Discussion Paper

Shifting Load through Space - the Economics of Spatial Demand Side Management Using Distributed Data Centers

by

Gilbert Fridgen, Robert Keller, Markus Thimmel, Lars Wederhake

October 2017

in: Energy Policy (2017), 109, p. 400-413

The final publication is available at: <https://doi.org/10.1016/j.enpol.2017.07.018>
Abstract
Demand-side flexibility (DSF) in the electricity grid has become an active research area in recent years. While temporal flexibility (e.g., load shedding, load shifting) is already discussed intensively in literature, spatial load migration still is a little researched type of DSF. Spatial load migration allows to instantly migrate power-consuming activities among different locations. Data centers (DCs) are power-intensive and process information goods. Since information goods are easily transferable through communication networks, the power-intensive processing of information goods is not necessarily tied to a specific location. Consequently, geographically distributed DCs inherit - in theory - a considerable potential to migrate load globally without using power lines. We analyze the economics of spatially migrating load to provide balancing power using geographically distributed DCs. We assure that neither of the participating electricity grids shall be burdened by this mechanism. By using historic data to evaluate our model, we find reasonable economic incentives to migrate positive as well as negative balancing power. In addition, we find that current scenarios favor the migration of negative balancing power. Our research thus reveals realistic opportunities to virtually transfer balancing power between different market areas worldwide.

Keywords:
Load migration, economic potential, balancing power, simulation-based case study, demand response, demand-side management
1. Motivation

In power markets, supply and demand must be so tightly coupled that they line up all times (Müller and Rammerstorfer, 2008; Rammerstorfer and Wagner, 2009). However, it is difficult to know a priori what the exact demand for power will be (Flinkerbusch and Heuterkes, 2010). Considering the ever-growing power generation from intermittent renewables, the supply side introduces even more uncertainty (Vandezande et al., 2010). The resulting high levels of volatility demand for elevated levels of flexibility (Ehrlich et al., 2015; Strbac, 2008). Flexibility is the potential to balance deviations from the scheduled power generation or demand caused by prediction errors (Eurelectric, 2014). The “intelligent control” (Buhl and Jetter 2009) of demand side resources by information and communication technologies (ICT) increases a power grid’s ability to react on higher levels of volatility (Fridgen et al., 2015; Strbac, 2008).

Unsurprisingly, examining the potential of demand-side flexibility (DSF) has become an active field of research in recent years, e.g. electric vehicles (Fridgen et al., 2014; Lujano-Rojas et al., 2012), heating and cooling-systems (Ehrlich et al., 2015; Goddard et al., 2014; Grein and Pehnt, 2011), and commercial and industrial processes (Jang et al., 2016). As these examples illustrate the vast majority of today's approaches to providing flexibility are variants of temporal flexibility (load shifting and load-shedding).

Another approach to providing flexibility is to spatially migrate load. Typically, this is done by import/exporting load through power lines. The exchange, in general, is favorable for both importers and exporters of power for two reasons: first, excess power in one location can neutralize a deficit in another, and second, some markets can provision flexibility more cost-efficiently and/or in a less carbon-emitting manner than others (Van Hulle et al., 2010; Vennemann et al., 2011). Constituents for provisioning flexibility efficiently are the mix of energy generation, efficient storage facilities, and the potential for adjusting load. The potential of the first two of these components, however, is heavily influenced by geographical realities (Vennemann et al., 2011). Therefore, it is imperative to interconnect markets to reap the benefits from reducing market inefficiencies. Nevertheless, many power line construction projects fail because of excessive initial costs (Kishore and Singal, 2014), insecure return on investment (Buijs et al., 2011), protests by local citizens (Lütticke, 2017), and high transmission costs (Vennemann et al., 2011).

An alternative concept to spatially migrate load without power lines is to instantly migrate power-consuming activities between different locations. Processing and delivering information goods are such a
power-consuming activity. This is so, because it is in theory perfectly location independent (Krcmar, 2015). Data centers (DCs) processing and delivering information goods are very power-intensive. As an example, all US-based DCs contribute some 2% to the country's total electricity consumption (Koomey, 2011). Accordingly, a setting consisting of geographically distributed DCs could enable the spatial migration of load over long distances relying on communication networks instead of the power grid.

Migrating load requires the participation of economic entities. The actions of economic entities are mainly driven by economic rationale (Simon, 1979). Therefore, economic entities will only adopt the migration of load, if it is economically feasible. Since the intentional in-/decrease of load is one possible source for flexibility, geographically distributed DCs can deliver flexibility in one location by migrating load from this location to a remote location. However, this might also result in unintentional in-/decreases of load in the remote location. Consequently, a major challenge of spatially migrating load is to avoid that one power grid improves grid stability at the expense of another. This may be caused by potentially introduced additional power imbalances.

Thus, the objective of this paper is to analyze the economic feasibility (cash flow from operating activities) of spatially migrating load as to provide flexibility, burdening neither of the participating power grids by potentially introduced additional power imbalances. We demonstrate the economic feasibility of spatially migrating load enabled by geographically distributed DCs.

2. Demand-side flexibility approaches

Despite recent research efforts examining the potential of DSF, according to the Energy Policy’s manuscript by Feuerriegel and Neumann (2014) “little is known about the economic dimension [of DSF] and further effort is strongly needed to realistically quantify the financial impact”. Because there are different approaches to DSF, economic analyses must take care of the decisive differences.

DSF is generally considered to be based on two types of approaches – load shedding and load shifting (Derakhshan et al., 2016; Feuerriegel and Neumann, 2014). Load shifting refers to the concept of postponing or putting forward an energy-consuming activity as to reduce load in peak hours – e.g. a charging process of an electric vehicle (Fridgen et al., 2016). Load shedding refers to the concept of ceasing
or not starting a planned energy-consuming activity without resuming it later in time – e.g. switching off a street lightning (Papagiannis et al., 2008).

