantaneously. That the adaptation is frequency-based and not shifting-based can be deduced from the curve of non-idling nodes in the lower part of the picture: For all three scenarios, these remain constant for a couple of minutes, before peaking and then being sharply reduced. This is, because jobs with a low energy consumption (i.e., small and short) are preponed, even though only a few jobs are shifted away from the demand response window. The reason is that inside the activation period, as explained in "Implemented demand response strategies," the job schedule is optimized to keep the load as steadily reduced as possible. The figure also shows, that only in the case of MaxSLA0.6 a considerable part of the workload is shifted away from the demand response window, i.e., the number of active nodes remains well reduced. At the end of the event, as described in many other works (e.g., Palensky and Dietrich 2011), there is a tiny peak, also called "rebound effect" that goes beyond the baseline load recapturing the shifted jobs partially, just before the next activation starts.

The other notable event in Fig. 5, from 1815 to 2000 hours (6:15–8 pm) seems more disruptive. Even though it is a shorter than the first one (i.e., less

energy), when the activation begins, the adaptive data center is still struggling with two issues: One is the great number of former adaptations, and the second is a sharp increase of the baseline job power, i.e., an increase of real demand of a real data center, directly before this last activation. The MaxSLA0.6 simulation is therefore still dealing with shifted jobs and reduces power just by frequency scaling. MaxSLA0.2 and MaxSLA0.4, to the contrary, increased nodes in order to respond to the sudden increase in workload. Therefore, they also react to the reduction request by frequency scaling without job shifting.

On the whole, technically and economically, frequency scaling is much preferred to workload shifting: Because on the one hand, due to the heterogeneity of the workload trace with regards to size and duration, only the small fraction of jobs that are submitted but have not yet been started can be shifted out of the demand response window. And also, with regards to cost, depending on the event duration, frequency scaling is a more fine-grained technique, as it only *scales* the power and runtime of a job instead of reducing its total power and inducing a *"total"* delay. This is also reflected in the SLA cost as can be seen in Table 3, which for all scenarios sums up energy and SLA cost, the power and energy rewards (PP and EP) as well as the gross benefit. The gross benefit is calculated as the sum of the difference between the baseline energy costs and the scenario energy costs, the EP benefit and the PP benefit. The gross benefit percentage is calculated as

$$GrosBenefit/BaselineEnergyCost \times 100 \qquad (9)$$

As only the MaxSLA0.6 run includes shifting to a considerable degree, it results in SLA cost that overcompensates the benefit from the secondary reserve market. This is why the MaxSLA0.5 run maximizes the benefit in this setting: the revenue on the secondary

Table 3 Comparison of costs: MaxSLA scenarios

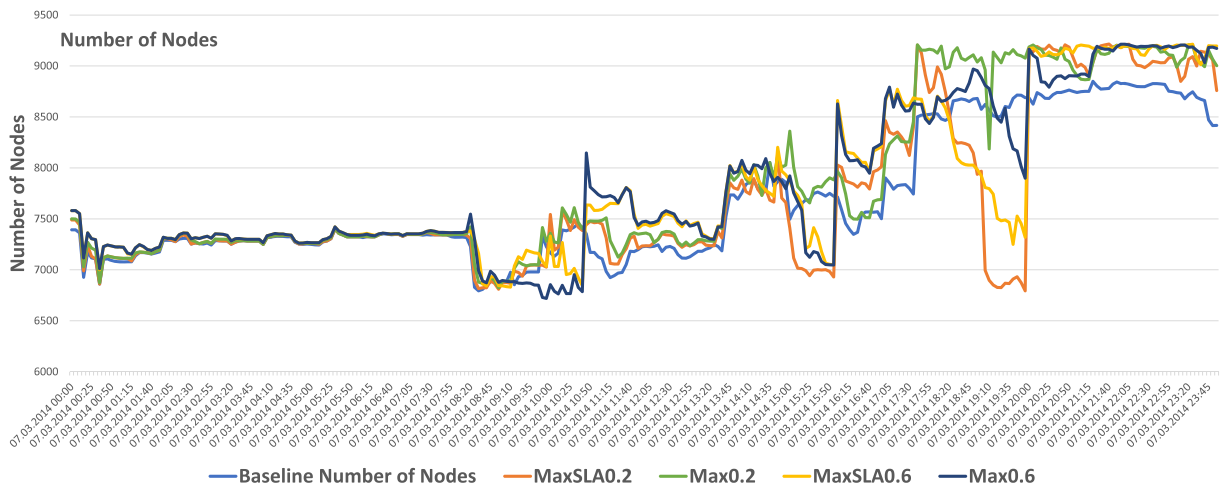|	Baseline	MaxSLA0.2	MaxSLA0.5	MaxSLA0.6
Energy cost	53,283.6	52,760.57	52,371.10	52,327.74
SLA cost	0	176.23	699.12	2,501.33
Benefit EP	0	283.90	709.90	851.90
Benefit PP	0	76.40	191.00	229.20
Gross benefit	0	707.11	1,114.29	−464.35
Gross benefit %	0%	1.3%	2.1%	−0.9%

