

Energie- und kostenorientiertes, auf Virtualisierung basiertes Last- und Powermanagement in einem Rechenzentrumsverbund

Dissertation zur Erlangung des Grades eines
Doktors der Naturwissenschaften

von

Dipl.-Inform. Kiril Schröder

Gutachter:

Prof. Dr.-Ing. Wolfgang Nebel

Prof. Dr. Michael Sonnenschein

Prof. Dr. Lorenz M. Hilty

Tag der Disputation: 29. Februar 2016

Abstract

The increasing energy demand of the ICT sector (information and communication technology) as well as the increasing energy price are influencing the operation of IT devices. In particular, data centers are affected because of additional effort in outsourcing, which has become more important with the advent of cloud computing. Thus, services and the energy demand has been centralized in big or mega data centers.

There are various methods for reducing the energy demand of data centers. Consolidating servers by using virtualization technology is one of them. With this method, services which have been operated on dedicated servers for isolation purposes are now operated in virtual machines (VMs). Encapsulated in VMs, several services can now be operated on a single physical server—still isolated from each other. By using live migration, one can also transfer VMs at runtime resulting in adapting the energy demand to the latest resource demand.

This flexibility can be used in a group of data centers to transfer loads to sites that can be operated most beneficially. The outdoor temperature and the local electricity price are parameters among others which influence the operational costs of data centers. The required energy or the costs can be reduced when there are some fluctuations in the parameter values. The greater the fluctuations are and the more the order of the most beneficial sites is altering, the higher the savings potential for site comprehensive load and power management (LPM) usually gets.

Such a group of data centers can be built by providers which operate several data centers or by cooperating operators of single data centers. Especially in the latter case, there are issues concerning standardizing provisioning and billing as well as data security and control mechanisms. Cloud computing also has to address these issues, so it could be seen as an enabler. When the technical aspects are solved, the flexible infrastructure can be used for demand side management as well.

Previous works determining the potential of site comprehensive LPM used large scale services for their research. These services are operated on several servers and have to be replicated on each operated site. However, the energy demand for transferring IT loads and for replicating data has not been taken into account. Moreover, the savings potential has only been determined for specific scenarios whereby no general statement can be derived. The influence of different parameters on the savings potential has not been pointed out either. Additionally, the scaling behavior of the proposed methods is not practical, not allowing them to be used for planning purposes (capacity and site planning).

Therefore, the architecture of the developed method in this dissertation is designed to be multilayered as a multi agent system. In order to be able to leave the management in control of the data center operator, the behavior of the local LPM regarding the influenced energy demand has been modelled. Together with further energy demand models for the components in the data center, which impact the energy demand most, an entire data center energy demand model has been developed enabling setting up forecasts for operating a specific set of VMs.

A local agent uses this model with a selection method, also developed, to determine appropriate VMs for reaching a targeted load profile. In coordination with other local agents and considering a model to determine the energy demand for migrating VMs the whole group

of data centers can be optimized attaining cost or energy demand minimization. For each agent, a runtime complexity, which depends on the VM and data center count, has been accomplished: $O(\#VM \cdot \#DC)$.

Based on measured data for the resource demand of services, the outdoor temperatures of more than 30 globally distributed cities, and the day ahead electricity prices of three European electricity exchanges, many different scenarios have been built in order to be able to derive a more general statement about the typical savings potential. Savings up to 10% or 40% for the total energy demand or the total operational costs respectively have been determined. Furthermore, the parameter values have been purposefully modified to point out their influence.

A great factor compensating the savings is caused by the replication costs. Reducing the count of replicas decreases this portion of costs but also the savings potential because of the reduced variety of possible sites for transferring loads. The developed site selection method solves this trade off by determining appropriate replicas for an overall optimization.

All methods in total are solving the mentioned challenges and issues and, therefore, they can be used in a planning tool. Another field of application is assigning costs in a data center group optimized operation. By analyzing the operation, a transparent cost model for operating VMs can be created, which is generally beneficial for operators and customers.

Kurzfassung

Der steigende Energiebedarf der IKT-Branche (Informations- und Kommunikationstechnologie) sowie auch der zugleich steigende Energiepreis beeinflussen den Betrieb von IT-Geräten in zunehmendem Maße. Insbesondere in Rechenzentren lässt sich diese Entwicklung beobachten, da durch den vermehrten Einsatz von Outsourcing, welches durch das Cloud Computing eine noch bedeutendere Rolle erhalten hat, Dienste und damit der Energiebedarf auf Groß- bis Megarechenzentren zentralisiert bzw. gebündelt werden.

Es existiert eine Reihe von Maßnahmen, die die Senkung des Energiebedarfs von Rechenzentren bewirken. Eine davon ist die Konsolidierung von Servern durch den Einsatz von Virtualisierungstechnologie. Hier werden Dienste, die vormals aus Isolationsgründen auf einem dedizierten Server betrieben wurden, in sogenannten virtuellen Maschinen (VMs) gekapselt und können dadurch – immer noch isoliert von anderen Diensten – gemeinsam auf einem Server betrieben werden. Mit Hilfe der Live-Migrationstechnik ist auch die Verschiebung von VMs zur Laufzeit möglich, wodurch eine Anpassung des Energiebedarfs an den aktuellen Ressourcenbedarf erfolgen kann.

Diese Flexibilität kann in einem Verbund von Rechenzentren genutzt werden, um Lasten an günstigste betreibbare Standorte zu verlagern. Parameter, die die Betriebskosten eines Rechenzentrums beeinflussen, sind z. B. die Umgebungstemperatur und der Strompreis. Energie und Kosten können eingespart werden, wenn die entsprechenden Parameter eine gewisse Fluktuation aufweisen: Je größer diese ist und je öfter sich die Ordnung der günstigsten Betriebsstandorte ändert, desto größer ist i. d. R. das Einsparpotential eines standortübergreifenden Last- und Powermanagements (LPM).

Ein solcher Verbund kann von Betreibern mehrerer Rechenzentren gebildet werden oder auch durch Kooperation von Betreibern einzelner Rechenzentren. Gerade in letzterem Fall kommen jedoch verstärkt Problemstellungen bzgl. Standardisierung beim Provisioning und Billing sowie bzgl. Datensicherheit und Kontrollmechanismen zutage. Mit diesen Aspekten muss sich auch das Cloud Computing auseinandersetzen, so dass es als ein Vorreiter bzw. Enabler angesehen werden kann. Sind die technischen Randaspekte gelöst, kann eine solche flexible Infrastruktur auch z. B. für Demand Side Management genutzt werden.

Bisherige Arbeiten, die sich mit dem Einsparpotential eines standortübergreifenden LPM auseinandergesetzt haben, stützen ihre Untersuchungen hauptsächlich auf Large-Scale-Services, welche auf mehreren Servern betrieben und an unterschiedliche Standorte repliziert werden. Der Energiebedarf für den Verlagerungsprozess sowie für das Replizieren des Datenbestandes wird jedoch nicht berücksichtigt. Außerdem wurden die Einsparpotentiale nur für bestimmte Szenarien ermittelt, wodurch keine allgemeine Aussage abgeleitet werden kann. Der Einfluss der auf das Einsparpotential wirkenden Parameter wurde nicht herausgestellt. Ferner mangelt es den Verfahren an einer praxistauglichen Skalierbarkeit, die den Einsatz nicht nur im Produktivbetrieb, sondern auch zu Planungszwecken (Kapazitäts- und Standortplanung) ermöglicht.

Das in dieser Dissertation entwickelte Verfahren wurde zu diesem Zweck in seiner Architektur mehrschichtig und als ein Multi-Agenten-System konzipiert. Damit das Management eines Rechenzentrums in der Hand des jeweiligen Betreibers verbleibt, wird von dem eingesetzten rechenzentrumsbezogenen LPM bzgl. seines Einflusses auf den Energiebedarf abstrahiert.

Zusammen mit weiteren erstellten Energiebedarfsmodellen für die im Rechenzentrum für den Energiebedarf hauptverantwortlichen Komponenten konnte ein Gesamtmodell geschaffen werden, das bei gegebenen Umgebungstemperaturen für eine Menge an zu betreibenden VMs eine Energiebedarfsabschätzung gibt. Ein für den Standort vorgesehener Agent kann mit diesem Modell und einem ebenfalls entwickelten Selektionsverfahren geeignete VMs ermitteln, um einen gewünschten Lastgang nachzustellen. In Koordination mit weiteren Standortagenten und unter Berücksichtigung eines Netzwerkmodells zur Bestimmung des Energiebedarfs bei der Migration von VMs wird der Betrieb im gesamten Verbund entsprechend der Zielstellung Energieminimierung, Kostenminimierung oder Lastanpassung optimiert. Es konnte je Agent eine zur VM- und Rechenzentrumsanzahl jeweils linear abhängige Laufzeitkomplexität erreicht werden: $O(\#VM \cdot \#DC)$.

Basierend auf realen Daten bzgl. des Ressourcenbedarfs von Diensten, den Umgebungstemperaturen von mehr als 30 über die Welt verteilten Städten und den Day-Ahead-Strompreisen von drei europäischen Strombörsen, wurden unterschiedlichste Szenarien gebildet und die Parameter in ihrer Ausprägung gezielt variiert, um zum einen eine allgemeingültigere Aussage über das Einsparpotential zu erhalten und zum anderen auch den jeweiligen Parametereinfluss herauszustellen. Es konnten Einsparungen von bis zu 10 % bzw. 40 % bezogen auf den Gesamtenergiebedarf bzw. die Gesamtkosten ermittelt werden.

Ein großer Faktor, der den Einspareffekt kompensiert, ist der Kostenanteil durch Replikation. Durch Reduzierung der Replikatanzahl kann dieser Kostenanteil zwar gesenkt werden, jedoch mindert dies auch den potentiellen Einspareffekt durch Lastverlagerung aufgrund der geringeren Auswahl an möglichen Standorten. Das entwickelte Standortselektionsverfahren löst diesen Trade-Off, indem es geeignete Replikate zur Gesamtoptimierung ermittelt.

Das Gesamtverfahren löst damit die benannten Herausforderungen und Problemstellungen und kann als Planungswerkzeug eingesetzt werden. Es ist auch dazu verwendbar, um den Kosteneinfluss im optimierten Verbundbetrieb herauszustellen. Hieraus lässt sich ein transparenteres Kostenmodell für den Betrieb von VMs entwickeln, was sich i.d.R. positiv auf Betreiber und Kunden auswirkt.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.1.1	Maßnahmen zur Energie- und Kosteneinsparung	6
1.1.2	Globales Last- und Powermanagement	10
1.2	Ziele der Dissertation	15
1.3	Aufbau der Arbeit	16
2	Stand der Forschung und Technik	17
2.1	Technologischer Kontext	17
2.1.1	Rechenzentrum	17
2.1.2	Server-Virtualisierung und Live-Migration	23
2.1.3	Last- und Powermanagement auf Systemebene	29
2.1.4	Cloud Computing	31
2.2	Verwandte Arbeiten	38
2.2.1	Klassifizierender Vergleich	39
2.2.2	Abgrenzung	42
2.3	Zusammenfassung	46
3	Technische Voraussetzungen	47
3.1	Datenbasis & adäquate Vorhersagemodelle	47
3.1.1	Vorhersagemethodik für den Ressourcenbedarf von Diensten	48
3.1.2	Wettervorhersage	50
3.1.3	Strompreise	51
3.2	Konzept eines rechenzentrumsinternen Last- und Powermanagements	53
3.3	Zusammenfassung	56
4	Konzeption samt Komponentenevaluation	57
4.1	Architektur	58
4.1.1	Konzeption	59
4.1.2	Einordnung	61
4.1.3	Systemüberblick	62
4.2	Energiebedarfsmodell eines Rechenzentrums	63
4.2.1	Modelle regelbarer Komponenten im Rechenzentrum	64
4.2.2	Lineares Gesamtenergiebedarfsmodell	67
4.3	Rechenzentrumsinternes Last- und Powermanagement	69
4.3.1	Konzept der iterativen Migrationsentscheidung	69
4.3.2	Heterogene Hardware	70
4.3.3	Dynamisierungserweiterungen	74
4.3.4	Auswirkungen der Modifikationen	81
4.4	Abstraktion des LPM-Konsolidierungsverhaltens	85
4.4.1	Konzeptionierung eines linearen Regressionsmodells	86
4.4.2	Modellermittlung und -analyse	89

4.4.3	Evaluation der LPM-Abstraktion	96
4.5	Lastganggenerierung	99
4.5.1	Problembeschreibung	99
4.5.2	GRASP	100
4.5.3	Evaluation der Lastganggenerierung	103
4.6	Netzwerkmodell	109
4.7	Globales Last- und Powermanagement	111
4.7.1	Lastgangvorgabe	112
4.7.2	Migrationsentscheidung	113
4.7.3	Migrationsdurchführung	115
4.7.4	Evaluation der Migrationsentscheidung	115
4.7.5	Laufzeitkomplexität	117
4.8	Replikation – Standortselektion	119
4.8.1	Verwandte Arbeiten	120
4.8.2	Heuristik zur Standortselektion	127
4.8.3	Evaluation der Standortselektion	132
4.8.4	Beschleunigtes Suchverfahren	135
4.9	Zusammenfassung	138
5	Gesamtevaluation und Analyse	139
5.1	Simulationsumgebung & Datenbasis	139
5.2	Einsparpotential – Sensibilitätsanalyse	145
5.2.1	Migrationsgeschwindigkeit	146
5.2.2	Strompreisverläufe	147
5.2.3	Temperaturverläufe	149
5.2.4	VM-Profil-Korrelation	150
5.2.5	Netzwerkenergiebedarf & Optimierungsintervall	150
5.2.6	Freie Rechenzentrumskapazität	154
5.2.7	Standortanzahl	155
5.3	Diskussion	156
5.3.1	Standort- bzw. Verbundoptimierung	157
5.3.2	Bepreisung & Profil-Klassifizierung von virtuellen Maschinen	158
5.3.3	Globales LPM beim Demand Side Management	162
5.4	Zusammenfassung	165
6	Fazit	167
7	Ausblick	169
	Glossar und Abkürzungsverzeichnis	171
	Abbildungsverzeichnis	177
	Literaturverzeichnis	181

1 Einleitung

Die Entwicklung von Rechenzentren führte aufgrund sowohl technisch administrativer als auch ökonomischer Gesichtspunkte zu sogenannten Cloud-Rechenzentren, in denen der Bedarf an Ressourcen und die allgemeine Bereitstellung von Diensten flexibel angepasst werden kann. Diese Flexibilität ermöglicht es, innerhalb eines globalen Prozesses auf standortbedingte Parameter wie die Außentemperatur und den Strompreis zur Minimierung des Gesamtenergiebedarfs bzw. der Gesamtkosten einzugehen. Im Rahmen der vorliegenden Dissertation wird ein entsprechendes Verfahren zum Last- und Powermanagement (LPM) konzipiert, welches speziell die Verteilung von Server-Sized-Services, also solchen, die auf einem einzigen Server betrieben werden können, berücksichtigt. Ausgangslage ist ein kooperativer Verbund von Cloud-Rechenzentren. Für die Betreiber der Rechenzentren gilt es festzustellen, welche Einsparpotentiale unter welchen Bedingungen, also z. B. bei welchen Veränderungen der Strompreise und Temperaturen gehoben und inwieweit diese an Endkunden weitergegeben werden können. Weitere Ziele, die durch ein globales LPM verfolgt werden können, sind der CO_2 -neutrale Betrieb, die Erhöhung der Flexibilität und Verfügbarkeit sowie eine Partizipation am Smart Grid.

Im Folgenden werden zunächst die Motivation und der Themenkontext der Arbeit gegeben und im Anschluss die Ziele der Dissertation konkretisiert.

1.1 Motivation

Durch den technologischen Fortschritt wird kontinuierlich eine Erhöhung der Lebensqualität erreicht. Dies bezieht sich insbesondere auf den Konsum von Luxusgütern. Wurden einst das Radio oder der Fernseher als Luxusgüter betrachtet, gehören diese bereits seit langem zu den Grundbedarfsgütern. Elektronische Geräte für alle Lebenslagen, sogar auch in mehrfacher Ausführung, gehören nun zum Alltag. Hierzu zählen Notebooks, Tablets, Smartphones, aber auch GPS-Uhren, Digitalkameras und die Heimkinoanlage. All diese Geräte benötigen nicht nur zum Betrieb Energie, sondern auch ihre Herstellung ist mit hohem Energie- und Ressourcenaufwand verbunden. Insgesamt konnte in den vergangenen Jahren ein wachsender Energiebedarf festgestellt werden, wie in Abbildung 1.1 zu erkennen ist.

Auch der Bedarf der IKT-Branche (Informations- und Kommunikationstechnologie) an elektrischer Energie hat sich in den vergangenen Jahren deutlich erhöht. Wie bereits aus einer Studie von AMD aus dem Jahre 2005 hervorgeht, lag allein der Strombedarf der US-amerikanischen Rechenzentren bei 45 TWh, was Kosten in Höhe von 2,7 Milliarden

1 Einleitung

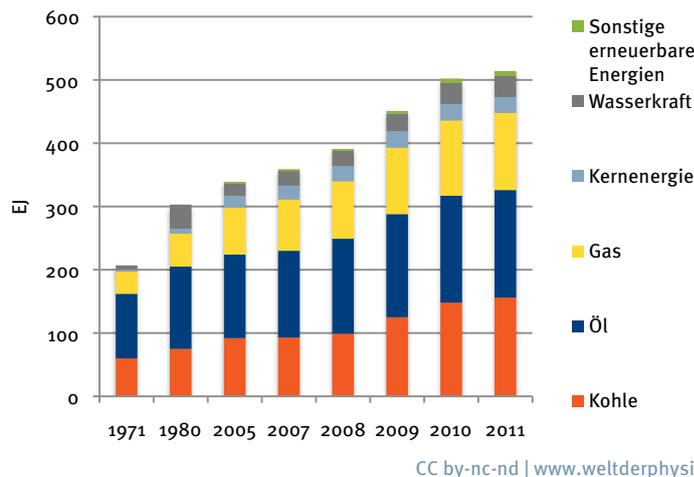


Abbildung 1.1: Der Weltenergiebedarf hat sich von 1971 bis 2011 mehr als verdoppelt.

Dollar entsprochen hat [9]. Greenpeace prognostiziert, dass der globale Strombedarf von Rechenzentren und Telekommunikationsnetzen von 623 TWh im Jahre 2007 auf über 1900 TWh im Jahre 2020 anwachsen werde [94]. Hierbei ist jedoch auch zu beachten, dass nur 3,6 % der benötigten Energie für die eigentliche Rechenarbeit der Server verwendet wird, was sich durch die Verluste in der Stromgenerierung, Transport und Transformationen begründet [115].

In Deutschland zählt die IKT-Branche mittlerweile mit einem Anteil von 10,5 % [213] zu den Großverbrauchern elektrischer Energie. Allein dieser Anteil hätte nur zu 70 % durch die gesamte in Deutschland durch Windkraft erzeugte Energie versorgt werden können.¹ Durch den Ausbau von Windkraftanlagen konnte die Versorgung zwar erhöht werden,² jedoch steigt auch der Energiebedarf der IKT.

In Abbildung 1.2 ist eine Prognose der Fraunhofer Institute IZM sowie ISI für den deutschen IKT-Energiebedarf angegeben [213]. Die Konsumelektronik, allen voran die Unterhaltungselektronik (Fernseher etc.), nimmt den größten Anteil ein. Aus den Prognosen ist allerdings eine Stagnation des Strombedarfs in diesem Bereich zu erkennen. Kritischer sieht die Lage in den Bereichen der Rechenzentren und des Internets aus, in denen durch eine exponentiell wachsende Nachfrage auch der Strombedarf entsprechend ansteigen wird.

Aber nicht nur der Strombedarf, sondern auch die Strompreise sind kontinuierlich angestiegen,³ wodurch sich die Entwicklung der anwachsenden Stromkosten verstärkt bemerkbar macht. Mit 9,7 TWh jährlichem Strombedarf betragen die Stromkosten von Servern nach einer Kurzstudie von Borderstep [103] für das Jahr 2011 rund 1,2 Mrd. Euro. Weltweit wird von einem Anteil von 20 % des IT-Budgets allein für die Stromkosten ausgegangen [116].

¹Stand 2007: IKT-Strombedarf: 55 TWh [213], Einspeisung durch Windkraft: 39,5 TWh [128]

²Einspeisung durch Windkraft, 2014: 51,45 TWh; Quelle: https://www.energy-charts.de/energy_de.htm, letzter Zugriff: 16.07.2015

³<http://www.kwh-preis.de/strompreis-dossier-teil-3-vergleich-strompreise-industrie-und-privathaushalt>, letzter Zugriff: 25.04.2015

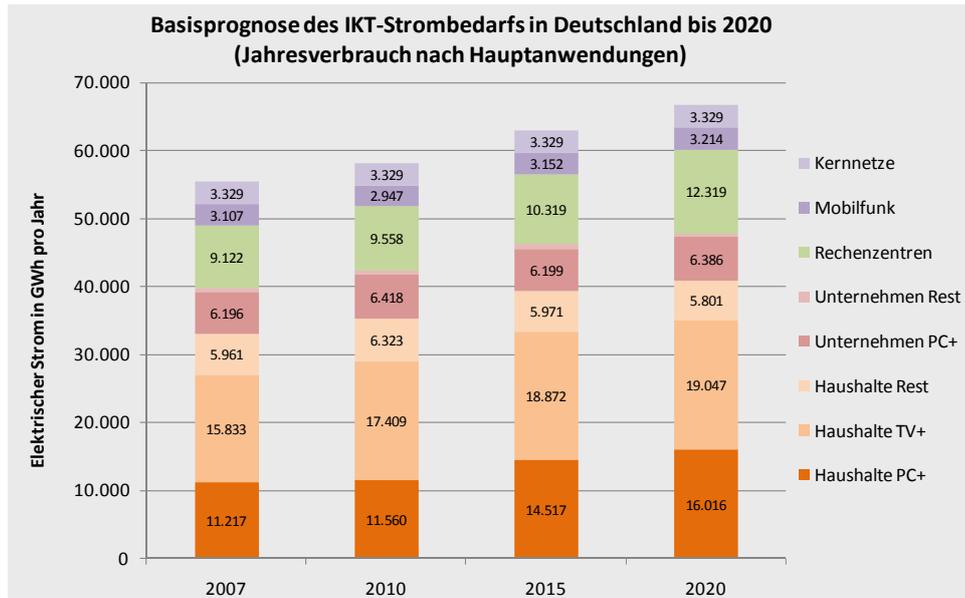


Abbildung 1.2: Prognose über den IKT-bedingten Strombedarf in Deutschland bis 2020 [213]

Der steigende Strombedarf und die steigenden Strompreise tragen zu einer erhöhten wirtschaftlichen Belastung von Industrie und Gewerbe bei. Auch in der IKT-Branche hat diese Entwicklung zu notwendigen Gegenmaßnahmen geführt. So kann durch fortlaufenden Technologiefortschritt die Effizienz von Hardware, die den Zusammenhang von Rechenleistung pro Watt angibt, gesteigert werden [130]. Die Recheneffizienz wird auch in den nächsten Jahren weiter verbessert werden, jedoch steht demgegenüber das Wachstum des Rechenleistungsbedarfs, welcher, wie in Abbildung 1.3 dargestellt, deutlich stärker ansteigt [23]: In den Jahren von 1999 bis 2007 konnte die Energieeffizienz von Hardware um das 16-Fache erhöht werden, im gleichen Zeitraum ist allerdings die Rechenleistung um das 75-Fache angestiegen.

Die Möglichkeiten der automatischen elektronischen Verarbeitung steigen mit der Leistungsfähigkeit der Server, so dass auch die Menge an Datenverarbeitungsaufträgen immer weiter zunimmt. Zu den aktuellen Beispielen zählt u.a. das in Chile geplante Large Synoptic Survey Telescope (LSST), welches nächtlich 30 TB an Daten generieren soll.⁴ Die entsprechenden Systeme sind dabei auf Performanz und Verfügbarkeit optimiert, was durch entsprechende Backup- und Replikationssysteme bewerkstelligt wird. Doch dies ist folglich auch mit erhöhten Kosten verbunden.

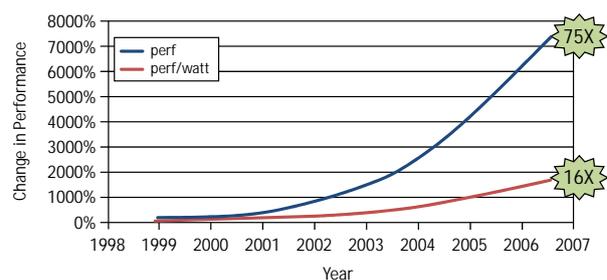


Abbildung 1.3: Die Rechenleistung und Rechenleistung pro Watt von Servern steigen unterschiedlich schnell [23].

⁴<http://www.lsst.org/lsst/science>, letzter Zugriff: 25.11.2014

1 Einleitung

Die höhere Effizienz und Rechenleistung führt in Rechenzentren zu einer Steigerung der Energiedichte von Servern und Racks. Während vor 20 Jahren Racks eine Leistungsaufnahme von 2 bis 3 kW hatten, kommen heutzutage Racks mit einer Leistungsaufnahme im unteren zweistelligen Kilowattbereich zum Einsatz [114, 113]. Beim Supercomputer Blue Gene/Q werden sogar mehr als 80 kW pro Rack erreicht [148]. Diese Leistungsdaten liegen deutlich höher als jene einer Heizungsanlage für ein Einfamilienhaus (bis zu 18 kW) [39]. Um der thermischen Mehrbelastung Herr zu werden, sind entsprechend leistungsfähige Kühlungstechniken notwendig, wodurch der Strombedarf abermals steigt.

Sogenannte Megarechenzentren, in denen zehntausende von Servern oder mehr betrieben werden [92], werden heutzutage dahingehend konzipiert, dass ihre Leistungsaufnahme in voller Ausbaustufe bis in den dreistelligen MW-Bereich hineinreicht: Microsoft Data Center in Chicago (60 MW), NSA Data Center in Utah (65 MW), QTS Metro Data Center in Atlanta (80 MW), 350 East Cermak in Chicago (100 MW), The Supernap in Las Vegas (250 MW).⁵

Der hohe Strombedarf hat nicht nur ökonomische Nachteile, sondern verursacht auch hohe CO_2 -Emissionen. Dies ist dadurch begründet, dass die Stromgenerierung (noch) zu großen Teilen auf der Verbrennung fossiler Energieträger (Braun-/Steinkohle und Erdgas) basiert. Den CO_2 -Emissionsfaktor für den Strommix in Deutschland zugrunde gelegt, lagen die Emissionen durch die IKT im Jahr 2010 bei 31,4 Mio. t CO_2 . Zum Vergleich: Der gesamte deutsche Flugverkehr liegt mit emittierten 52 Mio. t CO_2 in der gleichen Größenordnung [70]. Um diese ökologische Belastung zu senken, ist es möglich, den Strom aus regenerativen Energiequellen zu gewinnen, wie z. B. bei einem Rechenzentrum von Google in Schweden⁶ und bei zwei geplanten Rechenzentren von Apple in Dänemark und Irland⁷.

Zur Erhöhung der ökologischen Verträglichkeit eines Rechenzentrums gibt es neben den Möglichkeiten zur Anpassung der Eingangsparameter auch Maßnahmen bzgl. der Ausgangsparameter. Hiermit ist hauptsächlich die von den Rechenzentren in Wärme umgewandelte Energie gemeint. So kann die Abwärme für die Beheizung von Wohnsiedlungen⁸ oder sogar ganzen Städten wie Helsinki⁹ genutzt werden, was die Gesamtumweltbelastung reduziert.

Green-IT

Unter dem öffentlichkeitswirksamen Begriff Green-IT wurden die Auswirkungen der IKT auf die Umwelt herausgestellt. Neben der verursachten Umweltbelastung konnten aber auch unterschiedlichste Anwendungen (Hausautomatisierung, Energiespeicherung und -management, IP-Video-Konferenzen anstelle von Flugreisen etc.) gefunden werden, die zu einer drastischen Senkung der Umweltbelastung führen können: Eine Reduzierung um 207 Mio. t CO_2 sei bis 2020 möglich [163]. Dies fällt unter den Begriff IT-for-Green [82]. Unternehmen, die IT

⁵<http://www.datacenterknowledge.com/special-report-the-worlds-largest-data-centers/>, letzter Zugriff: 25.11.2014

⁶<http://heise.de/-1341059>, letzter Zugriff: 16.07.2015

⁷<http://www.faz.net/-i2r-80672>, letzter Zugriff: 16.07.2015

⁸<http://www.cowo.de/a/1893098>, letzter Zugriff: 25.11.2014

⁹<http://www.cowo.de/a/1912083>, letzter Zugriff: 25.11.2014

gestützte Maßnahmen umsetzen – in dem Fall als Greening Company bezeichnet –, können dadurch ihr Unternehmensimage verbessern, was ihre Wettbewerbsposition bei Verbrauchern und Investoren stärkt. Der Einsatz von IT zur Senkung der Klimagase und Reduzierung des Ressourcenverbrauchs ist letztendlich wirtschaftlich motiviert [149, 241].

Es existiert eine Vielzahl an Richtlinien, Konsortien, Projekten und Förderprogrammen, die eine Erhöhung der Energieeffizienz bezwecken. Mit dem *European Code of Conduct on Data Centre Energy Efficiency* werden in einem Konsortium namhafter Firmen wie HP, Intel und Microsoft Best-Practice-Lösungen für den effizienten Rechenzentrumsbetrieb erarbeitet. Ein anderes Konsortium, das *Green Touch*, hat sich zur Aufgabe gemacht, die Energieeffizienz in der Netzwerktechnik um das Tausendfache zu verbessern. Einen zusätzlichen Anreiz bieten internationale Wettbewerbe wie die *DCS Awards*, *DatacenterDynamics Awards*, *DataCloud Awards* und auch der nationale *Deutsche Rechenzentrumspreis*, die in unterschiedlichen Kategorien rund um den Rechenzentrumsbetrieb Lösungen auszeichnen. Mit dem *Green-IT Best Practice Award* werden neben dem Einsatz energieeffizienter IT auch die Anwendung von IT-Systemen zur Optimierung von Prozessen, wie z. B. dem Verkehrsnetz, gewürdigt. Inzwischen ist eine hohe Akzeptanz Energieeffizienz steigender Maßnahmen zu beobachten.¹⁰ Einer Studie des Deutschen Forums für Hochverfügbarkeit nach sind lediglich zwei Prozent der mittelständischen Unternehmen in Deutschland nicht bereit, für „grüne“ IT mehr Geld auszugeben.¹¹

Ursache des hohen Strombedarfs

Die Erhöhung der Energieeffizienz und die Fokussierung auf die Verwendung regenerativer Energien adressieren jedoch nur den vernünftigen Umgang mit den Symptomen. Aber was sind die Ursachen? Weshalb ist der Strombedarf von Rechenzentren in den letzten Jahren so drastisch angestiegen und weshalb ist auch kein Abklingen dieses Trends zu erwarten? Die Antwort hierauf liegt in der exponentiell ansteigenden Flut an zu verarbeitenden Daten. Ein Indikator stellt hierbei das Internet dar. Laut BITKOM ist weltweit jeder fünfte Mensch online.¹² Internetknoten wie der AMS-IX und der DE-CIX sind gezwungen, ihre Kapazitäten kontinuierlich aufzustocken. Insbesondere TV- und Videoinhalte in HD-Auflösung tragen zu diesem Anstieg bei, wie in Abbildung 1.4 für den Zeitraum von 2006 bis 2012 dargestellt ist. Allein das Video-Portal Youtube hat im Jahre 2007 einen doppelt so großen Datenverkehr verursacht wie das gesamte Internet noch zwei Jahre zuvor.¹³ In den Jahren 2013 bis 2018 soll der globale IP-Datenverkehr laut Cisco um das Dreifache weiter ansteigen [65].

Diese Entwicklung führt zu einer erhöhten Belastung der Telekommunikationsnetze, weshalb ein Ausbau mit Glasfaserkabel – mit einer einzelnen Glasfaser lassen sich Geschwindigkeiten von theoretisch bis zu 14 Tbit/s erreichen [119] – forciert wird. Die Erhöhung des Datenauf-

¹⁰<http://www.cowo.de/a/1930305>, letzter Zugriff: 25.11.2014

¹¹<http://www.cowo.de/a/2368858>, letzter Zugriff: 25.11.2014

¹²http://www.bitkom.org/de/presse/49919_46069.aspx, letzter Zugriff: 25.11.2014

¹³<http://sz.de/1.347967>, letzter Zugriff: 25.11.2014

1 Einleitung

kommens hat aber auch Auswirkungen auf die für die Verarbeitung der Anfragen benötigten Server. Die Art der Verarbeitung hat sich ebenfalls gewandelt. So bestanden frühere Internet-Sessions daraus, zwischen statischen Webseiten zu surfen, wohingegen heutzutage aufwendig gestaltete multimediale, dynamisch generierte und oftmals personalisierte Webseiten betrachtet werden. Insgesamt konnte in Rechenzentren eine Erhöhung des Strombedarfs von jährlich 12 % beobachtet werden [205].

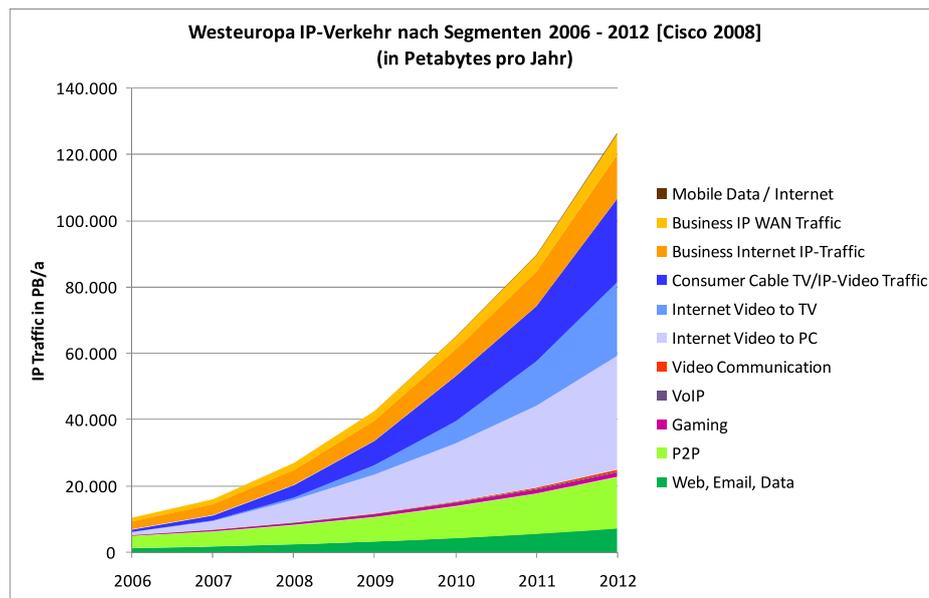


Abbildung 1.4: Entwicklung des IP-Verkehrs in Westeuropa

1.1.1 Maßnahmen zur Energie- und Kosteneinsparung

Bei der Entwicklung von Hardware ist die Leistungsaufnahme neben der reinen Rechenleistung zu einem weiteren primären Designziel geworden und sie sollte auf jedem Systemlevel betrachtet werden [47]. In gleichem Zuge muss allerdings auch darauf geachtet werden, dass beide Designziele geeignet durch Benchmarks und Metriken erfasst und damit vergleichbar werden [147].

Als Einsparmaßnahme bei Servern wurde die allgemeine Steigerung der Hardwareeffizienz bereits erwähnt, welche insbesondere durch die Miniaturisierung der Transistorgrößen und der damit verbundenen Verringerung der Versorgungsspannung erzielt wird. Aber auch durch Managementfunktionen wie das *Dynamic Voltage and Frequency Scaling* (DVFS) und die Multi-Core-Technologie, die den Energiebedarf leistungsgerechter ansteuern, lässt sich der Energiebedarfsanstieg eindämmen.

Neben der IT benötigt insbesondere die Kühlung Energie, denn – abgesehen von Lichtemittierung – wird die komplette Energie, die durch die IT verwendet wird, in Wärme umgewandelt und muss daher abgeführt werden. Techniken wie die Kalt- oder Warmgangeinhausung [22], gezielte Lüftersteuerung und die Nutzung von Freiluftkühlung etc. können den

Energiebedarfsanteil laut Gartner von aktuell 35 % - 50 % auf bis zu 15 % reduzieren.¹⁴

Ein Forschungsschwerpunkt zur Erzielung von Energieeffizienz besteht darin, einen energiebedarfsadaptiven Betrieb bzgl. der Rechenlast zu ermöglichen, d.h., dass bei hoher Auslastung zwar viel Energie, bei geringerer Auslastung allerdings auch entsprechend weniger Energie benötigt wird. Im besten Fall sollte bei keiner Auslastung auch keine Energie benötigt werden. Neuere USV-Module (unterbrechungsfreie Stromversorgung), z. B. mit dem *Energy-Saver-System* von Eaton mit einem maximalen Wirkungsgrad von 99 %, und auch neue Kühlungssysteme ermöglichen eine sehr gute Lastadaptivität.

Server lassen sich nicht so gut anpassen. Zwar kann durch DVFS der Energiebedarf geregelt werden, allerdings ist im Idle-Zustand der Energiebedarf immer noch bei 50 % - 66 % des max. Energiebedarfs (bei maximaler Auslastung) [19, 129, 55]. Neuere Server haben i.d.R. einen geringeren Energiebedarf im Idle-Zustand und weisen einen proportionaleren Energiebedarf bzgl. der Last auf, was über die EP-Metrik (Energy Proportionality) [194] ausgedrückt werden kann. Die Durchschnittsauslastung von Servern liegt nach McKinsey & Company bei ca. 10 %.¹⁵ Laut Equinix ist diese bei Mainframes zwar höher [204], jedoch liege die Zukunft in der Nutzung von Standardkomponenten, wie es bspw. Google und Amazon vorgemacht hätten,¹⁶ da diese kostengünstiger sind und eine leichtere Administration und Wartbarkeit aufweisen. Die geringe Auslastung ist dadurch begründet, dass selbst heutzutage noch viele Server auf Lastspitzen ausgelegt sind, d.h., dass die Server so viel Leistung bereitstellen müssen, dass die darauf laufenden Dienste auch bei Maximallast betrieben werden können. Sie sind somit die meiste Zeit überdimensioniert. Ein vergleichsweise junger Ansatz, dem beizukommen, besteht darin, in Zeiten der Ruhe (Idle) den Server in einen Schlafzustand zu versetzen. Dieser Vorgang soll sich im Millisekundenbereich realisieren lassen.¹⁷

Mit dem GreenIT RZ-Benchmarking¹⁸ wurde eine Möglichkeit geschaffen, den Status Quo eines Rechenzentrums bzgl. seiner Energieeffizienz zu bestimmen. Für die Energieeffizienz von Netzwerk- und Telekommunikationsgeräten hat die Energy Consumption Rating (ECR) Initiative eine gleichlautende ECR-Metric entwickelt. Server können anhand des SPECpower_ssj2008-, TPC-Energy- oder des Server-Efficiency-Rating-Tool-Benchmark (SERT), das Storage anhand des SNIA Emerald Power Efficiency Measurement Specification bzgl. der Effizienz bewertet werden.

In Leitfäden und anhand von Best-Practice-Beispielen (z. B. vom BITKOM [104] und BMU [85]) werden zusammenfassend Maßnahmen zur Verbesserung der Betriebseffizienz aufgezeigt und auch Hinweise dazu gegeben, wie diese am besten umgesetzt werden könnten. Setzt ein Rechenzentrum Energieeffizienz steigernde Maßnahmen ein, kann es sich dieses z. B. durch den Energy Star, den Blauen Engel oder den TÜV Rheinland zertifizieren lassen.

¹⁴<http://www.gartner.com/newsroom/id/799812>, letzter Zugriff: 25.11.2014

¹⁵<http://nyti.ms/18nZqnG>, letzter Zugriff: 25.11.2014

¹⁶<http://www.cowo.de/a/1927112>, letzter Zugriff: 25.11.2014

¹⁷<http://www.cowo.de/a/1889314>, letzter Zugriff: 25.11.2014

¹⁸<http://www.greenit-bb.de/projekte/rz-benchmarking>, letzter Zugriff: 16.07.2015

Virtualisierung

Ein erprobter und mittlerweile auch bewährter Ansatz zur Energiebedarfsreduzierung besteht in der Konsolidierung von Servern durch den Einsatz von Virtualisierungstechniken. Zusammen mit Cloud Computing seien hierdurch laut der Experton Group die größten Einspareffekte bei der Ressourcenoptimierung möglich.¹⁹

In klassischen Rechenzentren wurde jeder Dienst zur Isolation der Daten und Berechnung auf einem dedizierten Hardware-Server ausgeführt. Aufgrund der oben erwähnten Lastspitzendimensionierung der Server wird hierdurch viel Strom verschwendet. Durch die Virtualisierungstechnologie ist der Betrieb mehrerer Dienste auf einem Hardware-Server möglich, ohne das Isolationskriterium zu verletzen. Dazu wird eine sogenannte virtuelle Maschine (VM) gebildet, die die Hardware in einer Software-Umgebung abstrahiert und einen virtuellen Server für die darin laufende Applikation bzw. den darin betriebenen Dienst darstellt. Die VM kann also als ein dedizierter Software-Server angesehen werden. Hierin betriebene Dienste werden in ihrer Ausführbarkeit i.d.R. nicht durch andere VMs beeinflusst.

Auslastungsprofile von Diensten haben oftmals einen periodischen Verlauf [31]. Dieses Wissen kann dazu genutzt werden, um Dienste, deren Lastspitzen an unterschiedlichen Zeitpunkten auftreten, gemeinsam auf einem Server zu betreiben, wodurch eine hohe aggregierte durchschnittliche Auslastung geschaffen wird. Durch die Konsolidierung mehrerer Dienste auf weniger Server können überschüssige Server eingespart werden, was – kurzfristig – den Energiebedarf senkt und – mittelfristig – benötigte Ressourcen einspart. Beide Effekte führen zu einer geringeren Belastung des IT-Budgets.

Neben der direkten Energie- und Kosteneinsparung bietet die Virtualisierung einen weiteren Vorteil: Flexibilität. Eine VM stellt für den betriebenen Dienst eine Hardwareabstraktionsschicht dar, die unabhängig davon, auf welchem Hardware-Server sie betrieben wird, gleich bleibt. In Kombination mit der Live-Migrationstechnik [66] ist es möglich, die VM zur Laufzeit zwischen unterschiedlichen Servern zu verschieben, und dies mit einer Ausfallzeit von weniger als 100 ms. Voraussetzung hierzu ist ein zentral zugreifbares Stagesystem wie ein SAN (Storage Area Network). Auf diese Weise können Anpassungen an der Serverausstattung vorgenommen werden, ohne dass der Betrieb unterbrochen werden muss.

Die oben genannte Konsolidierung konnte bereits die durchschnittliche Auslastung von Servern auf über 50 % erhöhen [106]. Durch Live-Migration kann die Konsolidierungsrate auf dynamische Weise nochmals verbessert werden. Untersuchungen haben gezeigt, dass insgesamt 40 % - 80 % Energie gegenüber dem dedizierten Betrieb von Diensten auf Hardware-Servern eingespart werden kann [231, 145, 131].

Auch für das Storage lassen sich Virtualisierungstechniken anwenden. Mit dem Thin-Provisioning lassen sich Speichereinheiten entsprechend der derzeit tatsächlich benötigten Menge installieren und dynamisch aufstocken. Dies hebt nicht nur die Flexibilität an, sondern spart gegenüber dem jetzigen Modell, bei dem bereits vorab für zukünftige Anforderungen

¹⁹<http://www.experton-group.de/nc/home/monthly-news-dach/news/article/top-10-einsparmoeglichkeiten-fuer-das-jahr-2012.html>, letzter Zugriff: 25.11.2014

aufgestockt wird, Kosten ein. Noch mehr Flexibilität wird durch den Bezug von Cloud Storage erzielt. Hierdurch kann die Storage-Auslastung von derzeit 30 % - 40 % [236] auf theoretisch 100 % gesteigert werden.

Cloud Computing

Eine weitere Ursache, die zur Bildung von Groß- und Megarechenzentren geführt hat, liegt in der Bereitstellungsweise von Rechenkapazitäten begründet. Hier ist ein deutlicher Trend zum Outsourcing oder allgemein dem Bezug von Dienstleistungen über das Internet zu erkennen. Neben dem Bezug von Diensten sei auch eine Veränderung der internen IT-Infrastruktur in Unternehmen zu erkennen, so eine Studie des Unternehmensberaters Centracon.²⁰ Zentralisierung ist hier die Devise, d.h., dass die stromfressenden *Fat Clients* (Arbeitsplatz-PCs) durch energiesparende *Thin Clients* (Mini PCs, Notebooks) ersetzt werden. Die Rechenpower wird dabei verlagert von den verteilten Büros in das zentrale Rechenzentrum. Dass Unternehmen ihre Rechenzentren aufstocken,²¹ ist die Konsequenz dieser Entwicklung und führt schließlich zum Bau von Cloud-Rechenzentren [242].

Die Vorteile von Cloud Computing sind augenscheinlich: mehr Flexibilität [168] bei gleichzeitig geringeren Kosten und das Abrechnungsmodell richtet sich nach dem tatsächlichen Ressourcenverbrauch. Insbesondere kleine und mittlere Unternehmen (KMU) können so auf hohe Investitionskosten verzichten und Dienste sowie Infrastruktur nach Bedarf beziehen [209, 80]. Aber auch größere Unternehmen können laut einer Studie von Microsoft Einparungen beim Strombedarf von mind. 30 % erzielen.²² Durch den Betrieb föderierter Cloud-Rechenzentren kann die Erhöhung der Verfügbarkeit im Vergleich zur klassischen Architektur (siehe [2]) systembedingt auf kostengünstige Weise erreicht werden, was auch dem Anwendungsfall *Disaster Recovery* zugute kommt [11].

Doch es existieren auch Vorbehalte gegenüber dieser Technik, denn entsprechend der Idee des Cloud Computings ist der Ort der Arbeit unerheblich, einzig ihre Vollbringung steht im Mittelpunkt. Hierbei obliegt es dem Dienstleister, den Ort auch zur Laufzeit zu wechseln. Es werden also sowohl Daten als auch Prozesse in der sogenannten *Public Cloud* aus der Hand des eigenen Unternehmens gegeben. Fehlende Kontrollmöglichkeiten und Sicherheitsbedenken führen zur Zurückhaltung. Mit den Standards *ISO/IEC 27001* und *SAS 70* ist zwar eine Bewertung der Sicherheitsmechanismen und Kontrollfunktionen möglich, doch lösen diese Spezifikationen das Problem nicht. Tatsächlich sind bei den großen Cloud-Providern bereits Ausfälle vorgekommen, doch das Fraunhofer SIT kommt in seiner Studie zur Sicherheit von Cloud Computing zu einem anderen Schluss, denn hiernach würden KMU „ihre Sicherheit durch den Einsatz von Cloud-Services erhöhen“.²³ Das Cloud Computing scheint aber sei-

²⁰<http://www.lanline.de/fachartikel/centracon-sagt-tschüss-fat-client.html>, letzter Zugriff: 26.11.2014

²¹<http://www.cowo.de/a/1906253>, letzter Zugriff: 26.11.2014

²²<http://www.microsoft.com/de-de/news/pressemitteilung.aspx?id=533264>, letzter Zugriff: 26.11.2014

²³<http://www.fraunhofer.de/de/presse/presseinformationen/2009/09/cloud-computing-sicherheit.html>, letzter Zugriff: 26.11.2014

1 Einleitung

nen Wendepunkt erreicht zu haben: Der stetig steigende Bezug von Cloud-Services (siehe Abbildung 1.5), insbesondere von Cloud Storage, zeigt seine Akzeptanz.

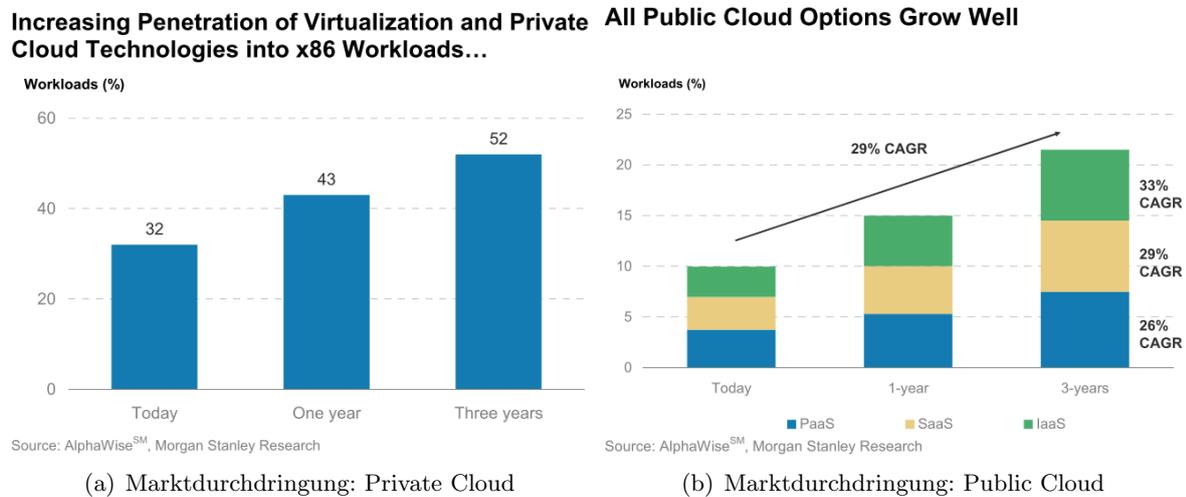


Abbildung 1.5: Steigender Einsatz von Cloud Computing prognostiziert [106], Basisjahr: 2011

Eine Basistechnologie für Cloud Computing ist die Virtualisierung [139]. Diese wird mittlerweile auch für geschäftskritische Anwendungen und Aufgaben eingesetzt²⁴. In Abbildung 1.5(a) ist dargestellt, wie verbreitet die Virtualisierungstechnologie und damit einhergehend *Private Cloud* ist. Aber auch *Public Cloud* erfreut sich immer mehr an Beliebtheit (siehe Abbildung 1.5(b)). (Auf die Ausprägungen von Cloud Computing wird in Abschnitt 2.1.4 eingegangen.) In den Augen des Analystenhauses techconsult²⁵ gilt das Cloud Computing als IT-Wachstumsmarkt der Zukunft. Intel spricht in seiner „Cloud 2015“-Vision von einer „Welt von verbundenen, miteinander kompatiblen Clouds“, die „Anwendungen maßgeschneidert für jeden Client“ ausführen.²⁶

1.1.2 Globales Last- und Powermanagement

Die Bildung von verteilten Cloud Services, welche also über unterschiedliche Standorte verteilt betrieben werden, bedeutet laut Google und Microsoft [20, 62] eine implizite Erhöhung der Redundanz und damit auch eine erhöhte Verfügbarkeit der Daten sowie auch eine erhöhte Ausfallsicherheit der Dienste. Es sei kein großes komplexes Gebilde mehr zu koordinieren und dessen Betrieb sicherzustellen, sondern viele kleine Gebilde, deren vereinzelter Ausfall sich nur auf einen kleinen Anteil bezieht. Die übrigen lauffähigen Standorte könnten dann die Aufgaben flexibel übernehmen.

Beim Cloud Computing, insbesondere bei Hybrid-Cloud- bzw. Inter-Cloud-Architekturen [209, 80, 11] ist eine flexible Infrastruktur gebildet, die es ermöglicht, Lasten in Form von VMs

²⁴<http://www.datacenter-insider.de/index.cfm?pid=8488&pk=415486>, letzter Zugriff: 26.11.2014

²⁵Verena Bunk, Analyst bei techconsult, ITK-Marktprognose 2011/12

²⁶<http://www.cowo.de/a/2356460>, letzter Zugriff: 26.11.2014

zwischen Rechenzentren unterschiedlicher Standorte zu migrieren. Hierdurch wird eine neue, höhere Systemebene geschaffen, die nach dem Khazzoom-Brookes Postulat²⁷ bei Energieeffizienzbetrachtungen wichtiger ist als die bisher betrachteten Systemebenen. Cloud Computing ist hierfür eine Art Vorreiter, da bereits jetzt Praktiken zum Ausgliedern von Unternehmensanwendungen eingesetzt und auch Fragestellungen zum Bezahlssystem, zur Sicherheit sowie zum Datenschutz behandelt werden.

Durch die Lastverschiebungen können der Energiebedarf sowie auch die Betriebskosten eines verteilten Rechenzentrumsverbundes minimiert werden. Dies ist dadurch motiviert, dass die Parameter, welche den Energiebedarf und die Betriebskosten beeinflussen, sich von Standort zu Standort unterscheiden und sich auch mit der Zeit ändern können.

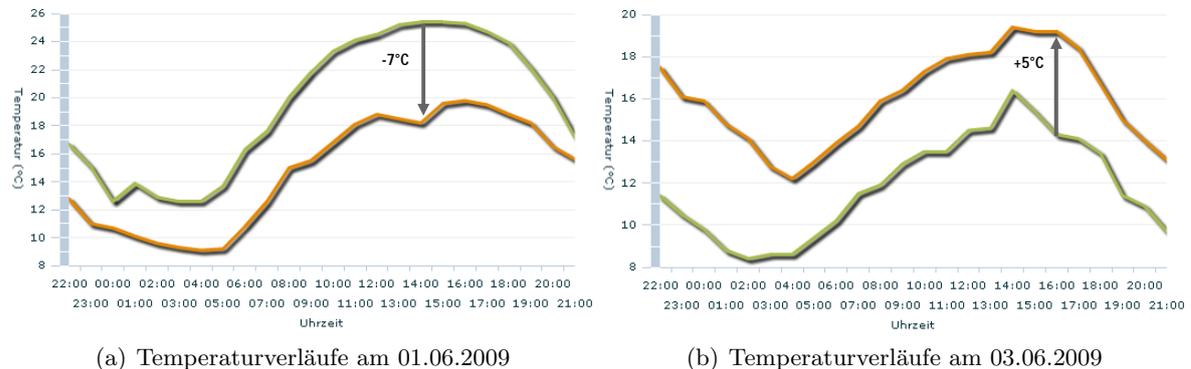


Abbildung 1.6: Temperaturverläufe von Stuttgart (grün) und Hannover (orange) für unterschiedliche Tage

Als erster Standortparameter sei hier die Effizienz des Rechenzentrums bzgl. der Rechenleistung pro Watt genannt. Diese ergibt sich nicht nur durch die verwendete IT, sondern auch durch die eingesetzte Infrastruktur. Einen großen Einfluss hat hier das Kühlungssystem, bei dem auch sogleich der zweite Parameter mit hinein spielt: die Außentemperatur. Je wärmer die Umgebungsverhältnisse sind, desto stärker müssen aktive Kühlkomponenten betrieben werden. Ein Beispiel für einen Temperaturumschwung ist in Abbildung 1.6 dargestellt. Über den Zeitraum von drei Tagen haben sich die Temperaturverhältnisse umgedreht, d.h., dass anfangs in Stuttgart, später aber in Hannover eine höhere Kühlleistung zu erbringen wäre. Einen weiteren Klimafaktor stellt die Luftfeuchtigkeit dar, denn bei trockenen Verhältnissen können Verdunstungstürme eingesetzt werden, welche Kühlleistung durch die durch Verdunstung abgegebene Energie erbringen.

Der nächste Parameter hat keinen Einfluss auf den Energiebedarf, sondern beeinflusst die Kosten direkt: der örtliche Strompreis. In aktuellen Vertriebsmodellen werden einheitliche Preise für die Nutzung von Strom erhoben, je nach vereinbartem Volumentarif. Werden jedoch die Preise an den Strombörsen betrachtet, so ergeben sich unterschiedliche Strompreise, die sich nicht nur über die Zeit verändern (siehe Abbildung 1.7), sondern sich auch zwischen den

²⁷ „Energy efficiency improvements that, on the broadest considerations, are economically justified at the micro level, lead to higher levels of energy consumption at macro level.“[198]

1 Einleitung

Strombörsen unterscheiden (siehe Abbildung 1.8). Dies liegt u.a. in den unterschiedlichen Stromgenerierungsmethoden begründet, die mitunter sehr volatil ausfallen können – Sonnen- und Windkraftwerke und hierzu adaptierend Gas- und Kohlekraftwerke. Auch der Strombedarf ist sehr fluktuativ, wodurch sich ein ständig verändertes Angebot-Zu-Nachfrageverhältnis ergibt.

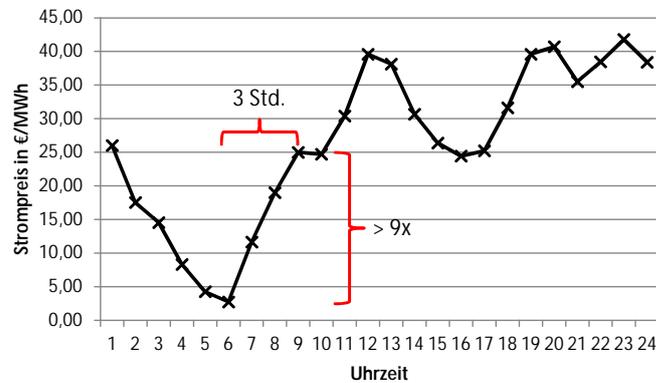


Abbildung 1.7: Strompreis für Stundenkontrakte, Spotmarkt EXAA 25.07.2010

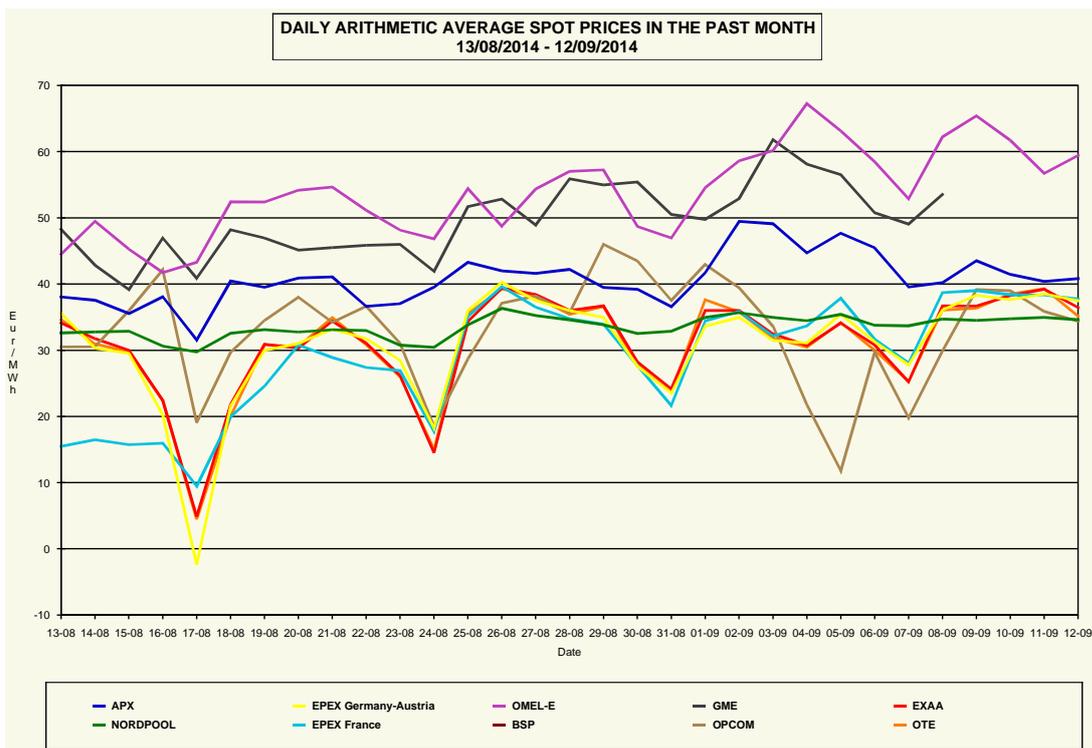


Abbildung 1.8: Tagesdurchschnittlicher Strompreis unterschiedlicher europäischer Strombörsen.²⁸

Werden die gehandelten Preise an die Endabnehmer weitergegeben, können diese ihr Abnahmeverhalten zur Senkung ihrer Stromrechnung anpassen [7]. Außerdem wird das Stromnetz

²⁸http://www.omel.es/informes_mercado/europex/MARKET_RESULTS.PDF, letzter Zugriff: 13.09.2014.

entlastet, denn die sich über den Tag ändernden Preise führen zu einer geringeren Spitzen- und Durchschnittslast sowie zu geringerer Lastvolatilität [32, 127]. Allgemein werden aber auch sowohl Investitionen in technologische Innovationen gefördert als auch Monopolstellungen vermindert oder gar verhindert, wie es bereits auf dem U.S.-amerikanischen Markt seit der Einführung dynamischer Elektrizitätspreise in den 90er Jahren zu beobachten ist [203].

Lasten lassen sich jedoch nicht nur örtlich verschieben. Bei einigen sogenannten Batch-Jobs ist die zeitliche Ausführung weniger kritisch, so dass sich diese zeitlich verschieben lassen [91, 6]. Es ist somit möglich, Lasten zeitlich wie örtlich entsprechend den energie- und kostenbeeinflussenden Faktoren zu verschieben. Voraussetzung hierfür ist der Betrieb (verteilter) flexibler Cloud Infrastrukturen, wie sie z. B. von Google, Amazon und Microsoft betrieben werden. Aber auch kleinere Unternehmen, die nur einen oder wenige Standorte betreiben, können hiervon profitieren, indem sie einen Verbund bilden. Neben einer kooperativen Variante ist auch eine Vermarktungsvariante möglich, bei der freie Kapazitäten angeboten werden, wie bspw. bei ZIMORY ECO. Für den Endanwender können so Preisvergünstigungen geschaffen werden.

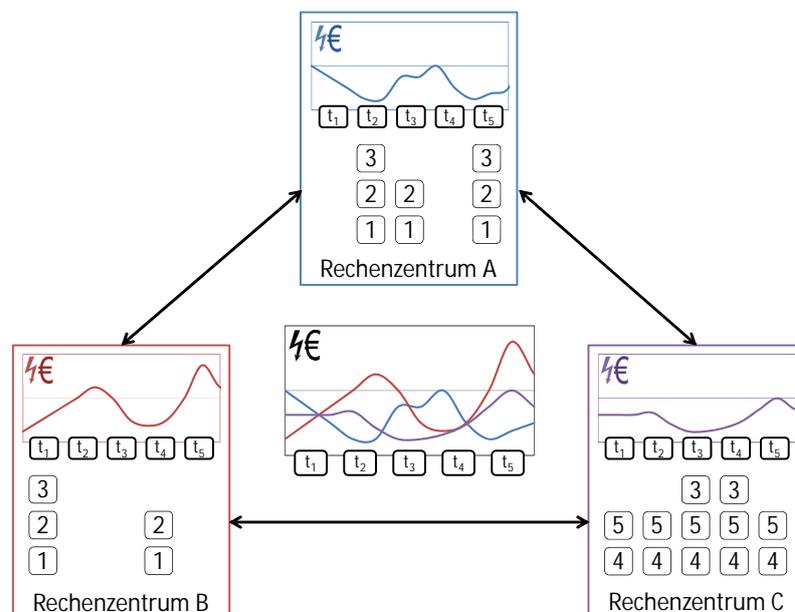


Abbildung 1.9: Beispiel für ein globales LPM: An jedem Zeitpunkt t_i werden die VMs (1 bis 5) zur Minimierung der Gesamtbetriebskosten zu den günstigsten Rechenzentren migriert.

In Abbildung 1.9 ist ein kleines Beispiel für ein globales LPM skizziert. Als Optimierungsziel werden hier die Kosten der beteiligten Rechenzentren minimiert. Hierzu muss der Betrieb von fünf VMs in den drei Rechenzentren, deren Kapazität auf maximal drei VMs beschränkt ist, geeignet geplant werden. Zu jedem Zeitpunkt t_i werden die VMs so alloziert, dass die Gesamtkosten minimal werden. Anhand der standortbedingten Strompreisverläufe lässt sich bestimmen, in welchen Rechenzentren der Betrieb aktuell am günstigsten ist. Beim Zeitschritt t_1 zu t_2 sollten z. B. die VMs 1, 2 und 3 von Rechenzentrum B nach A migriert werden, da dann

1 Einleitung

Rechenzentrum A am günstigsten und Rechenzentrum B am teuersten ist. In einem realistischen Szenario sind weit mehr Aspekte zu beachten. Hierzu gehören u.a., dass Dienste keine – wie hier suggeriert – konstante Last haben, es unterschiedliche Ressourcen mit individuellen Kapazitäten gibt, der Energiebedarf eines Rechenzentrums von der konkreten IT-Hardware und Infrastruktur abhängig ist, die klimatischen Bedingungen auf die Betriebseffizienz einwirken und bei einer Migration Netzwerkkosten entstehen.

In einer eigenen, durch das BMWi in Auftrag gegebenen Studie [162] wurde das Potential eines rechenzentrumsübergreifenden LPM zur Reduktion des Energiebedarfs ermittelt. Es konnten Kosten- und Energieeinsparungen von bis zu 40 % festgestellt werden.

Die flexible Infrastruktur kommt jedoch nicht nur der Kosteneinsparung zugute. Wird eine globale Infrastruktur aufgespannt, können neben den Lasten auch die Arbeitszeiten verschoben werden: Mitarbeiter aus der entsprechenden Zeitzone können auf die geografisch nachgeführten Dienste mit niedriger Latenz zugreifen. Eine Untersuchung von 2007 hat ergeben, dass hierdurch die Qualität einer Projektarbeit gesteigert werden konnte [96].

Partizipation am Smart Grid

Rechenzentren werden immer mehr dynamisch ihren Energiebedarf an den momentanen Rechenlastbedarf und den Standortbedingungen anpassen. Eine ähnlich gelagerte Dynamisierung ist im Energieversorgungsnetz zu beobachten, wo eine Veränderung weg von der Versorgungsrichtung von Kraftwerken in den Höchst- und Hochspannungsnetzen zu den Verbrauchern in den Mittel- und Niederspannungsnetzen hin zu einer ambienten Versorgung stattfindet, d.h., dass Energieeinspeiser nun auf jeder Netzebene anzutreffen sind. Aufgrund unzureichender Speichermöglichkeiten und um Transformationsverluste so gering wie möglich zu halten, müssen hierfür intelligente dynamische Systeme entwickelt werden, die die unterschiedlichen Energieeinspeiser mit -verbrauchern koordinieren. Insbesondere die hohe Volatilität der regenerativen Energien wie die Windkraft stellen eine große Herausforderung dar. Bis 2020 soll nach den Meseberger Beschlüssen von 2007 der Anteil regenerativer Stromerzeugung von ca. 15 % auf 25 % - 30 % gesteigert werden. Dies hat nach einem Sondergutachten über die Gesamtkosten einer regenerativen Vollversorgung im Jahre 2050 [195] auch ökonomische Vorteile. So soll der Preis je kWh um 7 Cent günstiger als bei CO_2 emittierenden Energieträgern sein, da steigende Brennstoffkosten und die Kosten für Emissionszertifikate vermieden werden. Allerdings geht dies mit einer noch höheren Volatilität einher.

Bei diesem intelligenten als Smart Grid²⁹ bezeichneten Netz können Lastanpassungsvorgänge für ein Demand Side Management³⁰ (DSM), nämlich *Load Reduction*, *Peak Capping* und

²⁹Begriffsbestimmung nach NIST: Ein Smart Grid ist ein modernisiertes, elektrisches Netz, das seinen eigenen Zustand kennt, sich schützt und automatisiert das Zusammenspiel der einzelnen Elemente koordiniert. Dies bezieht sich auf alle Elemente und alle Spannungsebenen – von der Höchstspannung zur Niederspannung, alle Erzeuger und Verbraucheranlagen und die Netzkomponenten.

³⁰Laut VDE [12] umfasst Demand Side Management die direkte Beeinflussung des Energiebedarfs von Verbrauchern, wohingegen beim Demand (Side) Response die Beeinflussung indirekt über ein Anreizsignal (z. B. unterschiedliche Strompreise zu unterschiedlichen Zeiten) erfolgt. Im Folgenden werden jedoch beide Aspekte unter dem Begriff Demand Side Management verstanden.

Valley Filling (siehe Abbildung 1.10), transparenter und effizienter umgesetzt werden. Durch das zeitliche Verschieben von Lasten wird ein *Load Shifting* durchgeführt.

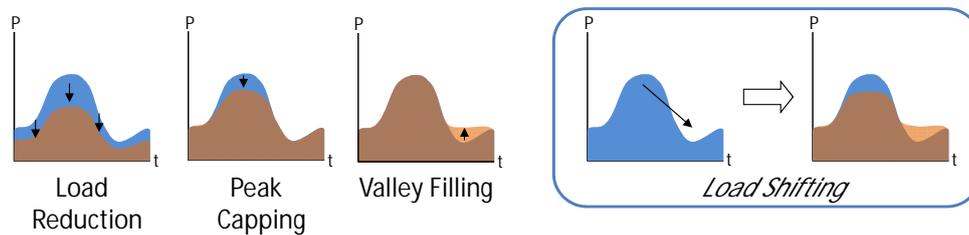


Abbildung 1.10: Lastanpassungsvorgänge

Rechenzentren als Energiegroßabnehmer können zur Lastregelung einen großen Beitrag leisten, da sie in ihrer Leistungsaufnahme dynamisch steuerbar sind [138]. Sie stellen damit ein virtuelles Kraftwerk dar, das als aktiver Teilnehmer im Smart Grid partizipieren kann [199].

1.2 Ziele der Dissertation

Im Rahmen der Dissertation wird die Technologie der Virtualisierung samt Live-Migration als Basistechnologien genutzt, um Lasten zur Minimierung des Energiebedarfs bzw. der Kosten gezielt über unterschiedliche Rechenzentrumsstandorte zu verteilen. Hierbei werden solche Dienste bzw. Prozesse, die wie in [193] definiert als einzelne Applikation eine kontinuierliche und zeitvariable Auslastung besitzen, betrachtet, welche auf einzelnen Servern betrieben werden können. Diese werden im folgenden als *Server-Sized-Services* bezeichnet, um sich begrifflich von *Large-Scale-Services* abzugrenzen, die zum Betrieb mehrere Server in einem (verteilten) Cluster benötigen.

Dienste, welche in Interaktion zu Nutzern stehen, können durch eine Migration in andere Rechenzentren negativ bzgl. der Zugriffslatenz beeinflusst werden, wodurch die Dienstegüte sinkt. Von diesem Aspekt wird in der Dissertation abstrahiert, da in erster Linie das mögliche Einsparpotential ermittelt werden soll.

Zur gezielten Lastverteilung werden die Rechenzentrumskomponenten einbezogen, also die IT und Infrastruktur sowie das bisher in Energiebedarfsmodellen vernachlässigte lokal eingesetzte LPM. Weiterhin müssen die Zusatzkosten bei der standortübergreifenden Migration, welche im Netzwerk entstehen, berücksichtigt werden. Als letzten, aber auch sehr wichtigen Punkt sind Replikationskosten einzubeziehen, da diese die Einsparpotentiale der optimierenden Lastverlagerung egalalisieren oder sogar überkompensieren können. Das zu entwickelnde Verfahren soll damit als Planungswerkzeug zur energie- und kostenorientierten Standort- und Kapazitätsplanung sowie als globales LPM im operativen Betrieb eingesetzt werden können.

Neben den energiebedarfs- und kostenminimierenden Ansätzen soll ebenso die Idee der Partizipation am DSM untersucht werden, bei dem ein gezielter Energiebedarf angestrebt wird. Alle drei Optimierungsansätze sollen im Verfahren berücksichtigt werden können, und aufgrund seiner Anwendung im Cloudbereich ist auch eine gute Skalierbarkeit zu empfehlen.

1 Einleitung

In das betrachtete System fließen unterschiedliche Parameter ein, die Auswirkungen auf ein mögliches Einsparpotential haben. Eine Sensibilitätsanalyse soll die einzelnen Einflüsse der Parameter kenntlich machen. Außerdem sollen sich die Evaluationsergebnisse nicht auf einzelne Szenarien beschränken, sondern möglichst allgemeingültig sein.

Handelt es sich bei den Rechenzentrumsbetreibern um Cloud-Provider, die keine eigenen Dienste betreiben, sondern den Betrieb als Serviceleistung anbieten, so ist es sowohl für den Betreiber als auch für den Endkunden wesentlich zu wissen, welche Kosten ein Dienst bzw. eine VM im optimierten Verbund verursacht, denn hierdurch kann ein transparenteres Kostenmodell aufgestellt und damit Endkunden potentiell günstigere Preise angeboten werden.

Zusammenfassend ergeben sich die folgenden wissenschaftlichen Fragestellungen:

- Unter welchen Gegebenheiten (Parameterbedingungen) kann ein globales LPM Einspareffekte bzgl. des Energiebedarfs und der Kosten bewirken und in welchen Größenordnungen belaufen sich diese?
- Welchen Einfluss hat ein im Rechenzentrum eingesetztes LPM auf den Energiebedarf?
- Inwiefern können die Kosten dienstespezifisch erhoben und die möglichen Einspareffekte bewertet werden?
- Welchen konkreten Beitrag können Rechenzentren beim DSM leisten?

Zur Beantwortung dieser Fragestellungen werden ein dem oben aufgeführten System entsprechendes Energiebedarfsmodell und ein Verfahren zum globalen LPM entwickelt sowie anhand weitreichender Simulationen verschiedenartiger Szenarien Evaluationsergebnisse generiert. Diese werden bzgl. der Fragestellungen analysiert und diskutiert.

1.3 Aufbau der Arbeit

Die Arbeit fängt im Folgenden mit der Aufbereitung des zum Themengebiet gehörenden Stand der Forschung und Technik an. Hierzu wird als erstes der technologische Kontext und im Anschluss verwandte Arbeiten näher durchleuchtet sowie eine inhaltliche Abgrenzung gegeben. Im darauffolgenden Kapitel werden das als Basis verwendete rechenzentrumsinterne LPM und weitere technische Voraussetzungen beschrieben. In Kapitel 4 werden die Hauptbeiträge der Dissertation erläutert. Das hierin vorgestellte Konzept für ein globales LPM ist in Unterkomponenten aufgeteilt. Diese werden einzeln bzgl. der jeweiligen Problemstellung vorgestellt und nach einem spezifischen Abschnitt zu den verwandten Arbeiten eine Lösung präsentiert. Diese wird in dem jeweiligen Unterabschnitt auch bereits evaluiert, um darauf aufbauend die weiteren Systemkomponenten abzustimmen. Das Gesamtkonzept wird im anschließenden Kapitel evaluiert und im Hinblick auf die formulierten Dissertationsziele analysiert. Zum Schluss werden ein Fazit zur Dissertation und zu den ermittelten Ergebnissen sowie ein Ausblick auf weiterführende Arbeiten gegeben.

2 Stand der Forschung und Technik

Nachdem die Ziele der Dissertation im Einleitungsteil herausgestellt wurden, wird in diesem Kapitel detaillierter auf den technologischen Kontext eingegangen, welcher den Bereich verwandter Arbeiten erweitert und so eine bessere Einordnung ermöglicht. Im Anschluss wird auf verwandte Arbeiten eingegangen, welche inhaltlich am nächsten zur vorliegenden Arbeit liegen, und eine Abgrenzung zu diesen geschaffen.

2.1 Technologischer Kontext

Bereits in der Einleitung hat sich herausgestellt, dass das Themengebiet der Dissertation weit gefächert ist. Technologisch betrachtet, handelt es sich um die Optimierung des *Rechenzentrums*betriebs, welche durch ein *Verfahren zum Last- und Powermanagement auf Systemebene* realisiert werden soll. Dieses Verfahren verwendet zum Management die Technologien der *Server-Virtualisierung und Live-Migration*, die eine flexible Lastverteilung auch über Standorte hinweg erlauben, was vom Grundsatz her auch bereits beim *Cloud Computing* genutzt wird. Alle vier Themenbereiche werden in der Dissertation tangiert und werden daher zum besseren Verständnis und Einordnung der Dissertation im Folgenden diskutiert.

2.1.1 Rechenzentrum

Das Rechenzentrum stellt in seiner Funktion Rechenleistung zur Verfügung. Die Größe eines Rechenzentrums hängt dabei vom Einsatzzweck ab. Die Grundelemente sind die IT selbst, bestehend aus Servern, dem Stagesystem und dem Netzwerk, sowie die Infrastruktur, die den Betrieb der IT-Komponenten sicherstellt. Bereits kleinere Unternehmen besitzen Rechenzentren, die lediglich mit wenigen Servern bzw. Racks ausgestattet sind. Häufig wird in dieser Größenordnung auch von Serverräumen gesprochen. Daher sei an dieser Stelle der Begriff Rechenzentrum anhand der Definition nach eco [181] näher spezifiziert:

Ein Rechenzentrum ist mindestens ein eigenständiger, baulich getrennter Raum mit einfacher Klimatisierung, einfacher Stromversorgung, einer USV – d.h. einer unterbrechungsfreien Stromversorgung (Qualitätsstrom geglättet, Überspannungsschutz, etc.) – ausgelegt mit fünf Minuten Überbrückungszeit zum Herunterfahren der Betriebssysteme, Einrichtungen zur Branderkennung (Rauchmelder) und zur Brandbekämpfung (Feuerlöscher), minimalem physikalischem Zugangsschutz und stabiler Netzanbindung (ein Provider, eine unabhängige Netzzuführung).

Je größer ein Unternehmen ist, desto größer ist auch i.d.R. der Bedarf an IT-Leistung. Die Energie wird jedoch nicht allein von der eingesetzten IT benötigt, sondern insbesondere auch von den Kühlungs- und Stromversorgungskomponenten.

Beim Betrieb von besonders wichtigen Diensten kommt den beiden letzteren Bereichen eine besondere Rolle zu, da hierbei die Zuverlässigkeit des Betriebs im Vordergrund steht. Diese wird durch redundante Ausführung der Komponenten erreicht. Das Uptime Institute unterteilt Rechenzentren diesbezüglich in vier Klassen (Tier-Klassifikation) [226]:

- Ein einfaches Rechenzentrum der Stufe I besitzt nicht-redundante Gebäudekomponenten und einfache nicht-redundante Verteilernetze, die die Server des Standorts versorgen.
- Ein Rechenzentrum der Stufe II besitzt redundante Gebäudekomponenten und einfache nicht-redundante Verteilernetze, welche die Server des Standorts versorgen.
- Ein Rechenzentrum der Stufe III besitzt redundante Gebäudekomponenten und multiple Verteilernetze, welche die Server des Standorts versorgen. Grundsätzlich versorgt zu jedem Zeitpunkt jeweils nur ein Verteilernetz den Server.
- Ein Rechenzentrum der Stufe IV ist ein fehlertolerantes Rechenzentrum, welches redundante Gebäudekomponenten und multiple Verteilernetze, die simultan die Server des Standortes versorgen, besitzt. Alle Server werden auf zwei Wegen mit Strom versorgt und so installiert, dass sie mit der Beschaffenheit der Standortarchitektur kompatibel sind.

Steigenden Kosten in der Produktion wird in Unternehmen i.d.R. entweder mit erhöhten Produktpreisen oder geringeren Ausgaben in Personal und Technik begegnet. Im Bereich von Rechenzentren meint Andreas Rüter, Leiter der europäischen Rechenzentren von Vodafone: „Energie-Effizienz bedeutet Kosten-Effizienz“.² Zwar sollen laut Gartner nur 12% der Rechenzentrumskosten dem Energiebedarf geschuldet sein, doch ist die Tendenz steigend.³

In Abbildung 2.1 ist dargestellt, wie sich der Energiebedarf in einem durchschnittlichen Rechenzentrum auf die einzelnen Komponenten verteilt. Hierin ist zu erkennen, dass die IT-Ausstattung den größten Bedarf hat, gefolgt von der Klimatisierung und der Stromversorgung. Je nach Tier-Klasse verschieben sich

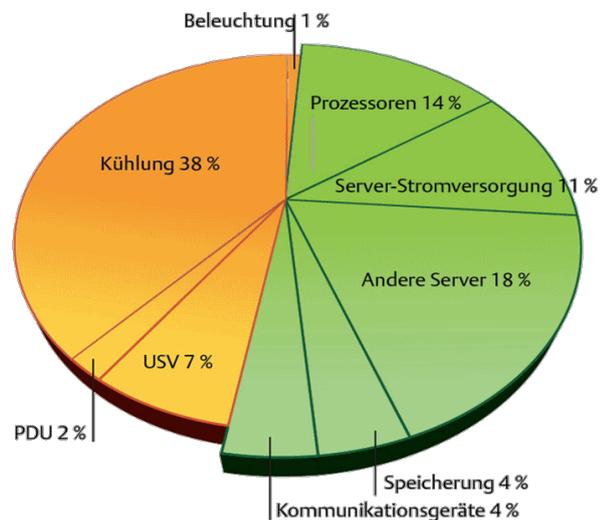


Abbildung 2.1: Aufteilung der Energiebedarfe in einem Rechenzentrum.¹

¹<http://www.datacenter-insider.de/index.cfm?pid=11511&pk=411514>, letzter Zugriff: 16.07.2015

²<http://www.datacenter-insider.de/index.cfm?pid=11509&pk=306372>, letzter Zugriff: 16.09.2014

³<http://heise.de/-1098823>, letzter Zugriff: 25.09.2014

die Anteile mehr in Richtung der IT bzw. der Infrastruktur. In allen drei Bereichen lassen sich Verbesserungen zur Senkung des Energiebedarfs erzielen.