In the domain of DC management the generally applicable two types of DSF approaches are extended by the idea of spatially “shifting” load by assigning requests to geographically distributed DCs, which is commonly referred to as load migration (e.g. Adnan et al., 2012; Wierman et al., 2014). In the following, we first briefly describe applications of load shifting and load shedding in the context of DCs before we, second, provide information about related works with regard to load migration.

2.1 Load shifting & load shedding
Load shifting requires load to be time-flexible. In the domain of DCs some types of requests are inflexible but some are delay-tolerant (Gmach et al., 2010). The latter type comprises scientific computing, routine tasks e.g. batch processing, and more recently especially bitcoin mining (Lewenberg et al., 2015; e.g. Nakamoto, 2008). Some DCs solely serve a single type of request. If this type is delay-tolerant the load shifting potential can be as high as the difference between maximum load and the load at idle state.

The possibility of creating DSF through load shedding typically comes with quality degradation under the constraints as quality-of-service (QoS) requirements and service level agreements (SLAs) as outlined by Wierman et al. (2014). As an example serve the growing number of big data algorithms – as they are used for targeting ads. A data center can reduce consumption by targeting ads less effectively (e.g. Baek and Chilimbi, 2010). Empirical studies by the Lawrence Berkeley National Laboratory found that 5% of the load can typically be shed in 5 minutes and 10% of the load can be shed in 15 minutes; and that these can be achieved without changes to how the IT workload is handled (Ghatikar et al., 2012, 2010).

2.2 Load migration
Geographically distributed DCs provide the opportunity of migrating load between locations of the DCs (Ghatikar et al., 2012, 2010; Kong and Liu, 2014) and thus contribute to grid stability by intelligently assigning (dispatching) the incoming workload to geographically distributed DCs.

There are a few contributions conducting economic analyses of load migration by DCs according to power price differences (e.g. Qureshi et al. 2009; Li et al. 2012; Zhang et al. 2012; Camacho et al. 2014). These
approaches indirectly contribute to medium and long term power grid stability, since the power market prices are, in general, determined by the available power supply and demand. However, price signals on power markets are not suitable for helping grid stability on short-term for two reasons: first, apart from (perhaps) real-time markets, lead time is too long, i.e. the time between gate closure and the time the contract becomes valid. Second, when trading on several markets, it cannot be assured that the DC’s primary objective of reducing procurement costs is aligned to the grid’s stability objective. This is because absolute price differences between markets can outplay relative (intra-market) price differences. So, the DC might benefit from migrating load to another market although its relative price is far above average. Chiu et al. (2012) propose a concept for grid balancing by intelligently dispatching incoming workload between geographically distributed DCs pursuant to local real-time price signals assuming they exist. However, in the majority of markets they do not.

The listed contributions illustrate that there are few options to marketing flexibility. To the best of our knowledge, previous research focusses on doing so on power markets and innovative (barely existing) bilateral products, only. Müller and Rammerstorfer (2008) show that delivering balancing power (BP) is not only another and already existing option for marketing flexibility but is specifically designed for this purpose. Rebours et al. (2007) illustrate that trades of BP are usually placed on so-called BP markets (BPM). In contrast to the referred articles, this paper provides flexibility by spatially migrating load and markets this flexibility on a BPM. We refer to this process as BP migration.

Yet, there is a potential major drawback of load migration, namely avoiding that one power grid improves its stability at the expense of another. For this reason, we base our analysis on the collaboration of balancing mechanisms (BM), e.g. pumped hydropower plants, and geographically distributed DCs. In the following, we describe a process based on that collaboration allowing BP migration without the potential drawback.

3. BP migration process

3.1 Setup

Generally, it is possible to apply the process presented in this section to a bidirectional migration of BP (between two locations). Since, in this paper, we merely strive to demonstrate the economic feasibility, we
nonetheless choose an irreducible setup considering a unidirectional BP migration setup. Mirroring the setup to the opposite direction facilitates bidirectional BP migration and may contribute to grid stability in both locations.

Our unidirectional setup consists of two DCs and one BM (figure 1). In our setup, both DCs can deliver the same services. Thus, each DC is capable of processing each individual incoming request. We consider a BPM design that distinguishes between positive and negative BP. Examples of this well-established market design are the regulation ancillary services in California (CAISO, 2015), the BPM in Finland (Fingrid, 2017) and the secondary and tertiary BP products of the German BPM (Consentec, 2014).

In our setup, the DCs are separated by distinct BPM areas. Only in this case price differences between these markets can be leveraged and the migration of BP can be economically feasible. In the first market area, there is $DC_{BPM}$ participating in the local BPM. In the second market area, there is $DC_{MEC}$ cooperating with a local BM. Furthermore, both DCs are charged by their local electricity providers at local power prices, e.g. power exchange prices of the respective location. This setup is especially relevant for large internet service providers such as Google, Amazon, and Microsoft, since they are known to operate geographically distributed cloud-scale DCs (Zhang et al., 2012).

Because geographically distributed DCs can often serve the same types of requests, each incoming request can be assigned to either of these DCs. Assigning the incoming requests is the main responsibility of the dispatcher which is a mandatory ICT component in today’s geographically distributed DC setups (e.g. Beloglazov and Buyya, 2010; Singh et al., 2010).
Figure 1 ICT-enabled spatial demand-side flexibility approach

3.2 Potential of DCs for offering BP

The maximum processing capacity (MPC) of a DC is the technically feasible upper bound to the number of incoming requests that a DC can process within a specific period (e.g. per second). On average only 40% of the MPC is required to serve the demand from incoming requests (Blagodurov et al., 2013). For reasons of energy management, servers can be switched off when idle (Tolia et al., 2008). The number of servers still running thus determines the current processing capacity (CPC), with $CPC \leq MPC$. The load of a DC, in turn, is greatly driven by the number of servers currently running.

