Making Data Centres Fit for Demand Response: Introducing GreenSDA and GreenSLA Contracts

Robert Basmadjian, Juan Felipe Botero, Giovanni Giuliani, Xavier Hesselbach, Sonja Klingert and Hermann de Meer

Abstract—**The power grid has become a critical infrastructure, which modern society cannot do without. It has always been a challenge to keep power supply and demand in balance; the more so with the recent rise of intermittent renewable energy sources. Demand response schemes are one of the counter measures, traditionally employed with large industrial plants. This paper suggests to consider data centres as candidates for demand response as they are large energy consumers and as they are able to adapt their power profile sufficiently well. To unlock this potential, we suggest a system of contracts that regulate collaboration and economic incentives between the data centre and its energy supplier (GreenSDA) as well as between the data centre and its customers (GreenSLA). Several presented use cases serve to validate the suitability of data centers for demand response schemes.**

Index Terms—**Demand response, Green Service Level (Supply Demand) Agreements, energy management, smart grid.**

I. INTRODUCTION

Traditionally, power systems were designed and operated based on the paradigm of supply following load. While being convenient to the load side, power systems also have proven to be intrinsically suitable for such a mode [1]. To a large extend, power systems adapt autonomously to changing loads, complemented by fast ramping further power plants that can be scheduled to step in, thus providing ancillary services.

More recently, however, some limitations of the supply-following-load paradigm have emerged. First of all, fast ramping power plants tend to be less desirable from emission and cost perspectives [2]. Secondly, an increasing fraction of nuclear plants or carbon based power stations phasing out calls for new sources of supply such as renewable ones.

With the advance of renewable power supplies, new challenges arise once a relative supply threshold is exceeded [3]. P/V-installations and wind mills, for example, are less suitable to the supply-following-load paradigm due to their intrinsic volatile nature and limited controllability.

Furthermore, such sources of power supply tend to be increasingly spread out at the fringes of the power grid, calling

R. Basmadjian and H. De Meer are with the Chair of Computer Networks and Communications, Universität Passau, Innstr. 43, 94032 Passau, Germany (email: robert.basmadjian@uni-passau.de).

J.F. Botero is with the Department of Electronic and Telecommunications Enginnering, Universidad de Antioquia, Calle 67 Nro 53- 108, Medellín, Colombia.

Giovanni Giuliani is with HPE Italy Innovation Center, Milano, Italy.

X. Hesselbach is with the Department of Networks Engineering, Universitat Politecnica de Catalunya, Jordi Girona Street, 1 and 3, Barcelona, Spain.

Sonja Klingert is with the University of Mannheim, A5, B101, Mannheim, Germany.

for more decentralized power management schemes within the distribution grid.

As a promising alternative to the supply-following-load paradigm, demand-response (DR) schemes [4], [5] have been gaining attention. While appearing economically and ecologically highly attractive, DR schemes (or demand side management, for that matter) also pose new challenges.

Most prominent among those challenges are the need for a comprehensive automation and a mitigation of potential risks by interfering with complex physical, economical or technical processes. More specifically, DR schemes require foremost an identification of flexibility potentials of the various processes in a comprehensive fashion and an exploitation of these potentials by means of automation technologies on the basis of economic incentives.

We have identified data centers as viable candidates for exposure within DR schemes. While being seen as energy intensive, data centers are typically equipped with an automation framework that can readily be used for energy and power management. Furthermore, extensive power consumption models exist for all sorts of components and constituents of data centers. And last but not least, substantial flexibility potentials have been identified over the years in the context of energy efficiency studies that could be adapted to a power management context as needed for DR schemes. What is even more of an advantage, formal service level agreements (SLA) are common practice between data center operators and their IT customers (ITC), in addition to specific power and energy supply contracts with a local energy supplier (ES).

We started off by capitalizing on these given favourable precursor conditions, exploring further towards a comprehensive solution for an incorporation of data centers into DR schemes.

The ultimate goal will be a fully automated process that would entail not only the data center with its physical and technological properties but also the extended technical and business contexts of (power) supplier and IT customers' perspectives.

More precisely, we investigate in this paper the concept of power adaptation collaboration in the whole eco-system consisting of ES, data centres (DC) and IT customers (ITC). Within this scenario, power adaptation can exploit the mechanisms of DR, in order to provide a framework where a set of rules agreed by the actors belonging to the eco-system can be defined. The rules are defined under the shape of agreements. Automation facilitates the activation of actions under the conditions defined in the agreements. Conventional methods do not consider any automatic handshake, but a manual reaction or with limited capacity to react, so automation

facilitates a rapid response to power events (such as surplus or shortage) and the possibility to manage a set of agreements under optimisation policies. To this end, we study the use case of private cloud computing from the view of DCs as the location where servers can be stored. We conduct several experiments by configuring numerous scenarios with different energy optimisation policies. The results show the high potential of the proposed concept in playing a major role to foster power adaptation collaboration between ES and DCs. Note that in this paper, we consider medium- to large-sized DCs having more than 200 kW of power consumption. Furthermore, details on the architectural overview of the proposed concept can be found in [6].

Our work makes the following contributions as shown in Fig. 1:

- To foster power adaptation collaboration between ES and DC, we propose a novel Green Supply Demand agreement (GreenSDA), which defines the energy-related flexibilities that these parties grant each other.
- To enable power adaptation collaboration between DC and ITC, we suggest new Green Service Level agreement (GreenSLA), which reflects the agreed scope for the DC to operate in an energy-aware manner while guaranteeing a certain level of quality of services (QoS) for the ITC.
- We promote DCs as excellent candidates to participate in an intelligent energy management eco-system as they can (1) store energy by increasing their power demand (e.g. cooling down the DC, de-consolidating the workload) during power surplus periods, and (2) decrease their power consumption (e.g. heating up the DC till allowed set-points, consolidating and shifting workload) during power shortage. The experiments conducted on a real private cloud computing DC confirm our approach.

It is worth pointing out that the proposed scheme of Fig. 1 is suitable for Distribution System Operators[1] (DSO) but is not only limited to. Other examples of potential end users of such a scheme are the Data Centre Operators (to reduce their energy costs) and IT clients in promoting new "green" services and hence increasing their competitiveness in the market.

The rest of this paper is organised as follows: Section III presents related work concerning GreenSDA and GreenSLA. Contractual terms between ES and DCs (GreenSDA) are given in Section IV. The power adaptation collaboration between DC and ITCs (GreenSLA) are covered in Section V. The configured environment, workload description as well as the obtained results of the experimental analysis are presented in Section VI. The paper is concluded in Section VII.

II. PRELIMINARIES

A. Algorithmic Perspective

In this section, we demonstrate that the algorithms implemented for the purpose of the proposed approach do not need to deal with scalability issues. More precisely, this approach, as stated in Section I, deals with medium to large-sized DCs.

[1]http://userwikis.fu-berlin.de/display/energywiki/distribution+system+operator
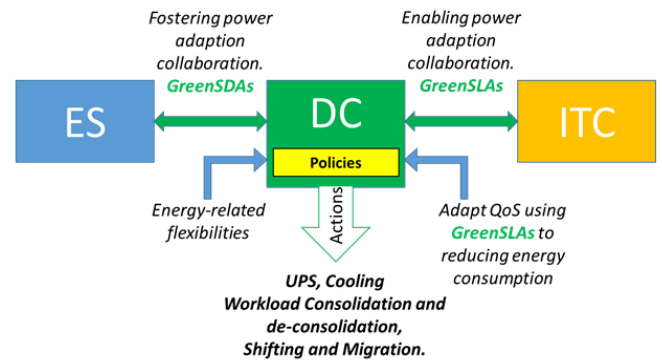

Fig. 1. Overview of the proposed scheme's different parties

Consequently, the number of such type of DCs in a given geographic location of an ES do not exceed more than 100 [7]. Furthermore, all the algorithms are based on the fact that during energy shortage situation, the ES sends a request to each participating DC separately and waits a reply from them within a pre-specified period of time. Consequently, two schemes were adopted: parallelism and timeout.

Therefore, based on the aforementioned schemes together with the considered number of DCs, we can confirm that all the implemented algorithms scale reasonably well.

Concerning the computational burden of the implemented algorithms, for the use case analysed in Section VI, typical execution time of the full control loop when the proposed framework was running inside a VM with 1 core and 2 GB RAM was less than 30 seconds, and peak processor load for that VM was less than 50%.