Fig. 6 Comparison of number of nodes between baseline and demand response events with and without SLA, 7th of March

reserve market amounts to 2.1% of the baseline energy cost.

Therefore, as a sensitivity analysis, the same runs were carried through without activating SLA cost. As expected, the picture does not change much (see Fig. 6). Only as the challenges from the increasing number of activations build-up, there are slight differences between the runs with (MaxSLA0.2 and MaxSLA0.6) and without SLA (Max0.2 and Max0.6). The reason is that only workload shifting leads to noteworthy delays and thus SLA cost—and only in the Max0.6 run, the necessary reduction in the second event is big enough that additionally to frequency scaling, shifting is evoked. However, even then, differences are small due to the technical feasibility of shifting which is impossible for jobs that have already started.

The cost (see Table 4), however, do change; and as again expected, in this case, the highest power reduction offer (Max0.6 scenario) is the most beneficial one, creating an income which is worth 3.8% of the energy cost. As mentioned, the benefit from the secondary reserve market compensation (EP and PP) must be shared with the aggregator (-30%). In the case of the MaxSLA0.5 scenario, that means that the net benefit would be 780.00€ instead of 1114.29€ and thus 1.5% of the total energy costs. Also important to note is that the benefits from the reduced energy cost might be partially compensated by possible rebound effects after the last demand response event.

In order to assess the whole range of possible benefits or losses, the data center could have made by participating in the secondary reserve market, a third set of runs includes the "Min bid" scenarios that were

Table 4 Comparison of costs: Max scenarios

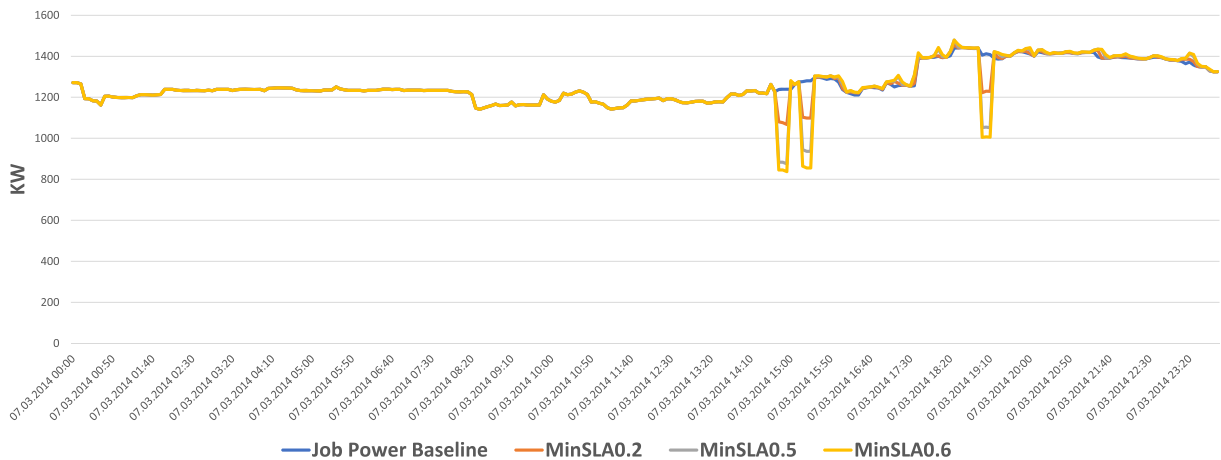|	Baseline	Max0.2	Max0.5	Max0.6
Energy cost	53,283.61	52,470.26	52,303.75	52,320.57
SLA cost	0	0	0	0
Benefit EP	0	283.90	709.90	851.90
Benefit PP	0	76.40	191.00	229.20
Gross benefit	0	1,173.65	1,880.77	2,044.05
Gross benefit %	0%	2.2%	3.5%	3.8%