In research and practice, there are several promising approaches demonstrating and further improving the proportionality between a DC’s CPC and its load (Gandhi et al., 2010; Ganesh et al., 2013; Lin et al., 2013; Tolia et al., 2008; Whitney and Delforge, 2014). Consequently, we assume:

**Assumption 1**: The load and the CPC of a DC are proportional.
While a positive dependence between the CPC and a DC’s load is indisputable, we build upon the findings by Gandhi et al. (2010) regarding the proportional dependence as it allows for a more tractable economic analysis.

In the following, \( \tau \) denotes the power demand required in a DC for processing one request within a specific period. Accordingly \( \tau \) is a measure of a DC’s energy efficiency. Due to the assumed proportionality between the load and the CPC of a DC, our first assumption implies that the energy efficiency of a DC is independent of the number of requests the DC processes within a specific period. Consequently, \( \tau \) is assumed to be independent of the DC’s CPC. Hence, the load of a DC can be calculated by \( \tau \cdot CPC \).

Additional circumstances influencing the energy efficiency of a DC are, for example, the installed hardware, the IT equipment density, and the outdoor temperature (Avelar et al., 2012). These circumstances are highly individual for each DC. Thus, the energy efficiency of different DCs can vary widely.

Google being one of the world’s largest cloud-scale DC operators regularly publishes the energy efficiency scores of its DCs (Google, 2017). Google’s DC are located in different climate zones and are equipped with different hardware. The published information about Google’s DCs show that the energy efficiency of cloud-scale DCs varies only marginally despite the different circumstances. Based on this fact and for reasons of simplicity, we assume the two DCs considered in this paper having the same energy efficiency: \( \tau_{DC_BPM} = \tau_{DC_MEC} = \tau \). This allows to abstract from DC specific circumstances and thus enables the demonstration of the economic feasibility of load migration in a more general manner. Nevertheless, in our evaluation, we also illustrate how the consideration of varying energy efficiencies for the two DCs affects the economic feasibility.

### 3.3 Amount of BP available for offering

In many BPMs, the offered BP must be readily available during the predefined offer period, i.e. the period during which the offered BP can be called. A typical period length is one week. As a result, the dispatcher has to ensure that the load, and thus the CPC of the two DCs is always appropriate to migrate the offered BP. However, the CPC of the DCs depends on the number of incoming request per period (e.g. per second) that must be processed by the two DCs. In general, the exact number of incoming requests is volatile and unknown a priori. Therefore, it is stochastic and must be forecasted. There is a number of scientific
contributions forecasting the number of incoming requests (e.g. Gmach et al. 2007; Jheng et al. 2014). Because of this and the fact that our main objective of this research is to provide an economic analysis of ICT-enabled spatial demand-side flexibility rather than on forecasting algorithms, we focus on the process of migrating BP and assume:

**Assumption 2**: The number of incoming requests can be forecasted with sufficient precision.

As a consequence, we treat in the following the amount of BP available for offering as known.

### 3.4 Bid prices

Next to determining the amount of BP available for offering, the dispatcher has to identify the bid prices. However, the remuneration methods of the existing BPMs vary greatly, e.g. a pay-as-bid or uniform price auction (Rebours et al., 2007). Furthermore, the remuneration structure also varies depending on whether or not the BP provider is paid for holding available and/or on call eventually delivering BP. Consequently, the determination of bid prices is highly market specific. However, there are several articles that consider pricing strategies for various remuneration methods (e.g. Swider 2006; Bajpai and Singh 2007; Krishna 2010). Since we strive to demonstrate the economic feasibility of the spatial demand-side flexibility approach, we apply a pricing strategy provided by Swider (2006) in our economic analysis to forecast bid prices. The strategy applied herein is based on established straightforward autoregressive moving average models and allows to determine a lower bound of the economic feasibility. Moreover, we also determine an upper bound of the economic feasibility by applying a second pricing strategy, which is based on perfect information.

### 3.5 Migration of BP collaboration

After having determined the amount and the bid prices of BP, the dispatcher places an offer for BP at the BPM. Provided that this offer is accepted, the dispatcher has to serve every call for BP during the offer period. In some BPMs the providers have to gradually adjust the BP actually delivered to the BP called. However, in order to make the economic analysis tractable, we state the following assumption:

**Assumption 3**: The BPM allows the abrupt rise of BP delivery to the total called amount without considering a gradual adjustment of the currently delivered BP (compare figure 2).
In the following, we illustrate the activities required to migrate positive and negative BP separately. In order to migrate BP, the dispatcher controls the CPC of the two DCs, adjusts the number of incoming requests dispatched to the DCs. Accordingly, this requires an ICT-enabled information exchange between the dispatcher and the DCs regarding their load. Note that the activities of our dispatcher described in this section are inspired by several contributions investigating how a dispatcher can migrate workload between geographically distributed DCs (e.g. Qureshi et al. 2009; Li et al. 2012; Zhang et al. 2012; Camacho et al. 2014). However, the dispatchers introduced in those approaches do neither migrate BP, nor control a local BM.

3.5.1 Migration of positive BP

In this section, we assume that a placed offer for positive BP in the amount of $BP_{bid}^+$ was accepted. Consequently, during the entire offer period, the BPM can call the dispatcher to deliver positive BP of at most $BP_{bid}^+$. Whenever there is a call in the amount of $BP_{called}^+ \leq BP_{bid}^+$, the dispatcher must reduce the load in $DC_{BPM}$ by $BP_{called}^+$ in order to counteract to the deficit of electricity. Typically, large power consumers, e.g. large-scale DCs, must report their expected load to the respective local electricity provider in advance. Accordingly, as long as the call lasts, the load of $DC_{BPM}$ must stay constant at the reported load reduced by $BP_{called}^+$, and the surplus power supply is passed on as BP (figure 2) to stabilize the power grid.