B. Data Centre Operations

In this section, we describe the DC's operational considerations for participating in DR. The following mechanisms were identified that provide flexibility for DCs in increasing/decreasing its power demand:

1) Workload consolidation

- Technique: Consists of packing as much workload as possible on a single IT equipment and shutting down idle IT equipment.
- Requirement: DC to be endowed with virtualisation technology.
- Impact: Reduced performance and power demand.
- Reaction time: Depending on the size of the workload. This can take in order of one minute.

2) Workload shifting

- Technique: Consists of rescheduling IT jobs to a time outside of the DR event window, followed by idling or shutting down the corresponding IT equipment.
- Requirement: The corresponding workload needs to be not sensitive to time (e.g. backup, anti-virus).
- Impact: Reduced power demand.
- Reaction time: Instantaneous.

3) Workload migration

- Technique: Refers to geographic shifting of workload to another DC from a DC that is participating in DR event.
- Requirement: The corresponding workload needs to be executable on the hosting DC. Both DCs must be equipped with virtulisation technologies.
- Impact: Reduced power demand but increased downtime.
- Reaction time: Depends on the workload's size, network's bandwidth as well as the mode (e.g. cold, warm, live) of migration. For instance live migration can have in optimal cases a downtime of only order of seconds.

4) HVAC temperature

- Technique: Consists of altering the temperature set-points of the IT infrastructure. Heating up during power shortage and cooling down during surplus.
- Requirement: DC to be equipped with different cooling technologies (e.g air- and liquid-based, free cooling). The different set-points need to be between $18^oC - 30^oC$ in compliance with the suggestions of American Society of Heating and Air-Conditioning Engineers (ASHRAE) for DCs.
- Impact: Reduced/increased power demand.
- Reaction time: Depends on the size of the DC as well as the number of cooling devices. This can take up to 5 minutes to cool down or heat up the DC.

5) Uninterrupted Power Supply (UPS)

- Technique: Consists of charging (surplus) or discharging (shortage) the battery of UPS.
- Requirement: UPS needs to be controllable from automation systems of DC.
- Impact: Reduced/increased power demand but quick aging of UPS's battery.
- Reaction time: Instantaneous depending on the status of the battery.

6) Dynamic Voltage and Frequency Scaling (DVFS)

- Technique: Refers to clocking down the frequency of the processor during low to medium workload utilisation periods.
- Requirement: Operating System (OS) of the corresponding IT equipment must support DVFS.
- Impact: Reduced performance and power demand.
- Reaction time: Instantaneous.

It is worth pointing out that mechanisms 4 and 5 can serve as energy buffering or storing during power surplus situations. Furthermore, such mechanisms can be used to reduce power consumption with no additional impact on the quality of service (QoS) of the IT operations. Also, the aforementioned mechanisms have very reasonable reaction time (order of minutes) suitable for most DR schemes.

III. RELATED WORK

The current paper is based on earlier work of some of the authors: In [6] they give an introduction on the presented contract based approach to DR with DCs in the context of the All4Green eco-system. Instead of going into more detail with regards to the GreenSDA and GreenSLA contract schemes, the former paper focused on the general architecture, working modes and communication within the All4Green eco-system. Also at that stage, the trial results were quite preliminary; in the current paper we present mature evaluation results. In [8] the presented approach is put into the context of a case study, using grid data of a German ES as well as the energy consumption and cost information of a small German DC. It could be shown that with only 6 small DCs all peaks in 2011 could have been avoided. The most efficient incentive scheme to do this would have been a combination of a comparably high monthly fixed payment per kW and a low variable reward per kWh.

The term demand side management was coined by C. Gellings and J. Chamberlin in the late 1980s [9]. It is the basis for the demand response concept introduced above and about flexibilising power demand as part of an integrated resource planning comprising both supply and demand. Demand side management is aimed at the power management of utilities influencing its load shape by changing its power usage pattern and/or magnitude. Demand response, on the other hand, deals with the interaction between a utility (or a different actor on the energy supply side) and a consumer, especially regarding the question of how to incentivise consumer power adaptation. Consequently, in [5] demand response was defined by the energy consumers' changes of demand side patterns as a reaction to dynamic prices or other incentives. An overview of demand response programs is given in [10]. There are basically three different kinds of motivation for demand response and demand side management: economic/market driven, environmental driven, and network (efficiency) driven [10]; the current work contains elements of all three.

GreenSDA is a novel concept proposed within the context of [11] in order to foster power adaptation collaboration between ES and DCs. To the best of our knowledge, there have been no efforts in proposing new supply demand agreements between ES and DCs for DR. Having said that, the feasibility of DCs to participate in DR programs was first analysed in [12], [13]. The authors conducted field tests to improve the understanding of the DR opportunities in DCs. The study evaluated an initial set of control and load migration strategies and economic feasibility. The findings show that with minimal or no impact to DC operations, a demand savings of 25% at the DC level or 10 – 12% at the whole building level can be achieved with strategies for cooling and IT equipment, and load migration. However, unlike this paper, no contractual terms for supply demand agreement between ES and DCs were proposed which enables power adaptation collaboration.

To the best of our knowledge, these works are the only elaborated empirical data that can be found in literature. Other papers either cite them [14] or provide rough estimations from DC managers: In [15] a workload flexibility between 5-20% is estimated, depending on the situation in individual DCs.

GreenSLA concept is being subject of recent attention by the academic community [6], [16], [17], [18], [19], [20]. In [17], authors propose a novel approach for the reduction of

energy consumption in data centres that will yield benefits both in terms of running costs and its environmental impact, where the method is based on the introduction of collaborative interactions and flexibility clauses in contracts between all the DC ecosystem entities. Considering these ideas, the concept is extended in this paper by suggesting a generalization based on automated demand response using in particular a system of contracts provided by the GreenSLA and the GreenSDA concepts to regulate the collaboration. Besides, a performance evaluation based on a proof of concept implementation is presented in this paper. Recent proposals [18], [19], [20] tackles the GreenSLA as a contract between the ITC and DC where several strategies are proposed inside the DC to reduce energy consumption and increase the use of renewable energy when available. The specific strategies used in the DC depends on the nature of its systems. In the testbed used for the experimental analysis in this work, the cloud computing DC is based on cloud nodes with virtual machines. However, these approaches neglect the ES and the ITC as active actors of the ecosystem. The GreenSLA approach proposed in this paper comes from the work developed in [11], [17], [21] by considering the interaction among the different players.

The research area of pricing demand response resources has been thoroughly studied for different market structures and for various DR schemes, basically differentiated into time-based rates (electricity price varies with time) and "incentive-based" DR (contractual, event-based reward) [22]. Especially the so-called real time pricing, – a spot price balancing supply and demand – has the economic advantage that it theoretically leads to a perfect market clearing at a perfect price. However, in reality, the electricity market is far from being perfect. This applies to DR with DC as to other DR actors: one major problem is e.g. to predict electricity prices in order to be able to react in time [23]. Also, the forecast of reactivity to price signals is highly error-prone and thus risky for the ES [24]. In reality, incentive-based programs are much more efficient [25] than time-based retail rates: [26] analysed that nearly all (about 93%) of the peak reduction realised in the U.S. is due to incentive-based programs. The challenge we are dealing with in this paper is to optimise the trade-off between flexibility and reliability of the incentive scheme. Pricing Service Level Agreements (SLAs) is often connected with the area of congestion-based pricing in networks estimating a traffic profile and attributing estimated cost based on this traffic ([27], [28]), thus calculating a price for future services. As an alternative usage-based pricing was introduced with the advent of cloud computing [29], [30]. To our knowledge, there exists only one paper that deals with pricing or rewarding GreenSLA [16]. The authors differentiate between prices for regular IT services and GreenSLA.

Hence, the major differentiator of our approach with respect to the existing ones is the *leveraging of power adaptation collaboration between ES – DC – ITCs.* Consequently, our novel concept of *Green Supply Demand Agreement* as well as *Green Service Level Agreement* foster this collaboration and make DR mechanisms viable for the case of DCs and their customers. Also, another major differentiator is that, unlike the existing proposed DR solutions, our approach takes into account adaptation collaboration to increase power consumption. This condition is suitable for power surplus situations in the case of renewable energy sources. Next, we introduce our novel GreenSDA and GreenSLA concepts.