Fig. 7 Comparison of job power consumption between baseline and remand response events of all "Min bid" scenarios, 7th of March

constructed by the combination of the minimum PP and minimum EP that were activated during the considered week in March 2014 (third line in Table 2). Compared to the 90 events that took place in the "Max bid" scenarios, in the "Min bid" scenarios, there were only four events. The last three of these four events happened on 7 March and are shown in Fig. 7. It can be seen that the job power consumption of the MinSLA scenarios equals the baseline power consumption until the second event on 1445 hours. This is, because the first event happened very early on the 3 March and since then, the data center had enough time to compensate the adjustment it faced in the first event. Similarly, also immediately after the second and third event there is much less difference between the baseline power consumption and the power consumption in the MinSLA scenarios than for the MaxSLA scenarios. This is also caused by the fact that, due to the

heavily reduced amount of demand response events, the data center is much less stressed by the compensation of the effects of the adjustments. Another difference between the load profiles of the MinSLA scenarios and the MaxSLA scenarios is that the power profiles of the different MinSLA runs do not differ as much as they do for the MaxSLA runs. This can also be explained by the much lower compensation effort that the data center faces in the MinSLA scenarios.

Table 5 sums up the costs of these runs. Compared to the MaxSLA scenarios, the benefits of the respective MinSLA scenarios are generally smaller by roughly a factor of 10! The two main reasons for this are that, due to a higher EP, the constructed bid is activated only 4 times (instead of 90 times) and the PP is lower than for the MaxSLA scenarios. Out of the three scenarios, the MinSLA0.6 scenario is the most profitable scenario with 0.4% revenue. Again, please