In order to decrease the load of $DC_{BPM}$ by $BP_{called}^+$, the dispatcher assigns less incoming requests to $DC_{BPM}$ and more incoming requests to $DC_{MEC}$. Since, according to our second assumption, both DCs have the same energy efficiency, this migrates load in the amount of $BP_{called}^+$ from $DC_{BPM}$ to $DC_{MEC}$.
We defined the energy efficiency \( \tau \) as the load required in a DC to process one incoming request within a given period (e.g. one second). Thus, to enable the migration of the offered BP (\( BP_{\text{bid}}^+ \)), it must always be possible to decrease the CPC of \( DC_{\text{BPM}} \) and to increase the CPC of \( DC_{\text{MEC}} \) by at most \( \Delta \lambda + \) \( + \) \( \text{bid} \):

\[
\Delta \lambda + \text{bid} = \frac{BP_{\text{bid}}^+}{\tau}
\]

The dispatcher must reserve a constant part \( \Delta \lambda + \text{bid} \) of the incoming requests during the entire offer period in order to always be ready for the migration of the load equivalent to the offered positive BP. As long as the called positive BP is zero, i.e. the dispatcher does not have to deliver positive BP, the reserved constant part \( \Delta \lambda + \text{bid} \) is dispatched to \( DC_{\text{BPM}} \). The resulting load is covered by electricity provider 1. Whenever the called positive BP changes, the dispatcher has to adjust the CPC and thus the load of the two DCs, in order to deliver the recently called positive BP. We refer to the phase at which the load of the two DCs is adjusted as the adjustment phase of a BP call.

In order to describe the activities required to adjust the called positive BP \( BP_{\text{called}}^+ \) during the adjustment phase, we look at the two possible cases separately: Case A applies whenever the BPM calls \( DC_{\text{BPM}} \) in order
to deliver more positive BP as to cover an increasing BP demand. Case A is typically caused by a new BP call which the DC has to satisfy (figure 3). Case B applies whenever the BPM calls $D_{BPM}$ as to deliver less positive BP to cover a decreasing BP demand. Case B is typically caused by the end of a BP call (figure 3).

The activities required in both cases are summarized in figure 3.

**Figure 3** Activities of the dispatcher in order to migrate positive BP

**Case A – The called positive BP increases**

An increasing called BP signifies that the dispatcher has to reduce the load of $D_{BPM}$. The first activity in order to reduce the load of $D_{BPM}$ is to raise the CPC of $D_{MEC}$, e.g. by switching additional servers on (activity A1-1). This instantly increases the load of $D_{MEC}$. Therefore, the BM simultaneously covers the increasing load of $D_{MEC}$ (activity A1-2). The second activity (activity A2) begins as soon as the dispatcher can assign more incoming requests to $D_{MEC}$, e.g. as soon as the servers of $D_{MEC}$ have started up. The dispatcher then assigns less incoming requests to $D_{BPM}$ by assigning more requests to $D_{MEC}$. As the third activity (activity A3-1), the dispatcher initiates the reduction of the CPC of $D_{BPM}$, e.g. by switching idle servers off. As soon as the CPC of $D_{BPM}$ is reduced, the load of $D_{BPM}$ decreases and the adjustment phase terminates. The excessive power supply of electricity provider 1 is passed on to deliver the called BP (activity A3-2). Consequently, whenever the called positive BP increases, the load of $D_{MEC}$ increases already at the beginning and the load of $D_{BPM}$ decreases only at the end of the adjustment phase.
Case B – The called positive BP decreases

Whenever the called positive BP decreases, the dispatcher has to increase the load of $DC_{BPM}$. The first activity (activity B1) is to increase the CPC of $DC_{BPM}$. This increases the load of $DC_{BPM}$ that is covered by electricity provider 1, and as a result, less positive BP is delivered. As soon as the additional CPC of $DC_{BPM}$ is available, the dispatcher assigns more incoming requests to $DC_{BPM}$, and thus less to $DC_{MEC}$ (activity B2). Within the third activity (activity B3), the dispatcher decreases the CPC of $DC_{MEC}$. As soon as the CPC of $DC_{MEC}$ is reduced, the adjustment phase terminates. Thus, whenever the called positive BP decreases, the load of $DC_{BPM}$ increases already at the beginning of the adjustment phase. However, the load of $DC_{MEC}$ that is covered by the local BM only decreases at the end of the adjustment phase.

Irrespective of whether the called BP increases or decreases, after the termination of the adjustment phase, $DC_{BPM}$ delivers the called positive BP in one location and $DC_{MEC}$ receives the corresponding positive BP in the other location until the called BP changes again. Consequently, the BP delivered by the BM is migrated by the distributed DCs to stabilize the power grid in a possibly very distant location.

3.5.2 Migration of negative BP

In contrast to the previous section, in this section, we suppose that a placed offer for negative BP in the amount of $BP_{bid}$ was accepted. Thus, during the entire offer period, the BPM can call the dispatcher to deliver negative BP of at most $BP_{bid}$. Whenever there is a call for negative BP in the amount of $BP_{called} \leq BP_{bid}$, the dispatcher must increase the load of $DC_{BPM}$ by $BP_{called}$ (figure 4) in order to counteract to an electricity surplus.
Accordingly, migrating negative BP is the reversed case of migrating positive BP. Remark that $\tau$ describes the load required in a DC to process one incoming request within a given period. Thus, the dispatcher must always be able to increase the CPC of $DC_{BPM}$ and to decrease the CPC of $DC_{MEC}$ by $\Delta_{\lambda}^{\max}$ in order to adjust the load of the two DCs, whereby $\Delta_{\lambda}^{bid}$ correspond to:

$$\Delta_{\lambda}^{bid} = \frac{BP_{bid}^-}{\tau}$$

Since the activities required to increase or decrease the load of the two DCs in order to migrate negative BP are analogous with regard to migrating positive BP, we do not describe them in more detail.