Bei den IT-Komponenten wird der Großteil des Stroms von den Servern benötigt. Ältere Netzteile sind sehr ineffizient, so dass bei normalem Betrieb bis zu 30 % des Stroms verloren gehen. Dies entspricht einem Wirkungsgrad von 70 %. Neuere Netzteile haben einen Wirkungsgrad von bis zu 95 % [205]. Weitere Bestandteile eines Servers wie die CPU, RAM, Festplatte und die Netzwerkschnittstelle lassen sich ebenfalls optimieren. Bei der CPU, die für etwa 85 % des Mehrverbrauchs unter Last verantwortlich ist [225], werden bereits Energieeffizienz-Techniken wie das DVFS, die Lastbalancierung zwischen mehreren Cores [118] oder auch ein Power-Gating auf Core-Ebene [14] eingesetzt. Je nach Prozessortyp lassen sich dadurch hohe Einsparungen erzielen: beim AMD Quad-Core Phenom X4 9850 nach [137] bis zu 60 %. Beim Arbeitsspeicher bedeutet eine neuere Transistortechnologie nicht nur einen Geschwindigkeitszuwachs, sondern durch die geringeren Spannungen auch eine Senkung des Energiebedarfs. Ein größerer Technologiewechsel bahnt sich in der Festplattentechnik an. Gewöhnliche auf Magnet basierte Standardfestplatten haben einen hohen Fixpower-Wert von ca. 60 % der Leistungsaufnahme bei maximaler Auslastung [225]. Werden sie aktiv, schnellst der Wert bereits auf 80 %. Danach weisen sie wie auch Solid State Disks eine zur Last proportionale Leistungsaufnahme auf. Die durch die Last erhöhte Leistungsaufnahme ist bei beiden Technologien ähnlich, nur besitzt letztere (fast) keine statische Leistungsaufnahme. Techniken zum Energiemanagement von Festplatten und des Arbeitsspeichers wurden in [97, 42, 134, 121, 61] vorgestellt, kommen jedoch in der Praxis bisher kaum zum Einsatz.

Server sind, wie bereits erwähnt, die meiste Zeit über auch nur sehr gering ausgelastet, da sie für Lastspitzen ausgelegt sind. Durch Einsatz von Virtualisierungstechnik kann eine höhere Gesamtauslastung erreicht und dadurch eine Einsparung von bis zu 80 % erzielt werden [231, 145]. Erreicht wird dies durch Konsolidierung von Servern (Näheres hierzu in Abschnitt 2.1.2). Während sich hierbei unterschiedliche Dienste Server teilen, können sich Server wiederum auch andere Komponenten teilen. In Blade-Enclosures werden die Netzteile, Netzwerkanschlüsse und auch die Lüfter geteilt. Hierdurch kann eine höhere Auslastung je Komponente bewirkt und damit der Gesamtenergiebedarf gesenkt werden.

Für die Stromversorgung werden Transformatoren und Stromrichter benötigt, die den Strom auf die richtigen Spannungsbereiche transformieren bzw. eine Umwandlung zwischen Wechsel- und Gleichstrom vornehmen. Zusätzlich wird, wie bereits erwähnt, eine USV verwendet, welche im Falle eines Netzausfalls die Zeit bis zum Einsatz des Notfallgenerators überbrückt. Der hierzu erforderliche Strom wird i.d.R. in Akkumulatoren zwischengespeichert. Verluste entstehen in der USV und bei den Transformationen. Zur Vermeidung der doppelten Transformation wird die Lösung diskutiert, das komplette Rechenzentrum auf den Gleichstrombetrieb umzustellen [188]. Das zu erwartende Einsparpotential ist hier jedoch aufgrund bereits erhältlicher, sehr effizienter Doppelwandler mit einem Wirkungsgrad von 95 % und mehr⁴ gering.

⁴<http://www.perspektive-mittelstand.de/USV-Anlagen-Schwungrad-statt-Batterie-im-Rechenzentrum/management-wissen/4706.html>, letzter Zugriff: 06.10.2014

2 Stand der Forschung und Technik

Die Bewertung einer USV geschieht in der Fachliteratur zumeist über den Wirkungsgrad und dessen auslastungsabhängigen Verlauf, welcher sich mit höherer Auslastung dem maximalen Wirkungsgrad annähert. Dieser Verlauf ergibt sich aus dem auslastungsabhängigen absoluten Energiebedarf, welcher einen Basiswert aufweist und anschließend linear skaliert. Zur Energiebedarfsminimierung ist eine USV zwar grundsätzlich mit möglichst hohem Wirkungsgrad, also mit hoher Auslastung zu betreiben, bei einem LPM steht jedoch der tatsächliche Energiebedarf im Vordergrund, weshalb dieser und nicht der Wirkungsgrad zu betrachten ist. Aufgrund des linearen, auslastungsabhängigen Energiebedarfs ergibt sich somit kein Nachteil, wenn Rechenlasten zeitlich wie örtlich – identische Infrastruktur vorausgesetzt – verschoben werden und dabei eine USV (temporär) mit geringer Auslastung, also mit niedrigem Wirkungsgrad betrieben wird.

Die von den IT-Komponenten und auch von der USV benötigte Energie in Form von Strom wird nahezu vollständig in Wärme umgewandelt. Diese muss geeignet abgeleitet werden, um Hot-Spots, die zu Überhitzung und letztlich zum Ausfall von Komponenten führen können, zu vermeiden. Bei der Dimensionierung der Kühlung sollte die Richtlinienreihe VDI 4710 hinzugezogen werden, da mit ihrer Hilfe standortbezogene meteorologische Daten zur Gewährleistung von Temperatur und Feuchte genutzt werden können. Zur Kühlung kommen Luft- und Wasserkreisläufe zum Einsatz.

In der konventionellen Raumklimatisierung wird eine Kompressionskältemaschine (Chiller) benötigt. Dazu wird mittels Lüfter warme Luft aus dem Raum durch einen Wärmetauscher geführt. An diesen ist ein Kältekreislauf angeschlossen, welcher die thermische Energie aufnimmt. Der Kältekreislauf folgt dem Prinzip eines Carnot-Kreisprozesses, bei dem durch einen aktiven Verdichter das Kältemittel zuerst in einen flüssigen Zustand gebracht wird und im Anschluss verdampft. Durch den Verdampfungsprozess wird Energie aufgenommen und so die Kühlleistung erbracht. Mittels eines an den Wärmetauscher angeschlossenem Wasserkreislaufs, der durch Pumpen gesteuert wird, wird schließlich die Wärme abtransportiert.

Durch die Erhöhung der Raumtemperatur kann der benötigte Energiebedarf reduziert werden. Selbst führende Unternehmen wie IBM und Siemens empfehlen eine Erhöhung von den derzeit üblichen 22 °C auf 26 °C, was eine Einsparung von 16 % bringt.⁵ Eine weitere Erhöhung der Temperatur ist zwar erlaubt, wird allerdings nicht mehr als optimaler Betriebszustand angesehen [166]. Aufgrund der geringen Abfuhrkapazität von Luft, kommen bei high-density Racks mit 20 kW und aufwärts hybride Luft/Wasser-Schrankkühlssysteme zum Einsatz (z. B. Liquid Cooling Package, LCP [190]). Mit speziellen Wärmetauschern lässt sich sogar bis zu 1000-fach so viel Wärme abführen wie mit einer reinen Luftkühlung.⁶

Bei Luftkühlssystemen ist darauf zu achten, dass die Luftströmungen auf direktem Wege ohne hindernde Elemente (z. B. Kabel) und vor allem ohne Vermischung von kalter und warmer Luft realisiert sind. Eine Anordnung der Racks in Kalt- und Warmgängen (siehe [167])

⁵http://www.ifeu.de/energie/pdf/Server_Kuehlung_neu.pdf, letzter Zugriff: 25.11.2014

⁶http://www.tecchannel.de/server/hardware/462972/fluessigkeitskuehlung_fuer_server/index3.html, letzter Zugriff: 25.11.2014

sowie die Einhausung eines Typs dieser Gänge wird daher empfohlen [205].

Der Carnot-Kreisprozess wird als aktive Kühlung bezeichnet, da die Kühlleistung durch das aktive Hinzufügen von Energie erbracht wird. Die Kühlung lässt sich jedoch auch auf passive Weise erreichen, indem die Außenluft mit geeignet niedrigen Temperaturen genutzt wird. Laut Rittal ist „in mitteleuropäischen Breitengraden in 93 % der Fälle und damit an 339 Tagen im Jahr“ eine Kühlung allein durch Außenluft möglich.⁷ Dieses lässt sich auch anhand der Kühlungslandkarten von *The Green Grid* nachvollziehen.⁸ Für eine detailliertere Abschätzung bietet ebenfalls *The Green Grid* ein kostenfreies Online-Tool an.⁹ Die Wirksamkeit unterschiedlicher Kühlungstechniken ist in [22] untersucht worden.

Insgesamt sollten Techniken (weiter-)entwickelt werden, die entweder eine Konsolidierung von Komponenten oder einen besseren Energie-Last-proportionalen Betrieb der Hardware ermöglichen [225, 19]. Durch die Korrelation von Last und Energiebedarf würde der resultierende Energiebedarf größeren dynamischen Schwankungen unterliegen, was bei den alten und auch teilweise noch vorliegenden statischen Architekturen nicht der Fall ist, denn hier ist der Energiebedarf unabhängig von der Last auf einem nahezu konstanten (Lastspitzen-)Niveau [30, 22].

Viele dieser Ansätze erfordern jedoch aufwendige Eingriffe in die Struktur des Rechenzentrums und daher auch einen entsprechend hohen Kostenaufwand. Um einen möglichst niedrigen TCO (Total Cost of Ownership) und hohen ROI (Return on Investment) zu erlangen, müssen die Auswirkungen der Maßnahmen quantifizierbar sein. Hierzu können entsprechende Messungen durchgeführt werden. Der Strombedarf einzelner Komponenten kann z. B. durch die PDUs (Power Distribution Unit) von Raritan mit einer Genauigkeit nach ISO/IEC von $\pm 1\%$ gemessen werden¹⁰. Monitoring-Software, die solche Messungen überwacht und geeignet aufbereitet, findet sich z. B. in Rittals RiZone¹¹, Trellis von Emerson Network Power, Nlyte und CA Nimsoft Monitor sowie in Mansystems BlueIT¹², das auf dem OpenDCME Modell¹³ aufsetzt. Letzteres ist spezialisiert auf die Visualisierung von Kenndaten und Metriken, aus denen die Rechenzentrumseffizienz abgeleitet werden kann. Bekannte Metriken sind PUE (Power Usage Effectiveness) bzw. EUE (Energy Usage Effectiveness), DCiE (Data Center Infrastructure Efficiency) und CADE (Corporate Average Data center Efficiency).

In der ursprünglichen Form setzt die PUE die Gesamtleistungsaufnahme des Rechenzentrums ins Verhältnis zur Leistungsaufnahme der IT [189]:

$$PUE = \frac{power_{total}}{power_{IT}} \quad (2.1)$$

⁷<http://www.crn.de/server-clients/artikel-89390.html>, letzter Zugriff: 25.09.2014

⁸<http://www.cowo.de/a/2515558>, letzter Zugriff: 25.11.2014

⁹http://cooling.thegreengrid.org/EUROPE/WEB_APP/calc_index_EU.html, letzter Zugriff: 25.11.2014

¹⁰http://assets.raritan.com/resources/data_sheets/raritan-ds-PX_iPDU_high_res.pdf, letzter Zugriff: 25.09.2014

¹¹www.rittal.com/imf/none/5_606/, letzter Zugriff: 25.09.2014

¹²<http://www.mansystems.com/mansystems-blueit>, letzter Zugriff: 25.09.2014

¹³<http://www.opendcme.org/>, letzter Zugriff: 25.09.2014

Damit ist sie ein Maß für die Effizienz der Infrastruktur. Je geringer der Wert ist, desto weniger zusätzliche Leistung wird für die Infrastruktur verwendet. Der theoretisch optimale Wert 1 bedeutet, dass der gesamte Strom ausschließlich zum Betrieb der Server genutzt wird. Die PUE ist sehr einfach zu ermitteln und gibt eine sehr gute Einschätzung für Veränderungsmaßnahmen. In zahlreichen Abhandlungen wurden allerdings ihre Nachteile offengelegt. So ist die PUE nicht übertragbar auf andere Rechenzentren mit anderer IT-Ausstattung, denn die Leistungsfähigkeit der Hardware fließt nicht mit ein. Daher ist es durchaus möglich, dass ein Rechenzentrum mit einem geringen PUE-Wert mehr Energie für die Bereitstellung einer gewissen Rechenleistung benötigt als ein Rechenzentrum mit einem höheren PUE-Wert. Da die PUE nur eine Momentaufnahme darstellt, ist die Verwendung der EUE, welche sich analog berechnen lässt, sich aber auf den Energiebedarf (für ein Jahr) bezieht, besser geeignet. Inzwischen wird unter der PUE aber auch die EUE verstanden [15].

Die DCiE ist der Kehrwert der PUE. Damit beziehen sich diese Metriken auf einen beobachteten Betrieb, bei dem die ausgeführten Dienste eine bestimmte Auslastung bewirkt haben. Eine Abhängigkeit von der Auslastung besteht jedoch nicht, weshalb eine Angabe zur Effizienz bei einem anderen Workload nicht möglich ist. CADE ist eine Metrik, die auch die Auslastung berücksichtigt. Allerdings wird diese mit der IT-Effizienz multipliziert, welche sich nicht direkt messen lässt. Metriken, die sich konkret messen lassen und die die Auslastung berücksichtigen, sind z. B. CPE (Compute Power Efficiency) und LDEE (Load Dependent Energy Efficiency). Die Angabe der Auslastung ist an dieser Stelle jedoch nicht trivial, da Hardwarebestandteile (CPU, RAM, IO, HDD) nicht in gleichem Maße ausgelastet werden. Eine Übersicht und Vergleich unterschiedlicher Metriken findet sich in [200]. Aktuell eingesetzt wird hauptsächlich nur die PUE.

Inzwischen existieren mehrere Leitfäden vom Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, dem U.S. Department of Energy und weiteren mit Best-Practice-Lösungen zur Erhöhung der Rechenzentrumseffizienz [85, 37, 28, 179]. Werden die genannten Verbesserungen umgesetzt, können Rechenzentren mit einem PUE von bis zu 1,2 erreicht werden [205]. In der Tat weisen neuere Rechenzentren einen sehr geringen PUE-Wert auf: Das greenDatacenter Zürich West besitzt eine PUE von 1,4¹⁴, das PT-Rechenzentrum in Covilhã hat eine PUE von 1,25¹⁵ und Google behauptet sogar Rechenzentren mit einer PUE von weniger als 1,2 zu besitzen¹⁶.

Um die Auswirkungen von Verbesserungen bereits im Vorfeld quantifizieren zu können, ist es jedoch notwendig, den Rechenzentrumsbetrieb zu simulieren. Hierzu können Energiebedarfsmodelle von ganzen Rechenzentren, wie z. B. in [173, 218, 1, 157] vorgestellt, herangezogen werden. Diese geben generell die Realität vereinfacht wieder, reichen aber, um Veränderungen an der Rechenzentrumsarchitektur einschätzen zu können. Einige Modelle gehen detaillierter auf thermische Zusammenhänge ein, so dass z. B. auch die Topologie gezielt in eine Planung

¹⁴<http://www.greendatacenter.ch/enus/greenIT.aspx>, letzter Zugriff: 25.09.2014

¹⁵<http://www.datacenter-insider.de/index.cfm?pid=3739&pk=419630>, letzter Zugriff: 25.09.2014

¹⁶<http://www.infoworld.com/article/2629466/technology-business/facebook-boldly-predicts-pue-of-1-15-for-future-data-center.html>, letzter Zugriff: 25.09.2014

mit einbezogen werden kann. Da es jedoch sehr viele unterschiedliche (Kühl-)Techniken gibt (z. B. Raum-/Rackkühlung, Luft-/Wasserkühlung), existiert derzeit kein allgemeines Modell, das alle Techniken berücksichtigt. Ein Modell abstrahiert daher i.d.R. von einer gegebenen Infrastruktur.

Neben der Erhöhung der eigenen Effizienz, kann das Rechenzentrum auch als Teil eines größeren Systems betrachtet werden, das ineinandergreifend optimiert werden kann. So kann z. B. die durch das Rechenzentrum produzierte Abwärme in Bürogebäuden oder Hallenbädern zur Heizung dienen. Bis zu 90 % der generierten Wärmeenergie könnten hierbei weitergenutzt werden¹⁷. Das Rechenzentrum hätte nach dieser Rechnung, ökologisch betrachtet, nur ein Zehntel der abgerufenen Energie tatsächlich beansprucht.

2.1.2 Server-Virtualisierung und Live-Migration

Im Einleitungskapitel wurde bereits beschrieben, dass der überhöhte Energiebedarf von Rechenzentren der bisherigen Praxis geschuldet ist, einen Dienst aus Isolationsgründen dediziert auf einem Server zu betreiben. Dies beruht auf Kriterien zur Performanz, Sicherheit und Wartung. Da ein Server im praktischen Betrieb jedoch durchschnittlich nur zu 10 % - 20 % ausgelastet ist und im Leerlauf immer noch mehr als 50 % seiner maximalen Stromaufnahme besitzt [19, 129, 55], wird die meiste Zeit über Strom verschwendet. Durch Virtualisierung kann die durchschnittliche Auslastung erhöht und damit Energie effizienter genutzt werden, unter weitestgehender Beibehaltung der Performanz-, Sicherheits- und Wartungskriterien.

Als Virtualisierung versteht man im IT-Bereich die Abstraktion von Ressourcen [170]. Hierzu wird eine logische Schicht eingeführt, in der reale Ressourcen durch virtuelle ersetzt werden, die die gleichen Schnittstellen und Funktionen anbieten.

Bei der Server-Virtualisierung wird die gesamte Systemumgebung virtualisiert. Ein Betriebssystem, welches in einem virtualisierten Server, der sogenannten virtuellen Maschine (VM) betrieben wird, kann keinen Unterschied zu einem Betrieb auf einem dedizierten Hardware-Server feststellen. Die Ausprägungen einer VM (insbesondere CPU- und RAM-Werte) können softwareseitig eingestellt werden, wodurch ermöglicht wird, Anwendungen abstrahiert von der realen Serverhardware zu betreiben. Die Zugriffe auf die Serverhardware werden von einem Virtual Machine Monitor (VMM), auch Hypervisor genannt, als Schnittstelle zwischen Hardware und VM geregelt. Diese Architektur ist in Abbildung 2.2 dargestellt.

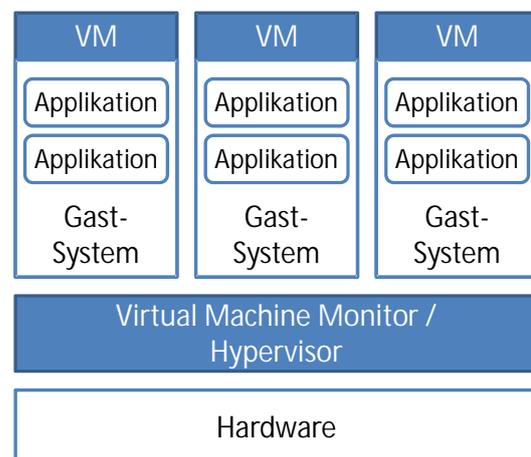


Abbildung 2.2: Virtualisierungsarchitektur

¹⁷<http://www.ibm.com/news/ch/de/2008/02/21/p079802o48711o43.html>, letzter Zugriff: 25.09.2014

2 Stand der Forschung und Technik

Das Gastsystem entspricht dem ursprünglichen System, das auf einem dedizierten Server betrieben wurde. Für dieses System erscheint der Betrieb unverändert: Es wird – nun virtuell – auf einem Server (Host) isoliert betrieben und hat keine Kenntnis darüber, ob andere Gastsysteme parallel betrieben werden. Der VMM erfüllt aber nicht nur die Aufgabe, den isolierten Betrieb der einzelnen VMs zu realisieren, sondern kennt auch alle internen Zustände der VMs, also auch bzgl. des Prozessors und des Arbeitsspeichers. Hierdurch ist ein Checkpointing und Rollback möglich, was die Architektur robust gegen z. B. einen Hardwareausfall macht.

Laut Gartner kommt die am meisten eingesetzte Virtualisierungssoftware von VMware (ESXi/vSphere) oder Microsoft (Hyper-V), siehe Abbildung 2.3. Lösungen von Citrix (Xen-Server) und Oracle (VM Server) waren in den letzten Jahren die nächsten Verfolger, verlieren aber an Boden. Weitere Alternativen kommen von Parallels, Red Hat und Huawei, doch diese spielten eine vergleichsweise unbedeutende Rolle.

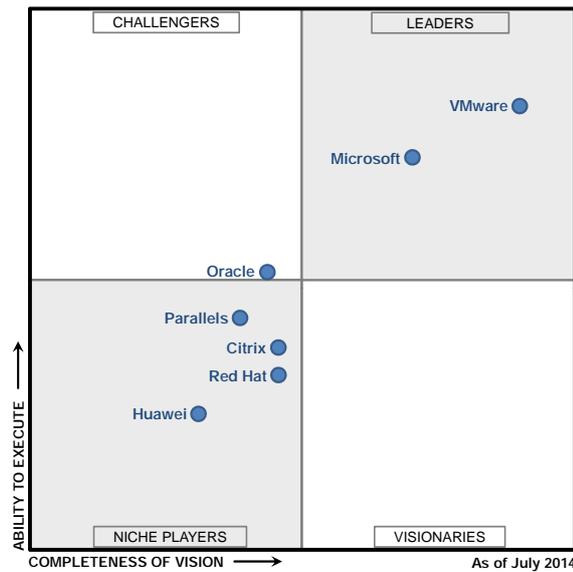


Abbildung 2.3: Magic Quadrant für x86-Server-Virtualisierung 2014. Quelle: Gartner

Statische Virtualisierung

Wie bereits erwähnt, können mehrere VMs isoliert voneinander auf dem gleichen physikalischen Server parallel betrieben werden. Auf diese Weise lassen sich niedrig ausgelastete Server auf eine geringere Anzahl an höher ausgelasteten Servern konsolidieren. Da die Ausprägungen von VMs softwareseitig eingestellt werden, ist es möglich, auf einen homogenen (virtuellen) Servercluster zurückzugreifen, auch wenn die realen Server heterogen sein sollten. Auf welchen dieser Server die VMs betrieben werden sollten, spielt bezüglich der Ausführbarkeit somit eine untergeordnete Rolle, sondern ist vielmehr von Merkmalen wie z. B. der Klimazone abhängig. Wurden die VMs auf eine geringere Anzahl an Hardwareservern konsolidiert, können die nicht mehr benötigten Server abgeschaltet werden. Neben den Energieeinsparungen durch die Verringerung der Anzahl aktiver Server kann auch an der Infrastruktur gespart werden, indem die Kühlkapazitäten und die Stromversorgung an die neuen Gegebenheiten angepasst werden.

Der VMM als zusätzlicher Ausführungslayer ist jedoch mit einem gewissen Overhead verbunden [178]. Die CPU-Auslastung auf dem Host-Server ist den Untersuchungen [180, 211] nach höher als die Summe der CPU-Auslastungen, wenn die Dienste auf einem dedizierten Server mit identischer Hardware betrieben worden wären. Aufgrund des Overheads ist nicht jeder Dienst dazu geeignet, einen Einspareffekt durch Konsolidierung zu erzielen. Soweit

keine neuere, effizientere Hardware beschafft wird, ist bei stark ausgelasteten Servern eine Konsolidierung – wenn überhaupt – nur schwer zu erreichen. Im schlimmsten Fall ist sogar ein erhöhter Energiebedarf zu verzeichnen. Zum Zwecke der Energieeinsparung sollten daher die entsprechenden Dienste eine geringe Auslastung haben. Vom Aspekt des Energiebedarfs abgesehen, bringt der Betrieb von Diensten mit hohem Rechenbedarf in einer VM aber auch Vorteile aufgrund der erhöhten Flexibilität mit sich.

Durch die sogenannte Paravirtualisierung kann der Overhead reduziert werden, indem die Zugriffe auf CPU, Arbeitsspeicher und Geräte-IO mittels spezieller Treiber realisiert werden. Speziell für High Performance Computing, wo die Performanz von Netzwerkverbindungen kritisch sein kann, ist es möglich, einen sogenannten VMM-Bypass durch gezielte Modifikationen zu realisieren, um einen Overhead fast vollständig zu vermeiden [112]. Beide Varianten verändern jedoch die Systemumgebung, weshalb die Sicht auf den betreibenden Server verändert und die Portabilität damit reduziert wird.

Dynamische Virtualisierung

Durch Schwankungen im Ressourcenbedarf der Anwendungen (z. B. wird der E-Mail-Server eines Unternehmens hauptsächlich zu Kernarbeitszeiten beansprucht) können die Server bei einer festen VM-Zuordnung nicht zu jeder Zeit optimal ausgelastet werden. Durch eine dynamische Zuweisung der VMs auf die Server kann die Anzahl der Server stets an den aktuellen Ressourcenbedarf angepasst werden, indem Server je nach Bedarf in einen Energiesparzustand versetzt bzw. aus diesem reaktiviert werden. Die Grundvoraussetzung für eine solche dynamische Virtualisierung ist mit der Live-Migrationstechnik [66] gegeben (siehe kommenden Abschnitt), welche die Verschiebung (Migration) einer VM zwischen zwei Servern während des Betriebs mit einer geringfügigen Ausfallzeit im Zehntelsekundenbereich ermöglicht. Die Ermittlung von sinnvollen Migrationen zur Minimierung des Energiebedarfs bei gleichzeitiger Vermeidung von Ressourcenengpässen übernimmt hierbei ein LPM.

Im Folgenden wird die Funktionsweise eines solchen LPM anhand eines kleinen Beispiels verdeutlicht, welches in Abbildung 2.4 dargestellt ist. Zum Zeitpunkt t_1 sind alle drei Server aktiv und es werden mehrere VMs auf diesen betrieben. Zum Zeitpunkt t_2 verringert sich der Ressourcenbedarf von VM 6, so dass für den Betrieb aller VMs nun auch zwei Server ausreichen würden. Hierzu werden die VMs von Server A per Live-Migration auf die anderen beiden Server migriert. Server A kann danach in einen Energiesparzustand versetzt werden. Zum Zeitpunkt t_3 ist der Ressourcenbedarf der VMs auf Server C gestiegen, so dass ein Ressourcenengpass entsteht. VM 1 wird wieder zurück auf Server A migriert, dieser aber zuvor aus dem Energiesparmodus geweckt. Die Live-Migration und die Energiebedarfszustandsänderungen der Server verursachen einen Energiebedarfsoverhead. Ist die Zeit zwischen den Zeitpunkten t_2 und t_3 so gering, dass sich der Overhead noch nicht amortisiert hat, ist die temporäre Konsolidierung nicht zweckmäßig gewesen. Das LPM muss somit mit einem genügenden Weitblick seine Entscheidungen bzgl. Migrationen und Energiebedarfszustandsänderungen der Server treffen.

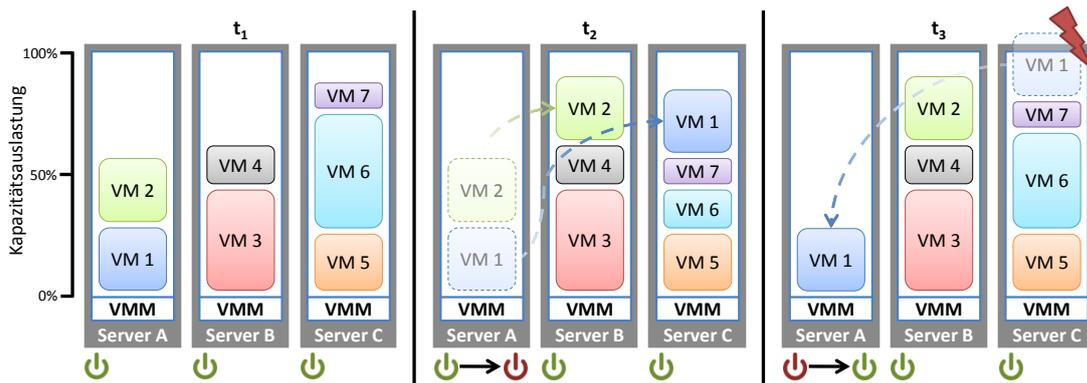


Abbildung 2.4: Beispiel für dynamische Virtualisierung zur Minimierung der Anzahl aktiver Server

Live-Migration

Die Technologie der Live-Migration erlaubt einer VM zur Laufzeit den Host zu wechseln – bei einer Ausfallzeit im Bereich von Zehntelsekunden. Der Ablauf lässt sich mit den folgenden Schritten grob beschreiben [66]:

0. Es wird sichergestellt, dass die benötigten Ressourcen am Zielhost zur Verfügung stehen.
1. Ein VM-Container mit gewünschten Ressourcen-Eigenschaften (ohne konkrete VM-Daten) wird am Zielhost reserviert.
2. Der Arbeitsspeicher wird kopiert. Iterativ werden weitere Pages nachgeladen, die sich zwischenzeitlich verändert haben (*dirty pages*).
3. Die VM wird angehalten, der Netzwerktraffic zur VM-Zielinstanz geleitet und die Prozessorwerte (Zustand, Cache) sowie die verbleibenden *dirty pages* kopiert. Es gibt nun eine komplette Kopie der VM. Falls ein Fehler auftreten sollte, ist der Betrieb am Starthost, der bis jetzt noch Primärhost ist, nach wie vor möglich.
4. Der Zielhost wird Primärhost. Die VM-Daten auf dem Starthost können somit gelöscht werden.
5. Die VM startet auf dem Zielhost und der normale Betrieb wird fortgesetzt.

Die Ausfallzeit tritt in den Schritten 3 und 4 auf. Auf weitere Aspekte der Live-Migration wird im Folgenden kurz eingegangen.

IP-Adresse

Beachtet werden muss, dass die VM nach einer Migration zur besseren Erreichbarkeit die gleiche IP-Adresse behält. Rechenzentrumsintern kann dies durch die Nutzung der Netzwerkarchitektur VL2 [93] gewährleistet werden. Durch ihre Agilität und hohe Skalierbarkeit ist sie insbesondere in Cloud-Umgebungen von Vorteil. Die konkrete Allokation einer VM hat bezogen auf die Netzwerkanbindung hierbei keinen Einfluss auf die Performanz.

Wird jedoch eine Migration in ein geografisch entferntes Rechenzentrum durchgeführt, so befinden sich aufgrund der geografisch bezogenen IP-Vergabe die beteiligten Rechenzentren in unterschiedlichen IP-Subbereichen. Eine Methode, um dieses Problem zu umgehen, besteht im dynamischen Routing auf Providerebene, vorgestellt in [184]. Der Router muss hierzu allerdings sämtliche VMs und Server kennen, was aufgrund der Größeneinschränkung der Routingtabellen nicht praktikabel ist. Außerdem werden in diesen Tabellen normalerweise Informationen zum Erreichen von Subnetzen und nicht von einzelnen Servern hinterlegt. Eingesetzt werden kann dieser Mechanismus auch nur dann, wenn die Rechenzentren am gleichen Netzprovider angeschlossen sind. Eine weitere Methode wird in [223] vorgestellt. Hier wird ein IP-Tunnel zwischen dem Client und der VM aufgebaut. Bei einer Migration wird der IP-Tunnel entsprechend rekonfiguriert. Hierzu wird allerdings clientseitig eine Software zum Aufbau eines IP-Tunnels benötigt. In [164] wird das Problem dadurch gelöst, dass ein sogenannter *Federation Stitcher* eingesetzt wird, welcher sämtliche Anfragen entgegen nimmt und sie an das jeweils aktuelle Rechenzentrum weiterleitet. Komplizierter wird dies bei der Migration zwischen Rechenzentren unterschiedlicher Anbieter, denn diese werden im Allgemeinen keine gemeinsame zentrale Anlaufstelle haben.

Mit VMware vCloud Director [234] existiert ein kommerzielles Produkt für das Migrieren von VMs zwischen unterschiedlichen Rechenzentrumsstandorten, das in Verbindung mit VMware vCloud Connector¹⁸ – VXLAN (Virtual Extensible LAN) vorausgesetzt – den Erhalt der IP-Adresse und Erreichbarkeit ermöglicht, zumindest im Firmennetz. Um von extern den direkten Zugriff auf migrierte VMs zu haben, könnten Software-Defined-Network-Technologien wie OpenFlow [227] aushelfen. Hierdurch lässt sich der Datenfluss (Routing) besser kontrollieren und gezielt managen, da die Steuerung nicht mehr durch den Router, sondern serverseitig, entweder manuell durch einen Anwender oder automatisch durch eine spezielle Applikation erfolgen kann. Vorbehalte gegenüber dieser Technologie werden jedoch durch mangelnde Skalierbarkeit und dem Single-Point-of-Failure¹⁹ begründet.

In dieser Dissertation wird von dieser Problemstellung abstrahiert, d.h. es wird von einer technischen Realisierung ausgegangen, die eine globale Migration samt IP-Erreichbarkeit ermöglicht.

Verzögerung

Um VMs migrieren zu können, muss laut [13] die Netzwerkanbindung genügend Bandbreite (622 Mbits/s) aufweisen und die Latenz gering genug (≤ 10 ms, was einer Entfernung von 500 km entspräche) sein. Die Migrationszeit hängt von diesen beiden Werten und der Größe des zu transferierenden Arbeitsspeichers ab [86]. Eine unter diesen Bedingungen in [64] durchgeführte Untersuchung ergab für die Migration einer VM mit 8 GB Arbeitsspeicher innerhalb eines Rechenzentrums eine Migrationszeit von 57 Sekunden und in ein 100 km

¹⁸<http://www.vmware.com/products/vcloud-connector>, letzter Zugriff: 04.08.2015

¹⁹Ein Single-Point-of-Failure ist ein Bestandteil eines Systems, bei dessen Ausfall die Funktion des gesamten Systems ausfällt.

entferntes Rechenzentrum eine Migrationszeit von 84 Sekunden.

In [136] wurde ermittelt, dass zur Erhaltung des Betriebs lediglich eine Bandbreite von 32 Mbit/s vorhanden sein muss. Die erforderliche Zeit, um VMs mit 4 GB RAM zu migrieren, betrage 1000 s. Dies würde allerdings der Nennbandbreite entsprechen, die zum einen in der Praxis nicht vollständig nutzbar ist und zum anderen müssen die in der Zwischenzeit veränderten Daten (Pages) nachgeladen werden. In [232] wurde die Migrationszeit für die gleiche RAM-Größe bei einer 10-Gbit/s-Netzanbindung gemessen: Mit 13 s liegt ein Faktor von 4 gegenüber der theoretisch erreichbaren Zeit vor. Diese Angabe bezieht sich auf VMs, die unter Volllast betrieben werden. Im Idle sinkt die Migrationszeit auf 10 s. Bei größerem Arbeitsspeicher fällt die Differenz auch in Relation größer aus. Auf einem Server können nicht alle VMs gleichzeitig unter Volllast betrieben werden. Ebenso werden sich auch nicht alle im Idle befinden. Daher wird von einer durchschnittlichen Auslastung ausgegangen. 30 % stellt bereits einen hohen Wert dar, soll hier aber zur Überabschätzung dienen. Nach [238] hängt die Geschwindigkeit aber auch von der geografischen Entfernung ab. So wird für eine Migration an einen 100 km entfernten Ort ca. 20 % mehr Zeit als bei einer lokalen Migration benötigt. Rechenzentrumsintern ist somit von einer Migrationszeit von etwa 11 s und rechenzentrenübergreifend von etwa 13 s auszugehen.

Während der Migration kann es aufgrund der erhöhten Serverbelastung zu geringfügigen Leistungseinbußen kommen [235]. Aber auch beim Umschalten auf den Zielservers soll die Performanz um ca. 10 % einbrechen [64]. Hier kommen im Vergleich zur rechenzentrumsinternen Migration grundsätzlich höhere Ausfallzeiten vor, und zwar von einigen Sekunden [224, 100, 235]. Durch geeignete Optimierungstechniken zur Reduzierung des notwendigen Datenübertragungsbedarfs können aber sowohl die Ausfallzeiten als auch die Gesamt migrationszeit dezimiert werden [216, 246, 21].

Daten

Grundsätzlich ist bei einer Migration einer VM zu unterscheiden, welcher Dienst-Typ darin betrieben wird [238]: Handelt es sich um einen session-bezogenen Dienst²⁰, muss der aktuelle Bearbeitungszustand, welcher im Arbeitsspeicher hinterlegt ist, auf den Zielservers kopiert werden. Bei einem session-losen Dienst, ist ein Kopiervorgang des Arbeitsspeichers nicht notwendig. Die VM kann somit am Zielservers neu instantiiert werden. Das Storage muss nicht zwangsläufig kopiert werden: Innerhalb eines Rechenzentrums sind die VM-Daten über ein SAN zugreifbar.²¹ Bei standortübergreifenden Migrationen führt eine zentrale Speicherung der Daten zu einem verzögerten Datenzugriff, was evtl. die Service-Level-Agreement (SLA)²² verletzen kann. In [105] wird ein nachgelagerter Kopiermechanismus vorgestellt, welcher die

²⁰Eine Session, zu Deutsch Sitzung, beschreibt einen zusammenhängenden Zugriff eines Clients auf einen Server. Sie wird i.d.R. durch ein Login begonnen und durch ein Logout abgeschlossen. Im Internet kann ein Zusammenhang auch automatisch mittels einer Session-ID gebildet werden.

²¹<http://www.ip-insider.de/index.cfm?pid=10336&pk=305996>, letzter Zugriff: 14.11.2014

²²SLA bezeichnet eine Vereinbarung zwischen Kunde und Anbieter für den Betrieb eines Dienstes, in der zugesicherte Leistungseigenschaften (Service-Level-Objectives) wie etwa Leistungsumfang, Reaktionszeit und Schnelligkeit der Bearbeitung festgehalten sind.

benötigten Daten im Nachhinein im Hintergrund kopiert. Dies erhöht die Performanz, kann aber dennoch zu SLA-Verletzungen führen.

2.1.3 Last- und Powermanagement auf Systemebene

Eine bereits angesprochene Methode, um eine höhere Energieeffizienz zu erhalten, besteht in der Verwendung dynamischer Virtualisierung. Das eingesetzte LPM hat die Aufgabe, zur Laufzeit die Anzahl aktiver Server zu minimieren und dabei sicherzustellen, dass zu jedem Zeitpunkt genügend Ressourcen bereitgehalten werden, indem es die VMs auf passende Server (re-)alloziert. Das LPM wird vor eine Herausforderung gestellt, wenn der Ressourcenbedarf plötzlich ansteigt und VMs auf einen freien Server migriert werden müssen. Ist dieser aufgrund einer vorangegangenen Energieeinsparmaßnahme erst noch hochzufahren, könnte die hierfür benötigte Zeit zu lang sein und daher ein Ressourcenengpass entstehen.

Eine Möglichkeit, dieses Problem zu umgehen, besteht darin, Schwellenwerte bzgl. der Ressourcenauslastung einzuführen, bei deren Überschreitung Anpassungen in der VM-Server-Zuordnung vorzunehmen sind. Hierdurch wird ein zeitlicher Puffer gebildet, um auf mögliche Engpässe zu reagieren. Notwendige Serverreaktivierungen und VM-Migrationen können so vor dem Auftreten von Engpässen eingeleitet werden. Engpassfreiheit wird hiermit nicht garantiert. Je größer der Puffer ist, desto geringer ist jedoch die Wahrscheinlichkeit, dass Engpässe auftreten. Allerdings sinkt hierdurch auch das Stromeinsparpotential. Verfahren, die diesen Mechanismus verwenden ([125, 247, 230]), gelten als reaktiv.

Diesen gegenüber stehen proaktive Verfahren ([31, 102, 249]), die (ergänzend) ein Vorhersageverfahren nutzen, um den zukünftigen Ressourcenbedarf von VMs abzuschätzen. Sie leiten den zukünftigen Bedarf aus einer Trendentwicklung ab. Je besser dabei die Prognose dem tatsächlichen Bedarf entspricht, desto effizienter kann ein Managementsystem vorhandene Kapazitäten verwalten, denn mit geringeren Abweichungen verringern sich auch die für Fehlprognosen bereitzuhaltenden Kapazitäten. In [249] wird beispielsweise eine Fuzzy basierte Vorhersage im Minutenbereich verwendet. Gegenüber statischer Virtualisierung wurde eine Einsparung von 44% ermittelt, bei Performanzeinbußen von 1%. Ein Verfahren mit ähnlichem Einsparpotential aber zugesicherter Performanz und mit Vorhersagen im Stunden- bis Tagebereich wird in [110] vorgestellt. Auf dieses wird in Abschnitt 3.2 näher eingegangen.

LPM-Verfahren unterscheiden sich nicht nur anhand ihres Aktivitätsmusters, sondern insbesondere in der Granularität der verschiebbaren Lasteinheiten. Die Verschiebung von Lasten durch Live-Migration ist eine vergleichsweise neue Technologie. In der Forschung wurden bisher die Lastverschiebungen durch Allokation von Jobs behandelt. Im Webserver-Bereich bedeutet dies, dass Anfragen über mehrere Server verteilt werden [69]. Für den Betrieb eines solchen Dienstes ist also ein Server-Cluster notwendig. Als Optimierungsziel kann abweichend zur Minimierung des Energiebedarfs auch die Balancierung der Lasten über alle Server gewollt sein. Dies dient z. B. zur Regulierung der Temperaturentwicklung [172, 156], um sogenannte Hot-Spots zu vermeiden [202]. Für gewöhnlich werden bei Scheduling-Verfahren, die die Thermik berücksichtigen, neben Energiebedarfsmodellen für Server und Racks auch

Kühlungsmodelle sowie Luft- bzw. Hitzezirkulationsmodelle eingesetzt [218].

Neben dem örtlichen Verschieben von Lasten ist es in einigen Fällen auch möglich, eine zeitliche Verschiebung durchzuführen [91, 6]. Hierzu muss die Ausführung der Dienste zu einem früheren bzw. späteren Zeitpunkt möglich sein, wie es z. B. bei Backup-Prozessen der Fall ist. Des Weiteren existieren neben dem Lastmanagement, welches sich auf den IT-Teil bezieht, auch solche für die Klimatisierung und USV. Auf diese wird hier nicht weiter eingegangen. Sie werden als gegeben angesehen und in den späteren Modellierungsprozessen in Kapitel 4 mit einbezogen.

Die Allokation von Jobs und das Migrieren von VMs hat technisch gesehen keine lokalen Beschränkungen. Dadurch sind Lastverschiebungen auch zwischen geografisch entfernt liegenden Rechenzentren möglich. Die Ziele einer globalen Optimierung liegen zumeist in der Balancierung der Lasten und in der Reduzierung der Antwortzeiten. Aber auch Standortfaktoren können für eine Optimierungsstrategie herangezogen werden. Dies sind die Witterungsbedingungen, die Auswirkungen auf den Energiebedarf der Klimatisierung besitzen, und die Energieversorgung. Der gelieferte Strom unterscheidet sich in seiner Herkunft (Atom, Kohle, Wind, Wasser etc.) und kann sich in seinem Preis ändern. Durch Verlagerung der Last wird zugleich eine Verlagerung des Energiebedarfs bewirkt, da durch die Lastmanagementsysteme (DVFS, Konsolidierung etc.) die Rechenleistung und damit auch der Energiebedarf den Lasten entsprechend angepasst wird. Es ist also möglich, gezielt den Energiebedarf von Standorten und Regionen zu beeinflussen, was z. B. beim intelligenten Steuern des Energienetzes (Smart Grid) oder zum Demand Side Management bzw. Demand Response genutzt werden kann.

Der Energiebedarf lässt sich aber auch lokal beeinflussen. Wie bereits herausgestellt wurde, ist der Energiebedarf eines Rechenzentrums sowohl von der eingesetzten IT-Hardware und den Infrastrukturkomponenten als auch von den eingesetzten Managementsystemen abhängig. Theoretisch kann bereits ein einzelner Server am Demand Response partizipieren und eine Energiekosteneinsparung von bis zu 30 % erzielen [57]. Bei Multi-Core-Prozessoren kann der Workload auf unterschiedliche Weise verteilt werden, so dass bei gleicher Rechenleistung der Energiebedarf variiert. Die Nutzung von Server-Power-Capping-Techniken und der gezielte Betrieb in unterschiedlichen Server-Energiezuständen erlaubt die Senkung des Strombedarfes ohne größere Einbußen an einem Service-Level-Objective (SLO) [56]. Zeitlich verschiebbare Dienste können variabel geplant werden, um die aktuelle Leistungsaufnahme anzupassen [91, 6]. Die unterbrechungsfreie Stromversorgung kann als Energiespeicher genutzt und der Betrieb der Klimatisierung innerhalb eines gewissen Rahmens verändert werden [27]. Wird Server-Virtualisierung genutzt, lässt sich der Strombedarf anhand der Konsolidierungsrate anpassen. In der in Abschnitt 4.3.2 durchgeführten Untersuchung für die bestmögliche Konsolidierung bei heterogenen Servern zeigt sich ein großer Unterschied im Strombedarf bei Verwendung unterschiedlicher LPM-Policies.

Die große lokale Energiebedarfsflexibilität lässt sich durch ein globales LPM für noch weiter reichende Anpassungen nutzen, was nach Ansicht der kalifornischen Energiekommission ein signifikantes Potential beim DSM besitzt [89]. In der Literatur konnte dies in zahlreichen auf

Simulationen gestützten Untersuchungen bestätigt werden [238, 3, 153, 27, 57, 56]. Hierbei werden hauptsächlich Large-Scale-Services betrachtet, welche über mehrere Rechenzentren verteilt betrieben werden. Die Datenbasis wird als vollständig repliziert vorausgesetzt, so dass eine Lastverschiebung durch simples Rerouting der Anfragen [237, 152] möglich ist. Für Netzbetreiber bewirkt dies die Minimierung von Netzverlusten [152], die Vermeidung von Netzausbau und die Stabilisierung des Netzes [12]. Zwar steigt hierdurch die Belastung des Datennetzes, allerdings sei diese laut [164] günstiger als die Belastung des Stromnetzes. Um aber hierdurch evtl. entstehende Engpässe im Datennetz zu vermeiden, sollten der Netztraffic im Vorfeld z. B. durch Nutzung des Network Weather Service [245] abgeschätzt werden.

2.1.4 Cloud Computing

Kurz beschrieben, stellt Cloud Computing laut [133] eine Lösung für KMU dar, um ihren gesamten IT-Bedarf auszulagern, und für große Unternehmen ist es eine Möglichkeit, bei Lastspitzen auf Ressourcen zurückzugreifen, ohne die eigene Infrastruktur ausbauen zu müssen. Um diese Aussage nachvollziehen zu können, wird im Folgenden der Begriff des Cloud Computings geklärt und seine Formen und Anwendungsmöglichkeiten diskutiert.

Was bedeutet Cloud Computing?

Cloud Computing ist eine Form des IT-Bezugs. Die Angabe einer einheitlichen Definition stellt sich aufgrund unterschiedlicher Sichtweisen als schwierig heraus. Eine gemeinsame Basis wird allerdings durch die Definition von Akamai wiedergegeben [4], wonach es beim Cloud Computing darum geht, IT-Ressourcen (*Computing*) skalierbar und nach Bedarf über ein Netzwerk (*Cloud*) bereitzustellen. In den Definitionen der BITKOM [150], National Institute of Standards and Technology (NIST) [146] und aus dem Open Cloud Manifesto [168] werden die Eigenschaften Kosteneffizienz, Transparenz bzgl. der Komplexität der dahinterliegenden Technologie, der bequeme Bezug, in Echtzeit und mit verbrauchsgenauer Abrechnung (Billing) hinzugefügt. In [169] werden die IT-Ressourcen konkreter beschrieben. Hier wird von Applikationen, Entwicklungsplattformen und Infrastrukturelementen gesprochen, was der üblichen Unterteilung der unterschiedlichen Cloud-Services in die Service-Ebenen *Infrastructure-as-a-Service* (IaaS), *Platform-as-a-Service* (PaaS) und *Software-as-a-Service* (SaaS) entspricht, auf die im nächsten Abschnitt eingegangen wird. In der Definition der European Network and Information Security Agency [51] wird die dahinterliegende Technologie konkretisiert: Virtualisierung und Techniken des verteilten Rechnens sollen das Cloud Computing ermöglichen.

Service-Ebenen

Auch in den Arbeiten von Buyya, Moreno-Vozmediano und weiteren [41, 158] wird die Virtualisierung als eine Basistechnologie für Cloud Computing angesehen. Hier werden Dienste, wie beschrieben, nicht auf dedizierten Servern betrieben, sondern in VMs, welche sich flexibel

erstellen, migrieren und wieder freigeben lassen. Diese Flexibilität ist ein Grundbaustein für IaaS [158].

Beim IaaS werden IT-Ressourcen wie z. B. Server und Storage angeboten. Aufgrund des hohen Einsatzes von Virtualisierung handelt es sich hierbei zumeist um virtuelle Server. Aber auch physische Server lassen sich teilweise im Angebot der Cloud-Provider finden. Der Kunde entscheidet darüber, welches Betriebssystem und welche Software auf den Servern betrieben werden. IaaS bietet Kostensenkung und Risikominimierung durch Vermeidung kostspieliger Anschaffungen, insbesondere zum Betrieb von Large-Scale-Services. Der Unterschied zu normalen Hosting-Anbietern besteht in der schnelleren Bereitstellung und feineren Abstufung der gewünschten Leistung sowie der verbrauchsabhängigen Preisgestaltung. Der Zugang zu den Angeboten wird über APIs oder Verwaltungswerkzeuge realisiert [222]. IaaS-Anbieter sind z. B.:

- Amazon EC2 (Server) bzw. S3 (Storage)
- Fujitsu Cloud Infrastructure as a Service (Server, Storage)
- GoGrid (Server, Storage)
- Oracle Cloud Infrastructure (Server, Storage)

PaaS liefert Anwendungsinfrastruktur. Hierzu gehören zum einen Entwicklungs- und Laufzeitumgebungen für z. B. Webanwendungen. Dabei werden häufig visuelle Werkzeuge eingesetzt, die die Entwicklung mit wenig Programmierkenntnissen unterstützen sollen [222]. Zum anderen werden Laufzeitumgebungen für Programmiersprachen (z. B. .Net, Python, Java) und Datenbanken angeboten. Der Kunde hat zwar Einfluss auf die Konfiguration und die zu betreibende Software, allerdings ist die Plattform vorgegeben. Der Betreiber realisiert dynamisch die bedarfsgerechte Anpassung der Leistungsfähigkeit. Neben der Investition in Infrastruktur wird hier also auch das Management übernommen. Insgesamt sinkt im Gegensatz zum IaaS die Flexibilität, und die Möglichkeit eines Vendor Lock-Ins²³ ist erhöht. PaaS-Anbieter sind z. B.:

- Akamai EdgeComputing (u.a. Leistungssteigerung von Webanwendungen)
- Google App Engine (Entwicklungsumgebung)
- Microsoft Azure (u.a. Datenbanken, Entwicklungsumgebung)
- T-Systems Cloud Integration Center (u.a. Datenbanken, Webserver)

Beim SaaS können Nutzer bequem und einfach (End-User-)Applikationen beziehen [4]. Um die Bereitstellung und Skalierung der Performanz kümmert sich vollständig der Cloud-Provider. Die Applikationen sind – wenn überhaupt – nur mit Einschränkungen anpassbar

²³Mit Vendor Lock-In wird der Effekt bezeichnet, dass es aufgrund von hohen Wechselkosten bzw. hohem Aufwand unwirtschaftlich ist, die aktuelle Lage (Produktnutzung) durch einen Wechsel (zu einem Produkt eines anderen Anbieters) zu ändern.

bzw. konfigurierbar [222]. Ein Vendor Lock-In ist hier sehr wahrscheinlich. SaaS-Anbieter sind z. B.:

- Adobe Connect, Cisco WebEx, Google Mail (Kommunikationsapplikationen)
- Demandware (E-Commerce)
- Netsuite (u.a. Accounting, ERP)
- Salesforce.com (CRM)

In Abbildung 2.5 sind diese drei Service-Ebenen nach Verwaltungs- bzw. Verantwortungsbereichen dargestellt.²⁴ Hierbei ist zu erkennen, dass bei der Bereitstellung von Services auf der SaaS-Ebene der Cloud-Anbieter den gesamten Verwaltungsaufwand besitzt. In Richtung IaaS erhöht sich der Verwaltungsaufwand für den Nutzer. Um die Bereitstellung der IT-Hardware muss sich dieser allerdings nicht kümmern, sondern erst, wenn von ihm selbstständig das Rechenzentrum betrieben wird.

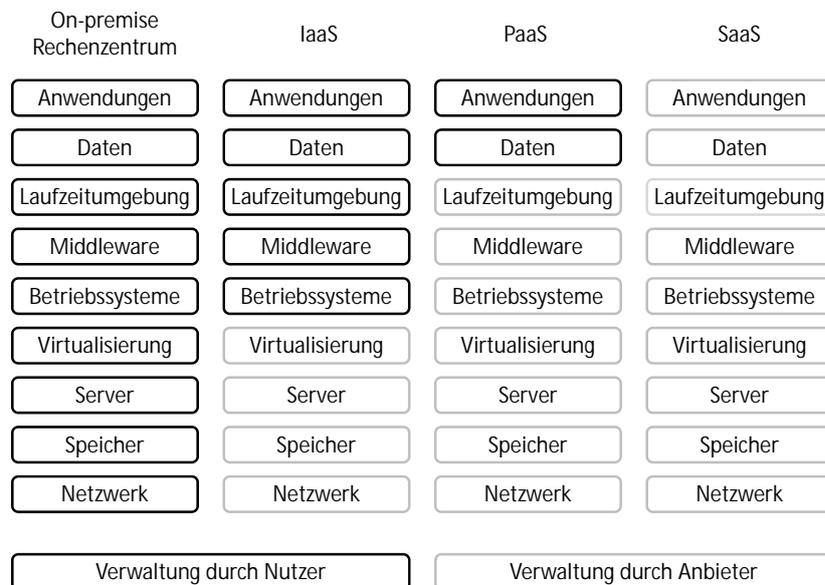


Abbildung 2.5: Verwaltungs- bzw. Verantwortungsbereiche bei den unterschiedlichen Service-Ebenen

Damit das Cloud-Angebot auch für bereits betriebene Dienste nutzbar ist, müssen diese einmalig in die Cloud migriert werden. Dies ist allerdings ein komplexer und ressourcenaufwendiger Prozess, welcher sich bis zu einem Jahr hinziehen kann. Aus diesem Grunde wurden Cloud Transformation Services eingerichtet, die den Prozess binnen weniger Wochen realisieren können [144]. Ob und in welchem Maße sich der Bezug von Services aus der Cloud im Vergleich zum Eigenbetrieb überhaupt rentiert, kann mit Hilfe des Cloud-Kalkulators²⁵ der Experten

²⁴<http://dailyitproblems.azurewebsites.net/cloud-computing-service-modelle>, letzter Zugriff: 15.01.2015

²⁵<http://www.business-cloud.de/calculator/calculator.html>, letzter Zugriff: 08.05.2015

Group abgeschätzt werden. Zu beachten ist auch, dass bei ausgelagerten Applikationen mit einem erhöhten Netzwerktraffic zu rechnen ist, welcher zu einer erhöhten Providerrechnung führen kann.

Cloud-Typ

Rechenzentren, in denen die Technologie für das Cloud Computing umgesetzt ist, können Dienste also dynamisch und flexibel bereitstellen. Diese Rechenzentren werden als sogenannte Clouds angesehen.

Bei unerwartet hohen Belastungen kann laut den Autoren von [11] eine einzelne Cloud an ihre kapazitiven Grenzen kommen. Daher wird von ihnen vorgeschlagen, eine übergeordnete Cloud-Organisationsschicht einzuführen, die ein Inter-Cloud-Computing realisiert, d.h. eine föderierte Cloud bildet. Ein Cloud-Verbund, welcher wiederum eine Cloud darstellt, erhöht auch die Zuverlässigkeit vor Dienstaussfällen. Die Realisierung föderierter Clouds stellt allerdings noch eine Herausforderung dar [158], da es aufgrund fehlender Standards Interoperabilitätsprobleme gibt, worauf später noch eingegangen wird.

Clouds werden in die drei folgenden Typen unterteilt [146]:

- *Public Cloud*: Die Angebotspalette einer Public Cloud ist über das Internet erreichbar. Jedes Unternehmen kann hier nach Bedarf Dienste etc. beziehen. Die Angebote profitieren von den Skaleneffekten der Cloud-Provider, was daher gerade für KMU interessant ist, ihre IT auszulagern.
- *Private Cloud*: Eine Private Cloud wird unternehmensintern gebildet, um die eigenen IT-Ressourcen leicht und flexibel für unterschiedliche Dienste und Abteilungen bereitzustellen [17]. Die hierin betriebenen Dienste werden ausschließlich vom Unternehmen verwendet und nicht der Öffentlichkeit zugänglich gemacht. Die flexible Architektur wird also zur Verbesserung der internen Abläufe genutzt. Um solch eine Infrastruktur aufzubauen, sind jedoch initial hohe Investitionen notwendig. Eine Private Cloud ist somit eher für große Unternehmen attraktiv, die vom Skaleneffekt profitieren können.
- *Hybrid Cloud*: In den meisten Fällen werden laut [4] allerdings keine ausschließlichen Private Clouds gebildet, sondern diese an eine Public Cloud angeschlossen, um z. B. bei Lastspitzen auf zusätzliche Ressourcen zurückgreifen zu können. Eine solche Cloud wird als Hybrid Cloud bezeichnet.

Cloud-Provider wie Google, Amazon, Microsoft, Facebook und Salesforce.com bauen weltweit sehr große Rechenzentren für ihre Cloud-Dienste auf. Dabei sind die Energieversorgung, die Energiepreise sowie günstige Rahmenbedingungen wie niedrige Außentemperaturen oft mitentscheidend für die Standortwahl. Auch in Deutschland werden Kapazitäten für Cloud-Computing erschlossen, wobei die Gründe hier vor allem darin liegen, dass Kunden ihre Daten nicht ins Ausland verlagern wollen oder dürfen (rechtliche Rahmenbedingungen). Hierzu wird oft die Infrastruktur von Colocation-Dienstleistern genutzt, da diese bereits

über große Hardware-Kapazitäten verfügen. Neben der Hardware wird auch Management-Software benötigt. Open-Source-Software, die hierfür eingesetzt werden kann, ist z. B. Apache CloudStack, HP Helion Eucalyptus und OpenStack.

Grid vs. Cloud

Beim Cloud Computing wird u.a. Rechenleistung als beziehbarer Dienst angeboten, was auch durch das Grid Computing ermöglicht wird. Daher bestand zur Geburtsstunde des Cloud-Hypes im Jahre 2009 Verwirrung darüber, ob es sich um die gleiche Technologie mit anderem Label, eine Evolution oder um etwas Neues handelt. Zum besseren Verständnis der in Abschnitt 2.2.2 kommenden Abgrenzung zu verwandten Arbeiten soll auch hier eine kurze Abgrenzung zwischen Grid und Cloud Computing gegeben werden.

Grob beschrieben kann man den Unterschied zwischen Grid und Cloud Computing daran festmachen, dass es beim Grid Computing oft darum geht, große Rechenlast durch große Ressourcenkapazitäten zu bedienen, während es beim Cloud Computing mehr darum geht, große Ressourcenkapazitäten vielen unterschiedlichen Applikationen zugänglich zu machen [79]. Genauer betrachtet, sind jedoch weitere Unterschiede feststellbar, wovon einige in Tabelle 2.1 festgehalten sind.

Kriterium	Grid Computing	Cloud Computing
Applikationstyp	batch	interaktiv
Zugriff	über Grid Middleware	über Standard-Web-Protokolle
Geschäftsmodelle	gemeinsame Nutzung	Preis basiert
Kontrolle	dezentral	zentral (Rechenzentrum), oder dezentral (Cloud-Föderation)
Nutzung	schwierig	einfach

Tabelle 2.1: Unterschiede zwischen Grid und Cloud Computing (angelehnt an [240])

Large-Scale-Applikationen, wie sie z. B. bei wissenschaftlichen Berechnungen eingesetzt werden, benötigen immense Rechen- und Speicherkapazitäten, so dass zu ihrem Betrieb ein Server-Cluster benötigt wird. Ist der Ressourcenbedarf allerdings so hoch, dass ein einzelner Standort nicht ausreicht, muss ein über mehrere Standorte verteilter Cluster gebildet werden. Mitte der 1990er Jahre, inspiriert vom elektrischen Übertragungsnetz (Power Grid), entsprang die Idee eines geografisch verteilten Rechen-Grids (Grid Computing) [40]. Ein Grid ermöglicht damit die gemeinsame Nutzung unterschiedlicher, geografisch verteilter Ressourcen, die darüber hinaus noch skalierfähig sind [240].

Der Nutzer ist jedoch beschränkt auf die bereitgestellten Funktionalitäten. Die Entwicklung benutzerspezifischer Erweiterungen ist nicht möglich [26], geschweige denn die Abarbeitung andersartiger Berechnungsjobs [240]. Beim Cloud Computing ist dagegen grundsätzlich jede Art von Dienst, Plattform oder Infrastruktur beziehbar – soweit sich diese im Portfolio befindet – und daher jeder Berechnungsauftrag möglich. Die angebotenen Dienste sind auch bequem und einfach über Standard-Web-Protokolle zugreifbar, während beim Grid Computing spezielles

Wissen über Schnittstellen der jeweiligen Grid-Plattform notwendig ist.

Ein weiterer Unterschied besteht im Geschäftsmodell. Wie oben bereits erwähnt, können die Services beim Cloud Computing nach Bedarf bezogen werden. Dieses wird dann verbrauchsgenau in Rechnung gestellt. Beim Grid Computing besteht das Geschäftsmodell in der gemeinsamen Nutzung der Ressourcen. Für die Koordinierung und ein faires Ressourcenscheduling wird eine Middleware benötigt. Die Optimierung des Grid-Betriebs erfolgt hinsichtlich der Erhöhung des Rechendurchsatzes, was ein gemeinsames globales Ziel darstellt. Beim Cloud Computing ist der jeweilige Provider selbst für die Optimierung seines Betriebs (Kostensenkung) verantwortlich, unter Berücksichtigung der SLAs ggü. seinen Kunden.

Fehlende Standards

Cloud Computing besitzt allerdings ein großes Problem bzgl. der Interoperabilität: Beim IaaS werden Ressourcen zwar über Standardsoftware bereitgestellt, jedoch dauert die Migration zu einem anderen Provider – identische VMMs vorausgesetzt – aufgrund fehlender Automatismen sehr lange [169]. Auf SaaS- und PaaS-Ebene werden anbieterspezifische Komponenten integriert. Die Portierung und Integration von anderen Entwicklungsmethoden und -werkzeugen stellt die Ausnahme dar [222]. Um Funktionalitäten und Daten über Anbietergrenzen hinweg auszutauschen, müssen spezielle Webservices eingesetzt werden (z. B. der Daten-Backend-Service von Caspio). Das Problem besteht darin, dass keine Cloud Computing Standards für Elemente und Prozesse wie APIs, Speicherung von Server-/VM-Images sowie Datenim- und export existieren [133]. Hierdurch ist ein Vendor Lock-In insbesondere beim SaaS und PaaS wahrscheinlich.

Die schlechte Interoperabilität sei laut [169] dem schnellen Wachstum des Cloud Computings geschuldet, denn dieses hätte die Entwicklung eines ganzheitlichen Standards verhindert. Es haben sich allerdings Gruppierungen gebildet, die Standards vorschlagen bzw. eine Standardisierung vorantreiben: Über den OVF-Standard (Open Virtualization Standard) und IEEE P2301/2 sollen Inter-Cloud-Migrationen wie innerhalb einer Private Cloud ermöglicht werden. Der CDMI-Standard (Cloud Data Management Interface) soll analog die Verschiebung von Daten ermöglichen. Das OCCI (Open Cloud Computing Interface) ist kompatibel zu diesen beiden Standards und ermöglicht so u. a. einen Providerwechsel ohne technische Hürden.

Mit dem Open Cloud Manifesto [168] wurden die Kernprinzipien von Cloud Computing verfasst. Die wichtigsten sind:

- hohe Transparenz und Kompatibilität bereitgestellter Dienste,
- standardisierte Lizenzierung und Verrechnung genutzter Ressourcen,
- Erstellung geeigneter Metriken zur Überwachung der Ressourcenauslastung.

Als weiterer wichtiger Punkt für eine erfolgreiche Standardisierung wird gefordert, dass Provider ihre marktbeherrschende Position nicht dazu ausnutzen, Kunden durch proprietäre

Strukturen an sich zu binden, sondern Interoperabilität zu unterstützen, damit die Angebote unterschiedlicher Provider genutzt und kombiniert werden können.

Kosteneinsparungen im Umfeld von Cloud Computing lassen sich nur bei weitreichender Automatisierung erreichen, welche sich über die gesamte Bandbreite der Management-Dienste wie Migration, Lastmanagement, Provisioning und Billing erstreckt.

Billing

Jeder Provider hat sein eigenes System zur Abrechnung der angebotenen Dienstleistung. Die Abrechnung reicht von Pauschalen für einen gewissen Zeitraum bis hin zu Kosten in Abhängigkeit von der Anzahl an Aufrufen, den Transaktionen und dem Datendurchsatz [76], wodurch nicht nur ein Vergleich, sondern auch z. B. die Bildung eines föderierten Cloud-Systems erschwert wird. In [79] werden die Eigenschaften eines geeigneten Kontoführungssystems vorgestellt. Dieses müsse u.a. in der Lage sein, zeitabhängige Werte wie die Lokalisierung und den Ressourcenbedarf sowie Verletzungen ggü. der SLA einzubeziehen.

Als einer der größten Cloud-Anbieter für Infrastruktur-Dienste bietet Amazon mit seinem Produkt *Elastic Compute Cloud* an, benötigte Rechenkapazität in sehr kleinen Abstufungen schnell an die eigenen Bedürfnisse anzupassen. Dennoch wird immer die gesamte reservierte Kapazität bezahlt und nicht nur der tatsächlich verwendete Anteil. Immerhin orientiert sich der Preis durch die kleinen Recheneinheiten – bei adäquatem Management – nahe an der Nutzung. Hinzu kommen monatliche Speicher- und Transaktionskosten, die sich nach dem transferierten Datenvolumen richten. Dass aber eine genaue verbrauchsabhängige Abrechnung technisch realisierbar ist, zeigt das Produkt IBM Tivoli Usage and Accounting Manager. Mit diesem lässt sich ermitteln, welcher Kunde bzw. welche VM in welchem Maße bestimmte IT-Ressourcen genutzt hat. Eine hiermit realisierte Live-Kostenanzeige, wie in [124] vorgeschlagen, würde die Transparenz erhöhen. Dies würde auch das Vertrauen in das Cloud Computing stärken, was insbesondere beim nächsten Aspekt sehr wichtig ist.

Sicherheit, Kontrolle und Datenschutz

Zu den größeren Hindernissen zur Nutzung von Cloud-Diensten gehören nach CIO Research²⁶ Bedenken hinsichtlich der Sicherheit und des Kontrollverlusts der Daten. Dieses löse Unbehagen aus, da der Betrieb auf Systemen, die nicht von einem selbst kontrolliert werden, ein erhöhtes Potential für unautorisierten Zugriff berge [168]. Um Vertrauen in die Cloud Computing Technologie aufzubauen, müssten nach [124, 123] Kontroll- und Sicherheitsmechanismen eingeführt werden. Dies schließt z. B. eine Zugriffskontrolle über Daten, die Zertifizierung der Sicherheitsmechanismen sowie die Anwendung von Verschlüsselungsmechanismen [222] mit ein.

Hintergrund der Bedenken zum Kontrollverlust ist die gesetzliche Verordnung, die laut [215] besagt, dass diejenigen, die personenbezogene Daten verarbeiten, dafür verantwortlich seien, dass die Verarbeitung gemäß den gesetzlichen Vorschriften geschieht. Diese Verantwortung

²⁶<http://www.cio.de/a/cloud-computing-polarisiert-cios,861652,2>, letzter Zugriff: 08.05.2015

könne nicht einfach weitergegeben werden. Wird die Datenverarbeitung ausgelagert, müsse daher dennoch für die Einhaltung der Vorschriften gesorgt werden. Ein einfaches Instruieren des Outsourcingunternehmens genüge hierbei nicht, sondern es müssten auch Kontrollen stattfinden.

Nach Dror-John Röcher (Computacenter) liegt an der kritischen Haltung mit der Argumentation des Kontrollverlusts jedoch die Fehlannahme zugrunde, dass Unternehmen heutzutage überhaupt noch die Kontrolle über ihre IT hätten.²⁷ Eine ähnliche Auffassung haben Lynn-Kristin Thorenz (IDC) und Karsten Sontow (Trovarit AG), wonach die Aussage, dass Daten im eigenen Haus sicherer seien, zumindest bei KMU unzutreffend sei, denn diese kämen nicht an die hohen, zertifizierten Sicherheitsstandards (z. B. ISO/IEC 27001) der professionellen Rechenzentren heran.²⁸

Die Bedenken haben in der Cloud einen weitreichenden Charakter, denn grundsätzlich können sich Daten, die in der Cloud gespeichert wurden, an einem beliebigen Standort befinden. Hier muss vertraglich geregelt und kontrolliert werden, dass die Daten nur an solchen Standorten vorgehalten werden, die über ein genügendes Datenschutzniveau verfügen. Denn Regierungen mit Gesetzen zum Schutze der Privatsphäre untersagen es beispielsweise, dass bestimmte persönliche Daten aus der Europäischen Union transferiert werden [133], es sei denn, es herrscht in den Ländern ein angemessenes Datenschutzniveau [150]. Dieses lässt sich von der EU-Kommission verbindlich feststellen. Auch Unternehmen in den USA, die sich den Safe Harbor Grundsätzen unterwerfen, werden anerkannt. Wobei aufgrund des Patriot Acts der US-Sicherheitsbehörde ein Datenzugriff ihrerseits rechtlich dennoch möglich und daher die Datenschutzlage zweifelhaft ist.

Auf weitere Aspekte zur Sicherheit in Cloud-Umgebungen gehen die Cloud Security Alliance Top Threats ein und werden durch einen Leitfaden der RSA²⁹ adressiert.

Cloud Computing beschäftigt sich also mit der Realisierung einer ambienten Servicebereitstellung. Die technischen Voraussetzungen hierfür ermöglichen auch das in der Dissertation angestrebte globale LPM. Außerdem werden beim Cloud Computing auch Fragestellungen zum Bezahlsystem, zur Sicherheit sowie zum Datenschutz behandelt, die für das globale LPM gleichermaßen gelten werden, weshalb das Cloud Computing als eine Art Vorreiter angesehen werden kann.

2.2 Verwandte Arbeiten

Das Themengebiet des verteilten Lastmanagements mit der Motivation, standortbedingte, sich unterscheidende Eigenschaften zur Minimierung des Energiebedarfs bzw. der Betriebskosten zu nutzen, ist ein vergleichsweise neues Forschungsgebiet. Cameron (EDS), Banerjee und weitere

²⁷<http://www.cowo.de/a/2363954>, letzter Zugriff: 08.05.2015

²⁸<http://www.cowo.de/a/1934969>, letzter Zugriff: 08.05.2015

²⁹<http://germany.emc.com/about/news/press/2009/20091104-02.htm>, letzter Zugriff: 14.09.2014

gehen in ihren Visionen aus den Jahren 2008/09 [44, 18] von flexiblen Rechenzentrumsarchitekturen aus, die nicht nur bzgl. der erforderlichen Rechenlast, sondern auch bzgl. des Energiebedarfs Agilität bei der IT-Hardware und der Kühl-Infrastruktur zeigen. Anpassungen an äußere Einflüsse wie die Außentemperatur und die regionale Generierung von regenerativen Energien werden ebenfalls bedacht. In [44] wird die Teilnahme am DSM mittels Preissignalen motiviert, auf die durch adäquates globales LPM reagiert werden sollte.

Wie in der nachfolgenden Beschreibung von verwandten Arbeiten hervorgeht, ist die konkrete Ausprägung des globalen LPM unterschiedlich. Abstrakt betrachtet geht es aber darum, lastverursachende Entitäten örtlich wie zeitlich zu planen, um den Energiebedarf bzw. die Kosten zu beeinflussen, wobei u.a. Kapazitätsbeschränkungen zu berücksichtigen sind. Damit handelt es sich um ein kombinatorisches Optimierungsproblem, das wie auch das rechenzentrumsinterne LPM zur Klasse der NP-schweren Problemen gehört.

2.2.1 Klassifizierender Vergleich

Verwandte Arbeiten, die sich mit der Ermittlung des Einsparpotentials eines globalen LPM oder der konkreten Umsetzung beschäftigt haben, fokussieren ein bestimmtes Ziel für die *Optimierung*. Sie unterscheiden sich in der verwendeten *Technologie*, die als Infrastruktur das LPM ermöglicht, dem betrachteten *Lastobjekt*, das zum Management geplant werden kann, und dem *Datenmanagement*, das für die Bereitstellung der von den Diensten benötigten Daten verantwortlich ist. Die zur Berechnung der Einsparung genutzten Rechenzentrumsenergiebedarfsmodelle unterscheiden sich zwar auch, doch lassen sich diese austauschen und stellen keinen Beitrag der Arbeiten selbst dar. Dieser liegt vordergründig im entwickelten *Verfahren* zur Findung einer Lösung für das Optimierungsproblem. In Tabelle 2.2 sind verwandte Arbeiten nach den aufgezählten Kriterien klassifiziert aufgelistet. Nicht aufgelistet sind solche Arbeiten, die sich zwar mit der Optimierung des Energiebedarfs in einem Verbund von mind. zwei Rechenzentren auseinandergesetzt haben, dieses allerdings unter statischen Parameter bzgl. der Standortbedingungen, wie z. B. im Fit4Green-Projekt.

Die hauptsächlichen Optimierungsziele sind die Senkung der Betriebskosten und des Energiebedarfs sowie die Steigerung der Nutzung regenerativer Energien. Die Koordinierung mit einem Smart Grid bzw. die Partizipation beim DSM geschieht hier über Preissignale, ist also ebenso kostenmotiviert. In [152] liegt das Ziel darin, die Belastung des Stromnetzes zu reduzieren.

Als Powermanagement ermöglichende Infrastruktur kommen verteilte Cluster/Grid und Virtualisierung zum Einsatz. Bei verteilten Clustern/Grid wird jeweils ein größerer Dienst betrieben. Die hierbei betrachteten Lastobjekte sind entweder Anfragen oder Tasks bzw. Jobs. In [173] wird hiervon abweichend Workload verteilt. Hierbei geht es nicht um konkrete Lastobjekte, sondern die Gesamtlast wird als solches – theoretisch – verteilt. Wird Virtualisierung eingesetzt, ist das Lastobjekt grundsätzlich eine VM. Da VMs jedoch nicht nur für den Betrieb einzelner Dienste, sondern auch als virtualisierte Server als Teil eines Clusters eingesetzt werden können, kommen als Lastobjekte hier auch Tasks vor.

Arbeit	Optimierung	Technologie	Lastobjekt	Datenmanagement	Verfahren
[182]	Kosten	verteilter Cluster	Anfrage	vollständige Replikation	n. a.
[138]	Kosten	verteilter Cluster	Anfrage	keine Betrachtung	MILP, $O(\exp \#SRV + \exp \#DC)$
[187]	Kosten	verteilter Cluster	Anfrage	keine Betrachtung	Angepasstes Verfahren mit $o(\#DC^2 \cdot \#SRV^4)$
[196]	Kosten	verteilter Cluster	Anfrage	keine Betrachtung	direktes Routing, keine Planung, $O(1)$
[38]	Kosten	verteilter Cluster	Batchjob (je Server)	keine Betrachtung	Heuristik, $O(\#DC^2)$
[101]	Kosten	n. a. (verteilter Cluster)	Task	keine Betrachtung	quadratische Programmierung, $O(\exp \#DC)$
[33]	Kosten	Virtualisierung	VM	Replikation: Storagebedarf minimiert durch Deduplikation	n. a., nur Konzept
[88]	Kosten, regenerative Energien	verteilter Cluster	Anfrage	keine Betrachtung	quadratische Programmierung, $O(\exp \#DC)$
[250, 140]	regenerative Energien	verteilter Cluster	Anfrage	keine Betrachtung	ILP, $O(\exp \#DC)$
[164]	regenerative Energien	Virtualisierung	VM	keine Betrachtung	n. a., nur Konzept
[173]	Energie (als Beispiel)	Grid / verteilter Cluster	Workload	keine Betrachtung	Formel für Workloadzuordnung, $O(\#DC)$
[43]	Energie (als Beispiel)	Virtualisierung	VM, Tasks darin	keine Betrachtung, da Beispiel nur für ein Rechenzentrum	Greedy (First-Fit-Decreasing), $O(\#VM \cdot \log \#VM)$
[152]	Stromnetzbelastung	verteilter Cluster	Anfrage	vollständige Replikation	lineare Programmierung, $O(\exp \#DC)$

Tabelle 2.2: Verwandte Arbeiten zum Themengebiet des globalen LPM, ohne Wertung entsprechend den Kriterien sortiert

Zum Datenmanagement äußern sich die meisten Arbeiten nicht. Theoretisch sind hierbei unterschiedliche Realisierungen möglich. Da jedoch für gewöhnlich die Zusicherung von SLOs aufrecht erhalten bleiben soll, ist hier von vollständiger Replikation auszugehen, zumal diejenigen, die sich zum Datenmanagement äußern, von eben dieser ausgehen. Die Arbeit in [33] stellt in diesem Kontext als einzige ein Verfahren zur Minimierung der Replikationskosten vor: Der Storagebedarf wird dadurch minimiert, dass die Replikate von VMs mit ähnlichen Daten geclustert werden. Hierdurch kann via Deduplikation der Speicherbedarf reduziert werden (näheres hierzu siehe Abschnitt 4.8). Es wird hier allerdings nicht auf ein konkretes globales LPM eingegangen, sondern lediglich konzeptionell die Anwendbarkeit beschrieben.

Auch in [164] wird nur das Konzept eines globalen LPM vorgestellt. In dem Paper werden vielmehr die ermöglichende Infrastruktur und das hierzu entwickelte Framework beschrieben. In [173] ist aufgrund der dortigen Fragestellung kein Verfahren notwendig. Der Hauptbeitrag der Arbeit liegt in der Vorstellung eines Energieeffizienzkoeffizienten für ein Rechenzentrum zum Betreiben von Lasten. Die Zuordnung des Workloads zu den Rechenzentren wird mit Hilfe einer Formel ermittelt.

Die Arbeit von Qureshi und weiteren [182] ist eine der ersten Arbeiten, in der das Potential eines standortübergreifenden LPM angewendet wird, um die standortbedingten Strompreise gezielt zu nutzen. Ein einfaches Energiebedarfsmodell bestehend aus einem Servermodell und einem PUE-Wert zur Berücksichtigung der Infrastruktur dient zur Ermittlung des Energiebedarfs. In der Arbeit werden Strompreisverläufe des US-amerikanischen Marktes evaluiert. Das eingesetzte globale LPM wird aber auch hier nicht erläutert.

In [187, 196] werden ähnliche Untersuchungen, aber mit anderen Last- und Strompreisverläufen durchgeführt. Während in [187] eine Kostenminimierung aufgrund des eingesetzten Energiebedarfsmodells durchgeführt wird, wird in [196] die Entscheidung, welches Rechenzentrum eine Anfrage zu bearbeiten hat, ausschließlich vom Strompreis abhängig gemacht, was eine deutlich schnellere Laufzeit aufweist. Im Gegensatz hierzu steht der Ansatz aus [138], bei dem zur Lösungsfindung ein Mixed Integer Linear Program (MILP) verwendet wird. Dieses besitzt eine exponentielle Laufzeitkomplexität bzgl. der Anzahl an Servern ($\#SRV$) und Rechenzentren ($\#DC$) und weist damit keine gute Skalierbarkeit auf. Gleiches gilt für die Verfahren aus [101, 88, 250, 140], auch wenn als laufzeitbeeinflussende Variable die Serveranzahl entfällt. Im Gegensatz zu den anderen Verfahren wird in [101] für die Strompreise nicht nur der Day-Ahead-Preis betrachtet, sondern zusätzlich eine gewisse Strafgebühr beim Überschreiten des auf dem Day-Ahead-Preis basierten angenommenen Verbrauchs.

Mit ebenso exponentieller Laufzeitkomplexität, aber durch die Formulierung als lineares Programm (LP) in der Praxis deutlich schneller lösbar, ist das Verfahren aus [152]. Durch den Einsatz von Heuristiken kann eine bessere Laufzeitkomplexität erzielt werden. Die in [38] vorgestellte Heuristik ist verteilt ausführbar, wodurch die Rechenlast je Rechenknoten nochmals reduziert werden kann. Die geringste Laufzeitkomplexität weist, abgesehen von [196], das Verfahren aus [43] auf – hier in Abhängigkeit von der VM-Anzahl ($\#VM$). Allerdings bezieht sich dieses Verfahren auf ein einzelnes Rechenzentrum. Durch Kombination mit dem

Energieeffizienzkoeffizienten aus [173] kann jedoch ein globales LPM konzipiert werden, das die Allokation der VMs am effizientesten Rechenzentrum beginnt. Die Laufzeitkomplexität wäre hierdurch unverändert. Diese Realisierung hätte jedoch einen sehr ausgeprägten Greedy-Charakter und würde höchstwahrscheinlich vergleichsweise schlechte Ergebnisse liefern, wie die Untersuchung aus [101] vermuten lässt, denn das dort eingesetzte genaue Verfahren hat gegenüber einer Greedy-Variante eine 14-fach höhere Einsparung erzielen können.

2.2.2 Abgrenzung

In diesem Abschnitt werden die bisher von den verwandten Arbeiten nicht adressierten bzw. nur unzulänglich behandelten Problemstellungen herausgearbeitet. Hieraus werden die eigenen Ziele der Dissertation formuliert, die gleichsam als Abgrenzung dienen und bereits in Abschnitt 1.2 zusammenfassend beschrieben wurden.