Table 5 Comparison of costs: MinSLA scenarios

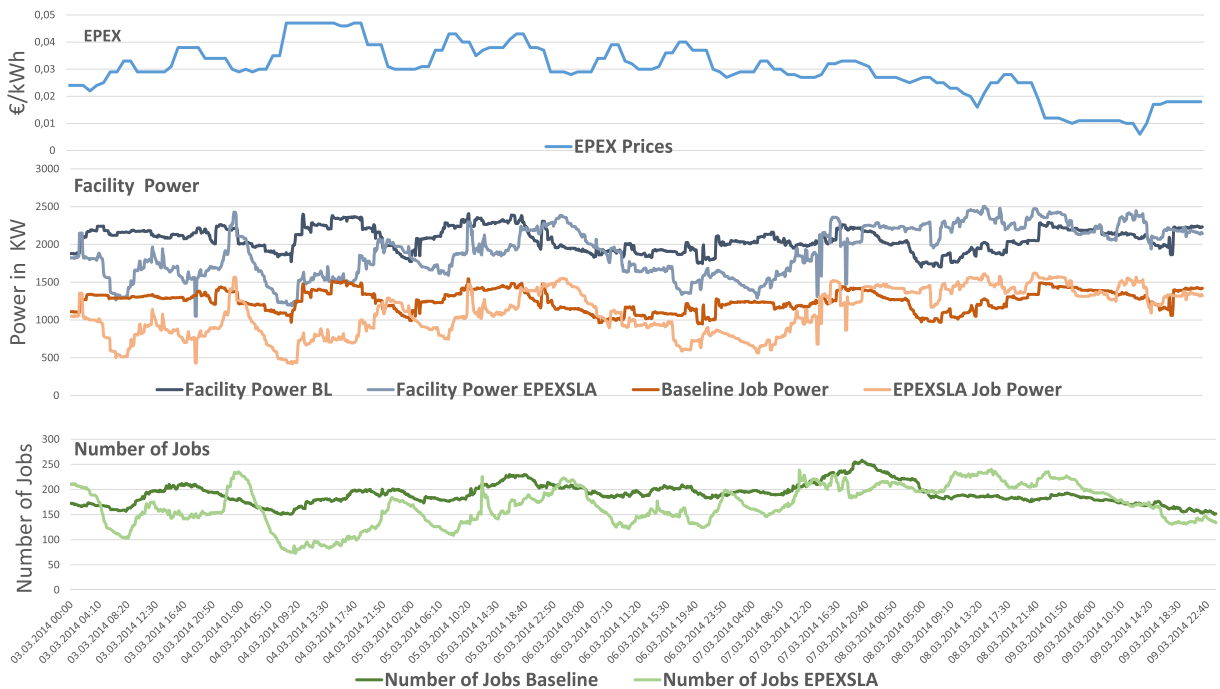|	Baseline	MinSLA0.2	MinSLA0.5	MinSLA0.6
Energy cost	53,283.61	53,259.96	53,243.28	53,240.03
SLA cost	0	10.01	13.87	10.35
Benefit EP	0	12.80	32.10	38.50
Benefit PP	0	54.20	135.50	162.60
Gross benefit	0	80.64	194.07	234.33
Gross benefit %	0%	0.2%	0.3%	0.4%

Fig. 8 Results of the EPEXSLA runs in the considered week in March 2014

note that all the benefits from secondary reserve market compensations are reduced by the aggregator's 30% fee. When this is applied to the benefit of the MinSLA0.6 scenario, the remaining benefit sums up to 164,03€ which equals 0.3% of the baseline energy cost.

EPEX Day Ahead market

Day Ahead market data traces for the EPEXSLA scenario are sourced from the EPEX Day Ahead website.[14] The potential benefit from participating in the EPEX spot market is evaluated for the same week as for the secondary reserve market in order to ensure the comparability of results. Contrary to the event-based simulation of the reserve market participation, here the data center participates by continuously adjusting the schedule of the workload to the dynamic EPEX spot energy price as shown in Fig. 8 using the procedure described in "Considered markets for power flexibility." That means, that rather than an event driven, temporary deviation of the adaptation from the

original schedule, the whole schedule is computed in dependence of the dynamic energy prices.

Results from sourcing on the EPEX Day Ahead market

The simulation run for the EPEX market started at January 1, 2014, and ended at May 15, 2014. However, due to the higher complexity of the considered scheduling procedure, it took a total of 48,629 s (approximately 810.48 min) to complete. The results of the EPEXSLA scenario are summarized in Table 6. For the considered week in March, the baseline energy cost is 10,339.5€, whereas the energy cost of the EPEXSLA run sums to only 9201.75€, albeit at the expense of some SLA costs, resulting in a net benefit of 756.6€. This equals to 7.3% of the baseline energy costs, where the baseline energy costs are calculated by the use of the original schedule of the considered HPC and the hourly changing EPEX price. However, as the used scheduling procedure does not use backfilling, it loses efficiency. Therefore, the data center cannot compute the same amount of workload as in the baseline run, thus also saving energy and thereby energy cost. This effect can be observed in the statistics of the average number of jobs: In the EPEXSLA

---

[14]http://www.epexspot.com/en/market-data/dayaheadauction/

Table 6 Comparison of costs: EPEX scenario

	Baseline	EPEX
Energy cost	10339.5	9201.75
SLA cost	0	381.1
Gross benefit	0	756.6
Gross benefit %	0%	7.3%

scenario, there were on average 11.9% less jobs active than in the baseline run (see Fig. 8). Compared to the benefit this is still reasonably efficient.