IV. GREENSDA

Currently, in Europe, there is hardly any direct feedback between the power grid provider experiencing pressure on the power grid and the grid users imposing this pressure: either through an uncontrollable, volatile power supply or through unexpected demand spikes from big energy consumers like a data centre. There are indirect feedback mechanisms for some consumers via dynamic energy tariffs and also for selected very big consumers via the direct participation in the wholesale energy market (e.g. EEX).

From the point of view of the grid provider, however, both concepts lack the needed reliability. Faced with the challenge of controlling the stability of the power grid, capacity markets are slowly being created and aggregators are emerging on the European power landscape. Most solutions that are currently in place, however, are not suitable for some energy consumers like data centres as they focus on granting the aggregator or utility an intervention right into specific production processes at an industrial site. In order to remove this stumbling block on the road towards an optimised power grid, we propose the so-called GreenSDA.

A GreenSDA is an agreement that replaces current energy tariffs and contracts and determines the conditions under which the ES may request power adaptation actions from the DC. The concept is generic and can be applied to any industry; for the current use case we created a domain-specific instantiation for the data centre industry. It accounts for specific characteristics in data centres which are highly flexible and reactive. The origin of the GreenSDA concept lies in technical contracts that specify guaranteed quality of a service like response times or reliability. The GreenSDA contract package differentiates between (technical) contractual terms and incentives (the reward and penalty scheme).

A. Contractual Terms

The major inconvenience of today's energy tariffs is their lack of flexibility. In other words, the ES and its energy consumers (in this case DCs) sign a contract specifying a fixed basic fee, a power charge varying with the highest power required in the billing period and a (flat rate or day/night based) electricity price for each consumed kWh, on a yearly basis regardless of the power grid's operational state. Such an inflexibility at contract level implies total elasticity of power supply. Furthermore, the current contracts do not take into account the fact that consumers should increase their power consumption, when there is a power surplus generated by the renewable energy sources. As a remedy for this imbalance, we suggest a Green Supply Demand Agreement between ES and DC.

In order to extract a meaningful set of contractual terms, we analysed existing energy tariffs as well as current DR schemes

together with representatives from an ES[2]. These terms are generic and not limited to the current use case:

Contractual Term 1. *minIncrease*

Sub Term 1.1. *minDuration*

Sub Term 1.2. *maxDuration*

Contractual Term 2. *maxIncrease*

Contractual Term 4. *maxDecrease*

Sub Term 4.1. *minDuration*

Sub Term 4.2. *maxDuration*

Contractual Term 5. *requestPeriod*

Contractual Term 6. *maxAdaptationTime*

Contractual Term 7. *maxReactionTime*

Sub Term 2.1. *minDuration*

Sub Term 2.2. *maxDuration*

Contractual Term 3. *minDecrease*

Sub Term 3.1. *minDuration*

Sub Term 3.2. *maxDuration*

Contractual Term 8. *maxRejectsPerMonth*

Contractual Term 9. *maxRejectsInSuccession*

Contractual Term 10. *maxRequestsPerMonth*

Contractual Term 11. *requestInterval*

Contractual Term 12. *minEnactionTime*

The Contractual Terms 1 – 4 correspond to power adaptation collaboration with respect to minimum and maximum power reduction/increase expressed in kW. Each of the above mentioned terms has two sub-terms that specify the minimum and maximum duration (in minutes) of the corresponding power adaptation. The Contractual Term 5 denotes the period when ES can send a power adaptation request to DC. Possible values can be noon, morning, night, etc. The Contractual Term 6 indicates the maximum time (in minutes) the DC needs to adapt its power consumption. The Contractual Term 7 represents the maximum time (in minutes) DC needs to reply back to an ES's power adaptation request. Typically, such a term should have a value of at most 5 minutes. The Contractual Term 8 grants DCs the right to reject ES's power adaptation requests on a monthly basis bounded by an upper limit. The Contractual Term 9 goes one step further and defines the maximum number of successive rejections allowed by DCs. In order not to overwhelm DCs with ES's power adaptation requests, the Contractual Term 10 specifies the maximum number of requests ES can send to DC on a monthly basis. Since recovering from a power adaptation needs to be performed in a reasonable amount of time, the Contractual Term 11 defines the minimum period of time (in minutes) between two successive ES's power adaptation requests to DC. Hence, we ensure that DC has enough time to recover from its previous power adaptation request and is not forced to reject a potential ES request thus exceeding the maximum agreed number of rejections. The Contractual Term 12 guarantees that the DC has enough time (in minutes) to receive a notice from ES regarding a potential power adaptation request. Note that every time one of the parties breaches one or more contractual terms of GreenSDA, a penalty may be applied. Also, the agreed contractual terms are connected to a financial reward whose details are out of the scope of this paper.

The Contractual Terms 8–11 need to be monitored in order to keep track of their actual execution in DCs. To this end, we propose the following *monitoring parameters*:

- currentRejectsPerMonth
- currentRejectsInSuccession
- currentRequestsPerMonth
- lastAdaptationStartTime

The first three parameters keep track of the actual number of rejections of DC to ES's requests, the actual number of successive rejections, and the actual number of power adaptation requests sent by ES to DC respectively. The fourth parameter denotes the last power adaptation request's start time. It is worth pointing out that at the beginning of every month, those monitored parameters (except for currentRejectsInSuccession) need to be reset to zero. Readers interested in analysis of GreenSDA under different power profiling can refer to [31].

B. Reward and Penalty Schemes

As pointed out in Section III, there are two main categories of reward schemes in DR programs: Time- and incentive-based schemes. The advantages of one are the disadvantages of the other. For instance, increased flexibility for the customers of dynamic time-based rates like real time pricing goes along with decreased reliability of the DR potential [24]. Through the contractual ties of incentive-based rates, these are more reliable for both partners, but less flexible from the point of view of the DC, the exact extent of flexibility depending on the specific GreenSDA agreement. In this paper, we combine the advantages of aforementioned two approaches by choosing an incentive-based reward scheme with critical peak pricing elements. One reason for this are the positive experiences regarding reliability and effectiveness of incentive-based schemes in the U.S. [25], [26]. Also, incentive-based schemes are more in-line with event-based DR approaches (as the one presented here) than dynamic pricing. And lastly, the differentiation into various pricing elements reflects the contract structure of GreenSDA (see Section IV-A) taking into account both complexity and flexibility of technical and business processes in a DC. Consequently, we propose the following two elements: a fixed reward and a variable reward[3].

1) Fixed Reward: Is granted for the general inclination of the DC to enter into a collaborative relationship with the ES as stated in the contractual terms 1–12. It rewards basic behaviour changes and it is higher the more the DC commits.

Two approaches were analysed: one is based on consistently representing each fine-grained decision of the customer in a fine-grained reward adaptation but it turned out to be too complex to understand. Therefore, we chose a different approach familiar to most users: For each contractual term i, a limited number of categories is created (e.g. bronze, silver, gold). Assigning points $P_i \in [3, 6]$ for each category (e.g. each gold is worth 6 points, each silver 4.5 points, and each bronze 3 points) the values of all contractual terms i in each category can be easily aggregated. For the gold, silver, bronze example, in the presented case (12 terms, max 6 points each) this approach can be reflected in the following formula:

---

[2]SWP: Stadt Werke Passau (http://www.stadtwerke-passau.de/)

[3]The presented approach has been partly tested in [8]

$$Reward_{Fixed} = Basic_{Fee} * \frac{36}{\sum_{i=1}^{12} P_i}, \qquad (1)$$

where $Basic_{Fee}$ denotes the power charge of the regular energy tariff . Thus a DC that commits to the "gold standard" in all contractual terms, the power charge is reduced to half.