Technische und regulatorische Voraussetzungen

Ein globales LPM, insbesondere in einem Verbund von Rechenzentren unterschiedlicher Betreiber, weist eine Reihe von Problemen auf, die erst noch gelöst bzw. durch politische Vorgaben geregelt werden müssen. Zum einen muss eine Standardisierung u.a. der Leistungsbereitstellung und ihrer Verrechnung erfolgen – Cloudanbieter wie Amazon, Rackspace, Salesforce.com etc. besitzen unterschiedliche Verrechnungsmodelle – und zum anderen stellt die Gewährleistung von Datensicherheit bzw. Datenschutz eine große Herausforderung dar. Fehlende Kontroll- und Handlungsmöglichkeiten erschweren zudem die Vermeidung und Auflösung von Servicestörungen. Hierbei handelt es sich um Randaspekte, die eine Voraussetzung zur Nutzung virtualisierungsbasiertem globalen LPM bilden. In Abschnitt 2.1.4 wurde bereits auf diese näher eingegangen. Für den eigentlichen Optimierungsprozess im Sinne der Energie- und Kosteneinsparung sind sie jedoch von minderer Bedeutung und werden daher hier und auch in den verwandten Arbeiten nicht weiter berücksichtigt.

Technologie und Lastobjekt

In dieser Arbeit werden Server-Sized-Services betrachtet, also solche Services, die auf einem einzelnen Server betrieben werden können (siehe Abschnitt 1.2). Diese werden jeweils in einer VM gekapselt und sind mit Hilfe der Live-Migration reallozierbar. Jeder Service (Dienst) bzw. jede Applikation läuft kontinuierlich und weist einen zeitvarianten Ressourcenbedarf auf. Das heißt, dass der Betrieb einer VM einen bestimmten Auslastungsverlauf und damit auch einen individuellen Energiebedarfsverlauf generiert. Dieser Sachverhalt unterscheidet sich von den verschiebbaren Lastobjekten der anderen Arbeiten, denn dort werden entweder finite Bearbeitungsprozesse (in Form von Tasks oder durch Anfragen ausgelöst) oder aber VMs betrachtet, welche eine ebenso begrenzte Lebenszeit oder einen kontinuierlich konstanten Ressourcenbedarf aufweisen. Das hier betrachtete verschiebbare Lastobjekt zum Management

des Workloads und des Leistungsbedarfs weist also eine aufwendiger zu behandelnde Form auf.

Modell des eingesetzten LPM

Wie bei den Maßnahmen zur Verbesserung der Energieeffizienz, z. B. durch den Einsatz von neuen Kühlungstechniken, müssen auch die Auswirkungen von Lastverschiebungen im Vorfeld bekannt sein. Die Höhe der prognostizierten Einspareffekte ist wesentlich, da eine Lastverschiebung i.d.R. mit Aufwand verbunden ist, den es erstmal zu kompensieren gilt.

In den oben genannten Arbeiten werden die Lastverlagerungen im Vorfeld analysiert, so dass die Auswirkungen und damit die Einspareffekte eingeplant werden können. Hierzu werden Energiebedarfsmodelle für Server, Racks und die Infrastruktur – Klimatisierung und evtl. auch Stromversorgung, teilweise allerdings nur vereinfacht als PUE oder „Kühleffizienzfaktor“ – herangezogen. In Abschnitt 2.1.1 wurden Arbeiten zu detaillierteren Rechenzentrumsmodellen erwähnt, mit denen eine genauere Analyse möglich ist. Was jedoch fehlt, ist die Berücksichtigung der Auswirkungen eines LPM insbesondere bei Verwendung dynamischer Virtualisierung.

Suzanne Rivoire und weitere [191] haben herausgestellt, dass Modelle notwendig sind, um die Effekte von dynamischem Power Management abschätzen zu können. Je höher der betrachtete Systemlevel ist, desto größer sind auch die potentiellen Einsparmöglichkeiten. Die Erstellung eines Modells, das die Auswirkungen dynamischer Virtualisierung kenntlich macht, ist somit von großer Bedeutung.

Dass sich der Energiebedarf aufgrund unterschiedlicher Policies für ein LPM verändert, wurde bereits in [90] für Job-Scheduling in einem Server-Cluster gezeigt. Um die Auswirkungen eines LPM offenzulegen, wurden daher in [159] die Effekte eines Job-Schedulers analytisch betrachtet. Für virtualisierte Serverumgebungen, in denen die Granularität des Lastmanagements auf VM-Level mittels Live-Migration besteht, ist dieser Ansatz jedoch nicht ohne Weiteres anwendbar. Diesen Anwendungsbereich betrachten Pelley und weitere [175]. Das dort vorgestellte Rechenzentrumsmodell geht auf die Effekte des LPM auf abstrahierte Weise ein. Es wird kein spezielles Verfahren zugrunde gelegt, sondern es wird angenommen, dass das eingesetzte LPM grundsätzlich eine bestimmte Konsolidierungsrate erzielt. Ob diese Annahme für ein LPM von VMs zutrifft, wird weder diskutiert noch anhand von Beispielen belegt.

Mit VMware Capacity Planner ist ein kommerzielles Produkt verfügbar, mit dem es zwar möglich ist, die benötigten Ressourcen zum Betrieb von VMs zu bestimmen, allerdings berücksichtigt das Programm nicht den Einsatz eines LPM, nicht einmal das firmeninterne vSphere DRS (Distributed Resource Scheduler) bzw. DPM (Distributed Power Management). Daher kann es auch nur zur Abschätzung des Energiebedarfs für die statische Virtualisierung herangezogen werden.

In der Dissertation sollen daher die Auswirkungen eines LPM bei Verwendung dynamischer Virtualisierung modelliert werden.

Netzwerkkosten

Die Migration einer VM über ein geografisch entferntes Netzwerk verursacht einen erhöhten Energiebedarf an mehreren Routern, bedingt auch durch die Leitungen. Sind die zu übertragenden Daten sehr groß, können die Netzwerkkosten schnell zum Angelpunkt werden [17], da sie den Einspareffekt durch die Migration überkompensieren und damit eine globale Lastverlagerung demotivieren können. Der Lastverlagerungsaufwand stellt also einen nennenswerten Faktor dar, so dass ein geeignetes Netzwerkenergiebedarfs- bzw. -kostenmodell eingebunden werden muss.

In [173] werden die Kosten durch eine Gewichtung der Energieeffizienzkoeffizienten behandelt. Eine Lastverlagerung wird dadurch also durch die erhöhten Kosten weniger attraktiv. [38] geht einen ähnlichen Weg. Dort wird ein Bandbreitenkostenquotient und ein Einsparungsfaktor verwendet, welche entgegengesetzt wirken. Über einen zusätzlichen Parameter kann der Greedy-Wert eingestellt werden, wie sehr die Einsparungen erzielt und dabei die Bandbreitenkosten vernachlässigt werden sollen. In [182] wird diesbezüglich die 95/5-Billing-Methode³⁰ berücksichtigt, d.h., dass die bisherige Netzwerkbandbreite nur insofern erhöht wird, solange das bisherige 95-Prozentil bestehen bleibt und damit die gleichen Konditionen beim Provider gelten. Insofern werden die Kosten berücksichtigt, der real verursachte, zusätzliche Energiebedarf jedoch nicht. In den anderen Arbeiten wird der Aspekt der Migrationskosten gänzlich ignoriert.

In der Dissertation soll auf den konkret anfallenden Energie- und Kostenaufwand eingegangen werden, unabhängig vom verwendeten Provider und Tarif.

Datenmanagement – Replikation

Die Migration von VMs zwischen Rechenzentren weist ein weiteres Problem auf. Neben dem Arbeitsspeicher benötigen Dienste i.d.R. Zugriff auf eine Datenbasis, welche im Storage des betreibenden Rechenzentrums bereitgehalten wird. Im Gegensatz zur rechenzentrumsinternen VM-Migration existiert bei der rechenzentrenübergreifenden jedoch kein zentrales Storage, über das auf die hinterlegten Daten ohne Verzögerung zugegriffen werden kann. Da diese Daten für gewöhnlich deutlich größer sind als der benötigte Arbeitsspeicher, würde das Kopieren des Datenbestandes zum einen höhere Netzwerkkosten verursachen und zum anderen zu einer deutlich höheren Migrationszeit führen, was die Flexibilität, um auf sich verändernde Energiebedarfs- und Kostenparameter zu reagieren, verringert. Dies würde also auch die Betriebskosten erhöhen. Daher werden die Daten im Zielrechenzentrum bereits vor der Migration der VM benötigt. Diese Notwendigkeit wird im Folgenden kurz diskutiert.

Wird die Datenhaltung einer VM auf einen gesonderten Standort festgelegt, so ist der Datenzugriff nur dann verzögerungsfrei, wenn die VM auch an diesem Standort betrieben wird. Wird sie an einen anderen Standort migriert, ist mit einem reduzierten Durchsatz zu rechnen.

³⁰In 5-Minuten-Abständen wird der Bandbreitenbedarf gemessen. Die 5% höchsten Werte werden gestrichen. Der verbleibende Höchstwert (95-Prozentil) wird zur Abrechnung beim Netzprovider herangezogen.

Dies kann vermieden werden, indem die Datenbasis mit an den Zielstandort kopiert wird. Allerdings ist dieses wiederum mit hohen Migrationszeiten verbunden. Einen Kompromiss stellt eine nachgelagerte Kopie dar, wie sie in [105] vorgestellt wird. Die Daten werden dabei erst nach Abschluss der Migration kopiert. Die VM könnte damit sofort am Zielstandort betrieben werden, wodurch die Migrationszeit minimal wäre. Je nach Kopierstrategie ist allerdings von einem anfänglichen reduzierten Datendurchsatz auszugehen.

Die komplette Kopie zur Migrationszeit beschränkt die Flexibilität des globalen LPM aufgrund der Verlängerung der Migrationszeit, die anderen beiden Varianten reduzieren den Datendurchsatz bzw. können den höchstmöglichen nicht garantieren. Dies könnte zur Verletzung von SLOs führen oder aber es wird grundsätzlich ein geringerer Durchsatz vereinbart, was jedoch die Lukrativität mindert. Um sowohl höchstmöglichen Datendurchsatz und größtmögliche Flexibilität zu gewährleisten, können die Daten repliziert werden. Zugleich wird dadurch auch die Verfügbarkeit der Daten erhöht [239], worauf in dieser Arbeit aber nicht näher eingegangen wird. Die erhöhte Flexibilität wird allerdings durch höhere Kosten für den Datentransfer erkaufte.

Da ein Verfahren zum globalen LPM nur solche Standorte heranziehen kann, an denen auf die benötigten Daten zugegriffen werden kann, muss eine Replikationsstrategie entwickelt werden, die geeignete Standorte ermittelt. In den verwandten Arbeiten wird von vollständiger Replikation ausgegangen oder das Datenmanagement außer Acht gelassen. Da aber auch im letzten Fall keine Einbußen beim Betrieb der Dienste entstehen, wird auch hier von vollständiger Replikation ausgegangen. Dies ermöglicht dem globalen LPM, mit höchstmöglicher Flexibilität die Lasten zu verteilen.

In [201, 95] wird gezeigt, dass das Hinzufügen von Replikaten bzgl. des Rechendurchsatzes einem asymptotischen Verlauf folgt. Der Aufwand für die Synchronisierungsoperationen nimmt nämlich überlinear zu. Daher wird zur Optimierung der Rechenleistung eine partielle Replikation empfohlen. Analog ist dieser Sachverhalt auch für den Einspareffekt denkbar, was sich in den durchgeführten Untersuchungen (siehe Abschnitt 4.8) auch bestätigt. Allgemein ist festzuhalten, dass die Flexibilität des globalen LPM mit steigender Anzahl an Replikaten zunimmt und damit die Kosten zum Betrieb der VMs sinken. Die Kosten für die Replikation steigen dabei jedoch. Zur Optimierung dieses Trade-Offs sollte daher ein Verfahren zur Bestimmung einer geeigneten Anzahl und Lokalisierung von Replikaten in den Gesamtoptimierungsprozess integriert werden.

Auch beim Data Grid, das im Gegensatz zum normalen Grid Computing viel Datenzugriff besitzt, ist eine gezielte Minimierung des Datentraffics ratsam [229]. Entsprechende Arbeiten werden in Abschnitt 4.8 bzgl. der vorliegenden Problemstellung diskutiert.

Optimierungsbereich

Als weitere Einschränkung ist bei den bisherigen Arbeiten zu erkennen, dass die entwickelten Verfahren ihre Optimierung auf jeweils nur ein Kriterium (Kosten, Energie, Lastgang, CO_2 -Emissionen etc.) beschränken. Zwar lässt sich eine Übertragung auf jeweils ein anderes

Kriterium sicherlich mit relativ geringem Aufwand erreichen, da aber in der Praxis immer alle Kriterien von Bedeutung sind, sollte ein Mechanismus zur Berücksichtigung aller Werte – gleichzeitig – geschaffen werden.

Skalierbarkeit

Das in den verwandten Arbeiten angegebene vollständige Verfahren für ein globales LPM mit der geringsten Laufzeitkomplexität ist eine Heuristik mit einer polynomiellen Laufzeit bzgl. der Anzahl an Rechenzentren. Dies erscheint aufgrund der Unabhängigkeit von der Anzahl an betriebenen Batchjobs bzw. Servern als sehr effizient. Mit Ausnahme von [38] ist jedoch allen verwandten Arbeiten ihre zentrale Ausführung gemein, was abgesehen vom Single-Point-of-Failure, der in dieser Arbeit nicht betrachtet wird, die Verantwortung für das Management in eine Hand legt. In einem Verbund gleichwertiger Rechenzentrumspartner wäre eine gleichgestellte Beteiligung aller an der Betriebsoptimierung wünschenswert. Aus diesem Grunde wird ein verteiltes Verfahren anvisiert, das zusätzlich maximal eine polynomielle Laufzeitkomplexität bzgl. der Skalierungsparameter Rechenzentrums- und VM-Anzahl aufweist.

Allgemeingültigkeit der Ergebnisse

Neben den unterschiedlichen Verfahrensaspekten ist festzustellen, dass sämtliche durchgeführten Simulationen und Einsparpotentialermittlungen lediglich auf einem überschaubaren Satz an Szenarien aufbauen, wodurch die ermittelten Potentiale auf eben diese Untersuchungen begrenzt sind und was somit die Allgemeingültigkeit der Verfahren in Frage stellt. Hier wäre es wünschenswert, eine allgemeine Aussage geben zu können. Aus diesem Grunde wird in dieser Arbeit eine große Ansammlung unterschiedlicher Szenarien evaluiert, indem die untersuchten Datensätze für Workload, Temperatur und Energiepreise variiert werden. Außerdem wird eine Sensibilitätsanalyse durchgeführt, um herauszufinden, welche Auswirkungen die Veränderung bestimmter Parameter auf das erzielbare Ergebnis hat.

2.3 Zusammenfassung

In diesem Kapitel wurde ein Überblick über den technologischen Kontext der Dissertation gegeben sowie verwandte Arbeiten diskutiert. In der zu diesen durchgeführten Abgrenzung hat sich herausgestellt, dass es große Unterschiede im Betrachtungsraum gibt und dass einige Problemstellungen bisher nur ungenügend oder gar nicht behandelt wurden. Aus diesen wurden die Ziele der Dissertation herausgearbeitet.

3 Technische Voraussetzungen

Die adressierte Problematik und die Eigenschaften des zu entwickelnden Systems wurden in den vorangegangenen Abschnitten konkretisiert. In diesem Kapitel werden die technischen Voraussetzungen für den Optimierungsvorgang in einem solchen System vorgestellt.

Motiviert werden die globalen Migrationen aufgrund von Veränderungen der Standortbedingungen. Diese müssen rechtzeitig im Voraus bekannt sein, um entsprechende Optimierungen durchführen zu können. Hierzu sind Prognoseverfahren erforderlich, auf die im Folgenden eingegangen wird. Des Weiteren wird das Konzept eines Konsolidierungsverfahrens vorgestellt, auf dessen Basis das verwendete lokale LPM realisiert wird.

3.1 Datenbasis & adäquate Vorhersagemodelle

Die Effizienz eines LPM hängt nicht zuletzt davon ab, auf welche Daten dieses zurückgreifen kann. Werden zur Analyse ausschließlich historische Werte herangezogen, so handelt es sich um ein reaktives Verfahren, das, wie bereits erläutert, ein geringeres Einsparpotential aufweist als ein proaktives Verfahren. Durch die Verwendung von Prognosewerten können Systemveränderungen mit weitgreifenden Operationen im Vorfeld vorgenommen werden. Dies setzt jedoch eine adäquate Qualität dieser Daten voraus, d.h., dass die Vorhersagen mit hoher Wahrscheinlichkeit eintreten sollten. Denn sind die Abweichungen zu groß, können die vorgenommenen Veränderungen gar zur Verschlechterung des Systems im Sinne der Kosteneinsparung führen.

Die Daten, auf denen die Optimierungen durchgeführt werden, unterteilen sich in zwei Kategorien: die dynamischen und die statischen (abgeleiteten) Daten. Letztere werden durch das System gebildet und geben das Systemverhalten wieder. Hiermit ist z. B. der Energiebedarf eines Servers gemeint. Dieser verändert sich zwar i.d.R. über die Zeit, allerdings liegt dies an den variierenden Rechenlasten, die durch die betriebenen Dienste verursacht werden. Bei gleicher Rechenlast hat ein Server unter gleichen äußeren Bedingungen den gleichen Energiebedarf. Die durch die Dienste verursachten variierenden Rechenlasten gehören zu den dynamischen Daten.

Anhand der statischen Daten kann ein Systemmodell erstellt werden, aus dem – mit den dynamischen Daten aufgerufen – das Systemverhalten ermittelt werden kann. Für die Modellierung eines Systems existieren unterschiedliche Ansätze und Verfahren. Die Güte eines Modells kann durch Abgleich mit der Realität in Form von Messdaten angegeben werden. Diese Güte ist statisch, also allgemein gültig (von unterschiedlichen Abweichungen je nach

Arbeitspunkt abgesehen). Auf die Güte der hier entwickelten und verwendeten Systemmodelle wird in Kapitel 4 eingegangen.

Dieser Abschnitt beschäftigt sich mit den dynamischen Daten, die zur Optimierung des Systems benötigt werden: Ressourcenbedarf von Diensten sowie Außentemperatur und Strompreise unterschiedlicher Standorte. Für diese sind Vorhersagen notwendig, wobei i.d.R. bei jedem Vorhersagemodell gilt: Je weiter der Blick in die Zukunft ist, desto geringer ist die Zuverlässigkeit der Vorhersage.

3.1.1 Vorhersagemethodik für den Ressourcenbedarf von Diensten

Es existieren viele Arbeiten, die sich mit der Modellierung des Ressourcenbedarfsverhaltens von Serverdiensten auseinandergesetzt haben. Die für die dort durchgeführten Analysen zugrundeliegenden Daten beziehen sich allerdings fast ausschließlich auf (very-)large-scale HTTP-Dienste¹. In [109, 110, 108] wurden Ressourcenbedarfszeitreihen von auf Virtualisierung basierten Rechenzentren selbst aufgenommen und analysiert, um ein Verfahren zu entwickeln, das eine Prognose des Ressourcenbedarfs eines Dienstes in einem Zeitraum von mehreren Stunden bis Tage ermöglicht.

Gemäß konventioneller Zeitreihenanalyse nach der Box-Jenkins-Methode (Identifikationsphase [221]) wurde zunächst eine grafische Darstellung einer Auswahl der ermittelten Ressourcenbedarfszeitreihen aller Ressourcentypen vorgenommen, was Folgendes ergab:

- Es liegt mitunter ein Langzeittrend vor, der z. B. einen linearen Verlauf aufweist.
- Häufig ist ein saisonaler Trend zu erkennen, der durchaus auch aus einer Überlagerung mehrerer Perioden bestehen kann. Häufige Perioden sind: 1, 2, 7, 14 Tag(e).
- Bereinigt um die Trends, bleibt ein als zufällig zu betrachtender Rauschanteil übrig. Dieser stellt im Vergleich zum Trend einen signifikanten Anteil dar.

Das ausgeprägte periodische Verhalten des saisonalen Trends legt eine periodische Modellierung des Verhaltens nahe. Klassische Ansätze trennen hierbei den saisonalen Trend vom Rauschen. Das Rauschen wird hierbei durch einen stochastischen stationären Prozess modelliert. Aus der obigen Analyse ging jedoch auch hervor, dass der Rauschanteil nicht-stationär ist und eine getrennte Modellierung damit keinen Sinn macht. Daher werden der saisonale Trend und das Rauschen gemeinsam durch einen stochastischen aber nicht zwingend stationären Prozess, wie in Abbildung 3.1 skizziert, modelliert.

Um einen betrachteten Zeitpunkt t_0 herum wird ein gewisser Zeitraum Δt^{avg} als stationär und alle gemessenen Werte für den Bedarf R einer Ressource i in diesem Zeitraum als statistisch unabhängig betrachtet. Aus dem umliegend beobachteten Verhalten wird eine Zufallsvariable abgeleitet, die das Verhalten zum betrachteten Zeitpunkt beschreibt. Diese Zufallsvariable wird durch Maximumbildung überabgeschätzt, um die benötigte Ressourcenkapazität abzuleiten.

¹ACM SIGCOMM. The Internet Traffic Archive. <http://ita.ee.lbl.gov/>, letzter Zugriff: 08.09.2014

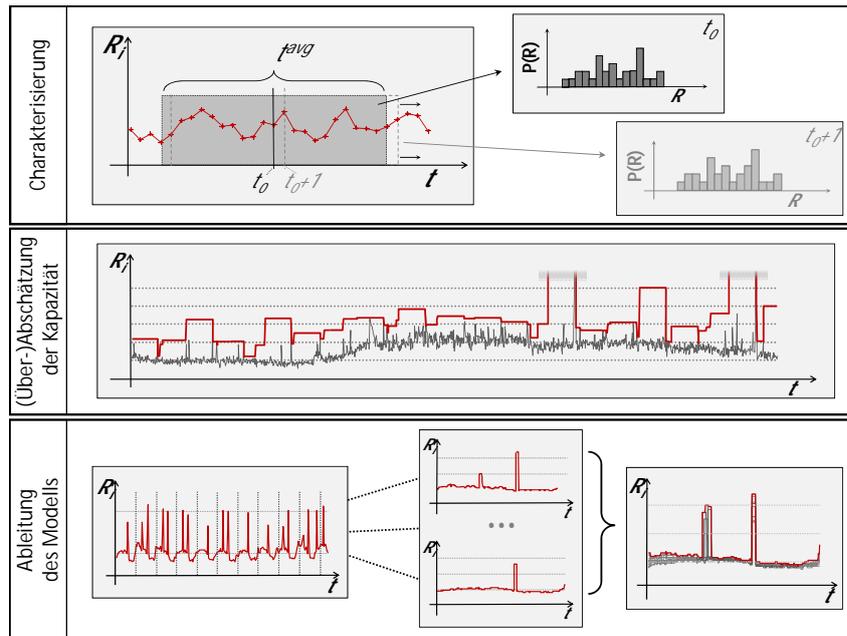


Abbildung 3.1: Die drei Ebenen des Modellierungsprozesses zur Prognose des Ressourcenbedarfs von Diensten nach [108]

Ggf. kann eine Reservekapazität aufgeschlagen werden. Das Modell wird dann durch Zerlegung der Zeitreihe in Instanzen einer dominierenden Periode und Maximumbildung eines jeden Zeitpunkts aller Instanzen abgeleitet. Die Erkennung von Perioden erfolgt mittels einer Autokorrelationsanalyse.

Der Langzeittrend sollte mit Methoden aus der Kapazitätsplanung (siehe [8]) behandelt werden. Entsprechend den hieraus generierten Abschätzungen muss eine Reservekapazität (zusätzliche Server) bereitgestellt werden. Auf diese kann das eingesetzte LPM bei steigendem Trend zurückgreifen. Das Vorhersageverfahren setzt somit auf einer um einen möglichen Langzeittrend bereinigten Zeitreihe auf.

Abweichungen vom prognostizierten Wert kommen nicht nur bei steigendem Trend vor, sondern können auch sporadisch auftreten. In diesen Fällen muss das Modell entsprechend korrigiert werden. In ungünstigen Konstellationen kann es jedoch vorkommen, dass ein Dienst zwar mehr Ressourcen benötigen würde, diese aber nicht bekommen kann, da die Kapazität des Servers zum Zeitpunkt bereits erschöpft ist. Ein von der Modellprognose höherer Wert lässt sich dann nicht messen. Allerdings entsteht hierbei ein Performanzproblem, welches durch ein Programm zur Performanzüberwachung (z. B. Munin oder Nagios) detektiert werden kann.

In einer Evaluation dieser Vorhersagemethodik in [108] konnte gezeigt werden, dass bei weniger als 3% der Fälle der Ressourcenbedarf unterabgeschätzt wurde. In diesen Fällen tritt also ein höherer als prognostizierter Wert auf. Das bedeutet, dass ein Ressourcenmanagement in 97% aller Fälle die Ressourcen vollständig hätte verplanen können, ohne dass Engpässe auftreten. Diese Güte hat, abgesehen von den oben genannten Fällen zur Modellanpassung, keine weitere, insbesondere keine zeitliche Beschränkung.

3.1.2 Wettervorhersage

Die Umweltparameter besitzen einen nicht zu vernachlässigenden Einfluss auf den Energiebedarf eines Rechenzentrums. Für die Realisierung eines globalen LPM muss daher ein genaues Umweltmodell, welches alle notwendigen Parameter und Vorhersagen über die zukünftige Entwicklung in allen Standorten enthält, herangezogen werden.

Den wichtigsten Parameter bei der Wettervorhersage stellt die Temperatur dar, da von ihr der Energiebedarf der Kühlsysteme abhängt, vor allem wenn zusätzlich zur aktiven Kühlung eine energiesparende Außenluftkühlung installiert ist. Für die Kühlung mit Hilfe eines Kühlturms, welcher die Wärmeenergie durch Verdunstung von Wasser an die Luft abgibt, ist die Luftfeuchtigkeit ein weiterer wichtiger Parameter, da trockenere Luft Wasser schneller aufnehmen kann. Weitere Daten zu Windstärke, Sonnenstunden und Sonnenintensität sind dann relevant, wenn die Nutzung von regenerativen Energien wie Windkraft und Sonnenenergie eine Rolle spielt.

Wettervorhersagen besitzen auch heutzutage noch keine absolute Treffsicherheit. Die Gründe hierfür liegen in einer wegen zu grober Messung unvollständigen Datenbasis über das atmosphärische Verhalten, so dass das Wetter einen gewissen Anteil an nicht vorhersehbarrem Verhalten besitzt. Außerdem reicht die aktuelle Rechenkapazität noch immer nicht dafür aus, die für noch genauere Vorhersagen extrem detaillierten Modelle zu verarbeiten. Aus diesem Grunde sinkt die Treffsicherheit der Vorhersagen ab vier Tagen in die Zukunft erheblich und eine Langzeitprognose des Wetters von mehr als zehn Tagen, eine sogenannte Witterungsprognose, gibt nur noch einen Trend an.

Prognostizierte Wetterdaten liefert z. B. der Deutsche Wetterdienst. Mit Hilfe der erhaltenen Wettervorhersagen und einem Modell für mögliche Abweichungswahrscheinlichkeiten lässt sich eine statistische Vorhersage geben. Ein Beispiel eines Modells zur Vorhersage der Temperaturabweichung ist in Abbildung 3.2 dargestellt. Die Höhe der Balken gibt für eine Vorhersage von bis zu zehn Tagen an, wie wahrscheinlich die jeweilige Temperaturabweichung in Kelvin ist. Die angegebenen Wahrscheinlichkeiten basieren auf empirisch ermittelten Abweichungen

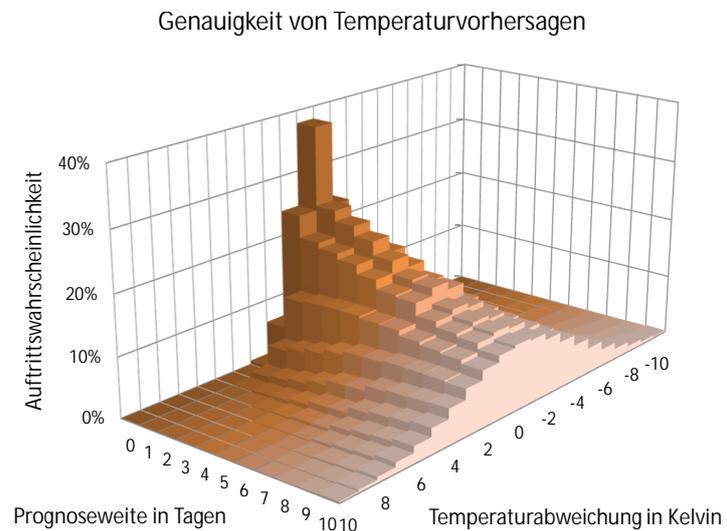


Abbildung 3.2: Beispielmodell der möglichen Temperaturabweichungen in Kelvin bei einer Vorhersage von bis zu zehn Tagen: Je weiter Prognosen in die Zukunft reichen, desto größer sind die Wahrscheinlichkeiten für größere Abweichungen.

von prognostizierten Temperaturen aus einer Studie der Freien Universität Berlin ([73]). In diesem Beispielmodell werden nur ganzzahlige Temperaturabweichungen betrachtet. Anhand des Diagramms ist gut erkennbar, dass für längerfristige Vorhersagen die Wahrscheinlichkeiten größerer Temperaturabweichungen ansteigen: Während eine Vorhersage bis zu 24 Stunden noch mit einer Wahrscheinlichkeit von 86 % zutrifft (± 1 Kelvin), besitzt eine 3-tägige Vorhersage nur noch eine 58 %-ige Genauigkeit.

3.1.3 Strompreise

Die Strompreise haben in Europa ein unterschiedliches Niveau. In Abbildung 3.3 sind die durchschnittlichen Preise bei einem Jahresverbrauch von 2500 kWh bis 5000 kWh für das zweite Halbjahr 2013 angegeben. Diese können, wie in Abbildung 1.8 dargestellt, auch sich überschneidende Verläufe aufweisen, so dass sich – im Durchschnitt betrachtet – teurere Standorte zu gewissen Zeitpunkten als betriebskostenminimal herausstellen. Der eingetragene mittlere Tageswert kann im Verlauf eines Tages noch deutlich stärker schwanken. So ist es möglich, dass er sich innerhalb weniger Stunden um das 9-Fache verändert (siehe Abbildung 1.7).

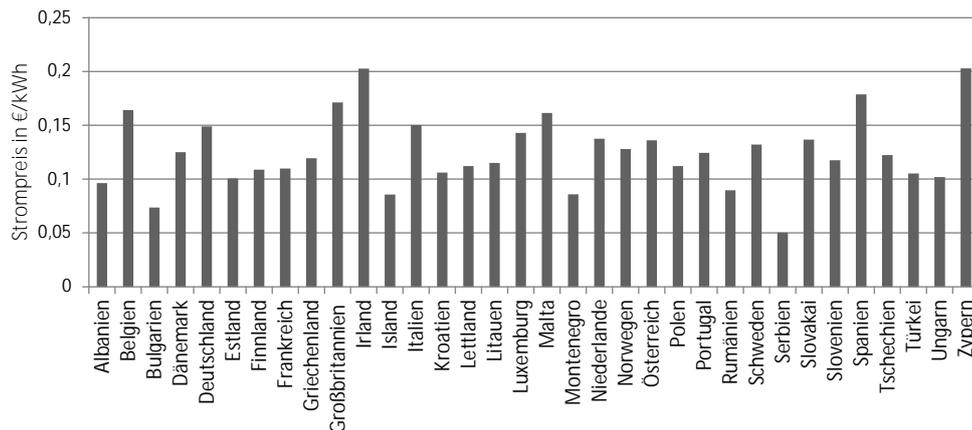


Abbildung 3.3: Europäische Strompreise (2. Halbjahr 2013)

In aktuellen Verträgen mit Stromlieferanten sind die Strompreise festgelegt. Die Kosten je kWh orientieren sich nach der Abnahmemenge und setzen sich aus folgenden Komponenten zusammen [251]: Für Haushalte machen diverse Steuern, Abgaben und Umlagen (Strom-/Öko- und Umsatzsteuer, Konzessionsabgabe und Umlagen zur Förderung von erneuerbaren Energien und Kraft-Wärme-Kopplung) einen Großteil des Preises aus. Einen weiteren großen Teil bilden die staatlich regulierten Entgelte der Netznutzung zur Übertragung des Stroms. Für Industriebetriebe ist der durchschnittliche Energiepreis deutlich niedriger als bei privaten Haushalten. Zum einen ist der Steuersatz ca. 40 % geringer und zum anderen wird durch die häufige Direktanbindung ans Hoch- oder Mittelspannungsnetz der Umweg über das teure Niederspannungsnetz vermieden, was in geringeren Netzkosten resultiert. Des Weiteren bekommen Industrie und Gewerbe als Großabnehmer weitere Rabatte. Der drittgrößte Anteil

3 Technische Voraussetzungen

des Preises gilt der Stromerzeugung und -lieferung. Dieser Anteil orientiert sich größtenteils an den Großhandelsmärkten für Strom, wie beispielsweise der Strombörse EPEX SPOT in Paris.

In solchen Strombörsen werden zeitlich abgegrenzte Mengen an Strom ähnlich wie an einer Wertpapierbörse gehandelt. Durch die Bündelung von Angebot und Nachfrage kann dabei eine bessere Ausnutzung der Kapazitäten erzielt werden. Der kurzfristige Handel innerhalb eines Tages oder für den nächsten Tag ermöglicht den Teilnehmern, Abweichungen vom prognostizierten Energiebedarf auszugleichen, indem entweder weitere Energiemengen hinzugekauft werden oder ein Überschuss verkauft wird. Die angebotenen Produkte unterscheiden sich hauptsächlich in der zeitlichen Länge der Lieferung bei einer konstanten Leistung: Es gibt einen „Baseload-Block“ für den täglichen Grundbedarf, welcher also eine Länge von 24 Stunden hat, und einen „Peakload-Block“ für die Stunden von 8 bis 20 Uhr, der den höheren Energiebedarf tagüber abdeckt. Außerdem gibt es für eine feinere Abbildung des täglichen Lastgangs die Möglichkeit, Einzelstundenkontrakte zu handeln. In Abbildung 3.4 ist eine Darstellung zur Deckung eines beispielhaften Energiebedarfs durch die erwähnten Blöcke gegeben.

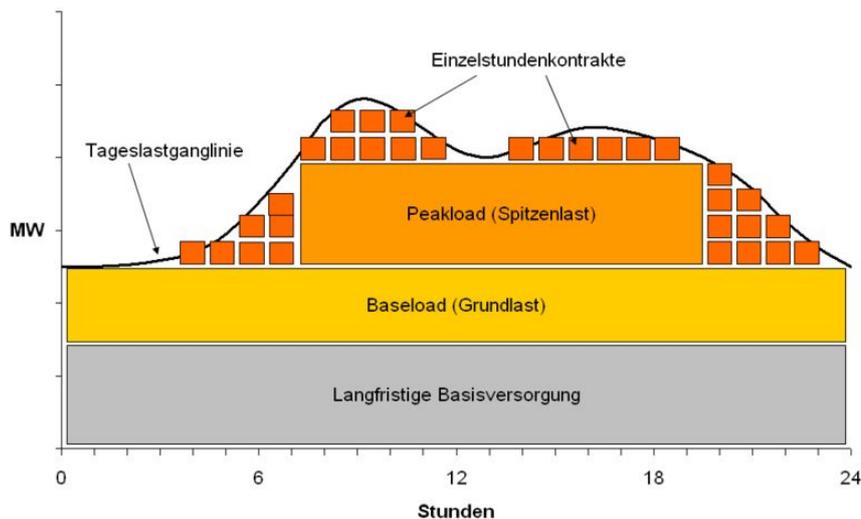


Abbildung 3.4: Die benötigte Energie (Tageslastganglinie) wird durch unterschiedliche Energieblöcke abgedeckt.²

Der Stromlieferant ist zuständig für die Bereitstellung des Stroms, unabhängig von der Tageszeit. Dabei können die realen Kosten sehr stark schwanken und den bezahlten Preis bei Peakzeiten übersteigen. Der Lieferant trägt hierbei das Risiko, das sich durch Kalkulationen und entsprechende Preisgestaltung abfangen lässt. Der hierbei entstehende, von Angebot und Nachfrage unabhängige, einheitliche Strompreis demotiviert allerdings davon, Energie forciert an bestimmten Zeitpunkten abzurufen.

Durch Weitergabe der realen Strompreise an den Kunden, wie es z. B. bei einem Feldversuch der Mannheimer MVV Energie AG im Jahre 2008 [212] durchgeführt wurde, oder auch durch Partizipation an der Strombörse können die dynamischen Kostenveränderungen gezielt genutzt

²<http://buhmannstrom.npage.de/stromboerse.html>, letzter Zugriff: 11.06.14

werden. Da zuverlässige Prognosen für Temperaturen nur für bis zu einem Tag in die Zukunft gegeben werden können, bietet sich ein Handel an den Strombörsen mit ähnlichem Zeitrahmen an: Day-Ahead- oder auch Intraday-Handel (Spotmarkt).

Laut [210] gilt der Day-Ahead-Strompreis als ein starker Vorhersageparameter für den Real-Time-Preis. Wird jedoch verstärkt DSM eingesetzt, habe dies laut [127] einen signifikanten Einfluss auf den Real-Time-Preis. Um den endgültigen Preis abschätzen zu können, wird daher in [127] der Einfluss von DSM auf den Strompreis analysiert. Die Anpassungen der Rechenzentrumslasten aufgrund der Strompreise entspricht der indirekten Teilnahme am DSM über ein Anreizsignal. In den Untersuchungen zum Einsparpotential sollen allerdings relative Zusammenhänge gebildet werden, weshalb nicht die absolute Höhe der Strompreise relevant ist, sondern die relativen Differenzen der Verläufe zueinander. Daher kann der Day-Ahead-Preis als ein erster Anhaltspunkt genommen werden.

3.2 Konzept eines rechenzentrumsinternen Last- und Powermanagements

Die Grundlagen des hier verwendeten Verfahrens für das lokale, rechenzentrumsinterne LPM werden in [110] vorgestellt. Das Verfahren ermöglicht eine dynamische Konsolidierung, die zu jedem Zeitpunkt (auf statistische Weise) gewährleistet, dass genügend Ressourcen vorhanden sind. Es agiert proaktiv, garantiert die Einhaltung von SLOs und überschreitet keine vorher definierte maximale Anzahl an benötigten Servern. Das Verfahren ist darauf ausgelegt, eine Mehrzieloptimierung für unterschiedliche Ressourcentypen (z. B. CPU, RAM, IO) durchzuführen. Bei den Migrationsentscheidungen fließen sowohl die Verzögerungszeiten als auch die Kosten für das Server-Powermanagement (an \leftrightarrow Stand-By) und für die Migrationen selbst mit ein.

Das Konzept des Verfahrens basiert auf der Unterscheidung zwischen sicheren und unsicheren Verteilungen von VMs auf Servern. Eine sichere Verteilung stellt jeder VM zu jedem Zeitpunkt genügend Ressourcen zur Abarbeitung einer maximal zu erwartenden Anfragelast bereit. Eine unsichere Verteilung erfüllt zwar das Kriterium, genügend Ressourcen zu einem bestimmten Zeitpunkt bereitzustellen, nicht jedoch zu jedem Zeitpunkt.

Initial wird eine sichere Verteilung hergestellt: Hierzu werden klassische Verfahren (Vektor-Bin-Packing) herangezogen, um eine Zuordnung von VMs und Servern zu ermitteln. Informationen über zu erwartende Anfragelasten werden mittels eines Vorhersageverfahrens gewonnen (siehe auch Abschnitt 3.1.1). Zur Laufzeit kann nun das LPM auf den Prognosen stützend Server frei räumen und in einen energiesparenden Modus (Stand-By oder aus) versetzen, wodurch eine unsichere Verteilung angenommen wird. Dies setzt jedoch voraus, dass keine Ressourcenengpässe auftreten und zwar

- während in die neue Verteilung gewechselt wird,
- solange Server zum Energiesparen deaktiviert sind und

3 Technische Voraussetzungen

- in der Zeit, die benötigt wird, um vollständig in die sichere Verteilung zurückzukehren.

Hierdurch wird sichergestellt, dass jeder beliebige drohende Ressourcenengpass verhindert werden kann, bevor er tatsächlich eintritt. Im schlimmsten Fall wird die sichere Verteilung vollständig wiederhergestellt, was per Definition der sicheren Verteilung jeden drohenden Ressourcenengpass auflöst. Verteilungen, die diese Bedingungen erfüllen, werden fortan als dynamische Verteilungen bezeichnet.

Einem Server werden also zwei unterschiedliche Mengen an VMs zugeordnet: die aktuell in der dynamischen Verteilung allozierten VMs und die VMs aus der sicheren Verteilung, die vom Server zu jedem Zeitpunkt genügend Ressourcen bereitgestellt bekommen. Für letztere stellt der Server den sogenannten Heimserver dar. Ist eine VM nicht auf ihrem Heimserver alloziert, sondern auf einem anderen Server, so stellt sie für diesen eine Fremd-VM dar.

Im Gegensatz zu einer intervall-bezogenen Optimierung agiert das Verfahren fortlaufend. Die oben genannten Bedingungen müssen daher zu jeder Zeit gelten. Hierzu werden zwei Konzepte eingeführt: ein dynamischer Betrachtungszeitraum und ein Migrationsgraph.

Dynamischer Betrachtungszeitraum

Damit die o.g. Bedingungen eingehalten werden können, muss für einen bestimmten Zeitraum Ressourcenengpassfreiheit gelten. Im Ausgangszustand, also der sicheren Verteilung, ist es per Definition nicht notwendig zu überprüfen, ob zukünftig Ressourcenengpässe aufkommen können. Der Betrachtungszeitraum hat also die Länge null. Sobald eine Migration ausgeführt werden soll, erhöht sich der Betrachtungszeitraum entsprechend den o.g. Bedingungen um die Zeit, die

- zur Migration,
- zur Deaktivierung, zum Energiesparen und zur Reaktivierung eines Servers – soweit dieser frei geräumt und deaktiviert werden kann – sowie
- zur Rückmigration in die sichere Ausgangslage benötigt wird.

Sowohl die Reaktivierung eines Servers (Δt_{up}) als auch die Migration einer VM (Δt_{mig}) benötigen eine gewisse Zeit. Damit auch im Worst-Case keine Ressourcenengpässe entstehen, müssen alle VM $i \in M$ (Menge der VMs, welche sich aktuell nicht auf ihren Heimservern befinden) zurückmigriert und alle Server $j \in N$ (Menge der deaktivierten Server) reaktiviert werden. Für eine beliebige dynamische Verteilung muss daher für den in der folgenden Formel angegebenen Zeitraum Δt_{safe} gelten, dass keine Ressourcenengpässe auftreten:

$$\Delta t_{safe} = \sum_{i \in M} \Delta t_{mig,i} + \sum_{j \in N} \Delta t_{up,j} \quad (3.1)$$

Soll eine dynamische Verteilung angenommen werden, zu deren Erreichung mehrere VMs (Menge O) migriert werden müssen, erweitert sich die Formel um die entsprechenden Migrationszeiten. Zusätzlich muss im Falle der Deaktivierung von Servern (Menge P) die hierfür

benötigte Zeit Δt_{down} und auch eine Break-Even-Zeit Δt_{BE} sichergestellt sein, damit sich der Aufwand rechnet. Zur Bestimmung der gesamten Rückmigrationszeit müssen diejenigen VMs aus O zusätzlich betrachtet werden, die nicht ihren Heimserver als Ziel haben (Menge O'). Die Menge der zu betrachtenden hochzufahrenden Server erweitert sich um P :

$$\Delta t_{check} = \sum_{i \in MUO'} \Delta t_{mig,i} + \sum_{j \in NUP} \Delta t_{up,j} + \sum_{k \in O} \Delta t_{mig,k} + \sum_{l \in P} (\Delta t_{down,l} + \Delta t_{BE,l}) \quad (3.2)$$

Zur Laufzeit ist grundsätzlich das Betrachtungsintervall Δt_{safe} und bei geplanten Migrationen das Betrachtungsintervall Δt_{check} zu überprüfen.

Die Länge der Betrachtungsintervalle ist proportional abhängig von der Anzahl an migrierten VMs. Dies ist eine pessimistische Betrachtungsweise der Netzwerkbelastung: Da keine konkrete Netzwerktopologie vorliegt, wird der Worst-Case angenommen, dass die Migrationen über einen Netzwerkbus erfolgen. Damit kann nur eine Migration zur Zeit durchgeführt werden.

Migrationsgraph

Wie bereits erwähnt, müssen im Worst-Case sämtliche VMs auf ihre Heimserver zurück migriert werden. Hierbei könnte sich jedoch das folgende Szenario bilden, das in Abbildung 3.5 dargestellt ist. Die VMs VM1 und VM2 befinden sich nicht auf ihren jeweiligen Heimservern. Der Heimserver von VM1 ist Server S1. Ein Zurückmigrieren ist aktuell allerdings nicht möglich, da nicht genügend freie Kapazitäten (grün eingefärbt) auf S1 vorhanden sind. Diese müssten erst noch geschaffen werden, indem zuerst von Server S1 VMs verschoben werden. VM2 ist hier die einzige Fremd-VM, weshalb diese verschoben werden sollte, und zwar wiederum auf ihren Heimserver. Dieser bietet aktuell jedoch ebenso nicht genügend freie Kapazitäten, da auch hier eine Fremd-VM, nämlich VM1 alloziert ist. Es ist ein sogenannter Deadlock entstanden. Es kann also keine der VMs zurückmigriert werden, ohne dass ein Ressourcenengpass entsteht.

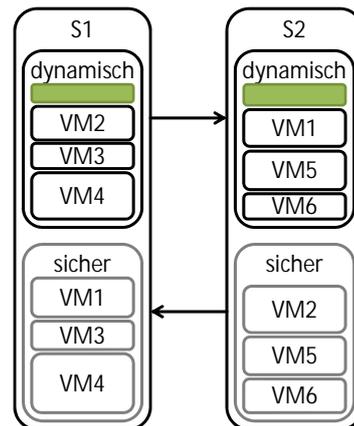


Abbildung 3.5: Deadlock durch Zyklenbildung: Eine direkte Rückmigration auf den jeweiligen Heimserver würde einen Ressourcenengpass verursachen.

Um einen Engpass zu vermeiden, kann nun auf zweierlei Weisen verfahren werden. Entweder werden andere VMs mit geringerem Ressourcenbedarf verschoben, so dass genügend freie Kapazitäten für VM1 oder VM2 geschaffen werden – im Beispiel würde dazu zuerst VM3 auf S2 verschoben werden, woraufhin dann genügend freie Kapazitäten für VM1 vorhanden sind – oder aber es wird ein weiterer Server mit ausreichend freien Kapazitäten hinzugezogen.

3 Technische Voraussetzungen

In beiden Fällen verlängert sich jedoch die notwendige Gesamtmigrationszeit. Die zuvor sichergestellte Ressourcenengpassfreiheit durch die Überprüfung der Betrachtungszeiträume Δt_{safe} bzw. Δt_{check} wird damit hinfällig.

Das Problem, das hier auftritt, wird durch den Zyklus verursacht, VMs auf ihre Heimserver zurückzumigrieren. Im Beispiel ist ein Zyklus der Länge 2 beschrieben worden, aber auch bei Zyklen größerer Länge ist ein solcher Sachverhalt möglich. Werden zyklische Migrationen vermieden, kann der beschriebene Fall ausgeschlossen werden. Anhand eines Migrationsgraphen, in dem (multiple) gerichtete Kanten von Servern, auf denen Fremd-VMs alloziert sind, zu den jeweiligen Heimservern dieser VMs zeigen, können Zyklen erkannt werden. In Abbildung 3.6 ist ein Migrationsgraph dargestellt, in dem das oben gegebene Szenario mit eingeschlossen ist und darüber hinaus ein weiterer, längerer Zyklus enthalten ist. Ein solcher Migrationsgraph lässt sich auch als Matrix darstellen (schwarze Ziffern) und ein Zyklus effizient durch Ableitung einer Adjazenzmatrix A (zusätzlich rote Ziffern) detektieren. Ein Zyklus besteht, wenn in der Diagonalen ein Wert gleich 1 ist, also $\exists i : A_{ii} = 1$.

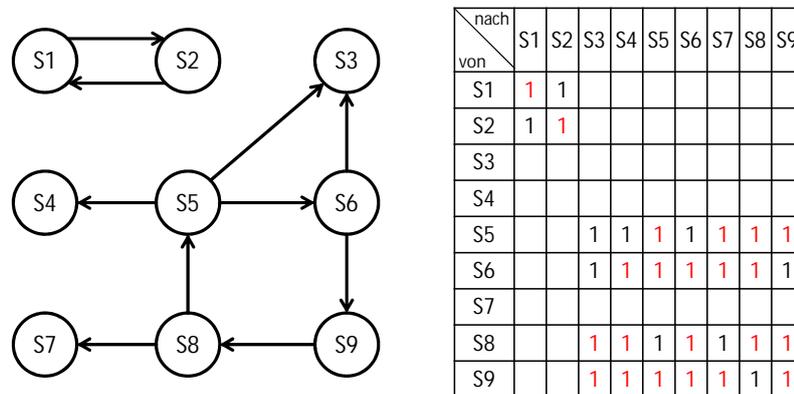


Abbildung 3.6: Migrationsgraph (links) und Matrixschreibweise (rechts, schwarze Ziffern) sowie Adjazenzmatrix (+ rote Ziffern) zur Zyklendetektion

Vor jeder Migration ist zu überprüfen, ob ein Zyklus entstehen würde. In diesem Fall ist die Migration zu verwerfen. Auf diese Weise wird für jede dynamische Verteilung ein sicherer Rückweg – also ohne Auftreten von Ressourcenengpässen – zur sicheren Verteilung garantiert.

3.3 Zusammenfassung

Die für die Optimierungen benötigten Werte bzgl. des Ressourcenbedarfs von Diensten, der Standorttemperaturen sowie der Strompreise lassen sich für einen Prognosehorizont von wenigen Stunden bis zu einem Tag mit hoher Genauigkeit vorhersagen.

Für das rechenzentrumsinterne LPM wird ein Verfahren verwendet, das unter Verwendung einer statischen Verteilung zur Absicherung der Ressourcenkapazität eine SLO-Garantie besitzt.

4 Konzeption samt Komponentenevaluation

Aufbauend auf den im vorherigen Kapitel beschriebenen technischen Voraussetzungen wird in diesem Kapitel das Konzept für das globale LPM vorgestellt, welches die in Abschnitt 2.2.2 herausgearbeiteten Anforderungen erfüllt. Hierzu wird zuerst die Architektur erläutert sowie ein Systemüberblick gegeben und im Anschluss auf die einzelnen Bestandteile näher eingegangen. Da es sich bei diesen mitunter um spezielle Problemstellungen handelt, werden jeweils eine Problembeschreibung, verwandte Arbeiten sowie die Lösungen gegeben. Die jeweils entwickelten Verfahren werden auch bereits evaluiert, um darauf aufbauend die weiteren Systemkomponenten abzustimmen.

Bei diesen Evaluationen werden die Güte und die Laufzeit der einzelnen Verfahren und Komponenten unabhängig von den jeweils anderen Bestandteilen analysiert, um die jeweilige Eignung zu ermitteln sowie das Potential für mögliche Verbesserungen herauszustellen. In Abgrenzung hierzu liegt der Fokus der in Kapitel 5 beschriebenen Gesamtevaluation in der Bestimmung des Nutzens des globalen LPM mit Schwerpunkt der Parameterabhängigkeitsermittlung. Die Güte des Gesamtverfahrens wird nicht evaluiert, da es keine Referenzimplementierung gibt und der Abstand zu einer optimalen Lösung auch nicht direkt abgeschätzt werden kann, wie es bei den einzelnen Verfahren durchaus möglich ist.

Zur Durchführung einer Evaluation wird eine Simulationsumgebung (siehe Abschnitt 5.1) eingesetzt, mit der basierend auf realen Daten unterschiedliche Szenarien analysiert werden können. Aus der Abgrenzung zu verwandten Arbeiten ging hervor, dass ihre Verfahren nur mit einigen wenigen Szenarien evaluiert wurden. Damit ist nicht nur das ermittelte Einsparpotential, sondern auch die Eignung der Verfahren selbst nur mit diesen Szenarien gezeigt worden. Dies schließt zwar eine allgemeine Eignung nicht aus, aber durch gezielte Parametrisierung können Verfahren auf bestimmte Probleminstanzen hin optimiert werden. Damit dieses bei den zu entwickelnden Verfahren nicht geschieht, werden Szenarien mit unterschiedlichen Ausprägungen evaluiert. Welche Parameter hierzu variiert werden, wird in den jeweiligen Evaluationen beschrieben. Eine vollständige Abdeckung des Parameterraums ist jedoch nicht möglich, weshalb mehrere zufällige Variationen evaluiert werden. Je untersuchten Wert werden mind. 25 unterschiedliche Szenarien herangezogen und ihr Durchschnittswert propagiert.

4.1 Architektur

Die meisten vorgestellten verwandten Arbeiten setzen Verfahren ein, die eine hohe Laufzeitkomplexität aufweisen und zusätzlich nur zentral ausgeführt werden. Für die vorliegende Problemklasse sind diese Verfahren aufgrund ihrer geringen Skalierbarkeit i.d.R. nicht praxistauglich. Hier ist der Einsatz von (Meta-)Heuristiken zu empfehlen, da diese zwar keine optimalen Ergebnisse liefern, dafür aber für gewöhnlich sehr gute in relativ kurzer Laufzeit. In [38] wird eine Heuristik genutzt, deren Laufzeit lediglich quadratisch von der Rechenzentrumsanzahl abhängt. Außerdem ist sie verteilt ausführbar. Das dort betrachtete System unterscheidet sich allerdings, wie bereits beschrieben, von dem in dieser Dissertation betrachteten. Der Systemkontext wird hier durch das Cloud Computing gebildet, das nicht nur ein technischer Wegbereiter für das globale LPM ist, sondern auch die gleichen Komponenten (im IaaS-Fall) betrachtet. Daher sollte eine Heuristik eingesetzt werden, die sich besonders in diesem Kontext eignet.

Von Cloud-Infrastrukturen wird verlangt, dass Rechenleistung, Verfügbarkeit, Fehlertoleranz, Skalierbarkeit, Sicherheit etc. auf flexible und vor allem automatische Weise bereitgestellt wird. Ein geeigneter Ansatz, um dieses umzusetzen, ist laut Dobson, Solomon und weiteren [75, 208] das Autonomic Computing, indem es die vier Self-Properties (Self-Configuring, Self-Optimizing, Self-Healing und Self-Protecting) [68] adressiert. Als Alternative zum Autonomic Computing kommt das Organic Computing in Betracht, da es hierbei um die Beherrschung hochkomplexer, adaptiver und selbstorganisierender Systeme geht. Es fehlen allerdings noch systematische Ansätze zum Entwurfsprozess solcher Systeme [160].

Für Autonomic Computing und für das Cloud Computing ist laut Chihi und weiteren [60] eine konkrete Realisierung mittels Multi-Agenten-Systemen zu empfehlen. Dieser Ansicht sind auch Liu, Brazier und weitere [139, 35], da sich die (Haupt-)Eigenschaften von Self-Management sowie von agenten-basierten und service-orientierten Systemen ähneln. Durch den Einsatz eines Multi-Agenten-Systems (MAS) wird ein hochkomplexes Problem auf kleinere Probleme aufgeteilt und durch einzelne Agenten verteilt lösbar. Gegenüber einer zentralen Lösung wird hierdurch höhere Robustheit und bessere Skalierbarkeit erreicht.

In [60] wird für eine föderierte Cloud eine konzeptionelle Umsetzung mit einem MAS vorgestellt, das für jede Self-Property einen Agenten auf jeder Ebene der Ausführungshierarchie (Region, Rechenzentrum, Server, VM) vorsieht. Die Erfüllung der vier Self-Properties wird anhand von Szenarien gezeigt. Das lokale LPM entspricht hier dem Self-Optimizing-Agenten auf Rechenzentrumsebene, das globale LPM dem auf regionaler bzw. globaler Ebene.

Das übergeordnete System im Problemkontext sollte demnach aufgrund der hohen Komplexität als MAS abgebildet werden. Das globale LPM stellt hier einen Bestandteil dar, welcher selbst eine hohe Komplexität aufweist. Die nachfolgende Konzeption legt nahe, dass das globale LPM selbst auch als MAS realisiert werden sollte.

Die Betrachtung der Replikationsaspekte erfolgt hiervon unabhängig (in Abschnitt 4.8). Zwar würden laut [46, 186] die besten Ergebnisse erzielt werden, wenn beim Scheduling

(hier: globales LPM) sowohl der aktuelle Workload als auch die Kosten für den Datenzugriff berücksichtigt werden, eine Kopplung von Scheduling- und Replikationsverfahren zu einem komplexeren Verfahren sei aber nicht notwendig.

4.1.1 Konzeption

Die Aufgabe des zu entwickelnden globalen LPM besteht darin, Lasten entsprechend den Betriebsbedingungen zwischen unterschiedlichen Rechenzentren zu verlagern. Hierzu werden wie auch beim rechenzentrumsinternen LPM einzelne VMs migriert. Allerdings wird aufgrund des Einbezugs mehrerer Rechenzentren eine deutlich höhere Anzahl an VMs betrachtet. Eine gute Skalierbarkeit ist hier also zwingend notwendig. In Abschnitt 3.2 wurde das hier verwendete rechenzentrumsinterne LPM vorgestellt. Es ist ein heuristisches Verfahren, das eine Laufzeit hat, die polynomiell mit der VM-Anzahl ($O(n^3)$, siehe Abschnitt 4.3.1) wächst. Dies entspricht zwar der Laufzeitkomplexitätsvorgabe, aber eine direkte Anwendung dieses Verfahrens auf einen ganzen Verbund von Rechenzentren ist nicht zu empfehlen, da zum einen durch die Vergrößerung des betrachteten VM-Pools die Effizienz des Verfahrens verringert wird [108] (siehe auch Abbildung 4.14) und zum anderen dadurch eine Vereinheitlichung der lokalen LPMs aufgezwungen wird, was aufgrund möglicher unterschiedlicher Betriebsvorgaben je Rechenzentrum impraktikabel sein kann.

Damit die Autonomie jedes einzelnen Rechenzentrums und damit auch die Eigenschaften des internen LPM beibehalten werden können, wird daher eine Hierarchisierung, wie in Abbildung 4.1 dargestellt, eingeführt. Für die darüber liegende, globale Schicht wird vom konkreten Rechenzentrum samt internem LPM abstrahiert. Hierdurch lässt sich auch von der internen Komplexität abstrahieren, so dass diese auf maximal $O(n)$ reduziert werden kann.

Grundlegend kann jeder Kunde bestimmen, mit welchen Standorten er zum Betrieb seines Dienstes einverstanden ist. Genauso sollte jedoch auch ein Rechenzentrumsbetreiber eine Auswahlmöglichkeit haben. Eine entsprechende Entscheidungsinstanz kann durch einen korrespondierenden Standortagenten (lokaler Agent) übernommen werden.

Neben der Skalierbarkeit ist ebenso die Flexibilität des Verfahrens bzgl. der Optimierungsziele wichtig. Als Optimierungsziele kommen abgesehen von der Kosten- und Energiebedarfsminimierung auch das Einhalten eines bestimmten Lastgangs, der sich bspw. aufgrund des anliegenden Smart Grids ergibt, sowie die Senkung der CO_2 -Emissionen in Frage. Denkbar wären auch unterschiedliche Optimierungsziele je Standort. Die meisten dieser Ziele beziehen sich auf den Energiebedarf, so dass dieser als generischer Zielwert festgelegt wird. Soll eine kostenbezogene Optimierung durchgeführt werden, muss aus dem Kostenwert ein Energiebedarfswert abgeleitet werden. Hierauf wird in Abschnitt 4.7.1 eingegangen.

Zur bestmöglichen Einhaltung von Lastgangvorgaben je Standort ist eine zentrale Kontrollinstanz ratsam. Zusammen mit den Standortagenten bietet sich hierzu das Prinzipal-Agent-Modell an: Auf globaler Ebene ermittelt ein arbitrierender Agent (Prinzipal/Arbiter) globale Vorgaben und übermittelt diese den Standortagenten. Diese wiederum besitzen die Aufgabe, die Vorgaben lokal umzusetzen, indem sie die notwendigen VM-Migrationen in

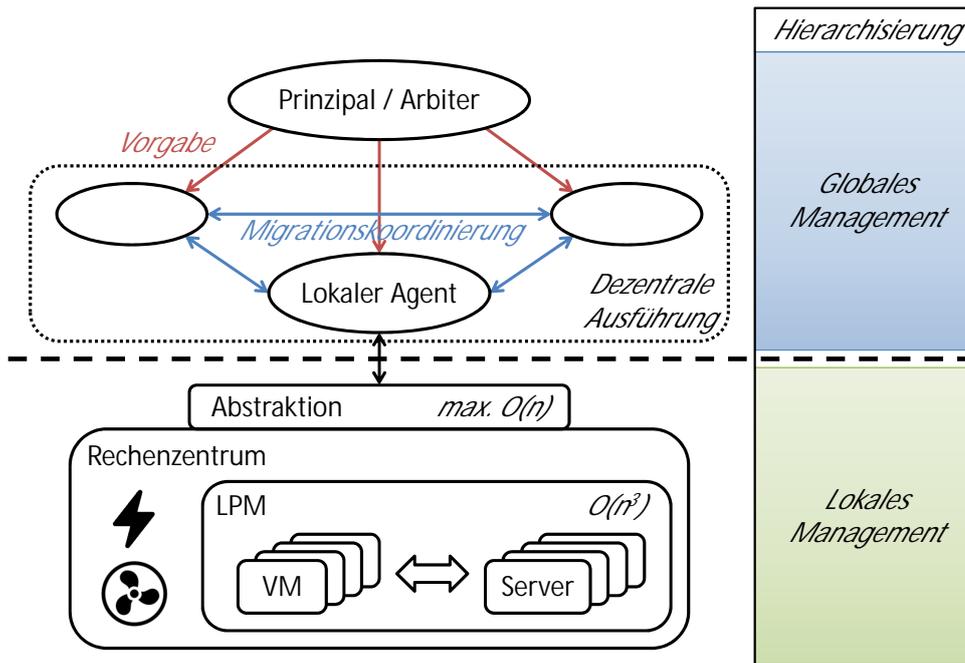


Abbildung 4.1: Logische Aufteilung des LPM in einen lokalen und globalen Part. Der globale Part sieht für jedes Rechenzentrum einen Agenten vor, der die lokalen Aspekte, nämlich die Zuordnung von VMs und Servern sowie die Betriebssicherstellung durch die Stromversorgung und Klimatisierung, abstrahiert betrachtet.

Abstimmung mit anderen Standortagenten durchführen. Sie können dabei dezentral und unabhängig voneinander agieren, wobei die Kommunikationstopologie vollvermascht ist.

Prinzipiell kann die Funktion des Arbiters auch von den lokalen Agenten übernommen werden (siehe Abschnitt 4.7.5), was im realen Betrieb auch so umgesetzt werden sollte. Der Übersichtlichkeit halber und zum besseren Verständnis des Konzeptes soll die Funktion hier aber im Arbitrer gebündelt beschrieben werden.

Ein Rechenzentrum stellt somit für das globale Lastmanagement ein Element dar, das sich bzgl. der verwaltbaren VMs, also der aufgenommenen Rechenlasten und des damit induzierten Energiebedarfs anpassen lässt. Von der internen Realisierung wird abstrahiert: Die konkreten Ausprägungen, also die IT-Hardware, die Infrastruktur und das LPM-System sind austauschbar. Damit ist kein einheitliches lokales LPM notwendig, was insbesondere in einem Kooperationsverbund unterschiedlicher Rechenzentrumsbetreiber wichtig ist. Interne Abläufe und der konkrete Aufbau eines Rechenzentrums müssen somit nicht preisgegeben werden.

Eine alternative Agentenstruktur für das globale LPM könnte für jede VM einen Agenten vorsehen, welcher selbstständig entscheidet, an welchem Standort die jeweilige VM betrieben werden sollte. Die Realisierung als Schwarmintelligenz würde sich in diesem Fall anbieten. Die Sicherstellung der Erfüllung einer standortspezifischen Lastgangvorgabe ist hierbei allerdings komplizierter, da sich hierzu hunderte bis tausende Agenten koordinieren müssten, um ein aggregiertes Ziel zu erlangen, das auch mit wenigen Agenten und damit auch mit weniger Koordinierungsbedarf ermöglicht werden kann.

Das andere Extrem, nämlich der Einsatz eines einzelnen Agenten zur Erfüllung aller Vorgaben hätte zwar den geringsten Kommunikationsaufwand, würde jedoch i.d.R. die Ausführung zentralisieren und die Skalierbarkeit reduzieren. Hierfür einsetzbare Verfahren wären bspw. ILPs und Greedy-Heuristiken, wie in den verwandten Arbeiten, oder auch evolutionäre Algorithmen. Letztere lassen sich parallel und damit auch verteilt ausführen, wodurch eine sich der obigen Agentenstruktur ähnelnde aufbauen lässt. Die einzelnen Rechenknoten (Agenten) hätten dann allerdings die Aufgabe, einen Teil der Gesamtrechenlast zu übernehmen, und nicht lokale Ziele zu verfolgen, was den Vorgang zur Erreichung lokaler Ziele verkomplizieren würde.

Schwarmintelligenz und evolutionäre Algorithmen stellen damit Alternativen zur vorgestellten Architektur dar. Entsprechend dem KISS-Prinzip (*Keep It Simple, Stupid*), das besagt, dass die meisten Systeme am besten laufen, wenn sie einfach gehalten werden, wird aber aufgrund der klaren und direkten Aufgabenpartitionierung das oben beschriebene Konzept gewählt.

4.1.2 Einordnung

Laut [52] ist es nicht notwendig, in einem MAS eine Entität mit einem globalen Ziel zu besitzen, um dieses dennoch zu erreichen. Es reicht hierzu aus, den individuellen Agenten gezielt Informationen zu geben, damit sie ihr Verhalten wie gewünscht anpassen. In der vorgestellten Architektur wird das Verhalten der lokalen Agenten nicht beeinflusst, sondern anhand der Lastgangvorgabe eine lokale Optimierungsrichtlinie gegeben, welche eine globale Optimierung bezweckt. Dies wird nach [52] als informativer Mechanismus bezeichnet. Abgegrenzt wird der informative Mechanismus gegenüber dem regulativen, bei dem in die Autonomie des Agenten eingegriffen wird.

Im Vergleich zu den in [158] vorgestellten Cloud-Föderationsarchitekturen entspricht die vorgestellte Architektur der Broker-Variante, bei der jede Cloud (hier: lokales LPM) für das eigene Management zuständig ist. Der Broker (hier: globales LPM) hat jedoch Einflussmöglichkeiten bzgl. Schedulingentscheidungen, die basierend auf Optimierungskriterien wie die Betriebskosten oder Energieaufnahme gefällt werden.

In [98] werden vier Eigenschaften beschrieben, nach denen eine Klassifikation von Multi-Agenten-Systemen durchgeführt werden kann. Diese Eigenschaften werden im Folgenden auf die vorgestellte Architektur bezogen diskutiert, indem bereits auf die konkrete Konzeptionierung eingegangen wird.

- Agentenmodelle: Besitzen Agenten Kenntnis über andere Agenten und auch sich selbst, z. B. durch ein Modell, und können dadurch das Verhalten abschätzen, wird von einem strukturierten und vorhersagbaren System gesprochen. In der hier eingesetzten Architektur verwenden die lokalen Agenten ein Modell vom korrespondierenden Rechenzentrum und damit auch quasi von sich selbst, denn sie streben nach der Erfüllung der Lastgangvorgabe. Dieses Modell nutzt auch der Arbitrer, um gesamtoptimierende

4 Konzeption samt Komponentenevaluation

Lastgangvorgaben ermitteln zu können. Somit besitzt dieser Kenntnis über andere Agenten. Insgesamt verhilft die Verwendung von Modellen zur besseren Koordinierung der Agenten.

- **Anpassung:** Das Verhalten eines Agenten kann grundsätzlich durch einen Lernprozess angepasst werden. Dieser findet hier nicht statt, d.h., dass das Verhalten der Agenten stets dem Ausgangsverhalten entspricht, wenngleich die Entscheidungsfindung die aktuellen Bedingungen berücksichtigt.
- **Entscheidungsfindung:** Ein Agent kann Entscheidungen treffen, indem er nur seine eigene Kenntnis einbringt oder durch Berücksichtigung anderer, z.B. mittels eines Gruppenvotums. Wie später in Abschnitt 4.7.2 detaillierter beschrieben wird, werden Entscheidungen hier in zwei Phasen getroffen. In der ersten werden nur die eigenen Interessen vertreten, in der zweiten werden die Interessen aller Agenten berücksichtigt. Erst dann wird entschieden, welche globalen VM-Migrationen tatsächlich durchgeführt werden.
- **Kooperationsbereitschaft:** Ist die Agentenstrategie egozentrisch ausgelegt, werden Entscheidungen nur zum Wohle der eigenen Ziele getroffen. Entsprechend der Entscheidungsfindung handeln die Agenten in der ersten Phase egozentrisch. Aber auch die zweite Phase ist egozentrisch, da nur – lokal betrachtet – die jeweilige beste Lösung (VM-Migration) genommen wird. Ein kooperativer Ansatz, der also eine beste Lösung für die Allgemeinheit sucht und dabei auch ein Zurückstecken der eigenen Ziele vorsieht (Altruismus), ist nicht direkt im Agentenverhalten verankert, sondern wird ausschließlich durch den Informationsmechanismus gesteuert und bewirkt.

4.1.3 Systemüberblick

Zur Realisierung des globalen LPM werden mehrere Komponenten und Verfahren entwickelt, deren Umsetzung in den folgenden Abschnitten erläutert wird:

- Energiebedarfsmodelle der Rechenzentrumskomponenten
- Rechenzentrumsinternes LPM
 - Migrationsrichtlinien (Policy), auch zur Berücksichtigung heterogener Hardware
 - Erweiterungen zum Betrieb in einem dynamischen Umfeld
- Abstraktion des LPM-Konsolidierungsverhaltens
- VM-Selektionsverfahren zur Lastganggenerierung
- Netzwerkmodell
- Globales LPM (Agentenverhalten)

- Verfahren zur Standortselektion von Replikaten

Das systemische Zusammenwirken dieser Bestandteile ist in Abbildung 4.2 skizziert: Je Rechenzentrumsstandort existieren die auf den Betrieb auswirkenden Standortparameter (Strompreis, Temperatur, regionaler Lastgang), welche sich über die Zeit verändern. Das Rechenzentrum ist in drei Hauptkomponenten unterteilt: das lokale LPM, das Energiebedarfsmodell der Rechenzentrumskomponenten samt Abstraktion des LPM-Konsolidierungsverhaltens und den lokalen Agenten, der die Vorgaben des LPM-Arbiters versucht umzusetzen, indem er aus dem globalen VM-Pool die notwendigen VMs bestimmt. Das lokale LPM koordiniert eine Menge an lokal zu betreibenden VMs auf den verfügbaren Servern (nicht dargestellt) und benötigt hierzu ein Datenbanksystem, das insbesondere für die Replikationsbetrachtungen zum globalen LPM herangezogen werden muss. Auf globaler Ebene agiert der LPM-Arbiter, welcher virtuell sämtliche VMs verwaltet. Ferner wird für die globalen Migrationsentscheidungen ein Energiebedarfsmodell für das verbindende Netzwerk benötigt, da dieses einen ökonomischen Widerstand darstellt.

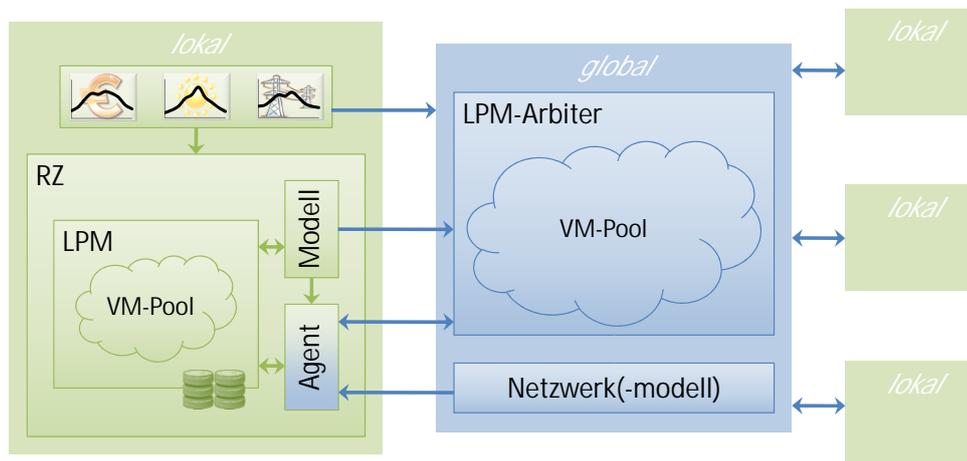


Abbildung 4.2: Übersicht über die Systembestandteile des Gesamtverfahrens und den durch die Pfeile gekennzeichneten Informationsfluss

4.2 Energiebedarfsmodell eines Rechenzentrums

Der Bau immer größerer Rechenzentren stellt nicht nur eine Herausforderung während der Planungsphase dar, sondern auch der Betrieb muss mit entsprechend redundanter Infrastruktur sichergestellt werden. Ein großer Aufwand muss hier betrieben werden, um einen Single-Point-of-Failure zu vermeiden. Eine Veröffentlichung der Microsoft Mitarbeiter Church, Greenberg und Hamilton [62] beschäftigt sich mit den unterschiedlichen Größen von Rechenzentren. Sie stellen fest, dass statt wenige große Rechenzentren eher viele kleine eingesetzt werden sollten. Eine erhöhte Anzahl und Geodiversifikation erhöhe auch die Redundanz [92]. Außerdem werden stark verteilte Architekturen auch von Content-Delivery-Netzwerken wie Akamai verwendet,

um die Latenz zu senken. Als standardisierte Größe bieten sich hierzu Containerrechenzentren an. Im Folgenden werden daher Energiebedarfsmodelle für die Komponenten, die in einer Container-Architektur eingesetzt werden, vorgestellt.

4.2.1 Modelle regelbarer Komponenten im Rechenzentrum

Wie bereits in Abschnitt 2.1.1 (Abbildung 2.1) gezeigt wurde, teilt sich der Energiebedarf in einem Rechenzentrum hauptsächlich auf die IT, die Kühlinfrastruktur und die Stromversorgung auf. Zu diesen werden entsprechende Energiebedarfsmodelle vorgestellt, wobei bei der IT Server und Chassis sowie Netzwerkkomponenten und das Storage erfasst werden.

Serverenergiebedarfsmodell

In der Literatur finden sich viele Untersuchungen zum Energiebedarf von Servern und geeignete Modellierungen. In einer Veröffentlichung von Rivoire und anderen [192] wurde festgestellt, dass Modelle, welche allein Daten zur CPU-Auslastung als Indikator besitzen lediglich einen durchschnittlichen Fehler von 10 % generieren. Auch in den Arbeiten von Fan [81] und Pedram [174] wird allein die CPU als Indikator verwendet. Dies liegt nahe, da die CPU den größten Einfluss auf die steigende Energieaufnahme unter Last hat. In [225] wurde im untersuchten System ein Anteil von 85 % ermittelt. In eigenen Untersuchungen zum Servertyp Dell PowerEdge 1950 konnte die große Abhängigkeit vom CPU-Wert bestätigt werden. Die Hinzunahme weiterer Einflussgrößen wie Speicher-, Netzwerk- und Festplattenzugriffe brachte keine nennenswerten Verbesserungen. Der durchschnittliche Fehler lag jeweils bei unter 5 %.

Die hier eingesetzten Server werden mit Virtualisierungstechnik betrieben. Eine Rechenleistungs- und Energiebedarfsuntersuchung solcher Systeme wurde durch Pedram and Hwang in [174] durchgeführt. Das dort eingesetzte Serverpowermodell bestätigt auch die eigenen Untersuchungen, in denen sich das folgende Modell als adäquat herausgestellt hat:

$$power_{SRV} = power_{idle} + usage_{CPU} \cdot (power_{max} - power_{idle}) \quad (4.1)$$

Wenn sich der Server im Idle-Modus befindet, also keine Berechnungen ausgeführt werden müssen, so hat er dennoch einen gewissen Energiebedarf ($power_{idle}$). Dieser Wert liegt je nach Servertyp bei 40 % bis 70 % des Energiebedarfs bei maximaler Auslastung ($power_{max}$). Ansonsten ergibt sich ein linearer Zusammenhang zwischen Energiebedarf und Auslastung der CPU. Aufgrund des quadratischen Zusammenhangs des Energiebedarfs bzgl. der genutzten Spannung, die durch den Einsatz von DVFS angepasst wird, hat die CPU-Auslastung eigentlich auch einen quadratischen Einfluss. Da damit die Verbesserung des Modells in eigenen Untersuchungen allerdings bei unter einem Prozent liegt, wird auf das einfachere, lineare Modell zurückgegriffen.

Eine alternative Modellierungsmöglichkeit bieten SPEC_{power} Benchmarks. Frei verfügbare Daten zu hunderten vermessener Server liefern Informationen über den Energiebedarf im

Idle und in 10%-Auslastungsschritten.¹ Zwischenwerte können durch lineare oder Spline-Interpolation ermittelt werden. Die Auslastungsabhängigkeit wird hier auch einzig auf den Prozessor bezogen. Die hierdurch erstellten Modelle haben eine höhere Präzision als das in Gleichung (4.1) angegebene Modell. Letzteres genügt jedoch für das hier betrachtete Gesamtsystem.

Beim Betrieb mehrerer VMs auf einem Server besteht kein einfacher Zusammenhang zwischen den Ressourcenbedarfen der VMs und der Auslastung des Servers. Wenn beispielsweise vier VMs – jeweils einzeln auf einem Server betrieben – eine Auslastung von jeweils 20% verursachen, dann verursachen diese vier VMs – zusammen auf einem Server betrieben – keine 80% Auslastung, sondern mehr. Eine Untersuchung über den konkreten Zusammenhang wurde von Srikantaiah, Kansal und Zhao [211] durchgeführt. Für den in der Dissertation gesetzten Untersuchungsschwerpunkt lässt sich hiervon allerdings abstrahieren, da es lediglich das Verhalten des rechenzentrumsinternen LPM beeinflusst. Es wird daher zur Vereinfachung von einem additiven Zusammenhang der Ressourcenbedarfe ausgegangen.

Chassisenergiebedarfsmodell

Ein gewöhnliches Serverchassis weist i.d.R. keinen Energiebedarf auf. Wird der Begriff weiter gefasst, so können hierzu auch Blade-Enclosures gezählt werden. Diese sind zum Betrieb von Blade-Servern notwendig und müssen daher als aktive IT-Komponenten erfasst werden. Sie weisen aufgrund der geteilten aktiven Komponenten (Netzteil, Lüfter, Netzwerk und weiteren Schnittstellen) einen beträchtlichen Energiebedarf auf. In [218] wird der konstante Anteil der Leistungsaufnahme des Dell PowerEdge 1855 Blade-Enclosures mit 820 Watt beziffert. Die jeweiligen Bladeserver haben im Idle eine Leistungsaufnahme von 120 Watt. Die dynamische Leistungsaufnahme eines Bladeservers liegt bei maximal zusätzlichen 50 Watt. In diesem Wert sind auch schon die dynamischen Anteile der geteilten Komponenten aus dem Enclosure enthalten. Für das Dell PowerEdge 1855 Blade-System, das 10 Server-Slots hat, wäre das Energiebedarfsmodell wie folgt:

$$power_{1855} = 820W + \sum_{i=1}^{10} (120W + usage_i \cdot 50W)$$

Verallgemeinert ergibt sich das folgende Energiebedarfsmodell für ein Blade-Enclosure/Chassis mit n bestückten Servern:

$$power_{chassis} = power_{idle} + \sum_{i=1}^n power_{SRV,i} \quad (4.2)$$

¹http://www.spec.org/power_ssj2008/results/, letzter Zugriff: 05.09.2014

Energiebedarfsmodell für Storage und Netzwerk

Ein Storage wird durch speziell hierfür ausgestattete Speicher-Server realisiert, weshalb hierzu grundsätzlich das Energiebedarfsmodell eines Servers herangezogen werden kann, wenngleich durch die Fokussierung auf Speicherhardware, die Charakterisierung anders ausfällt. Zur Erinnerung: Wie in Abschnitt 2.1.1 festgestellt wurde, weisen sowohl rotierende als auch Solid State Festplatten einen statischen und einen dynamischen, proportional zur Last auftretenden Energiebedarf auf.

Ähnlich verhält es sich mit dem Energiebedarf von Netzwerkkomponenten, hier am Beispiel eines Routers. Auch dieser besteht laut [233] aus einem Idle-Anteil und einem Anteil proportional zur Last (hier die Datenrate).

Für die Komponenten Storage und Netzwerk ergibt sich damit ebenfalls ein last-proportionales Energiebedarfsmodell:

$$power_{comp} = power_{idle} + load_{comp} \cdot (power_{max} - power_{idle}) \quad (4.3)$$

Stromversorgung

Bei der Stromversorgung wurde als Hauptverbraucher elektrischer Energie die unterbrechungsfreie Stromversorgung (USV) modelliert. In Zusammenarbeit mit dem Unternehmen Rittal konnten Messdaten für eine Zwei-Modul-USV mit einer Gesamtkapazität von 40 kW gewonnen werden. Diese Daten konnten durch eine Regressionsfunktion der Form

$$power_{USV} = \alpha_0 + \alpha_1 \cdot power_{IT} \quad (4.4)$$

mit einer Genauigkeit von durchschnittlich 99,3% und einem maximalen Fehler von 7% modelliert werden. Die Parameter α_i werden durch den Regressionsprozess bestimmt.

Kühlungskomponenten

Der hier betrachtete Kühlmechanismus beinhaltet die Komponenten Kaltwassersatz (Chiller), Pumpen, Lüfter und ein Freiluftkühlmodul. Auch hierzu konnten in Zusammenarbeit mit Rittal Messdaten gewonnen werden, und zwar für unterschiedliche IT-Lasten und Außentemperaturen.