Recently, Cupelli et al. (2018) developed the flexible optimizer for data center operations (FLODO) framework. To evaluate the performance of their framework in a price-based demand response scenario, they also considered the EPEX Day Ahead market in Germany and found that the participation in this market reduces the electricity costs of the considered testbed data center by 3.86%. The reasons for the rather large discrepancy between the results of Cupelli et al. and the results of this work (7.3%) are significant differences in the considered scenarios. Where in the here presented work solely batch workload is considered, Cupelli et al. consider three different classes of jobs, which also includes a class for interactive workload that has to be executed in real time. This is the most obvious factor for the discrepancy of benefits, as the introduction of real-time workload heavily reduces the flexibility of a data center in terms of power demand adaptation. In addition, Cupelli et al. use different knobs via which the power consumption of the data center can be controlled: Their framework can (de)charge onsite batteries, adjust the inlet air temperature, or shift workload (except real-time workload), but they do not consider the application of frequency scaling at it is done in this work. And finally, the size of the considered data centers differs. This might also account for the gap between the results due to the efficiency of the scale.

Conclusion and outlook

This work looked into options to support data centers in profiting from their inherent power flexibility by engaging in different power flexibility markets via demand response schemes. To this end, the simulation framework Sim2Win was developed that can be flexibly instantiated to simulate different power management techniques for different data center types to offer their power flexibility on different demand response markets. To the best of our knowledge, the Sim2Win framework is unique in offering this high degree of flexibility and therefore being able to represent a variety of different data centers and data center types. This framework was then instantiated for a specific data center in Germany and subsequently evaluated to assess this data center's individual demand response potential by participating on the EPEX Day Ahead spot market and the secondary reserve market in Germany.

Simulation results show that, had the considered data center bid into the positive secondary reserve market in the week of March 3, 2014, it could have created an income that compares to 2.1% of its total electricity bill of that week. Even bidding with a low power offer and requesting a high energy price into this market, the data center could have gained a modest income. The main risk involved in this strategy would have been to bid too much power, which, however, the simulation tool can help to avoid. An alternative covered by the simulator based on the Sim2Win framework would have been to source the needed power in the EPEX spot market. Here, the data center could have saved up to 7.3% with comparably low SLA cost of 381.1€. The simulator thus illustrates the different opportunities of the data center at hand under different realistic demand response schemes. It shows how the success of a data center's demand response strategy is impacted by concrete situational factors as the heterogeneity of the workload under the assumption that jobs cannot be interrupted. To our knowledge, this is the only simulation work that is based on the combination of a real HPC data center and its real power flex market environment. This concrete evaluation thus additionally evaluated the original approach of creating a simulation framework that empowers data center management to represent their own situation and context.

The main threats to the validity of the HPC simulator created are the quality and detail of data which needed a lot of data cleaning and clustering in order to be useful. Also, due to the level of detail of available data, only two different power management strategies could be evaluated in this work. Certainly, evaluating more strategies is an important point to consider in future work. Also, regarding the concrete evaluation of the Sim2Win framework, the EPEXSLA algorithm, implementing constant optimization, did not include backfilling, which reduces the profitability of indirect demand response. It is part of the envisioned future work to integrate backfilling into the optimization model.

However, the main advancement that could be made through this work is to consistently show the connection between a variable and generic simulation framework architecture of demand response with data centers and its application to a concrete scenario, thus demonstrating the value of the overall approach.