4. Evaluation of the BP migration’s economic feasibility

In the following section, we analyze the economic feasibility of BP migration for a representative real-world application reducing overhead capacities. Before illustrating the results, we provide an overview on the simulation setting and the simulation data set.
4.1 Simulation setting

We simulate the migration of BP from Finland to Germany to analyze the economic feasibility of geographical load migration. Therefore, we place BP offers on the German secondary BPM. As the offer period is one week the offered BP must be readily available during the entire week (Consentec, 2014). To offer BP in Germany, a BP provider has to bid a capacity price and an operating price for each offer period. Based on the bid prices the BPM decides through a pay-as-bid auction (Krishna, 2010), which BP providers are accepted to provision their BP. If the offer of one BP provider is accepted, the BPM pays the bid price irrespective of whether this BP provider must deliver BP within the offer period or not. Whenever BP is required, the BPM calls the accepted providers through a second pay-as-bid auction. The BP provider then is compensated by the bid operating prices.

Since the determination of the bid prices is not a core aspect of our paper, we use in this evaluation two pricing cases to identify a lower and an upper bound of the economic feasibility of migrating BP:

- Pricing case F: Forecasted prices (lower bound)
- Pricing case O: Optimal prices (upper bound)

In order to determine a lower bound of the economic feasibility, pricing case F is based exclusively on information about the past. When applying pricing case F we forecast the bid prices. In accordance with Swider (2006), we use a simple autoregressive model of order 1 that is refitted in every offer period by using historical data of the last two years. Accordingly, we use pricing case F to determine a lower bound of the economic feasibility of migrating BP in this real-world reproduction. More sophisticated forecasting models may result in considerably higher financial benefits, though.

In order to determine an upper bound of the economic feasibility, we apply pricing case O (optimal prices). Optimal prices are based on perfect information about not only historic but also future bid prices i.e. the bid prices result in the highest possible economic value.
DCs take a major share of the world’s power demand. However, the workload and the MPC of different DCs can vary greatly. In order to provide a generally valid and reproducible evaluation of the BP migration’s economic feasibility, we simulate two different scenarios. For each scenario, we conduct 100 simulation runs:

- Scenario $BP^+$: Offering 1MW of positive BP
- Scenario $BP^-$: Offering 1MW of negative BP

Furthermore, at the end of our evaluation we give a first estimation quantifying the economic value of load migration for two of Google’s DCs. Besides the scenarios in which BP is offered, we simulate a reference scenario at which no BP is offered. Given that scenario, the dispatcher assigns the incoming requests to this DC which can process them at the lowest cost without load migration:

- Scenario $BP^0$: The dispatcher offers no BP

The incoming request rate ($\lambda_{\text{incoming},t}$) corresponds to the number of incoming requests that must be processed by either of the two DCs at a given point of time $t$. $\lambda_{\text{incoming},t}$ can be divided into two parts: the reserved constant part necessary to migrate the offered positive or negative BP ($\Delta_{\lambda}^{bid}$) and the remaining part ($\lambda_t$) including volatility ($\lambda_{\text{incoming},t} = \Delta_{\lambda}^{bid} + \lambda_t$). As the second part is not affected by migrating positive or negative BP, it is omitted in this evaluation and we only consider the load caused by the constant incoming request rate required to deliver the offered BP ($\Delta_{\lambda}^{bid}$).

In this paper, we determine the economic value by the difference of the relevant costs and benefits in scenario $BP^0$ and $BP^+$ or $BP^-$. The relevant costs, as we define them herein, are the costs for the power required to process $\Delta_{\lambda}^{bid}$. Furthermore, we include the benefits from offering and delivering BP in Germany, and the opportunity costs for the BM in Finland. The opportunity costs are defined by the potential benefit a BM would receive when delivering the BP in Finland instead of migrating it to Germany.

### 4.2 Simulation data set

Regarding the timeframe of our simulation, we consider all offer periods of one entire year (from Monday, the 6th January 2014 to Monday, the 05th January, 2015) in order to compensate for seasonality effects. The
offer period is set to exactly one week (7 days). In the scenarios $BP^+$ and $BP^-$, we, consequently, offer positive and negative BP in 52 offer periods.

Regarding the BPM data, we use different sources. The German BPM operators publish real-world data about all accepted bids of the individual offer periods online (regelleistung.net, 2017). In addition, the amount of called BP is published on a quarter-hourly basis. However, according to Consentec (2014), the duration of a call for secondary BP lies in the range of 5 to 15 minutes. Since the actual duration of the individual calls is not published, we assume this duration to be equally distributed between 5 and 15 minutes. In order to determine the opportunity costs for the BM in Finland, we use the Finnish BP prices derived from Nord Pool Spot (2015). Regarding the local power prices in Finland and Germany, we use the hourly spot market prices of the respective power exchanges (EEX, 2017; Nord Pool Spot, 2015). Accordingly, we deliberately exclude (Brandt et al., 2014; Feuerriegel and Neumann, 2014; Ghatikar et al., 2012) taxes, levies, and subsidies as these are regularly adjusted and heavily influenced by political decisions. According to the underlying real-world data, the average local power price in Germany (mean: 32.66 €/MWh) is slightly lower than in Finland (mean: 35.96 €/MWh). Consequently, in the comparative scenario $BP^0$ all requests are processed in the German DC.