2) Variable Reward: The variable reward (or penalty) is granted at the end of a billing period for responding to specific ES requests. It expresses the realised grid support of the DC. Therefore, this support needs to be monitored in terms of MWh shifted and of the collaboration effort $Coll_{DC}$, i.e. the weighted positive reactions to requests from the ES:

$$Coll_{DC} = \frac{MWh_{AdapDC}}{MWh_{ReqES}} * \frac{T - currentRejectsPerMonth}{T}, \qquad (2)$$

with $Coll_{DC} \in [0,1]$ (1 indicating the highest collaboration). MWh_{AdapDC} and MWh_{ReqES} represent the power adaptation realised by the DC and requested by the ES respectively. $T > 0$ denotes the total number of requests sent by ES to DC, whereas $currentRejectsPerMonth$ is the number of rejects, both in the current billing period. The penalty scheme can be constructed in the same way with a penalty depending on $Coll_{DC}$ for $Coll_{DC} < a$ (a threshold specified in the GreenSDA). The formula for the variable reward is:

$$Reward_{Variable} = MWh_{AdapDC} * Coll_{DC} * Price_{MWh}, \qquad (3)$$

V. GREENSLA

This section presents and describes the terms necessary to leverage service adaptation and collaboration between DC and ITC, in a form similar to the GreenSDA. To this end, the GreenSLA is introduced: an agreement between the DC and the ITC that determines the DC's scope to operate the DC in an energy-aware way, thereby anyhow guaranteeing a flexible level of quality of service (QoS) [21], [11], [17]. Actually, the GreenSLA is an agreement between the customer and the provider of the IT service, considered under the shape of a DC, to take decisions regarding energy consumption that may affect the quality of the service provided. So, virtualisation and specifically virtual machines are one of the strategies where this concept can take place, but not limited to. Any service provider can exploit this concept regarding the type of systems used to provide services. First, the theoretical concepts underlying the GreenSLA are presented. Next, the contractual terms concerning GreenSLA are described and commented. The section concludes by proposing reward and penalty schemes based on the considered contractual terms (QoS).

A. GreenSLA Strategies

A GreenSLA, an SLA between DC and ITC, offers an extended scope of eco-optimisation to the service provider by: 1) relaxing traditional performance parameters (flexibility), 2) introducing novel energy performance parameters as classifying elements (Green Key Performance Indicator – GreenKPI), 3) introducing parameters that determine the level of collaboration and 4) offering incentives to the customers in the form of dynamic pricing schemes. ITCs accept an GreenSLA in order to get cheaper prices (and bonus) and eco-sustainability. Therefore, the guarantees to fulfill are clear for the ITC and DC in advance.

1) Flexibility: The concept of flexibility refers to the variability that the ITC and the DC are willing to accept in each of the service's running conditions based on a *context* dependent change. Context is the situation that creates opportunities to modify the service conditions in order to promote a more environmentally friendly behaviour of the eco-system.

In this work, three contexts are defined: time/calendar-related context, DC energy mode dependent context and load-based context. The former is related *to a time period*: hour(s), days of the week, among others. The second one refers to *a mode of operation of the DC*. By default the DC operates in regular mode. When receiving a request from its ES to reduce power consumption (e.g. to avoid tramping up diesel generators to cope with a peak in demand) or increase (energy surplus resulting from an unforeseen excess of renewable energy or due to a sudden decrease in demand from other customers), the DC may switch during a certain period to saving mode or to extra mode for a certain period [17]. In the saving mode, the DC could attempt to save energy by downgrading IT services and/or undertaking actions such as migrating or consolidating VMs, shifting tasks in time, etc. In the extra mode, the DC could promote IT services to higher performance levels, or anticipate the execution of maintenance tasks it is committed to. The *load-based* context considers changes in the configuration (switching off nodes, etc) of the system depending on the required real-time load.

2) GreenKPIs : Key Performance Indicators (KPI) evaluate the efficiency and effectiveness of a set of specified items, by means of collecting, analysing and reporting information on these items. The term GreenKPI refers to KPIs concerning a GreenSLA, so targeting energy savings and carbon emission reduction. Two main GreenKPIs are defined: The *Guaranteed Energy Consumption* refers to the guaranteed boundary in the kWh incurred by the execution of a service. This KPI is not applicable for all IT services, but it can be used for specific services such as the ones deployed only on a set of selected physical servers (i.e. HW power capping). The *Guaranteed Energy Mix* refers to the ratio of types of energy sources mixed for the energy provision and the values admitted for each service to run. The IT service is run on a DC that receives energy from a provider that has a guaranteed energy mix that provides a boundary on the relative CO_2 emissions generated by the execution of that service. The energy mix is composed of renewable and non-renewable sources. To reduce the CO_2 emissions, the guaranteed clean energy GreenKPI is defined in order to tie collaboration to the CO_2 emissions saved through this collaboration. This will steer the users' behaviour so as to consume energy if possible only when 100 % renewable energy sources are available. Consequently, services can be labelled according to their capacity being run with non-renewable energy sources: Non-adaptive 100 % renewable (the service must be run only with renewable energies) and Adaptive 100 % renewable.

Note that in this paper, how the energy mix is varying is out of scope, since this issue concerns to the energy supplier, who really knows how the mixing is made. What is important regarding our work is the need to keep updated the Guaranteed Energy Mix KPI and whether the energy supplier is able

to report the values. In the framework presented, this is an open issue out of scope. The proposal assumes that the real value of the energy mixing is known at the time of taking decisions. Since this value is used for the KPI, the system can decide appropriately if a service is not able to run under some conditions regarding this KPI.

3) Collaboration: Collaboration refers to the ITCs reactions to requests from the DC and vice versa. These requests are triggered by the change of DC mode or from decisions taken at the ITC: From regular situations, the DC might switch to states where it needs to reduce its current energy consumption or, on the contrary, switch to a state in order to consume extra available energy. This change of DC status comes from the interaction of the actors within the DC–ES sub-ecosystem, based on the GreenSDA terms. Each of these DC to ITC requests asks for changes in the performance/availability/execution/maintenance parameters of an IT service. This means that collaboration can occur between each pair of interacting entities. As specified in a GreenSDA, the ES may have permission to request the DC to decrease its power consumption during a certain period of time, or it may request from the DC to increase its power usage. The DC may have some freedom to accept or deny the request, but be obliged to handle it at least a certain number of times per month. In turn, as stated in a GreenSLA, the DC may be able to request an ITC's permission to downgrade a certain parameter of an IT service such as performance or availability; the ITC also having the freedom to deny the request but the obligation to honour it a certain number of times.

B. Contractual Terms

Each GreenSLA is tightly bounded to the IT service it is applied to. Consequently, contrary to GreenSDAs, contractual terms in GreenSLAs cannot be generalised. But, to give the reader an idea of how a GreenSLA might look like, we present the following example of an IaaS cloud service.

1) Service Description: The service being defined is the deployment of a virtual machine (VM). Thus, the DC rents VMs to its customers during a specific amount of time. The portfolio of the DC is described by VM attributes as follows:

VM *sizes*
a) **Small**: 1 VCpu, 1 GB RAM, 10 GB HD, b) **Medium**: 2 VCpu, 2 GB RAM, 20 GB HD c) **Large**: 4 VCpu, 8 GB Ram, 50 GB HD

VMs OS *images*
a) Ubuntu-64bit-v11, b) Windows2008-64bit, c) RedHatEnterprise-v6.2

Performance limits in ECUs[4]:
a) High Performance (**highPerf**) → 100%, b) Medium Performance (**midPerf**) → 85%, c) Low Performance (**lowPerf**) → 70%

Delay in the startup of a VM (in secs)
a) Regular Startup (**RegStartup**) → 30, b) Slow Startup (**SlowStartup**) → 120

---

[4]Amazon EC2 Compute Unit (ECU) provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor (http://aws.amazon.com/de/s3-sla/)

2) GreenSLA Additions: As already stated, a GreenSLA is no more than a regular SLA with three main additions: *Flexibility*, *GreenKPIs* and *Collaboration*. In the present example, these additions could be realised in the following way:

a) GreenKPIs: A GreenKPI on **emissions** (retrieved through the ES of the DC, as average) guaranteeing CO_2/service $<$ X gCO_2/kWh, meaning that all the services of this portfolio should not exceed X gCO_2/kWh on average. If this guarantee is breached, the DC must pay a high economic penalty per breach.

b) Flexibility: The **Time/Calendar** dependent flexibility can be applied in the following items.

- On week days between 9am and 9pm the performance should be equal or higher than *highPerf* (if the DC is in regular mode). If this guarantee is breached a medium economic penalty per breach must be paid by the DC.
- On week days before 9am and after 9pm and during weekends the performance should be equal or higher than midPerf (if the DC is in regular mode). If this guarantee is breached a medium economic penalty per breach must be paid by the DC.