References

Ahmed, K., Liu, J., Wu, X. (2017). An energy efficient demand-response model for high performance computing systems. In *2017 IEEE 25th international symposium on modeling, analysis, and simulation of computer and telecommunication systems (MASCOTS)* (pp. 175–186): IEEE.

Aksanli, B., & Rosing, T. (2014). Providing regulation services and managing data center peak power budgets. In *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014* (pp. 1–4): IEEE.

Arnone, D., Barberi, A., La Cascia, D., Sanseverino, E.R., Zizzo, G. (2017). Green data centres integration in smart grids: new frontiers for ancillary service provision. *Electric Power Systems Research*, *148*, 59–73.

ASHRAE (2011). Thermal guidelines for liquid cooled data processing environments. *ASHRAE TC*, *9*, 9.

Avgerinou, M., Bertoldi, P., Castellazzi, L. (2017). Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency. *Energies*, *10*(10), 1470.

Basmadjian, R., Niedermeier, F., Lovasz, G., De Meer, H., Klingert, S. (2013). Greensdas leveraging power adaption collaboration between energy provider and data centres. In *Sustainable Internet and ICT for Sustainability (SustainIT), 2013* (pp. 1–9): IEEE.

Bhattacharya, A.A., Culler, D., Kansal, A., Govindan, S., Sankar, S. (2013). The need for speed and stability in data center power capping. *Sustainable Computing: Informatics and Systems*, *3*(3), 183–193.

Calheiros, R.N., Ranjan, R., Beloglazov, A., De Rose, C.A.F., Buyya, R. (2011). CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience*, *41*(1), 23–50. ISSN 1097-024X.

Calheiros, R.N., Netto, M.A.S., Rose, C.A.F.D., Buyya, R. (2013). EMUSIM: an integrated emulation and simulation environment for modeling, evaluation, and validation of performance of Cloud computing applications. *Software: Practice and Experience*, *43*(5), 595–612. ISSN 1097-024X.

Capozzoli, A., Chinnici, M., Perino, M., Serale, G. (2014). Review on performance metrics for energy efficiency in data center: The role of thermal management. In *International Workshop on Energy Efficient Data Centers* (pp. 135–151): Springer.

Cioara, T., Anghel, I., Bertoncini, M., Salomie, I., Arnone, D., Mammina, M., Velivassaki, T.-H., Antal, M. (2018). Optimized flexibility management enacting Data Centres participation in Smart Demand Response programs. *Future Generation Computer Systems*, *78*, 330–342. ISSN 0167739X. http://linkinghub.elsevier.com/retrieve/pii/S0167739X16301200.

Coalition, S.E.D. (2014). Mapping demand response in europe today Technical report, Technical report, Brussels.

Coalition, S.E.D. (2017). Explicit demand response in europe - mapping the markets 2017 Technical report, Technical report, Brussels.

Consentec (2014). Beschreibung von regelleistungskonzepten und regelleistungsmarkt. whitepaper.

Cristian, P., Eugen, P., Marcel, A., Pop, C., Cioara, T., Anghel, I., Salomie, I. (2018). Coolcloudsim: Integrating cooling system models in cloudsim. In *2018 IEEE 14th international conference on intelligent computer communication and processing (ICCP)* (pp. 387–394): IEEE.

Cupelli, L.J., Schutz, T., Jahangiri, P., Fuchs, M., Monti, A., Muller, D. (2018). Data center control strategy for participation in demand response programs. IEEE Transactions on Industrial Informatics 14(11).

Dayarathna, M., Wen, Y., Fan, R. (2016). Data center energy consumption modeling: a survey. *IEEE Communications Surveys & Tutorials*, *18*(1), 732–794.

der Energie-und Wasserwirtschaft, B.-B. (2018). Bdew-strompreisanalyse februar 2018. Technical report. https://de.statista.com/statistik/daten/studie/252029/umfrage/industriestrompreise-inkl-stromsteuer-in-deutschland/.