Regarding the DCs, we consider the duration required to increase and decrease the CPC of a DC as the startup and shutdown delay of servers. As recommended by Gandhi et al. (2010), we set the startup delay to 200 sec and the shutdown delay to 0 sec. We consider the processing time of the incoming requests to be negligibly short. This results in a duration of the adjustment phase of 200 sec. Furthermore, we suppose that processing same workload consumes the same amount of power in both DCs. All input values of our simulation are summarized in table 1.
<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>German BP offers and calls</td>
<td>Real-world data</td>
<td>regelleistung.net (2017)</td>
</tr>
<tr>
<td>Duration of German BP calls</td>
<td>Uniform (300 sec, 900 sec)</td>
<td>Following Consentec (2014)</td>
</tr>
<tr>
<td>German power prices</td>
<td>Real-world data</td>
<td>EEX (2017)</td>
</tr>
<tr>
<td>Duration of the adjustment phase</td>
<td>$d_{\text{adjust}} = 200 \text{ sec}$</td>
<td>Following Gandhi et al. (2010)</td>
</tr>
</tbody>
</table>

Table 1 Simulation input

4.3 Simulation results

Figure 5 visualizes the economic value per offer period achieved in our simulation by using pricing case F. Our analysis identifies three types of outcomes.

- First, we find offer periods revealing very high economic value from migrating BP. Especially the last two offer periods (December 22th, 2014 to January 4th, 2015) are of this type of outcome. According to the used real-world data, this peak can be explained by a comparatively high BP demand in Germany between Christmas and New Year. In this case, there were significantly more BP calls. Additionally, calls were priced at higher operating prices. As a result, migrating BP turned out to be profitable.

- Second, there are several offer periods, in which migrating BP did neither result in added economic value, nor did it cause economic losses. Examples are the following offer periods: offer period 1 in scenario $BP^+$ and $BP^-$, offer period 36 in scenario $BP^+$, and offer period 48 in scenario $BP^-$. In these periods, the respective bids for BP were not accepted and thus no BP was migrated resulting in zero economic value.

- Third, there are offer periods, in which the migration of negative BP resulted in a negative economic value, e.g. offer periods 30-33. In these offer periods, the relevant costs and benefits in scenario $BP^-$ are higher than in scenario $BP^0$. Migrating positive BP, however, results in a non-negative economic value for all offer periods.
Figure 5 Weekly economic value for offering 1 MW of BP achieved by pricing case F

Figure 6 visualizes the economic value for offering 1 MW of BP per offer period achieved by using pricing case O. As might be expected in this pricing case the economic value is higher than in pricing case F, since optimal bid prices result in an upper bound for this economic analysis. Especially between Christmas and New Year, pricing case O enables a considerably higher economic value than pricing case F. Furthermore, in offer period 42 the economic value of migrating negative BP is exceptionally high. According to the real-world data, this is due to unusually high bidden capacity prices in Germany during this offer period. In the offer periods 31 and 32 the economic value of migrating negative BP (scenario $B_P^-$). Even optimal bid prices would result in these two periods in a negative economic value and thus no BP offer is placed.
Comparing the two pricing cases and the two scenarios, we summarize descriptive statistics in table 2 and visualize them in figure 7 by a boxplot. All four combinations of pricing cases and scenarios show a positive median indicating that, aggregated over the entire year, in all four combinations the migration of BP results in a positive economic value. The difference of the economic value achieved by the two pricing strategies illustrates that pricing case F does not exploit the entire potential provided by migrating BP in this scenario.

**Figure 6** Weekly economic value for offering 1 MW of BP achieved by pricing case O
<table>
<thead>
<tr>
<th>Pricing Case</th>
<th>Scenario</th>
<th>Minimum</th>
<th>Lower Quartile</th>
<th>Median</th>
<th>Upper Quartile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forecasted prices</strong></td>
<td>$BP^+$</td>
<td>0 €</td>
<td>898 €</td>
<td>1 160 €</td>
<td>1 442 €</td>
<td>2 279 €</td>
</tr>
<tr>
<td></td>
<td>$BP^-$</td>
<td>-687 €</td>
<td>0 €</td>
<td>421 €</td>
<td>1 138 €</td>
<td>4 693 €</td>
</tr>
<tr>
<td><strong>Optimal prices</strong></td>
<td>$BP^+$</td>
<td>880 €</td>
<td>1 525 €</td>
<td>1 821 €</td>
<td>2 100 €</td>
<td>4 192 €</td>
</tr>
<tr>
<td></td>
<td>$BP^-$</td>
<td>0 €</td>
<td>870 €</td>
<td>1 530 €</td>
<td>2 304 €</td>
<td>9 439 €</td>
</tr>
</tbody>
</table>

**Table 2** Summary of the weekly economic value for offering 1 MW of BP averaged over one entire year

Irrespective of the applied pricing case, in scenario $BP^-$ the spread of the maximum and minimum value as well as of the lower and upper quartile of the weekly economic value is considerable higher than in scenario $BP^+$. Accordingly, the economic value in scenario $BP^-$ is subject to a higher volatility than in scenario $BP^+$. Especially in scenario $BP^-$, applying more advanced and tailored forecasting methods better coping with volatility might narrow the gap between the two pricing strategies. Analogously, this will help closing the gap between the upper and the lower bound for the economic value.
Table 3 aggregates descriptive statistics of the yearly economic value calculated in 100 simulation runs. The yearly economic value is defined by the sum of the economic value of the 52 offer periods which each have a duration of one week. Offering 1 MW of positive BP by BP migration results in a yearly economic value of 59 967 € (forecasted prices) and 99 936 € (optimal prices) based on 100 simulation runs. The offering 1 MW of negative BP by BP migration results in a yearly economic value of 31 114 € (forecasted prices) and 110 281 € (optimal prices) based on 100 simulation runs. Consequently, summarized over the entire simulated year, both pricing cases result in a positive economic value for migrating positive as well as negative BP. As we illustrate in a third evaluation case below, the power demand of today’s cloud-scale DCs typically allows for the migration of more than a MW of BP.