Besides Time/Calendar dependent flexibility, the DC mode dependent flexibility can be applied in the following items. DC mode means that triggered by an event the whole DC operation is modified for a short period of time in order to adapt to a critical situation in the power grid:

DC in regular mode
- *Boot Delay*: In regular mode, the boot delay should not exceed *RegStartup* time in minutes. If this guarantee cannot be kept, a medium economic penalty per breach is applied to the DC.

DC in energy saving mode
Energy saving mode here does not include regular energy saving strategies that every DC is running in order to reduce its energy bill; it is rather a critical operation mode where the DC makes extra efforts (in some cases supported by its customers) in order to bring the power consumption of the DC down for a certain period of time.

- *Boot Delay*: In this mode, the boot delay should not exceed SlowStartup time in minutes. If this guarantee is breached, a medium economic penalty per breach is due.
- *Flexible VM performance*: When the DC is in energy saving mode, the performance should not be lower than *LowPerf*. If this guarantee is not adhered to, a medium economic penalty per breach is applied to the DC.

c) Collaboration constraints:
- A collaboration request might be to halt the virtual machine for a maximum time of two hours.
- The DC can ask a maximum of 10 collaboration requests per month to the ITC.
- The ITC can reject a maximum of 5 collaboration requests per month coming from the DC.
- The maximum time the DC can request a collaboration from the ITC (caused by a mode change to Energy Saving) is 2 hours.

- The maximum time the DC can request a collaboration from the ITC (due to a mode change to Energy Extra) is 2 hours.

Some of the collaboration contractual terms are directly inherited from the terms of the GreenSDA (see Section IV-A).

To sum up, the GreenSLA clauses may have the following effects on the cloud service: The time/calendar based flexibility guarantees high performance during working hours and mid performance otherwise. If the status of the DC changes to energy saving, the service can even be degraded to low performance. The boot delay should not exceed RegStartup minutes in regular mode, in energy saving mode it must not exceed LowStartup (so the tolerance in the delay is higher in energy saving mode). Finally, the collaboration clauses allow to pause the VM in energy saving mode, if the DC explicitly asks for it and the ITC agrees.

C. Pricing GreenSLA

Total cost of ownership (TCO) in a DC consists of capital (Capex) and operational (Opex) expenses. The former refers to investments (e.g. construction costs of DC, purchase of IT equipment, etc.) made that will be later subject to depreciation. The latter denotes to the recurring monthly expenses (e.g. electricity costs, repairs and maintenance, salaries, etc.) of running the DC services. It was shown in [32] that the dollars per watt cost of server hardware is trending down, whereas electricity and construction costs are increasing. Hence, over long term, the authors believe that the DC facility costs which are proportional to power consumption will become the prominent fraction of total costs.

Generally, the price of a good is dependent on its cost and on the market position of the producer. As mentioned above, a huge chunk of the cost for DC services is energy cost as one of the largest factors of Opex. Through a GreenSLA this formerly constant cost becomes variable indirectly it depends on the status of the power grid. This enables the DC to optimise its energy bill by additionally offering GreenSLAs to its customers. The GreenSLA price and reward offerings should reflect the structure of energy cost variations, support the incentive function of pricing to use resources according to their scarcity and be simple to understand as already mentioned in Section IV. Therefore, the following structure, mirroring the structure of IT service pricing, is suggested: A fixed reward that is reliable for both DC and ES: It should be calculated in a way that the aggregated sum of all fixed rewards granted does not exceed the fixed reward the DC receives from the ES. This part of the reward is risk-free from the DCs view. And a variable reward that reduces the variable service price taking into account monitored QoS modifications perceived by the DC costumer and specific GreenKPI.

VI. EXPERIMENTAL ANALYSIS

In this section, we present the results of the experimental analysis, for the implementation of the presented eco-system, conducted at Hewlett-Packard Italy Innovation Center (HPIC) cloud computing lab.

TABLE I
TESTBED'S HARDWARE CONFIGURATION

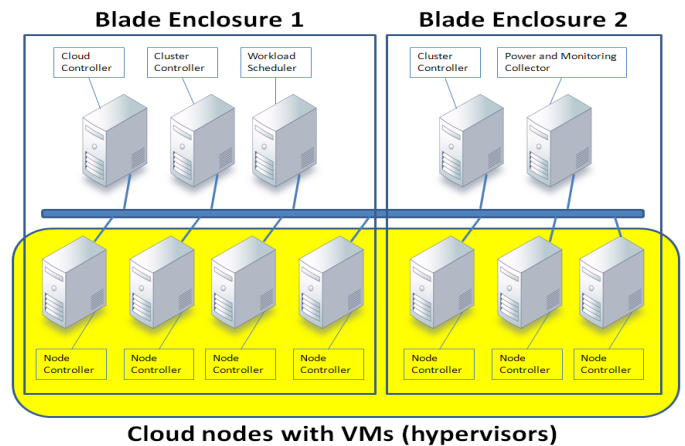|	
Processor	Quad-core Intel Xeon E5520, 2.26 GHz 8 MB L3 cache
Memory	24 GB (6 x 4GB DIMMs) DDR$_3$
Hard Disk	Two hot plug hard drives 2 x 300 GB
Network	Dual-port 10 gigabit Ethernet adapter NC532m
Power Rating	120 W (idle) and 235 W (full load)

Fig. 2.   Testbed's software configuration

A. Testbed Environment and Configuration

The realised cloud computing DC testbed is at HP Italy premises in Milan. It provides a computational environment implementing a cloud computing platform for Infrastructure and Platform as a Service (IaaS and PaaS) layers.

1) Hardware Configuration: The hardware equipment consist of two racks each hosting an *HP Blade System C3000*[5] enclosure where equivalent ISS (Industry Standard Server) blade servers are mounted inside. Enclosures 1 and 2 bear seven and five blade servers respectively belonging to *HP ProLiant BL460c G6*[6] series configured as in Table 1. Energy measurement is performed by an HP hardware component named *iLO*[7] (Integrated Lights-Out), accessible through the *Insight Control* software suite. *iLO* can read real-time electrical power consumption down to single server level every minute. For each 1-minute interval, the peak and average power usage is stored in a circular buffer. This measurement set-up is assumed to have an uncertainty of 1%.

2) Software Configuration: Fig. 2 illustrates the software configuration of the testbed. The *Cloud Controller* software, which is deployed on a physical server, in this case plays the role of existing DC automation framework. The cloud management system is based on OpenStack[8] platform.

The *Node Controller* software runs on physical machines providing virtualised environment to cloud platform clients. They are the physical servers on which VMs are created and instantiated by the Cloud Controller. The *Power and Moni-*

---

[5]http://h18000.www1.hp.com/products/blades/components/enclosures/c-class/c3000/?jumpid=reg_R1002_USEN
[6]http://h10010.www1.hp.com/wwpc/us/en/sm/WF25a/3709945-3709945-3328410-241641-3328419-3884098.html
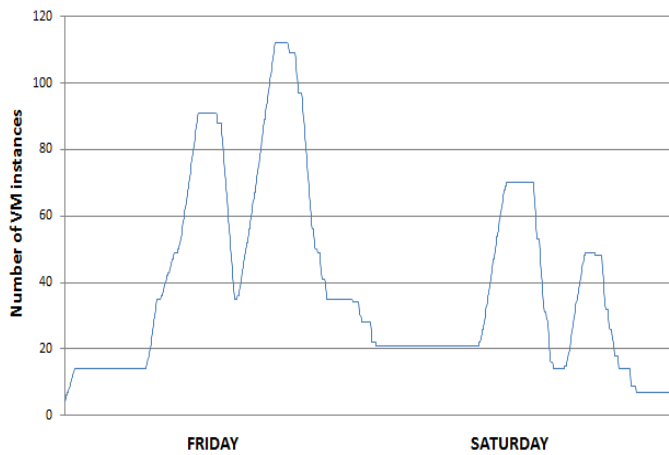[7]http://h18013.www1.hp.com/products/servers/management/ilo_table.html?jumpid=reg_R1002_USEN
[8]http://www.openstack.org

Fig. 3.   Workload shape

toring Collector software is implemented by a customisation of *collectd*. *Collectd* is an open source Linux daemon able to collect, transfer and store performance data of computers and network equipment. For this testbed, specific *collectd* plug-ins have been implemented, to interface with *iLO* and acquire power measurement data. The Power and Monitoring Collector is deployed on a physical server. The *Workload Scheduler* software is an energy-aware optimisation decision maker similar to the one proposed in [21] and is deployed on a dedicated physical server. Finally, the client systems are emulated by a custom software tool (see Section VI-B). The client load simulation is deployed inside virtual machines running an Ubuntu image, executed by Workload Scheduler.