Elnozahy, E.N., Kistler, M., Rajamony, R. (2002). Energy-efficient server clusters. In *PACS*, (Vol. 2325 pp. 179–196): Springer.

Etinski, M., Corbalán, J., Labarta, J., Valero, M. (2012). Understanding the future of energy-performance trade-off via dvfs in hpc environments. *Journal of Parallel and Distributed Computing*, *72*(4), 579–590.

Fernández-Montes, A., Fernández-Cerero, D., González-Abril, L., ÁLvarez-garcía, J.A., Ortega, J.A. (2015). Energy wasting at internet data centers due to fear. *Pattern Recognition Letters*, *67*, 59–65.

Garg, S.K., & Buyya, R. (2011). NetworkCloudSim: modelling parallel applications in cloud simulations. In *2011 Fourth IEEE International Conference on Utility and Cloud Computing* (pp. 105–113).

Garg, S.K., Toosi, A.N., Gopalaiyengar, S.K., Buyya, R. (2014). Sla-based virtual machine management for heterogeneous workloads in a cloud datacenter. *Journal of Network and Computer Applications*, *45*, 108–120.

Ghamkhari, M., & Mohsenian-Rad, H. (2012). Data centers to offer ancillary services. In *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)* (pp. 436–441): IEEE.

Ghasemi-Gol, M., Wang, Y., Pedram, M. (2014). An optimization framework for data centers to minimize electric bill under day-ahead dynamic energy prices while providing regulation services. In *2014 International Green Computing Conference (IGCC)* (pp. 1–9): IEEE.

Giacobbe, M., Celesti, A., Fazio, M., Villari, M., Puliafito, A. (2015). Towards energy management in cloud federation: a survey in the perspective of future sustainable and cost-saving strategies. *Computer Networks*, *91*, 438–452.

Gils, H.C. (2014). Assessment of the theoretical demand response potential in europe. *Energy*, *67*, 1–18.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H. (2009). The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10–18.

Hintemann, R. (2017). *Update2017: Rechenzentren in deutschland: Eine studie zur darstellung der wirtschaftlichen bedeutung und der wettbewerbssituation*. Technical report: Borderstep Institute.

Hintemann, R. (2018). Digitalisierung Treibt strombedarf von rechenzentren: Boom führt zu deutlich steigendem energiebedarf der rechenzentren in deutschland im jahr 2017. Technical report, Borderstep Institute for Innovation and Sustainability.

Hintemann, R., Beucker, S., Hinterholzer, S. (2017). *Energieeffizienz und rechenzentren in deutschland weltweit führend oder längst abgehängt? Technical report*. Berlin: Borderstep Institut für Innovation und Nachhaltigkeit.

HLRS (2018). Lesefassung der hlrs entgeltordnung 2018. online. https://www.hlrs.de/fileadmin/sys/public/solution_services/system_access/legal_requirements/HLRS_Entgeltordnung_Lesefassung_2018.pdf.

Kliazovich, D., Bouvry, P., Khan, S.U. (2012). Greencloud: a packet-level simulator of energy-aware cloud computing data centers. *The Journal of Supercomputing*, *62*(3), 1263–1283.

Klingert, S., & Becker, C. (2018). Economics-inspired modeling of data centre power flexibility. *Computer Science-Research and Development*, *33*(1-2), 247–249.

Kong, F., & Liu, X. (2015). A survey on green-energy-aware power management for datacenters. *ACM Computing Surveys (CSUR)*, *47*(2), 30.

Laganà, D., Mastroianni, C., Meo, M., Renga, D. (2018). Reducing the operational cost of cloud data centers through renewable energy. *Algorithms*, *11*(10), 145. https://www.mdpi.com/1999-4893/11/10/145.