Furthermore, the results of 100 simulation runs confirm that the volatility of the yearly economic value brought in by the only stochastic component, namely the uniform distributed duration of the calls, is low. Accordingly, a higher number of simulation runs does not significantly change results.
<table>
<thead>
<tr>
<th>Pricing Case</th>
<th>Scenario</th>
<th>Minimum</th>
<th>Lower Quartile</th>
<th>Median</th>
<th>Upper Quartile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecasted prices</td>
<td>(BP^+)</td>
<td>58 812 €</td>
<td>59 729 €</td>
<td>59 984 €</td>
<td>60 245 €</td>
<td>61 132 €</td>
</tr>
<tr>
<td></td>
<td>(BP^-)</td>
<td>28 692 €</td>
<td>30 752 €</td>
<td>31 355 €</td>
<td>32 025 €</td>
<td>33 882 €</td>
</tr>
<tr>
<td>Optimal prices</td>
<td>(BP^+)</td>
<td>99 149 €</td>
<td>99 804 €</td>
<td>99 965 €</td>
<td>100 143 €</td>
<td>100 725 €</td>
</tr>
<tr>
<td></td>
<td>(BP^-)</td>
<td>100 413 €</td>
<td>101 225 €</td>
<td>101 473 €</td>
<td>101 874 €</td>
<td>102 867 €</td>
</tr>
</tbody>
</table>

Table 3 Summary of the yearly economic value (100 simulation runs)

Summarizing the above, both pricing cases generate economic incentives to migrate - positive as well as negative - BP. Note that even if the last two offer periods (Christmas and New Year) are not considered, offering and migrating BP still generates a positive economic value in both scenarios. Furthermore, within our simulation, we migrate on average 4.97MWh (forecasted prices) and 19.15MWh (optimal prices) of positive BP per offer period in scenario \(BP^+\). In scenario \(BP^-\), we migrate on average 25.04MWh (forecasted prices) and 54.09MWh (optimal prices) of negative BP per offer period. Accordingly, we migrate considerably more negative than positive BP. A reason for this might be that the optimal prices (pricing case O) are comparatively high for positive and comparatively low for negative BP. Consequently, in scenario \(BP^+\) we bid rather high prices and thus are rarely called and in scenario \(BP^-\), we bid rather low prices resulting in many BP calls.

To extend our evaluation to DC scale, we also give a first estimation of the economic value being achievable by two of Google’s real-world DCs spatially migrating load. Google as one of the biggest operators of large-scale DCs is regularly publishing information about its DCs’ energy efficiencies (Google, 2017). Since we simulate the BP migration from Finland to Germany, we consider Google’s DC in Hamina, Finland as the DC receiving BP \((DC_{MEC})\). As the second DC which provides BP at the BP market in Germany we actually have to consider a DC in Germany. However, due to the lack of a large scale DC Google is currently operating in Germany, we assume Google’s DC in St. Ghislain, Belgium \((DC_{BPM})\) to be located in the German market area. In fact, the DC is only about 150 kilometers away from the German border. Google, measuring the energy efficiency of DCs by a power usage effectiveness factor (Yuventi and Mehdizadeh, 2013), specifies for \(DC_{MEC}\) a factor of \(\tau_{DC_{MEC}} = 1.12\) and for \(DC_{BPM}\) a factor of \(\tau_{DC_{BPM}} = 1.09\) as the energy efficiency. A
smaller value of $\tau$ means a lower power demand is required to process the same task (Avelar et al., 2012; Hintemann and Clausen, 2014). Regarding the power consumption, Google does not publish appropriate specifications, which makes it impossible to exactly determine the amount of BP available for offering in those DCs. Thus, in line with Hintermann and Clausen (2014), we assume the maximum power consumption of large scale DCs is in the order of 50 MW. According to Koomey and Taylor (2015) 30 percent of the servers in today’s enterprise data centers are “comatose”, meaning that they “have delivered no information or compute services in 6 months or more”. Thus, we suppose in that the real-world DC setting consisting of the two Google-scale DCs could provide 15 MW of positive or negative BP at the BP market.

<table>
<thead>
<tr>
<th>Pricing Case</th>
<th>Scenario</th>
<th>Minimum</th>
<th>Lower Quartile</th>
<th>Median</th>
<th>Upper Quartile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecasted prices</td>
<td>$BP^+$</td>
<td>900 796 €</td>
<td>908 788 €</td>
<td>912 494 €</td>
<td>916 571 €</td>
<td>930 127 €</td>
</tr>
<tr>
<td></td>
<td>$BP^-$</td>
<td>295 089 €</td>
<td>336 861 €</td>
<td>346 326 €</td>
<td>354 458 €</td>
<td>381 511 €</td>
</tr>
<tr>
<td>Optimal prices</td>
<td>$BP^+$</td>
<td>1 447 357 €</td>
<td>1 453 695 €</td>
<td>1 456 418 €</td>
<td>1 458 819 €</td>
<td>1 464 997 €</td>
</tr>
<tr>
<td></td>
<td>$BP^-$</td>
<td>1 319 116 €</td>
<td>1 330 997 €</td>
<td>1 337 801 €</td>
<td>1 342 943 €</td>
<td>1 363 017 €</td>
</tr>
</tbody>
</table>

Table 4 Summary of the yearly economic value considering two Google DCs (100 simulation runs)

The increase of the provided BP (from 1 MW to 15 MW) as well as the consideration of the energy efficiency of two Google DCs resulted in an almost proportionally increased yearly economic value (Table 4). Thus, applying the optimal pricing strategy to migrate positive BP results in an economic value of over 1.4 million Euro and to migrate negative BP in over 1.3 million Euro per year. Accordingly, within our evaluation setting, the BP migration with two real-world DCs offers an enormous economic potential. This economic potential incentivizes a BM and spatially distributed DCs to cooperatively migrate BP between Finland and Germany.

Based on this setting, a BM operator, thus, benefits from additional options. A BM operator might - depending on the considered markets - choose between a BP market where calls may happen more frequently but capacity prices are low and vice versa, or at best reach out to a previously inaccessible market.
where both operating and capacity are higher than its own market region. An operator then chooses the alternative which fits best the specific cost structure of its BM.