Die Kühlinfrastruktur besitzt ein Lastmanagement, so dass entsprechend diesen Parametern die inneren Stellgrößen, wie der Wasserdurchfluss, die Druckdifferenz und die Lüfterdrehzahl automatisch angepasst wurden. Die Messungen wurden für unterschiedliche Wasservorlauf-temperaturen durchgeführt. Für eine geeignete Nutzung des Freiluftkühlmoduls wurde die Wasservorlauf-temperatur auf 20 °C festgesetzt.

Als Modellierungstechnik wurde eine multiple lineare Parameterregression verwendet. Hierbei ergab eine polynomielle Abhängigkeit vierten Grades zur Außentemperatur t und eine lineare

Abhängigkeit von der zu kühlenden Last *power* ein adäquates Modell:

$$power_{Cooling} = \alpha_0 + \alpha_1 \cdot power + \sum_{i=1}^4 \alpha_{1+i} \cdot t^i + power \cdot \sum_{i=1}^4 \alpha_{5+i} \cdot t^i \quad (4.5)$$

Ein Regressionsmodell mit einem Polynom höheren Grades führt lediglich zu marginalen Verbesserungen, weist allerdings mitunter unpassende extrapolierende Kurvenverläufe auf, was auf ein Overfitting² hindeutet.

Bei fester Außentemperatur stellen die temperaturabhängigen Terme Konstanten dar. Die Formel vereinfacht sich damit zu einer lediglich linearen Abhängigkeit von der Wärmelast:

$$\begin{aligned} power_{Cooling, t_{const}} &= \alpha_0 + \alpha_1 \cdot power + term_0 + power \cdot term_1 \\ &= (\alpha_0 + term_0) + (\alpha_1 + term_1) \cdot power \\ &= \tilde{\alpha}_0 + \tilde{\alpha}_1 \cdot power \end{aligned} \quad (4.6)$$

Im reinen Freiluftkühlbetrieb lässt sich ein ebenso einfaches Regressionsmodell einsetzen, das allein von der zu kühlenden Leistung abhängig ist. Eine Abhängigkeit von der Außentemperatur konnte hierbei allerdings mangels Testdaten nicht ermittelt werden. Die modellierte Betriebskurve wurde für eine Außentemperatur von 10 °C erstellt. Für niedrigere Temperaturen, ist ein geringerer Energiebedarf zu erwarten. Für diese Temperaturen wird die gleiche Modellkurve genutzt, was einer Überabschätzung entspricht. Für höhere Temperaturen wird festgesetzt, dass der alleinige Betrieb der Freiluftkühlung nicht mehr möglich ist. Ein zusätzlicher Betrieb eines Kaltwassersatzes wird somit notwendig, dessen Energiebedarf deutlich über dem der Freiluftkühlung liegt. Der vermutliche Fehler am Freiluftkühlmodell durch die erhöhte Temperatur spielt daher in diesem Betriebsbereich eine untergeordnete Rolle.

Die erreichte Modellqualität der Gesamtkühlungskette liegt durchschnittlich bei 98 %, im schlechtesten Fall bei 84 %. Die höheren Modellabweichungen treten bei kleinerer Last und bei geringeren Außentemperaturen auf, was zum einen den absoluten Fehlerwert und zum anderen den Anteil der Kühlleistung im Vergleich zur zu kühlenden IT-Last vernachlässigbar erscheinen lässt. So erklärt sich auch, dass bezogen auf das Gesamtenergiebedarfsmodell eines Rechenzentrums (siehe Abbildung 4.4) der durch das Kühlungsmodell entstehende Fehler bei durchschnittlich 0,3 % und maximal 0,6 % liegt.

4.2.2 Lineares Gesamtenergiebedarfsmodell

In Abbildung 4.3 ist das strukturelle Zusammenwirken der Energiebedarfsmodelle skizziert. Das betrachtete Containerrechenzentrum kann – theoretisch – aus einer oder mehreren Kühlzonen bestehen. Innerhalb dieser Zone müssen sämtlich betriebene Anlagen gekühlt werden. Die abzuführende Wärmemenge wird durch die von der USV benötigten Energie

²Die Modellparameter werden auf die Modelldaten so stark spezialisiert/optimiert, dass eine Anwendung auf andere Daten als die Modelldaten zu erhöhten Fehlern führt. Das Modell verliert dadurch seine Generalisierbarkeit.

4 Konzeption samt Komponentenevaluation

bestimmt. Diese setzt sich aus dem Eigenbedarf der USV, abgebildet durch die USV-Module, und dem IT-Bedarf (blauer Bereich) zusammen. Dieser bildet sich wiederum aus den Energiebedarfen des Storages, des Netzwerks, der Racks bzw. Chassis/Enclosures sowie den in diesen eingesteckten Servern zusammen.

Die durch dieses Gesamtmodell ermittelte Leistungsaufnahme eines Containerrechenzentrums ist in Abbildung 4.4 dargestellt. Auf der rechten Achse ist die durch die IT-Last bewirkte zu kühlende Leistung, auf der linken Achse die Außentemperatur und auf der Z-Achse die Leistungsaufnahme des Rechenzentrums abgetragen. Im Diagramm sind die Leistungsaufnahmen der IT (schwarze Fläche), der IT samt USV sowie die Gesamtleistungsaufnahme inkl. Kühltungskette dargestellt.

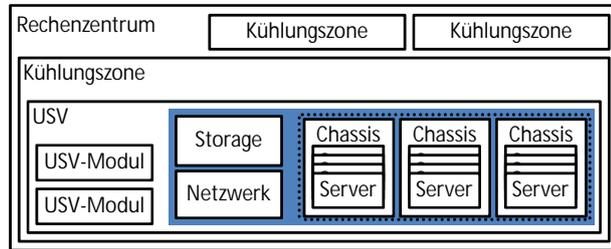


Abbildung 4.3: Hierarchischer Zusammenhang der Energiebedarfsmodelle

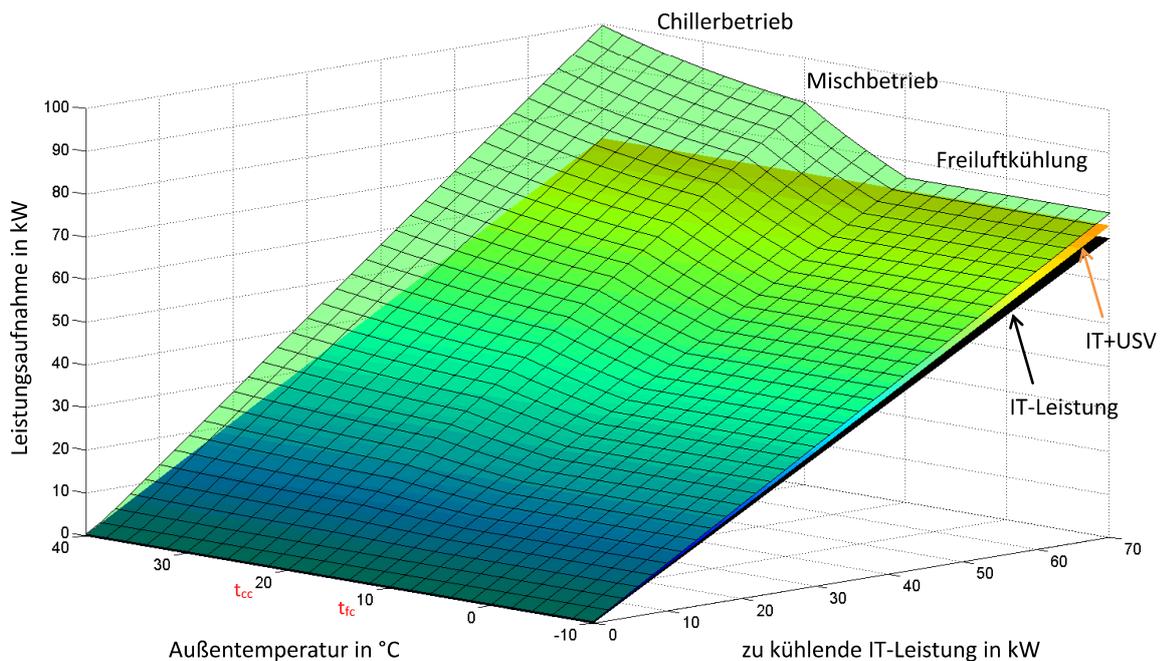


Abbildung 4.4: Modellierter Leistungsaufnahme eines Containerrechenzentrums

In Abhängigkeit von der Außentemperatur wird zwischen drei Betriebsbereichen unterschieden. Bis zu einer maximalen Temperatur t_{fc} kann die anfallende Kühlleistung durch die Freiluftkühlung (FK) bereitgestellt werden, die eine sehr geringe Leistungsaufnahme hat (hauptsächlich für Pumpen und Lüfter). Im Anschluss kann die FK nicht mehr die gesamte Kühlleistung aufbringen, aber zumindest Vorarbeit leisten. Die restliche Kühlleistung muss durch den Kaltwassersatz erfolgen. Ab der Temperatur t_{cc} kann die Freiluftkühlung keine Kühlleistung mehr aufbringen, so dass die erforderliche Kühlleistung nun komplett durch

den Kaltwassersatz erbracht werden muss. Im Mischbetrieb wird mit steigender Temperatur – entsprechend einer Studie der TU Berlin [22] – ein linearer Anstieg des Anteils für den Kaltwassersatz angenommen. Aufgrund fehlender Messdaten kann der genaue Zusammenhang leider nicht ermittelt werden.

Entsprechend den angegebenen Formeln ist die Gesamtleistungsaufnahme linear abhängig von der IT-Leistung. Im Chiller- und Mischbetriebsbereich ist die Leistungsaufnahme polynomiell abhängig von der Außentemperatur, im Freiluftkühlbereich ist sie aufgrund der Vereinfachung konstant. Der reale Zusammenhang ist hier vermutlich ebenso polynomiell.

4.3 Rechenzentrumsinternes Last- und Powermanagement

In Abschnitt 3.2 wurde das Konzept des hier exemplarisch eingesetzten Verfahrens für das interne LPM vorgestellt. Für die konkrete Umsetzung und den Einsatz im Rechenzentrumsverbund müssen jedoch noch weitere Konzepte entwickelt werden. Diese beziehen sich hauptsächlich auf die konkrete Selektion von Migrationen als auch auf die Anpassungen der zu koordinierenden VMs. Außerdem setzt das Verfahren in der ursprünglichen Implementierung [108] einen homogenen Server-Pool voraus, d.h., dass unterschiedliche Leistungsfähigkeiten und Energiebedarfe der Server nicht berücksichtigt werden.

4.3.1 Konzept der iterativen Migrationsentscheidung

Unter Einhaltung der in Abschnitt 3.2 beschriebenen Bedingungen zur Gewährleistung der zugesicherten Eigenschaften sind bei der Selektion der VM-Migration zwei Fragestellungen zu konkretisieren: 1. Von welchem Quellserver sollte eine VM migriert werden und 2. auf welchen Zielservers sollte diese verschoben werden? Zur Beantwortung dieser Fragen sei an das primäre Ziel des Verfahrenskonzeptes erinnert: Zur Minimierung des Energiebedarfs soll die Anzahl der benötigten aktiven Server minimiert werden. Zu diesem Zweck müssen Server möglichst viele VMs aufnehmen.

Zur Vereinfachung sollen in einem ersten Schritt homogene Server betrachtet werden. In diesem Fall kann ohne Berücksichtigung der spezifischen Leistungswerte der Hardware eine Heuristik angewendet werden, welche es vorsieht, VMs von den geringst befüllten Servern auf die meist befüllten Server zu migrieren. Im Konkreten wird eine Liste nach VM-Anzahl (#VM) aufsteigend sortierter Server durchlaufen und nacheinander getestet, ob VMs, wie in Abbildung 4.5 skizziert, verschoben werden können. Dieser Test läuft solange bis eine VM entsprechend den o.g. Bedingungen zur Migration freigegeben werden kann.

Hierdurch wird ein Server schrittweise leerräumt, was gegenüber der parallel entstandenen Implementierung in [108] den Vorteil hat, dass die Packrate erhöht wird, da eine Migration nicht dadurch beschränkt wird, dass sie nur dann durchgeführt wird, wenn alle VMs eines Servers migriert werden können. Als Nachteil stellt sich jedoch heraus, dass evtl. unnötige Migrationen durchgeführt werden, z. B. in dem Fall, wenn das Leerräumen eines ganzen Servers in nächster Zeit aufgrund der Ressourcenlage gar nicht möglich ist. Zur Minimierung bzw.

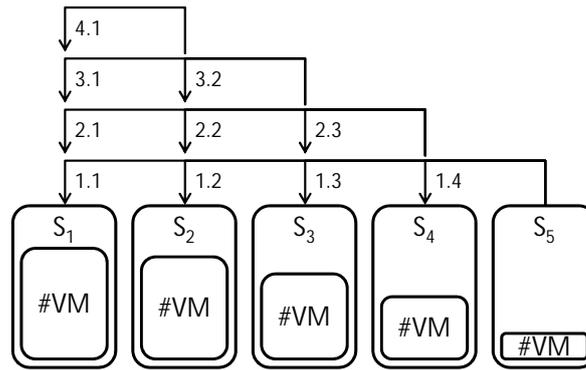


Abbildung 4.5: In der Reihenfolge der Bezifferung wird testweise versucht, VMs von Servern mit niedrigerer VM-Anzahl (#VM) auf Server mit höherer VM-Anzahl zu migrieren: $S_{i+k} \rightarrow S_i$

Verhinderung unnötiger Migrationen wurden Heuristiken eingesetzt, auf die an dieser Stelle jedoch nicht eingegangen wird.

Falls Engpässe in Zukunft auftreten, müssen eine oder mehrere VMs vom betreffenden Server auf andere Server migriert werden. Auch hier wird die Heuristik, möglichst volle Server weiter zu befüllen, angewendet. Wird kein passender Server gefunden, muss eine VM auf ihren Heimserver zurückmigriert werden. Ist dieser inaktiv, muss dieser folglich erst reaktiviert werden. Ist dieser allerdings aktiv, muss auf diesem erst einmal Platz geschaffen werden, denn der vorherige Test hat gezeigt, dass die VM nicht auf dem Heimserver alloziert werden kann, was bedeutet, dass sich auf diesem Fremd-VMs befinden. Mindestens eine dieser Fremd-VMs muss also verschoben werden, bevor die VM auf ihren Heimserver migriert werden kann. An dieser Stelle wird in eine Rekursion eingestiegen, bis eine mögliche Migration gefunden wird.

Die Laufzeitkomplexität der Konsolidierung und der Engpassauflösung ist jeweils $O(\#VM^3)$: Jede VM wird maximal einmal versucht auf einen anderen Server zu migrieren ($\rightarrow O(\#VM^1)$). Die Anzahl der Server ist abhängig vom Ressourcenbedarf der VMs, was im Worst-Case bedeutet, dass ein Server pro VM benötigt wird ($\rightarrow O(\#VM^2)$). Schließlich muss bei jeder Abfrage über den gesamten Betrachtungszeitraum die Ressourcenlage überprüft werden. Die Länge dieses Zeitraumes wird linear durch die VM-Anzahl beschränkt ($\rightarrow O(\#VM^3)$).

4.3.2 Heterogene Hardware

Bei der Verwendung von heterogenen IT-Komponenten wie Server und Chassis (im Falle von Serverblades) sollten zur Energieminimierung sowohl ihr Energiebedarf als auch ihre Leistungsfähigkeit mit in Betracht gezogen werden. An dieser Stelle kann statt der Metrik VM-Anzahl eine Bewertungsfunktion eingesetzt werden, die die unterschiedlichen Eigenschaften zusammenfasst. Hierarchische Abhängigkeiten, welche z. B. durch das zusätzliche Abschalten von aktiven Chassiskomponenten entstehen könnten, wenn alle Blades eines Chassis inaktiv sind, sollten grundsätzlich auch berücksichtigt werden. Zum Herausstellen des Einflusses komplexerer Strukturen reicht es allerdings allein heterogene Hardware zu betrachten, so dass

der Einfachheit halber an dieser Stelle von hierarchischen Strukturen abgesehen wird. Im Folgenden wird also untersucht, welche Auswirkungen die Verwendung von heterogener Hardware auf den durch das LPM erzielten bzw. verursachten Energiebedarf hat. Insbesondere soll eruiert werden, inwiefern die konkrete Ausprägung der Bewertungsfunktion den Energiebedarf beeinflusst.

Zur Erstellung der Bewertungsfunktion werden die Anzahl der betriebenen VMs sowie die Leistungsfähigkeit und der Energiebedarf des Servers betrachtet. Wie auch schon bei homogenen Servern deutet eine geringe VM-Anzahl darauf hin, dass der Server mit wenig Aufwand geleert werden kann. Zur Berücksichtigung der unterschiedlichen Leistungsfähigkeit kann zum einen die maximale Leistungsfähigkeit, also bei voller Auslastung – sie stellt damit eine Kapazitätsangabe dar – und zum anderen die Leistungsfähigkeit bei geringeren Auslastungen betrachtet werden. Letztere ist allerdings nur im Zusammenhang mit der Energieaufnahme aussagekräftig. Daher sollte zu diesem Zwecke die Leistungseffizienz im dynamischen Modus ($0\% < \text{Auslastung} < 100\%$) betrachtet werden. Diese kann aufgrund des linearen Energiebedarfsmodells (siehe Abschnitt 4.2.1) durch einen einzigen Effizienzwert abgebildet werden. Kennwerte zum maximal und dynamisch ausgelasteten Betrieb sind damit berücksichtigt, fehlt also noch ein Kennwert bei minimaler Auslastung. Hier ist der Energiebedarf im Idle charakteristisch.

Server mit großer Kapazität oder hoher Leistungseffizienz sollten eher als Ziel anvisiert werden, da sie mehr VMs aufnehmen können bzw. wenig Energie zum Betrieb weiterer VMs benötigen. Beim Idle-Energiebedarf ist ein hoher Wert jedoch nicht wünschenswert. In die Bewertungsfunktion sollte dieser Wert daher entgegengesetzt einfließen. Zusammenfassend werden die folgenden vier Werte bzgl. eines Servers *srv* berücksichtigt:

- $\#VM_{srv}$: Anzahl betriebener VMs
- $power_{srv, idle}$: Energiebedarf im Idle
- cap_{srv} : Leistungskapazität
- eff_{srv} : Leistungseffizienz

In der Bewertungsfunktion werden die Werte additiv gewichtet. Da sie in unterschiedlichen Dimensionen und Größenordnungen vorliegen, ist eine Gewichtung von den konkreten Größen abhängig. Die jeweilige Relevanz der Parameter lässt sich so anhand dieser Gewichte nicht ermitteln. Hierzu muss die Gewichtung im Bereich $[0;1]$ liegen, wofür eine Normalisierung der Parameter notwendig ist. Diese wird mit dem jeweiligen maximalen Wert aller betrachteten Server vollzogen. Bei der VM-Anzahl wird die Gesamtanzahl im betrachteten Pool herangezogen. Es ergibt sich somit die folgende Bewertungsfunktion *rate* für einen Server *srv*:

$$rate_{srv} = \alpha_1 \cdot \frac{\#VM_{srv}}{\#VM_{total}} + \alpha_2 \cdot \left(1 - \frac{power_{srv, idle}}{max_power_idle}\right) + \alpha_3 \cdot \frac{cap_{srv}}{max_cap} + \alpha_4 \cdot \frac{eff_{srv}}{max_eff} \quad (4.7)$$

4 Konzeption samt Komponentenevaluation

Um einen grundsätzlichen Eindruck des Einflusses der vier Parameter zu gewinnen, wird vorab eine isolierte Analyse durchgeführt. Der zu untersuchende Parameter i erhält die Gewichtung $\alpha_i = 1$, die Gewichtungen der anderen Parameter werden auf null gesetzt ($\alpha_{j \neq i} = 0$).

Um heterogene Hardware zu simulieren, wird der aktuell verwendete Standardserver, welcher eine normierte Rechenleistungsfähigkeit (CPU) von 100 % hat, in seiner Rechenleistung mit einem Faktor aus dem Intervall [1;4] und das zugrunde liegende Energiebedarfsmodell mit einem Faktor aus dem Intervall [0,5;2] variiert. In der Simulation können auf diese Weise (zufallsbedingt) heterogene Komponenten instantiiert werden. Insgesamt werden 100 verschiedene Szenarien mit 200 VM und einer simulierten Zeit von zehn Tagen ausgewertet. Die Szenarien unterscheiden sich in der konkreten Wahl der VMs und der eingesetzten Hardware (jeweils zehn unterschiedliche).

Das Ergebnis der Voranalyse ist in Abbildung 4.6 anhand eines Boxplots dargestellt. Hierin ist der Energiebedarf in Relation zum Ergebnis der ursprünglichen Metrik (VM-Anzahl) angegeben. Das beste Ergebnis wird durch den Effizienzparameter erreicht. Mit diesem lassen sich nahezu immer bessere Ergebnisse erzielen als mit der ursprünglichen Metrik. Mit dem Kapazitätsparameter wurden sowohl deutlich bessere als auch deutlich schlechtere Ergebnisse als mit dem Parameter VM-Anzahl erzielt. Im Durchschnitt liegen die beiden Parametereinflüsse aber gleich auf. Der Parameter für die Idle-Power hat den schlechtesten Einfluss, um ein energiebedarfsminimierendes Ergebnis zu erzielen.

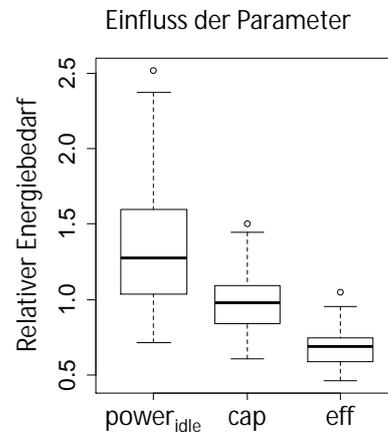


Abbildung 4.6: Einfluss bzw. Güte der einzelnen Parameter bzgl. des Energiebedarfs in Relation zum Energiebedarf bei Verwendung des Parameters VM-Anzahl

Insgesamt lässt sich eine große Streuung erkennen. Anhand der durchschnittlichen Ergebnisse hat der Effizienzparameter zwar den besten Einfluss, aber die Streuung legt nahe, dass eine Mitberücksichtigung der anderen Parameter für eine weitere Verbesserung sorgen wird. In welchem Verhältnis die Berücksichtigung erfolgen sollte, also welche konkreten Gewichtungen den einzelnen Parametern zukommen sollten, wird in der nun folgenden Analyse ermittelt.

Das Verfahren zur Ermittlung der optimierten Gewichtungen erfolgt durch kreuzweise Variation der Gewichtungen im Untersuchungsintervall [0; 1]. Initial beträgt die Schrittweite $\frac{1}{3}$. Der Durchlauf einer kreuzweisen Variation entspricht einer Iteration. Das beste Ergebnis wird als Ankerwert für eine nachfolgende Iteration genommen. In dieser wird das Untersuchungsintervall auf ein Viertel reduziert und um die aktuellen Gewichtungswerte zentriert. Die Schrittweite wird ebenso auf ein Viertel reduziert. Das Verfahren endet, wenn die erzielte Verbesserung von Iteration zu Iteration unter 5 % fällt.

In Abbildung 4.7 sind die optimierten Gewichtungen für die untersuchten Szenarien angegeben. Es ist eine sehr große Streuung zu erkennen. Eine detaillierte Betrachtung der Einzelergebnisse zeigt, dass die Gewichtung stark von der eingesetzten Hardware und den betriebenen VMs abhängt. Inwiefern eine allgemeine Gewichtung nützlich ist, lässt sich in Abbildung 4.8 erkennen. Die gleichgewichtige Betrachtung aller Parameter besitzt bereits eine im Durchschnitt deutliche Reduzierung des Energiebedarfs. Mit den optimierten Gewichtungen lässt sich der Energiebedarf noch um weitere 7% reduzieren. Die allgemeine Gewichtung, die dem Durchschnitt der Szenario spezifisch optimierten Gewichtungen entspricht, ordnet sich zwischen den beiden Ergebnissen ein und kann somit sinnvoll für eine allgemeingültige Bewertungsfunktion herangezogen werden.

Für einen zusammenfassenden Vergleich der unterschiedlichen Parametergewichtungen ist in Tabelle 4.1 das jeweilige Energiebedarfsverhältnis gegenüber dem Parameter VM-Anzahl angegeben. Es ist herauszustellen, dass die Parametergleichgewichtung schlechter abschneidet als die bloße Betrachtung des Effizienzwertes. Hier ist das Verhältnis der eher unbedeutenden Parameter zum Effizienzwert im Durchschnitt zu groß. Durch Korrektur der Verhältnisse lässt sich eine allgemeine Gewichtung mit verbessertem Ergebnis bilden, die nur noch durch die spezialisierte Optimierung übertroffen wird.

Die Untersuchung der unterschiedlichen Bewertungsfunktionen zeigt auch, dass die Effizienz eines LPM bzgl. des Energiebedarfs sehr von der Policy, wonach optimiert werden soll, abhängt. In dieser Analyse wurde gezielt versucht, den Energiebedarf zu senken, und dennoch streut dieser sehr stark. Kommen nun hierarchische Strukturen wie Chassis und auch Klimazonen hinzu oder werden gar Policies hinzugezogen, die z. B. Verfügbarkeitsaspekte berücksichtigen, kann sich der Energiebedarf auf völlig unterschiedlichen Niveaus befinden. Möchte man also den Energiebedarf eines Rechenzentrums, das sich dynamisch an die aktuellen

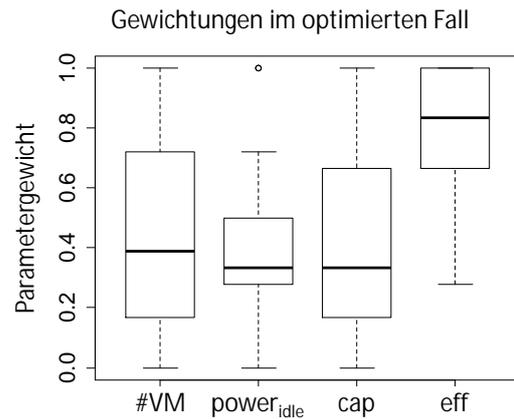


Abbildung 4.7: Optimierte Gewichte für die einzelnen Parameter

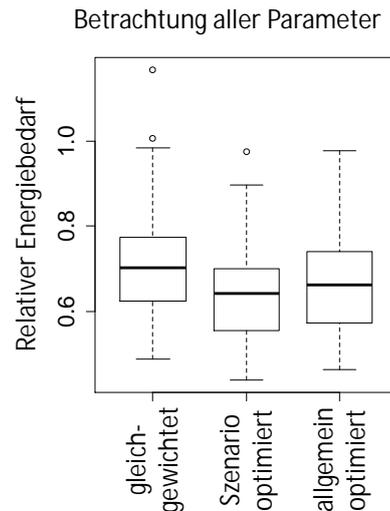


Abbildung 4.8: Güte bei gleicher Gewichtung, bei für das jeweilige Szenario optimierten Gewichtungen und bei Anwendung der allgemein optimierten Gewichtungen

Parameter/-gewichtung	$\frac{\text{Energiebedarf}_{\text{param}}}{\text{Energiebedarf}_{\text{VM-Anzahl}}}$
VM-Anzahl	1,00
Idle-Power	1,33
Kapazität (CPU)	0,98
Leistungseffizienz	0,68
Gleichgewichtung	0,71
Szenario optimierte Gewichtung	0,64
Allgemein optimierte Gewichtung	0,66

Tabelle 4.1: Durchschnittliche Güte bei unterschiedlichen Parametergewichtungen (je geringer, desto besser)

Erfordernisse anpasst, abschätzen, so muss neben den Hardwarekomponenten auch das LPM einbezogen werden. Die konkrete Auswahl an zu betreibenden VMs hat ebenso einen großen Einfluss auf den Energiebedarf. In flexiblen Cloud-Umgebungen wirkt sich dieser Umstand verstärkt aus, da sich hier der Pool an zu betreibenden VMs jederzeit ändern kann. Um dieses berücksichtigen zu können, muss das betrachtete LPM allerdings noch erweitert werden, was im folgenden Abschnitt beschrieben wird.

4.3.3 Dynamisierungserweiterungen

Das bisher vorgestellte Konzept des internen LPM sieht keine Anpassungen der Auslastungsprofile bzw. der im Pool befindlichen zu koordinierenden VMs zur Laufzeit vor. Insbesondere für die Live-Migration zwischen Rechenzentren muss die Entnahme und das Hinzufügen von VMs möglich sein. Hierbei muss darauf geachtet werden, dass die Eigenschaften (SLO-Garantien) des Verfahrens erhalten bleiben.

Wenn sich das Auslastungsprofil einer VM ändert, müssen die Prognosewerte aktualisiert werden, wofür in der aktuellen Implementierung des Verfahrens bereits eine Fehlerkorrektur eingesetzt wird, soweit statistisch relevante Abweichungen auftreten. Auf die aktualisierten Werte kann die dynamische Verteilung direkt reagieren. Auf die sichere Verteilung hat die Fehlerkorrektur jedoch keinen Einfluss, so dass sich über die Zeit nicht nur die Packrate verschlechtern kann, sondern auch insbesondere die zugesicherten Eigenschaften der sicheren Verteilung verletzt werden können. Gleiches gilt, wenn VMs hinzugefügt oder entfernt werden. Die sichere Verteilung sollte also bei Bedarf aktualisiert werden.

Die Problemstellung besteht hier also darin, die Zuordnungen für eine sichere Verteilung zu verändern und dabei die aktuelle dynamische Verteilung zu berücksichtigen. Die Bedingungen über die zeitlichen Einschränkungen und Migrationskanten (siehe Migrationsgraph in Abschnitt 3.2) müssen dabei wie bei der dynamischen Verteilung berücksichtigt werden. Im Gegensatz zum dynamischen Verfahren können hier allerdings für mehrere VMs gleichzeitig neue Zuordnungen bestimmt werden, da keine Daten verschoben werden müssen, sondern lediglich ein (neuer) logischer Zusammenhang festgelegt werden muss.

In Abbildung 4.9 wird der Zusammenhang von sicherer (S) und dynamischer (A) Verteilung

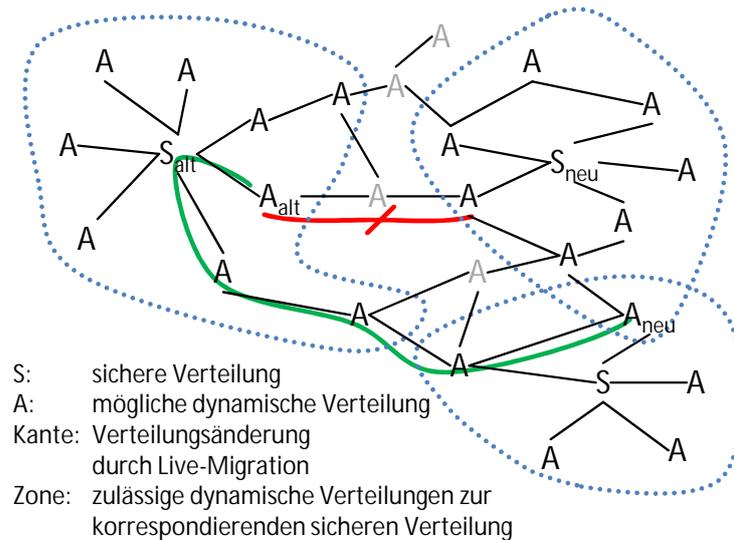


Abbildung 4.9: Schematische Darstellung der Überführung des Systems zu einer neuen sicheren Verteilung (Momentaufnahme)

mit der Beschränkung des sicheren Rückwegs skizziert. Für eine Menge an zu betreibenden VMs sind grundsätzlich mehrere sichere Verteilungen möglich. Von diesen ausgehend können eine bestimmte Anzahl an Veränderungen (Migrationen) ausgeführt worden sein, was zu einer dynamischen Verteilung führt. In der Abbildung sind die dynamischen Verteilungen, bei denen zum aktuellen Zeitpunkt der sichere Rückweg garantiert ist, eingekreist. Wird nun eine andere sichere Verteilung angestrebt (S_{neu}), so muss in jedem Verteilungszustand ein sicherer Rückweg garantiert werden. In der Abbildung ist ein gültiger Pfad an Migrationen dargestellt, welcher über eine weitere sichere Verteilung führt. Da es, wie auch hier im Beispiel dargestellt, nicht immer möglich sein wird, die angestrebte sichere Verteilung innerhalb eines Zeitschritts zu erreichen, ist es notwendig, das zeitliche Verhalten zu berücksichtigen. Denn mit jeder Migration vergeht Zeit, wodurch sich der Systemzustand ändert. Der Pfad im Beispiel könnte somit nach Durchführung einiger Migrationen nicht mehr gültig sein.

Problemformalisierung

Die Problematik zur Veränderung der sicheren Verteilung bei gegebener dynamischer Verteilung wird im Folgenden anhand einer ILP-Beschreibung formalisiert. In dieser Form wird zugleich ein Komplexitätseindruck für ein Lösungsverfahren vermittelt.

Das Verfahren basiert auf zwei unterschiedlichen Verteilungen, die jeweils eine Zuordnung der m VMs zu den n Servern definiert. Diese werden in der ILP-Beschreibung durch die binären Matrizen S (sichere Verteilung) und A (dynamische Verteilung) umgesetzt, wobei ein 1-Eintrag in der i -ten Zeile und j -ten Spalte bedeutet, dass VM i Server j zugeordnet ist:

$$S^{m \times n}, (\text{analog für } A)$$

$$\forall i : \sum_j S_{ij} = 1, \forall j : \sum_i S_{ij} = 1$$

4 Konzeption samt Komponentenevaluation

Damit sichergestellt werden kann, dass die beiden Verteilungen gültig sind, müssen zum einen genügend Ressourcen bereitgestellt werden und zum anderen müssen sie in Bezug zueinander stehen.

Der Bezug zwischen der sicheren und dynamischen Verteilung wird durch die Migrationen gebildet. Zur Abbildung dieser wird eine Migrationsmatrix M verwendet, die festhält, wie viele VM ausgehend von der aktuellen Verteilung A von einem Server r zu einem Server c zurück migriert werden müssen, damit die sichere Verteilung S erreicht wird. Die Anzahl lässt sich dadurch ermitteln, dass für jede VM i überprüft wird, ob sie in der sicheren Verteilung auf Server c (S_{ic}) und in der dynamischen Verteilung auf Server r (A_{ir}) zugeordnet ist. In einem ILP kann dies bestimmt werden, indem die Summe der beiden Werte gebildet und um eins verringert wird. Ist die Migration zutreffend, ist das Ergebnis 1, ansonsten ≤ 0 . In Verbindung mit einer Maximumfunktion mit 0 werden negative Werte ausgeschlossen:³

$$M^{n \times n}(A, S), \\ M_{rc} = \sum_i \max(0, A_{ir} + S_{ic} - 1), M_{rr} = 0$$

Für die Deadlockvermeidung dürfen keine Zyklen bestehen. Um dies zu detektieren, wird die binäre Adjazenzmatrix Ad verwendet, in der transitiv auf Grundlage der Migrationsmatrix M die möglichen Migrationsabhängigkeiten aufgezeigt werden. Initial werden dazu die Erreichbarkeiten der Migrationsmatrix übernommen. Da in dieser allerdings Werte größer 1 vorkommen können, werden die Werte mit der Gesamt-VM-Anzahl normiert. Der jeweilige Eintrag in der Adjazenzmatrix muss damit nur größer gleich dem normierten Wert sein, damit auf 1 aufgerundet wird. Eine transitive Verbindung kann analog zum obigen Vorgehen bei der Fragestellung, ob zwei Werte eine 1 haben, durch die Summe der beiden Werte und anschließenden Inkrementierung ermittelt werden:

$$Ad^{n \times n}(M), \\ Ad_{rc} \geq \frac{M_{rc}}{n}, Ad_{rc} \geq Ad_{rj} + Ad_{jc} - 1$$

Damit keine Zyklen im Migrationsgraphen entstehen, muss die Bedingung $\forall j : Ad_{jj} = 0$ erfüllt sein.

Die aufgeführten Matrizen beziehen sich auf einen einzelnen Zeitpunkt. Um das zeitliche Verhalten einzubringen, müssen mehrere Zeitpunkte betrachtet und damit entsprechend viele Matrizen definiert werden. Eine Verbindung zwischen den Matrizen wird durch den Zusammenhang geschaffen, dass aufeinanderfolgende dynamische Verteilungen sich in maximal einem Wert unterscheiden dürfen. In der folgenden Formulierung wird sichergestellt, dass nur maximal eine neue VM-Zuordnung stattfindet, indem eine Subtraktion zum nachfolgenden

³Der Filterungseffekt durch die Maximumfunktion wird dadurch erreicht, indem eine geeignete Zielfunktion gewählt wird, durch die die entsprechenden Variablen minimiert werden.

Wert durchgeführt wird. Die Maximumfunktion filtert auch hier negative Werte heraus:

$$\forall t : \sum_i \sum_j \max(A(t+1)_{ij} - A(t)_{ij}, 0) \leq 1$$

Eine Abfrage über entfernte Zuordnungen ist nicht notwendig, da diese Information redundant zur obigen Feststellung ist.

Die Kapazität eines Servers muss zu jedem Zeitpunkt t eingehalten werden. Die summierte Last der auf Server j allozierten VMs muss dementsprechend kleiner gleich seiner Kapazität cap_j sein:

$$\forall t, j : \sum_i A(t)_{ij} \cdot load_i(t) \leq cap_j$$

Analog gilt dieser Zusammenhang auch für die sichere Verteilung bzgl. der maximal zu erwartenden Last. Die Prognosen für den benötigten Ressourcenbedarf einer VM i werden durch das Vorhersageverfahren geliefert.

Im LPM wird zu jedem Zeitpunkt sichergestellt, dass die aktuelle Verteilung für eine gewisse Zeitspanne ($[t, t+k]$) keine Ressourcenengpässe verursacht. Streng genommen müssten daher zu jedem Zeitpunkt weitere Matrizen $A(t, k)$ verwendet werden. Allerdings reicht es, dies nur für den letzten Zeitpunkt anzuwenden, denn die Verfahrensbedingungen können während der Koordination durch das ILP entfallen. Die Kernaufgabe besteht darin einen Migrationspfad zu finden, der zur optimierten sicheren Verteilung führt. Hierbei muss das ILP für die genügende Bereitstellung von Ressourcen sorgen. Erst zum Schluss, wenn also das LPM das Management wieder übernimmt, muss ein den Bedingungen valider Zustand garantiert sein. Auch die Matrizen der sicheren Verteilung sowie der Migrationen und Zyklusdetektion sind damit nur zum Schluss relevant.

Eine ILP-Beschreibung besitzt eine fest definierte Anzahl verwendeter Variablen (bzw. Matrizen). Das bedeutet, dass nur ein begrenztes Zeitintervall analysiert werden kann, um das Ziel einer sicheren Verteilung mit minimal benötigter Anzahl an Servern zu erreichen. Wie groß dieses Intervall hierzu sein muss, kann nicht im Voraus festgestellt werden, so dass die Größe entweder großzügig gewählt oder iterativ bestimmt werden müsste. Aufgrund der Berechnungskomplexität sollte das Intervall jedoch möglichst klein sein, denn mit steigender Größe wächst die Anzahl an ILP-Variablen und -Beschränkungen. Hiervon abgesehen ist der gesamte Optimierungsvorgang an weitere, noch nicht vorgestellte Bedingungen zur erwarteten Ersparnis geknüpft.

Repacking Heuristiken

Wie oben skizziert stellt es ein komplexes Problem dar, den Betrieb von einer sicheren Verteilung zu einer anderen (optimierten) zu führen. Um dennoch einer Verschlechterung der sicheren Verteilung über die Zeit vorzubeugen, die durch viele Veränderungen der zu betreibenden VMs hervorgerufen sein kann, werden in diesem Abschnitt Heuristiken vorgestellt, die eine

Gütegarantie⁴ besitzen. Hierdurch kann der Nutzen eines optimalen Verfahrens abgeschätzt werden (siehe Abschnitt 4.3.4).

Für die heuristische Optimierung wird die Problemstellung vereinfacht: Es wird nicht versucht, ein optimales Ziel zu erreichen, das erst durch mehrere Migrationen erreicht werden kann, sondern – ausgehend von der aktuellen Verteilung – eine neue, „bessere“ sichere Verteilung gebildet, die auch zur aktuellen Verteilung korrespondiert. In Bezugnahme auf das Beispiel in Abbildung 4.9 bedeutet dies, dass für eine gegebene Verteilung A eine andere, „bessere“ sichere Verteilung S_{neu} gesucht wird, in deren eingekreister Region sich A momentan auch befindet. Die neue sichere Verteilung ist somit bereits gültig. Es müssen keine Migrationen durchgeführt werden, sondern lediglich die Zuordnungen für die sichere Verteilung aktualisiert werden. Auch hierbei müssen die Verfahrensbedingungen betrachtet werden. Es dürfen also keine Zyklen im Migrationsgraphen entstehen. Um die Komplexität der Heuristiken einfach zu halten, wird darauf geachtet, dass sich der Betrachtungszeitraum (siehe Gleichung (3.1)) nicht verlängert. Dadurch ist es bei Neuuzuordnungen nicht zusätzlich notwendig, auf mögliche Ressourcenengpässe zu prüfen.

Hintergrund-Repacking

Während des Betriebs befindet sich das System i.d.R. in einer dynamischen Verteilung, in der sich mehrere Server im Stand-by befinden. Da auf diesen keine VMs aktiv alloziert sind, existieren im Migrationsgraphen keine gerichteten Migrationskanten zwischen ihnen oder zu aktiven Servern (siehe Abbildung 4.10, grau eingekreist). Das Verändern der sicheren Zuordnungen auf den inaktiven Servern führt somit weder zur Verlängerung des Betrachtungsintervalls noch ist es möglich, dass Zyklen im Migrationsgraphen entstehen. Somit kann die sichere Verteilung auf den inaktiven Servern komplett neu durchgeführt und optimiert werden. Wie auch initial bei Erstellung der gesamten sicheren Verteilung, kann auch für diese Servergruppe ein offline Bin-Packing-Verfahren verwendet werden. Das Hintergrund-Repacking kann jederzeit eingesetzt werden.

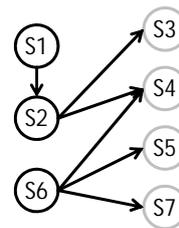


Abbildung 4.10: Im Migrationsgraphen existieren auf ausgeschalteten Servern (grau eingekreist) keine ausgehenden Migrationskanten.

Hinzufügen von VMs

Beim Hinzufügen von neuen VMs gilt es, diese in die bereits bestehende sichere Verteilung so einzubinden, dass möglichst wenig weitere Server reserviert werden müssen. Auch hierbei kann ein Bin-Packing-Verfahren eingesetzt werden. Im Gegensatz zum initialen Packing, wo sämtliche Server noch frei sind, sind diese nun bereits zu großen Teilen befüllt. Daher

⁴Mit der Gütegarantie wird der maximale Abstand zum Optimum beschrieben, welcher eine Lösung des betreffenden Verfahrens haben kann. Dieser Abstand wird als Quotient vom Ergebnis des eingesetzten Verfahrens zum Optimum angegeben.

spricht man hier von online Bin-Packing. In der Praxis hat dies allerdings nur Einfluss auf die zugesicherte Güte des Packings, nicht jedoch auf die Anwendbarkeit. Für das Hinzufügen von VMs kann somit das gleiche Bin-Packing-Verfahren eingesetzt werden wie beim initialen Packing.

Entfernen von VMs

Das Entfernen von VMs hat zur Folge, dass Kapazitäten, die vorher für die sichere Verteilung reserviert wurden, nun frei sind. Im schlimmsten Fall könnten die Freiräume so entstehen, dass auf allen Servern nur eine VM zugeordnet ist, was einen großen Überschuss an zu reservierenden Servern darstellen würde. Die Server könnten für evtl. andere Einsatzzwecke nicht freigegeben werden. Freie Kapazitäten sollten daher bestmöglich durch andere VMs genutzt werden.

Der Heimserver einer zu entfernenden VM kann aktiv oder auch ausgeschaltet sein. Ist dieser ausgeschaltet, so kann das Hintergrund-Repacking die freien Kapazitäten wieder effizienter verteilen. Bei aktiven Servern ist diese Heuristik nicht anwendbar. Hierbei kann die folgende eingesetzt werden.

Direktes Repacking

Um freie Kapazitäten für eine sichere Zuordnung auf einem aktiven Server besser zu nutzen, kann wie folgt verfahren werden: Aus den zur Zeit allozierten Fremd-VMs können ein oder mehrere VMs (z. B. via Best-Fit-Verfahren) ermittelt werden, um den betreffenden Server zu ihrem neuen Heimserver zu machen. Der Betrachtungszeitraum würde sich hierdurch verkürzen, da Migrationskanten entfallen würden. Die Entstehung von evtl. Migrationszyklen wäre damit ausgeschlossen.

Der Vorgang ist beispielhaft in Abbildung 4.11 dargestellt: VM2 wird aus dem betrachteten Gesamt-VM-Pool entfernt. Dadurch entstehen auf dem Heimserver S1 Freiräume für die sichere Zuordnung. Die VM4 ist auf Server S1 eine Fremd-VM. Ihr Heimserver ist aktuell Server S2. Wird Server S1 ihr neuer Heimserver, so kann die durch VM2 freigegebene Kapazität durch VM4 belegt werden. Die Migrationskante von Server S1 zu S2 würde entfallen, da sich auf S1 keine Fremd-VMs mehr befinden.

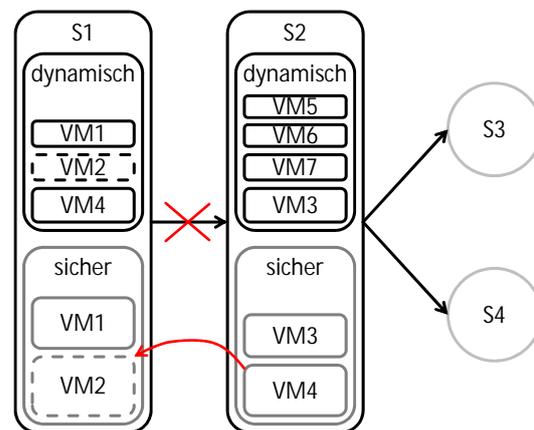


Abbildung 4.11: Durch Entfernen von VMs (hier VM2) aus dem VM-Pool entstehen in der sicheren Verteilung freie Kapazitäten, die durch partielle Neuallokation wieder belegt werden können.

Rekursives Repacking

Im Beispiel in Abbildung 4.11 wird die freie Kapazität durch eine VM (VM4) gefüllt, die

aktuell einem anderen aktiven Server sicher zugeordnet ist. Hierdurch wird die Problematik auf diesen Server verlagert, da nun hier wiederum ein Freiraum entsteht. Für diesen Fall kann das direkte Repacking rekursiv auf diesen und alle nachfolgenden Server angewendet werden.

Um den Effekt der rekursiven Entstehung weiterer freier Kapazitäten zu vermeiden, sollten vorzugsweise solche VMs bestimmt werden, deren Heimserver aktuell inaktiv sind. Die auf den inaktiven Servern entstehenden freien Kapazitäten können in einem anschließenden Prozess, nachdem alle gewünschten VMs aus dem VM-Pool entfernt wurden, durch das Hintergrund-Repacking neu belegt werden.

Veränderung eines VM-Auslastungsprofils

Wenn sich das Auslastungsprofil einer VM über die Zeit verändert, dann wird dies i.d.R. zur Folge haben, dass sich die für die sichere Verteilung zu reservierende Kapazität ebenfalls verändert. Wenn sie kleiner werden sollte, dann entspricht dies dem Vorgang, dass eine VM aus dem VM-Pool entfernt wurde. Wird sie jedoch größer und die Kapazität des Heimservers wird überschritten, dann müssen VMs anderen Servern zugeordnet werden – vorzugsweise inaktiven. Zwar können hierbei neue Migrationskanten entstehen, allerdings keine Migrationszyklen. Neue Migrationskanten bedeuten jedoch zwangsläufig auch eine Vergrößerung des Betrachtungszeitraumes, weshalb hierbei eine Engpassüberprüfung durchgeführt werden muss.

Nach der Neuordnung entstehen auf dem aktuellen Server jedoch wieder freie Kapazitäten. Auch dies entspricht dem Entfernen von VMs, so dass in beiden Fällen, also sowohl bei der Verringerung als auch bei der Vergrößerung des Ressourcenbedarfs, das in Abschnitt *Entfernen von VMs* beschriebene Vorgehen verwendet werden kann.

Um einen besseren Überblick zu erhalten, in welchen Fällen die Heuristiken angewendet werden sollten, ist in Abbildung 4.12 resümierend ein entsprechendes Flussdiagramm gegeben. Grundsätzlich sollte nach Abschluss der Veränderungsprozesse das Hintergrund-Repacking ausgeführt werden, da dieses einen globaleren Blick auf die Situation besitzt als die speziellen Heuristiken.

Die vorgestellten Erweiterungen erlauben an unterschiedlichen Zeitpunkten Veränderungen. Diese können Einfluss auf den Betrieb haben und so möglicherweise das Verhalten des LPM ändern (z. B. Peaks, „level switches“). Da dies i.d.R. schädliche Auswirkungen auf die Güte eines Modells hat [74], ist bei ihrer Umsetzung darauf zu achten, dass das Verhalten nicht zu stark vom Normalbetrieb abweicht. Insbesondere die Aufnahme von neuen VMs muss mit Bedacht ausgeführt werden, denn das LPM muss jede VM in das bisherige System integrieren, indem es evtl. zusätzliche Kapazitäten bereitstellt und dabei weiterhin die Zusicherungsbedingungen erfüllt. Zu beachten gilt hier die Reihenfolge der Abarbeitung von aufkommenden Engpässen, der Konsolidierung und der Aufnahme bzw. dem Entfernen von VMs. Da aufkommende Engpässe kritisch sind, müssen diese priorisiert abgearbeitet werden. Bei der Fragestellung, wann konsolidiert und wann der VM-Pool angepasst werden soll, ist jedoch keine einfache Antwort möglich: Erhält die Konsolidierung den Vorzug, kann die VM-Pool-Anpassung

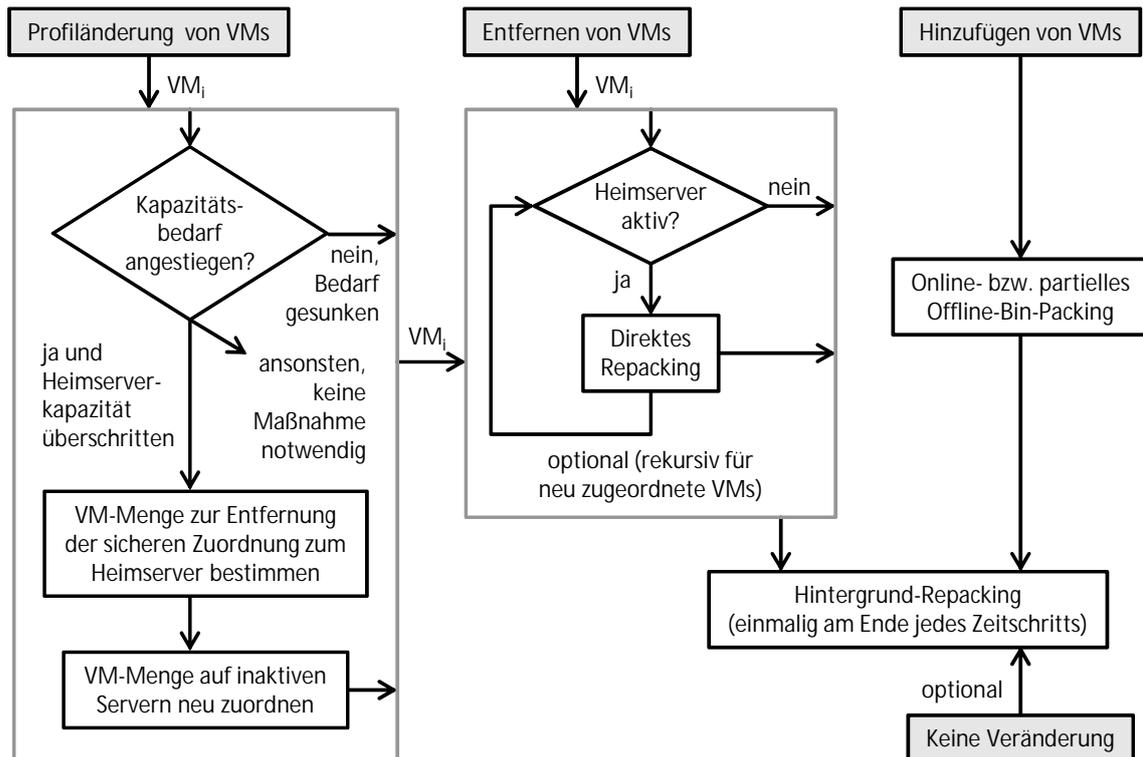


Abbildung 4.12: Flussdiagramm für die Anwendung der Repacking-Heuristiken zur Anpassung und Optimierung der sicheren Verteilung

mitunter sehr viel Zeit beanspruchen, was zu verzögerten globalen Migrationen führt und mögliche Einsparungen reduziert. Erhält die VM-Pool-Anpassung jedoch den Vorzug, wird das LPM-Verhalten sehr stark beeinflusst, so dass es beispielsweise durch viele VM-Aufnahmen dazu kommen könnte, dass zuerst viele Server hochgefahren werden und erst im Anschluss wieder konsolidiert wird (Peak). Um hier einen „fairen“ Kompromiss zu bilden, wird eine Regelung eingesetzt, die die Priorisierung der aktuellen Lage anpasst und dadurch sowohl für eine zügige Anpassung des VM-Pools sorgt als auch Konsolidierungen zulässt.

4.3.4 Auswirkungen der Modifikationen

Die Modifikationen zum Aufnehmen und Entfernen von VMs sollen nun im Folgenden evaluiert werden. Insbesondere der Einfluss der vorgestellten Heuristiken zur Anpassung der sicheren Verteilung und die Auswirkungen dessen auf die sichere Verteilung selbst und den dynamischen Betrieb sollen eruiert werden. Ferner soll untersucht werden, welchen Einfluss die Modifikationen auf das Skalierungsverhalten haben.

Zur Evaluation wurden homogene Server und Datensätze realer VMs verwendet, worauf in Abschnitt 5.1 eingegangen wird. Insgesamt wurden 100 unterschiedliche Simulationsläufe durchgeführt. Jeder Lauf simuliert 10 Tage und beinhaltet einen VM-Pool von 300 VM (initial 150 VM, maximal 300 VM). Um eine Dynamik im VM-Bestand zu simulieren, wurden Opera-

4 Konzeption samt Komponentenevaluation

tionsverläufe definiert, die alle 4 bis 8 Stunden VMs hinzugefügt oder entfernt haben, wobei die Veränderungsrate bei bis zu 50 % der derzeit beinhalteten VMs liegt. Die Simulationsläufe unterscheiden sich in der Auswahl des VM-Pools und in der Anwendung der Operationsverläufe (jeweils zehn unterschiedliche).

Auswirkungen auf die Verteilungen

Die Heuristiken sollen bewirken, dass die Anzahl benötigter Server in der sicheren Verteilung reduziert wird. Zu erhoffen wäre hierdurch auch, dass die dynamische Verteilung davon profitiert und damit eine Senkung des Energiebedarfs möglich ist. Um diese Zusammenhänge festzustellen, wurden die Evaluationsszenarien mehrfach simuliert – einmal ohne Repacking und jeweils einmal mit den folgenden sechs Heuristiken und -kombinationen:

- [b] Hintergrund-Repacking,
- [cb] permanentes Hintergrund-Repacking – Durchführung in jedem Zeitschritt, ungeachtet dessen, ob Veränderungen vorkommen oder nicht –,
- [d] direktes Repacking,
- [db] direktes und Hintergrund-Repacking,
- [dr] direktes und rekursives Repacking,
- [drb] direktes, rekursives und Hintergrund-Repacking.

In Abbildung 4.13 sind die Ergebnisse dieser Untersuchung als Boxplot veranschaulicht. Es wird die durchschnittliche Anzahl benötigter Server in Relation zur Anzahl, wenn kein Repacking eingesetzt wird, angegeben.

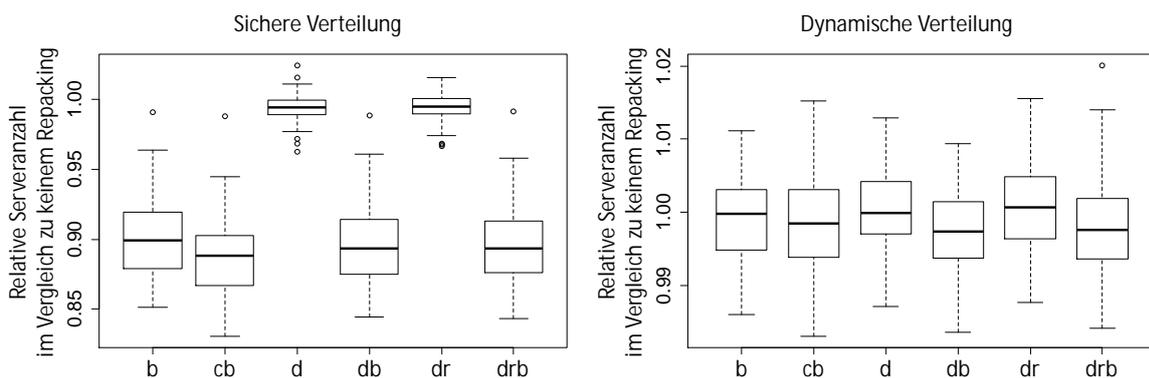


Abbildung 4.13: Auswirkungen der unterschiedlichen Repacking-Heuristiken bzgl. der Anzahl benötigter Server

Die Anzahl der für die sichere Verteilung reservierten Server kann bis zu durchschnittlich 10 % gesenkt werden, allerdings hat dies keinen nennenswerten Einfluss auf den dynamischen

Betrieb. Die Anzahl der dynamisch benötigten Server ist nahezu identisch. Eine wichtige Erkenntnis hieraus ist, dass die sichere Verteilung einen nur sehr geringen Einfluss auf die Effizienz der dynamischen Verteilung hat. Als beste Heuristiken bzw. -kombinationen stellen sich das permanente Hintergrund-Repacking (*cb*) und das direkte Repacking in Kombination mit dem Hintergrund-Repacking (*db*) heraus.

Ein theoretisches Potential zur Verbesserung der dynamischen Verteilung durch Optimierung der sicheren Verteilung ist nicht erkennbar, womit eine Verbesserung des aktuellen Energiebedarfs voraussichtlich nicht möglich ist. Sehr wohl ist ein Verbesserungspotential in der Ressourcenreservierung (sichere Verteilung) zu erkennen. Die Minimierung des Bedarfs an Ressourcenreservierung erlaubt es, bei gleichbleibender Serveranzahl mehr VMs zu betreiben. Daher wird im Folgenden kurz abgeschätzt, welchen Nutzen ein optimales Repackingverfahren haben könnte.

In der anschließenden Untersuchung zum Skalierungsverhalten des LPM wird festgestellt, dass durchschnittlich nur 20 % der für die sichere Verteilung benötigten Server aktiv sind. Damit sind rund 80 % der Server inaktiv. Bei den aktiven Servern wurde das direkte Repacking angewendet. Hierzu wurde das Best-Fit-Verfahren eingesetzt, das eine Gütegarantie von 1,7 besitzt. Bei den inaktiven Servern wurde das Hintergrund-Repacking eingesetzt, das durch ein First-Fit-Decreasing-Verfahren umgesetzt wurde. Dieses weist eine asymptotische Gütegarantie von $\frac{11}{9} \approx 1,22$ auf. Ein optimales Verfahren würde maximal eine Verbesserung um den Wert $(1 - \text{Gütegarantie}^{-1})$ bewirken. Grob gerechnet würde somit eine Verbesserung von ca. 23 % ($\approx 0,2 \cdot (1 - \frac{10}{17}) + 0,8 \cdot (1 - \frac{9}{11})$) möglich sein, was einer Verbesserung von ca. 30 % gegenüber keinem Repacking entspricht.

Laufzeit- und Qualitätsbetrachtung

Durch die Möglichkeit der Anpassung des VM-Pools verändert sich die Ausgangslage für das LPM. Dies kann sich sowohl auf die Laufzeit als auch auf die erzielbare Güte, also die Konsolidierungsrate auswirken. Um evtl. Einflüsse festzustellen, wurden unterschiedliche Szenarien mit unterschiedlicher VM-Anzahl simuliert. Hierdurch wurde zugleich auch ein Skalierungstest durchgeführt.

Um die Veränderungen kenntlich zu machen, wurden die Simulationen einmal mit nicht verändertem VM-Pool, einmal mit dynamischen VM-Pool aber ohne Anwendung einer Repacking-Heuristik sowie einmal mit dynamischen VM-Pool und dem Einsatz der Heuristikkombination *db* durchgeführt. Die Ergebnisse sind in Abbildung 4.14 dargestellt.

Die in Abhängigkeit von der VM-Anzahl aufgetragenen Laufzeiten geben relative Laufzeiten bezogen auf die Laufzeit von jeweils 100 VM wieder. Die jeweiligen Regressionskurven legen für alle drei Varianten ein quadratisches Verhalten nahe. Die Erweiterungen haben zu keiner Veränderung der Laufzeitklasse geführt, da sie selbst nur einen linearen Aufwand darstellen. Der quadratische Aufwand liegt im Prozess der Konsolidierung und der Engpassauflösung. In der Theorie ist zwar eine Komplexität mit einer Potenz von 3 analysiert worden, jedoch tritt der entsprechende Part (Betrachtungszeitraum, siehe Abschnitt 4.3.1) aufgrund von Caching

4 Konzeption samt Komponentenevaluation

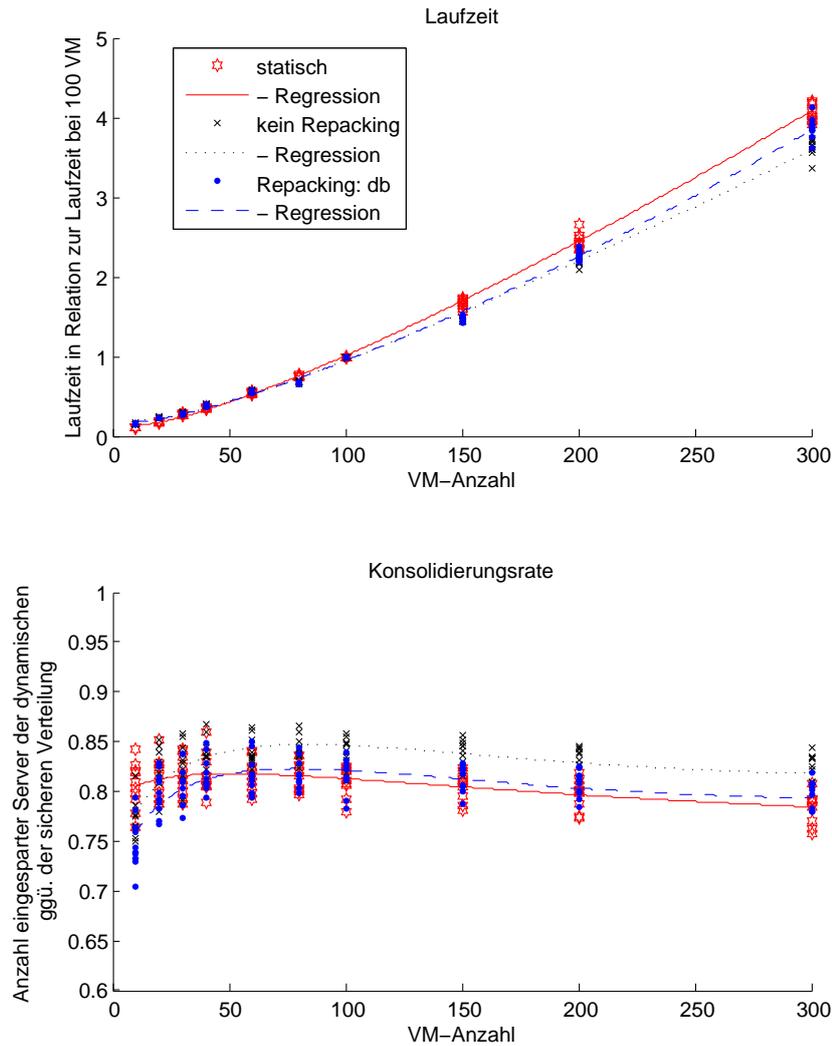


Abbildung 4.14: Skalierungsverhalten des LPM bzgl. der Laufzeit und der Konsolidierungsrate in Abhängigkeit von der VM-Anzahl

relativ selten auf und hat daher in der Praxis nur einen marginalen Einfluss.

Die dargestellte Metrik für die Qualität gibt an, wie hoch die Reduktionsquote (Konsolidierungsrate) an benötigten Servern ggü. der sicheren Verteilung ist. Die Regressionskurven für den dynamischen VM-Pool liegen über der Regressionskurve für den statischen VM-Pool. Dies lässt sich dadurch erklären, dass durch Veränderungen im VM-Pool eine weniger gute sichere Verteilung erreicht wird. Davon ausgehend kann das Verfahren mehr bewirken, als wenn bereits eine bessere sichere Verteilung vorliegt – besser im Sinne, dass weniger Server benötigt werden. Analog lässt sich auch erklären, wieso die Kurve für die dynamische Variante mit Repacking-Heuristik niedriger ist als die ohne Heuristik, denn bei der erstgenannten gibt es durch die verbesserte sichere Verteilung weniger Potential, sie zur Laufzeit noch zu verbessern. Die erzielte Konsolidierungsrate liegt in den untersuchten Szenarien bei etwa 80% und fällt mit steigender VM-Anzahl ab.

4.4 Abstraktion des LPM-Konsolidierungsverhaltens

Das erstellte Energiebedarfsmodell berücksichtigt die drei größten Verbraucher elektrischer Energie. Bei den hier modellierten und charakterisierten Infrastrukturkomponenten ist allerdings zu erkennen, dass ihr Anteil vergleichsweise gering ausfällt. Dies entspricht zwar nicht der allgemeinen Beobachtung, dass der Infrastrukturanteil 50 % - 100 % des Energiebedarfs der IT hat, allerdings bezieht sich diese Angabe auf ältere Rechenzentrumsarchitekturen, bei denen eine Energiebedarfsoptimierung allenfalls eine mindere Rolle gespielt hat. Die hier modellierten Komponenten sind für einen energieeffizienten Betrieb ausgelegt. Insgesamt ergibt sich damit ein verhältnismäßig großer Energiebedarfsanteil für die IT selbst, was für eine hohe Infrastruktureffizienz spricht. Bei einer IT-Last von 30 kW und einer Außentemperatur von 20 °C ergibt sich eine PUE von ca. 1,3.

Die Maßnahmen mit dem größten Potential zur Reduzierung des Energiebedarfs beziehen sich unter diesen Umständen auf die Optimierung des Energiebedarfs der für die IT eingesetzten Geräte, also Server und Chassis. Ein Server allein wird bereits durch unterschiedlichste Techniken energieeffizient und lastadaptiv betrieben. Der Energiebedarfsoffset für den alleinigen Betrieb eines Servers im Idle-Modus bleibt jedoch bestehen. Dieser lässt sich, wie oben beschrieben, nur durch seine Abschaltung einsparen. Das LPM bewirkt durch die dynamische Konsolidierung ein bedarfsgerechtes Abschalten und rechtzeitiges Hochfahren, wobei in der in Abschnitt 4.3.2 durchgeführten Untersuchung zu heterogenen Servern festgestellt wurde, dass der Energiebedarf maßgeblich durch die LPM-Strategie beeinflusst wird. Folglich muss für das LPM als weiterer Einflussfaktor ebenfalls ein Modell erstellt werden, das mit den aktuellen Policies berücksichtigend charakterisiert wird.

Da ein Verfahren zur Umsetzung eines LPM sich in seinem Verhalten aufgrund unterschiedlicher Funktionsweisen und Strategien sehr unterschiedlich äußern kann, muss grundsätzlich für jedes LPM-Verfahren ein Modell erstellt werden. In dieser Dissertation wird dies beispielhaft am Verfahren, das in den Abschnitten 3.2 und 4.3 vorgestellt wurde, durchgeführt. Für die hier betrachtete Rechenzentrumsgröße (Container) ist eine homogene Serverbestückung sehr wahrscheinlich. Daher werden bei der nun folgenden Modellierung zur Vereinfachung und ohne Beschränkung der Allgemeinheit nur homogene Server betrachtet. Außerdem wird nur von Servern als aktive IT-Komponenten ausgegangen. Chassis bzw. Blade-Enclosures werden also nicht berücksichtigt. Der IT-Energiebedarf hängt damit größten Teils von der Konsolidierungsrate des LPM ab.

Das genaue Konsolidierungsverhalten eines LPM lässt sich zwar durch Simulationen ermitteln, aber nicht zuletzt die Untersuchung in Abschnitt 4.3.4 hat gezeigt, dass das betrachtete LPM ein polynomielles Laufzeitverhalten bzgl. der VM-Anzahl aufweist. Während für einzelne Abfragen die Laufzeit einer Simulation noch praktikabel ist, führt dies bei verteilten Lastoptimierungen zu immensen Verzögerungen, da der Energiebedarf bzw. das Konsolidierungsverhalten für jede Konstellation als eine elementare Information abgerufen wird. Um eine beschleunigte Abfrage zum LPM-Verhalten zu erhalten, ist also eine Abstrahierung bzw.

Modellierung dessen notwendig.

Das zu erstellende Modell hat damit die Funktion anzugeben, wie viele Server für den Betrieb einer bestimmten Menge an zu betreibenden VMs zur Laufzeit benötigt werden, und damit eine Abschätzung über den hierdurch entstehenden Energiebedarf zu ermöglichen. Das Modell stellt insoweit auch eine Vorhersagemethode dar. Eine Referenz in der Modellierung durch Zeitreihenanalysen stellt das Buch *Time Series Analysis: Forecasting and Control* [34] dar. Hierin wird u.a. eine Übersicht zur Erstellung von stochastischen Modellen durch (Auto-)Korrelations- und Regressionszusammenhänge gegeben und auf den Umgang mit aufkommenden Fehlern eingegangen.

Im näheren Kontext der Dissertation, und zwar beim Grid Computing, werden nach der Arbeit von Dobber, van der Mei und Koole [74] für gewöhnlich lineare Modelle aufgrund ihrer Einfachheit und schnellen Ausführung eingesetzt. Es wird herausgestellt, dass sich Autoregression besonders bei periodischen Verläufen eignet. Die für das LPM bereitgestellten Auslastungsverläufe der VMs weisen zwar einen periodischen Verlauf auf, allerdings liegen diese Perioden im Bereich mehrerer Tage bis Wochen. Das Regressionsmodell, das diese Perioden abdeckt, wäre somit sehr komplex. In der Arbeit von Vazhkudai und Schopf [228] wird der Netzwerktraffic betrachtet und der Mehrwert von multivariaten Daten herausgestellt. Neben den bloßen Daten zur Zeitreihe des abzuschätzenden Wertes sollten auch solche berücksichtigt werden, die das System aktuell näher spezifizieren. Das dort empfohlene lineare Modell wird durch multiple Regression erstellt.

Bei der Modellierung der Hardwarekomponenten (in Abschnitt 4.2) wurden auch bereits (multiple) Regressionsmodelle genutzt, um das Systemverhalten zu modellieren. Die erzielten Ergebnisse sind sehr gut. Die Modellierung mittels multipler Regression erscheint damit als ein probates Mittel, um eine gute Prognosequalität bei zugleich schneller Berechnung zu realisieren. Dies ist jedoch an die Bedingung geknüpft, dass sich die gewählten Variablen ebenfalls auf lineare Weise ermitteln lassen.

Alternativ könnten zur Modellierung des Konsolidierungsverhaltens auch neuronale Netze verwendet werden, deren typische Anwendung auch im Bereich der Trend- und Verhaltensvorhersagen liegt. Auch bei ihnen ist eine hohe Verarbeitungsgeschwindigkeit gegeben. Sie werden häufig verwendet, wenn das Problem entweder nicht vollständig verstanden ist oder das zu modellierende System so komplex ist, dass die Bildung eines mathematischen Zusammenhangs nur sehr schwierig ist. Hierzu müssen sie allerdings geeignet trainiert werden. Dies erscheint für die vorliegende Problemstellung, insbesondere in Anbetracht verwandter Arbeiten, als übertrieben, weshalb im Folgenden auf den simpleren Ansatz mittels Regressionsmodellen eingegangen wird.

4.4.1 Konzeptionierung eines linearen Regressionsmodells

Die Funktion des Regressionsmodells besteht darin, eine Abschätzung darüber zu geben, wie viele Server durch das LPM eingeplant werden, um den durch die VMs generierten Rechenleistungsbedarf abzudecken. Da – in einem ersten Schritt – von homogenen Servern

4.4 Abstraktion des LPM-Konsolidierungsverhaltens

ausgegangen wird, spielt es keine Rolle, wie stark tatsächlich jeder einzelne Server ausgelastet wird, denn die Auswirkungen auf den Energiebedarf sind in Summe gleich. Für einen Server müssen damit nur zwei Energiezustände bestimmt werden: an und aus.

Als abhängige Variable bietet sich somit die Anzahl der aktiven Server ($\#SRV_a$) an. Die Leistungsaufnahme des Servers ergibt sich dann durch Multiplikation mit der Leistungsaufnahme im Idle zuzüglich der auslastungsabhängigen Leistungsaufnahme, welche sich aufgrund der homogenen Struktur aggregiert über alle n VMs betrachten lässt. Ohne Beschränkung der Allgemeinheit soll zur Vereinfachung auch erst einmal nur der Ressourcentyp CPU betrachtet werden, da dieser den größten Einfluss auf die Leistungsaufnahme des Servers hat. Mit der abgeschätzten Anzahl aktiver Server kann somit anhand folgender Formel die Leistungsaufnahme der IT (Serveranteil) ermittelt werden:

$$power_{IT} = \#SRV_a \cdot power_{SRV,idle} + (power_{SRV,max} - power_{SRV,idle}) \cdot \sum_{i=1}^n load_{VM_i,CPU} \quad (4.8)$$

Die Anzahl aktuell benötigter Server in Relation zu einer ursprünglichen Anzahl, sei es die Anzahl benötigter Server in der sicheren Verteilung oder im dedizierten Betrieb, ist ein Maß für die Konsolidierungsrate bzw. -effizienz. In dem Paper von Pelley [175] wird die Auswirkung von Serverkonsolidierung mittels der folgenden Formel ausgedrückt:

$$u_{SRV} = \frac{U}{U + (1 - U) \cdot l}, \quad U, l \in [0, 1] \quad (4.9)$$

Berechnet wird die Auslastung, eines individuellen Servers. Mit der Variablen U wird die Gesamtauslastung des Rechenzentrums angegeben. Der Parameter l ist ein Maß dafür, wie stark die Lasten durch das LPM auf die Server verteilt sind: Gilt $l = 0$, so wird eine optimale Konsolidierung angenommen, die aktiven Server sind voll ausgelastet. Gilt $l = 1$, so werden die Lasten gleichmäßig über alle Server verteilt. Der Nenner gibt somit die auf die maximale Serveranzahl normierte Anzahl aktiver Server an. Für die übrigen Server wird davon ausgegangen, dass auf ihnen keine VMs laufen und sie abgeschaltet sind.

Zur Betrachtung der Anzahl aktiver Server $\widetilde{\#SRV}_a$ (normiert auf die maximal benötigte Serveranzahl) wird der Nenner in den folgenden Term umgeformt:

$$\widetilde{\#SRV}_a = (1 - l) \cdot U + l, \quad U, l \in [0, 1]$$

Hiermit lässt sich leichter erkennen, dass der Parameter l auch als ein Maß für einen möglichen Overhead zur Organisation des VM-Pools angesehen werden kann. Mit dem Steigungsfaktor $(1 - l)$ wird zwischen diesem initialen Overhead (l) und dem Maximum (1) mit zunehmender Auslastung linear skaliert. In Abbildung 4.15 ist dieser Zusammenhang visuell verdeutlicht.

Mittels einer Parameterregression kann sowohl der Overhead als auch die lineare Abhängigkeit von der Gesamtauslastung ermittelt werden. Das hieraus resultierende Regressionsmodell

4 Konzeption samt Komponentenevaluation

hat die folgende Form:

$$\#SRV_a = \alpha_0 + \alpha_1 \cdot U$$

Die Angabe der Serveranzahl ist hier nun absolut gegeben, da die entsprechende Skalierung durch den Regressionsprozess durchgeführt wird: α_i sind die Regressionsparameter, die durch den Regressionsprozess automatisch ermittelt werden. Zu klären wäre nun, woraus sich die Gesamtauslastung U bildet.

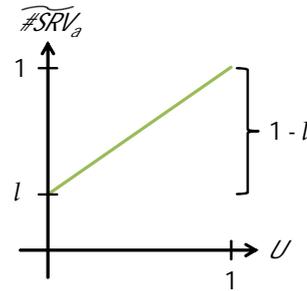


Abbildung 4.15: Serverkonsolidierungsrate nach Pelley [175]

Die VMs werden in der sicheren Verteilung entsprechend ihrer statistisch maximalen Auslastung alloziert. Die Anzahl an Servern in

der sicheren Verteilung $\#SRV_s$ entspricht somit der maximal benötigten Anzahl. Wenn die Summe der CPU-Lasten aller VMs $SoL (= \sum_{i=1}^n load_{VM_i,CPU})$ gleich der Summe der maximalen CPU-Lasten aller VMs $SoML (= \sum_{i=1}^n load_{VM_i,CPU,max})$ ist, wird die maximale Anzahl an Servern benötigt:

$$SoL = SoML \rightarrow \#SRV_a = \#SRV_s$$

$$SoL = 0 \rightarrow \#SRV_a = 0$$

Die Anzahl benötigter Server ist minimal, wenn keine VM CPU-Leistung benötigt. Zwischen diesen Extremen skaliert die Anzahl in Abhängigkeit von SoL . Damit lässt sich U durch SoL substituieren:

$$\#SRV_a = \alpha_0 + \alpha_1 \cdot SoL \quad (4.10)$$

Eine Abstufung in ganzen Servern ist hier nicht gegeben, aber an dieser Stelle auch nicht erforderlich, da eine nachgelagerte Rundungsoperation dieses übernehmen kann.

Das hier verwendete LPM verwendet zum Packen einen Bin-Packing-Ansatz. Grundsätzlich hängt die erreichbare Güte eines Bin-Packing-Verfahrens nicht nur von der Reihenfolge der einzusortierenden Elemente ab, sondern auch davon, wie viele Elemente einzusortieren sind und insbesondere welche konkreten Größen vorliegen. Außerdem handelt es sich beim LPM um einen dynamischen Prozess, in dem auch zukünftige Werte berücksichtigt werden müssen. Um diese Aspekte einzubinden, sollten noch weitere Variablen berücksichtigt werden.

Hierbei sollten nur Werte genommen werden, die aus dem gesamten VM-Pool aggregiert werden können. Die Summe der einzelnen CPU-Lasten ist solch ein aggregierter Wert. Es macht aber z. B. keinen Sinn, die einzelnen CPU-Lasten der VMs in die Variablenliste aufzunehmen, da die Variablenanzahl dann von der VM-Menge abhängig wäre. Das Regressionsmodell könnte somit nicht auf einen veränderten VM-Pool angewendet werden. Die folgenden aggregierten Werte, die ihrerseits mit höchstens linearem Aufwand ermittelt werden können, sollen auf ihre Tauglichkeit als Schätzer-Variable untersucht werden:

- $[SoL]$ die (summierte) Gesamtauslastung aller VMs,
- $[SoML]$ die (summierte) maximale Gesamtauslastung aller VMs,
- $[\#VM]$ die Anzahl an VMs im VM-Pool,
- $[\#VM_{LC}_i]$ die Anzahl an VMs, die mit ihrer aktuellen Auslastung einer bestimmten Auslastungsklasse angehören (zehn Klassen mit 10 %-Schrittweite) – dies gibt die Unterschiedlichkeit der Lasten wieder.

Zusätzlich werden die Potenzen und Vorhersagewerte dieser Variablen betrachtet. Historische Werte werden nicht herangezogen, da das LPM seine Entscheidungen aufgrund der zukünftigen Werte fällt. Ein autoregressiver Ansatz wird nicht weiter verfolgt, da der eintretende Fehler bei Autoregressionsmodellen mit zunehmender Prognoseweite laut Dobber [74] signifikant ansteigt.

Neben der Modellierung des LPM, das für den dynamischen Part zuständig ist und damit den Energiebedarf beeinflusst, soll auch modelliert werden, wie viele Server in der sicheren Verteilung alloziert werden, da dies die zu reservierenden Ressourcen festlegt. Beim Planen der Migrationen muss schließlich nicht nur der Energiebedarf angepasst, sondern auch sichergestellt werden, dass vor Ort genügend Ressourcen vorhanden sind. Für dieses Regressionsmodell kommen die Variablen in Frage, die sich über die Zeit nicht ändern, also $SoML$ und $\#VM$.

4.4.2 Modellermittlung und -analyse

In diesem Abschnitt werden auf Simulationen basierend als erstes die vorgenannten Variablen bzgl. ihrer Einflussgröße untersucht und darauf stützend das zu verwendende Regressionsmodell festgelegt. In den weiteren Untersuchungen werden die Randbedingungen herausgestellt, die berücksichtigt werden müssen, um das Modell zu Prognosezwecken nutzen zu können. Zur Regression wird die Methode nach dem geringsten quadratischen Fehler verwendet.