B. Workload Description

The *HPIC* is actively involved in cloud computing pilot projects with real customers. Recently, a private IaaS was run at one of these customers' premises[9]. On top of the IaaS infrastructure, the *HPIC* instrumented a custom auditing system, that captured and stored a detailed log of the transactions occurring among the end users and the IaaS, together with the monitoring of all system parameters[10].

After the end of Proof of Concept (PoC), the log files were carefully analysed in order to extract a perceptible track of the actual user activities. As final outcome, we got crisp and content-relevant activity profiles of 7 different usage patterns. The chosen profiles span a sufficient timeframe and content to get a significant variance of activity profiles, and a sufficient amount of dynamical context changes. These profiles were the basis for designing and implementing the workload patterns used in this analysis. Finally, in order to emulate the workload pattern, the *HPIC* designed and developed a custom workload generator. This tool can generate a sequence of actions and direct them to the Cloud Controller. The tool is using an open source scheduler (jobscheduler[11]), which allows running all the tests at any wanted time and with specified time sequencing.

Fig. 3 depicts the total number of simultaneous instances of Virtual Machines (VM) running during the simulation of a typical Friday and Saturday. In the workload, instances have different sizes (e.g. Small, Medium, Large, etc.), corresponding to different virtual resources, in line with the statistical distribution of the PoC usage. Moreover, the processor load inside each VM is modelled using different patterns (constant low/medium/high, fluctuations with predefined period and slopes), to make the overall resource consumption as similar as possible to the originally recorded patterns.

C. Configured GreenSLA Parameters

In this testbed, there is a single contract template for all the VMs. Therefore the same contractual terms are applied to all clients, and thus all VMs are treated equally from the QoS. The specific GreenSLAs in place in the test state that:

- During week-days between 9:00–20:00, the VM-guaranteed performance is **HIGH** (100%[12]).
- During week-days between 0:00–9:00 and 20:00–24:00, the performance is **MEDIUM** (80%).
- During week-ends, the performance is **MEDIUM**.
- If the DC is in *Energy Saving* mode (for a max of 2 hours, max once per day), the startup of new virtual machines can be delayed up to 2 hours, and the active ones get **LOW** performance (50%).

D. Obtained Results

Before presenting the results, it is worthwhile to mention the following facts for this testbed:

- The temperature set-points were not altered and kept within the respected operational range of ASHRAE.
- On-going customer operations were not deferred or rescheduled and were executed based on the signed GreenSLA. This was validated through SLA manager which reported no SLA violations, thus reflecting customers' satisfaction for QoS.
- During power shortage situation and upon the request of ES, new customer operations were either rejected or delayed.
- Experimental data were gathered in 1-minute intervals at the server level. However, for clarity, Figures 3, 4 and 6 are presented in 3-hour intervals.
- The time resolution of the GreenSLA and GreenSDA is 1-minute intervals. Note that such a resolution should not be more than 5 minutes.
- The main goals of the All4Green project are both the usage of GreenSLA to foster energy saving, as well as to take advantage of the collaboration between the different actors in the system (ES, DC and ITC) to help the ES in situations of a power shortage or surplus. In this paper we focus on the second one, using the results related to the first as a baseline to show the additional improvement due to collaboration.

---

[9] The name of this customer is under Non Disclosure Agreement
[10] Both at single server and virtual machine level
[11] http://sourceforge.net/projects/jobscheduler/

[12] As percentage of ECU (Amazon WS EC2 Compute Unit) where 1 ECU is the equivalent CPU capacity of a 1.0–1.2 GHz 2007 Opteron or 2007 Xeon processor.
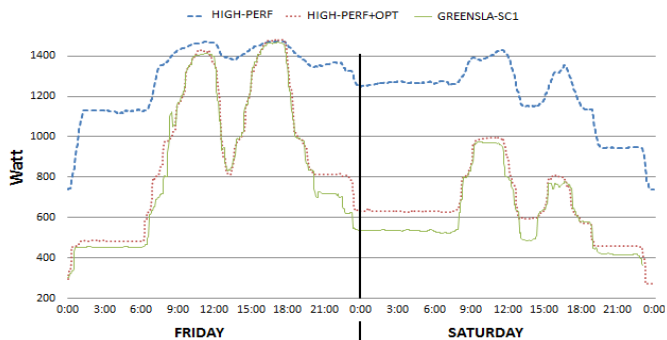
Fig. 4.   Power consumption (Watt) of the tested scenario

Fig. 5.   Energy consumption (Wh) of the tested scenario

1) Scenario 1: The first set of results is related to a scenario of the typical Friday-Saturday load patterns. It is applied first to the DC without any energy optimiser and no GreenSLAs (square doted line: HIGH–PERF), then with the energy optimiser configured without GreenSLAs (round doted line: HIGH–PERF+OPT) and finally with the energy optimiser and activated GreenSLAs (line GREENSLA–SC1). It is worth noting that the energy optimiser's main objectives are (1) to consolidate workload (e.g. in this case VMs) on minimal number of most *energy-efficient* servers, (2) turn off unutilised servers, and (3) turn them back on when the workload grows. To achieve this, the optimiser is equipped with power demand estimator that predicts the power consumption of the whole DC given a specific hardware and workload configuration. It is hence then up to the optimiser to find the optimal configuration to reduce the overall power demand. Further details can be found in [21].

In this first scenario, the DC is supposed to stay in regular mode for the whole time period. The major objective of this scenario is to demonstrate the potential of GreenSLA with respect to the normal SLA without shifting any workload. This means that the achieved power reduction results in an energy reduction.

The experimental results obtained from this scenario assess our theoretical observation. Fig. 4 shows the average of the results collected during several 2-day runs of the test. The X-axis represents the time of the day starting from Friday at 0:00 and ending up on Saturday at 24:00, the Y-axis the power (in Watt) consumed by the 7 hypervisor nodes (the ones in the bottom of Fig. 2).

Fig. 5 summarises the results of the tests in terms of energy consumed during the 2-day run. The results show savings introduced by the optimiser with respect to the original configuration of approximately 38%, while the GreenSLA setting further saves 5.5% with respect to the optimised version (with standard SLAs) due to the extra flexibility related to a time period. The difference is already visually evident in the line graph of Fig. 4, where the green line is lower during nights and week-ends.

2) Scenario 2 and 3: In the next two scenarios, the main potential of the presented approach is evaluated in the case where at approximately 9:30am the ES requests the DC to run in *Energy Saving* mode for about 2 hours. Here, the power is shifted, not shed, as the workload is partially delayed, not just
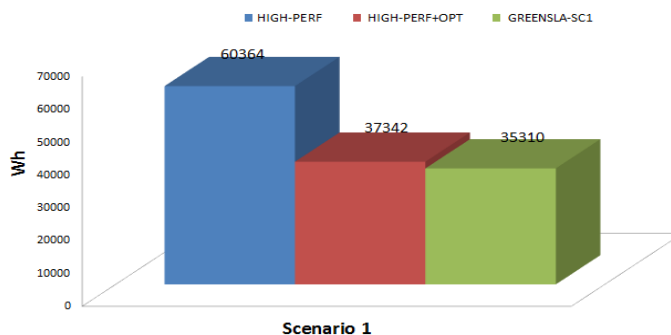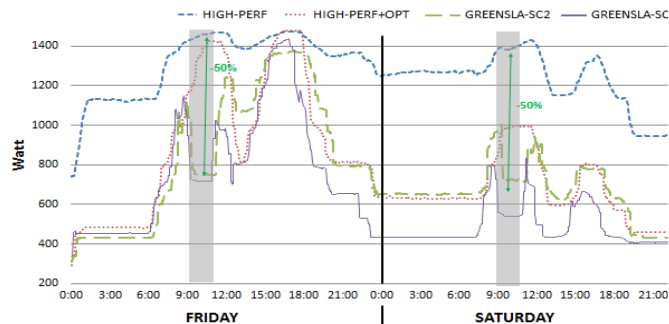

Fig. 6.   Power consumption (Watt) of the tested scenarios

reduced. Fig. 6 illustrates the average of the results collected during several 2-day runs of the test. The X- and Y-axis denote the same as in Fig. 4. The gray vertical bars represent the time when the DC is in *Energy Saving* mode. In these scenarios, HIGH-PERF and HIGH–PERF+OPT represent the same policy as in Fig. 4. The third line shows the energy optimiser and GreenSLAs in action for a non-federated case (GREENSLA–SC2), while the fourth line is the federated[13] case (GREENSLA–SC3). From the obtained results, we note the following:

- The power demand reduction is around 50% (see the green arrow line in Fig. 6) during the *Energy Saving* mode, thanks to the flexible GreenSLAs vs rigid SLA.
- In the non-federated case (green long dashed line GREENSLA-SC2), after the end of the *Energy Saving* mode period, there is a secondary peak due to the restart of the delayed VMs in addition to the regular workload for that time period. In the federated case (purple line GREENSLA-SC3), this does not happen since the VMs have been allocated to a different (federated) DC.