5. Conclusion and policy implications

5.1 Policy implications

The economic analysis in the previous section provides a number of new insights into the economic value of spatial load migration in specific and DSF in general. The results of our analysis indicate economic feasibility for both positive and negative BP.

Germany serves as a reference case for a market region with high imbalance volume, while Finland serves as a reference for a market region with low volume but comparatively low grid balancing costs (van der Veen and Hakvoort, 2009). Both markets are well-suited to be connected in order to mitigate this differential. As mentioned in the introduction, transmission lines and market coupling are means of first choice. However, in absence of transmission lines, enough capacity or a lack of market integration, our analysis suggests to base decisions regarding the site selection of large-scale and energy-intensive data centers not only on wholesale market prices but also on the cost for balancing as indicated by e.g. BP prices. Our findings suggest that sites with either very high or very low BP prices are of particular relevance.

Well aware of the current pre-qualification criteria (i.e. the BPM’s requirements regarding a BM having it as a BP provider), policy-makers need to articulate, if they advocate load migration as a favorable concept. In that case the pre-qualification criteria have to be extended. The criteria have to be extended in a way so that they allow market-area-spanning BP providers to participate. And, even major DCs will probably provide few megawatts capacity, only. In order to seize the full potential of flexibility lifted through DCs, it is imperative to lower the minimum bidding size. As a reference, the European Energy Exchange serves bids from 100kW for some time now (EPEX SPOT SE, 2017).

Reconsidering the existing approaches to load migration, spot market prices are the primary factor for determining the market region one wants migrate to (e.g. Ghatikar et al., 2012). These load migration approaches, however, might import volatility to the receiving grid, if there is no counterbalancing BM. If penalties only base on peak loads (Kleindorfer and Fernando, 1993), they might not accurately capture the
incurred costs from the imported volatility. Thus reassessing the penalty schemes also taking volatility into account might help mitigating asymmetrical cost-benefit profiles for grid operators and DC operators.

5.2 Conclusion
In this paper, we have presented an economic analysis of spatial load migration as an alternative form of DSF compared to load-shifting and load-shedding. We found that spatially migrating load provides an interesting alternative to economically balancing a grid which has previously only been attributed to transmission lines. We studied in detail the use-case of distributed DCs participating in a BPM in one location, and allowing a BM to access another market region. For this use-case, we have proposed a BP migration process facilitating this. In the course of the analysis it became apparent that the process is not only influenced by but in fact enabled by ICT.

By our evaluation, we found that a setup (Germany, Finland) deploying the BP process is economically feasible for both positive and negative BP. This is because generally offering BP as a DC opens up a new stream of revenue and in the evaluation setting the gained benefits outweigh its costs. BM operators benefit from alternative and perhaps more suitable options to market their services (flexibility). While this result, of course, supports the paper’s objective, we like to believe that our approach contributes to a broader discussion relating to spatial DSF as well as current efforts to integrate markets across regions and nations. The approach as presented in this paper provides a process that in a way “virtualizes” BP and allows BMs such as hydro power plants to participate in remote market areas possibly even globally.

On a market region level, consumers of BP might profit from cheaper prices (because of ceteris paribus more supply). Additionally, integrating DCs in the current array of facilities to provide grid ancillary services reduces need of back-up (standby) capacities. Besides, standby capacities run in majority on fossil-fuels because they necessarily have to be quickly dispatchable. Reducing overhead capacities hence leaves plenty room for cutting CO2-emissions.

5.3 Directions for further research
This study explores the economics of a use-case for spatial DSF. It is supposed to serve as a starting point for a larger policy-related discourse on its benefits and potentials. In order to spur interest and stimulate
further discussion on the topic of spatial load migration as a third type of DSF, we especially invite academics to devote a share of their time thinking out-of-the-box and start identifying additional use-cases for load migration apart from the DC-case as presented herein. From this use-case alone follow a number of directions for further research. First, we see a number of immediate extensions of the approach presented in this work to follow-up on with. Second, on a more general and global level, it might be relevant for research to study the economic and political implications of our approach.

Considering the more immediate extensions of this work, the process as presented in this paper integrates the steps “determining the amount of BP” and “setting bid prices”. While we base our analysis on established methods, the results obtained from the optimal pricing case signal room for improvement. Thus it could be interesting from both an academic as well as practical point view to tailor forecasting approaches for this specific setting. This might further increase the potential of the BP migration process. The second immediate extension of this work might even be more profound. Our approach might be well-suited to be adapted to improve congestion management. While we strive to virtually connect two market areas, congestion management extends the idea to virtually connect two (non-)neighboring balancing groups on a distribution grid level. We estimate the DCs’ impact to be even more beneficial in relative terms, because balancing group’s deviations might be smaller in magnitude than deviations on a market area level.

Considering the more extensive and global extensions of this work, on a market area level, we expect that virtually connecting more than just two market areas (and considering alternate forms of ancillary services) yields elevated levels of flexibility based on ICT-managed complexity. The approach proposed with this paper requires DCs to be of significant size. In order to seize a greater share of the potential, we suggest to discuss the idea of pooling DCs in a market region – quite similar to the concept of virtual power plants (Saboori et al., 2011). In this regard, we like to carry on the economic analysis as to better understand how market prices or CO2-emissions in the respective regions might develop. Eventually, we will continue with and appreciate to see future studies, experiments, implementations evolving around the idea of globalized power and ancillary services markets e.g. based on - in theory - worldwide distributed DCs.
References


Eurelectric, 2014. Flexibility and Aggregation Requirements for their interaction in the market.


Fingrid, 2017. Fingrid - Balancing power market [WWW Document]. URL


Koomey, J., Taylor, J., 2015. New data supports finding that 30 percent of servers are “Comatose”, indicating that nearly a third of capital in enterprise data centers is wasted.