Für die durchgeführten Simulationen lag ein Pool von insgesamt 10000 unterschiedlichen VMs vor (siehe Abschnitt 5.1), die auf realen VM-Auslastungsdaten basieren. Es wurden insgesamt 100 unterschiedliche Szenarien gebildet, in denen 10 Tage simuliert wurden. Die Szenarien unterscheiden sich in der Menge der betriebenen VMs (10 bis 200) und der konkreten Auswahl der VMs, welche pseudo-randomisiert durch Angabe eines sogenannten *Seed* aus dem Gesamtpool selektiert wurden. Der verwendete Zufallsgenerator erzeugt bei Angabe des gleichen *Seed* die gleiche Pseudo-Zufallszahlenfolge und damit die gleiche VM-Auswahl, was hier zu Vergleichbarkeitszwecken eingesetzt wird. Als Fehlermetriken wurden die maximale (relative), die durchschnittliche (relative) und die Standard-Abweichung zwischen den Modelldaten und den durch die Simulation ermittelten Daten ausgewertet.

Variablenwahl

Der Einfluss der Variablen wird durch den Vergleich unterschiedlicher Regressionsmodelle ermittelt. Es werden Modelle erstellt, welche neben dem Konstantanteil nur aus der entsprechenden Variablen bestehen und solche, die Variablen kombinieren. Das jeweils ausgewertete Regressionsmodell wurde anhand der gesamten Messdaten erstellt. Das Ergebnis der Untersuchung ist in Abbildung 4.16 dargestellt. Neben den einzelnen Variablen, deren Prognosewerte sich auf den jeweils abgefragten Zeitpunkt beziehen, ist auch der Einfluss der weiter in die Zukunft reichenden Prognosewerte dargestellt. Es wurde sowohl die Anzahl der zusätzlichen Werte als auch der zeitliche Abstand zwischen ihnen (Schrittweite) variiert. Als Gütemetrik ist hier die durchschnittliche relative Abweichung abgetragen.

Bereits die alleinige Verwendung der Auslastungswerte oder der VM-Anzahl liefert brauchbare Werte (durchschnittliche Abweichung unter 10 %). Durch die Kombination der Variablen können bessere Ergebnisse erzielt werden, bis zu einer durchschnittlichen Abweichung von 7 %. Für die Variablen $\#VM_LC_i$ ist die Variable $\#VM$ redundant, da die Gesamtanzahl auf die einzelnen Klassenwerte aufgeteilt ist. Es ist anzumerken, dass die hier dargestellten Abweichungen punktueller Natur sind, d.h., dass die Abweichungen pro Zeitpunkt gelten.

Üblicherweise wird der Energiebedarf für einen Zeitraum betrachtet, so dass der Durchschnittswert über mehrere Zeitpunkte gebildet wird. Da der Durchschnittsfehler für die in der Regression verwendeten Werte einen Erwartungswert von 0 besitzt, ist diese Metrik an dieser Stelle jedoch nicht aussagekräftig.

Eine Erhöhung der Anzahl zusätzlicher Prognosewerte sowie auch der Schrittweite führt jeweils zu einem besseren Ergebnis. Je mehr Werte berücksichtigt werden, desto komplexer wird jedoch die Modellierung und der nachträgliche Berechnungsaufwand. Bei der Erhöhung der Schrittweite steigt der Berechnungsaufwand nicht und kann daher frei gewählt werden, soweit entsprechend in die Zukunft reichende Prognosewerte vorliegen. Wird das beste Modell betrachtet, nämlich das, welches alle vorgestellten Variablen berücksichtigt, ergibt sich bei einer Schrittweite von fünf und ebenfalls fünf prognostizierten Werten eine Sättigung. Weitere Verbesserungen sind dann nur noch marginal.

Als drittes wurde untersucht, welchen Mehrwert Potenzen der Variablen bringen. Bis zu einer Potenz dritten Grades können Verbesserungen bis zu 2 % erzielt werden, jedoch hat

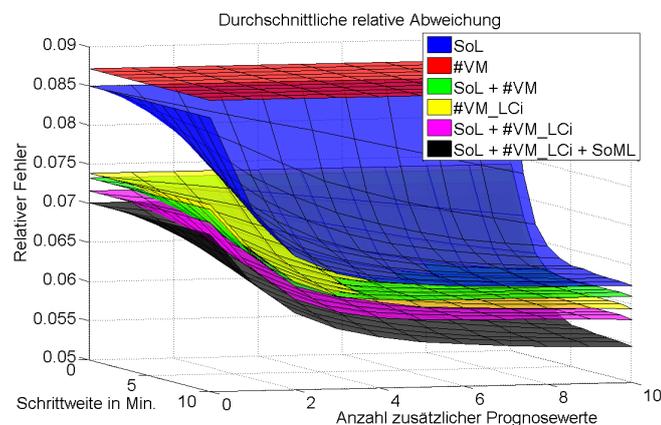


Abbildung 4.16: Auswirkungen der unterschiedlichen Variablen sowie ihrer prognostizierten Werte auf die Güte des Regressionsmodells

sich in nachfolgenden Evaluationen gezeigt, dass dies zu einem Overfitting des Modells führt, weshalb diese nicht weiter betrachtet werden.

Entsprechend den Beobachtungen werden fortan die beiden folgenden Regressionsmodelle verwendet:

$$\#SRV_a(t) = \alpha_0 + \sum_s \left(\begin{array}{l} \alpha_{1,s} \cdot SoL(t+s) + \\ \alpha_{2,s} \cdot SoML(t+s) + \\ \sum_{i=0}^9 (\alpha_{(3+i),s} \cdot \#VM_LC_i(t+s)) \end{array} \right)$$

$$s \in \{0\} \tag{4.11}$$

$$s \in \{0, 5, 10, 15, 20, 25\} \tag{4.12}$$

Im Modell (4.11) werden lediglich die Werte zum jeweiligen Zeitpunkt verwendet. Beim Modell (4.12) kommen noch jeweils fünf Vorhersagewerte bis zu einer Vorhersagelänge von 25 Min. hinzu.

Generalisierbarkeit

Diese Analyse soll Aufschluss darüber geben, wie gut sich die charakterisierten Regressionsmodelle auf andere VM-Mengen übertragen lassen. Dazu werden für jeden Simulationsdurchgang vier unterschiedliche Regressionsmodelle herangezogen, die durch die folgenden unterschiedlichen Datensätze charakterisiert werden:

- die Daten des jeweiligen Simulationsdurchgangs,
- die Daten der Simulationsdurchgänge mit gleicher VM-Anzahl,
- die Daten der Simulationsdurchgänge mit gleichem Seed,
- die Daten aller übrigen Simulationsdurchgänge.

Nur im ersten Fall werden die Daten des zu untersuchenden Durchgangs in der Modellcharakterisierung verwendet. Das korrespondierende Regressionsmodell dient damit als Referenz. In den übrigen drei Fällen fehlen diese Daten. Hier muss also eine Generalisierung basierend auf den übrigen Daten erfolgen. Wie gut dieses in den untersuchten Szenarien funktionierte, ist in Abbildung 4.17 dargestellt.

Das beste Ergebnis liefert die Referenz, das schlechteste das Regressionsmodell für alle Messreihen. Dazwischen liegen die Regressionsmodelle für die Messreihen mit gleicher VM-Anzahl bzw. gleichem Seed. Bei diesen beiden ist herauszustellen, dass das Regressionsmodell, welches sich auf die Messreihen mit gleichem Seed bezieht, bessere Ergebnisse liefert, wenngleich hier die Streuung größer ist. Durch die Hinzunahme von weiteren Prognosewerten in die Regression (komplexes Modell (4.12)) verbessert sich das Ergebnis insgesamt, allerdings bleiben die Verhältnisse zueinander bestehen.

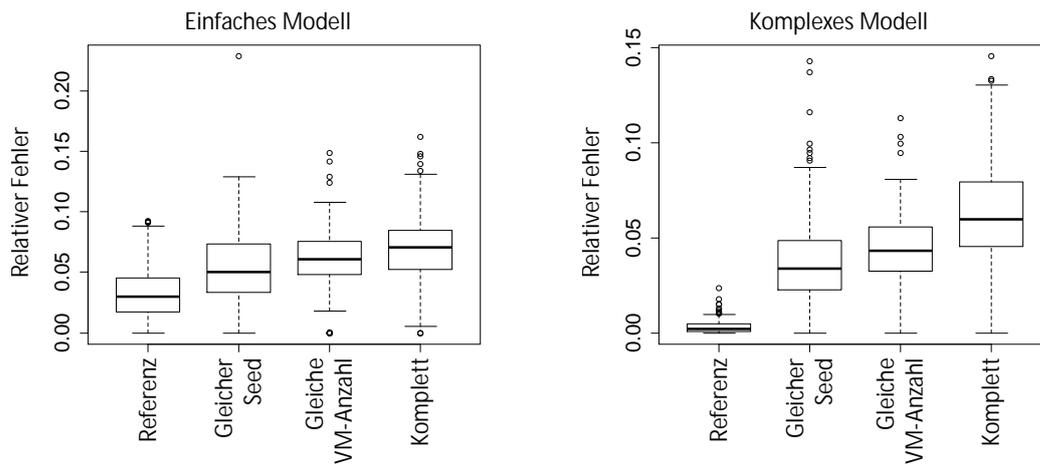


Abbildung 4.17: Güte der Regressionsmodelle bei unterschiedlichen Messreihen/-gruppen

Es zeigt sich, dass die lineare Modellierung des Allokationsverhaltens von den konkreten VMs abhängig ist. Um möglichst gute Prognosen zu erhalten, sollte daher kein allgemeines Modell herangezogen werden, sondern stets ein zur Zeit aktuelles. Das Modell sollte demnach vor seiner Konsultation aktualisiert werden, indem es neu charakterisiert wird. Je nach Komplexität des verwendeten Modells – in dem Vorliegenden ist es relativ simpel – kann dieser Prozess zeitaufwendig sein. Eine Möglichkeit, das Modell durch einen iterativen Prozess zu aktualisieren, ergibt sich durch die „Boosting“-Technik. Das Grundprinzip dieser Technik besteht nach Cao und weiteren [48] darin, dass die zuvor ermittelten Parametergewichtungen anhand von Fehlerbeobachtungen bei Anwendung auf einem aktuelleren Untersuchungsdatensatz modifiziert werden. Für tiefer gehende Betrachtungen zum generellen Online-Lernen mit Boosting sei auch die Arbeit von Freund und Schapire [87] nahegelegt.

Trainingslänge

Die bisherigen Untersuchungen haben die Daten des gesamten Messzeitraums zur Modellierung verwendet. Für eine Untersuchung der Prognosegüte soll nun nur ein begrenzter Zeitraum herangezogen werden. Es soll festgestellt werden, wie lang der Trainingsdatensatz sein muss, um eine bestimmte Prognoseweite mit adäquater Qualität zu erhalten. Hierzu werden unterschiedliche Trainingslängen und Prognoseweiten entsprechend der Abbildung 4.18 analysiert.

Beste Ergebnisse werden ab einem Trainingsintervall von 24 Stunden (1440 Min.) erzielt. Die durchschnittliche absolute Abweichung liegt hier bei unter 0,7 Servern. Soweit der Vorhersagezeitraum nur wenige Stunden beträgt, reichen auch geringe Trainingslängen aus, um gute Ergebnisse zu erhalten (blauer Bereich). Anhand der Standardabweichung lässt sich zusätzlich erkennen, dass mit Zunahme des Vorhersagezeitraumes die Streuung der Abweichungen in den ersten Stunden stark zunimmt, danach allerdings stagniert.

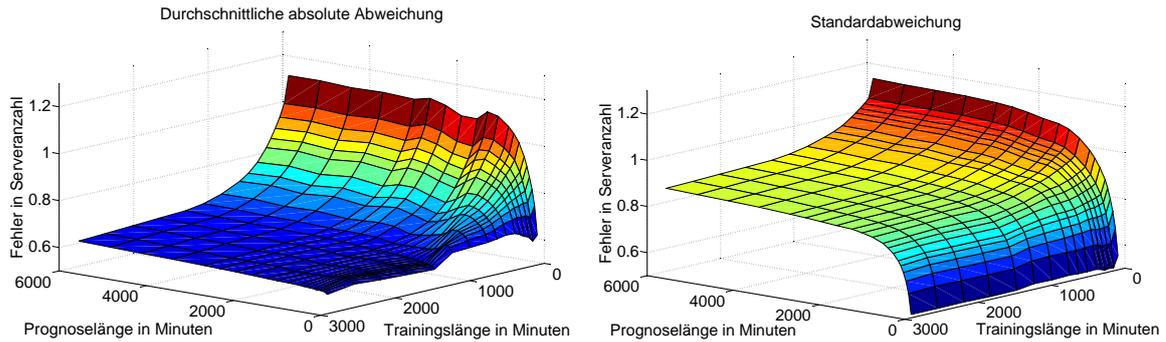


Abbildung 4.18: Prognosegüte des Modells (4.11) für unterschiedliche Vorhersagezeiträume in Abhängigkeit von der Trainingslänge

In Abbildung 4.19 wird zum Modell (4.11) zusätzlich die Prognosegüte des Modells (4.12) dargestellt. Hierin ist zu erkennen, dass das Regressionsmodell (4.12) bei Trainingslängen unter einem Tag schlechtere Ergebnisse liefert, im Bereich einiger Stunden sogar deutlich schlechtere. Bei diesen kurzen Zeiträumen führen die zusätzlichen Variablen zu einem Overfitting. Erst wenn die Trainingslänge mindestens einen Tag beträgt, ist die Vorhersagegüte besser.

Bei beiden Modellen ist somit festzuhalten, dass sie mit Daten von mindestens 24 Stunden charakterisiert werden sollten.

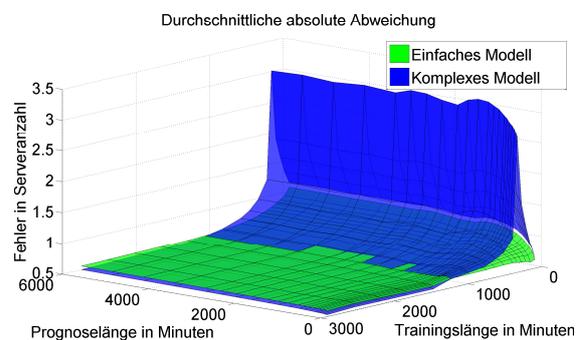


Abbildung 4.19: Prognosegüte in Abhängigkeit von der Trainingslänge: Gegenüberstellung der Regressionsmodelle (4.11) und (4.12)

Veränderungsauswirkung

Als nächstes werden die Auswirkungen von Veränderungen (das Hinzufügen bzw. Entfernen von VMs) auf die Prognosegüte untersucht. Entsprechend obiger Feststellung wird zuerst über einen Zeitraum von 24 Stunden simuliert. Die hierbei ermittelten Daten werden zur Charakterisierung der Regressionsmodelle verwendet. Im Anschluss wird die Simulation für weitere 48 Stunden fortgeführt.

Zur Prognose wurden Modelle herangezogen, die mit den Simulationsdaten unterschiedlicher VM-Anzahlen charakterisiert wurden, wobei die Auswahl der VMs über den gleichen Seed erfolgt ist wie für die VM-Mengen, für die die Prognosen abgegeben werden sollen. Hierdurch ist also ein Teil der VM-Mengen (aus der Modellierung und für die Prognosen) gleich. Die Ergebnisse hierzu sind in Abbildung 4.20 dargestellt. Bei der Fehlermetrik zur absoluten Abweichung lässt sich erkennen, dass das für die jeweilige VM-Anzahl entsprechende Modell die besten Ergebnisse liefert. Je größer die Veränderung ist, desto schlechter wird das Ergebnis.

4 Konzeption samt Komponentenevaluation

Die gleiche Beobachtung lässt sich auch bei der Fehlermetrik zur relativen Abweichung machen. Ausreißer stellen hier die Prognosen für eine sehr geringe VM-Anzahl dar, denn die Anzahl allozierter Server ist hier mit ziemlich konstant 1 leicht abschätzbar.

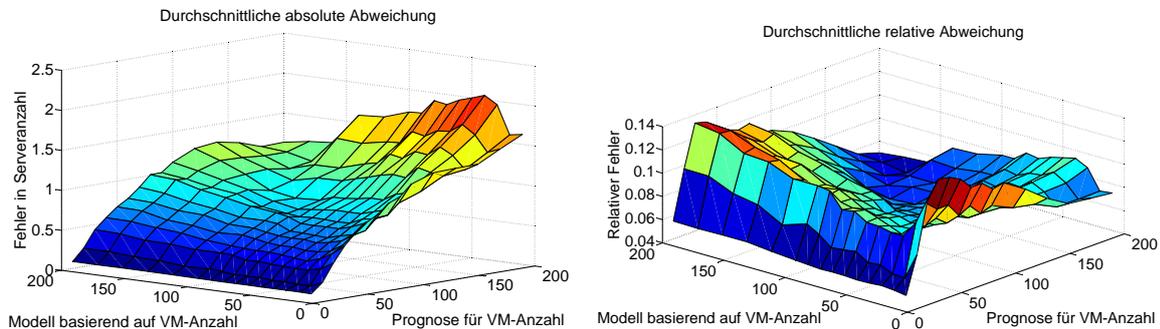


Abbildung 4.20: Fehlerentwicklung bei Veränderung des VM-Pools – Ausgangslage: unterschiedliche VM-Anzahlen

Da die Modellierungen nur anhand von Daten mit jeweils konstanter VM-Anzahl charakterisiert wurden, ist keine Abhängigkeit von der VM-Anzahl gegeben. Ohne die konkrete Anzahl und Auswahl an VMs nach Durchführung einer Veränderung zu kennen, lässt sich jedoch ein allgemeines Modell generieren, dass die Abhängigkeit von der VM-Anzahl abbildet. Hierzu können Simulationen mit beliebigen anderen VMs durchgeführt und daraus das allgemeine Modell abgeleitet werden. Da dieses nicht auf die aktuelle Situation spezialisiert ist, sollte das aktuelle Modell auch betrachtet werden. Über eine Gewichtung kann der Einfluss der jeweiligen Modelle eingestellt werden. In Abbildung 4.21 sind die durchschnittlichen Abweichungen für unterschiedliche Gewichtungen und Veränderungsgrade angegeben. Die Ausgangslage ist hier bei einer VM-Anzahl von 100.

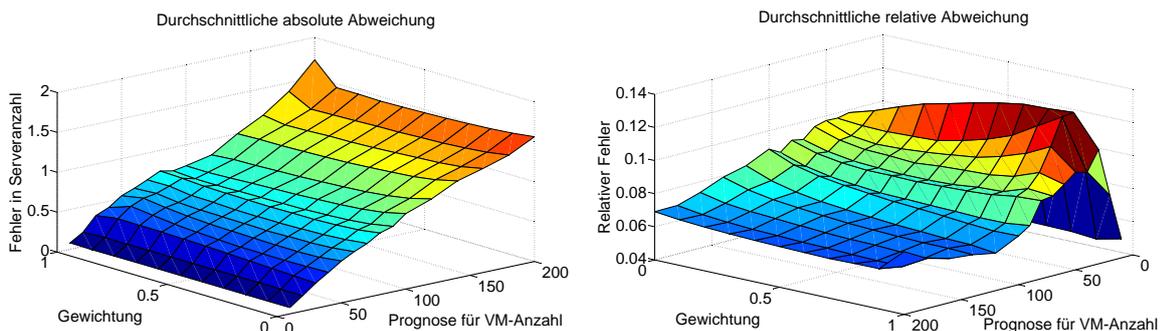


Abbildung 4.21: Fehlerentwicklung bei Veränderung des VM-Pools – Ausgangslage: 100 VM, unterschiedlich gewichtete Berücksichtigung eines allgemeinen Modells

Vergleicht man die Ergebnisse für die beiden Extremen „nur das aktuelle Modell“ (Gewichtung = 1) und „nur das allgemeine Modell“ (Gewichtung = 0), so lässt sich ein flacherer Verlauf bei dem allgemeinen Modell erkennen – dies spricht für die bessere Generalisierung. Allerdings ist der Unterschied gering. Die Abhängigkeit von der Gesamt-VM-Anzahl fällt

damit nicht so stark ins Gewicht, da sie grundsätzlich als Information in den Auslastungsklassen vorliegt. Für eine gute Generalisierung müssen jedoch genügend Informationen zu den einzelnen Klassen vorliegen. Ist dies gegeben, kann basierend auf dem aktuellen Modell eine Prognose für veränderte VM-Pools gegeben werden. Bei größeren Veränderungen sollte jedoch ein allgemeingültiges Modell – zusätzlich – herangezogen werden, um die Prognosegüte zu verbessern. Die Ergebnisse zum Modell (4.12) sind ähnlich, wobei eine stärkere Spezialisierung (größerer Fehler bei Veränderungen) zu erkennen ist.

Bei einem Blick auf die Ergebnisse der Untersuchung zur Generalisierbarkeit fällt auf, dass umso bessere Ergebnisse erzielt werden, je mehr VMs der aktuellen VM-Menge im Modellierungsprozess herangezogen werden. Diese Aussage trifft allerdings nur für die zur Modellierung verwendeten Daten zu. Wenn also Prognosen über VM-Mengen erstellt werden sollen, in denen auch andere VMs beinhaltet sind, kann eine starke Spezialisierung sich negativ auswirken. Für diese unbekannt VMs sind die Informationen des allgemeinen Modells von Vorteil.

Modell für die sichere Verteilung

Zur Berechnung der zu reservierenden Kapazität muss die Serveranzahl in der sicheren Verteilung ermittelt werden. Aus den oben genannten Variablen kommen nur *SoML* und *#VM* in Frage. Das hierzu gehörende Regressionsmodell ist in Gleichung 4.13 angegeben. Mit den Daten aller Messreihen charakterisiert liefert es eine Modellgüte von durchschnittlich 97%. Im Gegensatz zur Modellierung der dynamischen Allokation ist hier das Potential zur Verbesserung des Ergebnisses durch ein an die aktuelle Situation spezialisiertes Modell sehr gering. Es reicht für diesen Zweck somit ein allgemeines Modell.

$$\#SRV_s(t) = \beta_0 + \beta_1 \cdot SoML(t) + \beta_2 \cdot \#VM(t) \quad (4.13)$$

In Abbildung 4.22 ist eine detailliertere Angabe des Ergebnisses dargestellt. Die durchschnittliche Güte ist zwar sehr gut, allerdings existieren viele Ausreißer die jenseits der 15% Abweichung liegen. Eine genauere Analyse hat ergeben, dass es sich hierbei um die Abschätzungen in den kleineren Regionen der Serveranzahl handelt. In Abbildung 4.23 sind die Ergebnisse für jeweils eine Intervallgröße von 10 Servern zusammengefasst. Bei der absoluten Abweichung ist ein mit der Intervallgruppe ansteigender Fehler zu erkennen. Bei der relativen Abweichung wird der Fehler geringfügig kleiner. Die Ergebnisse für das Intervall (0,10] für die Abschätzung der Serveranzahl stechen gegenüber dem Trend negativ heraus. Die großen relativen Abweichungen liegen allesamt in diesem Intervall. Dies hat allerdings keine Tragweite, denn das Modell wird zur Kapazitätsabschätzung genutzt, bei der die Grenzen eines Rechenzentrums relevant sind. Diese liegen in den höheren Bereichen, bei denen das Modell zuverlässig gute Ergebnisse liefert.

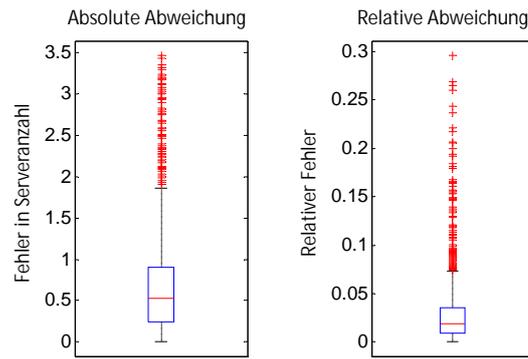


Abbildung 4.22: Güte des Modells (4.13) zur Abschätzung der benötigten Anzahl an Servern in der sicheren Verteilung

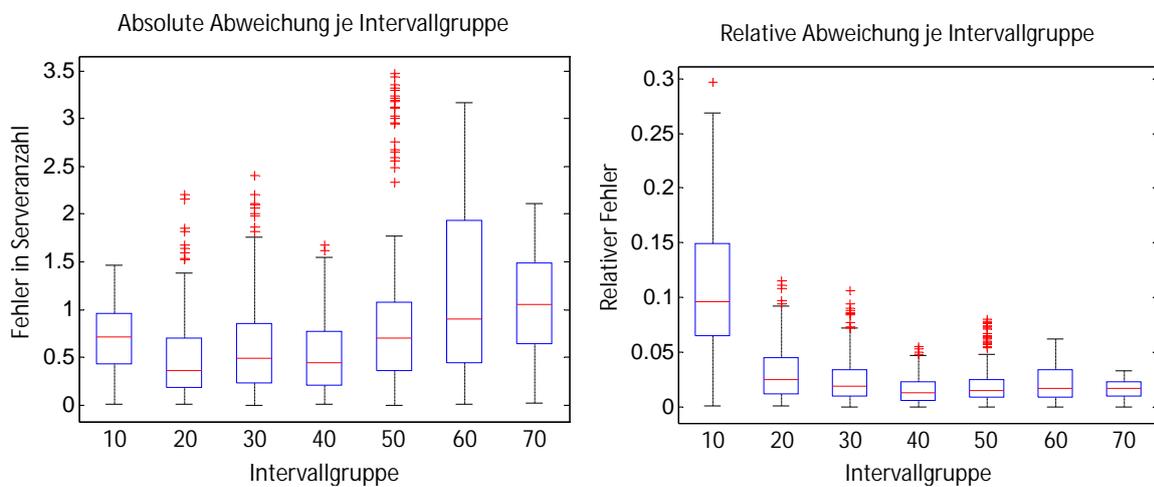


Abbildung 4.23: Güte des Modells (4.13) bei unterschiedlichen Intervallgruppen für die Serveranzahl. Eine Intervallgruppe hat eine Betrachtungsbreite von 10 Servern und als rechte einschließliche Grenze die im Diagramm angegebene Serveranzahl.

4.4.3 Evaluation der LPM-Abstraktion

In dieser Evaluation wird die Prognosequalität des Regressionsmodells ermittelt, wenn es in einem simulierten Betrieb an den Veränderungszeitpunkten aktualisiert, also mit den Daten bis zu diesem Zeitpunkt charakterisiert und zur Prognose bis zum nächsten Veränderungszeitpunkt genutzt wird. Es wurden wie in Abschnitt 4.3.4 homogene Server und Datensätze realer VMs verwendet, worauf in Abschnitt 5.1 eingegangen wird, und insgesamt 100 unterschiedliche Simulationsläufe durchgeführt. Jeder Lauf simuliert 10 Tage und beinhaltet einen VM-Pool von 300 VM (initial 150 VM, maximal 300 VM). Um eine Dynamik im VM-Bestand zu simulieren, wurden Operationsverläufe definiert, die alle 4 bis 8 Stunden VMs hinzufügen oder entfernen, wobei die Veränderungsrate bei bis zu 50% der aktuellen VM-Anzahl liegt. Die Simulationsläufe unterscheiden sich in der Auswahl des VM-Pools und in der Anwendung der Operationsverläufe (jeweils zehn unterschiedliche).

Die Allokationsabschätzung bezieht sich immer auf das Intervall bis zur nächsten Veränderung, was also einer 4 bis 8 Stunden Prognose entspricht. Diese wird dann mit den durch die Simulation ermittelten Werten verglichen. In Tabelle 4.2 sind die Ergebnisse für die beiden Modelle (4.11) und (4.12) bzgl. des durchschnittlichen absoluten und relativen Fehlers als auch bzgl. der Standardabweichung angegeben. Die Modelle wurden auf zwei Weisen charakterisiert: zum einen durch einen allgemeinen Datensatz, welcher sich durch die Werte der ersten zwei simulierten Tage aller Simulationen bildet, und zum anderen durch den Datensatz der jeweils betrachteten Simulation bis zur letzten Veränderung vor dem Prognosezeitpunkt.

	Modell	absoluter Fehler	relativer Fehler	Standardabweichung
(4.11)	allgemein	1,03	7,76 %	1,36
	individuell	0,96	7,60 %	1,26
	beide gemittelt	0,93	7,19 %	1,22
(4.12)	allgemein	0,90	6,72 %	1,20
	individuell	0,91	7,11 %	1,40
	beide gemittelt	0,84	6,42 %	1,16

Tabelle 4.2: (Durchschnittlicher) Prognosefehler der Regressionsmodelle im simulierten Betrieb

Grundsätzlich lässt sich festhalten, dass die allgemeinen Modelle bereits sehr gute Ergebnisse liefern und nur geringfügig durch die anderen untersuchten Modelle übertroffen werden. Beim Modell (4.11) ergibt die individuelle Charakterisierung nur eine marginale Verbesserung. Durch Kombination beider Modelle (Bildung des arithmetischen Mittels der Prognosewerte) wird die beste Prognosegüte erreicht. Das Modell (4.12) erreicht durchweg bessere Ergebnisse. Eine Ausnahme stellt die individuelle Charakterisierung dar. Hier deutet die erhöhte Standardabweichung auf eine höhere Streuung der Werte und damit eine geringere Zuverlässigkeit hin. Die allgemeine Charakterisierung liefert hier bessere Ergebnisse. Allerdings führt die Kombination beider Modelle dennoch zu einer Verbesserung, wenngleich nicht in dem Maße wie beim Modell (4.11). Sämtliche Modellvariationen liefern hohe Prognosequalitäten von durchschnittlich über 92 % Genauigkeit, bei einem Unterschied zueinander von maximal 1,3 %. Für alle nachfolgenden Untersuchungen wird daher zur Vereinfachung das allgemein charakterisierte Modell (4.11) verwendet.

In Verbindung mit den Energiebedarfsmodellen aus Abschnitt 4.2 kann nun der Energiebedarf des Rechenzentrums abgeschätzt werden. Es ist auch möglich den Energiebedarf direkt durch das Modell (4.11) abschätzen zu lassen. Die erzielte Güte ist vergleichbar, was vornehmlich darin begründet ist, dass die Energiebedarfsmodelle einen linearen Zusammenhang zwischen Auslastung und Energiebedarf aufweisen. Komplizierter stellt sich die Abschätzung heraus, wenn heterogene Server eingesetzt werden. Zwar besitzen sie für sich selbst genommen jeweils ein lineares Energiebedarfsmodell, wenn sie jedoch in der Gesamtheit betrachtet werden, dann ist ein einfacher linearer Zusammenhang wie bei homogenen Servern nicht mehr erkennbar.

In Abbildung 4.24 ist die Leistungsaufnahme von Servern in Abhängigkeit von der Gesamtauslastung (*SoL*) abgetragen. Zur Ermittlung der Daten wurde ein sich dynamisch

4 Konzeption samt Komponentenevaluation

verändernder VM-Pool mit einer Größe von bis zu 500 VM über 10 Tage simuliert. Im Graphen für die homogenen Server sind kleine Sprünge vorhanden, die das Hinzuschalten von Servern, also den Anstieg um die Idle-Power kennzeichnen. Bei heterogenen Servern sind zwar auch einzelne lineare Verläufe durch einzelne Server erkennbar, aber da die Sprünge unterschiedlich groß sind und die Gradienten unterschiedliche Steigungen aufweisen, hat die Auswahl eines konkreten Servers für den Betrieb einen großen Einfluss auf die Leistungsaufnahme. Für die Auswahl ist die Bewertungsfunktion (4.7) zuständig, die in Abschnitt 4.3.2 eingeführt wurde. Wie in der dort durchgeführten Untersuchung zeigt sich auch in Abbildung 4.24, dass die Bewertungsfunktion für unterschiedliche Leistungsaufnahmen sorgt. Neben dem immer noch sehr großen Einflussfaktor *SoL* müssen somit noch weitere detektiert werden, was allerdings nicht mehr Bestandteil dieser Arbeit ist.

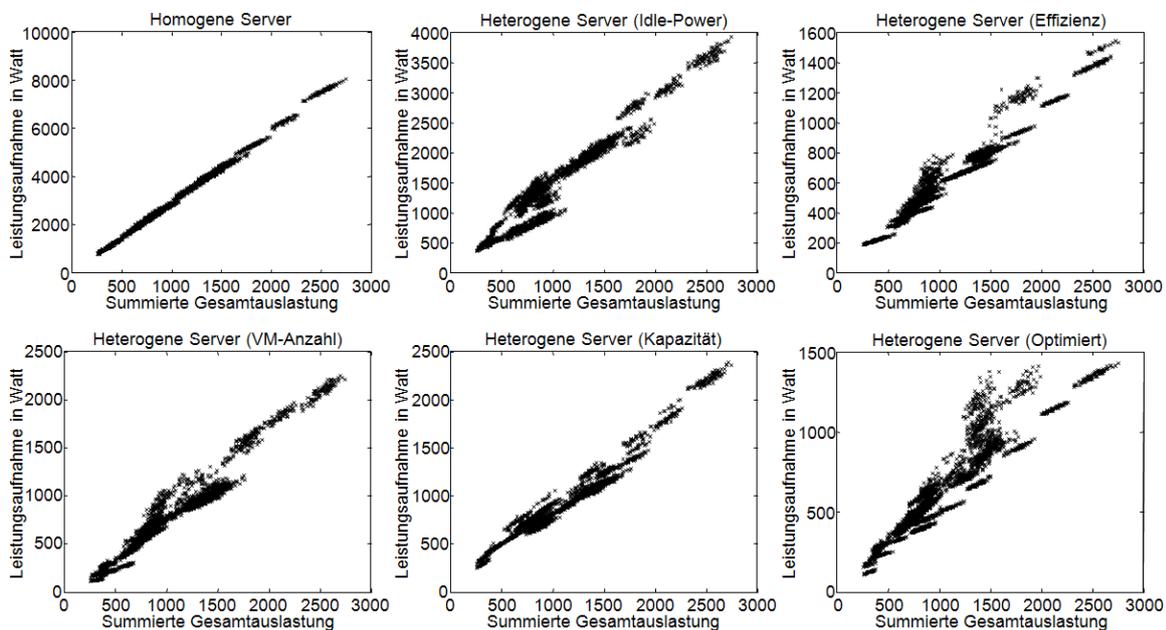


Abbildung 4.24: Bezug zwischen Leistungsaufnahme und summierter Gesamtauslastung der VMs bei homogenen sowie heterogenen Servern mit unterschiedlichen LPM-Policies

Das erstellte Verhaltensmodell ist damit für das hier betrachtete LPM mit jeweils homogenen Servern innerhalb eines Rechenzentrums einsetzbar. Für andere LPM-Verfahren und eben auch für heterogene Server müssen, wie bereits beschrieben, entsprechend passende Modelle erstellt werden. Innerhalb der Systemarchitektur sind diese jedoch austauschbar. Einen Einfluss auf ein globales LPM haben unterschiedliche LPM-Modelle insofern, als dass die Effizienz des globalen LPM von der Effizienz der lokalen LPMs abhängig ist, denn wenn das lokale LPM nur eine geringe Energiebedarfsdynamik bewirkt, kann das globale LPM nur auf ein geringes Verschiebepotential bzgl. der Leistungsaufnahme zurückgreifen.

4.5 Lastganggenerierung

In den vorangegangenen Abschnitten wurde einerseits das Energiebedarfsmodell eines Rechenzentrums in Abhängigkeit von der Hardwareauslastung und andererseits ein LPM-Verhaltensmodell erstellt, das eine Abschätzung über die Anzahl benötigter Server in Abhängigkeit von den zu betreibenden Diensten bzw. VMs gibt. Beide Modelle zusammen ermöglichen ein Energiebedarfsmodell in einem dynamisch optimierten System, welches den Energiebedarf zum Betrieb bestimmter VMs ermittelt.

Entsprechend den variierenden Auslastungsprofilen der VMs und der durch das LPM ermöglichten dynamischen Anpassung ergibt sich für den Betrieb von VMs ein bestimmter Energiebedarfslastgang für das Rechenzentrum. Um für den Rechenzentrumsverbund die Lastgänge energie- und kostenminimierend zu verändern, müssen Modifikationen in der Auswahl der konkreten VMs an unterschiedlichen Standorten erfolgen.

4.5.1 Problembeschreibung

Die Optimierung des Lastgangs setzt voraus, dass Lasten in einem Rechenzentrum generiert werden können. Hierzu müssen VMs gezielt zum Betrieb selektiert werden. Es handelt sich hierbei um ein kombinatorisches Optimierungsproblem, das eine spezielle Variante des Rucksackproblems ist. Beim Rucksackproblem werden in einen Behälter mit endlichem Fassungsvermögen Objekte hineingesteckt, die neben ihrer Größe auch eine gewisse Wertigkeit aufweisen. Diese Wertigkeit gilt es unter Einhaltung der Größenbeschränkung zu maximieren. Gibt es mehr als nur eine Dimension für die Größenbeschränkung, so spricht man von einem multidimensionalen (z. B. bei Flächen oder Raumgrößen) oder Vektor-Rucksackproblem (z. B. Höhe und Gewicht). Wenn die zu berücksichtigende Größe und die Wertigkeit identisch bzw. linear voneinander abhängig sind, dann handelt es sich nach Martello und Toth [143] um das Subset-Sum-Problem. Auch hierbei können mehrere Dimensionen vorkommen.

Beim Problem der VM-Selektion handelt es sich um ein Vektor-Subset-Sum-Problem. Das Rechenzentrum in seiner Gesamtkapazität zum Betrieb von IT-Diensten entspricht dem Behälter. Als Größenangabe ist jedoch nicht die Aufnahmekapazität von IT-Rechenleistungen, sondern der resultierende Energiebedarf anzusehen. Damit hat es das gleiche Maß wie die Vorgabe einer Energiebedarfskurve. Diese wird für unterschiedliche, diskrete Zeitpunkte vorgegeben und besteht damit aus einem Vektor von zu erreichenden Energiebedarfswerten. Die VMs stellen die Objekte dar, die als Eigenschaft einen Ressourcenbedarfsverlauf besitzen. Mit Hilfe des Rechenzentrumsenergiebedarfsmodells und der LPM-Abstraktion kann der Energiebedarfsanteil jeder einzelnen VM ermittelt werden.

Das Subset-Sum-Problem gehört zur Klasse der NP-schweren Probleme [143]. Naive Verfahren zur Ermittlung einer optimalen Lösung dieses Problems haben eine exponentielle Laufzeitkomplexität zur Anzahl der Objekte: $O(2^n)$. Durch geeignete Vorverarbeitung bzw. Sortierung der Elemente kann die Laufzeitkomplexität auf $O(\sqrt{2}^n)$ gesenkt werden [243]. Mit dynamischer Programmierung kann sogar ein Lösungsverfahren für das Subset-Sum-Problem

realisiert werden, das die pseudo-polynomielle Laufzeitkomplexität $O(\text{Summe} \cdot n)$ hat.⁵ Der Ansatz in diesem Verfahren besteht darin, Submengen zu finden, die summiert eine bestimmte Summenzahl ergeben, die wiederum inkrementell bis auf die gesuchte Summe erhöht wird. Die Zwischenergebnisse werden dabei in einem Array gespeichert und für die weiteren Iterationen genutzt. Dies führt zu einer Speicherkomplexität von $O(\text{Summe})$.

Das Verfahren ist allerdings nur auf ganzzahlige Werte anwendbar. Die Vorgaben des Energiebedarfs bzw. der Leistungsaufnahme sind i.d.R. aus \mathbb{R} . Mit 1 Watt als kleinste Einheit ließe sich eine Quantisierung mit vernachlässigbarem Fehler durchführen. Bei größeren Rechenzentren im dreistelligen kW-Bereich würde dies allerdings zu sehr hohem Speicheraufkommen führen, ganz zu Schweigen bei Megarechenzentren im MW-Bereich. Um die Berechenbarkeit gewährleisten zu können, müsste unter Umständen eine gröbere Quantisierung gewählt werden. Dies verändert den Grad an Gründlichkeit p bei der Lösungsfindung und hat direkten Einfluss auf die Lösungsqualität. Das größere Problem stellt jedoch dar, dass beim Vektor-Subset-Sum-Problem neben der Anzahl an Elementen n als zusätzliche exponentielle Einflussgröße die Länge l des Vektors hinzukommt.

Zur Lösung des Vektor-Subset-Sum-Problems existiert kein effizientes Verfahren bzw. keine spezielle Heuristik. Daher soll eine Metaheuristik eingesetzt werden, die auch weitestgehend den Ansatz des zuletzt diskutierten Verfahrens umsetzt. Abstrakt betrachtet, besteht der Lösungsansatz im inkrementellen Aufbau einer Lösung, die sich in ihrer Güte immer weiter der Vorgabe annähert. Des Weiteren werden die oben genannten Einflussparameter sicherlich auch in der einzusetzenden Heuristik Einfluss auf die Skalierbarkeit haben. Daher sollten sie einstellbar sein, um einen guten Trade-Off zwischen Berechnungszeit und Lösungsqualität ermitteln zu können.

Während sich die Einflussgröße p direkt auf die Lösungsqualität auswirkt, liegt für den Parameter l ein indirekter Zusammenhang vor: Die Verringerung der Vektorlänge für die Zeitpunkte zur Einhaltung eines bestimmten Energiebedarfs käme einer Reduzierung der Abtastung gleich. In den nicht berücksichtigten Zeitpunkten kann allerdings theoretisch ein unbegrenzt hoher Fehler auftreten. Der Parameter l sollte daher nicht verändert werden.

Bekanntes Greedy-Verfahren wie der Bergsteiger-Algorithmus und die Tabu-Suche sowie auch evolutionäre Algorithmen lassen sich modifizieren, so dass n und p verändert werden können. Der gezielte inkrementelle und qualitätsapproximierende Aufbau einer Lösung ist jedoch nicht ausdrücklicher Bestandteil dieser Verfahren. Eine Heuristik, in der ein solcher Aufbau speziell vorgesehen ist, ist GRASP.

4.5.2 GRASP

GRASP steht für *Greedy Randomized Adaptive Search Procedures*. Es handelt sich um eine Metaheuristik, die in einem Artikel von Feo und Resende [83] vorgestellt wurde. Eine Zusammenfassung über die Grundlagen des Verfahrens ist in einem Paper von Festa zu finden [84].

⁵<http://www.geeksforgeeks.org/dynamic-programming-subset-sum-problem/>, letzter Zugriff: 08.08.2014

Hierin werden auch Anwendungen vorgestellt, in denen GRASP mit anderen Metaheuristiken wie die Tabu-Suche oder Genetischen Algorithmen kombiniert wurden. Hierdurch konnten die bisher erzielten Lösungen zum Teil noch verbessert werden.

Funktionsweise

Das Verfahren ist in zwei Phasen aufgeteilt: Aufbauphase und lokale Suche. In der ersten Phase wird mit randomisiertem Einfluss eine initiale Lösung erstellt. Von dieser ausgehend wird in der zweiten Phase eine lokale Suche gestartet, um die initiale Lösung zu verbessern. Das Ergebnis der zweiten Phase stellt die Lösung einer GRASP-Runde dar. Insgesamt werden mehrere Runden durchgeführt und die beste Lösung aus allen Runden stellt die finale Lösung dar.

Phase 1 – Randomisierter Aufbau Zu Beginn ist die Lösungsmenge leer. Es wird als erstes die sogenannte RCL-Liste (Restricted Candidate List) aufgebaut. Diese hat eine beschränkte Größe und wird mit zufälligen möglichen Elementen gefüllt. Diese Elemente werden entsprechend einer Bewertungsfunktion sortiert. Aus den besten Elementen wird zufällig ein Element gewählt und in die Lösung aufgenommen. Die Anzahl der besten Elemente ergibt sich anhand eines festen RCL-Quotienten bzgl. der Größe der RCL-Liste. Die übrigen Elemente werden zurück zu den anderen möglichen Elementen gelegt. Danach beginnt die nächste Iteration mit dem Aufbau der RCL-Liste. Die Bewertungen berücksichtigen immer die aktuelle Lösung, weshalb sich die Bewertungen für ein Element bei unterschiedlichen Iterationsschritten i.d.R. unterscheiden. Es werden so lange Elemente der Lösung hinzugefügt, bis die Abbruchbedingung erfüllt ist: Neben dem Erreichen der Vorgabe könnte dies z. B. auch die Überschreitung einer Kapazitätsschranke sein.

Phase 2 – Lokale Suche In der zweiten Phase wird ausgehend von der durch Phase 1 ermittelten initialen Lösung eine lokale Suche durchgeführt. Jedes Element aus der aktuellen Lösung kann durch ein anderes Element, welches sich in einer definierten Nachbarschaft zum auszutauschenden Element und sich derzeit nicht in der Lösung befindet, ausgetauscht werden. Es ist nicht vorgesehen, dass weitere Elemente der Lösung hinzugefügt werden. Auch hier müssen die Elemente bei jeder Veränderung der Lösung neu bewertet werden. Im Gegensatz zu Phase 1 wird nun jedoch immer das beste Element genommen und mit dem aktuellen Element ausgetauscht – soweit es besser als dieses bewertet ist. Die lokale Suche terminiert für gewöhnlich, wenn keine bessere Lösung mehr gefunden werden kann oder eine vorher festgelegte maximale Iterationsanzahl erreicht wird.

Gegenüber dem exakten Verfahren für das Subset-Sum-Problem mit pseudo-polynomieller Laufzeitkomplexität wird hier beim Lösungsaufbau nicht der Zielwert inkrementell erhöht, sondern Elemente mit besten Werten inkrementell der Lösung hinzugefügt. Dadurch wird zwar eine schnelle Annäherung an den Vorgabewert erzielt, aber ob dieser mit minimalem Abstand

erreicht werden kann, wird nicht sichergestellt. Hierzu sieht GRASP vor, eben nicht immer die Elemente mit besten Werten zu nehmen, sondern jeweils aus den besten Elementen in der RCL-Liste zufällig ein Element zu wählen. Mit Durchführung mehrerer Runden entstehen so unterschiedliche Initillösungen, welche durch die lokale Suche noch verbessert werden können.

Anpassungen

Als Elemente werden hier die einzelnen VMs betrachtet. Die Lösung besteht somit aus einer Menge an zu betreibenden VMs, die in ihrer Gesamtheit die gewünschte Soll-Leistungsaufnahme verursacht. Zur Adaption der Problemstellung müssen einige im Folgenden aufgeführte Anpassungen vorgenommen werden.

Bewertung Mit Hilfe des Rechenzentrums- bzw. LPM-Modells kann zu jedem Zeitpunkt t die Differenz zwischen der prognostizierten Leistungsaufnahme $power_{model}(S, t)$ einer bestimmten Lösung S und der Soll-Leistungsaufnahme $power_{target}(t)$ ermittelt werden. Die Bewertung R einer VM geht davon aus, dass die VM Teil der Lösung ist und berechnet für diese Lösung die summierten Differenzen für jeden betrachteten Zeitpunkt (siehe Gleichung (4.14)). Je geringer der Wert ist, desto besser ist die Bewertung. Bei der Generierung der initialen Lösung wird der Fokus stärker darauf gelegt, die Verlaufsform nachzubilden. Hierzu werden die Differenzen zum Quadrat betrachtet.

$$R(VM, S) = \sum |power_{model}(S + VM, t) - power_{target}(t)| \quad (4.14)$$

Kapazitätsberücksichtigung Das hier eingesetzte LPM verwendet für die SLO-Zusicherung eine sichere Verteilung. Durch diese wird eine bestimmte Anzahl an Servern entsprechend der prognostizierten Maximalauslastung reserviert. Diese Serveranzahl wird durch die zur Verfügung stehenden Server im Rechenzentrum begrenzt. Mit Hilfe des Konsolidierungsmodells (4.13) kann die benötigte Serveranzahl für die sichere Verteilung einer VM-Menge mit hoher Genauigkeit abgeschätzt werden und somit entschieden werden, ob eine VM-Auswahl die Kapazitätsschranken überschreitet. Wird diese Abfrage nur zu Abbruchzwecken genutzt, kann der bis dahin noch bestehende Fehler entsprechend der Bewertung R verhältnismäßig groß sein. Besser ist es, die Bewertung der VM in Bezug zur benötigten Kapazität zu setzen. Hierzu wird ein Quotient cq genutzt, welcher den verbleibenden Fehler in Relation zur verbleibenden Kapazität stellt. Beim Hinzufügen einzelner VMs wird darauf geachtet, dass der Quotient einen zuvor definierten Grenzquotienten cq_{limit} möglichst nicht übersteigt. Hierdurch wird heuristisch der Fehler bei Erreichen der Kapazitätsgrenzen minimiert.

Abbruchbedingung für die Generierung der initialen Lösung Als Abbruchbedingung wird entweder der Fall genommen, dass die Hinzunahme von weiteren VMs keine Verbesserung der aktuellen Lösung ergibt, oder, dass das Hinzufügen weiterer VMs die Kapazitäten überschreitet.

Nachbarschaft Bei der lokalen Suche wird i.d.R. nicht versucht, bei jedem Element zu überprüfen, ob aus der Gesamtmenge der verbliebenen Elemente ein besseres existiert, sondern es wird nur in der Nachbarschaft des untersuchten Elementes gesucht. Diese Nachbarschaft muss jedoch auf geeignete Weise definiert werden. Es wurde festgelegt, diese anhand der Durchschnittsauslastung über den betrachteten Zeitraum zu bestimmen. Die Idee hierbei ist folgende: Durch das Hinzufügen von VMs wird, in der Gesamtheit betrachtet, der durchschnittliche Energiebedarfslevel erhöht. Bei der initialen Lösung ist dieser Level auf dem Niveau der Vorgabe. Werden Nachbarn mit ähnlicher Durchschnittsauslastung betrachtet, kann eine VM zur Verbesserung der Verlaufsform gewählt werden, ohne dabei den Energiebedarfslevel stark zu beeinflussen. Zum Vergleich: Eine über die Verlaufskorrelation definierte Nachbarschaft liefert für den gleichen Level eine geringere Anzahl guter Alternativen. Dies führt zu einer niedrigeren Konvergenzgeschwindigkeit.

Abbruchbedingung der lokalen Suche In der gewählten Realisierung wurde ein auf die Qualität der Lösung sich beziehendes Abbruchkriterium verwendet. Hierbei wird zuerst versucht, innerhalb einer Iteration jedes einzelne Element der aktuellen Lösung durch ein besseres auszutauschen. Danach wird überprüft, ob die Verbesserung gegenüber der letzten Iteration über einem gewissen Schwellenwert liegt. Falls ja, wird auch weiterhin ein hohes Verbesserungspotential angenommen und somit eine weitere Iteration durchgeführt. Andernfalls wird die lokale Suche beendet.

Der Parameter n , welcher die Anzahl der zu untersuchenden Elemente angibt, entspricht in dieser Heuristik der Größe der RCL-Liste. Der Parameter p , welcher den Grad an Gründlichkeit bei der Lösungssuche angibt und damit die Ergebnisqualität direkt beeinflusst, ist in die Abbruchbedingung für die lokale Suche integriert. Beide Parameter sind veränderbar, womit die Vorgabe erfüllt ist.

4.5.3 Evaluation der Lastganggenerierung

Wie auch bereits bei der Evaluation des LPM-Verhaltensmodells wurden für die Untersuchung dieses Verfahrens Daten realer VMs genutzt (siehe Abschnitt 5.1). Zusätzlich standen eine drei-monatige Aufzeichnung mit einer Auflösung von 15 Minuten über den Strombedarf einer Siedlung⁶ zur Verfügung.

Das hier untersuchte Szenario ist wie folgt: In der unmittelbaren Umgebung einer Siedlung befindet sich ein Rechenzentrum, das Zugriff auf einen globalen VM-Pool besitzt. Dieser könnte bspw. durch eine föderierte Cloud bereitgestellt werden. Um den Energiebedarf der Siedlung zu glätten und dadurch insbesondere Lastspitzen abzufangen, sollte das Rechenzentrum einen zur Siedlung inversen Energiebedarf aufweisen. Die zur Generierung dieses Lastgangs benötigten VMs können aus dem globalen Pool selektiert werden.

⁶Datenquelle: Erhebung im Rahmen eines Kooperationsprojektes zwischen HTEE der TU Braunschweig und BSenergie

Parametereinstellung

Aus der obigen Beschreibung des Verfahrens lassen sich einige Parameter entnehmen, an denen sowohl die Laufzeit als auch die Qualität des Verfahrens variiert werden können. Um einen bestmöglichen Trade-Off zu erzielen, wurden diese nachfolgend aufgeführten Parameter in ihren Auswirkungen untersucht:

- *Größe der RCL-Liste (Parameter n):* Da die Größe für gewöhnlich kleiner ist als die VM-Anzahl im globalen Pool, wird je Iteration nur eine Teilmenge für den Lösungsaufbau herangezogen.
- *Größe des RCL-Quotienten:* Je größer der RCL-Quotient ist, desto mehr Varianz ist bei den Ergebnissen möglich. Führt der Quotient zur Betrachtung von nur einem möglichen Element, wird immer das beste Element genommen.
- *Größe der Nachbarschaft:* Bei der lokalen Suche beeinflusst die Nachbarschaftsgröße u.a. die Auftrittswahrscheinlichkeit lokaler Minima und die Konvergenzgeschwindigkeit der Lösungsfindung.
- *Qualitätsverbesserungslimit für die lokale Suche (Parameter p):* Je geringer der Wert ist, desto länger wird nach Verbesserungen gesucht. Im Falle des Qualitätslimits von 0% wird erst abgebrochen, wenn keine Verbesserungen mehr erzielbar sind.
- *Anzahl an GRASP-Runden:* Legt fest, wie oft das Verfahren durchlaufen werden soll. Es soll hierbei festgestellt werden, wie viele Iterationen mit typischen Datensätzen i.d.R. für gute Ergebnisse notwendig sind.

Eine Betrachtung aller Parameterkombinationen durch kreuzweise Variation wäre nicht nur sehr zeitaufwendig, sondern insbesondere kompliziert zu analysieren: In einem 5-dimensionalen Raum ist es schwierig, anhand der Kurvenverläufe Trade-Off-Entscheidungen zu treffen. Daher werden die Parameter voneinander unabhängig variiert. Um dennoch eventuelle Seiteneffekte einzubeziehen, soll das folgend beschriebene, iterative Vorgehen angewendet werden.

Initial erhält jeder Parameter aus den jeweiligen Betrachtungsintervallen einen Standardwert. Die Parameter werden nun einzeln variiert, während die übrigen ihren Standardwert behalten. Nachdem alle Parameter untersucht wurden, wird anhand einer Analyse der Laufzeit- und Qualitätsverlaufskurven der jeweils beste Wert ermittelt und dieser zum neuen Standardwert bestimmt. Die Untersuchung wird erneut durchgeführt und wiederholt sich, bis an den Ergebnissen keine Veränderungen mehr zu beobachten sind. Aufgrund der Trade-Off-Entscheidungen können Zyklen entstehen. In diesem Fall wird die Parameterbelegung mit der besten erzielten Güte genommen.

Für die Untersuchung wurde ein globaler VM-Pool mit 1000 VM festgelegt. Die Energiebedarfskurve der Siedlung, die durch den Energiebedarf des Rechenzentrums in gegensätzlichem Verlauf nachgestellt werden soll, hat eine Länge von 24 Stunden. Es wurden insgesamt 10

unterschiedliche Zeitabschnitte untersucht, wobei die Vorgabekurve auf einen durchschnittlichen Leistungslevel von 5 kW skaliert wurde, damit der Wert auch mit den vorhandenen VMs generiert werden kann.

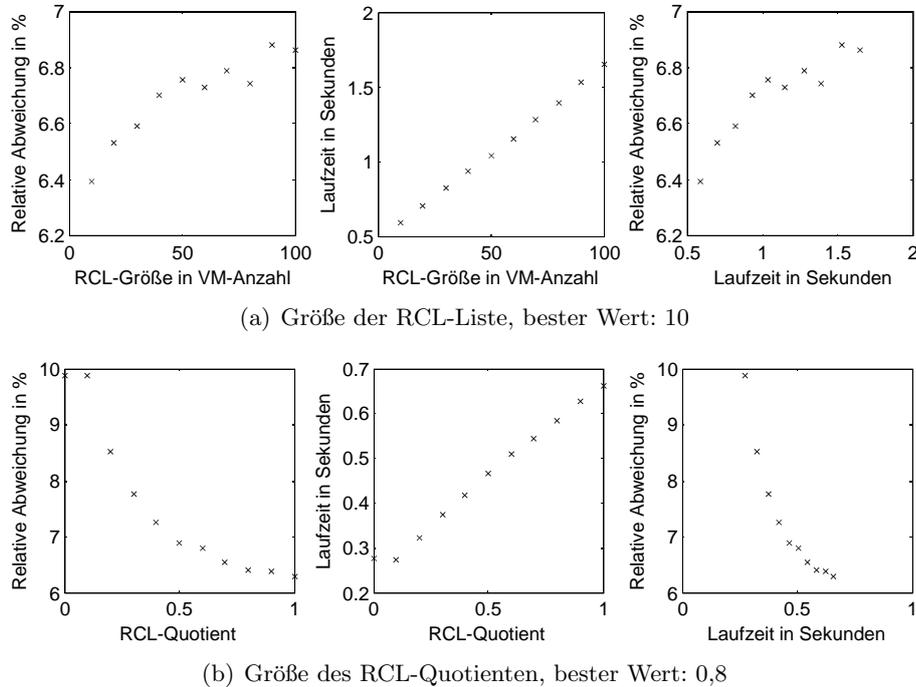
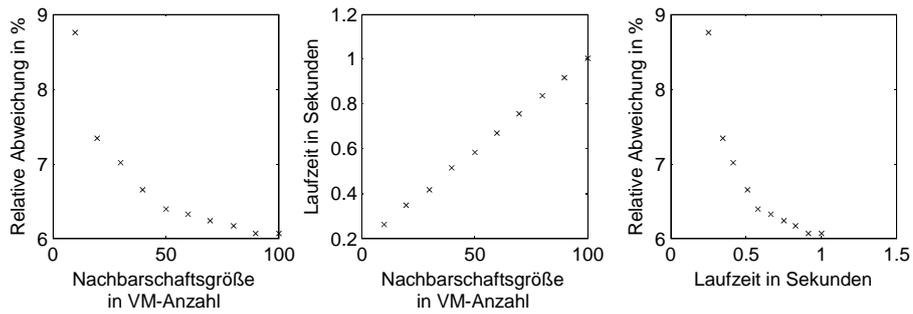


Abbildung 4.25: Ergebnisse des finalen Durchlaufs der GRASP-Parameteruntersuchung (1/2)

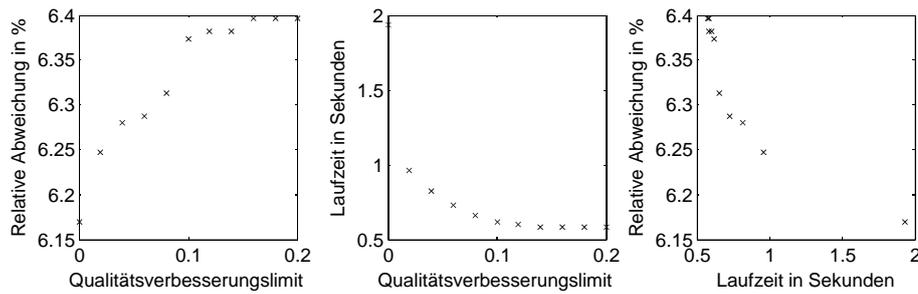
In den Abbildungen 4.25 und 4.26 sind die Ergebnisse des finalen Durchlaufs dargestellt. Für jeden Parameter sind die Ergebnisse in drei Diagrammen abgetragen. Die ersten beiden geben die Auswirkungen auf die Güte und die Laufzeit wieder, im dritten ist die Güte gegen die Laufzeit abgetragen. Die Güte wird durch die relative Abweichung (bezogen auf die über das gesamte Vorgabeintervall summierte Last) ausgedrückt. Die jeweils besten Parameterwerte stehen im jeweiligen Bildtext.

Bis auf das Qualitätsverbesserungslimit besteht zwischen den jeweiligen Parametern und der Laufzeit ein linearer Zusammenhang. Dadurch ähneln sich der Verlauf des jeweils ersten und dritten Diagramms. Die größten Auswirkungen auf die Güte haben der RCL-Quotient und die Nachbarschaftsgröße. Da der Güteverlauf bei beiden dem einer Hyperbel gleichkommt, lohnt sich insbesondere bei der Nachbarschaftsgröße ein hoher Wert in Anbetracht des Laufzeitanstiegs nicht. Beim RCL-Quotienten sind die Ergebnisse für den Wert 0,0 und 0,1 identisch. Dies liegt daran, dass bei der RCL-Listengröße von 10 beide Werte dazu führen, dass nur 1 Element für die zufällige Wahl betrachtet wird. Dass der Wert für die RCL-Größe so klein ist, erscheint auf den ersten Blick widersprüchlich zum Verlauf des RCL-Quotienten zu sein, denn dieser deutet darauf hin, dass eine hohe Varianz besser ist. Durch die kleine Listengröße wird bei der zufälligen Wahl die mögliche Varianz jedoch klein gehalten. An dieser Stelle kommt allerdings der folgende Effekt zum Tragen: Durch die geringe Listengröße werden nur wenige

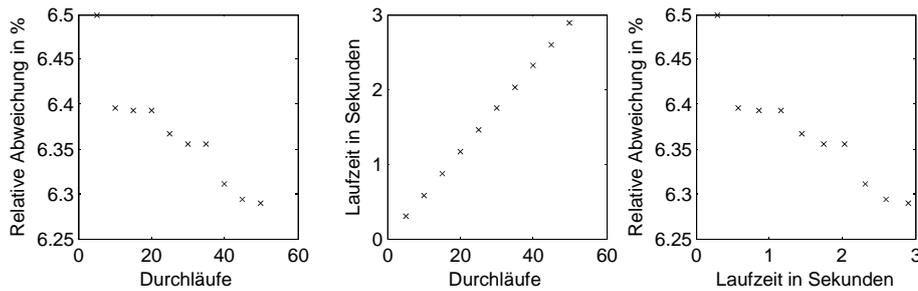
4 Konzeption samt Komponentenevaluation



(a) Größe der Nachbarschaft, bester Wert: 50



(b) Qualitätsverbesserungslimit, bester Wert: 8%



(c) Anzahl an GRASP-Runden, bester Wert: 10

Abbildung 4.26: Ergebnisse des finalen Durchlaufs der GRASP-Parameteruntersuchung (2/2)

Elemente aus der Gesamtmenge vorsortiert. Potentiell werden damit also im Durchschnitt schlechtere Elemente gewählt und nicht die Varianz verringert. Dadurch wird verhindert, dass die Vorgabe mit möglichst wenigen VMs erreicht wird, was einem Greedy-Vorgehen entsprechen würde. Die anschließende lokale Suche hätte dann nämlich einen zu geringen Spielraum für Verbesserungen.

Güteevaluation

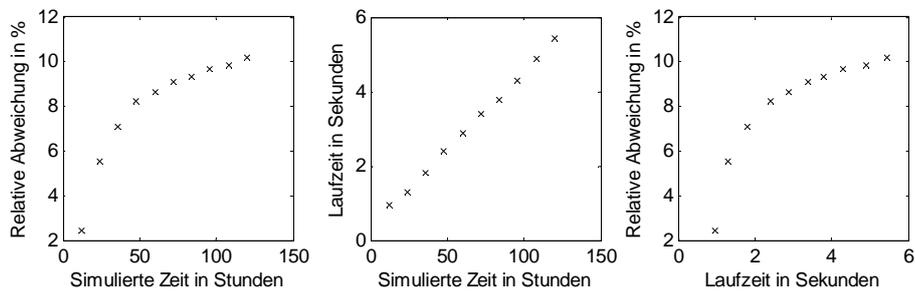
Mit der besten Parametereinstellung soll nun die Güte des Verfahrens im Allgemeinen und das Skalierungsverhalten bzgl. der Qualität und Laufzeit ermittelt werden. Es gibt drei Eingangsparameter, mit denen sich das Problem skalieren lässt:

- *Länge der Energiebedarfsvorgabe*: Eine Energiebedarfsvorgabe von einer Stunde hat aufgrund der Auflösung von 15 Min. vier Betrachtungszeitpunkte. Je Betrachtungszeit-

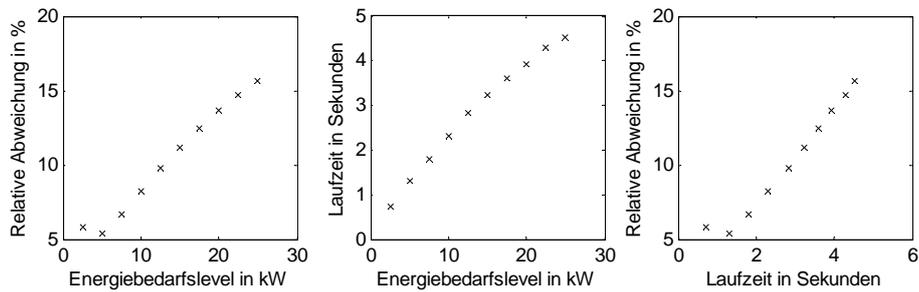
punkt wird die Vorgabe als durchschnittliche Leistungsaufnahme angegeben, um die Lastanpassung auch für eine feinere Auflösung zu ermöglichen. (48 Stunden)

- *Energiebedarfslevel*: Je höher der Level ist, desto mehr VMs müssen i.d.R. für die Lösung gefunden werden. Die Angabe erfolgt der Konsistenz halber wieder als Leistungsaufnahme. (10 kW)
- *Größe des globalen VM-Pools*: Durch Vergrößerung des Pools wird ein größerer Verbund simuliert. (2000 VM)

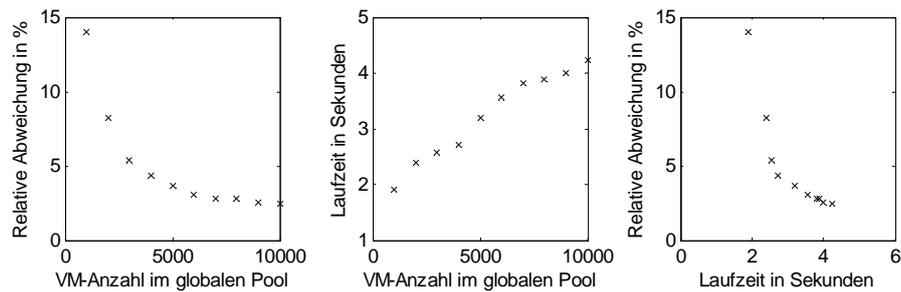
Die Parameter wurden wieder einzeln variiert. Die Standardwerte sind oben in Klammern angegeben.



(a) Länge der Energiebedarfsvorgabe



(b) Energiebedarfslevel



(c) Größe des globalen VM-Pools

Abbildung 4.27: Ergebnisse der Skalierungsuntersuchung

Aus den in Abbildung 4.27 dargestellten Ergebnissen geht hervor, dass die beobachtete Laufzeit von allen Parametern maximal linear abhängig ist. Bei der Länge der Energiebedarfsvorgabe erklärt sich dies durch die linear vergrößerte Anzahl an unabhängig voneinander

zu berücksichtigenden Werten. Ein erhöhtes Energiebedarfslevel bedeutet, dass zu seiner Erreichung im Durchschnitt mehr VMs selektiert werden müssen, was den Selektionsprozess entsprechend verlängert. Die Vergrößerung des VM-Pools kann diesen Prozess auch beeinflussen, indem VMs selektierbar werden, die zu einer schnelleren Konvergenz führen. Größeren Einfluss hat die VM-Pool-Vergrößerung allerdings auf die lokale Suche, da hier der Nachbarschaftsbaum erweitert wird. Damit kann länger nach potentiell besseren VMs gesucht werden, was sich auch positiv auf die Lösungsgüte auswirkt.

Mit steigender simulierter Zeit bzw. Vorgabelänge steigt die relative Abweichung linear, und zwar in drei Abschnitten: $[0,24]$, $[24,48]$ und $[48,120]$, jeweils in Stunden. Die Steigungen reduzieren sich, was sich anhand der periodischen Verläufe der Energiebedarfe und der VM-Auslastungsdaten erklären lässt: Bei den VM-Auslastungsdaten haben die Perioden eine Länge von einem, zwei, sieben oder 14 Tagen, bei den Energiebedarfsverläufen – innerhalb einer Arbeitswoche – eine Länge von einem Tag. Ist eine Lösung für eine bestimmte Periodenlänge gefunden, passt sie auch recht gut für die folgenden Perioden. Der zusätzliche Fehler durch eine längere Vorgabe verringert sich dadurch.

Beim Energiebedarfslevel ist mit einer Ausnahme ein vollständig linearer Zusammenhang zur relativen Abweichung zu beobachten. Dass die Abweichung mit steigendem Level überhaupt zunimmt, liegt daran, dass nicht genügend VMs zur Auswahl vorhanden sind, um sowohl die Höhe als auch den Verlauf nachzubilden. Wird nämlich der globale VM-Pool vergrößert, so verringert sich die Abweichung.

In Abbildung 4.28 ist beispielhaft eine Energiebedarfskurve mit den Parameterwertbelegungen 48 Stunden, 10 kW und 4000 VM angegeben. Die hierzu gefundene Lösung hat eine relative Abweichung von 6,80 %. Die Abweichung kann zwar durch die Vergrößerung des globalen VM-Pools verringert, aber nicht vollständig eliminiert werden. Das hat nicht nur damit zu tun, dass es sich beim Lösungsverfahren um eine Heuristik handelt, sondern insbesondere damit, dass sich die Vorgabe nicht im Lösungsraum befindet. Mit den vorhandenen VMs ist es schlichtweg nicht möglich, die Vorgabe zu erfüllen.

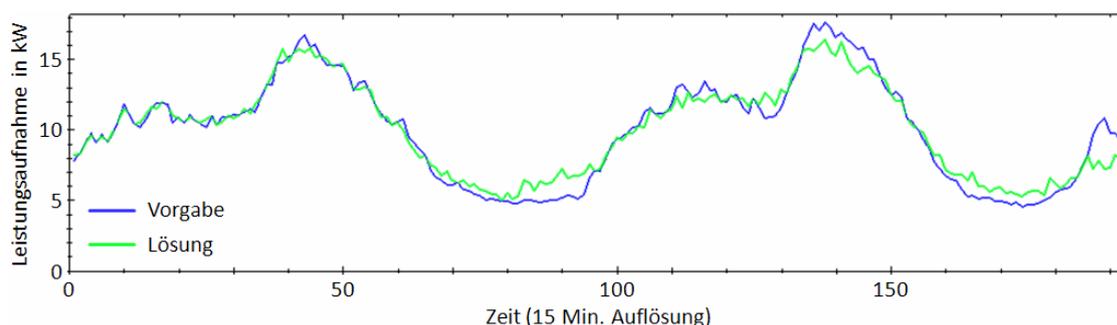


Abbildung 4.28: Lastgangenerierung bei Vorgabe einer (realen) Energiebedarfskurve

Um die Güte der Heuristik zu ermitteln, muss eine Vorgabe geschaffen werden, die durch die VMs tatsächlich nachgebildet werden kann. Hierzu wurden in 100 verschiedenen Simulationen mit den gleichen Parameterbelegungen des Beispiels Vorgaben anhand zufällig aus dem globalen

VM-Pool gewählter VMs generiert, die in etwa den angegebenen Energiebedarfslevel erreichen. Die Lösungen hatten durchweg eine Güte von mindestens 99 %. Eine Gegenüberstellung zu einer Greedy-Methode, welche eine Lösung aufbaut, indem sie iterativ aus dem Gesamtpool die jeweils bestbewertete VM der Lösung hinzufügt, und zu einer zufälligen Auswahl, die zufällig VMs hinzufügt, solange sich das Ergebnis dadurch nicht verschlechtert, zeigt jedoch, dass dieses Szenario nicht zur Ermittlung der Güte herangezogen werden kann, da auch diese Verfahren – insbesondere die zufällige Auswahl – ähnliche Ergebnisse erzielen.

Da somit der Abstand zu einem optimalen Lösungsverfahren auf diese Weise nicht abgeschätzt werden kann, soll an dieser Stelle zumindest eine vergleichende Gegenüberstellung mit den obigen beiden Verfahren (Zufall, Greedy) durchgeführt werden, diesmal mit realen Lastgängen. Mit den Parameterwertbelegungen 48 Stunden, 10 kW und 4000 VM wurden 100 unterschiedliche Lastgänge versucht nachzubilden. Die Ergebnisse dieser Untersuchung sind in Tabelle 4.3 dargestellt. Es sind die Durchschnittswerte der relativen Abweichung und der Laufzeit angegeben.

Verfahren	relative Abweichung	Laufzeit
Zufall	27,89 %	0,19s
Greedy	16,25 %	24,05s
GRASP	6,43 %	4,30s

Tabelle 4.3: Gegenüberstellung vom entwickelten GRASP-Verfahren zur zufälligen Auswahl und zur Greedy-Methode

Die zufällige Auswahl liefert am schnellsten Lösungen. Die Güte ist allerdings am schlechtesten. Bessere Ergebnisse erzielt die Greedy-Methode, benötigt hierzu aber die längste Laufzeit, was sich dadurch erklärt, dass sie in jeder Iteration beim Hinzufügen einer VM sämtliche VMs aus dem gesamten VM-Pool analysiert. Die besten Ergebnisse werden durch das entwickelte (GRASP-)Verfahren erzielt, wobei die Laufzeit sich zwischen der zufälligen Auswahl und der Greedy-Methode einordnet.

Während der Aufbauphase verfährt das entwickelte Verfahren nach einem greedy-randomisierten Einfluss. Bei der lokalen Suche wirkt ausschließlich ein greedy Einfluss. Die Vorteile beider Einflüsse wurden vereint und konnten letztlich zu einem besseren Ergebnis führen. Insgesamt lässt sich also beobachten, dass das entwickelte Verfahren gute Ergebnisse in kurzer Laufzeit liefert.

4.6 Netzwerkmodell

Die bisherigen Abschnitte haben sich mit den Objekten innerhalb eines Rechenzentrums befasst. Um Rechenzentren untereinander zu verbinden, bedarf es eines Kommunikationsnetzwerks. Dieses kann je nach Anbindung unterschiedliche Strukturen aufweisen. Allgemein spricht man vom Internet, das aus einem Access, Metro/Edge und Core Bereich besteht. Größere Rechenzentren besitzen jedoch einen Gateway Router, welcher direkt mit einem Core Router

4 Konzeption samt Komponentenevaluation

eines Providers verbunden ist (siehe Abbildung 4.29) [17]. Provider sind untereinander mit Edge Routern oder auch über Internetknoten verbunden. Über Transit- (der Provider, mit dem kleineren Netz zahlt an den, mit dem größeren Netz) und Peering-Abkommen (Provider mit etwa gleich großen Netzen erheben keine gegenseitigen Abgaben) wird der Traffic zwischen den Netzen finanziell geregelt.

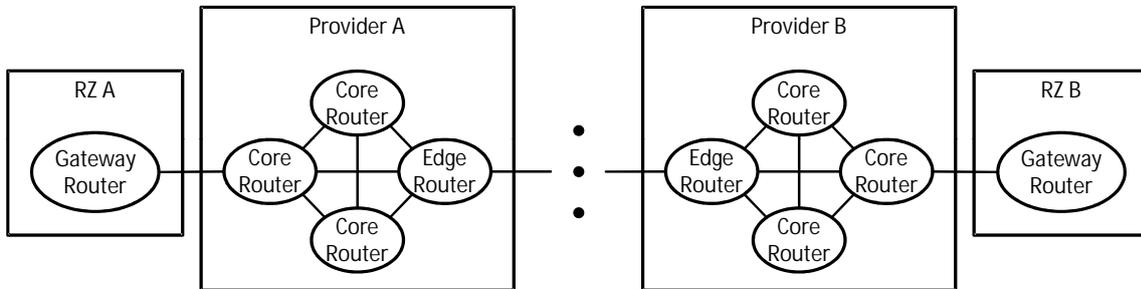


Abbildung 4.29: Netzwerkstruktur (angelehnt an [17]).

Die Verbindung zwischen zwei Rechenzentren erfolgt i.d.R. über einen oder mehrere Provider und damit auch über mehrere Router. Da der Energiebedarf betrachtet werden soll, wird an dieser Stelle nicht auf die konkreten Provider, sondern auf die Infrastruktur eingegangen. Jeder Router stellt einen sogenannten Hop dar. Laut [17] werden für den Versand eines Internetpakets im Durchschnitt 13 Hops benötigt. Mit entsprechenden Energiebedarfsmodellen zu den jeweiligen Komponenten, also auch der Verkabelung, wurde der Energiebedarf des Internets ermittelt. Dieser wird durch Router dominiert, was auch die Analysen in [67] ergeben. Das Ergebnis der Untersuchungen ist ein gewisser Energiebedarf zur Übertragung einer bestimmten Datenmenge: der Netzwerkenergiebedarfskoeffizient. Dieser erlaubt eine abstrahierte Darstellung des Netzwerks.

Für die Bestimmung der Kosten zur Übertragung von Daten sind damit nur der Netzwerkenergiebedarfskoeffizient γ_{Net} , die Kosten pro Energie δ_L am Standort L sowie die Datenmenge D selbst zu berücksichtigen:

$$c(D) = D \cdot \gamma_{Net} \cdot \delta_L \quad (4.15)$$

Der Energiebedarf tritt je Datenübertragung nur einmal auf, aber sowohl Sender i als auch Empfänger j sind an der Übertragung beteiligt. Der Fairness halber wird daher der Energiebedarf auf beide Teilnehmer aufgeteilt, wodurch sich die Formel (4.15) zu folgender modifiziert:

$$c(D_{ij}) = D_{ij} \cdot \gamma_{Net} \cdot \frac{1}{2}(\delta_i + \delta_j), \quad i \neq j \quad (4.16)$$

Veranschaulicht ist dieser Zusammenhang in Abbildung 4.30 und soll anhand eines kleinen Beispiels verdeutlicht werden. Der Datenverkehr D_{AB} zwischen den Rechenzentren RZ A und RZ B wird durch ein homogenes Netzwerk geleitet, welches hierzu eine gewisse Menge an Energie benötigt, die durch den Netzwerkenergiebedarfskoeffizienten γ_{Net} ermittelt wird. Wer Sender oder Empfänger ist, spielt keine Rolle. Die durch die Datenübertragung entstehenden

Kosten sind damit von beiden Rechenzentren in gleichem Maße zu übernehmen. Die Kosten ergeben sich anhand der für die Rechenzentren jeweils gültigen Strompreise δ_A und δ_B .

In den in dieser Arbeit durchgeführten Untersuchungen wird von einer Netzwerkanbindung mit einer Bandbreite von 10 Gbit/s ausgegangen, was z. B. auch in den Analysen in [164] zugrunde gelegt wird und sich aktuell im Produktportfolio des IT- und Telekommunikationsanbieters Colt befindet⁷. Der Einfachheit halber wird die zur Verfügung stehende Bandbreite nicht durch anderweitigen Netzwerkverkehr beeinflusst. Dies bedeutet, dass Netzüberlastungen lediglich durch das rechenzentrumsspezifische Sende- bzw. Empfangsverhalten auftreten können. Der in realen Netzwerken auftretende Traffic, lässt sich z. B. durch den Network Weather Service prognostizieren [244]. Um eine praxisnähere Simulation durchführen zu können, bieten sich z. B. Netzwerkeulatoren wie Apposite Linktropy oder PacketStorm 4XG IP Network Emulator an. Offene software-basierte Netzwerk-Simulatoren finden sich beispielsweise in ns-3 und SNDlib.

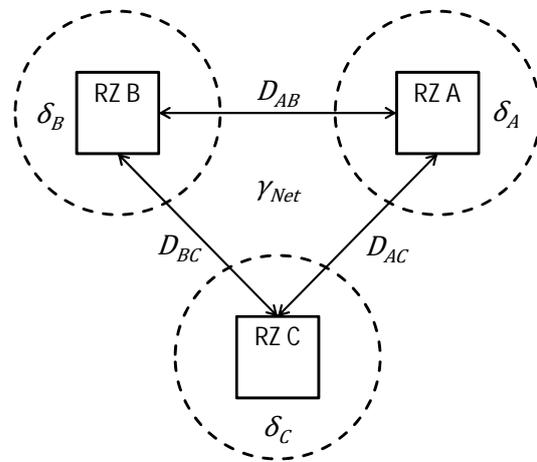


Abbildung 4.30: Netzwerk(-kosten-)modell (abstrahiert)

4.7 Globales Last- und Powermanagement

Die einzelnen Systembestandteile, die ermöglichen, Lasten gezielt an einem Ort aufzunehmen und zu koordinieren, wurden beschrieben. In diesem Abschnitt wird nun erläutert, wie diese im Zusammenspiel zu einem globalen LPM verbunden werden. Das entwickelte globale Verteilungsverfahren wird im Anschluss bzgl. seiner Güte evaluiert. Zusätzlich wird im Zuge der Skalierbarkeitsanforderung die Laufzeitkomplexität des gesamten Verfahrens diskutiert.

Entsprechend der agentenbasierten Aufteilung aus Abbildung 4.1 wird in einem ersten Schritt durch den Arbiter ermittelt, wie die Lastverteilung bei gegebenen und zukünftigen Betriebskonditionen optimaler Weise aussehen sollte. Es wird also für ein bestimmtes Planungsintervall eine optimale Vorgabekurve für jedes Rechenzentrum ermittelt. Dieses Vorgehen entspricht der Intervall-Optimierung in [5], die gegenüber einer kontinuierlichen Optimierung zum Vorteil hat, dass Zusicherungen bzgl. evtl. vereinbarter Lastgänge gegeben werden können. Insofern herrscht Planungssicherheit. Anhand dieser Kurve ermittelt nun jeder Agent für sein Rechenzentrum mit Hilfe der Modelle für den Energiebedarf und des LPM-Konsolidierungsverhaltens sowie des Lastganggenerierungsverfahrens, welche Gruppe an VMs benötigt wird, um der Vorgabe möglichst genau zu entsprechen. In Abbildung 4.31

⁷<http://www.colt.net/de/de/products-services/data-networking/colt-lanlink-de.htm>, letzter Zugriff: 04.09.2014

4 Konzeption samt Komponentenevaluation

ist dieser Vorgang anhand eines Beispiels skizziert und wird im Nachfolgenden detaillierter beschrieben.