Even though the focus of the work is on sharply reducing (or increasing) the DC power for a limited amount of time as reaction to an external trigger, the results also show reductions in the overall energy consumption of the described test runs: Fig. 7 shows the energy spent on the tested scenarios during the 48-hours run (in Wh). In this case, the additional savings introduced by the GreenSLAs are slightly lower (4.15%) because of the overhead of the secondary peak mentioned

---

[13]By federation, we mean migration of the workload (in this case VMs) to a more energy-efficient DC.
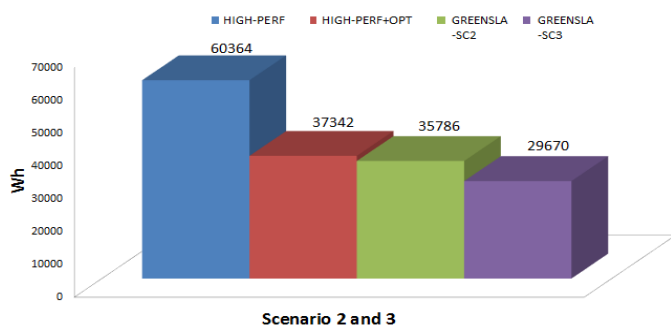
Fig. 7.   Energy consumption (Wh) of the tested scenarios

before: as the power is both shifted and shed, the reaction to the request from the ES is partially fulfilled by saving energy, partially by delaying workload. The federated case (last bar) shows a lower value, since it does not include the energy spent in the federated DC. Here, the amalgamation of shifting and shedding is overlaid by the fact that two DCs are affected.

VII. Conclusions

Demand Response (DR) mechanisms serve to preserve the stability of the power grid by shedding the electricity load of the consumers during power shortage situations in order to match power generation to demand. Data centres have been identified as excellent candidates to participate in such mechanisms. In this paper, we presented a proof-of-concept of a power adaptation collaboration between energy supplier (ES), data centres (DC) and IT customers (ITC). In order to foster this collaboration, we used two novel concepts of agreements: GreenSDA and GreenSLA. The former defines the contractual terms to be signed between the ES and DCs, whereas the latter is necessary to state the agreed terms between DCs and ITCs. The main goals are to demonstrate the feasibility as well as the potential of such a power collaboration from technical perspective, while the economical view has been narrowed to considering incentives based on reward and penalty schemes. To this end, we conducted experimental analysis on real cloud computing testbed by taking into account different types of service level agreements as well as DC energy modes. In this environment, proposed management schemes for the green contracts had been tested successfully, using GreenKPIs to evaluate the efficiency and effectiveness of the system in three scenarios. The obtained results show the high potential of DCs participating in power adaption schemes for demand response. Interestingly, as our results show, one can capitalize on the flexibility of data centers both for adapting power profiles as well as for energy savings.

Lessons learnt from this work are the following:

- In order to enable DR mechanisms, the participating parties need to provide some form of flexibilities. By doing so, the parties should be rewarded in order to create incentives in joining the DR mechanisms.
- The flexibilities should be mapped into agreements. Hence, the proposed GreenSDA and GreenSLA of this paper are the cornerstones of creating power adaptation

collaboration. Those agreements need to also be coupled with corresponding reward and penalty schemes in order to incentivise the different parties in joining a DR mechanism.

- Data centres are excellent candidates to participate in DR mechanisms due to their significant energy demand as well as their automated infrastructure providing different forms of flexibilities. Furthermore, those flexibilities have in general very quick reaction time (e.g. order of minutes).
- The flexibilities of DCs can be divided into two categories: internal and external.
  - Internal flexibilities are the ones that can be applied without affecting the operation of the clients. Those are (1) running the DC on Uninterrupted Power Supply (UPS), (2) heating up or cooling down the DC within the temperature set-points of 18°C and 30°C, and (3) migrating the workload to another DC which operates more efficiently or the ES it supplies energy does not suffer from shortage situation. For all those flexibilities, only GreenSDA is required.
  - External flexibilities are the ones that can be applied by affecting the operation of the clients. Those are (1) workload shedding or rescheduling, (2) workload shifting, (3) degrading the performance of IT infrastructure on which the clients' workload are running. All those flexibilities require the acknowledgement of the IT client and hence need GreenSLA concept.

For practical reasons, we have limited our test scenarios to smaller data centers as a proof of concept. But this is by no means a restriction of the applicability of our approach to larger scenarios of data centers of medium or large size, perhaps extending into MW or 10s of MW range. The scalability analysis and deployment procedure in big systems, however, is out of the scope of this paper and has been left for future work.

References

[1] S. Klingert, A. Berl, M. Beck, R. Serban, M. Di Girolamo, G. Giuliani, H. de Meer, and A. Salden, "Sustainable energy management in data centres through collaboration," in *1st Int'l workshop on Energy-Efficient Data Centres*, 2012, pp. 13–24.

[2] P. Graichen and C. Redl, "The german energiewende and its climate paradox," Agora Energiewende, Tech. Rep. 038/04-A-2014/EN, 2014.

[3] F. Katiraei and J. R. Agero, "Solar pv integration challenges," *Power and Energy Magazine, IEEE*, vol. 9, no. 3, pp. 62–71, 2011.

[4] "Openadr 2.0 profile specification b profile," OpenADR Alliance, Tech. Rep., 2013.

[5] "Assessment of demand response and advanced metering," Federal Energy Regulatory Commission, Tech. Rep., 2010.

[6] R. Basmadjian, G. Lovasz, M. Beck, H. D. Meer, X. Hesselbach-Serra, J. F. Botero, S. Klingert, M. P. Ortega, J. C. Lopez, A. Stam, R. van Krevelen, and M. D. Girolamo, "A generic architecture for demand response: The all4green approach," in *3rd Int'l Conf. on Cloud and Green Computing*, 2013, pp. 464–471.

[7] S. Klingert, "Final market analysis d2.4," DC4Cities: An environmentally sustainable data centre for Smart Cities, Tech. Rep., Nov 2015.

[8] A. Berl, S. Klingert, M. T. Beck, and H. D. Meer, "Integrating data centres into demand-response management: A local case study," in *39th Conf. of the IEEE Industrial Electronics Society*, 2013, pp. 4762–4767.

[9] C. Gellings and J. H. Chamberlin, *Demand-Side Management: Concepts and Methods*. PennWell Books, 1987.

[10] J. Aghaei and M.-I. Alizadeh, "Demand response in smart electricity grids equipped with renewable energy sources: A review," *Renewable and Sustainable Energy Reviews*, vol. 18, p. 6472, 2013.

[11] http://www.all4green-project.eu/.

[12] G. Ghatikar, M. A. Piette, S. Fujita, A. McKane, J. H. Dudley, and A. Radspieler, "Demand response and open automated demand response opportunities for data centers," Lawrence Berkeley National Laboratory, Tech. Rep., 2010.

[13] G. Ghatikar, V. Ganti, N. Matson, and M. A. Piette, "Demand response opportunities and enabling technologies for data centers: Findings from field studies," Lawrence Berkeley National Laboratory, Tech. Rep., 2012.