Le, T.N., Liu, Z., Chen, Y., Bash, C. (2016). Joint capacity planning and operational management for sustainable data centers and demand response. In *Proceedings of the Seventh International Conference on Future Energy Systems* (p. 16): ACM.

Liu, Z., Wierman, A., Chen, Y., Razon, B., Chen, N. (2013). Data center demand response: avoiding the coincident peak via workload shifting and local generation. *Performance Evaluation*, *70*(10), 770–791.

Meisner, D., Wu, J., Wenisch, T.F. (2012). BigHouse: a simulation infrastructure for data center systems. In *2012 IEEE International Symposium on Performance Analysis of Systems Software* (pp. 35–45).

Ostermann, S., Plankensteiner, K., Prodan, R., Fahringer, T. (2011). Groudsim: an event-based simulation framework for computational grids and clouds. In Guarracino, M.R., Vivien, F., Träff, J.L., Cannataro, M., Danelutto, M., Hast, A., Perla, F., Knüpfer, A., Di Martino, B., Alexander, M. (Eds.) *Euro-Par 2010 Parallel Processing Workshops* (pp. 305–313). Berlin: Springer.

Palensky, P., & Dietrich, D. (2011). Demand side management: demand response, intelligent energy systems, and smart loads. *IEEE Transactions on Industrial Informatics*, *7*(3), 381–388.

Postema, B.F., & Haverkort, B.R. (2018). Evaluation of advanced data centre power management strategies. *Electronic Notes in Theoretical Computer Science*, *337*, 173–191.

Rahmani, R., Moser, I., Seyedmahmoudian, M. (2018). A complete model for modular simulation of data centre power load. arXiv:1804.00703.

Sargent, R.G. (2004). Validation and verification of simulation models. In *Proceedings of the 36th conference on Winter simulation* (pp. 17–28): Winter Simulation Conference.

Schulze, T., Kessel, M., Giuliani, G., Hesselbach, X. (2012). D6.2 all4green simulation tool descriptio. Technical report, EU Project All4Green. http://all4green-project.gfi.es/sites/default/files/documents/D6.2_final.pdf. accessed 2018-12-20.

Shoukourian, H., Wilde, T., Auweter, A., Bode, A., Tafani, D. (2015). Predicting energy consumption relevant indicators of strong scaling hpc applications for different compute resource configurations. In *Proceedings of the Symposium on High Performance Computing* (pp. 115–126): Society for Computer Simulation International.

Tang, C.-J., Dai, M.-R., Chuang, C.-C. (2013). Exploring the potential benefits of demand response in small data centers. In *Inproceedings of the International MultiConference of Engineers and Computer Scientists*, Vol. 1.

Tipley, R., Cader, T., Froedge, J., Tozer, R., Wofford, R. (2009). Pue scalability metric and statistical analyses. Technical report.

Tran, N.H., Pham, C., Ren, S., Han, Z., Hong, C.S. (2016). Coordinated power reduction in multi-tenant colocation datacenter: an emergency demand response study. In *2016 IEEE International Conference on Communications (ICC)* (pp. 1–6): IEEE.

Wang, R., Kandasamy, N., Nwankpa, C. (2012). Data centers as demand response resources in the electricity market: some preliminary results. In: Intl. Workshop on Feedback Computing.

Whitney, J., Delforge, P., Masanet, E., Peridas, G., Stamas, M., Jimenez, N., Remick, P., Clinger, J., Brown, S.M.T. (2014). Scaling up energy efficiency across the data center industry: evaluating key drivers and barriers. Issue Paper No IP, pp. 14–08.

Wilde, T., Clees, T., Shoukourian, H., Hornung, N., Schnell, M., Torgovitskaia, I., Alvarez, E.L., Labrenz, D., Schwichtenberg, H. (2015). Increasing data center energy efficiency via simulation and optimization of cooling circuits-a practical approach. In *DA-CH Conference on Energy Informatics* (pp. 208–221): Springer.