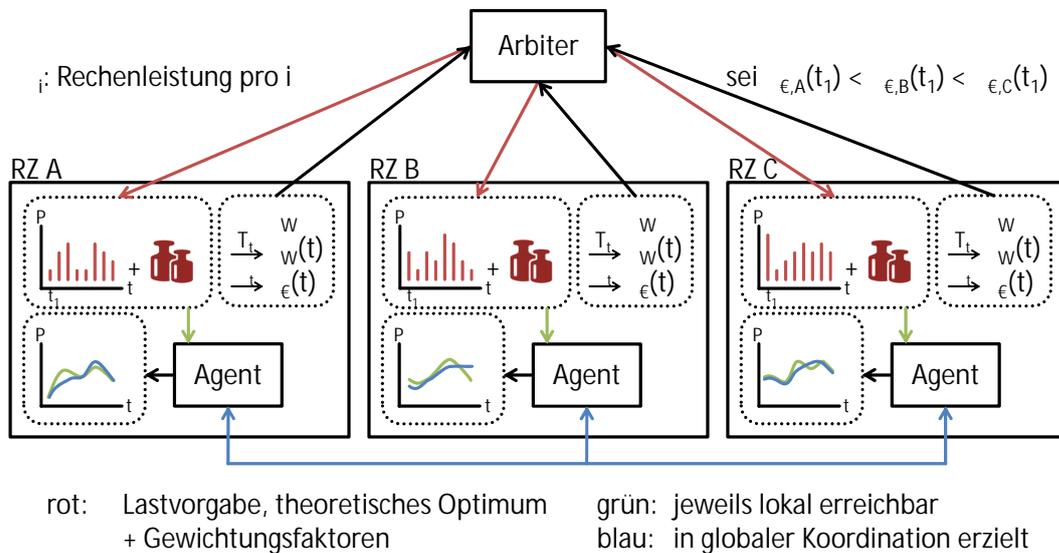


Abbildung 4.31: Übersicht zur Funktionsweise des globalen LPM anhand eines Beispiels

4.7.1 Lastgangvorgabe

In regelmäßigen Abständen wird für den nächsten Planungszeitraum berechnet, wie die Lastverteilung zwischen den unterschiedlichen Standorten für einen kosten- bzw. energieoptimierten Betrieb sein sollte. Hierzu wird zu jedem der betrachteten Zeitpunkte eine Bewertung aller Rechenzentren – angelehnt an den Energieeffizienzkoeffizienten aus [173] – durchgeführt. Als Bewertungsmetrik dient ein aus der Hardware des Rechenzentrums ableitbarer, durchschnittlicher Rechenleistungswert η_W , welcher angibt, wie viel Rechenleistung z. B. in Flops pro Watt bereitgestellt wird. Hieraus ergibt sich die allgemeine Effizienz des Rechenzentrums.

Da dies ein Durchschnittswert ist, könnte der Fall eintreten, dass ein Rechenzentrum mit homogener, aber nur durchschnittlich effizienter Hardware besser bewertet wird, als ein Rechenzentrum mit besonders effizienter Hardware, das aber auch ineffiziente Hardware betreibt. Durch die höhere Präferenzierung des ersten Rechenzentrums würde die effiziente Hardware evtl. nicht genutzt werden. Für diesen Fall könnten virtuelle Rechenzentren eingeführt werden: Das zweite Rechenzentrum würde in zwei virtuelle Rechenzentren aufgeteilt werden, einem gut und einem schlecht bewerteten. Wie in Abschnitt 4.4 begründet, wird hier aber von einer homogenen Hardwareausstattung je Rechenzentrum ausgegangen.

Durch das Einbinden des Klimamodells werden die Temperaturbedingungen mit verrechnet, was zu einem Rechenleistungswert in Relation zum Energiebedarf $\eta_W(t)$ für einen bestimmten Zeitpunkt t führt. Dieser kann zur Energieminimierung herangezogen werden. Für eine Kostenminimierung müssen noch die standortbezogenen Energiekosten einbezogen werden ($\rightarrow \eta_{\epsilon}(t)$). Für jedes Rechenzentrum ergibt sich auf diese Weise zu jedem Zeitpunkt eine energie- bzw. kostenbezogene Recheneffizienz.

Im optimalen Fall sollten Rechenzentren mit der besten Recheneffizienz möglichst voll ausgelastet werden. Um eine solche Lastvorgabe zu berechnen, werden sämtliche VMs, am effizientesten Rechenzentrum angefangen, simulativ alloziert. Sobald dieses keine weiteren VMs aufnehmen kann, werden die übrigen VMs auf das nächst-effizienteste Rechenzentrum alloziert usw. Entscheidend ist hier also die Kapazität des Rechenzentrums, deren Einhaltung mit Hilfe des LPM-Modells für die statische Allokation sichergestellt werden kann.

Algorithmisch ist an dieser Stelle zu beachten, dass gewährleistet ist, dass alle VMs auch tatsächlich alloziert werden. Dies hat den Zweck, einen – zumindest theoretisch – nachstellbaren Lastwert zu ermitteln. Da das LPM-Modell für die sichere Verteilung ein heuristisches Verfahren modelliert (First-Fit Bin-Packing), werden die vorhandenen Ressourcen i.d.R. nicht vollständig ausgenutzt. Durch Neuallokationen können sich Situationen ergeben, in denen der Ressourcennutzungsgrad geringer ausfällt als zuvor, was im Worst-Case dazu führen kann, dass die zuvor in den gesamten Verbund bereits allozierten VMs nun nicht mehr vollständig aufgenommen werden können. Um diesen Fall zu vermeiden, werden die VMs ausgehend von den weniger effizienten Rechenzentren zu den effizientesten simulativ verschoben. Auf diese Weise besteht während des gesamten simulativen Allokationsprozesses eine gültige Verteilung, in der alle VMs in einem Rechenzentrum platziert sind.

Das Ergebnis dieses Vorgangs ist eine diskrete Lastgangvorgabe für jedes Rechenzentrum (in Abbildung 4.31 rot dargestellt).

4.7.2 Migrationsentscheidung

Jeder lokale Agent erhält die für seinen Standort zu erfüllende Lastgangvorgabe. Da diese eine rein theoretisch optimale Vorgabe ist, wird der erreichbare Lastgang sich dieser lediglich annähern können.

Mit Hilfe des GRASP-Verfahrens bestimmt der Agent unter Berücksichtigung der evtl. Migrationskosten die gewünschten VMs aus dem globalen VM-Pool, woraus sich ein lokal erreichbarer Lastgang ergibt (in Abbildung 4.31 als grüne Kurve dargestellt). Migrationskosten ergeben sich durch das Verschieben über das Netzwerk und sind auch abhängig von der Größe der zu verschiebenden Daten.

Die Betriebskosten der Rechenzentren verändern sich über die Zeit. Aus diesem Grund verändern sich nicht nur die günstigsten Standorte, sondern in erster Linie variiert die Höhe des Einsparpotentials. Abweichungen zur Vorgabe haben damit eine unterschiedliche Tragweite. Um dem Rechnung zu tragen, wird ein Gewichtungsfaktor eingeführt (in Abbildung 4.31 als Gewichte dargestellt), welcher bei der Lösungsfindung mit dem Fehlerwert multipliziert wird. Bei zu befüllenden Rechenzentren ergibt sich dieser Faktor aus der Differenz zwischen dem betrachteten Rechenzentrum und dem effizientesten, welches optimaler Weise gar nicht bzw. nur zu Teilen ausgelastet werden sollte. Dieses wäre nämlich das erste Rechenzentrum, in dem die nicht zugewiesenen VMs betrieben werden würden. Der Gewichtungsfaktor für die zu leerenden Rechenzentren ergibt sich analog als betragsmäßige Differenz zum ineffizientesten voll bzw. teils auszulastenden Rechenzentrum.

Werden Lastgänge von extern vorgegeben, wie beim DSM (in diesem Fall entfällt die Berechnung der Lastgangvorgabe durch den Arbitrer), können Gewichtungsfaktoren dazu genutzt werden, zeit- und standortabhängige Prioritäten festzulegen. Durch die Verwendung von Gewichtungsfaktoren lassen sich auch unterschiedliche Optimierungsziele gemeinsam betrachten. Wie die Berechnung der Gewichtungsfaktoren in diesem Fall erfolgen sollte, ist jedoch nicht mehr Bestandteil der Dissertation.

Soweit kann jeder Agent selbstständig agieren. Da jedoch die VM-Selektionen ohne Beachtung der anderen Agenten ermittelt werden, tritt das Problem auf, dass VMs mehrfach, nämlich von unterschiedlichen Agenten selektiert werden. In diesen Fällen entscheidet der jeweilige Agent desjenigen Rechenzentrums, in dem die betreffende VM aktuell betrieben wird, welcher Agent den Zuschlag erhält. Dies ist davon abhängig, wodurch die jeweiligen Vorgaben am besten erreicht werden. Da hierdurch die Lösung der übrigen Interessenten verschlechtert wird, muss für diese eine weitere Suche angestoßen werden. Der Pool an zur Verfügung stehenden VMs wird dabei um die bereits zugewiesenen VMs verringert. Dadurch, dass das GRASP-Verfahren auch eine systemerhaltende Suche durchführen kann, d.h., dass die Ausgangslösung auch mit Veränderungen als Ausgangsbasis für eine neue Suche genutzt werden kann, ist eine neue Lösung in einem Bruchteil der Ausführungszeit der initialen Lösungsfindung ermittelbar. Der Vorgang der VM-Selektion wiederholt sich so lange, bis alle VMs für den nächsten Allokationszeitraum eingeplant sind.

Für jeden Zeitpunkt im Planungsintervall wird also zuerst durch den Arbitrer eine theoretisch optimale Lastverteilung ermittelt und im Anschluss durch die Koordinierung der lokalen Agenten eine an die optimale Verteilung angenäherte und praktisch erreichbare Lastverteilung gebildet (in Abbildung 4.31 als blaue Kurven dargestellt).

In dieser Realisierung handelt es sich um ein recht simples MAS, das mehr einem verteilten Algorithmus entspricht und noch dazu auch sequentiell realisierbar ist. Die Simplität lässt sich auch mit einem Bezug zum BDI-Modell (Belief-Desire-Intention) zeigen. Das Weltwissen (Beliefs) wird zu Beginn jedes Planungsintervalls aktualisiert, damit die Agenten ihre Entscheidungen qualifiziert fällen können. Die Ziele (Desires) jedes Agenten bleiben immer gleich: Der globale Agent verfolgt die globale, der lokale Agent die lokale Optimierung. Die Absichten (Intentions) zur Erreichung der Ziele sind die Ermittlung der theoretisch optimalen Lastgangvorgaben sowie die Annäherung an diese. Die Mittel, um dieses umzusetzen, sind fix. Ein Agent hat somit keinen Handlungsspielraum darüber, welche Aktion er durchführen sollte.

Eine alternative Umsetzung der Migrationsentscheidungen mittels eines Marktplatzes würde den einzelnen Agenten einen größeren Handlungsspielraum einräumen. Hiermit würde das MAS auch eine globalere Sichtweise bei der Problematik der Mehrfachselektionen erhalten. In der aktuellen Variante wird hierzu für jede einzelne VM die jeweils beste Lösung gewählt. Da dies unabhängig von anderen, ebenfalls mehrfach selektierten VMs erfolgt, ist die finale Lösung i.d.R. nicht optimal. Eine Realisierung des Marktplatzmechanismus besitzt somit Potential, ein besseres Ergebnis zu liefern. Um welches Verbesserungspotential es sich hierbei handelt, zeigt die in Abschnitt 4.7.4 durchgeführte Evaluation. Das grundsätzliche Einsparpotential könnte

hiermit also evtl. erhöht werden, aber an den Parameterabhängigkeiten, deren Ermittlung den Fokus der Arbeit darstellt, würde sich hiermit grundsätzlich nichts ändern. Daher und aufgrund der höheren Komplexität wurde im Rahmen der Dissertation auf diese Realisierungsvariante verzichtet.

4.7.3 Migrationsdurchführung

Nach Abschluss der Planungsphase erfolgt die Durchführung der Migrationen zur Laufzeit. Für die einzelnen Migrationen muss entschieden werden, wann sie erfolgen sollen. Wie bereits in Abschnitt 4.3.3 zu den Dynamisierungserweiterungen erläutert, können VMs nicht sofort zwischen den Rechenzentren migriert werden, da die lokalen LPMs erst den Raum für die VM-Aufnahme schaffen müssen und evtl. bevorstehende Engpässe priorisiert auflösen müssen. Für die Durchführung einer bestimmten Migration müssen beide beteiligten Rechenzentren bereit sein.

Die Entscheidung, wann eine Migration durchgeführt werden sollte, lässt sich auch anhand des möglichen Einspareffektes ordnen. Da die VMs unterschiedliche Energiebedarfsverläufe implizieren, können die hierzu sensiblen Einspareffekte durch eine entsprechende Migrationsreihenfolge optimiert werden. Im Rahmen dieser Arbeit wurde das Potential einer optimierten Reihenfolge jedoch nicht untersucht.

4.7.4 Evaluation der Migrationsentscheidung

Das vorgestellte Verfahren zur Durchführung des globalen LPM beinhaltet die Bestimmung der theoretisch optimalen Lastgangvorgaben und die Ermittlung der durchzuführenden Migrationen, um den Lastgangvorgaben möglichst zu entsprechen. In dem hier beschriebenen Prozess der Lastgangvorgabe wurde der Fall behandelt, dass eine Minimierung der Gesamtkosten bzw. des Gesamtenergiebedarfs erfolgen soll. Ist jedoch eine Lastgangvorgabe von extern gegeben, dann ist ihre Bestimmung nicht mehr Bestandteil des globalen LPM. Die Güte des globalen LPM hängt in der obersten Verfahrensschicht somit nur noch von den Agentenentscheidungen bezüglich der Migrationen ab. Diese Güte soll in diesem Abschnitt evaluiert werden.

Wie auch bei der Evaluation der Lastganggenerierung, besteht hier die Problematik, Szenarien zu erschaffen, aus denen die Güte abgeleitet werden kann. Denn ist die Lastgangvorgabe nicht erreichbar, muss der Abstand zu einer optimalen Lösung bestimmt werden, die ihrerseits erst noch ermittelt bzw. abgeschätzt werden muss. Da hierfür kein entsprechendes Verfahren vorliegt, werden erreichbare Vorgaben anhand von Partitionen der Gesamtmenge der im Verbund betriebenen VMs abgeleitet. Der Aspekt der Migrationskosten wird hierbei ausgespart.

Bei der Evaluation der Lastganggenerierung hat das Bestimmen einer Lastgangvorgabe anhand zufällig gewählter VMs aus dem Gesamtpool zu einem Szenario geführt, aus dem keine Güte abgeleitet werden kann, da ein Vergleich mit einer zufälligen Auswahl keine Unterschiede gab und die Lösungen nahezu optimal waren. In dieser Problemstellung sind jedoch durchaus

4 Konzeption samt Komponentenevaluation

nennenswerte Unterschiede zu beobachten.

Um den Prozess der Migrationsentscheidung unabhängig von den Fehlereinflüssen der Modelle zum Energiebedarf und zur LPM-Konsolidierung bewerten zu können, wird das Abstraktionsmodell als Ausgangsbasis genommen. Die Energiebedarfe werden also nicht mit den einzelnen Modellen der Komponenten nach Simulation mit dem lokalen LPM ermittelt, sondern anhand des Abstraktionsmodells, sowohl für die Lastgangvorgabe als auch für die Lastganggenerierung.

Zur Bestimmung der Güte wird somit wie folgt vorgegangen: Der Verbund besteht aus einer bestimmten Menge an homogenen Rechenzentren und betreibt eine gewisse Gesamtmenge an VMs. Die Kapazitäten der Rechenzentren sind bei Gleichverteilung der VMs zu 50 % belegt. Nun werden zur Bestimmung der Lastgangvorgaben die VMs zufällig auf die Rechenzentren verteilt. Dabei wird allerdings darauf geachtet, dass einige Rechenzentren eher gefüllt und einige andere eher geleert sind. Für ein bestimmtes Intervall werden nun mit Hilfe des Abstraktionsmodells die Energiebedarfe für jedes Rechenzentrum bestimmt. Dies sind die Lastgangvorgaben. Die Gewichtungsfaktoren sind konstant 1. Ein optimales Verfahren würde genau die gleichen VM-Mengen-Partitionen ermitteln (von VMs mit gleichen Auslastungsprofilen abgesehen).

In Abbildung 4.32 sind die Ergebnisse dieser Untersuchung für unterschiedliche Rechenzentrumsanzahlen und Intervallgrößen angegeben. Die VM-Anzahl entspricht der Rechenzentrumsanzahl multipliziert mit dem Faktor 200. Jeder dargestellte Messwert entspricht dem Durchschnitt der Ergebnisse von 100 Szenarien, die sich aus der Variation von 10 unterschiedlichen Gesamt-VM-Pools und 10 unterschiedlichen Zufallsverteilungen für die Bestimmung der Lastgangvorgaben ergeben.

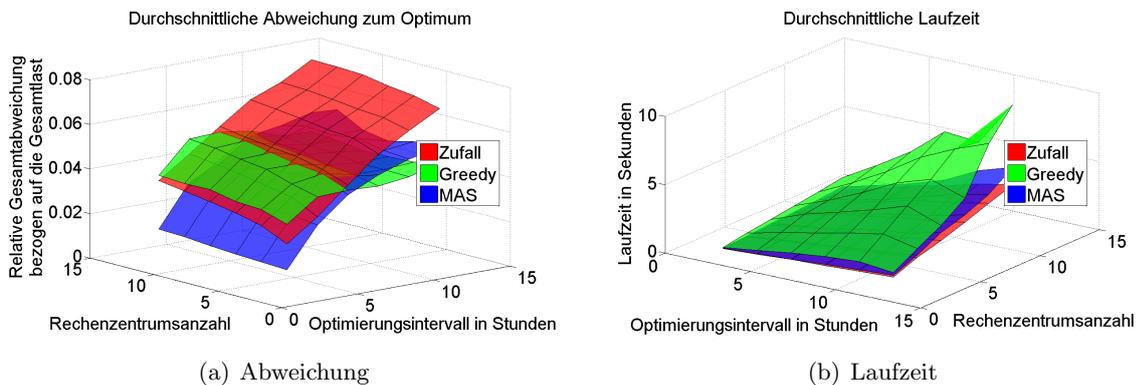


Abbildung 4.32: Evaluation des Verfahrens zur Migrationsentscheidung unter Vorgabe einer erreichbaren Lösung

Zusätzlich zum entwickelten MAS wurden eine zufällige Verteilung und eine Greedy-Methode evaluiert. Bei der zufälligen Verteilung werden die VMs zufällig auf die Rechenzentren verteilt, soweit die betreffende VM den Abstand zur Lastgangvorgabe reduziert. Trifft diese Einschränkung für kein Rechenzentrum zu, wird ein zufälliges – nur diesmal ohne Einschränkung –

gewählt. Bei der Greedy-Methode werden die VMs iterativ dem Rechenzentrum zugewiesen, bei dem sich der Abstand zur Lastgangvorgabe am meisten reduziert bzw. dieser am wenigsten anwächst. Es handelt sich um ein zentrales Verfahren.

Die in Abbildung 4.32(a) dargestellten Werte für die Abweichung zur Lastgangvorgabe sind relative Angaben bezogen auf die über das gesamte Intervall summierte Last. Die Greedy-Methode liefert bei großen Intervallen die besten Lösungen, bei kleinen jedoch die schlechtesten. Hier schneidet das entwickelte MAS am besten ab. Gegenüber der zufälligen Verteilung liefert das MAS in allen Bereichen bessere Lösungen. Der Fehler wächst mit der Anzahl an Rechenzentren linear, allerdings nur marginal. Mit größerem Optimierungsintervall steigt der Fehler beim MAS unterlinear.

Das entwickelte Verfahren vereint in gewisser Hinsicht die Eigenschaften der zufälligen und der Greedy-Methode. Ähnliches konnte auch bereits bei der Evaluation der Lastganggenerierung beobachtet werden. Dort konnte durch die Vereinigung ein grundsätzlich besseres Ergebnis erzielt werden, indem ein Optimierungsprozess mittels lokaler Suche angeschlossen wurde. Eine solche anschließende Suche würde jedoch im Kontext der globalen Migrationen das Verfahren zentralisieren oder zu Blockierungen bei den Agenten führen, was höhere Laufzeiten bedeuten würde.

Die in Abbildung 4.32(b) dargestellten Laufzeiten der Verfahren – zur besseren Sichtbarkeit sind hier die Achsen verändert – zeigen eine lineare Abhängigkeit von der Rechenzentrumsanzahl und von der Größe des Optimierungsintervalls auf. Für das MAS sind die durchschnittlichen Laufzeiten für einen einzelnen Agenten angegeben. Der Overhead für die Kommunikation ist nicht mit eingeschlossen. Die Gesamtlaufzeit entspricht somit der Multiplikation mit der Rechenzentrumsanzahl und ist damit deutlich höher als die der anderen Verfahren. Im nachfolgenden Abschnitt erfolgt eine detailliertere Analyse der Laufzeitabhängigkeiten.

Die Laufzeit der zufälligen Verteilung ist erwartungsgemäß am geringsten. In der Abbildung mit der höchsten Laufzeit dargestellt, ist die Greedy-Methode. Zwar ist ihre Laufzeit geringer als die Gesamtlaufzeit des MAS, jedoch lässt sich die Greedy-Methode in dieser Form nicht sinnvoll verteilt berechnen, ohne wiederum ihre Gesamtlaufzeit nennenswert zu erhöhen.

4.7.5 Laufzeitkomplexität

Entsprechend den formulierten Zielen sollte das Verfahren zum globalen LPM zum einen verteilt ausführbar sein und zum anderen maximal eine polynomielle Laufzeit in Abhängigkeit von den im Verbund relevanten Skalierungsparametern VM-Anzahl ($\#VM$) und Rechenzentrumsanzahl ($\#DC$) aufweisen.

In der vorgestellten Multi-Agenten-Architektur erfolgt die Ermittlung der konkreten Migrationen durch die lokalen Agenten und ist damit verteilt realisierbar. Jeder Agent führt für sich als Hauptaufgabe die Lastganggenerierung aus, welche in Abschnitt 4.5.3 bzgl. der Laufzeit analysiert wurde. Hier ergab sich eine lineare Laufzeitabhängigkeit von der VM-Anzahl des globalen Pools, was durch die lokale Suche begründet ist. Die Laufzeit der Aufbauphase besitzt eine lineare Abhängigkeit von der vorgegebenen Lasthöhe. Im vorliegenden Szenario ergeben

sich die Lasten allerdings allein durch die im Verbund betriebenen VMs, weshalb also auch die Laufzeit der Aufbauphase linear abhängig von der VM-Anzahl ist.

Bei der Selektion von VMs durch mehrere Agenten muss die Lastganggenerierung mehrfach ausgeführt werden, wobei die Anzahl der zu betrachteten VMs aufgrund der Zuweisungen mindestens linear zurückgeht. Der Worst-Case tritt dann ein, wenn alle Rechenzentren die gleiche Last vorgegeben bekommen, weil sie z. B. in ihrer Effizienz identisch sind. Die Anzahl der zu selektierenden VMs müsste hierzu gleich dem Quotienten aus VM-Anzahl durch Rechenzentrumsanzahl sein, was also einer gleichmäßigen Verteilung entspricht. Zusätzlich müssten die Agenten die gleichen VMs selektieren und die Migrationszuweisungen über alle Rechenzentren gleich verteilt sein. Wenn dieses in jedem Durchgang geschieht, ist die Anzahl an Durchgängen gleich der Anzahl an Rechenzentren. Da in diesem Fall die VM-Anzahl je Durchgang linear abnimmt, ergibt sich für jeden Agenten die Laufzeitfunktion $f^{Ag} = \frac{1}{2} \cdot (2 \cdot \#VM) \cdot \#DC = \#VM \cdot \#DC$.

In der durchgeführten Evaluation wird mit Erhöhung der Rechenzentrumsanzahl auch die VM-Anzahl linear erhöht. Aus der zu beobachtenden Laufzeitveränderung lässt sich die (Worst-Case) Laufzeitfunktion f^{Ag} allerdings nicht erkennen. In der Praxis liegt vielmehr ein linearer Zusammenhang zu einem der beiden Werte $\#VM$ und $\#DC$ vor. In Beobachtungen zum Lösungsprozess ging hervor, dass der erste Durchgang länger dauert als alle Folgedurchgänge zusammen. In der Praxis gilt damit für den gesamten Selektionsprozess die gleiche Laufzeitabhängigkeit wie für einen einzelnen Durchgang, nämlich eine lineare Abhängigkeit von der VM-Anzahl.

Der Arbitrer berechnet die theoretisch optimalen Lasten je Standort. Dazu werden als erstes die Rechenzentren nach ihrer Energieeffizienz bzw. Kosteneffizienz bewertet und sortiert. Der Bewertungsschritt skaliert linear mit der Anzahl an Rechenzentren. Er kann vollständig durch die jeweiligen Rechenzentrumsagenten übernommen werden, was dort jeweils eine konstante Laufzeit erfordert. Der Sortierprozess hat die Laufzeitfunktion $f_S^{Ar} = \#DC \cdot \log \#DC$. Durch Parallelisierung lässt sich die Laufzeit zwar reduzieren, jedoch sollte ein anderer Mechanismus zur Verteilung der Rechenlast genutzt werden. Aufgrund der Tatsache, dass die Funktion des Arbiters für mehrere Zeitpunkte unabhängig voneinander ausgeführt wird, bietet es sich an, die Bearbeitung der einzelnen Zeitpunkte grundlegend auf die Agenten zu verteilen. Hierdurch teilt sich die Laufzeit des Arbiters auf die Agenten auf, wovon insbesondere die Ermittlung der konkreten Lastgangvorgabe profitiert. Bei diesem Vorgang werden VMs schrittweise, simulativ auf die Rechenzentren alloziert ($f_L^{Ar} = \#VM$). Dies lässt sich nur sequentiell ausführen. Der Arbitrer hat damit insgesamt die Laufzeitfunktion $f^{Ar} = \#DC + \#DC \cdot \log \#DC + \#VM$.

Abbildung 4.33 gibt eine Übersicht über die Laufzeitabhängigkeiten im Verfahren. Die Funktion des Arbiters lässt sich komplett auf die lokalen Agenten aufteilen, was durch die grauen Bereiche beim Agenten verdeutlicht ist. Die Laufzeitkomplexität jedes einzelnen Agenten dominiert jedoch die des Arbiters, denn grundsätzlich gilt, dass die Anzahl an VMs größer ist als die an Rechenzentren. Durch Übernahme der Arbitrerfunktion ändert sich somit nichts an der Laufzeitkomplexität der Agenten: $f^{Ag} + f^{Ar} \in O(\#VM \cdot \#DC)$. Die Laufzeitkomplexität

des gesamten Verfahrens ist damit $O(\#VM \cdot \#DC^2)$, womit die Zielsetzung zur Skalierbarkeit erfüllt ist.

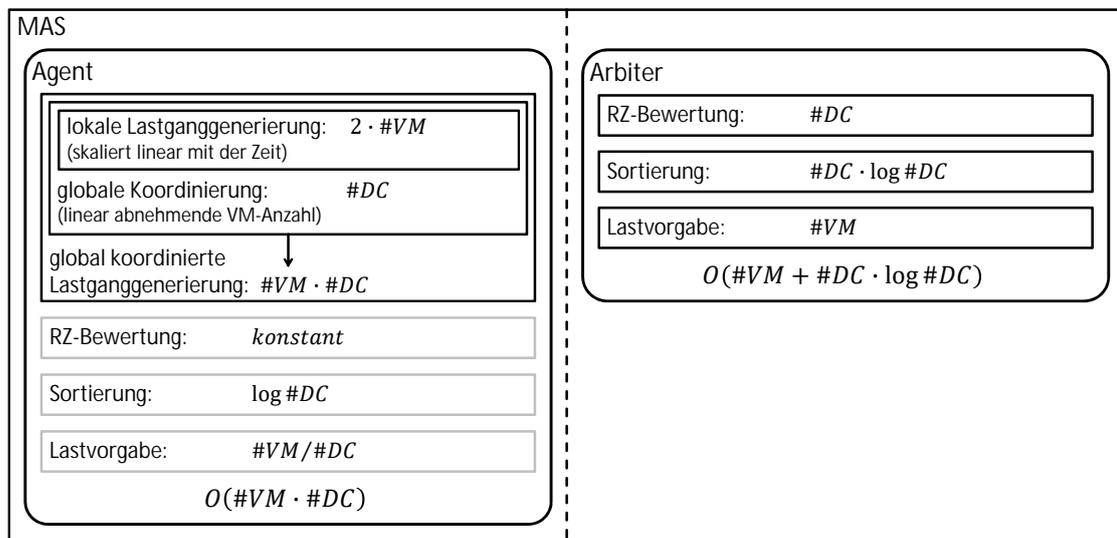


Abbildung 4.33: Laufzeitkomplexität der Agenten – die Funktion des Arbiters kann durch die lokalen Agenten übernommen werden (grauer Bereich).

4.8 Replikation – Standortselektion

Beim vorgestellten globalen Lastmanagement können Migrationen bisher zu jedem Standort, also ohne Einschränkungen durchgeführt werden, da davon ausgegangen wird, dass die notwendige Datenbasis vollständig repliziert wird. Das globale LPM hat damit die größtmögliche Flexibilität und kann die Betriebskosten, welche sich aufgrund des Betriebs der VMs ergeben und durch das vorgestellte Rechenzentrumsenergiebedarfsmodell abgebildet sind, auf maximale Weise optimieren. Allerdings verursacht eine vollständige Replikation der Daten einen großen Energie- und Kostenaufwand, welcher die Einsparungen vermindern bzw. überkompensieren kann. Je weniger Replikationsstandorte betrieben werden, desto geringer sind zwar die Replikationskosten, jedoch reduziert dies auch die Flexibilität für das globale LPM, so dass im Gegenzug die möglichen Einsparungen bei den Betriebskosten sinken können. Es besteht somit ein Trade-Off zwischen Betriebs- und Replikationskosten.

Zur gemeinsamen Optimierung dieser Kosten müssen sie bereits in der Planungsphase bestimmbar bzw. abschätzbar sein. Die Betriebskosten lassen sich durch Simulation und Auswertung der Simulationsdaten unter Verwendung des Rechenzentrumsenergiebedarfsmodells ermitteln. Dies setzt allerdings voraus, dass die Standortauswahl für die einzelnen VMs feststeht, da das globale LPM keine Erwägungsentscheidungen bzgl. möglicher Replikationsstandorte fällen soll. Das System bleibt auf diese Weise modular, d.h., dass das globale LPM – unter Einbezug der vorherrschenden Bedingungen – lediglich zur Optimierung der globalen Betriebskosten dient.

Auch die Replikationskosten lassen sich mittels Simulationen ermitteln. Eine Abschätzung

kann zwar auch ohne Durchführung einer Simulation gegeben werden, wenn sich innerhalb einer Simulation die Replikationsstandorte nicht ändern, aber wie sich in der später vorgestellten Kostenformel zeigt, geht der Standort, an dem aktuell eine VM betrieben wird, gesondert ein. An diesem Betriebsstandort befindet sich nämlich das Master-Replikat, das einzige Replikat, an dem Lese- und Schreibzugriffe stattfinden. Da dieser Standort durch das LPM bestimmt wird, ist für eine genauere Energie- und Kostenangabe auch hier die Durchführung einer Simulation notwendig.

Die in diesem Kontext zur Ermittlung der Kosten durchgeführten Simulationen beinhalten nicht die Ausführung der lokalen LPMs, sondern nur die des globalen LPM. Die Energiebedarfe und Kosten pro Standort werden mit Hilfe der entwickelten Energiebedarfsmodelle und des LPM-Verhaltensmodells berechnet.

Die Granularität der betrachteten Daten zur Replikation entspricht dem Datensatz zum Betrieb einer VM. Eine feinere Granularität würde bedeuten, den Datensatz unterschiedlicher VMs und damit evtl. auch unterschiedlicher Endverbraucher in einem gemeinsam betriebenen Container (siehe auch [33]) zu vermengen. Die Isolation der Daten wäre damit nicht mehr gegeben. Eine gröbere Granularität würde hingegen die Flexibilität bei der Auswahl an Standorten einschränken. Bei dem hier betrachteten Problem der Standortselektion gilt es also, für jede einzelne VM zu entscheiden, an welchen Standorten entsprechende Replikate platziert werden sollten. Unter einem Replikat wird somit fortan ein Tupel von VM-Datenbasis und Standort verstanden.

Um nun die minimalen Gesamtkosten zu ermitteln, müssten Simulationen mit allen möglichen Replikationsverteilungen durchgeführt werden. Die Anzahl dieser Verteilungen $\#dist_{all}$ ist abhängig von der VM-spezifischen Anzahl der Replikate und den Platzierungen an konkreten Standorten. Wie in Gleichung (4.17) angegeben, kann dazu die Datenbasis zu jeder VM an einem bis maximal der Gesamtstandortanzahl $\#DC$ vorgehalten werden. Da dies für jede VM gilt, ergibt sich in Kombination ein exponentiell großer Such- und Lösungsraum bzgl. der VM-Anzahl:

$$\#dist_{all} = \left[\sum_{k=1}^{\#DC} \binom{\#DC}{k} \right]^{\#VM} \quad (4.17)$$

Für eine vollständige Auswertung ist dieser Suchraum zu groß. Daher werden im folgenden Abschnitt verwandte Arbeiten bzgl. ihrer Ansätze und Heuristiken sowie ihre evtl. Anwendbarkeit auf das vorliegende Problem beleuchtet. Im Anschluss wird die Konzeption des entwickelten Standortselektionsverfahrens vorgestellt.

4.8.1 Verwandte Arbeiten

Verfahren zur Selektion von geeigneten Replikationsstandorten kommen beispielsweise im Kontext von Peer-to-Peer und Content Delivery Netzwerken sowie beim Grid Computing zum Einsatz. In diesen Anwendungsgebieten werden die Systeme auf Verfügbarkeit und Datenverarbeitungsdurchsatz bzw. Antwortzeiten optimiert, was mit Hilfe von Analysen der

Nutzungsstatistiken bewerkstelligt wird. Eine Berücksichtigung der Energiekosten erfolgt in diesem Kontext bisher nicht, was sich auch in den nachfolgenden Ausführungen zeigt.

Eine Ausnahme stellen MANETs (mobile ad hoc networks) dar, bei denen Energiekosten explizit betrachtet werden. Jedoch bezieht sich dies darauf, Replikationen über mehrere mobile Geräte dahingehend zu verteilen, dass eine für die begrenzte Energiemenge aller Geräte möglichst lange Laufzeit erzielt wird. Im Gegensatz zu festen Replikationsstandorten können sich hier die Geräte spontan dem Netzwerk anschließen, sich bewegen und sich davon wieder entkoppeln. Die Problemstellung unterscheidet sich damit zu sehr von der vorliegenden. Dem interessierten Leser sei eine Übersicht zu diesem Themenkomplex sowie eine Klassifizierung über die in MANETs eingesetzten Verfahren in [171, 71] nahegelegt.

Abschätzung der zukünftigen Zugriffshäufigkeit

Grundlegend ist die historische Anfragestatistik ein guter Indikator für Anfragen in der nahen Zukunft, sowohl für statische als auch dynamische Inhalte [206, 217]. In [54] werden die vergangenen Zugriffe gewichtet betrachtet: Je jünger ein Zugriff ist, desto höher ist die jeweilige Gewichtung. Gegenüber der Least-Frequently-Used-Strategie (LFU) [99] wird auf diese Weise die Netzwerkbelastung reduziert, die durchschnittliche Job-Ausführungszeit steigt jedoch.

In [49] wird die zu erwartende Zugriffshäufigkeit eines Replikats an einem bestimmten Standort statistisch ermittelt. Es wird dabei ein Zusammenhang beim Zugriff auf unterschiedliche Dateien zugrunde gelegt. Sowohl die Verwendung einer *Binomialverteilung* als auch einer *Normalverteilung* führten bei simulierten, zufälligen Anfragemustern – den Autoren nach – zu sehr guten Vorhersagewerten, wenngleich dieses aus dem beispielhaft angegebenen Ergebnisgraphen nicht hervorgeht. Eine Metrik, die die Güte quantifiziert, wurde nicht angegeben. Gleichwohl lässt sich diese Vorhersagefunktion sinnvoll für eine Replikationsersatzungsstrategie nutzen. In einer in [25] veröffentlichten Evaluation für ein konkretes Grid-Netzwerk (EU DataGrid Test-Bed 1) konnten mindestens genauso gute Ergebnisse erzielt werden wie mit den konventionellen Strategien Least-Recently-Used (LRU) [99] und LFU. Die Bildung eines Zusammenhangs von Replikaten unterschiedlicher VMs ist im vorliegenden Problem jedoch nicht zielführend, da in diesem Fall VM-Gruppen gebildet werden müssten, was eine Verringerung der Flexibilität zur Folge hätte.

In [50, 45] wird ein Marktplatzsystem, das die Kosten für den Datenzugriff minimiert, vorgestellt und formal analysiert. Als Vorhersagefunktion für die zukünftige Zugriffshäufigkeit und damit zur Bewertung eines Replikats wird u.a. die oben erwähnte Binomialverteilungsfunktion verwendet. Die in [24] durchgeführte Evaluation zeigt allerdings, dass das System nicht grundsätzlich bessere Ergebnisse liefert als eine simple LRU-Strategie, denn die Güte hängt entscheidend vom Anfragemuster ab. Ein noch schlechteres Resultat ergibt die Evaluation in [135]. Hier liefern LFU und LRU grundsätzlich bessere Ergebnisse bzgl. der Job-Ausführungszeit und Datenverfügbarkeit. Diesen überlegen ist jedoch eine Vorhersagefunktion, die auch die Größe der Datei und die aktuelle Anzahl an Replikaten berücksichtigt, denn diese Parameter

beeinflussen die zu optimierenden Werte.

Bisher hat sich die LRU-Strategie als sehr konkurrenzfähig herausgestellt, trotz der vergleichsweise simplen Idee. In [120] wird diese u.a. drei weiteren einfachen, klassischen Replikationsstrategien gegenübergestellt. Folgende Rangliste für die Güte bzgl. der benötigten Netzwerkbandbreite und der Zugriffszeit wurde ermittelt (beste Strategie zuerst):

1. Wahrscheinlichkeit für ein Replikat, ersetzt zu werden, ist proportional zur Quadratwurzel der Zugriffshäufigkeit,
2. Wahrscheinlichkeit ist proportional zur Zugriffshäufigkeit,
3. Least-Recently-Used,
4. Wahrscheinlichkeit ist immer gleich, was zur zufallsbedingten Auswahl führt.

Zusammenfassend sollten bei der Bewertung eines Replikats eine proportionale Abhängigkeit von der Zugriffshäufigkeit bestehen sowie sämtliche Parameter berücksichtigt werden, die den zu optimierenden Wert beeinflussen.

Daten-/Replikationskosten

Bei den bisher vorgestellten Verfahren handelt es sich um dynamische Replikationsverfahren. Das heißt, dass Replikate dynamisch neu erstellt und gelöscht bzw. ersetzt werden. Kosten für diese Operationen wurden allerdings nicht behandelt.

In [132] wird ein Kostenmodell für die Erstellung und das Betreiben von Replikaten herangezogen, das die Lese- und Schreibstatistik der Datei sowie die Distanzen zwischen den Replikaten berücksichtigt. Die Storagekapazität wird allerdings komplett genutzt, was bedeutet, dass das Kostenmodell in erster Linie nicht zur Kostenreduktion genutzt wird, sondern vielmehr als Metrik zur Reduzierung der Zugriffszeit dient.

In [63] geht es um die Allokation von Daten auf Servern, die topologisch in einem hierarchischen Baum angeordnet sind. Es handelt sich hierbei nicht um Replikation. Allerdings wird ein Kostenmodell für die Datenhaltung und die Kommunikation formalisiert, in dem der Speicherort und das Netzwerk als Kostenfaktoren identifiziert werden.

GlobeDB ist ein Verfahren, das eine partielle Replikation von Web-Dokumenten auf Servern durchführt [206]. Die Anfragebearbeitung wird sowohl ggü. keiner als auch voller Replikation beschleunigt – letzteres aufgrund geringerer Belastung der Server durch Datensynchronisierungen. In die Funktion zur Bewertung von Replikaten fließt die Lese- und Schreibverzögerung sowie die Netzwerkbandbreite mit ein. Im betrachteten System der Dissertation wird durch die Replikationsoperationen die Rechenleistung der operativen Server nicht beeinflusst, da das Storage ein dediziertes System darstellt.

In [219] werden ebenso Daten auf Servern repliziert. Dabei werden die Kosten für die (Neu-) Erstellung und das Aktualisieren von Replikaten eingebunden – in der Evaluation aber rein formell, also ohne authentische Werte. Ziel der dort vorgestellten Heuristiken ist

die Minimierung der Kosten unter Einbehaltung von Quality-of-Service (QoS). Dazu wird eine Lösung aufgebaut, indem entweder inkrementell zusätzliche Replikate hinzugefügt bzw. ersetzt werden, die den höchsten Zuwachs an QoS pro Replikationskosten bedeuten, bis die notwendige QoS erreicht ist, oder aber, ausgehend von einer vollständigen Replikation, analog Replikationen entfernt werden, soweit die QoS erfüllt bleibt. Die Laufzeitkomplexität ist hierbei polynomiell. Die Ergebnisse liegen im Bereich von bis zu 30 % über einem sogenannten Super-Optimum, welches in diesem Fall eine untere Schranke darstellt. Die Abweichung zum realen Optimum ist damit geringer.

In [239] wird ein Cloud Storage betrachtet, das aus mehreren Knoten besteht. Es wird ein Modell zwischen der Verfügbarkeit und der Anzahl an Replikaten durch eine asymptotische Kurve gebildet – je höher die Anzahl, desto höher die Verfügbarkeit –, deren Steigung sich durch den Parameter Ausfallwahrscheinlichkeit eines Knotens ändert. Unter Berücksichtigung der Storagekapazitäten und der variierenden Lasten werden Replikate an unterschiedlichen Standorten platziert. Ein Ziel ist die Minimierung der Kosten durch Reduktion der Anzahl an Replikaten unter Einbehaltung von Verfügbarkeitsanforderungen. Die Replikatanzahl wird durch eine Formel ermittelt, welche die Ausfallwahrscheinlichkeit der einzelnen Knoten einbezieht. Ein weiteres Ziel besteht in der Anpassung der Replikatanzahl zur Erhöhung der Gesamtrechenleistung durch Lastbalancierung (siehe auch [111, 141, 248]).

In [33] werden – wie in der Dissertation – die Daten von VMs repliziert. Die Kostenminimierung erfolgt durch Anwendung von Deduplikation. Das Verfahren versucht hierzu VM-Gruppen auf Basis ähnlicher Daten zu bilden. Das LPM migriert die jeweils aktive VM aufgrund von Veränderungen der Betriebskosten im Zeitintervall von Stunden, für die Auswahl der Replikate bleiben diese Kosten jedoch unberücksichtigt. Die Größenordnung für die Reallokation der Replikate liegt bei Monaten. Eine Einschränkung bzgl. der minimalen oder maximalen Anzahl an Replikaten gibt es nicht, sondern sie ist allein von der Storagekapazität begrenzt. Daher werden hier zur Simulation zufällige Werte angenommen. Kosten für die Live-Migration und den Betrieb der Replikate werden nicht berücksichtigt. Wie oben bereits beschrieben, wird eine Zusammenführung von Storage unterschiedlicher VMs aufgrund geringerer Flexibilität, erhöhter Replikationskomplexität und aufgrund der Verletzung des Isolationskriteriums in der Dissertation ausgeschlossen.

Insgesamt wurden als Kostenfaktoren die Zugriffsstatistik, der Speicherort und das Netzwerk genannt sowie Verfahrenskonzepte zur Erreichung der gesteckten Ziele vorgestellt, auf die in den folgenden Abschnitten Bezug genommen wird.

Klassifikation

In den meisten bisher aufgeführten Verfahren werden Replikate dynamisch erstellt. Beim letztgenannten Verfahren ([33]) wird in bestimmten Zeitabständen eine Replatzierung durchgeführt. Je größer das Intervall ist, desto schlechter kann auf sich veränderte Gegebenheiten reagiert und damit optimiert werden. Dieser Ansatz wird zur Abgrenzung als statisch bezeichnet. In [183] wurde festgestellt, dass die dynamische Replikation, bei der eine Replatzierung jederzeit

erfolgen kann, eine Verbesserung bis zum 3-Fachen ggü. der statischen Replikation erreicht. Das in [217] vorgestellte dynamische Verfahren konnte bezogen auf die Antwortzeit und den Datendurchsatz jedoch maximal eine 2-fache Verbesserung erzielen. In den meisten Fällen liegt der Unterschied sogar nur im unteren 2-stelligen Prozentbereich. Im Rahmen der Dissertation soll für einen ersten Schritt der statische Ansatz genügen.

Neben dem Berechnungszeitpunkt für Replatzierungen (statisch, dynamisch) wird in [78] eine weitere Klassifizierungseigenschaft vorgestellt: zentrale oder dezentrale Berechnung. Die zentrale Berechnung hat ggü. der dezentralen eine geringere Komplexität, stellt allerdings einen Single-Point-of-Failure dar. Bei der dezentralen Berechnung wird eine höhere Fehlertoleranz und eine i.d.R. höhere Skalierbarkeit erzielt, die Komplexität steigt jedoch. Um diese zu reduzieren, werden zumeist nur partielle Informationen lokal genutzt, wodurch allerdings suboptimale Lösungen generiert werden. Da das globale LPM, wie in Abschnitt 4.7.5 beschrieben, verteilt ausführbar ist, können automatisch auch die einzelnen Simulationen innerhalb des Replikationsverfahrens von der erhöhten Skalierbarkeit profitieren, ohne auf Informationen zu verzichten. Das eigentliche Replikationsverfahren sollte allerdings auch die Möglichkeit haben, verteilt ausführbar zu sein, oder zumindest parallelisierbar sein.

In der Regel wird eine Replikationsstrategie auf das gesamte System angewendet. In [176] konnte jedoch festgestellt werden, dass der Einsatz unterschiedlicher Caching- bzw. Replikationsstrategien je Datei bessere Ergebnisse erzielt. Durch Simulationen wird die jeweils beste Strategie ermittelt und auch dynamisch optimiert. In der dort durchgeführten Evaluation werden allerdings nur zwei Dateien behandelt und es wird auf Skalierungsprobleme bei diesem Vorgehen hingewiesen. Daher wurden in [207] Heuristiken vorgestellt, die nur unter bestimmten Bedingungen Aktualisierungsanalysen durchführen. Dabei wird ein Trade-Off zwischen Berechnungsbeschleunigung und Güte eingegangen. In der Dissertation wird ein globales Ziel anvisiert: Es geht nicht darum, die Verfügbarkeit oder Bearbeitung einzelner Dateien zu verbessern, sondern die Kosten der Gesamtheit zu minimieren.

In [46, 186] werden unterschiedliche einfache Scheduling- und dynamische Replikationsalgorithmen evaluiert und miteinander kombiniert. Es werden u.a. LFU, LRU und die Binomialverteilungsfunktion aus [49] (siehe oben) betrachtet. Eine den anderen überlegene Methode ist zwar nicht feststellbar, allerdings wurden die besten Ergebnisse erzielt, wenn beim Scheduling sowohl der aktuelle Workload der Berechnungsressource als auch die Kosten für den Datenzugriff berücksichtigt werden. Eine Kopplung von Scheduling- und Replikationsverfahren zu einem komplexeren Verfahren sei aber nicht notwendig.

In [142] wird angemerkt, dass sich die Verfahren zur Replikationsplatzierung in Simulationen von denen in Produktivsystemen meist dahingehend unterscheiden, dass einfache Datenmodelle angenommen (teils auch nur Lesezugriffe) und wohl ausgesuchte Zugriffsmuster untersucht werden sowie vollständige Informationen über die Infrastruktur vorliegen. Außerdem wird dort ein Klassifikationsschema vorgestellt, das Replikationsverfahren nach den folgenden Eigenschaften unterteilt, die sich zum Teil mit den zuvor beschriebenen decken. Jede dieser Eigenschaften wird auf das in dieser Dissertation betrachtete Problem hin diskutiert.

- **Datenmodell:** Im hier betrachteten System wird nicht zwischen einzelnen Dateien unterschieden, sondern die Gesamtheit aller Dateien, die eine VM verarbeitet, als eine Einheit betrachtet – nach der Klassifizierung von [229] als Container bezeichnet. Eine Modifizierung dieses Datensatzes wird mit einem gewissen Schreibzugriffsfaktor berücksichtigt. Da eine VM jeweils nur an einem Standort gleichzeitig aktiv ist, können Dateien auch nur von einem Standort aus verändert werden, wenngleich sich dieser dynamisch ändern kann. Der Lesezugriff wird nicht gesondert behandelt.
- **Ressourcenmodell:** Durch den Einsatz der lokalen Abstraktionsschicht wird Transparenz geschaffen, die eine Unwissenheit auf globaler Ebene bzgl. der konkreten, lokal vorhandenen Ressourcen ermöglicht. Bekannt jedoch ist, dass jeder Standort eine gewisse Storagekapazität sowie eine bestimmte Anzahl an Servern und damit Rechenleistung besitzt, die jeweils je Standort variieren können. Eine Unterscheidung zwischen den einzelnen Standorten bzgl. ihrer Funktionalität findet nicht statt. Dies wird nach [142] als Single-Tier Topologie angesehen.
- **Optimierungsziel und -methode:** Das Ziel der Optimierung ist die Minimierung der Kosten, und zwar nicht die der einzelnen Dienste, sondern die des Betriebs des gesamten Rechenzentrumsverbundes.
- **Verfahrensumsetzung:** Das globale LPM geht von einer festen Verteilung der Replikate auf Standorte aus, zwischen denen es Migrationen durchführen kann. Das Replikationsverfahren soll in verteilter Manier und der vorliegenden Situation dynamisch angepasst agieren. Dieses soll jedoch nicht kontinuierlich, sondern in bestimmten Intervallen geschehen. Damit löst das Verfahren abschnittsweise ein statisches Problem basierend auf den aktuell vorliegenden Informationen. Aufgrund der gewollten modularen Trennung besteht damit keine direkte Interaktion mit dem globalen LPM, jedoch wird zum einen das Verhalten des LPM in das Replikationsverfahren einbezogen und umgekehrt werden VM-Migrationen – durch die Entscheidungen des Replikationsverfahrens beschränkt – durchgeführt. Dies führt zu einer (losen) Kopplung.
- **Evaluationsmethodik:** Zur Evaluation der Verfahrensqualität werden Simulationsreihen unterschiedlicher Szenarien untersucht.

Die in [142] klassifizierten Verfahren unterscheiden sich von dem hier zu entwickelnden Verfahren in mind. einer Eigenschaft, wie in Tabelle 4.4 zu sehen ist. Hierin sind die Arbeiten mit der größten Ähnlichkeit zum zu entwickelnden Verfahren diesem gegenübergestellt. In [197] wird das Problem durch ein ILP gelöst, was gerade in föderierten Cloud-Rechenzentren, die für sich allein genommen bereits zur Klasse der Megarechenzentren zählen können, eine schlechte Skalierbarkeit verspricht. Es besteht keine Interaktion mit dem Job-Scheduler und ein möglicher Schreibzugriff wird nicht berücksichtigt. Letzteres gilt auch für das Verfahren in [217], das bereits oben erwähnt wurde. Bei diesem Verfahren handelt es sich um ein dynamisches, welches aber wie die anderen beiden Verfahren zentral ausgeführt wird. In [155]

	[197]	[217]	[155]	eigenes Verfahren
Datenmodell	statistischer Lesezugriff von unterschiedlichen Standorten aus initiiert, kein Schreibzugriff	statistischer Lesezugriff von unterschiedlichen Standorten aus initiiert, kein Schreibzugriff	Datengruppen berücksichtigt, statistischer Schreib- und Lesezugriff von unterschiedlichen Standorten aus initiiert	Datencontainer, impliziter Lesezugriff und statistischer Schreibzugriff von nur einem aktiven, sich dynamisch wechselnden Standort aus initiiert
Ressourcenmodell	Ressourcenkapazitäten sind bekannt, Single-Tier (Gleichbehandlung der Standorte)	Ressourcenkapazitäten sind bekannt, Single-Tier (Gleichbehandlung der Standorte)	Ressourcenkapazitäten sind bekannt, Multi-Tier (standortspezifische Rollenverteilung beim Datenmanagement)	Ressourcenkapazitäten sind bekannt, Single-Tier (Gleichbehandlung der Standorte)
Optimierungsziel, -methode	lokale Optimierung der globalen Zugriffszeit und „Kosten“	globale Optimierung der lokalen sowie globalen Verfügbarkeit, Zugriffszeit und „Kosten“	globale Optimierung der globalen Verfügbarkeit, Zugriffszeit und „Kosten“	globale Optimierung der globalen Kosten
Verfahrensumsetzung	dynamisch, zentral, keine Kopplung zum Scheduler	dynamisch, zentral, Kopplung zum Scheduler	dynamisch, zentral, Kopplung zum Scheduler	statisch, (de)zentral, (lose) Kopplung zum Scheduler (globales LPM)
Evaluationsmethodik	formale Berechnung, Simulation und Einsatz an realer Hardware	formale Berechnung und Simulation	Einsatz im Produktivbetrieb	Simulation

Tabelle 4.4: Gegenüberstellung von Replikationsverfahren

werden Replikate online analysiert und anhand ihrer Popularität aktualisiert. Dadurch entsteht eine enge Verzahnung mit dem Job-Scheduler. Für das Datenmanagement werden zum einen den Standorten unterschiedliche Rollen zugeteilt und zum anderen die Storagekapazitäten vollständig genutzt. Letzteres zeigt, dass – im Gegensatz zur Einordnung – keine Kosten für die Replikation betrachtet werden. Diese Missinterpretation des Optimierungsziels „Kosten“ ist auch bei den anderen Verfahren zu erkennen, denn auch bei diesen wird nicht auf finanzielle Kosten eingegangen, sondern eine definierte Kostenfunktion zur Optimierung z. B. des Datendurchsatzes herangezogen.

In [248] wird ein weiterer Anwendungsfall für Replikation vorgestellt: die Minimierung der Replikanzahl bei Kenntnis des maximalen Rechenleistungsbedarfs. Im betrachteten System ist dies dann der Fall, wenn es nur ein Original gibt, also keine weiteren Replikate. Hierbei sind die Replikationskosten minimal und dementsprechend die Betriebskosten maximal. Dieser Fall ist also das Gegenstück zur vollständigen Replikation und bedeutet für das globale LPM minimale bzw. keine Flexibilität.

4.8.2 Heuristik zur Standortselektion

Im vorangegangenen Abschnitt wurden basierend auf verwandten Arbeiten einige Ansätze für das vorliegende Problem identifiziert sowie gewünschte Eigenschaften des zu entwickelnden Verfahrens herausgestellt. Vor seiner konkreten Konzeptionierung wird herausgearbeitet, woraus sich die Replikationskosten zusammensetzen.

Replikationskosten

Bei Datenreplikationen müssen zwei Aspekte betrachtet werden: die Selektion von geeigneten Replikationsstandorten und die Replikationsdurchführung selbst. Bei Letzterem ist zuerst über den Synchronisierungstyp zu entscheiden. Während eine synchrone Replikation bei fehlertoleranten Systemen zum Einsatz kommt, um Datenkonsistenz zu gewährleisten, ist im vorliegenden System einzig die Aktualität zum Migrationszeitpunkt gefragt. Daher reicht hier eine asynchrone Replikation, die im Vergleich zur synchronen einen geringeren Overhead hat und damit effizienter und kostengünstiger ist. Die asynchrone Replikation kann periodisch in bestimmten Synchronisierungsintervallen oder aperiodisch aufgrund von Events (z. B. das Überschreiten einer gewissen Veränderungsmenge) erfolgen. Hier wird zur Vereinfachung von einer periodischen Synchronisierung ausgegangen. Von weiteren Details über Replikationsstrategien (Merge-, Transaktionsreplikation etc. [154]) und, ob die Replikation daten- oder blockbasiert [29] durchzuführen ist, wird abstrahiert.

Um zu gewährleisten, dass das globale LPM von den Replikationsaspekten isoliert agieren kann, müssen bei allen Migrationen die Daten der VMs am Zielstandort vorhanden sein. Anderenfalls würde nämlich die Migrationszeit länger dauern und das globale LPM müsste in Abhängigkeit von der Datenaktualität Migrationsentscheidungen fällen. Zur Vereinfachung wird von einer konstanten Datenveränderungsrate ausgegangen. Hierdurch kann abgeschätzt werden, wie viel Daten bei einem bestimmten Synchronisierungsintervall und auch während einer Migration zu übertragen sind.

Neben den bereits erwähnten Parametern der Standortanzahl und den Standort bedingten Strompreisen sind die Replikationskosten demnach vom Synchronisierungsintervall und der konkreten Menge an zu synchronisierenden Daten abhängig. Ferner ist die verwendete Hardware sowohl lokal als auch netzseitig zu berücksichtigen. Auf diese einzelnen Aspekte wird nun im Folgenden eingegangen.

Der Energiebedarf, welcher durch Datentransfer generiert wird, ist davon abhängig, wie viel Daten übertragen werden (siehe Netzwerkmodell in Abschnitt 4.6). Bei Datensynchronisierungen werden nur solche Daten übertragen, welche sich seit der letzten Synchronisierung verändert haben. Dies hängt von der Veränderungsrate der Daten und dem Synchronisierungsintervall ab. Ist das Intervall klein, werden weniger Daten zu aktualisieren sein als in einem großen. Jedoch kann es vorkommen, dass in einem größeren Intervall die gleichen Daten mehrfach überschrieben werden. Über den gesamten Betrachtungszeitraum müssen bei kleineren Synchronisierungsintervallen damit insgesamt mehr Daten zur Synchronisierung

4 Konzeption samt Komponentenevaluation

übertragen werden.

In einer Studie des BMWi [162] wurde ein logarithmischer Zusammenhang für den Mehraufwand bei kleineren Synchronisierungsintervallen zugrunde gelegt: Bei einer gegebenen Storagegröße $cap_{Storage,vm_i}$ einer VM vm_i und einer täglichen Gesamtaktualisierungsrate UR_{24} lässt sich die Datenmenge zur Aktualisierung innerhalb eines Synchronisierungsintervalls Δt_{synch} (in Stunden) durch die Formel in Gleichung (4.18) berechnen. Dieser Zusammenhang wird in Abbildung 4.34 verdeutlicht, in der das anteilige Datenaufkommen ggü. einer täglichen Synchronisierung dargestellt ist.

$$D_{synch}(vm_i, \Delta t_{synch}) = cap_{Storage,vm_i} \cdot UR_{24} \cdot \frac{\log(0,3295 \cdot (\Delta t_{synch} + 3,0349))}{\log(0,3295 \cdot (24 + 3,0349))} \quad (4.18)$$

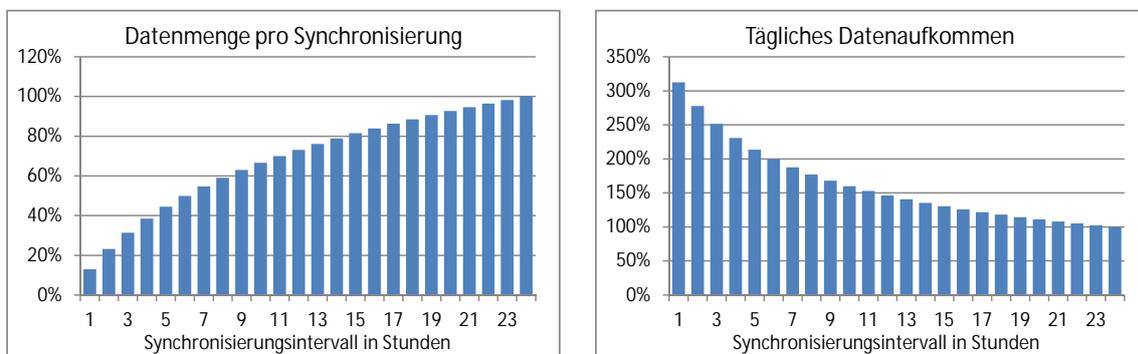


Abbildung 4.34: Durchschnittliche Datenvolumina bei Verwendung asynchroner Replikation in Relation zum Datenvolumen bei einem Synchronisierungsintervall von 24 Stunden

Hitachi Data Systems berichtet in einem ihrer Storage-Projekte von einer täglichen Update-rate von 10%.⁸ Kevin Adams hat in seinen Untersuchungen festgestellt, dass blockbasierte Systeme mit weniger als 4% täglicher Änderungsrate nur halb so viel Traffic verursachen wie filebasierte Systeme [2]. Für die in der Dissertation durchgeführten Untersuchungen wird $UR_{24} = 5\%$ gesetzt.

Um die Replikationskosten zu minimieren, muss ein möglichst langes Synchronisierungsintervall gewählt werden. Dieses darf aber, wie bereits erwähnt, nicht zur Erhöhung der Zeit für eine globale Migration führen. Diese ist mit 60 Sekunden angesetzt, damit sie der lokalen Migrationszeit entspricht (siehe Abschnitt 5.1). In Abschnitt 2.1.2 wurde gezeigt, dass die Übertragung einer VM mit 4 GB RAM über eine 10 Gbit/s-Leitung 13s dauert. Die übrigen 47s können dazu genutzt werden, um die Daten aus dem Festplattenspeicher nachzusynchronisieren.

Der Festplattenspeicher, auf den eine VM zurückgreifen kann, liegt bei maximal 410 GB (siehe Abschnitt 5.1). In [16] wird angegeben, dass der zur Verfügung stehende Festplattenspeicherplatz durchschnittlich nur zu 30% - 35% genutzt wird. Bei einer Speicherplatzbelegung

⁸www.hds.com/assets/pdf/electro-motive-diesel-success-story.pdf, letzter Zugriff: 28.08.2014

von 35 % werden somit durchschnittlich nur 143 GB genutzt. Innerhalb der Migrationszeit liegt nach der angegebenen Formel (4.18) der zu synchronisierende Datensatz bei 90 MB, was in weniger als einer Sekunde übertragbar ist. Die noch verbleibende Zeit wird hiermit auf 46 s reduziert.

Voraussetzung hierfür ist jedoch, dass der Datensatz bis zu diesem Zeitpunkt aktuell war, was gleichbedeutend damit ist, dass die Replikationssynchronisierungen direkt vor den globalen Migrationen durchgeführt werden müssten. Dies ist automatisch dann der Fall, wenn das globale Migrationsintervall in Vielfachen größer oder gleich dem Synchronisierungsintervall ist. Sollte das Synchronisierungsintervall aber größer sein, um Replikationskosten zu sparen, müssen die bis zur globalen Migration vorgekommenen Veränderungen nachsynchronisiert werden. Geht man davon aus, dass in der Praxis die Nennbandbreite zu 50 % erreicht wird, können in 46 s noch 28,75 GB übertragen werden. Der Formel (4.18) nach wird diese Datenmenge bei etwa 14,5 Std. erreicht. Um das globale LPM in seinen Entscheidungen nicht einzuschränken, sollte die letzte Synchronisierung vor den globalen Migrationen damit nicht vor 14,5 Std. erfolgt sein. Im Folgenden wird jedoch zur Vereinfachung von identischen Intervallgrößen ausgegangen.

Mit WAN-Optimierungssystemen wie die Riverbed Steelhead Appliance können die Daten speziell zur WAN-weiten VM-Synchronisierung komprimiert werden. Hierdurch würden Übertragungsbeschleunigungen bis zum 14-Fachen erreicht werden [21]. Für die Berechnungsanalysen wird ein geringerer Wert (10-fach) angenommen, was einer 10-fachen Kompression entspricht ($cr = 0,1$).

Die zu übertragenden Daten verursachen einen Mehraufwand nicht nur im Netzwerk, sondern auch in den Speichersystemen, welche die Aktualisierungen verarbeiten müssen. In [117] wird der Energiebedarf für den Datenzugriff auf Festplatten von gewöhnlichen Servern analysiert. Hieraus kann ein durchschnittlicher Energiebedarfswert zur Verarbeitung von Daten von aufgerundet $\gamma_{Storage} = 0,0003$ kWh/GB abgeleitet werden. SAN-Systeme, wie sie hier für das lokale LPM vorausgesetzt werden, sind auf die Verarbeitung von Storageanfragen spezialisiert und weisen damit einen proportional geringeren Energiebedarf auf. Die in der Veröffentlichung präsentierten Messwerte stellen damit eine Überabschätzung dar.

Die Kosten für die Synchronisierung einer VM vm_i , welche aktuell am Standort j betrieben wird und an n Standorten repliziert ist, lässt sich schließlich durch die folgende Kostenformel ermitteln, die aus einem Netzwerk- und einem Storageanteil besteht:

$$\begin{aligned}
 costs_{synch}(vm_i, t_k) &= costs_{Network}(vm_i, t_k) + costs_{Storage}(vm_i, t_k) \\
 &= D_{synch}(\Delta t_{synch}, vm_i) \cdot cr \cdot \gamma_{Network} \cdot \frac{1}{2} \sum_{r=1}^n (\delta_{j,t_k} + \delta_{r,t_k}) \\
 &\quad + D_{synch}(\Delta t_{synch}, vm_i) \cdot \gamma_{Storage} \cdot \left(\delta_{j,t_k} + \sum_{r=1}^n \delta_{r,t_k} \right) \quad (4.19)
 \end{aligned}$$

δ ist gemäß dem Netzwerkkostenmodell aus Abschnitt 4.6 der Strompreis der jeweiligen

Standorte j (Quelle) und r (Ziel/Replikat) zum betrachteten Zeitpunkt t_k und $\gamma_{Network}$ der Netzwerkenergiebedarfskoeffizient. Der benötigte Gesamtstoragebedarf einer VM soll hier zur Vereinfachung als konstant angesehen werden.

Ordnungsmetrik

Mit der in Gleichung (4.19) angegebenen Kostenformel und dem in Abschnitt 4.2 präsentierten Energiebedarfsmodell können nun die Betriebs- und Replikationskosten und somit die Gesamtkosten berechnet werden, wenn mit der entsprechenden Replikationsverteilung das globale LPM simuliert wird. Der durch die Gleichung (4.17) aufgezeigte Suchraum ist jedoch zu groß, um die kostengünstigste Verteilung zu ermitteln. In den verwandten Arbeiten wurden Replikate auf Basis ihrer Zugriffsstatistik bewertet, um ihren zukünftigen Nutzen abzuschätzen. Mit dieser Überlegung, sollen nun die Auswirkungen einer Standortselektion herausgestellt und davon eine Bewertungsmetrik abgeleitet werden.

Für die Standortselektion hat nicht nur die bloße Anzahl an Replikationsstandorten, sondern insbesondere auch die konkrete Auswahl Einfluss auf die Gesamtkosten. Wie bereits festgestellt wurde, erhöht die Anzahl der Replikationsstandorte grundsätzlich die Flexibilität des globalen LPM, welches dadurch effizienter kostenvermindernde Migrationen veranlassen kann. Je nach den konkreten Verläufen der Strompreise und Temperaturen ist der Anteil für den Einspareffekt zwischen den Standorten unterschiedlich. So ist es möglich, dass nur eine geringe Anzahl an Standorten relevant ist, um die Kosteneinsparungen zu erzielen. Um welche Standorte es sich dabei handelt, ließe sich durch Analyse der Verläufe für Temperaturen, Energiepreise sowie Ressourcenbedarfe sämtlicher VMs ermitteln. Aber nicht zuletzt hängt es davon ab, ob das globale LPM die Standorte auch nutzt: Wenn ein Standort vom globalen LPM genutzt wird, dann hat dieser Standort dazu beigetragen, die Gesamtkosten bzw. den Gesamtenergiebedarf zu minimieren.

Ein geeigneter Standort für ein Replikat ergibt sich also daraus, dass er durch das LPM genutzt wurde. Dieser Zusammenhang ist ähnlich den in den verwandten Arbeiten zugrunde gelegten Ansätzen des Nutzerzugriffs. Der Nutzer, welcher den Zugriff auf die Daten haben möchte, ist in dem hier vorliegenden Fall jedoch nicht derjenige, der den in einer VM betriebenen Dienst nutzt, sondern das LPM, das möchte, dass die VM an bestimmten Standorten betrieben wird. Wie auch die Gesamtkosten können die Zugriffsstatistiken für jeden Standort durch eine Simulation ermittelt werden. Ein Zugriff bedeutet, dass eine VM für eine bestimmte Zeit an einem konkreten Standort betrieben wurde. Die Verwendung der Verweildauer als Metrik zur Replikatbewertung entspricht, analog betrachtet, der LFU-Strategie.

Neben der bloßen Verweildauer bringen aber sowohl der Standort als auch die VM noch weitere Aspekte mit ein, die den Grad, wie ausschlaggebend der jeweilige Betrieb ist, beeinflusst. Wenn z. B. ein Standort zu einem gegebenen Zeitpunkt zwar vollständig genutzt wurde, aber ggü. einem ebenfalls vollständig genutzten Standort eine geringere Einsparung bewirkt, so ist der Wegfall dieses Standortes weniger gravierend. Besitzt der Standort diese geringere Einsparung über einen deutlich längeren Zeitraum, so kann es sein, dass sein Wegfall nun doch

eine größere Auswirkung hat. Bei der Standortauswahl ist also die Höhe der Kostenersparnis (siehe Gewichtungsfaktor in Abschnitt 4.7.2) einzubeziehen.

VMs unterscheiden sich untereinander im Ressourcenbedarfsprofil: Eine VM mit hohem Ressourcenbedarf hat an einem günstigen Standort eine höhere Auswirkung auf die Energieersparnis als eine VM mit niedrigerem Ressourcenbedarf. Und analog, wenn letztere an einem teureren Standort betrieben werden würde, verursacht sie weniger Mehrkosten als die VM mit höherem Ressourcenbedarf.

Die Zugriffsstatistik wird damit mit der Kostenersparnis des Standortes und dem Ressourcenbedarf der VM zu einem Nutzenwert use zusammengefasst. In Gleichung (4.20) wird seine Berechnung für die VM vm_i und den Standort loc_j für das diskrete Zeitintervall $[t_1; t_n]$ mit dem Elementarzeitschritt ts , der Kostenersparnis $cs(loc_j, t_k)$ und die durch den VM-Ressourcenbedarf generierte Auslastung $load(vm_i, t_k)$ angegeben. Es werden nur solche Zeitpunkte t_k berücksichtigt, in denen die VM auch am jeweiligen Standort betrieben wurde.

$$use_{vm_i, loc_j}(t_1, t_n) = \sum_{t_k=t_1}^{t_n} (ts \cdot cs(loc_j, t_k) \cdot load(vm_i, t_k)), \forall t_k : loc(vm_i, t_k) = loc_j \quad (4.20)$$

Die Funktion loc liefert den Betriebsstandort einer VM. Der Elementarzeitschritt ts ist in der gesamten Betrachtung konstant und kann komplett aus der Gleichung genommen werden, da er sich bei einem Vergleich herauskürzen lässt.

Alle Replikate lassen sich nun nach dem Nutzenwert ordnen. Hierdurch kann die Flexibilität des globalen LPM eindeutig variiert werden. Es ist dabei heuristisch davon auszugehen, dass das Replikat mit dem höchsten Nutzenwert zur höchsten Betriebskosteneinsparung führt. Mit absteigendem Nutzenwert werden die Betriebskosteneinsparungen geringer. Zur Erhöhung der Flexibilität sollten daher die Replikate nach dem Nutzenwert absteigend sortiert betrachtet werden. Die hierzu korrespondierende Liste wird im Folgenden Flexibilitätsliste genannt.

Im idealisierten Fall ist die Betriebskosteneinsparung anfangs hoch und nimmt dann asymptotisch ab. Die Kosten für die Replikationen skalieren weitgehend linear mit der Anzahl an Replikaten bzw. der Flexibilität. Durch Berücksichtigung der konkreten Standortstrompreise, könnte sich ein quadratischer Verlauf ergeben. Die minimalen Gesamtkosten liegen an dem Flexibilitätspunkt, bevor die Replikationskosten die Betriebskosteneinsparung eines zusätzlichen Replikats kompensieren. Dieser Zusammenhang ist beispielhaft in Abbildung 4.35 dargestellt.

Bei der Befüllung eines Rechenzentrums gemäß der Flexibilitätsliste ist jedoch grundsätzlich zu beachten, dass das Storage in einem Rechenzentrum eine begrenzte Kapazität hat. Wird die für die Replikate bereitgestellte Kapazität überschritten, so müssen Replikate entfernt bzw. außen vor gelassen werden. Dies sollte die Replikate mit den geringsten Nutzenwerten treffen. Zu berücksichtigen gilt es hier, dass die Daten jeder einzelnen VM mindestens an einem Standort vorgehalten werden. Um die Heterogenität der VM-Speichergröße zu berücksichtigen, kann der Nutzenwert mit der Speichergröße skaliert werden. In den hier durchgeführten Simulationen wird allerdings davon ausgegangen, dass die Storagekapazität jedes Standorts

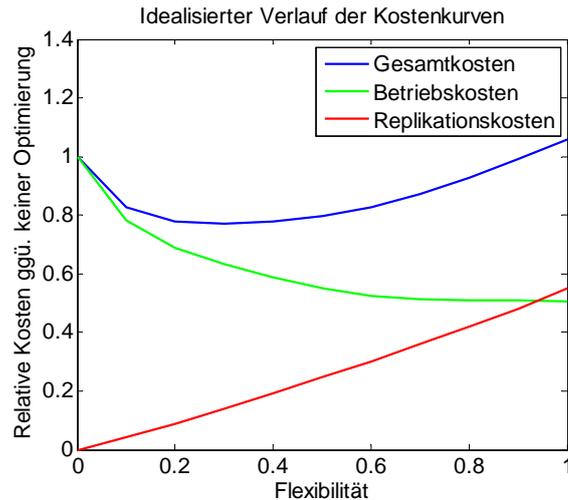


Abbildung 4.35: Idealisierte Verlaufsform der Kurven für die Gesamt-, Betriebs- und Replikationskosten

zur Aufnahme sämtlicher Replikate ausreicht.

Die Ordnungsmetrik *use* benötigt Informationen zu den Entscheidungen des LPM, welche sich erst durch eine Simulation ergeben. Werden bestimmte Annahmen zur Entscheidungsfindung getroffen, so können auch Ordnungsmetriken erstellt bzw. verwendet werden, die diese konkreten Informationen nicht benötigen:

- Im Prozess des globalen LPM werden die Live-Migrationen durch die lokalen Agenten ermittelt. Diese agieren aufgrund der Lastgangvorgaben des Arbiters, welche grundsätzlich unabhängig vom konkreten VM-Standort erstellt werden können. Wird eine VM einer leeren Lösung bei der Lastganggenerierung für einen Standort hinzugefügt, reduziert sich die Abweichung zu seiner Lastgangvorgabe. Dieser Fehlerreduzierungs Wert lässt sich als Ordnungsmetrik ERR verwenden.
- Neben den Fehlerreduzierungs Werten können mit der Lastgangvorgabe und den VM-Auslastungsprofilen auch Korrelationswerte bestimmt werden (Ordnungsmetrik COR).
- Die bisherigen Metriken gehen auf den Betrieb der VMs ein. Da die Gesamtkosten aber eben auch von den Replikationskosten abhängen, können diese auch gezielt minimiert werden. Eine solche Metrik REP ordnet die Replikate entsprechend ihren verursachenden Kosten. Hierbei handelt es sich um eine vollständige Entkopplung vom globalen LPM.

4.8.3 Evaluation der Standortselektion

Die erstellten Ordnungsmetriken sollen in diesem Abschnitt auf ihre Qualität hin evaluiert werden, wie gut mit ihnen die Gesamtkosten minimiert werden können. Hierzu sollen sie miteinander verglichen werden. Eine sinnvolle Angabe zur Abweichung vom Optimum kann

nicht gegeben werden, da hierzu das Optimum bestimmt werden müsste. Aufgrund des bereits erwähnten exponentiell großen Lösungsraumes ist seine Ermittlung nicht praktikabel.

In [219] wurde ein Super-Optimum als Schranke zur Abweichung vom Optimum berechnet, indem die Originalproblembeschreibung als ILP zu einem LP relaxiert wurde. Angewendet auf das vorliegende Problem würde dies zwar den Suchvorgang zwischen den unterschiedlichen Replikationskonstellationen beschleunigen, der Löwenanteil liegt jedoch in der Durchführung der Simulation zur Bestimmung der Betriebskosten je Konstellation. Eine geeignete Abschätzung dieser Kosten ist aber nicht vorhanden.

Es werden die folgenden Ordnungsmetriken gegenübergestellt – beim Nutzenwert werden drei Varianten untersucht:

- *Zufall* (RND) – eine zufällige Ordnung.
- *Fehlerreduzierung* (ERR) – Auswirkung für das Hinzufügen einer VM zu einer leeren Lösung bzgl. der Lastgangvorgaben.
- *Auslastungskorrelation* (COR) – die standortbezogenen Lastgangvorgaben werden mit den VM-Auslastungsdaten korreliert.
- *Replikationskosten* (REP) – Ordnung entsprechend den anfallenden Replikationskosten. Die Replikationskosten je Flexibilität sind hierbei minimal.
- *Use-Heuristik V1* (UV1) – Ordnung entsprechend der Verweildauer.
- *Use-Heuristik V2* (UV2) – wie UV1 und Berücksichtigung der Kostenersparnis.
- *Use-Heuristik V3* (UV3) – wie UV2 und Berücksichtigung der prognostizierten VM-Lasten (gemäß Gleichung (4.20)).

Zum Vergleich werden mit einer Abtastung von 0,1 unterschiedliche Flexibilitäten untersucht: 0 bedeutet, dass keine Replikat erstellt werden; die VMs können also nur am jeweils aktuellen Standort betrieben werden. 1 bedeutet eine vollständige Replikation.

In den Simulationen werden 500 VM in fünf homogenen Rechenzentren betrieben. Die freie Kapazität ist auf 50% eingestellt, d.h., dass in den Rechenzentren die doppelte Anzahl an VMs betrieben werden könnte (die gleichen VMs ein weiteres Mal). Simuliert wird ein Zeitraum von zwei Tagen. Zur Ermittlung der Zugriffsstatistiken und Ableitung der Metriken wird ein gleichlanger Zeitraum vor der simulierten Zeit analysiert.

Der Energiebedarfskoeffizient für das Netzwerk wurde auf $\gamma_{Network} = 0,01$ kWh/GB festgelegt, was einen realistischen Wert in den kommenden Jahren bis 2020 darstellt (siehe Abschnitt 5.1). Das Optimierungs- bzw. Migrationsintervall wurde auf zwei Stunden gesetzt. Die weiteren Simulationseinstellungen betreffen die VM-Mengen- und Strompreisverlaufsvariation. Hier wurden jeweils fünf unterschiedliche zufällig zusammengestellte VM-Mengen bzw. Strompreisverläufe verwendet. Mit diesen Einstellungen befinden sich die Kosten für eine vollständige Replikation in der Größenordnung der Betriebskosten, so dass sich eine

Trade-Off-Situation ergibt, wie sie im idealisierten Verlauf in Abbildung 4.35 zu sehen ist und anhand ihrer man die Auswirkungen der Metriken erkennen kann.

In Abbildung 4.36 ist das Ergebnis dieser Untersuchung in Form des Mittelwerts aller Szenarien anhand der Gesamtkosten sowie der Betriebs- und Replikationskosten dargestellt. Die Metriken ohne Information über die LPM-Entscheidungen liefern schlechtere Minimalwerte. Erstaunlicherweise erzielt die zufällige Ordnung RND ein besseres Ergebnis als die Metriken ERR und COR, die sich somit als unbrauchbar erweisen. Herauszuheben sind die Ergebnisse mit der Metrik REP. Diese verursacht von den „unwissenden“ Metriken durchweg die geringsten Kosten. Wie erwartet, sind mit ihr die Replikationskosten minimal.