[14] A. Wierman, Z. Liu, I. Liu, and H. Mohsenian-Rad, "Opportunities and challenges for data center demand response," in *Green Computing Conference (IGCC), 2014 International*, Nov 2014, pp. 1–10.

[15] S. Klingert, T. Schulze, T. Moller, N. Wesemeyer, W. Duschl, F. Niedermeier, H. de Meer, F. Lombart Badal, F. Casaus, E. Martin, S. Vives Sanjoaquin, E. Pantoja Alonso, G. J. Diaz Velez, M. Sheikhalishahi, F. Wauters, M. Perez Ortega, and M. Chinnici, "D2.4 - final market analysis," EU project DC4Cities, Deliverable, 2015. [Online]. Available: http://www.dc4cities.eu/en/?smd_process_download=1&download_id=4642

[16] M. E. Haque, K. Le, I. Goiri, R. Bianchini, and T. D. Nguyen, "Providing green slas in high performance computing clouds," in *4th Int'l Conf. on Green Computing*, 2013.

[17] D. Rincón, A. Agustí-Torra, J. F. Botero, F. Raspall, D. Remondo, X. Hesselbach, M. T. Beck, H. de Meer, F. Niedermeier, and G. Giovanni, "A novel collaboration paradigm for reducing energy consumption and carbon dioxide emissions in data centres," *Computer Journal*, vol. 56, no. 12, 2013.

[18] A. Amokrane, R. Langar, M. F. Zhani, R. Boutaba, and G. Pujolle, "Greenslater: On satisfying green slas in distributed clouds," *IEEE Trans. on Network and Service Management*, vol. 12, no. 3, pp. 363–376, 2015.

[19] S. Hasan, Y. Kouki, T. Ledoux, and J. L. Pazat, "Exploiting renewable sources : when green sla becomes a possible reality in cloud computing," *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp. 1–14, 2015.

[20] L. Wang, S. U. Khan, D. Chen, J. Kolodziej, R. Ranjan, C. zhong Xu, and A. Zomaya, "Energy-aware parallel task scheduling in a cluster," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1661–1670, 2013.

[21] http://www.fit4green.eu/.

[22] "Assessment of demand response and advanced metering," FERC Staffreport, Tech. Rep., 2013.

[23] Z. Liu, I. Liu, S. Low, and A. Wierman, "Pricing data center demand response," *SIGMETRICS Perf. Eval. Rev.*, vol. 42, pp. 111–123, 2014.

[24] J. L. Mathieu, D. S. Callaway, and S. Kiliccote, "Examining uncertainty in demand response baseline models and variability in automated responses to dynamic pricing," in *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 2011, pp. 4332–4339.

[25] J.-H. Kim and A. Shcherbakova, "Common failures of demand response," *Elsevier Energy*, vol. 36, no. 2, pp. 873–880, 2011.

[26] P. Cappers, C. Goldman, and D. Kathan, "Demand response in u.s. electricity markets: Empirical evidence," Lawrence Berkeley National Laboratory LBNL-2124E, Tech. Rep., 2009.

[27] C. Courcoubetis and V. A. Siris, "Managing and pricing service level agreements for differentiated services," in *7th International Workshop on Quality of Service (IWQoS99)*. IEEE, 1999, pp. 165–173.

[28] C. Bouras and A. Sevasti, "Sla-based qos pricing in diffserv networks," *Computer Communications*, pp. 1868–1880, 2004.

[29] I. Menache, A. Ozdaglar, and N. Shimkin, "Socially optimal pricing of cloud computing resources," in *Proc. 5th Int'l ICST Conf. on Performance Evaluation Methodologies and Tools*, 2011, pp. 322–331.

[30] C. Weinhardt, A. Anandasivam, B. Blau, N. Borissov, T. Meinl, W. Michalk, and J. Stößer, "Cloud computing – a classification, business models, and research directions," *Business & Information Systems Engineering*, vol. 1, no. 5, pp. 391–399, 2009.

[31] R. Basmadjian, L. Müller, and H. D. Meer, "Data centres power profile selecting policies for demand response: Analysis of greensda," *Ad Hoc Networks: Energy-Aware Data Centers: Architecture, Infrastructure, and Communication*, vol. 25, no. B, pp. 581–594, 2014.

[32] L. A. Barroso, J. Clidaras, and U. Hölzle, *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines*, 2nd ed., M. D. Hill, Ed. Morgan & Claypool, 2013, chapter 6 page 96.

Robert Basmadjian holds a Ph.D. from University of Toulouse on data replication. After completing his doctorate, in 2009 he joined University of Passau and since then he has been as a postdoctoral fellow, where his main research interests are large-scale energy management systems (Smart Grid), and performance modelling of computing systems (queuing theory). He has more than 25 scientific publications in the respective fields. He was a scientific and technical contributor to EU FP7 FIT4Green and ALL4Green projects related to Demand Response in data centers. Moreover, he was an active member of WG 2 and 3 of COST ACTION 804, EURONF and EINS.

Juan Felipe Botero is an assistant professor of the University of Antioquia, Colombia. In 2006 he received his Computer Science Degree from the University of Antioquia, Colombia, and the M.Sc. degree in Telematics Engineering in 2008 from the telematics department of the Technical University of Catalonia, UPC, in Barcelona, Spain. In 2013 he received the Ph.D. degree in Telematics Engineering also from the telematics department of the Technical University of Catalonia, UPC. In 2013, he joined the applied telecommunications research group at the University of Antioquia. His main research topics are: Energy Efficiency, Network Virtualization, Software Defined Networking and Network Functions Virtualization.

Giovanni Giuliani graduated in Electronics Engineering with honors in 1982, he has over 30 years of working experience in RD, IT Consulting and System Integration in major Computer Companies (Olivetti, Digital, Compaq, HP - now HPE). Back in 2005 he joined the HP Italy Innovation Center, where hes been leading the Cloud Computing Initiative since 2009. Starting from 2013 he serves as Lead Architect of the HP LIFE platform. During the last decade, he has been involved in several research projects funded by European Commission in the area of Mobility, e-Collaboration, Cloud Computing and more recently he is technical coordinator in a sequence of 3 Green-IT research projects (FIT4Green, All4Green and DC4Cities).

Xavier Hesselbach associate Professor at the Department of Networks Engineering at the UPC, and IEEE member since 1992, received the M.S. degree with honors in Telecommunications Engineering in 1994 and the PhD. degree with honors in 1999, from the Universitat Politcnica de Catalunya (UPC). His research interests include networks virtualization, resources management, broadband networks, quality of service and green networking. He has been involved in several national and international projects, and is author in 4 books and more than 60 national and international publications in conferences and journals. In 1994 he received the award from the COIT/AEIT of Spain for the best Master Thesis on Networks and Telecommunication Services. He has taken part in several European and Spanish research projects, such as the EuroNGI/FGI/NF network of Excellence, COST293, Mantychore and All4Green.

Sonja Klingert has been with the University of Mannheim since January 2010 as a researcher and local manager for European research projects (STREP) like FIT4Green, All4Green, DC4Cities and currently ELECTRIFIC. Before joining University of Mannheim, she was coordinator of the European research project (STREP) GridEcon at the International University in Germany. Prior to that she worked as a researcher for the Wuppertal Institute for Climate, Environment and Energy, Germany. Her research interest is in demand response with data centres and corresponding metrics, in green business models, eco-aware contracts like GreenSLA and GreenSDA as well as EV grid integration. Among others, she has been co-chair of the international workshop series Energy Efficient Data Centres (E2DC) and was part of the organization team of the international conference ACM e-Energy. S. Klingert received her M.S. degree (Diplom) in Economics from the University of Karlsruhe in 1996.

Hermann De Meer received his PhD degree in Computer Science from the University of Erlangen-Nuremberg (Germany). He held postdoctoral research positions at Hamburg University (Germany), University of Texas at Austin (USA), Duke University (USA) and Columbia University (USA). After his Readership at University College London (UCL) (UK) he was appointed as Professor at the University of Passau (Germany) in 2003. His area of research comprises Computer Networking and Energy Systems. Special focus has been on Network Virtualization, IT-Security of the Smart Grid, Demand Side Management, E-Mobility, Industry Automation, Resilience and Risk Management of Distributed Systems.