Bei den „wissenden“ Metriken sind UV1 und UV2 in etwa gleich gut (UV1 geringfügig besser). Die Berücksichtigung der Kostenersparnis hat hier also keinen positiven Einfluss, die der VM-Auslastung in UV3 dagegen schon. Die Betriebskostenkurve besitzt hier zu Anfang eine steilere negative Steigung. Mit den Metriken UVX können zwar die geringsten Gesamtkosten im Testfeld bestimmt werden, jedoch nicht über den gesamten Flexibilitätsbereich. Ab einem Flexibilitätswert von 0,5 führt die Metrik REP zu teils deutlich besseren Ergebnissen. Durch Analyse der Kurvenentwicklung lässt sich dies folgendermaßen erklären:

Die UVX-Metriken stellen die zur Betriebskosteneinsparung relevanten Replikat heraus. Ab etwa einem Flexibilitätswert von 0,4 sind die Betriebskosten beinahe konstant, also nicht mehr verbesserbar. Die Metrikerwerte der übrigen Replikat sind dann nahezu identisch, genauso wie ihre Auswirkung auf die Betriebskosten. Ihre Replikationskosten unterscheiden sich allerdings, was aber nicht von den UVX-Metriken berücksichtigt wird. Relevant

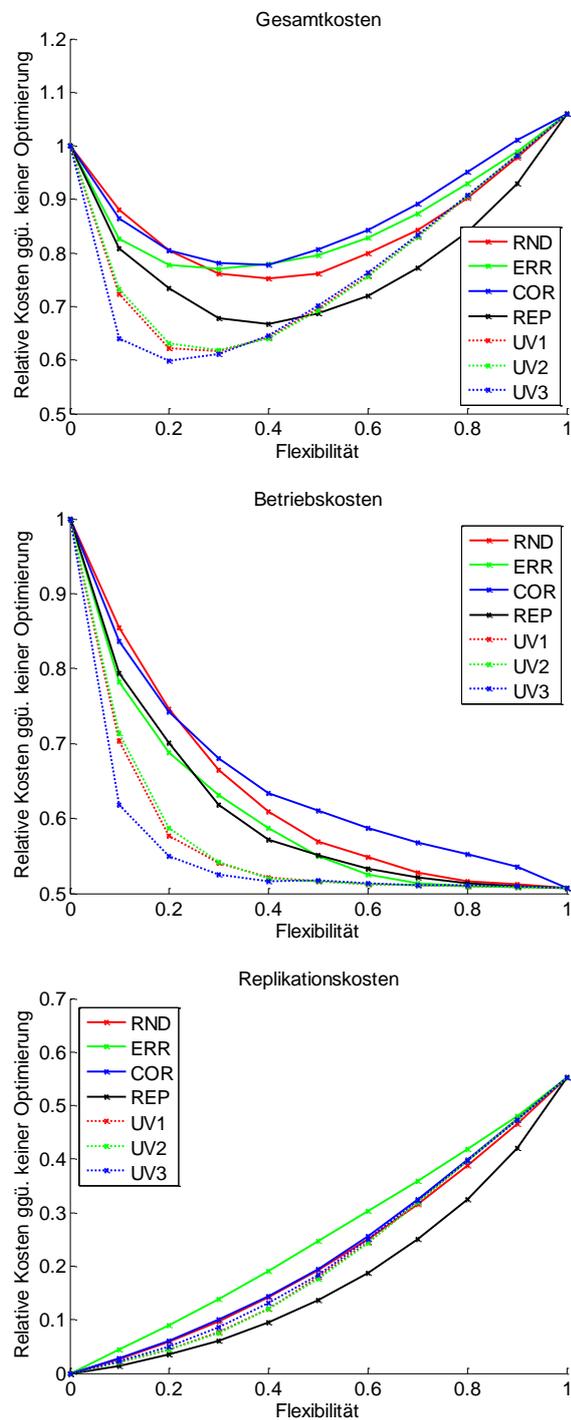


Abbildung 4.36: Gegenüberstellung der Ordnungsmetriken

wird dieser Fall zur Minimierung der Kosten, wenn eine bestimmte Anzahl an Replikaten z. B. aufgrund der SLA vorgeschrieben ist. Eine Kombination mit der Metrik REP wäre dann sinnvoll.

Beim Flexibilitätswert 0 werden die VMs an den Standorten betrieben, an denen sie auch ohne globales LPM betrieben werden würden, daher beginnen alle Kurven bei den gleichen Kosten. Wenn jedoch die Metriken dazu eingesetzt werden, um auch initial die VMs an unterschiedlichen Standorten betreiben zu lassen, nämlich an den entsprechend den Metriken besten Standorten, dann ergibt sich ein anderes Ergebnis, das in Abbildung 4.37 dargestellt ist. Hier zeichnet sich die Unbrauchbarkeit der Metriken ERR und COR umso stärker ab.

Allein die Optimierung der Startstandorte ermöglicht bereits sehr hohe Einsparungen. In der Praxis ist ein Wechsel des Startstandortes nicht immer ohne Weiteres möglich. Insbesondere dann, wenn es sich um unterschiedliche Rechenzentrumsbetreiber handelt, kann die Wahl, dass sich ein Replikate am Startstandort befindet, zu einer Voraussetzung gehören.

Im Folgenden werden keine initiale Optimierung sowie nur noch die Metriken REP und UV3 betrachtet. Ihre Kurvenverläufe sind in beiden Untersuchungen konvex. Dies entspricht auch dem idealisierten Verlauf in Abbildung 4.35. Die hier dargestellten Ergebnisse sind allerdings Mittelwerte aus den oben benannten Variationen der VM-Mengen und Strompreisen. Für einzelne Szenarien ergeben sich auch nicht konvexe Kurvenverläufe.

4.8.4 Beschleunigtes Suchverfahren

In der Güteevaluation wurde die Ordnungsmetrik gemäß Gleichung (4.20) als beste der untersuchten Metriken herausgestellt, um die Flexibilität mit den geringsten Gesamtkosten

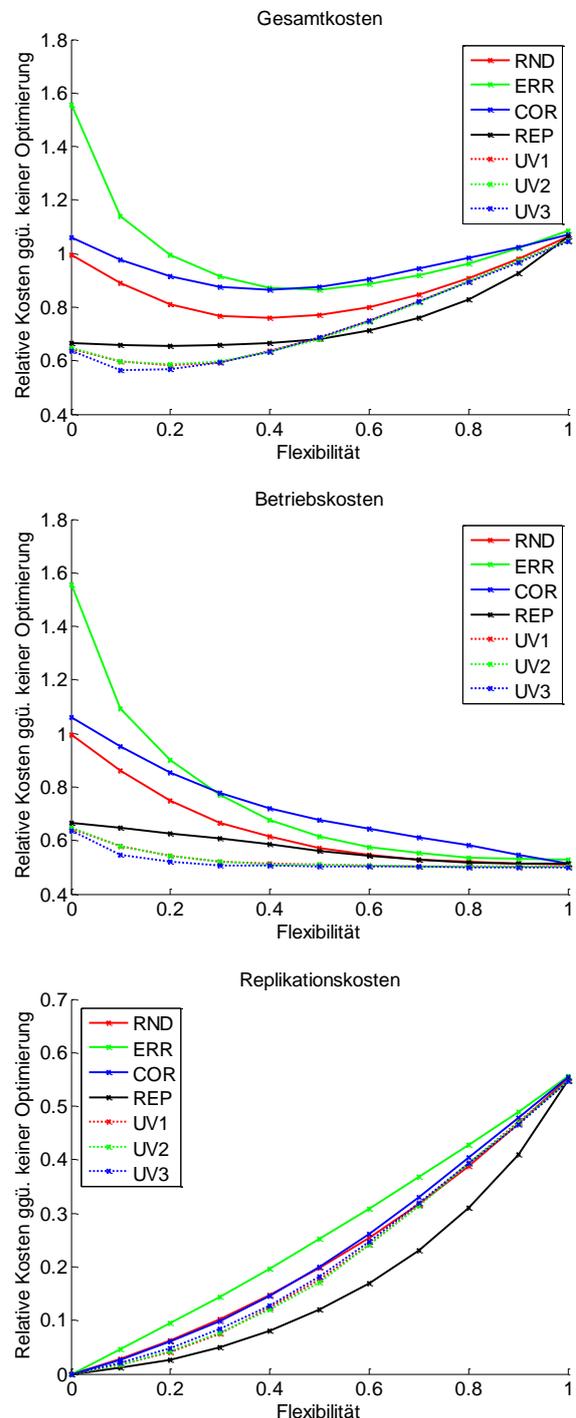


Abbildung 4.37: Gegenüberstellung der Ordnungsmetriken mit Bestimmung des initialen Standortes

4 Konzeption samt Komponentenevaluation

zu ermitteln. An welchem exakten Wert sich das Optimum befindet, müsste allerdings erst durch eine entsprechende Suche gefunden werden. Im Vergleich zur Formel zur Berechnung der Anzahl aller möglichen Replikationsverteilungen in Gleichung (4.17) hat die Flexibilitätsliste deutlich weniger Einträge. Diese besteht nämlich aus den einzelnen Replikaten. Die hierbei maximale Anzahl an zu simulierenden Replikationsverteilungen $\#dist_{list}$ entspricht damit dem Produkt der VM- und Standort-Anzahl:

$$\#dist_{list} = \#VM \cdot \#DC \quad (4.21)$$

Aber auch diese Anzahl ist für eine vollständige Analyse nicht praktikabel. Denn für jede Verteilung muss eine Simulation des gesamten Systems (globales LPM) durchgeführt werden. Setzt man den konvexen Verlauf der idealisierten Gesamtkostenkurve aus Abbildung 4.35 voraus, so könnte die Flexibilitätsliste schrittweise durchlaufen und die entsprechende Verteilung jeweils simuliert werden, bis die Gesamtkosten nicht mehr weiter absinken. Bei der Güteevaluation hat sich allerdings bereits gezeigt, dass auch nicht konvexe Verläufe entstehen, was beim schrittweisen Durchlaufen zu lokalen Minima führen würde, wengleich diese nur sehr selten vorkommen.

Der Kurvenverlauf wurde daher näher analysiert. Es wurden 500 weitere Szenarien untersucht, bei denen neben der VM-Menge und den Strompreisverläufen auch die Datenübertragungskosten und die Standortanzahl variiert wurden. In 16 % der Fälle traten lokale Minima auf. Diese waren entweder auf Niveau des globalen Minimums (maximale Kosteneinsparungsdifferenz unter 1 %), was als Güte (im Vergleich zu den anderen Systemkomponenten) ausreichend ist, oder aber das Tal des lokalen Minimums ist im Vergleich zu dem des globalen Minimums klein bzw. steil zulaufend. Das Tal, in dem sich das globale Minimum befindet, ist, wie in Abbildung 4.38 beispielhaft dargestellt, relativ großflächig, was bedeutet, dass ein großer Flexibilitätsbereich eine höhere Kostenminimierung aufweist als das lokale Minimum. Mit einer adäquaten Abtastung lässt sich das richtige Tal finden und dann z. B. durch ein angepasstes Intervallhalbierungsverfahren das globale Minimum ermitteln. Aufgrund der unterschiedlichen Parametrisierung ergaben sich auch Verläufe, in denen sich das Kostenminimum am Rand, also bei maximaler oder minimaler Flexibilität befand.

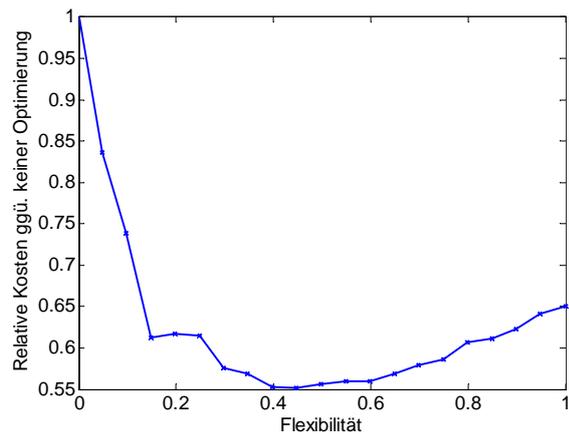


Abbildung 4.38: Beispiel für den Gesamtkostenverlauf bei einer 5%-Abtastung des Flexibilitätswertes

In den untersuchten Fällen hatte eine Abtastung in 10% Schritten genügt, um das Tal des globalen Minimums zu finden. Das erwähnte Intervallhalbierungsverfahren sucht im Bereich des aktuell gefundenen Minimums mit kleiner werdender Intervallgröße nach besseren Werten. In Abbildung 4.39 ist das Suchverfahren skizziert. Nach der ganzflächigen Abtastung des Flexibilitätsspektrums, wird iterativ die nähere Umgebung der aktuell besten Flexibilität untersucht: Je Iterationsschritt wird das Suchintervall auf die Hälfte reduziert und zwei weitere Punkte, nämlich die neuen Intervallgrenzen werden untersucht.

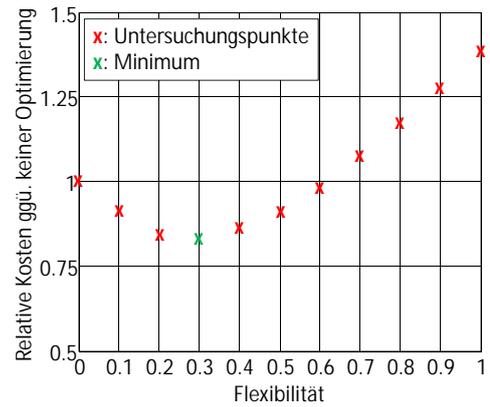
Der Suchbaum hat maximal $k = 2 \cdot \#dist_{list} / (C - 1) - 1$ Knoten, wobei die Wurzel bereits bei der initialen Abtastung ausgewertet wurde. C bezieht sich auf die initiale Anzahl an Abtastwerten (hier 11). Da es sich um einen binären Baum handelt, beträgt seine Höhe $\log_2(k + 1) - 1$. Die Anzahl der zu simulierenden Replikationsverteilungen ($\#dist_{search}$) kann durch dieses Intervallsuchverfahren entsprechend der Gleichung (4.22) weiter verringert werden. Das Suchverfahren hat damit, von der Simulation des globalen LPM abstrahiert, eine logarithmische Laufzeitkomplexität bzgl. der VM- und Standort-Anzahl.

$$\#dist_{search} \leq C + 2 \cdot \lceil \log_2 \frac{\#dist_{list}}{C-1} \rceil \quad (4.22)$$

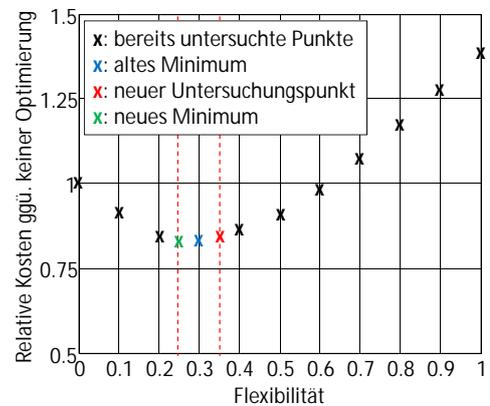
$$\rightarrow O(\log \#dist_{list}) = O(\log (\#VM \cdot \#DC))$$

Die Anzahl an Iterationen kann in der Praxis nochmals verringert werden, indem ein Limit für die Verbesserungen je Iterationsschritt festgelegt wird. In dem realisierten Verfahren wurde eine minimale Verbesserung von 2% als Voraussetzung für eine weitere Untersuchungsiteration verwendet.

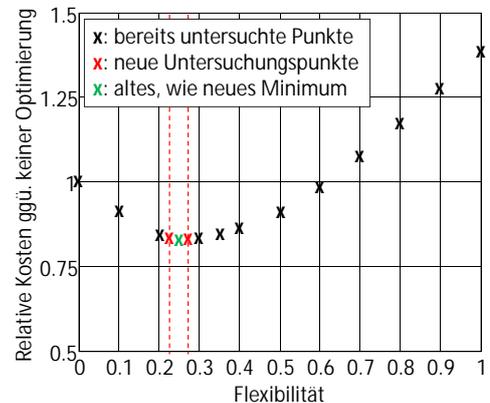
Die Untersuchung der Flexibilitätswerte kann unabhängig voneinander erfolgen, wodurch auf triviale Weise eine Parallelisierung des Suchverfahrens realisiert werden kann. Für eine dezentrale Umsetzung



(a) Schritt 1: Initiale Abtastung.



(b) Schritt 2: Untersuchung der näheren Umgebung des Minimums (halber Abtastschritt).



(c) Schritt i: Wiederholte Untersuchung um neues bzw. altes Minimum mit inkrementell halbiertes Intervallgröße.

Abbildung 4.39: Modifiziertes Intervallhalbierungsverfahren zur Optimierung der Standortselektion

müssten sich die Rechenknoten (lokalen Agenten) noch bzgl. der Untersuchungsergebnisse koordinieren. Bei gleichmäßiger Aufteilung der Untersuchungspunkte kann i.d.R. die Suchtiefe reduziert werden. Dazu kann z. B. die Anzahl der Untersuchungspunkte mit dem Vielfachen der Agentenanzahl gleichgesetzt werden. Bei der initialen Abtastung heiße das $C = \min(\{x | x \cdot \#DC \geq 11\}_{\mathbb{N}})$, und für jeden weiteren Iterationsschritt wäre die Anzahl der neu zu untersuchenden Werte gleich der Agentenanzahl.

Im Rahmen der Dissertation wurde nur die zentrale Ausführung des Verfahrens implementiert. Dieses wird initial – vor der Simulation der zu betrachtenden Zeit – durchgeführt, indem nach Analyse historischer Daten eine bestmögliche Replikationsverteilung ermittelt wird. Ein solches Vorgehen setzt voraus, dass sich das Nutzungsverhalten in näherer Zukunft nicht gravierend ändert. Langfristig ist jedoch mit solch einer Änderung zu rechnen, so dass die Replikationsstandorte aktualisiert werden sollten. Das vorgestellte Verfahren kann zwar auch hierfür gute Standorte ermitteln, jedoch ist zu bedenken, dass die Daten bereits an bestimmten Standorten repliziert sind. Die Erstellung neuer Replikate ist aufgrund der hohen zu übertragenden Datenmenge kostspielig. Ein solches Reallokationsverfahren wurde nicht entwickelt.

4.9 Zusammenfassung

Aufbauend auf den herausgearbeiteten Anforderungen an das zu entwickelnde Verfahren wurde in diesem Kapitel ein MAS konzipiert. Dieses ist darauf ausgelegt, dass die Berechnungslast auf alle beteiligten Rechenzentren verteilt werden kann, und es besitzt eine Laufzeitkomplexität, welche linear mit der Gesamtanzahl aller VMs und quadratisch mit der Anzahl an Rechenzentrumsstandorten skaliert.

Das verwendete Replikationsverfahren ist nicht vollständig in das globale LPM integriert, sondern optimiert die Replikate, die das globale LPM nutzen kann, aus einer Beobachterrolle. Hierdurch bleiben die Verfahren voneinander unabhängig und sind daher austauschbar. Diese modularisierte Austauschbarkeit gilt auch für die einzelnen Systemkomponenten. Hierzu gehören das lokale LPM, Modelle zum Energiebedarf des Rechenzentrums sowie zum Konsolidierungsverhalten des lokalen LPM und der lokale Agent, der mit Hilfe eines Selektionsverfahrens die benötigten VMs zur Erfüllung einer Lastgangvorgabe ermittelt.

5 Gesamtevaluation und Analyse

In Kapitel 4 wurde das Gesamtsystem zum globalen LPM vorgestellt und seine einzelnen Komponenten bereits auf ihre Güte hin evaluiert. In diesem Kapitel wird nun das Gesamtsystem evaluiert. Die wissenschaftlichen Fragestellungen dieser Arbeit zielen auf die Brauchbarkeit und den Nutzen eines globalen LPM ab und darauf, welche Auswirkungen dies auf den Betrieb von Diensten sowohl für den Betreiber als auch den Endkunden hat. Zur Beantwortung dieser Fragen wird eine Sensibilitätsanalyse durchgeführt, bei der unterschiedliche prägnante Parameter variiert werden und ihr Einfluss auf den Einspareffekt untersucht wird.

Vor der Analyse wird kurz auf die verwendete Simulationsumgebung eingegangen und die Herkunft der verwendeten Daten angegeben.

5.1 Simulationsumgebung & Datenbasis

Zur Ermittlung der Güte der einzelnen Systemkomponenten als auch für die Sensibilitätsanalyse des Gesamtsystems werden Simulationen durchgeführt. In diesen werden die einzelnen Modelle zum Rechenzentrum miteinander kombiniert und nach einem simplen Zeitmodell mit unterschiedlichen Daten zu den Ressourcenbedarfen, Strompreisen etc. evaluiert.

Das lokale LPM analysiert in jedem Zeitschritt die aktuelle Rechenzentrumssituation und entscheidet, ob eine Migration aufgrund eines Engpasses, eines Konsolidierungsvorhabens oder einer rechenzentrenübergreifenden Verlagerung durchzuführen ist. Die Länge des Zeitschritts richtet sich nach der kleinsten im Betrachtungsintervall vorkommenden zu berücksichtigenden Zeit (siehe Gleichungen (3.1) und (3.2)): die Migrationszeit. In Abschnitt 2.1.2 wurde für die hier betrachteten VMs eine benötigte Migrationsdauer von 11 Sekunden ermittelt. Zur vereinfachten Handhabung wird jedoch die Migrationszeit und damit die Auflösung auf 1 Minute gesetzt.

Das globale LPM agiert mit einer Auflösung entsprechend der Länge des Optimierungsintervalls. Hier wird das Gesamtsystem analysiert und es werden die entsprechenden Lastverlagerungsmigrationen in die Wege geleitet. Die Migrationszeit für eine globale Migration wird auch auf eine Minute festgelegt, damit sie der lokalen Migrationszeit entspricht. Dies vereinfacht die Verarbeitung im lokalen LPM vereinfacht. Der gesamte Ablauf ist schematisch in Abbildung 5.1 skizziert.

Zur Simulation dieses Systems wurde auf einem C#-Framework aufgebaut, das auch im Projekt AC4DC¹ zur Bestimmung des Einsparpotentials dynamischer Virtualisierung in Re-

¹<http://www.it2green.de/en/779.php>, letzter Zugriff: 28.08.2014

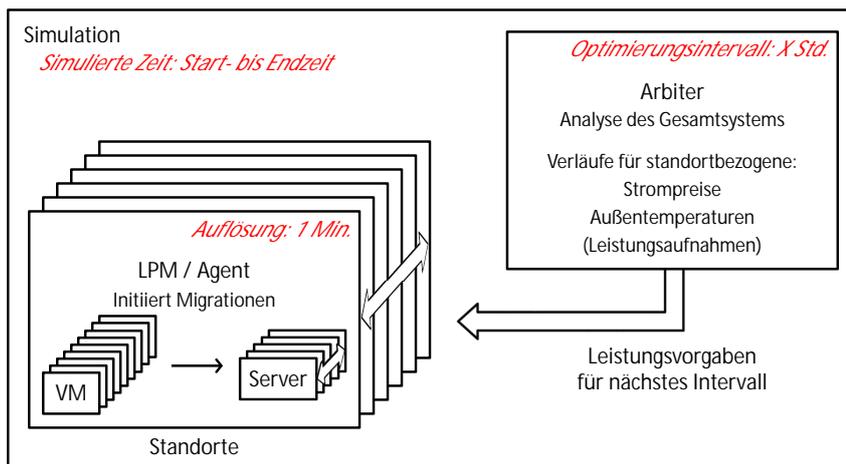


Abbildung 5.1: Zeitliche Größen in einer Simulation

chenzentren genutzt wurde. Auf den Einsatz eines MAS-Frameworks (wie z. B. Boris.NET) wurde verzichtet, da zum einen das genutzte Framework bereits über eine Simulationskomponente verfügt und zum anderen die Kommunikation zwischen den Agenten sowie deren Ausführung überschaubar und sequentiell realisierbar sind.

Innerhalb einer Simulation wird ein Szenario gebildet, das sich anhand der folgenden Eigenschaften variieren lässt:

- Simulierte Zeit
- Optimierungsintervall
- Anzahl genutzter Rechenzentrumsstandorte
- Anzahl betriebener VMs sowie die konkrete Auswahl
- Freie Rechenzentrumskapazität
- VM- und Modellcharakterisierung
- Strompreisverläufe
- Temperaturverläufe
- Energiebedarf für die Datenübertragung

Auf diese wird nun im Folgenden kurz eingegangen sowie die Herkunft der Daten für z. B. die Modellcharakterisierungen angegeben.

Simulierte Zeit bzw. Simulationslänge

Die Datensätze zu den Verläufen der Strompreise, Temperaturen und VM-Auslastungen stammen aus unterschiedlichen Messzeiträumen. In einer Simulation wird eine bestimmte

Zeitspanne simuliert. Die Zeitstempel der jeweiligen Datensätze werden hierzu ausgerichtet. Es ist auch möglich einen gewissen Versatz anzugeben, so dass unterschiedliche Konstellationen simuliert werden können. In den durchgeführten Simulationen wurde der Betrieb über wenige Tage bis Wochen simuliert.

Optimierungsintervall

Das Optimierungsintervall gibt an, in welchen Abständen die Verteilung der VMs auf globaler Ebene optimiert werden soll. Während geringere Abstände eine flexible Anpassung an veränderte Situationen erlauben, verursacht dies unter Umständen einen höheren Migrationsaufwand (siehe anschließende Sensibilitätsanalyse). Ein größeres Optimierungsintervall kann durch den größeren Weitblick eine für eine längere Zeitspanne bessere Verteilung ermitteln, was Migrationskosten einspart. Untersucht wurden Intervallgrößen in der Größenordnung von Stunden bis Tagen.

Für die Optimierung innerhalb eines Intervalls müssen Prognosewerte herangezogen werden. Je größer das Optimierungsintervall ist, desto geringer ist die Qualität der prognostizierten Werte. In Abschnitt 3.1.2 wurde gezeigt, dass die Fehlerwahrscheinlichkeit bei Temperaturprognosen über einen Tag hinaus bei über 10 % liegt und daher entsprechende Optimierungsintervalle vermieden werden sollten.

Standortanzahl

Die Standortanzahl des Verbundes gibt an, über wie viele Rechenzentren das globale LPM eingesetzt werden soll. An einem Standort werden die Verläufe für die Strompreise, Temperaturen und evtl. auch des regionalen Leistungsbedarfs der Siedlung um das Rechenzentrum zusammengeführt. Dies erfolgt in quasi zufälliger Weise durch Verwendung eines Seeds.

Anzahl und Auswahl der betriebenen VMs

Mit der Anzahl der zu betreibenden VMs wird die Größe des globalen VM-Pools bestimmt. Diese VMs werden initial über die Standorte gleich verteilt. Über einen Seed wird bestimmt, aus welchen konkreten VMs dieser Pool aus der Gesamtmenge der verfügbaren VMs gebildet wird.

Freie Rechenzentrumskapazität

Die Rechenzentrumskapazität wird durch die Anzahl und Rechenleistung der verwendeten Server bestimmt. In den hier durchgeführten Simulationen werden homogene Server eingesetzt, so dass sich die Kapazität allein aus der Anzahl ergibt. Für den Betrieb von VMs wird eine bestimmte Anzahl an Servern benötigt. Als Beispiel werden für den Betrieb von 500 VM insgesamt 200 Server reserviert, die im Falle maximaler Auslastung genutzt werden. Aktiv werden jedoch weit weniger Server benötigt (Stichwort: sichere und dynamische Verteilung).

Bei Gleichverteilung über fünf Standorte sind pro Rechenzentrum 40 Server das Minimum. In diesem Falle könnte jedoch keine Lastverlagerung stattfinden, da keine freien Kapazitäten vorliegen. Wenn jedes Rechenzentrum 200 Server hätte, könnten alle VMs zu jedem Zeitpunkt immer ins günstigste Rechenzentrum migriert werden. Dies würde in diesem Beispiel einer freien Kapazität von 80 % entsprechen.

Modell- und VM-Charakterisierung

Die erstellten Modelle werden mit Daten realer Hardware charakterisiert. Beim Rack bzw. Blade-Enclosure, der Klimatisierung und der USV liegt nur jeweils ein Datensatz vor, welcher auch bei der jeweiligen Modellierung in Abschnitt 4.2 verwendet wurde. Es kann hieraus kein allgemeingültiges Modell abgeleitet werden. Daher werden diese Modelle auch nur mit dem jeweiligen Datensatz charakterisiert.

Server

Mit dem SPECpower_ssj2008 Benchmark wird der Energiebedarf und die Effizienz von Servern ermittelt. Eine Sammlung von Ergebnissen ist auf der entsprechenden Webseite² veröffentlicht. Für die durchgeführten Simulationen wurde der folgende Servertyp verwendet: Dell Inc. PowerEdge R520 (Intel Xeon E5-2470, 2.30 GHz).

Storage und Netzwerk

In [233] wird ermittelt, dass der statische Energiebedarfsanteil eines Routers bei mehr als 90 % des maximalen Energiebedarfs liegt. Für die hier durchgeführten Untersuchungen ist der statische Anteil auf 90 % und der dynamische auf 10 % gesetzt. Die gleiche Aufteilung ist auch für den Energiebedarf des Storages festgelegt. Konkrete Auslastungsdaten liegen nicht vor. Deshalb wird von einer Auslastung ausgegangen, die proportional von den Auslastungen der betriebenen VMs beeinflusst wird. Insgesamt entspricht die Auslastung damit der Gesamtauslastung des Rechenzentrums *SoL* genormt mit der maximal möglichen Auslastung.

In [36] und in Abbildung 2.1 wird der Energiebedarf des Storages und der Netzwerkkomponenten auf 4 % - 5 % des Energiebedarfs des gesamten Rechenzentrums beziffert. Da es sich bei den dort untersuchten Rechenzentren um voraussichtlich statische Strukturen handelt, bei denen der Energiebedarf nicht mit der Last skaliert, sondern auf einem konstanten Peak-Niveau liegt [30, 22], wird davon ausgegangen, dass sich der bezifferte Wert auf den maximalen Rechenzentrumsenergiebedarf bezieht. Bezogen auf den IT-Anteil wird also von einem Energiebedarf ausgegangen, welcher jeweils etwa 10 % des maximalen IT-Energiebedarfs entspricht.

Die Bestimmung der aktuellen Leistungsaufnahme von Storage und Netzwerk soll nun an einem kleinen Beispiel verdeutlicht werden: Hat ein Rechenzentrum insgesamt 20 Server, die je eine maximale Leistungsaufnahme von 200 W aufweisen, so liegt die maximale Leistungsaufnahme aller Server bei 4 kW. Storage und Netzwerk haben einen Anteil von insgesamt

²http://www.spec.org/power_ssj2008/results/, letzter Zugriff: 05.09.2014

20 %, so dass ihre maximale Leistungsaufnahme bei $4 \text{ kW} \cdot \frac{0,2}{1-0,2} = 1 \text{ kW}$ liegt. Der statische Anteil von Storage und Netzwerk liegt damit bei 0,9 kW und der dynamische bei 0,1 kW. Die maximal mögliche Auslastung bzgl. der Gesamt-CPU-Performanz liegt in diesem Beispiel bei $20 \cdot 100 \% = 20000 \%$.

Dienste bzw. virtuelle Maschinen

Um die entwickelten Verfahren zu evaluieren, werden sie anhand realer VM-Datensätze simuliert. Hierzu wurden von zwei im Raum Oldenburg ansässigen KMU über mehrere Wochen hinweg die Auslastungsdaten von Diensten bzgl. der CPU und des Hauptspeichers aufgezeichnet. Die zeitliche Auflösung der Daten beträgt 5 Minuten. Insgesamt konnten Daten für rund 200 Dienste bzw. VMs erhoben werden. Um diese auch vergleichbar zu machen, wurden sie an einer Standard-VM-Klasse³ normiert: M1 Small/Medium. Diese Instanzen stellen VM-Leistungsdaten für den allgemeinen Gebrauch dar. Sie verfügen über einen virtuellen Rechenkern sowie 1,7 GB bzw. 3,75 GB Arbeitsspeicher. In den Simulationen wird von einer Arbeitsspeichergröße von 4 GB ausgegangen, was bspw. auch in den Analysen in [164] zugrunde gelegt wird. Die Festplattengröße liegt bei 160 GB bzw. 410 GB. Die Aufzeichnungen bzgl. der CPU-Auslastungen sind skaliert auf den jeweils eingesetzten Server, d.h., dass eine Auslastung von 100 % die gesamte CPU-Performanz des Servers benötigt.

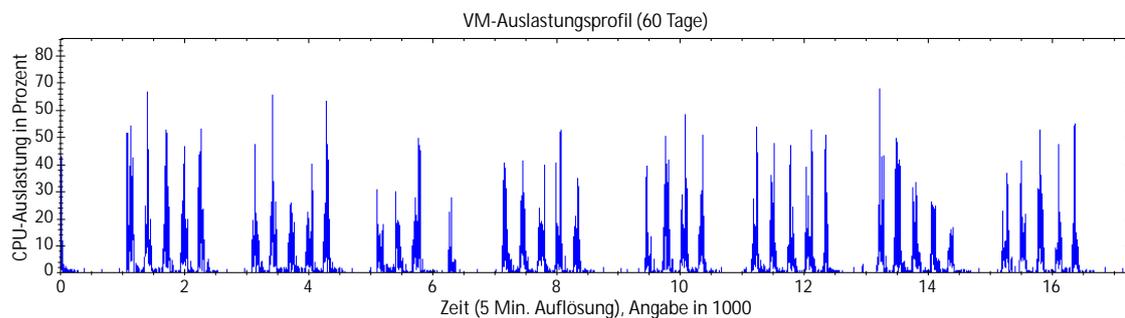


Abbildung 5.2: Beispiel für ein VM-Auslastungsprofil

Die meisten Auslastungsprofile haben einen Verlauf, der dem in Abbildung 5.2 ähnelt: periodisch mit einigen Lastspitzen. Das periodische Verhalten gleicht oft einem wöchentlichen Rhythmus. Zur Vergrößerung der Heterogenität wurden die Rohdaten der rund 200 VM unter randomisiertem Einfluss modifiziert. Diese Modifikationen beinhalteten additive und multiplikative Veränderungen der Ausgangswerte sowie zeitliche Verschiebungen. Außerdem wurden die modifizierten Auslastungsdaten von bis zu fünf Diensten gewichtet kombiniert. Auf diese Weise wurde ein VM-Pool mit insgesamt 10000 unterschiedlichen VM-Auslastungsdaten generiert.

³gemäß der ersten Generation von VMs im Amazon EC2

Strompreise

Für die Simulationen werden die Strompreise des Day-Ahead-Handels der Strombörsen EXAA, GME und NordPool herangezogen. Der verwendete Datensatz umfasst die Jahre 2008 bis einschließlich 2010 und hat eine stündliche Auflösung. Um den Einfluss des Real-Time-Pricings mit hinein zu bringen, werden die Verläufe geringfügig verrauscht. Zusätzlich werden die Verläufe wie bei den Auslastungsdaten der Dienste modifiziert, um eine größere Variation unterschiedlicher Verläufe simulieren zu können.

Temperaturverläufe

Die verwendeten Wetterdaten entstammen von den Wetterdiensten von Google und Yahoo. Es wurden über sechs Wochen (April bis Mai 2012) für insgesamt 31 Städte Werte zur Temperatur, Luftfeuchtigkeit und Windgeschwindigkeit aufgezeichnet. Die betrachteten Städte lassen sich in die folgenden drei Gruppen einteilen:

- Deutschland: Berlin, Chemnitz, Cuxhaven, Frankfurt am Main, Köln, München, Oldenburg, Rostock, Trier
- Europa: Bergen, Bukarest, Edinburgh, Le Mans, Madrid, Moskau, Palermo, Rotterdam, Stockholm, Zagreb
- Welt: Agades, Brisbane, Manaus, Maun, Mexico City, Mountain Village, Nagano, New York, Rio Grande, Santiago, Vancouver, Wuhan

Die Auflösung der Daten beträgt eine Minute, die Änderungsrate liegt allerdings bei mindestens 10 Minuten.

Datenübertragungskosten

In den Untersuchungen [220, 161] zum allgemeinen Energiebedarf des Internets wurden Energiebedarfskoeffizienten von 9 bis 16 kWh/GB bzw. 1 kWh/GB für das Jahr 2006 bzw. 2008 ermittelt. Die hohen Abweichungen ergeben sich nicht nur aufgrund des zeitlichen Unterschieds, sondern auch aufgrund der unterschiedlichen betrachteten Regionen (USA, Europa) sowie der Granularität bei der Abschätzung bzgl. der betrachteten Bestandteile des Internets. In beiden Vorgehensweisen wurden jedoch Hochrechnungen durchgeführt. Eigene Untersuchungen [162] konnten basierend auf diesen Daten Prognosen für die Jahre 2014 (0,1 kWh/GB) bis 2020 (0,01 kWh/GB) geben. Dies entspricht einer negativen jährlichen Wachstumsrate von -32 %. In [17] wird mit einer betragsmäßig geringeren Rate von -14 % bis -20 % gerechnet. Allerdings wird dort bereits für das Jahr 2009 ein sehr geringer Energiebedarfskoeffizient von 6,44 Wh/GB zugrunde gelegt.

In einer aktuelleren Untersuchung [67] wurde der konkrete Energiebedarf für eine Interkontinentalübertragung gemessen. Der hier ermittelte Wert von 0,2 kWh/GB ergibt sich ebenfalls für die Netzwerktechnologie aus dem Jahre 2009. Dieser Wert ist auch konform mit einer

im Rahmen der GreenTouch-Initiative entstandenen Veröffentlichung [126]. Den Ergebnissen zufolge wurde für 2009 ein Energiebedarf von ca. 0,18 kWh/GB ermittelt und für 2014 ca. 0,028 kWh/GB prognostiziert. Aus dem Versuchsaufbau von [67] lässt sich entnehmen, dass es sich bei dem Wert 0,2 kWh/GB bezogen auf eine allgemeine durchschnittliche Datenübertragung um eine pessimistische Angabe handelt, denn es wurden unterausgelastete und redundante Router verwendet.

Allgemein lässt sich festhalten, dass sich der Energiebedarf pro zu übertragendem Bit linear mit steigender Performanz des Routers [77] reduziert. Entsprechend der üblichen Provisionierung von IT-Leistung werden aber auch Router die meiste Zeit nicht voll ausgelastet und verbrauchen dadurch unnötig Strom. Um eine bessere Energiebedarfsadaptivität zur anfallenden Last zu erzielen, werden Managementsysteme entwickelt, die zum einen die Netzwerkkomponente selbst dem nötigen Bedarf anpassen und zum anderen auf höherer Ebene das Routing energieeffizient realisieren [165, 10, 59, 53, 5]. Als Potentialabschätzung wurde in [5] festgestellt, dass mit idealen energie-last-proportionalen Routern eine Einsparung von bis zu 70 % erzielt werden kann. Dies deckt sich mit einer Studie des VDE [151], wonach eine Einsparung von 60 % durch intelligente und last-adaptive Netze möglich ist. Allerdings sind hierzu kooperierende Netze notwendig, statt der aktuell vorliegenden konkurrierenden Lage. Aber auch Verbesserungen in der Kodierung sind möglich: 20 % haben Chen und weitere [58] mit ihrer neuen Kodierung erzielt, wobei sich dieser Wert auf kabelgebundene Lösungen bezieht, denn hier sei im Gegensatz zu Funkverbindungen der Energiebedarf für die Kodierung höher als der für die Übertragung. Insgesamt steigt die Übertragungseffizienz jährlich aufgrund der fortschreitenden Technologien exponentiell und soll durch die GreenTouch-Initiative im Vergleich zum bisherigen Trend noch schneller verbessert werden.

Die forcierte Steigerung der energiebezogenen Übertragungseffizienz kommt einem globalen LPM zugute, wie die anschließende Evaluation zeigt. Um sowohl ältere Technologie als auch aktuelle und zukünftige zu berücksichtigen, werden hier Energiebedarfskoeffizienten von 0,2 bis 0,001 kWh/GB untersucht.

5.2 Einsparpotential – Sensibilitätsanalyse

Das Einsparpotential eines globalen LPM ist von vielen unterschiedlichen Parametern abhängig. Wenn z. B. die Strompreisverläufe konstant sind, sollten die VMs permanent an den günstigsten Standorten betrieben werden. Ein globales LPM ist hier nicht notwendig. Eine gewisse Änderungsrate in den Strompreisverläufen ist daher notwendig. Gleiches gilt für die Temperaturverläufe. Die hieraus sich ergebenden Einsparpotentiale müssen allerdings durch Migrationen erkauft werden, welche zusätzliche Kosten verursachen.

Die folgenden Parameter werden bzgl. ihres Einflusses auf das Einsparpotential untersucht:

- Migrationsgeschwindigkeit
- Temperaturverläufe

5 Gesamtevaluation und Analyse

- Strompreisverläufe
- VM-Profil-Korrelation
- Netzwerkenergiebedarf
- Optimierungsintervall
- Freie Rechenzentrumskapazität
- Standortanzahl

In der Regel beeinflussen sich diese Parameter gegenseitig, so dass für eine vollständige Analyse eine Kreuzvariation notwendig wäre. Da diese allerdings 1. zu komplex wäre und 2. der Fokus auf dem generellen Einfluss liegt, werden sie separat betrachtet. Einzig das Optimierungsintervall wird vereinzelt mit anderen Parametern kreuzweise variiert. Die jeweils übrigen Parameter werden auf einen Standardwert gesetzt, der im jeweiligen Variationsbereich liegt und eine plausible Ausgangslage beschreibt.

Bei allen Untersuchungen werden die Ergebnisse normiert zum Kosten- bzw. Energiewert angegeben, welcher ohne Anwendung des globalen LPM erzielt wird. Es werden jeweils zwei Tage simuliert und für jeden Messwert – soweit nicht anders angegeben – fünf unterschiedliche Strompreisverläufe und fünf unterschiedliche Selektionen für die Erstellung der betrachteten VM-Menge untersucht. Hierdurch beschränken sich die abgeleiteten Aussagen nicht auf ein bestimmtes Szenario, sondern gewinnen an Allgemeingültigkeit.

5.2.1 Migrationsgeschwindigkeit

Die Migrationsgeschwindigkeit bezieht sich hier nicht allein auf die Bandbreite des Netzanschlusses, sondern umfasst auch die Fähigkeit des lokalen LPM, VMs aufzunehmen bzw. abzugeben. In einem ersten Schritt wurde daher die durchschnittliche Anzahl an vom lokalen LPM durchführbaren Migrationen ermittelt. Rein konzeptionell ist, wie beschrieben, im lokalen LPM nur eine Migration pro Zeitschritt (hier eine Minute) vorgesehen. Dies ist die maximal erreichbare Migrationsquote. In unterschiedlichen Simulationen konnte jedoch nur eine durchschnittliche Verarbeitungsgeschwindigkeit von 0,67 globalen Migrationen je Standort und Minute ermittelt werden. In 33 % der Zeit, wenn globale Migrationen anstehen, muss sich also das lokale LPM um das rechenzentrumsinterne Management kümmern, was zu einer Verzögerung der globalen Migrationen um knapp 50 % führt.

Diese Untersuchung soll das Potential des globalen LPMs zeigen, das daher möglichst wenig von den spezifischen lokalen Beschränkungen betroffen sein sollte. Mit anderen Verfahren für das lokale LPM ist es durchaus möglich, mehr Migrationen zu verarbeiten. Da allerdings kein anderes konkretes Verfahren vorliegt, wird das theoretisch mögliche Einsparpotential ermittelt, indem eine Grenzwertbetrachtung durchgeführt wird. Hierzu wird das lokale LPM in den Modus *instantane Optimierung* versetzt, so dass es beliebig viele Migrationen aufnehmen kann, ohne dabei Engpässe aufkommen zu lassen. Dieser Modus ist real nicht durchführbar, sondern soll an

dieser Stelle lediglich aufzeigen, welches Potential durch die verzögerte Migrationsverarbeitung ungenutzt bleibt und somit durch ein in diesem Sinne besseres lokales LPM noch gehoben werden könnte.

In dem hier untersuchten Szenario wird eine VM-Menge mit 200 VM betrachtet, zu deren Betrieb zwei Rechenzentren mit 50 % freien Kapazitäten zur Verfügung stehen. Dadurch kann der Betrieb sämtlicher VMs in einem einzelnen Rechenzentrum erfolgen. Für die VM-Menge und die Strompreisverläufe werden jeweils zehn unterschiedliche Selektionen untersucht. Da es hierbei um die Ermittlung der theoretischen Verbesserung der Migrationsgeschwindigkeit geht, wird die Berücksichtigung der Migrationskosten gezielt ausgespart.

In Abbildung 5.3 ist das Untersuchungsergebnis für unterschiedliche Optimierungsintervalle angegeben. Die Nachteile des normalen Betriebs, hier als *inkrementelle Optimierung* benannt, zeigt sich insbesondere bei kleinen Optimierungsintervallen. Die Ergebnisse der instantanen Optimierung sind in der Realität zwar nicht erreichbar, jedoch kann man sich ihnen annähern, was auch bereits in dieser Untersuchung bei höheren Optimierungsintervallen zu sehen ist. Der große Unterschied bei kleineren Intervallen beruht darauf, dass die Zeit nicht ausreicht, alle vorgesehenen Migrationen durchzuführen. Je mehr Zeit zur Verfügung steht, desto mehr nähern sich die erzielten Einsparungen einander an. Übrig bleibt die nicht erzielte Einsparung während des Migrationsvorgangs. Das theoretische Einsparpotential weist einen logarithmus-ähnlichen Verlauf in Abhängigkeit vom Optimierungsintervall auf.

Eine nähere Analyse der Ergebnisse führt zu einer weiteren Erkenntnis. In Abbildung 5.4 sind die Ergebnisse in Abhängigkeit von den Variationen der Strompreisverläufe und der VM-Mengen-Selektion angegeben. Hierin ist zu erkennen, dass die VM-Mengen-Selektion einen geringeren Einfluss auf den Einspareffekt hat. Während sich hier Unterschiede bei den Einsparereffekten von maximal 30 % ergeben, können diese bei unterschiedlichen Strompreisverläufen auch im Bereich von Vielfachen liegen.

5.2.2 Strompreisverläufe

Da die konkreten Strompreisverläufe einen hohen Einfluss auf den erzielbaren Einspareffekt haben, soll dieser Einfluss nun näher untersucht werden – wiederum für zwei Rechenzentren. Für diese Untersuchung werden nicht die aufgezeichneten Strompreise genutzt, sondern Verläufe generiert, welche sich in den beiden Parametern *maximale Differenz* und *Vorzei-*

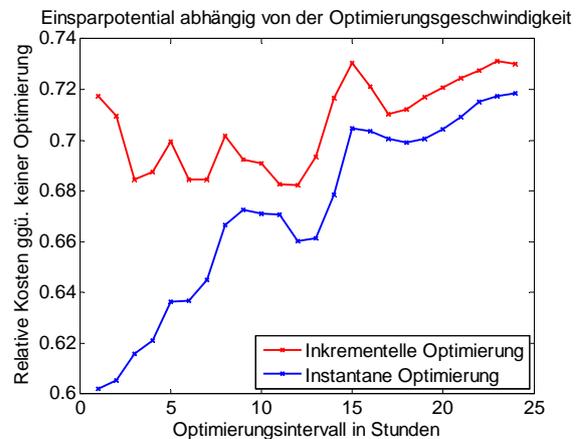


Abbildung 5.3: Migrationsgeschwindigkeit: Je schneller die globalen Migrationen durchgeführt werden können, desto höher ist das Einsparpotential.

5 Gesamtevaluation und Analyse

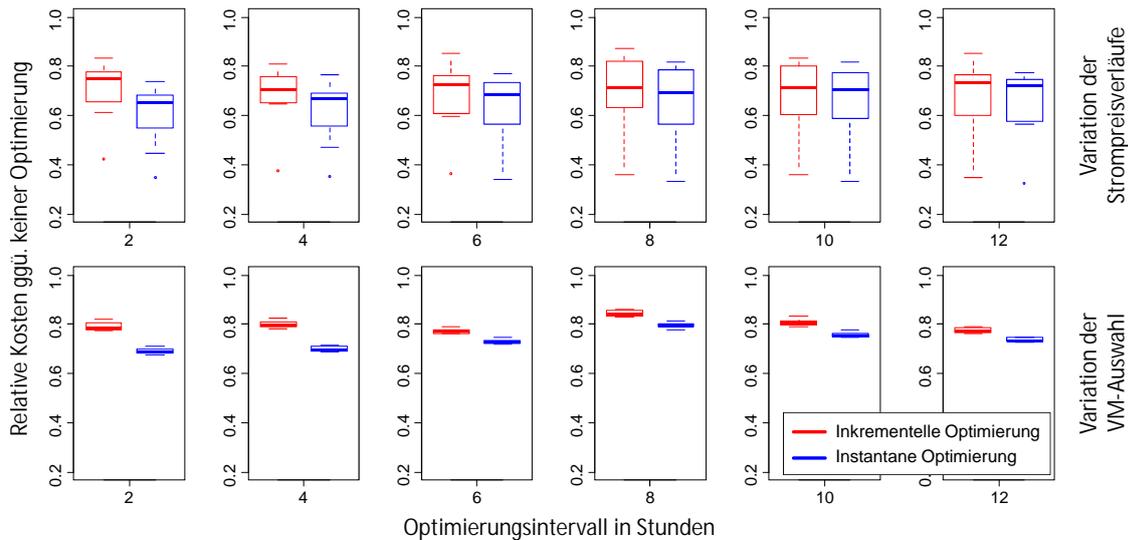


Abbildung 5.4: Die Auswahl der konkreten VM-Menge hat einen geringeren Einfluss auf das Einsparpotential als die der konkreten Strompreisverläufe, zu erkennen an den kleineren Boxen → geringere Streuung.

chenwechselwahrscheinlichkeit unterscheiden. Diese Parameter charakterisieren die Relation zwischen zwei Strompreisverläufen. Die Vorzeichenwechselwahrscheinlichkeit bezieht sich auf das Vorzeichen bei der Differenzbildung zweier aufeinander folgenden Strompreisangaben. Die maximale Differenz ist in Vielfachen der jeweils kleineren Zahl angegeben. Es wird von Preisen größer null ausgegangen. Ausgehend von einem aufgezeichneten Strompreisverlauf wird ein zweiter Verlauf unter Beeinflussung dieser beiden Parameter generiert. Die quasi zufällige Generierung wird mit einem Seed festgelegt. Für jede Parametervariation werden zehn unterschiedliche Seeds angewendet. Die übrigen Szenarioeinstellungen gleichen denen der Untersuchung zur Migrationsgeschwindigkeit.

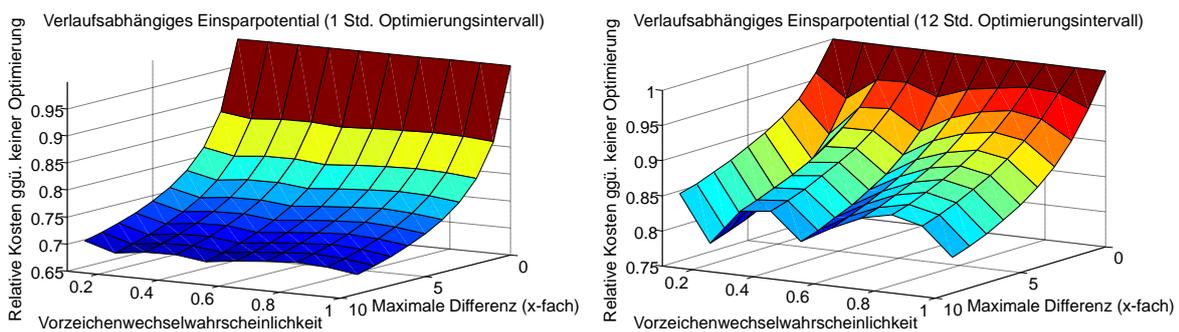


Abbildung 5.5: Einfluss unterschiedlicher Strompreisverläufe

In Abbildung 5.5 sind die Ergebnisse für ein Optimierungsintervall von einer Stunde und von zwölf Stunden dargestellt. In beiden Fällen steigt das Einsparpotential mit steigender maximaler Differenz logarithmisch an. Eine Beeinflussung durch die Vorzeichenwechselwahrscheinlichkeit ist zumindest beim 1-stündigen Optimierungsintervall nicht zu erkennen, was

nicht verwunderlich ist, da sich die Strompreise auch nur stündlich ändern. Bei einem solch kurzen Optimierungsintervall kann auf diese Veränderungen also komplett reagiert werden. Beim 12-Stunden-Intervall kann auf diese Veränderungen nicht so gut eingegangen werden. Dies ist grundsätzlich – wie auch in der vorherigen Untersuchung – am geringeren Einsparpotential zu erkennen. Die Ergebnisse für die unterschiedlichen Vorzeichenwechselwahrscheinlichkeiten sind nicht so konstant wie beim 1-stündigen Intervall. Dies liegt an der Zufallsgenerierung der Strompreisverläufe: Einige Verlaufspaare können mit dem langen Intervall besser optimiert werden als andere.

5.2.3 Temperaturverläufe

Neben den Strompreisen haben auch die Temperaturverläufe einen indirekten Einfluss auf die Kosten, da sie den Energiebedarf beeinflussen. Um diesen herauszustellen, wird ausschließlich der Energiebedarf untersucht. Die Strompreise aller Standorte werden dazu konstant und gleich gesetzt.

Die Temperaturen von geografisch naheliegenden Standorten ähneln sich mehr als von weiter entfernt liegenden. Dies hat nicht nur mit der unterschiedlichen Tageszeit zu tun, sondern auch mit den unterschiedlichen klimatischen Bedingungen. In dieser Untersuchung werden drei Regionsgrößen untersucht: national, kontinental und interkontinental. Für die nationale Größe wird Deutschland und für die kontinentale Größe Europa stellvertretend betrachtet.

Im untersuchten Szenario wird ein Verbund von zwei Rechenzentren mit dem Betrieb von 200 VM sowie einem Optimierungsintervall von einer Stunde simuliert. Zur Variation der Wetterdaten werden jeweils fünf unterschiedliche Konstellationen für zufällig gewählte Standorte und Startzeiten durch einen Seed bestimmt. Die Ergebnisse sind in Abbildung 5.6 dargestellt.

Je nach betrachteter Region und dem damit resultierenden unterschiedlichen Grad an Temperaturunterschieden können Einsparungen bis zu 10% erreicht werden. Je größer die betrachtete Region ist, desto mehr Energie kann potentiell eingespart werden. Allerdings steigt hierbei in den untersuchten Fällen der absolute Energiebedarf, da die Temperaturen in den betrachteten Standorten im Durchschnitt höher sind.

Die Verläufe der Strompreise und der Temperaturen können über die Standorte hinweg gesehen gegenläufig sein. Dadurch können sich die Einspareffekte teilweise bis ganz aufheben. In einem Simulationslauf werden die Verläufe zufällig einem Standort zugeordnet. Daher kann der nivellierende Effekt willkürlich auftreten. Dies erschwert die Herausstellung der Sensibilität auf die einzelnen Parameter. Für die weiteren Untersuchungen werden daher

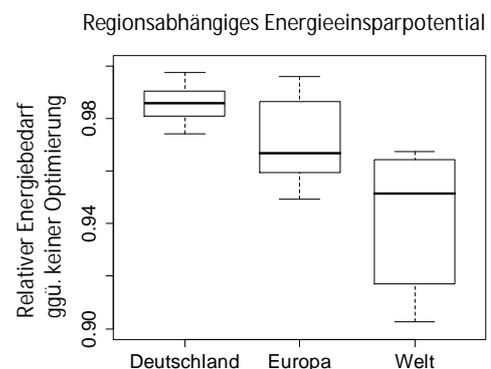


Abbildung 5.6: Einfluss unterschiedlicher Temperaturverläufe (regionsabhängig)

ohne Beschränkung der Allgemeinheit die Temperaturverläufe ignoriert und eine einheitliche Temperatur (20 °C) zur Berechnung des Energiebedarfs angenommen.

5.2.4 VM-Profil-Korrelation

In der Untersuchung zur Migrationsgeschwindigkeit hat sich zwar gezeigt, dass die konkrete VM-Auswahl einen eher geringen Einfluss auf den Einspareffekt hat, jedoch wurden dabei durch die zufällige Auswahl tendenziell ausgeglichene VM-Mengen gebildet. Das heißt, dass die Auslastungsprofile der VMs sich im Durchschnitt nivellieren. Es besteht damit in der gesamten Menge eine geringe Korrelation zwischen den Auslastungsprofilen. Die Nachstellung von gewünschten Ziellastgängen ist hiermit leichter zu realisieren, als wenn die Profile miteinander korrelieren würden. Inwiefern korrelierende Auslastungsprofile den Einspareffekt reduzieren, ist Bestandteil dieser Untersuchung.

Der Korrelationswert ist zwischen jeweils zwei Datensätzen definiert, nicht jedoch für eine Menge von Datensätzen. Daher wird an dieser Stelle von Gleichheit gesprochen, um ein Maß für die mengenbezogene Korrelation auszudrücken. Die Gleichheit wird dadurch verändert, dass eine bestimmte VM mehrfach in der VM-Menge instantiiert wird. Eine Gleichheit von 1 bedeutet, dass die Menge zu 100 % aus der betreffenden VM besteht. Bei der hier untersuchten Mengengröße von 200 VM wäre eine bestimmte VM also 200 mal instantiiert. Bei einem Gleichheitswert von 0 wäre die VM nicht in der VM-Menge vertreten.

Neben fünf unterschiedlichen Strompreisverläufen wurden auch fünf unterschiedliche VMs für die Gleichheitsveränderung selektiert. Der gemessene Mittelwert ist in Abbildung 5.7 dargestellt. Hier ist zu erkennen, dass die Selektion der konkreten VM-Menge durchaus einen großen Einfluss auf den Einspareffekt hat. Mit steigender Gleichheit der VM-Profile sinkt das Einsparpotential. Die Steigung ist bei kleineren Optimierungsintervallen steiler, was daran liegt, dass die unterschiedlichen Verläufe hier besser ausgenutzt werden können.

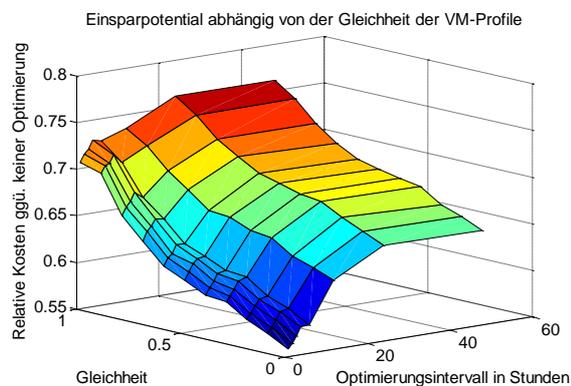


Abbildung 5.7: Geringeres Einsparpotential bei korrelierenden VM-Auslastungsprofilen

5.2.5 Netzwerkenergiebedarf & Optimierungsintervall

In den vorangegangenen Untersuchungen wurden keine Migrationskosten berücksichtigt. Daher konnten alle Migrationen durchgeführt werden, die zu einer Senkung der Betriebskosten führen. Die Berücksichtigung von Migrationskosten reduziert das mögliche Einsparpotential. Die Kosten ergeben sich nach dem vorgestellten Netzwerkmodell aus dem Energiebedarf zur Übertragung der Daten und den aktuellen Kosten der jeweils beteiligten Standorte.

In dieser Untersuchung werden sechs Rechenzentren mit insgesamt 600 VM simuliert. Aus den in Abbildung 5.8 dargestellten Ergebnissen wird das erste Mal eine andere Abhängigkeit vom Optimierungsintervall erkennbar. Bei geringem Energiebedarf für die Datenübertragung ist noch der logarithmische Zusammenhang erkennbar. Bei höherem Energiebedarf stellt sich bei kleinen Intervallen jedoch ein geringerer Einspareffekt ein. Insgesamt werden zwar immer noch Kosten ggü. dem Fall fehlender Optimierung eingespart, jedoch wird durch den geringen Weitblick eine verschlechterte Ausgangsposition für die weiteren Optimierungsintervalle geschaffen. Während bei langem Optimierungsintervall die Verteilung über eine lange Periode minimiert wird, wird bei kurzen Intervallen immer die aktuell – für die kurze Periode – kostenminimale Verteilung gesucht. Um diese allerdings jeweils zu erreichen, müssen immer wieder Migrationskosten aufgewendet werden. Der Verlauf der Migrationskosten selbst hat keine größere Bedeutung, sondern ist als Produkt der Betriebskostenminimierung anzusehen. Es zeigt sich aber, dass mit relativ geringen Migrationskosten hohe Einsparungen bei den Betriebskosten (inkl. Migrationskosten) erzielt werden können.

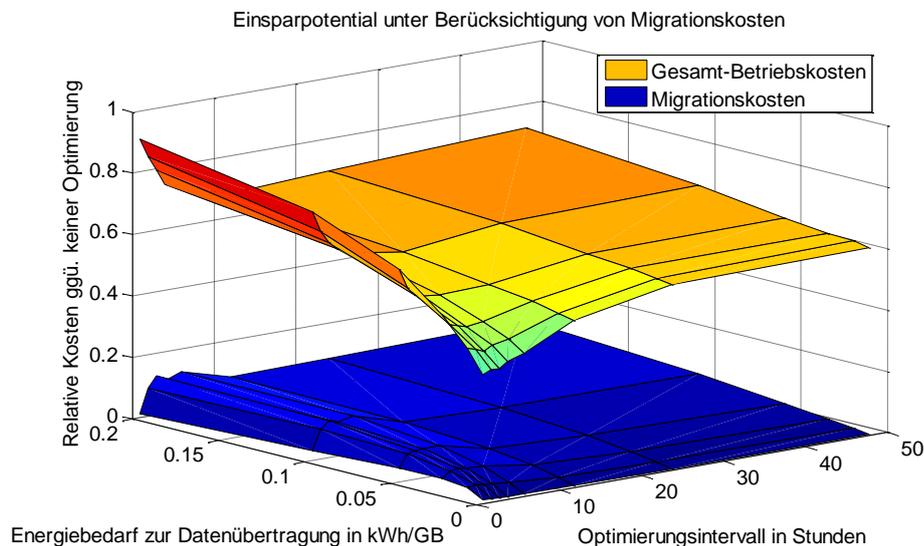


Abbildung 5.8: Einfluss der Datenübertragungskosten und des Optimierungsintervalls

Insgesamt lässt sich aus der Untersuchung ableiten, dass bereits mit dem für das Jahr 2009 geltenden Wert für den Energiebedarf zur Datenübertragung von 0,2 kWh/GB ein signifikantes Einsparpotential vorliegt. Die für die längeren Optimierungsintervalle notwendigen Daten müssen durch entsprechend lange Prognosen bzw. gültige Werte gestützt werden. Prognosen über einen Tag hinaus sind unzuverlässig, aber, wie sich hier herausgestellt hat, auch nicht notwendig. Die höchsten Einsparungen werden bei Optimierungsintervallen mit einer Länge von unter 12 Stunden erzielt.

In Abbildung 5.9 sind zu den Betriebskosten auch die Replikationskosten und die damit sich ergebenden Gesamtkosten dargestellt. Die Replikationskosten steigen mit dem Energiebedarf zur Datenübertragung linear und stellen den Großteil der Gesamtkosten dar. Nur bei einem geringen Energiebedarfskoeffizienten von maximal 0,01 kWh/GB und einem Optimierungsin-

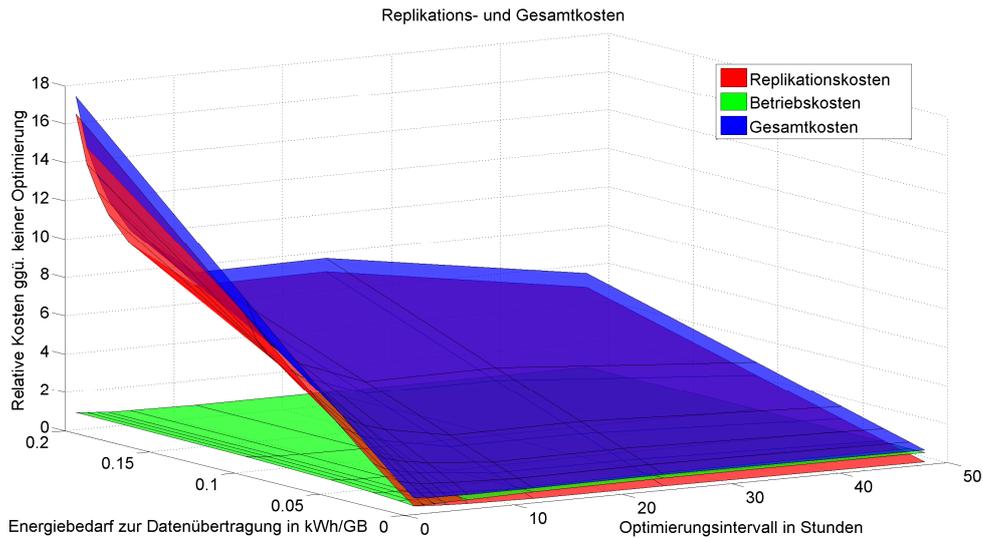


Abbildung 5.9: Einfluss der Datenübertragungskosten und des Optimierungsintervalls auf die Replikations- und Gesamtkosten bei vollständiger Replikation

tervall von mindestens sechs Stunden liegen sie unter den Betriebskosten. Dargestellt sind die Kosten für die vollständige Replikation. In den meisten Fällen ist eine vollständige Replikation somit nicht ökonomisch. Soweit diese also nicht aus anderen Gründen notwendig ist, sollte mit dem vorgestellten Standortselektionsverfahren eine kostenoptimierende Verteilung der Replikate gefunden werden.

Zur Ermittlung der besten Replikate wurden die Ordnungsmetriken UV3 und REP herangezogen. Zwar hat sich UV3 als die beste Metrik herausgestellt, allerdings liegt bei ihr der Fokus auf den Betriebskosten und bei der Metrik REP auf den Replikationkosten. Es bestand somit die Annahme, dass REP bessere Ergebnisse liefert, wenn die Kosten für Replikationen grundsätzlich höher liegen als die des Betriebs. In Abbildung 5.10 sind die minimierten Gesamtkosten für beide Metriken zusammen mit dem Ergebnis der vollständigen Replikation bei einem Optimierungsintervall von drei Stunden angegeben. Hieran lässt sich erkennen, dass die Gesamtkosten deutlich reduziert werden können.

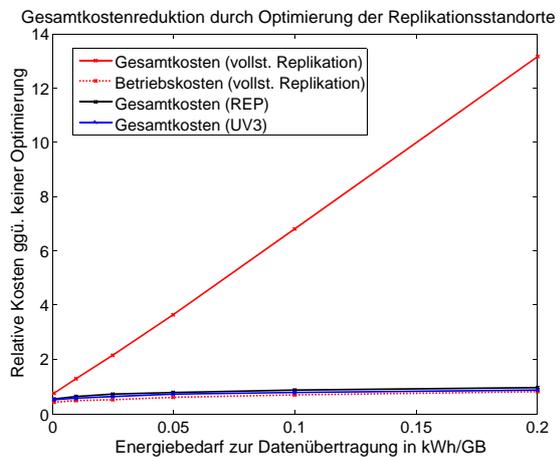


Abbildung 5.10: Einfluss der Datenübertragungskosten bei optimierter Standortselektion (1)

Zum Vergleich der beiden Heuristiken bietet sich die Abbildung 5.11 besser an, in der auf die Darstellung der Gesamtkosten bei vollständiger Replikation verzichtet wird. Einspareffekte lassen sich auch bei einem hohen Energiebedarfskoeffizienten erzielen. Die Ordnungsmetrik

UV3 zeigt sich über den gesamten Bereich als die überlegene.

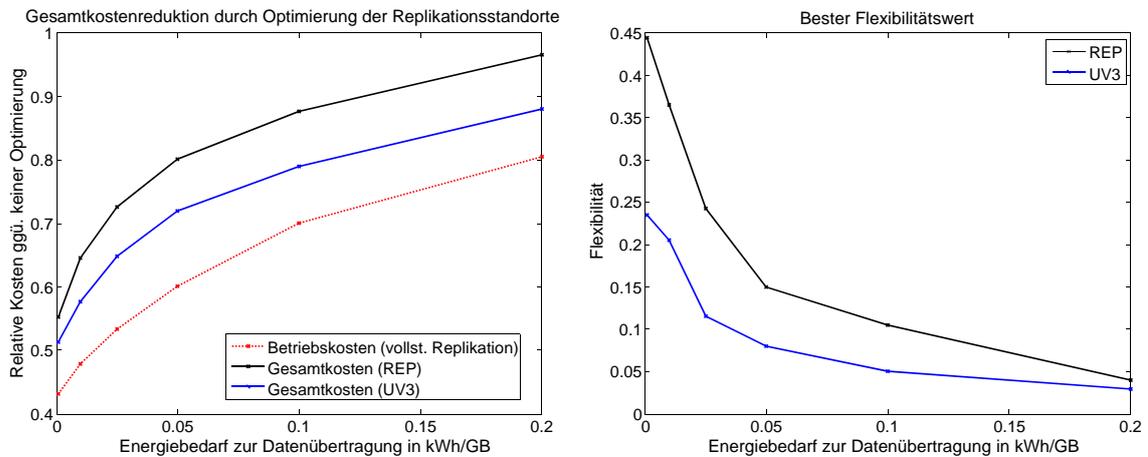


Abbildung 5.11: Einfluss der Datenübertragungskosten bei optimierter Standortselektion (2)

In der Abbildung 5.11 rechts ist dargestellt, welche Flexibilität zur jeweils gesamtkostenminimierenden Replikatverteilung führt. Bei einem hohen Energiebedarfskoeffizienten sind die Datenübertragungskosten relativ hoch, so dass vergleichsweise wenige Replikate erstellt werden. Je geringer der Energiebedarf ist, desto mehr Replikate können angelegt werden, was zur Senkung der Betriebskosten führt und somit auch die Gesamtkosten verringert.

Da sich die Ordnungsmetrik UV3 als grundsätzlich überlegen gezeigt hat, wird im Folgenden auf die Optimierung mittels der Ordnungsmetrik REP verzichtet.

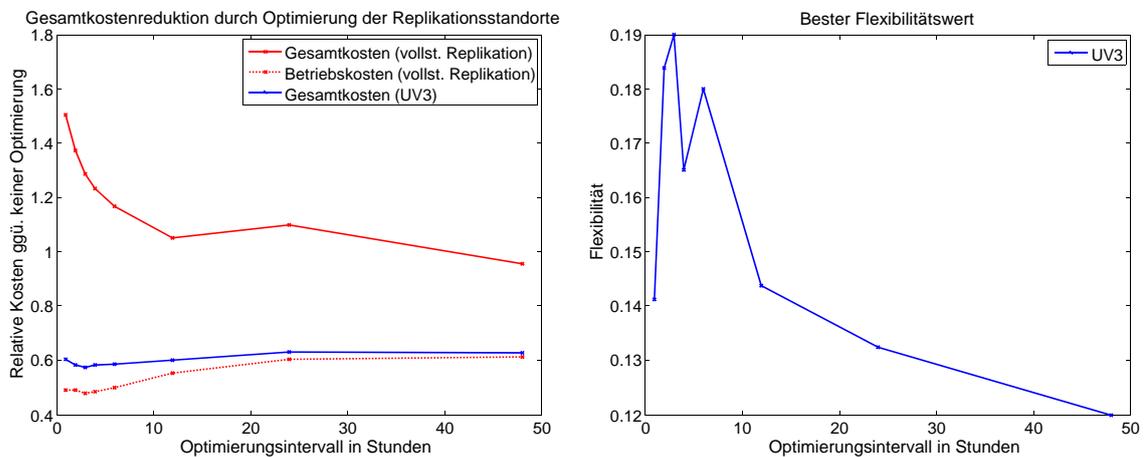


Abbildung 5.12: Einfluss des Optimierungsintervalls bei optimierter Standortselektion

In Abbildung 5.12 ist der Einfluss des Optimierungsintervalls bei optimierter Standortselektion dargestellt. Hier ist der Energiebedarfskoeffizient 0,01 kWh/GB zugrunde gelegt. Auch hier zeigt sich deutlich, dass durch einen relativ geringen Replikationsaufwand hohe Kosteneinsparungen erzielt werden können. Die Differenz zu den Betriebskosten bei vollständiger Replikation ist bei kurzen Optimierungsintervallen allerdings höher als bei längeren. Dies liegt

vornehmlich daran, dass bei vollständiger Replikation gerade bei kleineren Intervallen, das Einsparpotential sehr gut ausgeschöpft werden kann. Ein kürzeres Optimierungsintervall kommt zwar auch bei Berücksichtigung der Replikationskosten der Minimierung der Gesamtkosten zugute, allerdings nicht in gleichem Umfang, da nicht alle Standorte genutzt werden können.

5.2.6 Freie Rechenzentrumskapazität

Je mehr freie Ressourcen ein Rechenzentrum besitzt, desto mehr VMs kann es zusätzlich aufnehmen. Dies vergrößert grundsätzlich das Verlagerungspotential und damit auch das Einsparpotential, wie es in Abbildung 5.13 zu sehen ist. Untersucht ist hier wieder das Szenario mit sechs Rechenzentren und 600 VM. Der Energiebedarf für die Datenübertragung ist auf 0,01 kWh/GB und das Optimierungsintervall auf drei Stunden festgelegt.

Wenn die freie Kapazität sehr groß ist, werden relativ wenige VMs betrieben. Das bedeutet, dass das Potential zum Verschieben aufgrund der geringen Anzahl an Lastobjekten mäßig ausfällt. In dem untersuchten Szenario besteht ein gewisser statischer Energiebedarf, welcher in diesem Fall stärker zur Geltung kommt. Das relative Einsparpotential sinkt daher bei sehr großer freier Kapazität wieder ab.

In Abbildung 5.14 sind die dazugehörigen Replikations- und Gesamtkosten abgebildet. Die vollständige Replikation zeigt sich hier unter keiner Parameterbelegung als ökonomisch sinnvoll. Erst die Anwendung der Standortselektionsoptimierung ermöglicht einen ökonomisch verbesserten Betrieb (siehe Abbildung 5.15). Wie auch bei der Untersuchung des Optimierungsintervalles steigt der Abstand zwischen den Gesamtkosten im optimierten Fall und den Betriebskosten bei vollständiger Replikation an, wenn die Kosteneinsparung grundsätzlich größer wird. Dies geschieht hier aus demselben Grund: Das mögliche Verschiebe- und Einsparpotential steigt und kann bei vollständiger Replikation besser genutzt werden als bei einer Teilreplikation.

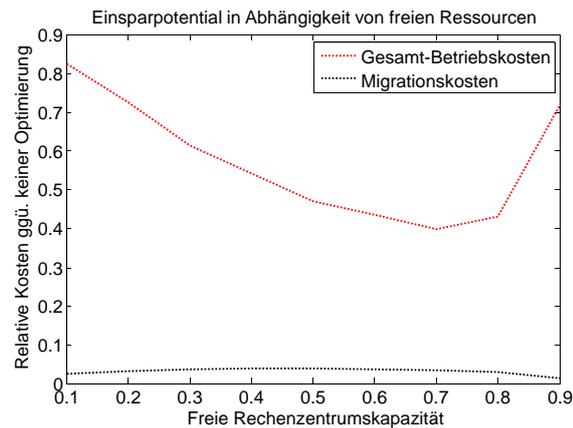


Abbildung 5.13: Einfluss des Grades freier Ressourcen

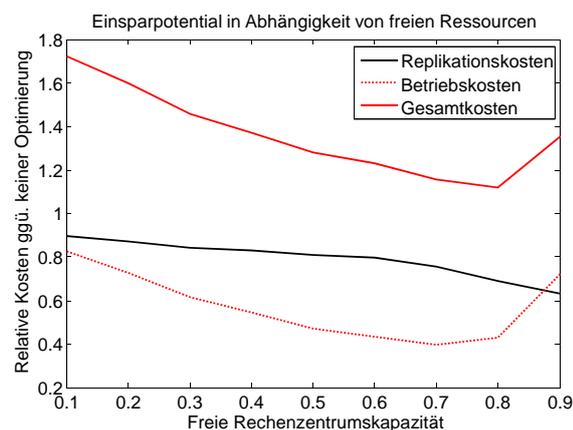


Abbildung 5.14: Einfluss des Grades freier Ressourcen auf die Replikations- und Gesamtkosten bei vollständiger Replikation

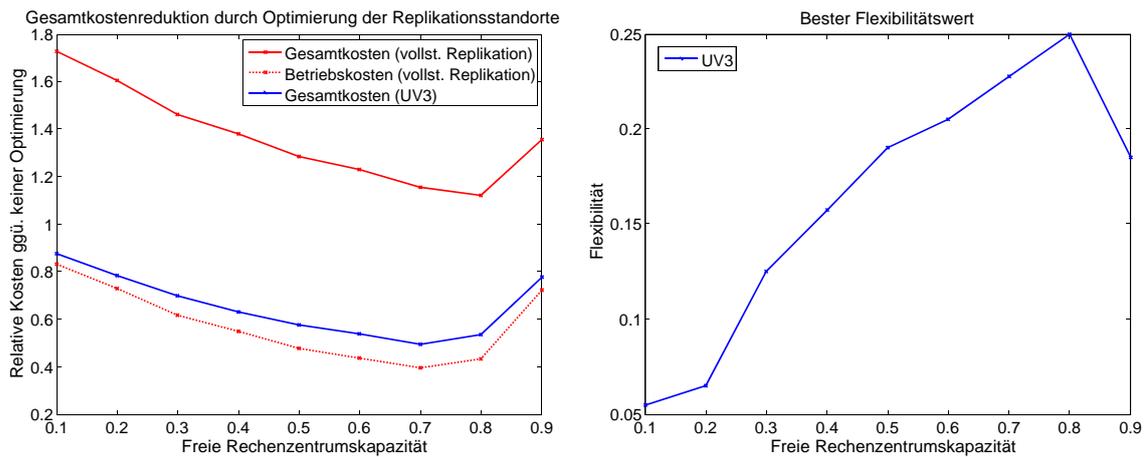


Abbildung 5.15: Einfluss des Grades freier Ressourcen bei optimierter Standortselektion

5.2.7 Standortanzahl

In dieser Untersuchung wird die Standortanzahl variiert. Die übrigen Szenarioeinstellungen entsprechen den vorangegangenen: sechs Rechenzentren, 600 VM, drei Stunden Optimierungsintervall, 0,01 kWh/GB sowie 50 % freie Rechenzentrumskapazität. Da sowohl die VM-Anzahl als auch die freie Kapazität gleich bleibt, bedeutet die Erhöhung der Standortanzahl zugleich auch eine Verkleinerung der einzelnen Rechenzentren bzgl. ihrer Kapazität. Würden die Kapazitäten gleich bleiben, wäre dies gleichbedeutend mit der Vergrößerung der freien Kapazität, was allerdings bereits oben untersucht wurde.

Aus Abbildung 5.16 lässt sich entnehmen, dass der Einspareffekt bis zu einer Standortanzahl von sechs ansteigt, danach aber nahezu konstant bleibt. Die Kurven für die Replikations- und Gesamtkosten, zu sehen in Abbildung 5.17, nehmen einen quadratischen Verlauf an. Der Verlauf der Replikationskostenkurve entspricht somit dem, der beim Durchlaufen unterschiedlicher Flexibilitätswerte entsteht (siehe z. B. Abbildung 4.36). Dies ist nachvollziehbar, da die Flexibilität ein normiertes Maß für die Anzahl an Replikaten ist, welches bei fester VM-Anzahl auch auf eine Standortanzahl abgebildet werden kann. Bei einer geringen Standortanzahl (bis zu vier) können Einsparungen selbst bei vollständiger Replikation erzielt werden.

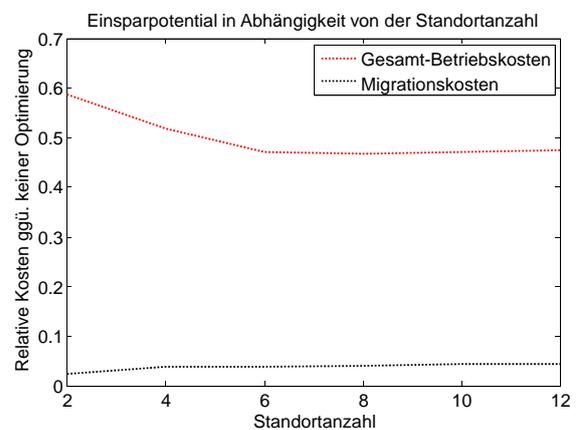


Abbildung 5.16: Einfluss der Standortanzahl

In Abbildung 5.18 ist das Ergebnis der Standortselektionsoptimierung dargestellt. Wie auch bei den vorangegangenen Untersuchungen zeigt sich, dass sich mit vergleichsweise geringen Replikationskosten hohe Gesamteinsparungen erzielen lassen. Der Flexibilitätswert fällt mit

5 Gesamtevaluation und Analyse

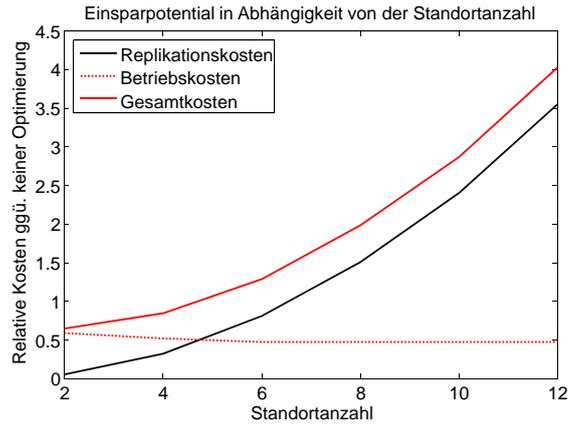


Abbildung 5.17: Einfluss der Standortanzahl auf die Replikations- und Gesamtkosten bei vollständiger Replikation

steigender Standortanzahl hyperbelförmig. Die absolute Replikatanzahl steigt jedoch von ca. 330 auf ca. 530 an.

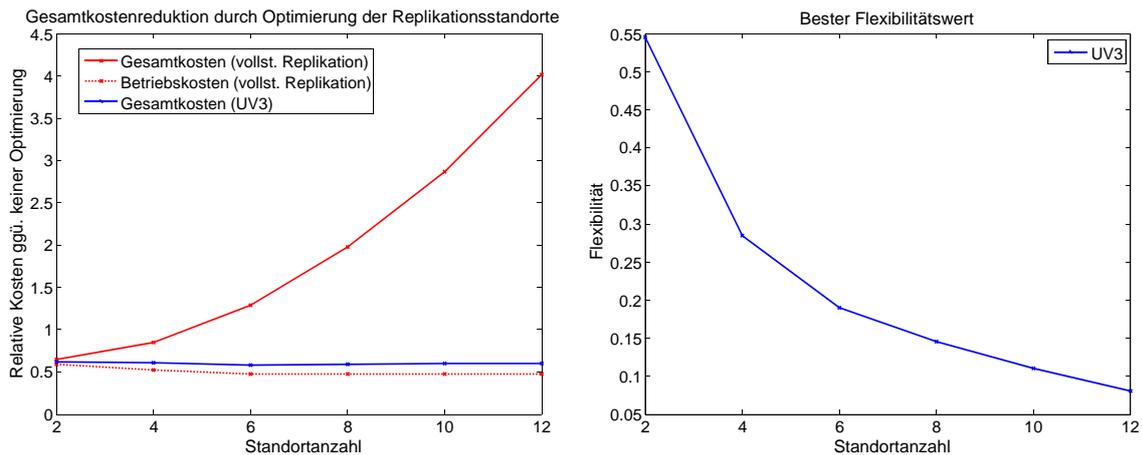


Abbildung 5.18: Einfluss der Standortanzahl bei optimierter Standortselektion

5.3 Diskussion

Die Sensibilitätsanalyse hat gezeigt, dass die Höhe des Einsparpotentials bei Nutzung eines globalen LPM von vielen Parametern abhängig ist. Grundlage für einen erzielbaren Einspar-effekt sind die unterschiedlichen Betriebsvoraussetzungen an den Standorten. Dies bezieht sich sowohl auf die Außentemperatur als auch auf den Strompreis. Dabei ist es notwendig, dass sich diese über die Zeit ändern, so dass sich eine Verlagerung der Lasten überhaupt als sinnvolle Maßnahme anbietet. Abhängig von der Änderungsrate sollte ein adäquates Optimierungsintervall gewählt werden, damit die Änderungen entsprechend genutzt werden können.

Der erzielbare Einspareffekt wird durch VM-Migrationen erkauft: Jede Migration ist mit einem Energieaufwand und damit auch mit Kosten verbunden. Durch einen hohen Energiebedarf für die Datenübertragung werden die Anzahl an Migrationen und zugleich die theoretisch erzielbaren Einspareffekte reduziert. Umgekehrt wird bei einem kleinen Energiebedarfskoeffizienten die Hürde für Migrationen verringert, so dass das Einsparpotential ansteigt.

Neben dem Energiebedarf für Migrationen wird der Einspareffekt auch durch die Migrationsgeschwindigkeit beschränkt. Bei einer 10 Gbit/s-WAN-Anbindung sind in etwa vier globale Migrationen pro Minute möglich. Das lokale LPM verursacht eine Verzögerung um knapp 50 %. Gerade bei kurzen Optimierungsintervallen führt dies zu einem hohen Einbruch am möglichen Einsparpotential. Um die optimalen Verteilungen schnell einzunehmen und damit möglichst viel vom theoretischen Einsparpotential heben zu können, sollte daher neben der Netzanbindung insbesondere die Bearbeitung von globalen Migrationen durch das lokale LPM beschleunigt werden.

Eine weitere beschränkende Größe wird durch die Replikationen gegeben. Die hierdurch verursachten Kosten sind i.d.R. höher als die, die durch Migrationen verursacht werden. Insbesondere bei Verwendung von vollständiger Replikation können zusätzliche Kosten im Bereich von Vielfachen der Ausgangslage entstehen. Erst bei einem sehr geringen Energiebedarf zur Datenübertragung stellt sich auch die vollständige Replikation als ökonomisch dar. In allen anderen Fällen sollte eine Teilreplikation eingesetzt werden.

Die für die Teilreplikation geeigneten Replikate können mittels des entwickelten Standortselektionsverfahrens ermittelt werden. Durch dieses werden solche Replikate identifiziert, die maßgeblich zur Kosteneinsparung genutzt werden können. In allen untersuchten Fällen, also auch bei hohem Energiebedarf zur Datenübertragung, konnten hierdurch Kosteneinsparungen gegenüber der Ausgangslage erzielt werden.

Bei den ermittelten Einsparwerten ist zu berücksichtigen, dass sie anhand der Prognosewerte abgeleitet wurden. Eine statistische Betrachtung über die Fehlerwahrscheinlichkeit wurde nicht durchgeführt. Daher wird in einem realen Umfeld von geringeren Einsparwerten ausgegangen.

5.3.1 Standort- bzw. Verbundoptimierung

Beim Bau eines neuen Rechenzentrums sind die räumlichen Kapazitäten darauf ausgelegt, dass eine Aufstockung der IT-Systeme über die Jahre möglich ist. Dies bedeutet, dass die Kapazitäten anfangs nicht vollkommen ausgeschöpft werden. Aber nicht nur die räumlichen Kapazitäten, sondern auch die gesamte Infrastruktur ist überdimensioniert, damit bei nachträglichen IT-Erweiterungen die notwendigen Voraussetzungen für die Stromversorgung und Kühlung gegeben sind. Grundsätzlich ist dieses Vorgehen nicht kosteneffizient. Ökonomisch sinnvoller wäre es, genau so viel Hardware und Infrastruktur bereitzustellen, wie es aktuell zum Betrieb notwendig ist. Nun ist es natürlich nicht möglich, ein bestehendes Rechenzentrum beliebig größer werden zu lassen und die Infrastrukturkomponenten gleichsam höher skalieren zu lassen. Und wenn die Rechenzentrumskapazitäten erschöpft sind, muss sogar ein komplett neues errichtet werden, das allerdings wieder nur gering ausgelastet sein dürfte.

Durch Einführung von Containerrechenzentren hat sich eine Rechenzentrumsklasse gebildet, welche mit vergleichsweise geringen Kosten Erweiterungen ermöglicht. Auch bei diesen ist i.d.R. eine anfängliche Überdimensionierung vorhanden. Jedoch ist diese aufgrund der kompakteren Bauweise vergleichsweise gering. Mit einem Containerrechenzentrum, das ein geschlossenes Rechenzentrumssystem darstellt, kann somit auf feingranularere Weise der IT-Bedarf und die dafür nötige Infrastruktur aufgestockt werden.

Hierdurch entsteht eine Flexibilität, die auch beim globalen LPM gezielt genutzt werden kann. Denn sowohl die Untersuchung zur freien Kapazität als auch die zur Standortanzahl zeigen, dass eine Erhöhung dieser Parameter für eine bessere Kostenminimierung genutzt werden kann. Der Effekt ist bei freien Kapazitäten besonders hoch. Werden mehrere Standorte zu einem Verbund zusammengeschlossen, sollte also darauf geachtet werden, dass die Standortparameter miteinander synergieren und ausreichend freie Kapazitäten vorhanden sind, um diese Synergien auch nutzen zu können.

Das (flexible) Vorhalten freier Ressourcen geht selbstverständlich mit erhöhten Kosten bei der Beschaffung einher. Wird jedoch mit einkalkuliert, dass mit der Zeit eine Aufstockung sowieso notwendig sein wird, würde es sich hierbei nicht um Mehrkosten handeln (von Kreditzinsen abgesehen), sondern um vorgezogene Aufwendungen. Allerdings hat die erstandene Hardware einen geringeren Wert, da sie im Vergleich zum späteren Kauf veraltet sein wird. Je nach Adaptivität der Komponenten ist auch noch mit einem erhöhten statischen Grundenergiebedarf zu rechnen, welcher gegen die erzielbaren Einspareffekte gerechnet werden muss.

Die Größendimensionierung und die konkrete Platzierung von Rechenzentren haben also einen signifikanten Einfluss auf die Gesamtkosten einer föderierten Cloud. Dies wurde auch bereits von Greenberg und weiteren [92] festgestellt. Welche Konstellation im Endeffekt die kostengünstigste ist, muss im Einzelfall ermittelt werden und ist damit auch relevant für die Wettbewerbsfähigkeit. Sowohl bei der Erweiterung eines bestehenden Verbundes von Rechenzentren als auch bei der Erschließung eines solchen kann das entwickelte System zur Planung unterstützend herangezogen werden. In Simulationen können die Auswirkungen von Veränderungen im Vorfeld ermittelt und dadurch die am besten geeigneten Maßnahmen identifiziert werden.

5.3.2 Bepreisung & Profil-Klassifizierung von virtuellen Maschinen

Der gezielte Ausbau eines Rechenzentrumsverbundes oder die Bildung eines solchen anhand bereits gegebener Rechenzentren lässt sich zu einem kosten-vergünstigenden Betrieb nutzen. Dies hat direkten Einfluss auf die Betriebskosten von VMs, wodurch sich entsprechende Wettbewerbsvorteile generieren lassen. Wie sich diese im Konkreten ergeben, soll in diesem Abschnitt beleuchtet werden, indem einführend die aktuelle Bepreisung von VMs vorgestellt wird und im Anschluss Aspekte bei der Erstellung von Kostenmodellen diskutiert werden.

Cloud-Provider wie Amazon oder RackSpaceCloud bieten Rechenleistung gekapselt innerhalb einer VM an. Die Kosten variieren je nach gewählter VM-Klasse, die sich in der Anzahl an Rechenkernen sowie Größe des Arbeits- und Festplattenspeichers unterscheiden. Abgerechnet

wird für diskrete Zeitintervalle – bei diesen beiden Anbietern im Stundentakt.⁴ Weitere Kosten entstehen für Datenübertragungen beim Überschreiten eines Inklusivkontingents.

Die verwendeten Kostenmodelle beziehen sich auf eine feste Leistungsspezifikation. Der Cloud-Provider selbst sorgt dafür, dass diese Leistung erbracht wird. Ob die tatsächlich abgerufene Leistung diese Kapazität ausschöpft oder nur einen Bruchteil davon darstellt, spielt bei der Kostenansetzung keine Rolle. Es stellt eine Herausforderung dar, ein Kostenmodell zu entwickeln, bei dem die tatsächlich verursachten Kosten für den Betrieb bepreist werden, denn hierzu müssten die Kosten auf faire Weise den einzelnen Verursachern, hier den VMs zugeordnet werden. Eine solche Zuordnung soll in einem ersten Schritt auf Serverebene betrachtet werden.

In [72] stellen Dhiman und weitere für die auf einem Server betriebenen VMs eine Zuordnung zum verursachten Energiebedarf vor. Hierbei wird zwischen einem statischen und einem dynamischen Anteil am Energiebedarf unterschieden. Ihre und die in [122] durchgeführten Untersuchungen fokussieren die Zuordnung und insbesondere die Ermittlung der dynamischen Anteile. Der Energiebedarf des Servers ist messbar, die Auslastungsanteile der VMs lassen sich aus dem Hypervisor entnehmen. Wie beim Serverenergiebedarfsmodell ist hier der CPU-Wert maßgeblich für die Anteilsberechnung. Damit lässt sich der dynamische Anteil anhand des Auslastungswerts zuordnen.

Für ein Kostenmodell dieser Abstraktionsschicht stellt jedoch die Zuordnung der statischen Anteile die größere Herausforderung dar. Wie sollte diese aufgeteilt werden? Entsprechend dem Auslastungsgrad (wie beim dynamischen Anteil) oder zu gleichen Teilen, da jede VM für die bloße Bereitstellung des Servers heranzuziehen ist? Für diese beiden Fälle ergeben sich zwei unterschiedliche Modelle zur Berechnung des Energiebedarfs für eine VM vm_i , ein proportionales und ein pauschales:

$$power_{proportional,vm_i}(t) = load_{vm_i}(t) \cdot (power_{static}(t) + power_{dynamic}(t)) \quad (5.1)$$

$$power_{flat,vm_i}(t) = \frac{1}{n} \cdot power_{static}(t) + load_{vm_i}(t) \cdot power_{dynamic}(t) \quad (5.2)$$

Die Angabe der Auslastung ergibt sich aus dem Anteilsquotienten, welcher sich anhand der Hypervisorangaben zu den CPU-Auslastungen der einzelnen VMs berechnen lässt:

$$load_{vm_i}(t) = \frac{CPU_{vm_i}(t)}{\sum_{j=1}^n CPU_{vm_j}(t)} \quad (5.3)$$

n gibt an, wie viele VMs aktuell auf dem Server betrieben werden.

An einem Server, welcher mit 13 VM betrieben wurde, konnten Messungen zum Energiebedarf durchgeführt und die Auslastungen der VMs bestimmt werden. Der Gesamtenergiebedarf des Servers betrug im Messzeitraum 6,7 kWh. Wie in Abbildung 5.19 zu sehen ist, ordnen die beiden VM-Energiebedarfsmodelle den einzelnen VMs sehr unterschiedliche Werte zu. Dies

⁴<http://t3n.de/magazin/uberblick-cloud-hosting-flexible-datenspeicher-232775/2/>, letzter Zugriff: 01.09.2014

5 Gesamtevaluation und Analyse

liegt an dem sehr hohen Anteil des statischen Energiebedarfs, der bei diesem Server bei etwa 65 % seines maximalen Energiebedarfs liegt. Je geringer dieser Anteil ist, desto geringer fällt auch der Teil ins Gewicht, welcher in den beiden Modellen unterschiedlich ist. Die Angleichung würde in Richtung des proportionalen Modells erfolgen.

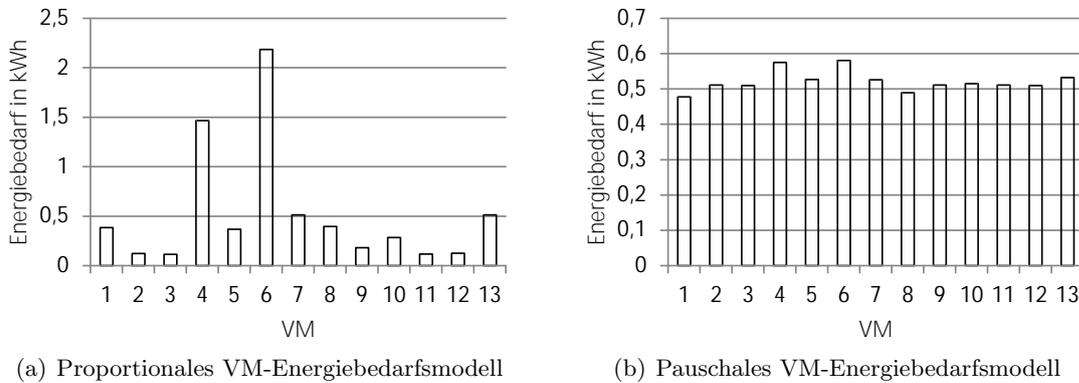


Abbildung 5.19: Energiebedarfe unterschiedlicher VMs

Werden diese Modelle auf eine Cloud übertragen, so ist das Bezugssystem, aus dem ein konkreter Energiebedarf ermittelt werden kann, nicht mehr der einzelne Server, sondern die Cloud selbst, also alle in ihr betriebenen Bestandteile, um die Dienste bereitzustellen, inklusive der lokalen und globalen LPM-Systeme. In den Abschnitten 4.2 und 4.4 wurde ein Rechenzentrumsenergiebedarfsmodell vorgestellt, mit dem anhand der VM-Auslastungsdaten der Gesamtenergiebedarf abgeschätzt werden kann. Da sich das Modell jedoch auf ein Rechenzentrum beschränkt, sind die Migrations- und Replikationskosten nicht abgebildet. Diese werden allerdings durch das globale LPM in der Gesamtkostenkalkulation berücksichtigt, so dass der globale Betrieb i.d.R. günstiger ist als der isolierte Betrieb jedes einzelnen Rechenzentrums. An dieser Stelle ist zu unterscheiden, ob ein gewisser Energiebedarf den einzelnen VMs zugeordnet werden soll oder die verursachten Kosten. Da sich die Eingangsfragestellung auf die Bepreisung bezieht, wird im Folgenden dieser Fall weiter diskutiert.

Die Kosten, die im isolierten Betrieb der Rechenzentren entstehen, stellen eine obere Grenze dar. Anhand des Energiebedarfsmodells können die Kostenanteile der einzelnen VMs bestimmt werden, wobei der statische Anteil auf dieser Systemebene eine untergeordnete Rolle spielt. Dies lässt sich aus den Regressionsmodellen ableiten, denn hier macht der statische Anteil einen vergleichsweise geringen Part aus. Die durch das globale LPM erzielten Einsparungen können entsprechend der gleichen Anteilsverrechnung an die VMs weitergegeben werden. In einem eingeschwungenen System können so die verursachten Kosten bestimmt und die gebotenen Leistungen entsprechend bepreist werden.

Die realen Preise ergeben sich damit erst während des Betriebs. Dies scheint zwar in gewisser Hinsicht fair zu sein, jedoch ist dies für einen Kunden nicht planbar. Ihm müssen Preise bzw. maximal anfallende Kosten zugesichert werden. Entweder werden hierzu die Kosten des isolierten Betriebs genommen, was keinen Wettbewerbsvorteil verspricht, oder das erzielbare

Einsparpotential wird mit einberechnet. Hierbei ist zu berücksichtigen, dass die Aufnahme von zusätzlichen VMs den Anteil freier Kapazität verringert und damit den Gesamteinspareffekt reduziert; was ebenso durch die Erhöhung der Gleichheit der VM-Auslastungsprofile geschieht. Das heißt, dass der Betrieb zusätzlicher VMs den Betrieb der aktuellen VMs evtl. verteuert.

Bei der Erstellung von Preiszusagen ist somit entweder ein entsprechender Puffer mit einzuberechnen oder der Verbund muss ausgebaut werden, um möglichst mit den Bedingungen betrieben werden zu können, die die höchsten Kosteneinsparungen versprechen. Ein weiteres mögliches Preismodell, welches der allgemeinen marktwirtschaftlichen Natur entsprechen würde, wäre ein Marktplatz für IT-Ressourcen, bei dem sich der Preis aufgrund der aktuellen Belastung in Folge des Angebots (Ressourcenkapazität) und der Nachfrage (bereits betriebene VMs) bilden würde.

Als Grundlage für die Berechnung einer Preiszusage benötigen die eingesetzten Energiebedarfsmodelle Informationen über den Dienst in Form prognostizierter Auslastungen. Diese liegen jedoch im Vorfeld i.d.R. nicht vor. Es müssten also spezielle Kenndaten zum zu betreibenden Dienst gefunden werden, die den Kostenanteil abschätzbar machen. Diese Kenndaten sollten statistisch einfach erhebbar sein, wie z. B. die erwartete maximale Auslastung, die Durchschnittsauslastung (\emptyset) und evtl. auch die Standardabweichung der Auslastungswerte (σ).

In einer Evaluation, in der 100 unterschiedliche VM-Mengen mit einer Größe von jeweils 500 VM für einen Zeitraum von sieben Tagen simuliert wurden, hat sich gezeigt, dass diese Kenndaten sehr gut dazu geeignet sind, den Gesamtenergiebedarf abzuschätzen. Die in Tabelle 5.1 angegebenen Ergebnisse für entsprechende Regressionsmodelle zeigen, dass bereits durch die bloße Verwendung der maximalen Auslastung der Energiebedarf den einzelnen VMs sehr gut zugeordnet werden kann.

	Modell basierend auf Auslastungsdaten		
	Max.	Max.+ \emptyset	Max.+ \emptyset + σ
Durchschnittliche relative Abweichung	3,5 %	0,9 %	0,8 %

Tabelle 5.1: Güte der Energiebedarfszuordnung für einen Zeitraum von 7 Tagen

Für einen Cloud-Provider kann also der in der aktuellen Konstellation erwartete Energiebedarf als Grundlage genommen werden, um eine Abschätzung zum anteilmäßigen Energiebedarf weiterer Dienste geben zu können. Da sich die Abschätzung auf einen längeren Zeitraum bezieht und somit einen Durchschnittswert darstellt, können zwischenzeitlich auch höhere Energiebedarfe vorliegen. Hier gilt es, adäquate Kompensationsmodelle einzusetzen, z. B.: Der Preis reduziert sich mit der Länge der Teilnahme approximativ bis zu einem gewissen Zielwert.

5.3.3 Globales LPM beim Demand Side Management

Grundsätzlich hat die Evaluation zur Lastganggenerierung in Abschnitt 4.5.3 gezeigt, dass ein gewünschter realer Lastgang sehr gut durch den Betrieb gezielter VMs nachgestellt und auch mit inversem Verlauf ausgeregelt werden kann. Die Güte hängt davon ab, welche VMs mit ihren entsprechenden Auslastungsprofilen vorliegen. In einem Verbund bedeutet die Verlagerung von VMs, um an einem Standort einen bestimmten Lastgang zu erreichen, dass an anderen Standorten der Lastgang abgeändert wird. Sollten an diesen ebenso bestimmte Lastgänge gewünscht sein, müsste wie bei den Strompreisen eine Gewichtung eingeführt werden, über die die Wichtigkeit eines bestimmten Zielwerts für einen bestimmten Standort und Zeitpunkt angegeben werden kann.

Wie schnell ein bestimmter Lastgang erreicht werden kann, ist u.a. von der Migrationsgeschwindigkeit abhängig. In der Sensibilitätsanalyse wurde hier das aktuell eingesetzte lokale LPM als „Flaschenhals“ identifiziert. Sollte dies beschleunigt werden können, ist die nächst beschränkende Größe die Bandbreite der Netzanbindung. Eine VM stellt hier das verschiebbare Lastobjekt dar. Sie verursacht jedoch keinen Energiebedarf, zumindest nicht direkt, denn erst durch den Betrieb einer VM auf einem Server wird Energie benötigt. Die Lastverlagerungsgeschwindigkeit hängt daher auch von der Leistungsaufnahme der Server ab.

Zur Bestimmung der verschiebbaren Leistungsaufnahme im gesamten Rechenzentrum müssen zusätzlich das Storage, das Netzwerk und die Infrastruktur-Komponenten herangezogen werden. Grundsätzlich skalieren diese zwar in ihrer Leistungsaufnahme mit der Serverauslastung, aber im Gegensatz zum Server lassen sie sich nicht ausschalten, so dass ihr Idle-Energiebedarf bestehen bleibt. Je effizienter diese Komponenten jedoch funktionieren, desto geringer fällt dieser Anteil aus. Für die hier betrachtete Berechnung wird der theoretisch bestmögliche Fall angenommen, also ohne Idle-Anteil. Die Leistungsaufnahme des Storages und des Netzwerks, die 20 % an der IT-Last ausmacht, und die der Infrastruktur-Komponenten skalieren damit vollständig mit der Serverauslastung.

In Abschnitt 2.1.2 wurde gezeigt, dass eine rechenzentrenübergreifende Migration einer VM mit 4 GB RAM bei einer 10 Gbit/s-Netzanbindung in etwa 13s dauert, wobei die genaue Dauer von der tatsächlichen Entfernung der Rechenzentren beeinflusst wird. Werden auf einem Server bspw. 24 VMs betrieben, müssen insgesamt 96 GB RAM übertragen werden, bevor der Quellserver ausgeschaltet werden kann und damit die Verschiebung der Leistungsaufnahme auf den Zielservers (vollständig) abgeschlossen ist. Ein effizienter Server aus dem Jahre 2013 ist der PowerEdge C5220 mit 48 Kernen und 96 GB RAM.⁵ Dieser benötigt unter Vollast 672 W. Das Storage und das Netzwerk würden hier dem Modell nach 168 W benötigen. Wird nun ein PUE-Wert moderner Rechenzentren von 1,3 zugrunde gelegt, der auch für das hier modellierte Containerrechenzentrum zutrifft, ergibt sich für das gesamte Rechenzentrum schließlich eine Verlagerungsgeschwindigkeit von 210 W pro Minute.

Moderne voll ausgerüstete Racks haben eine Leistungsaufnahme von 30 kW und mehr.

⁵http://www.spec.org/power_ssj2008/results/res2013q2/, letzter Zugriff: 09.09.2014

Die Verlagerung dieser Leistungsaufnahme würde damit einige Stunden in Anspruch nehmen. Damit fällt das Ergebnis deutlich geringer aus als in einer eigenen im Jahre 2009 erstellten Studie [162]. Basierend auf den damals aktuellen, – im Vergleich zu heute – ineffizienten Servern konnte eine Verlagerungsgeschwindigkeit von 3 kW/Min. erreicht werden.

In [232] wurde die Beschleunigung der Live-Migration durch eine schnellere Netzwerktechnologie herausgestellt. Die Nutzung einer 10-Gbit/s-Netzanbindung hat gegenüber einer 1-Gbit/s-Leitung eine Beschleunigung um das Vierfache ergeben – *nur*, muss hierbei erwähnt werden. Allgemein lässt die Weiterentwicklung der Netzwerktechnik das globale LPM bei einer Teilnahme am DSM relevanter werden. Die Weiterentwicklung von Servern wirkt dem jedoch entgegen.

Mögliche Szenarien, in denen eine Veränderung der Leistungsaufnahme notwendig werden könnte, motivieren sich aufgrund schwankender Werte bei der Stromgenerierung sowie bei der Stromabnahme. Bei der Stromgenerierung führt insbesondere die Generierung mittels Windkraft zu volatiler Verfügbarkeit. Auf Erzeugerseite lässt sich dies teilweise durch den variablen Betrieb von z. B. Gasturbinenkraftwerken ausgleichen. Eine weitere Maßnahme besteht in der Speicherung der Energie in Pumpspeicherkraftwerken. Beides weist jedoch einen vergleichsweise geringen Wirkungsgrad auf und führt damit zu einem erhöhten Primärenergiebedarf sowie erhöhtem CO_2 -Ausstoß [107]. Es wird daher eine variable Anpassung auf Verbraucherseite angestrebt, die zumindest als eine unterstützende Maßnahme zur Netzstabilisierung beitragen kann. Dies wird dem Verbraucher im Rahmen von Demand Response Programmen auch entsprechend vergütet [238, 3, 153, 27, 57, 56]. Realisiert werden die Angebote und Zusagen über den offenen Kommunikationstandard OpenADR [177].

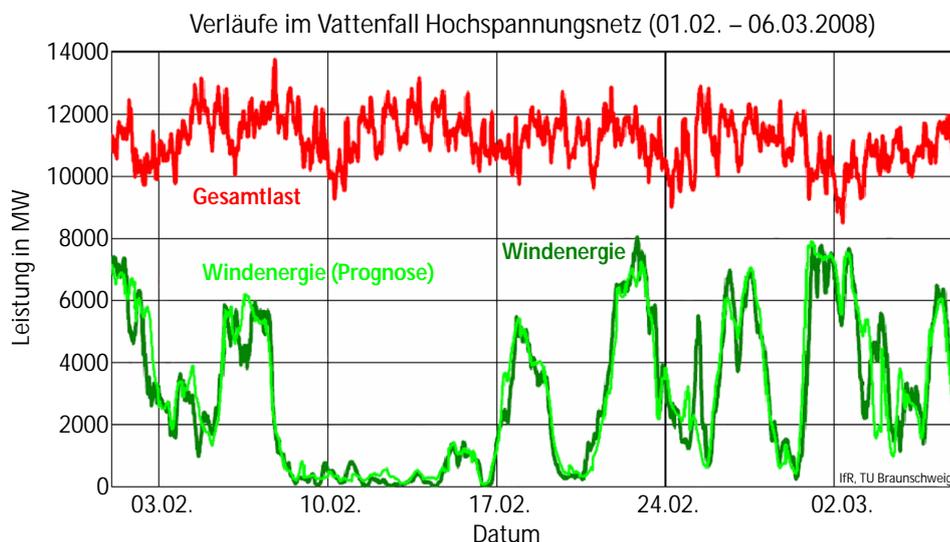


Abbildung 5.20: Die Stromgenerierung mittels Windkraft ist sehr volatil.

Für gewöhnlich geht es hierbei um die Senkung des Strombedarfs. Aber auch der erhöhte Bedarf an Strom ist für die Stabilisierung des Netzes bei sprunghaft ansteigender Generierung notwendig. In Abbildung 5.20 ist beispielhaft der Verlauf der aus Windkraft generierten

Energie angegeben. Hier entstehen mitunter hohe Generierungssprünge sowohl steigend wie auch fallend. Wird die erzeugte Windkraft jedoch über ein größeres Gebiet betrachtet, ergibt sich eine stabilere Kurve [3]. Das bedeutet, dass wenn die Erzeugung an einem Ort gerade ansteigt, diese andernorts wieder fällt. Eine Verlagerung von Lasten zu Standorten, an denen die Erzeugung aktuell hoch ist, hat, wie bereits erwähnt, den Nebeneffekt, dass die Lasten andernorts fallen. Dieses Verhalten erweist sich somit als synergetisch. Zwar ließe sich der Strom ebenso verlagern, allerdings ist dies mit höheren Kosten verbunden als die Verlagerung von IT-Lasten [136].

Die Gesamtlast der Verbraucher stellt auch bereits eine hohe Volatilität dar. In Abbildung 5.21 ist beispielhaft die Last der Verbraucher für eine Siedlung mit ca. 100 Haushalten und zwei Industrieunternehmen angegeben. Der hier dargestellte tägliche Verlauf hat einen sehr ausgeprägten periodischen Charakter, weshalb eine Prognose dieses Verlaufes eine höhere Güte besitzt als bei Windkraft. Ohne Verwendung von Prognosen müsste spontan (in stärkerem Maße) versucht werden, die Leistungsaufnahme zu glätten. Die mittlere Veränderungsrate liegt in diesem Beispiel bei 400 Watt/Min. Befinden sich zwei wie oben beschrieben ausgestattete Rechenzentren in der Siedlung, so könnte der Verlauf teilweise oder sogar komplett geglättet werden. Bei Verwendung von Prognosen, die einen Tag oder länger in die Zukunft reichen, besteht genügend Zeit die notwendigen Migrationen durchzuführen, um die Vorgabe eines Stromverbrauchsprofils nachzustellen.

Diese allgemeine Anpassung des Energiebedarfs mittels DSM bzw. Demand Response entspricht der Definition vom VDE [12]. Dem gegenüber steht die Auffassung, dass Demand Response nur zum Zwecke der Spitzenlastreduktion dient [210], was z. B. auch in [238] aufgegriffen wird. Dort wird davon ausgegangen, dass der Strombedarf über mehrere Stunden gesenkt werden können muss. Hierfür habe der Rechenzentrumsverbund ebenso mehrere Stunden Zeit, um entsprechende Anpassungen vorzunehmen. Wie in der Dissertation werden in der dortigen Analyse VMs migriert. Die verwendeten Spezifikationen für Server, auf denen VMs betrieben werden, sind jedoch unüblich bis unrealistisch: Es werden Server mit maximal 8 GB Arbeitsspeicher betrachtet. Der zu übertragende Datenbestand ist damit sehr gering, wodurch die Verlagerungsgeschwindigkeit gesteigert und die Kosten von Migrationen gesenkt werden.

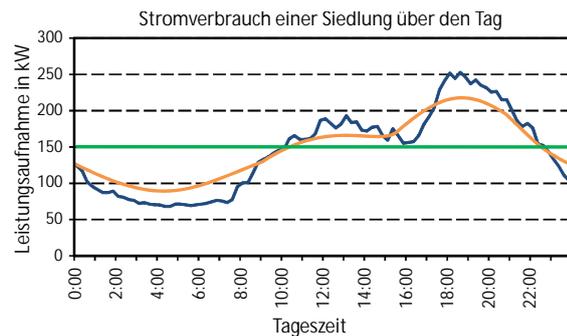


Abbildung 5.21: Beispiel für die Leistungsaufnahme einer Siedlung (blau),⁶ mit geglättetem Verlauf (Vermeidung von Peaks) (orange) und vollständiger Glättung (grün)

⁶Datenquelle: Erhebung im Rahmen eines Kooperationsprojektes zwischen HTEE der TU Braunschweig und BSenergie

Je nach konkreter geografischer Migration hat eine Lastverlagerung aber nicht nur Einfluss auf die regionalen Lasten, sondern auch der Strompreis würde sich dadurch ändern. In [127] wird untersucht, welche Auswirkungen die Effekte von DSM auf den Elektrizitätsmarkt haben. Dazu wird zuerst abgeschätzt, welche Strombedarfselastizität sich bildet. Mit Hilfe dieser Schätzung wird ein neuer Strompreis für den Day-Ahead-Handel berechnet. Allgemein wurde ermittelt, dass der Einsatz von DSM den Strompreis signifikant reduziert. Dies bezieht sich nicht nur auf die Teilnehmer am DSM, sondern auch für alle anderen Endverbraucher reduziert sich der Strompreis. Eine Abschätzung der Elastizität bietet darüber hinaus den Vorteil, unvorhersehbares Verhalten beim Real-Time-Pricing einzudämmen, denn aufgrund fehlender Synchronisierung unter den Teilnehmenden können unerwartete Lastspitzen entstehen, weil auf Preissignale ähnlich reagiert wird [185].

5.4 Zusammenfassung

Die entwickelten Verfahren ergeben als Gesamtsystem ein globales LPM, das eine Flexibilität schafft, um sich an die standortbedingten, variierenden Außentemperaturen und Strompreise anzupassen. In einer Simulationsumgebung wurde das Gesamtsystem hinsichtlich des erzielbaren Einsparpotentials evaluiert. Es konnte anhand unterschiedlicher Szenarien gezeigt werden, dass sowohl der Energiebedarf als auch die Gesamtkosten um bis zu 10 % bzw. 40 % gesenkt werden können – auch unter Berücksichtigung von Migrations- und Replikationkosten.

Die Auswirkungen von Veränderungen der Randbedingungen können zuvor simuliert werden. Dadurch lassen sich bestmögliche Maßnahmen zur Optimierung des Betriebs bereits im Vorfeld identifizieren.

Als weiteren Punkt wurden Aspekte für ein Kostenmodell von VMs diskutiert, die auf den Einspareffekt durch den Betrieb in einem Rechenzentrumsverbund eingehen. Außerdem konnten die in der Literatur bereits ermittelten Potentiale für die Partizipation am DSM grundsätzlich bestätigt und anhand der hier vorliegenden Datenbasis aktualisiert werden.

6 Fazit

Die Flexibilität von Cloud-Rechenzentren bzgl. der Ressourcenprovisionierung ermöglicht es, Lasten in Form von virtuellen Maschinen (VMs) zwischen unterschiedlichen Rechenzentrumsstandorten zu verschieben. Hierdurch kann ein geeignetes Verfahren zum Last- und Powermanagement (LPM) auf standortbedingte Parameter wie die Außentemperatur und den Strompreis gezielt eingehen, um den Gesamtenergiebedarf bzw. -kosten eines Rechenzentrumsverbundes zu minimieren. In der Dissertation wurde ein solches Verfahren konzipiert und der erzielbare Einspareffekt sowie der Einfluss prägnanter Parameter ermittelt.

Für ein globales LPM müssen die Veränderungen der Standortbedingungen rechtzeitig im Voraus bekannt sein, was entsprechende Prognoseverfahren erfordert: Die für die Optimierungen benötigten Werte bzgl. des Ressourcenbedarfs von Diensten, der Standorttemperaturen sowie der Strompreise lassen sich für einen Prognosehorizont von wenigen Stunden bis zu einem Tag mit hoher Genauigkeit vorhersagen.

Um die Arbeit besser einordnen zu können, wurde detaillierter auf den technologischen Kontext – hier hat sich das Cloud Computing als technischer Wegbereiter herausgestellt – und auf verwandte Arbeiten eingegangen. Zu diesen wurde eine Abgrenzung geschaffen, aus der hervorging, dass es große Unterschiede im Betrachtungsraum gibt und dass einige Problemstellungen ungenügend bzw. gar nicht behandelt wurden. Diesem Umstand abhelfend wurden entsprechende Eigenschaften für das zu entwickelnde System und damit auch die Ziele der Dissertation herausgearbeitet.

Als Architektur wurde ein Multi-Agenten-System gewählt, das darauf ausgelegt ist, die Berechnungslast auf alle beteiligten Rechenzentren zu verteilen. So konnte ein globales LPM mit der Laufzeitkomplexität $O(\#VM \cdot \#DC)$ für jeden einzelnen Rechenzentrumsagenten und damit mit der Gesamtlaufzeitkomplexität $O(\#VM \cdot \#DC^2)$ realisiert werden.

Das verwendete Replikationsverfahren ist nicht vollständig in das globale LPM integriert, sondern optimiert die Replikat, die das globale LPM nutzen kann, aus einer Beobachterrolle. Hierdurch bleiben die Verfahren voneinander unabhängig und sind daher austauschbar. Dies gilt auch für die weiteren Komponenten im entwickelten System. Durch den modularisierten Aufbau wird gewährleistet, dass jeder Teilnehmer im Verbund seine eigenen Verfahren für die lokale Lastganggenerierung und insbesondere für das lokale LPM einsetzen kann. Interne Abläufe und der konkrete Aufbau eines Rechenzentrums müssen damit nicht preisgegeben werden.

Die wissenschaftlichen Fragestellungen dieser Arbeit zielen zum einen auf die Brauchbarkeit und den Nutzen eines globalen LPM ab und zum anderen darauf, welche Auswirkungen dies auf

den Betrieb von Diensten sowohl für den Betreiber als auch den Endkunden hat. Innerhalb einer simulierten Umgebung wurde das Verfahren einer Sensibilitätsanalyse unterzogen, durch die das erzielbare Einsparpotential für Szenarien mit unterschiedlichen Bedingungen bestimmt werden konnte. Für diese Szenarien bildeten reale Datenbestände für die Ressourcenauslastung von Diensten, Temperaturen unterschiedlicher Standorte, Strompreise europäischer Strombörsen und Lastgänge einer Siedlung die Grundlage.

Einsparungen von bis zu 10 % bzw. 40 % bezogen auf den Gesamtenergiebedarf bzw. die Gesamtkosten – auch unter Berücksichtigung von Migrations- und Replikationkosten – konnten ermittelt werden. Änderungen an der IT und Infrastruktur beeinflussen diesen Einspareffekt. Da die Auswirkungen solcher Änderungen durch Simulationen ermittelt werden können, lassen sich bestmögliche Maßnahmen zur Optimierung des Betriebs bereits im Vorfeld identifizieren.

Der erzielte Einspareffekt lässt sich an den Endkunden weitergeben, wofür ein geeignetes Kostenmodell von VMs notwendig ist. Die zur Erstellung eines solchen Kostenmodells zu berücksichtigenden Aspekte, speziell für den Betrieb in einem Rechenzentrumsverbund, wurden diskutiert. Es hat sich gezeigt, dass bereits anhand von wenigen Informationen über den Dienst eine adäquate Preisgestaltung ermöglicht werden kann.

Für eine Partizipation am DSM konnten die in der Literatur bereits ermittelten Potentiale – auch für Lastverschiebungen mittels Live-Migration – bestätigt werden. Der technologische Fortschritt in der Netzwerktechnik erhöht dieses Potential, wohingegen der Fortschritt in der Servertechnik das Potential aufgrund der geringer werdenden VM-bezogenen Leistungsaufnahme reduziert.

7 Ausblick

Das entwickelte Gesamtsystem zur Durchführung des globalen LPM besteht aus mehreren Verfahren, die sich auf heuristische Methoden stützen. Außerdem besitzen die entwickelten Modelle einen gewissen Modellfehler. Die hierauf basierenden Entscheidungen für globale Migrationen können daher suboptimal sein, weshalb der theoretisch erzielbare Einspareffekt größer ausfallen kann. Weitere Aspekte, welche den Einspareffekt beeinflussen (können), beziehen sich auf den eingeschränkten Betrachtungshorizont:

- In den Simulationen wurden nur homogene Rechenzentren und zur Vereinfachung mit der CPU nur ein einzelner Ressourcentyp untersucht. Durch Betrachtung von Server-Heterogenität und mehr Ressourcentypen kann sich der Einspareffekt verändern, im positiven wie auch im negativen Sinne.
- Bei den betriebenen Diensten wurde die Betrachtung auf Server-Sized-Services eingeschränkt, da diese im Themenkontext noch nicht berücksichtigt wurden. Für eine vollständige Potentialabschätzung sollten jedoch auch Large-Scale-Services hinzugezogen werden. Da ein solcher Service in seinem eigenen (virtuellen) Cluster über einen ebenso eigenen Lastmanager verfügt, muss an dieser Stelle eine Koordination zwischen diesem und evtl. mehreren lokalen LPMs erfolgen.
- Ein weiterer Aspekt bezieht sich auf die Einflussparameter. In dem betrachteten System ergeben sich die Kosten bzw. der Energiebedarf aufgrund der Standortparameter und der Ressourcenbedarfe der Dienste. Letztere sind jedoch in großen Teilen bedingt durch Clientanfragen, welche Traffic sowie Kosten im Rechenzentrum und im Netzwerk (Internet) verursachen. Clientanfragen lassen sich somit als zusätzlicher Einflussparameter für die Kosten- bzw. Energiebedarfsminimierung heranziehen. Das betrachtete System könnte hierzu mit einem detaillierteren Netzwerkmodell erweitert werden, um nicht nur die VMs und Replikate zu reallozieren, sondern auch um den Traffic geeignet zu leiten. In Content-Delivery-Netzwerken werden die Informationen über die Clientanfragen dazu genutzt, um die Latenz zu reduzieren. Von dieser wurde in der Dissertation abstrahiert. Für den Einsatz im Produktivbetrieb ist allerdings eine Berücksichtigung der Latenz aufgrund von Vereinbarungen bzgl. der Dienstegüte notwendig.

Neben einer weiter reichenden Betrachtung gibt es bei den einzelnen entwickelten Verfahren auch Anknüpfungspunkte, um diese hinsichtlich der folgend genannten Aspekte zur erweitern. Diese betreffen insbesondere das Standortselektionsverfahren für die Ermittlung geeigneter Replikate zur Minimierung der Gesamtbetriebskosten.

- Im betrachteten System werden Datenveränderungen asynchron propagiert. Dies geschieht in regelmäßigen Abständen an alle Replikate, was nicht notwendig ist, denn in Abschnitt 4.8.2 wurde gezeigt, dass Migrationen auch ohne Verzögerung durchgeführt werden können, wenn innerhalb der letzten 14,5 Std. keine Datensynchronisierung erfolgte. Hier könnte also eine On-Demand-Synchronisierung, wie sie in [214] vorgestellt ist, den Replikationsaufwand reduzieren.
- Bei der Evaluation der Ordnungsmetriken für die Bewertungen der Replikate hat sich gezeigt, dass mit der Metrik UV3 die beste Gesamtkostenminimierung erzielt werden kann, jedoch sind die Gesamtkosten nicht mit jeder Flexibilität minimal. Bei SLA bedingter Anforderung einer höheren Anzahl an Replikaten, schneidet die Metrik REP zumeist besser ab. Hier sollten weitere Untersuchungen eine adäquate Kombination schaffen, die die Vorteile beider Metriken zusammenführt.
- Durch Veränderung der Nutzungsstatistiken kann eine Neuauswahl der Replikatstandorte ratsam sein. Diese kann zwar auch mit dem entwickelten Standortselektionsverfahren getroffen werden, jedoch sind im laufenden Betrieb bereits Replikate vorhanden. Eine Reallokation dieser ist mit erhöhten Kosten verbunden, welche daher innerhalb eines erweiterten Ansatzes berücksichtigt werden müssen.

Für die globalen Migrationen hat sich das lokal eingesetzte LPM als „Flaschenhals“ herausgestellt: Es beschränkt die Migrationsgeschwindigkeiten, weil es aufgrund der lokalen Aspekte (Konsolidierung, Engpassauflösung) nur relativ langsam Veränderungen des VM-Pools verarbeiten kann. Hier sollte ein Mechanismus geschaffen werden, welcher eine nebenläufige Verarbeitung ermöglicht, ohne dabei die Eigenschaften des lokalen LPM zu beeinflussen.

Eine grundsätzliche Erweiterung sollte auch hinsichtlich der Betrachtung von Prognosefehlern erfolgen. Aktuell wird vom 100%-igen Eintreten der Prognosewerte ausgegangen. Treten jedoch hohe bzw. im Vergleich zu den Einsparungen hohe Abweichungen auf, so können die vermeintlich optimierenden Migrationen sogar zu einem Energiebedarfs- bzw. Kostenanstieg führen.

Neben den Erweiterungen sollte in nachfolgenden Untersuchungen auf eine reale Hardwareplattform für die Simulationen umgestiegen werden, um eine bessere Praxistauglichkeit zu gewährleisten. Hierzu bietet sich bspw. die Framework-Testumgebung BonFIRE¹ an. Bezüglich eines dynamischen Strompreisverhaltens wäre eine Co-Simulation mit dem PowerTAC-Simulator² anzustreben, in welchem sowohl Verbraucher, Erzeuger, sog. Pro-Sumer (Verbraucher und Erzeuger) und Broker als auch Haushalte, Gewerbe, Windparks, elektrische Autos usw. abgebildet sind.

¹www.bonfire-project.eu, letzter Zugriff: 14.08.2014

²www.powertac.org/, letzter Zugriff: 14.08.2014

Glossar und Abkürzungsverzeichnis

Cloud ist ein über das Netzwerk (Internet) erreichbarer Bereich, aus dem IT-Ressourcen und Dienste flexibel bezogen werden können. Die Flexibilität wird üblicherweise durch Techniken der Virtualisierung und des verteilten Rechnens realisiert.

Cloud Computing ist ein Servicemodell, bei dem IT-Ressourcen (*Computing*) skalierbar und nach Bedarf über ein Netzwerk (*Cloud*) bereitgestellt werden. Die Provisionierung erfolgt in nahezu Echtzeit und das Billing verbrauchsgenau.

Cloud-Provider ist der Betreiber einer Cloud und damit Anbieter der offerierten Services.

Deadlock bezeichnet einen Zustand, bei dem weitere Aktionen nicht erfolgen können, da diese zyklisch auf das Ereignis jeweils anderer Aktionen warten.

Demand Side Management umfasst laut VDE [12] die direkte Beeinflussung des Energiebedarfs von Verbrauchern, wohingegen beim Demand (Side) Response die Beeinflussung indirekt über ein Anreizsignal (z. B. unterschiedliche Strompreise zu unterschiedlichen Zeiten) erfolgt. In der Dissertation werden beide Aspekte unter dem Begriff Demand Side Management verstanden.

Disaster Recovery (Notfallwiederherstellung) bezeichnet Maßnahmen, die nach einem Disaster (z. B. Naturkatastrophe) eingeleitet werden. Zu diesen Maßnahmen zählen die Datenwiederherstellung und das Ersetzen von defekter Hardware sowie Infrastruktur.

DSM ist die Abkürzung für Demand Side Management.

DVFS ist die Abkürzung für Dynamic Voltage and Frequency Scaling.

Dynamic Voltage and Frequency Scaling ist eine Technologie zur Anpassung der Frequenz und Spannungsversorgung, wodurch der Strombedarf und damit auch die Wärmeentwicklung auf Kosten der Verarbeitungsgeschwindigkeit geregelt werden können.

Dynamische Verteilung bezeichnet in dem eingesetzten rechenzentrumsinternen LPM eine Verteilung der VMs auf die Server, die rechtzeitig durch Migrationen in eine sichere Verteilung gebracht werden kann, ohne dass Ressourcenengpässe entstehen.

Dynamische Virtualisierung bezeichnet in Abgrenzung zur statischen Virtualisierung die dynamische Reallokation von VMs mit Hilfe der Live-Migrationstechnik.

Flexibilität drückt im Kontext von Replikationen aus, welcher Grad an möglichen Migrationen dem globalen LPM zur Verfügung steht. Dieser wird durch die Anzahl an Replikaten im Gesamtverbund bestimmt. Werden die Daten jeder VM an jedem Standort repliziert, ist die Flexibilität maximal. Werden sie jedoch jeweils nur an einem Standort vorgehalten, lässt sich keine globale Migration durchführen. Die Flexibilität ist damit gleich Null.

Fremd-VM ist für einen Server eine VM, für die er nicht ihr Heimserver ist.

GRASP ist die Abkürzung für Greedy Randomized Adaptive Search Procedures. Es handelt sich um eine Metaheuristik für kombinatorische Optimierung, die durch einen greedy randomisierten, iterativen Aufbauprozess und durch einen Verbesserungsprozess mittels lokaler Suche eine Lösung generiert.

Greedy (zu deutsch gierig) beschreibt Algorithmen, welche iterativ Folgezustände auswählen, die von der aktuellen Situation den größten Gewinn bzgl. einer Bewertungsmetrik darstellen.

Grid Computing ist die Realisierung von Large-Scale-Services über lose gekoppelte Server zu einem verteilten Rechencluster.

Heimserver ist der Server einer VM, welcher ihr für die sichere Verteilung zugeordnet ist.

Heuristik bezeichnet eine Methode, komplexe Probleme, die sich nicht oder nur sehr schwierig vollständig lösen lassen, mit Hilfe einfacher Regeln und unter Zuhilfenahme nur weniger Informationen zu entwirren.

Hybrid Cloud ist ein Zusammenschluss einer Private Cloud mit einer Public Cloud, wobei die Autonomie der Clouds bestehen bleibt.

Hypervisor ist ein alternativer Begriff für VMM.

ILP ist die Abkürzung für Integer Linear Programming (ganzahlige lineare Optimierung). Hierbei geht es um die Optimierung linearer Zielfunktionen über einer Menge von ganzzahligen Variablen, die durch lineare Gleichungen und Ungleichungen eingeschränkt sind.

IaaS ist die Abkürzung für Infrastructure-as-a-Service.

IKT ist die Abkürzung für Informations- und Kommunikationstechnologie.

Infrastructure-as-a-Service ist die Service-Ebene beim Cloud Computing, bei der IT-Ressourcen wie z. B. Server und Storage angeboten werden. Aufgrund des hohen Einsatzes von Virtualisierung handelt es sich hierbei zumeist um virtuelle Server.

Inter-Cloud oder auch Inter-Cloud-Computing bezeichnet den Zusammenschluss von unterschiedlichen Clouds zu einer größeren gemeinsamen Cloud.

KMU ist die Abkürzung für kleine und mittlere Unternehmen.

Konsolidierung bezeichnet einen Verdichtungsvorgang bei der Zuordnung von Elementen zu Behältern, so dass nach Abschluss der Konsolidierung weniger Behälter notwendig sind.

Large-Scale-Services sind Dienste bzw. Applikationen, die zum Betrieb mehrere Server in einem (verteilten) Cluster benötigen.

Last- und Powermanagement ist ein System bzw. Verfahren, das die Last bzw. den Energiebedarf entsprechend der aktuellen Situation und einem bestimmten Ziel anpasst.

Lastobjekt ist das Element, welches durch das LPM für seine Anpassungen genutzt wird.

Laufzeitkomplexität gibt die Anzahl an Rechenschritten in Abhängigkeit einer oder mehrerer Bezugsgrößen an. In der Regel wird hierbei vom Worst-Case ausgegangen. Die Angabe erfolgt in Landau-Notation.

Least-Frequently-Used ist eine Ersetzungsstrategie beim Caching, bei der derjenige Eintrag ersetzt wird, welcher am wenigsten häufig genutzt wurde.

Least-Recently-Used ist eine Ersetzungsstrategie beim Caching, bei der derjenige Eintrag ersetzt wird, welcher am längsten nicht genutzt wurde.

LFU ist die Abkürzung für Least-Frequently-Used.

LRU ist die Abkürzung für Least-Recently-Used.

Live-Migration ist eine Technik zur Migration einer aktiven VM zwischen physischen Servern, ohne dass die VM oder die darin laufenden Dienste bzw. Applikationen angehalten werden müssen.

LPM ist die Abkürzung für Last- und Powermanagement.

MAS ist die Abkürzung für Multi-Agenten-System.

Migrationsgraph bezeichnet beim eingesetzten rechenzentrumsinternen LPM den gerichteten Graphen, bei dem die Knoten die betriebenen Server und die Kanten die durchzuführenden Migrationen darstellen, um in die sichere Verteilung zurückzukehren.

Migrationskante bezeichnet im Migrationsgraphen eine Kante, welche aussagt, dass zur Erreichung der sicheren Verteilung eine Migration vom Ausgangspunkt der Kante zum Ziel der Kante zu erfolgen hat.

Multi-Agenten-System wird dazu eingesetzt, um ein hochkomplexes Problem auf kleinere Probleme aufzuteilen, so dass diese durch einzelne Agenten verteilt lösbar sind. Gegenüber einer zentralen Lösung wird hierdurch höhere Robustheit und bessere Skalierbarkeit erreicht. Eine Gütegarantie kann hierbei allerdings i.d.R. nicht gegeben werden.

Ordnungsmetrik stellt im entwickelten Standortselektionsverfahren eine Bewertungsmetrik für die möglichen Replikate dar, wodurch diese in einer festen Reihenfolge iterativ durchlaufen werden können, um die Flexibilität anzupassen.

PaaS ist die Abkürzung für Platform-as-a-Service.

Platform-as-a-Service ist die Service-Ebene beim Cloud Computing, bei der Anwendungsinfrastruktur geliefert wird. Hierzu gehören zum einen Entwicklungs- und Laufzeitumgebungen für z. B. Webanwendungen. Dabei werden häufig visuelle Werkzeuge eingesetzt, die die Entwicklung mit wenig Programmierkenntnissen unterstützen soll [222]. Zum anderen werden Laufzeitumgebungen für Programmiersprachen (z. B. .Net, Python, Java) und Datenbanken angeboten.

Private Cloud wird unternehmensintern gebildet, um die eigenen IT-Ressourcen leicht und flexibel für unterschiedliche Dienste und Abteilungen bereitzustellen [17]. Die hierin betriebenen Dienste werden ausschließlich vom Unternehmen verwendet und nicht der Öffentlichkeit zugänglich gemacht. Die flexible Architektur wird also zur Verbesserung der internen Abläufe genutzt. Eine Private Cloud ist aufgrund der hohen Investitionskosten eher für große Unternehmen attraktiv, die vom Skaleneffekt profitieren können.

Public Cloud ist eine über das Internet erreichbare Cloud, deren Services jeder beziehen kann. Die Angebote profitieren von den Skaleneffekten der Cloud-Provider, die hierzu große Rechenzentren betreiben. Services aus der Public Cloud sind insbesondere für KMU interessant ist.

PUE ist die Abkürzung für Power Usage Effectiveness. Sie setzt die Gesamtleistungsaufnahme des Rechenzentrums ins Verhältnis zur Leistungsaufnahme der IT und stellt damit ein Maß für die Infrastruktureffizienz in einem Rechenzentrum dar.

Regressionsmodell ist ein Modell zur Bestimmung der Beziehung zwischen einer sogenannten abhängigen Variable und einer oder mehreren unabhängigen Variablen (auch Schätzer genannt).

REP ist eine Ordnungsmetrik zur Ermittlung jener Replikate, die auf den Einspareffekt den größten Einfluss haben. In dieser Metrik werden die anfallenden Replikationskosten berücksichtigt, weshalb die Replikationskosten je Flexibilität hierbei minimal sind.

Replikation bezeichnet die Vervielfältigung von Daten oder Diensten (an unterschiedlichen Standorten) zur Erhöhung z. B. des Datendurchsatzes oder der Verfügbarkeit.

SaaS ist die Abkürzung für Software-as-a-Service.

Seed wird im Zusammenhang mit der Verwendung eines Zufallsgenerators genutzt. In der Arbeit wird ein Zufallsgenerator benötigt, um unterschiedliche Ausprägungen für z. B. heterogene Server zu generieren. Zur Vergleichbarkeit von Algorithmen muss jedoch auf dieselben Ausprägungen zurückgegriffen werden können. Hierzu wird ein festgelegter Seed verwendet, über den immer die gleichen quasi zufälligen Ausprägungen generiert werden können.

Server-Cluster ist ein Zusammenschluss von mehreren (verteilten) Servern. Werden diese durch eine Middleware miteinander verbunden und koordiniert, können hierauf Large-Scale-Services betrieben werden.

Server-Sized-Services sind Dienste bzw. Applikationen, die auf einzelnen Servern betrieben werden können.

Service-Level-Agreement bezeichnet eine Vereinbarung zwischen Anbieter und Kunde für den Betrieb eines Dienstes, in der zugesicherte Leistungseigenschaften (Service-Level-Objectives) wie etwa Leistungsumfang, Reaktionszeit und Schnelligkeit der Bearbeitung festgehalten sind.

Service-Level-Objective ist ein Vereinbarungselement der SLA.

Sichere Verteilung wird vom eingesetzten rechenzentrumsinternen LPM als eine Verteilung angesehen, bei der zu keinem Zeitpunkt Ressourcenengpässe auftreten können, da hier die reservierten Ressourcen den statistisch maximal zu erwarteten Ressourcenbedarfen genügen.

Single-Point-of-Failure ist ein Bestandteil eines Systems, bei dessen Ausfall die Funktion des gesamten Systems ausfällt.

SLA ist die Abkürzung für Service-Level-Agreement.

SLO ist die Abkürzung für Service-Level-Objective.

Smart Grid ist ein modernisiertes, elektrisches Netz, das seinen eigenen Zustand kennt, sich schützt und automatisiert das Zusammenspiel der einzelnen Elemente koordiniert. Dies bezieht sich auf alle Elemente und alle Spannungsebenen – von der Höchstspannung zur Niederspannung, alle Erzeuger und Verbraucheranlagen und die Netzkomponenten. (Begriffsbestimmung nach NIST)

Software-as-a-Service ist die Service-Ebene beim Cloud Computing, bei der (End-User-) Applikationen wie ERP, CRM oder Kommunikationssysteme bezogen werden können.

Statische Virtualisierung bezeichnet den Betrieb von einer oder mehreren VMs auf einem Hardwareserver, was durch einen VMM realisiert wird. Der Betrieb ist statisch – eine Reallokation geschieht hier also nicht.

Unterbrechungsfreie Stromversorgung wird in Rechenzentren verwendet, um im Falle eines Netzausfalls die Zeit bis zum Einsatz des Notfallgenerators zu überbrücken. Der hierzu erforderliche Strom wird i.d.R. in Akkumulatoren zwischengespeichert. Außerdem erfüllt sie die Aufgabe der Stromglättung und des Überspannungsschutzes.

USV ist die Abkürzung für unterbrechungsfreie Stromversorgung.

UV3 ist die Ordnungsmetrik *Use-Heuristik V3* zur Ermittlung jener Replikate, die auf den Einspareffekt den größten Einfluss haben. In dieser Metrik werden die Aufenthaltsdauer und die prognostizierten Auslastungswerte der VMs sowie die Kostenersparnis der Standorte gemäß Gleichung (4.20) berücksichtigt.

Virtualisierung wird im IT-Bereich als die Abstraktion von Ressourcen verstanden. Hierzu wird eine logische Schicht eingeführt, in der reale Ressourcen durch virtuelle ersetzt werden, die die gleichen Schnittstellen und Funktionen anbieten. Bei der Paravirtualisierung kommen Veränderungen an der Schnittstelle vor.

Virtuelle Maschine ist eine Softwareumgebung, die virtuell die Schnittstellen und Funktionen eines realen Servers bereitstellt.

Virtual Maschine Monitor ist eine Software, die den Betrieb von mehreren VMs auf einem einzelnen physischen Server ermöglicht.

VM ist die Abkürzung für virtuelle Maschine.

VMM ist die Abkürzung für Virtual Machine Monitor.

WAN ist die Abkürzung für Wide Area Network (Weitverkehrsnetz). Hierbei handelt es sich um ein Netzwerk, welches sich über einen sehr großen geografischen Bereich erstreckt (länder-, kontinenteübergreifend).

Zyklus oder geschlossene Kantenfolge bezeichnet in der Graphentheorie eine Kantenfolge, deren Start- und Endknoten identisch sind.

Abbildungsverzeichnis

1.1	Der Weltenergiebedarf hat sich von 1971 bis 2011 mehr als verdoppelt. . . .	2
1.2	Prognose über den IKT-bedingten Strombedarf in Deutschland bis 2020 . . .	3
1.3	Die Rechenleistung und Rechenleistung pro Watt von Servern steigen unterschiedlich schnell.	3
1.4	Entwicklung des IP-Verkehrs in Westeuropa	6
1.5	Steigender Einsatz von Cloud Computing prognostiziert	10
1.6	Temperaturverläufe von Stuttgart und Hannover für unterschiedliche Tage . .	11
1.7	Strompreis für Stundenkontrakte, Spotmarkt EXAA 25.07.2010	12
1.8	Tagesdurchschnittlicher Strompreis unterschiedlicher europäischer Strombörsen	12
1.9	Beispiel für ein globales LPM: An jedem Zeitpunkt t_i werden die VMs (1 bis 5) zur Minimierung der Gesamtbetriebskosten zu den günstigsten Rechenzentren migriert.	13
1.10	Lastanpassungsvorgänge	15
2.1	Aufteilung der Energiebedarfe in einem Rechenzentrum	18
2.2	Virtualisierungsarchitektur	23
2.3	Magic Quadrant für x86-Server-Virtualisierung 2014	24
2.4	Beispiel für dynamische Virtualisierung zur Minimierung der Anzahl aktiver Server	26
2.5	Verwaltungs- bzw. Verantwortungsbereiche bei den unterschiedlichen Service-Ebenen	33
3.1	Die drei Ebenen des Modellierungsprozesses zur Prognose des Ressourcenbedarfs von Diensten	49
3.2	Beispielmodell der möglichen Temperaturabweichungen in Kelvin bei einer Vorhersage von bis zu zehn Tagen	50
3.3	Europäische Strompreise (2. Halbjahr 2013)	51
3.4	Die benötigte Energie (Tageslastganglinie) wird durch unterschiedliche Energieblöcke abgedeckt.	52
3.5	Deadlock durch Zyklbildung: Eine direkte Rückmigration auf den jeweiligen Heimserver würde einen Ressourcenengpass verursachen.	55
3.6	Migrationsgraph und Matrixschreibweise sowie Adjazenzmatrix zur Zyklendetektion	56
4.1	Logische Aufteilung des LPM in einen lokalen und globalen Part	60
4.2	Übersicht über die Systembestandteile des Gesamtverfahrens und den Informationsfluss	63
4.3	Hierarchischer Zusammenhang der Energiebedarfsmodelle	68
4.4	Modellierte Leistungsaufnahme eines Containerrechenzentrums	68
4.5	In der Reihenfolge der Bezifferung wird testweise versucht, VMs von Servern mit niedrigerer VM-Anzahl auf Server mit höherer VM-Anzahl zu migrieren.	70

4.6	Einfluss bzw. Güte der einzelnen Parameter bzgl. des Energiebedarfs in Relation zum Energiebedarf bei Verwendung des Parameters VM-Anzahl	72
4.7	Optimierte Gewichte für die einzelnen Parameter	73
4.8	Güte bei gleicher Gewichtung, bei für das jeweilige Szenario optimierten Gewichtungen und bei Anwendung der allgemein optimierten Gewichtungen	73
4.9	Schematische Darstellung der Überführung des Systems zu einer neuen sicheren Verteilung (Momentaufnahme)	75
4.10	Im Migrationsgraphen existieren auf ausgeschalteten Servern keine ausgehenden Migrationskanten.	78
4.11	Durch Entfernen von VMs aus dem VM-Pool entstehen in der sicheren Verteilung freie Kapazitäten, die durch partielle Neuallokation wieder belegt werden können.	79
4.12	Flussdiagramm für die Anwendung der Repacking-Heuristiken zur Anpassung und Optimierung der sicheren Verteilung	81
4.13	Auswirkungen der unterschiedlichen Repacking-Heuristiken bzgl. der Anzahl benötigter Server	82
4.14	Skalierungsverhalten des LPM bzgl. der Laufzeit und der Konsolidierungsrate in Abhängigkeit von der VM-Anzahl	84
4.15	Serverkonsolidierungsrate nach Pelley	88
4.16	Auswirkungen der unterschiedlichen Variablen sowie ihrer prognostizierten Werte auf die Güte des Regressionsmodells	90
4.17	Güte der Regressionsmodelle bei unterschiedlichen Messreihen/-gruppen	92
4.18	Prognosegüte des Modells (4.11) für unterschiedliche Vorhersagezeiträume in Abhängigkeit von der Trainingslänge	93
4.19	Prognosegüte in Abhängigkeit von der Trainingslänge: Gegenüberstellung der Regressionsmodelle (4.11) und (4.12)	93
4.20	Fehlerentwicklung bei Veränderung des VM-Pools – Ausgangslage: unterschiedliche VM-Anzahlen	94
4.21	Fehlerentwicklung bei Veränderung des VM-Pools – Ausgangslage: 100 VM, unterschiedlich gewichtete Berücksichtigung eines allgemeinen Modells	94
4.22	Güte des Modells (4.13) zur Abschätzung der benötigten Anzahl an Servern in der sicheren Verteilung	96
4.23	Güte des Modells (4.13) bei unterschiedlichen Intervallgruppen für die Serveranzahl	96
4.24	Bezug zwischen Leistungsaufnahme und summierter Gesamtauslastung der VMs bei homogenen sowie heterogenen Servern mit unterschiedlichen LPM-Policies	98
4.25	Ergebnisse des finalen Durchlaufs der GRASP-Parameteruntersuchung (1/2)	105
4.26	Ergebnisse des finalen Durchlaufs der GRASP-Parameteruntersuchung (2/2)	106
4.27	Ergebnisse der Skalierungsuntersuchung	107
4.28	Lastganggenerierung bei Vorgabe einer (realen) Energiebedarfskurve	108
4.29	Netzwerkstruktur	110
4.30	Netzwerk(-kosten)-modell (abstrahiert)	111
4.31	Übersicht zur Funktionsweise des globalen LPM anhand eines Beispiels	112
4.32	Evaluation des Verfahrens zur Migrationsentscheidung unter Vorgabe einer erreichbaren Lösung	116
4.33	Laufzeitkomplexität der Agenten	119
4.34	Durchschnittliche Datenvolumina bei Verwendung asynchroner Replikation in Relation zum Datenvolumen bei einem Synchronisierungsintervall von 24 Stunden	128

4.35	Idealisierte Verlaufsform der Kurven für die Gesamt-, Betriebs- und Replikationskosten	132
4.36	Gegenüberstellung der Ordnungsmetriken	134
4.37	Gegenüberstellung der Ordnungsmetriken mit Bestimmung des initialen Standortes	135
4.38	Beispiel für den Gesamtkostenverlauf bei einer 5%-Abtastung des Flexibilitätswertes	136
4.39	Modifiziertes Intervallhalbierungsverfahren zur Optimierung der Standortselektion	137
5.1	Zeitliche Größen in einer Simulation	140
5.2	Beispiel für ein VM-Auslastungsprofil	143
5.3	Migrationsgeschwindigkeit: Je schneller die globalen Migrationen durchgeführt werden können, desto höher ist das Einsparpotential.	147
5.4	Die Auswahl der konkreten VM-Menge hat einen geringeren Einfluss auf das Einsparpotential als die der konkreten Strompreisverläufe.	148
5.5	Einfluss unterschiedlicher Strompreisverläufe	148
5.6	Einfluss unterschiedlicher Temperaturverläufe (regionsabhängig)	149
5.7	Geringeres Einsparpotential bei korrelierenden VM-Auslastungsprofilen	150
5.8	Einfluss der Datenübertragungskosten und des Optimierungsintervalls	151
5.9	Einfluss der Datenübertragungskosten und des Optimierungsintervalls auf die Replikations- und Gesamtkosten bei vollständiger Replikation	152
5.10	Einfluss der Datenübertragungskosten bei optimierter Standortselektion (1)	152
5.11	Einfluss der Datenübertragungskosten bei optimierter Standortselektion (2)	153
5.12	Einfluss des Optimierungsintervalls bei optimierter Standortselektion	153
5.13	Einfluss des Grades freier Ressourcen	154
5.14	Einfluss des Grades freier Ressourcen auf die Replikations- und Gesamtkosten bei vollständiger Replikation	154
5.15	Einfluss des Grades freier Ressourcen bei optimierter Standortselektion	155
5.16	Einfluss der Standortanzahl	155
5.17	Einfluss der Standortanzahl auf die Replikations- und Gesamtkosten bei vollständiger Replikation	156
5.18	Einfluss der Standortanzahl bei optimierter Standortselektion	156
5.19	Energiebedarfe unterschiedlicher VMs	160
5.20	Die Stromgenerierung mittels Windkraft ist sehr volatil.	163
5.21	Beispiel für die Leistungsaufnahme einer Siedlung, mit geglättetem Verlauf (Vermeidung von Peaks) und vollständiger Glättung	164

Literaturverzeichnis

- [1] ABBASI, Z. ; VARSAMOPOULOS, G. ; GUPTA, S. K. S.: Thermal aware server provisioning and workload distribution for internet data centers. In: *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*, ACM, 2010 (HPDC '10). – ISBN 978-1-60558-942-8, S. 130–141
- [2] ADAMS, K. : Geographically Distributed System for Catastrophic Recovery. In: *Proceedings of the 16th USENIX Conference on System Administration*. USENIX Association (LISA '02), 47–64
- [3] AIKEMA, D. ; SIMMONDS, R. ; ZAREIPOUR, H. : Data Centres in the Ancillary Services Market. In: *Proceedings of the 2012 International Green Computing Conference (IGCC)*. Washington, DC, USA : IEEE Computer Society, 2012 (IGCC '12). – ISBN 978-1-4673-2155-6, S. 1–10
- [4] AKAMAI: Akamai and Cloud Computing a Perspective from the Edge of the Cloud / Akamai. 2009. – White Paper. – 10 S.
- [5] AKSANLI, B. ; ROSING, T. ; MONGA, I. : Benefits of green energy and proportionality in high speed wide area networks connecting data centers. In: *Design, Automation Test in Europe Conference Exhibition (DATE), 2012*, 2012. – ISSN 1530-1591, S. 175–180
- [6] AKSANLI, B. ; VENKATESH, J. ; ZHANG, L. ; ROSING, T. : Utilizing Green Energy Prediction to Schedule Mixed Batch and Service Jobs in Data Centers. In: *Proceedings of the 4th Workshop on Power-Aware Computing and Systems*, ACM, 2011 (HotPower '11). – ISBN 978-1-4503-0981-3, S. 5:1–5:5
- [7] ALLCOTT, H. : Rethinking real-time electricity pricing. In: *Resource and Energy Economics* 33 (2011), Nr. 4, S. 820 – 842. – DOI 10.1016/j.reseneeco.2011.06.003. – ISSN 0928-7655. – Special section: Sustainable Resource Use and Economic Dynamics
- [8] ALLSPAW, J. : *The Art of Capacity Planning: Scaling Web Resources*. O'Reilly Media, Inc., 2008. – ISBN 0596518579, 9780596518578
- [9] AMD: *Study Reveals Scope of Annual U.S. Data Center Energy Consumption*. http://www.powerpulse.net/story.php?storyID=16697&source=1_. Version: Februar 2007, Abruf: 04.08.2015. – Pressemitteilung
- [10] ANTONAKOPOULOS, S. ; FORTUNE, S. ; ZHANG, L. : Power-aware routing with rate-adaptive network elements. In: *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*, 2010, S. 1428–1432
- [11] AOYAMA, T. ; SAKAI, H. : Inter-Cloud Computing. In: *Business & Information Systems Engineering* 3 (2011), Nr. 3, S. 173–177. – DOI 10.1007/s12599-011-0158-4

- [12] APEL, R. ; AUNDRUP, T. ; BUCHHOLZ, B. M. ; DOMELS, H. P. ; FUNKE, S. ; GESING, T. ; GLAUNSINGER, W. ; GRONSTED, P. ; HEINS, F. ; HÖSLE, A. ; KNACK, V. ; KÜTER, T. ; STÖTZER, M. ; STRUWE, C. ; STYCZYNSKI, Z. A. ; SCHEVEN, A. von ; SEIDL, H. : Demand Side Integration – Lastverschiebungspotenziale in Deutschland / VDE. 2012. – Forschungsbericht
- [13] ARDICA, M. ; HOWERING, J. : Data Center Interconnect, Design Guide for Virtualized Workload Mobility with Cisco, Netapp, and VMware / Cisco. 2011. – Forschungsbericht. – 68 S.
- [14] ARORA, M. ; MANNE, S. ; ECKERT, Y. ; PAUL, I. ; JAYASENA, N. ; TULLSEN, D. : A Comparison of Core Power Gating Strategies Implemented in Modern Hardware. In: *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*, ACM, 2014 (SIGMETRICS '14). – ISBN 978-1-4503-2789-3, S. 559–560
- [15] AVELAR, V. ; AZEVEDO, D. ; FRENCH, A. : PUE?: A COMPREHENSIVE EXAMINATION OF THE METRIC / The Green Grid. 2012. – White Paper
- [16] BAKER, J. : Worldwide Storage Management Software 2008-2012 Forecast: Balanced Growth Throughput the Segment / IDC. 2008. – Survey
- [17] BALIGA, J. ; AYRE, R. ; HINTON, K. ; TUCKER, R. : Green Cloud Computing: Balancing Energy in Processing, Storage, and Transport. In: *Proceedings of the IEEE* 99 (2011), Jan, Nr. 1, S. 149–167. – DOI 10.1109/JPROC.2010.2060451. – ISSN 0018-9219
- [18] BANERJEE, P. ; PATEL, C. D. ; BASH, C. ; RANGANATHAN, P. : Sustainable Data Centers: Enabled by Supply and Demand Side Management. In: *Proceedings of the 46th Annual Design Automation Conference*, ACM, 2009 (DAC '09). – ISBN 978-1-60558-497-3, S. 884–887
- [19] BARROSO, L. A. ; HÖLZLE, U. : The Case for Energy-Proportional Computing. In: *Computer* 40 (2007), Nr. 12, S. 33–37. – DOI 10.1109/MC.2007.443. – ISSN 0018-9162
- [20] BARROSO, L. A. ; DEAN, J. ; HÖLZLE, U. : Web Search for a Planet: The Google Cluster Architecture. In: *IEEE Micro* 23 (2003), März, Nr. 2, S. 22–28. – DOI 10.1109/MM.2003.1196112. – ISSN 0272-1732
- [21] BEEVOR, M. ; BIRBECK, S. ; O'CONNOR, R. : Disaster Recovery White Paper: Reducing the Bandwidth to Keep Remote Sites Constantly Up-to-date / DataCore, Riverbed, waterstons. Version: 2008. http://www.datacore.com/downloads/DataCore_Riverbed_DR_22Sep08.pdf, Abruf: 04.08.2015. – White Paper
- [22] BEHRENDT, F. ; SCHAEFER, M. ; BELUSA, T. ; ZIEGLER, F. ; LANSER, W. ; ERDMANN, G. ; DITTMAR, L. ; KLESCHIN, S. : Konzeptstudie zur Energie- und Ressourceneffizienz im Betrieb von Rechenzentren / Technische Universität Berlin Innovationszentrum Energie. Version: 2008. http://www.energie.tu-berlin.de/uploads/media/Konzeptstudie_EiR_20090807.pdf, Abruf: 04.08.2015. – Studie
- [23] BELADY, C. L.: In the Data Center, Power and Cooling Costs More than the IT Equipment It Supports. In: *Electronics Cooling Magazine* 13 (2007), February, Nr. 1, 24–27. <http://www.electronics-cooling.com/2007/02/in-the-data-center-power-and-cooling-costs-more-than-the-it-equipment-it-supports/>, Abruf: 04.08.2015

- [24] BELL, W. ; CAMERON, D. ; CARVAJAL-SCHIAFFINO, R. ; MILLAR, A. ; STOCKINGER, K. ; ZINI, F. : Evaluation of an Economy-Based File Replication Strategy for a Data Grid. In: *Cluster Computing and the Grid, 2003. Proceedings. CCGrid 2003. 3rd IEEE/ACM International Symposium on*, 2003, S. 661–668
- [25] BELL, W. H. ; CAMERON, D. G. ; CAPOZZA, L. ; MILLAR, A. P. ; STOCKINGER, K. ; ZINI, F. : Simulation of Dynamic Grid Replication Strategies in OptorSim. In: *Journal of High Performance Computing Applications*, Springer-Verlag, 2002, S. 46–57
- [26] BENMERAR, T. Z. ; OULEBSIR-BOUMGHAR, F. : From Grids to Cloud - The Pathway for Brain dMRI Cloud Services. In: *CLOSER 2012 - Proceedings of the 2nd International Conference on Cloud Computing and Services Science, Porto, Portugal, 18 - 21 April, 2012*, 2012, S. 141–146
- [27] BERL, A. ; KLINGERT, S. ; BECK, M. ; MEER, H. de: Integrating data centres into demand-response management: A local case study. In: *Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE*, 2013. – ISSN 1553–572X, S. 4762–4767
- [28] BERLIN, I. : Best Practice Beispiele für energieeffiziente Rechenzentren / ITDZ Berlin. 2010. – White Paper. – 16 S.
- [29] BHAGWAN, R. ; MOORE, D. ; SAVAGE, S. ; VOELKER, G. M.: Replication Strategies for Highly Available Peer-to-Peer Storage. Version: 2003. In: SCHIPER, A. (Hrsg.) ; SHVARTSMAN, A. (Hrsg.) ; WEATHERSPOON, H. (Hrsg.) ; ZHAO, B. (Hrsg.): *Future Directions in Distributed Computing* Bd. 2584. Springer Berlin Heidelberg, 2003. – DOI 10.1007/3-540-37795-6_28. – ISBN 978-3-540-00912-2, S. 153–158
- [30] BIANCHINI, R. ; RAJAMONY, R. : Power and Energy Management for Server Systems. In: *Computer* 37 (2004), Nov., Nr. 11, S. 68–74. – DOI 10.1109/MC.2004.217. – ISSN 0018–9162
- [31] BOBROFF, N. ; KOCHUT, A. ; BEATY, K. A.: Dynamic Placement of Virtual Machines for Managing SLA Violations. In: *Integrated Network Management*, 2007, S. 119–128
- [32] BORENSTEIN, S. : The long-run efficiency of real-time electricity pricing. In: *The Energy Journal* (2005), S. 93–116
- [33] BOSE, S. ; BROCK, S. ; SKEOCH, R. ; SHAIKH, N. ; RAO, S. : Optimizing live migration of virtual machines across wide area networks using integrated replication and scheduling. In: *Systems Conference (SysCon), 2011 IEEE International*, 2011, S. 97–102
- [34] BOX, G. E. P. ; JENKINS, G. M. ; REINSEL, G. C.: *Time Series Analysis: Forecasting and Control*. fourth. Wiley, 2008
- [35] BRAZIER, F. ; KEPHART, J. ; VAN DYKE PARUNAK, H. ; HUHS, M. : Agents and Service-Oriented Computing for Autonomic Computing: A Research Agenda. In: *Internet Computing, IEEE* 13 (2009), May, Nr. 3, S. 82–87. – DOI 10.1109/MIC.2009.51. – ISSN 1089–7801
- [36] BROWN, R. : Report to Congress on Server and Data Center Energy Efficiency: Public Law 109-431 / Lawrence Berkeley National Laboratory. 2008. – Report. – 137 S.

- [37] BRUSCHI, J. ; RUMSEY, P. ; ANLIKER, R. ; CHU, L. ; GREGSON, S. : Best Practices Guide for Energy-Efficient Data Center Design / U.S. Department of Energy. 2010. – White Paper. – 28 S.
- [38] BUCHBINDER, N. ; JAIN, N. ; MENACHE, I. : Online Job-migration for Reducing the Electricity Bill in the Cloud. In: *Proceedings of the 10th International IFIP TC 6 Conference on Networking - Volume Part I*. Springer-Verlag (NETWORKING'11). – ISBN 978-3-642-20756-3, 172–185
- [39] BURGER, H. ; ROGATTY, W. : Überschlägige Ermittlung der erforderlichen Kesselleistung. In: *IKZ-HAUSTECHNIK* 18 (2004), 44 ff. <http://www.ikz.de/1996-2005/2004/18/0418044.php>, Abruf: 04.08.2015
- [40] BUYYA, R. ; ABRAMSON, D. ; VENUGOPAL, S. : The Grid Economy. In: *Proceedings of the IEEE* 93 (2005), March, Nr. 3, S. 698–714. – DOI 10.1109/JPROC.2004.842784. – ISSN 0018-9219
- [41] BUYYA, R. ; YEO, C. S. ; VENUGOPAL, S. : Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities, in. In: *Department of Computer Science and Software Engineering (CSSE), The University of Melbourne, Australia. He*, 2008, S. 10–1016
- [42] CAI, L. ; LU, Y.-H. : Power reduction of multiple disks using dynamic cache resizing and speed control. In: *ISLPED '06: Proceedings of the 2006 international symposium on Low power electronics and design*, ACM, 2006. – ISBN 1-59593-462-6, S. 186–190
- [43] CALHEIROS, R. N. ; RANJAN, R. ; BELOGLAZOV, A. ; DE ROSE, C. A. F. ; BUYYA, R. : CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms. In: *Softw. Pract. Exper.* 41 (2011), Jan., Nr. 1, S. 23–50. – DOI 10.1002/spe.995. – ISSN 0038-0644
- [44] CAMERON, A. : *Taming IT Energy Usage Toward Future Environmental Sustainability Has Become a Global Imperative*. <http://www.eds.com/-news/-features/-4316/>. Version: 2008, Abruf: 07.11.2009
- [45] CAMERON, D. G. ; CARVAJAL-SCHIAFFINO, R. ; NICHOLSON, C. ; STOCKINGER, K. ; ZINI, F. ; MILLAR, A. P. ; SERAFINI, L. : Formal Analysis of an Agent-based Optimisation Strategy for Data Grids. In: *Multiagent Grid Syst.* 2 (2006), März, Nr. 2, 149–162. <http://dl.acm.org/citation.cfm?id=2699931.2699936>. – ISSN 1574-1702
- [46] CAMERON, D. ; MILLAR, A. ; NICHOLSON, C. ; CARVAJAL-SCHIAFFINO, R. ; STOCKINGER, K. ; ZINI, F. : Analysis of Scheduling and Replica Optimisation Strategies for Data Grids Using OptorSim. In: *Journal of Grid Computing* 2 (2004), S. 57–69. – DOI 10.1007/s10723-004-6040-6. – ISSN 1570-7873
- [47] CAMERON, K. W.: Green Introspection. In: *Computer* 42 (2009), Nr. 1, S. 101–103. – DOI 10.1109/MC.2009.18. – ISSN 0018-9162
- [48] CAO, D.-S. ; XU, Q.-S. ; LIANG, Y.-Z. ; ZHANG, L.-X. ; LI, H.-D. : The boosting: A new idea of building models. In: *Chemometrics and Intelligent Laboratory Systems* 100 (2010), Jan., Nr. 1, S. 1–11. – DOI 10.1016/j.chemolab.2009.09.002. – ISSN 01697439

- [49] CAPOZZA, L. ; STOCKINGER, K. ; ZINI, F. : Preliminary Evaluation of Revenue Prediction Functions for Economically-Effective File Replication / CERN. 2002. – Forschungsbericht. – DataGrid-02-TED-020724
- [50] CARMAN, M. ; ZINI, F. ; SERAFINI, L. ; STOCKINGER, K. : Towards an Economy-Based Optimisation of File Access and Replication on a Data Grid. In: *Proceedings of the 2Nd IEEE/ACM International Symposium on Cluster Computing and the Grid*. IEEE Computer Society (CCGRID '02). – ISBN 0-7695-1582-7, 340–
- [51] CATTEDDU, D. ; HOGBEN, G. : Cloud Computing – Benefits, risks and recommendations for information security / European Network and Information Security Agency (ENISA). 2009. – Forschungsbericht. – 125 S.
- [52] CENTENO, R. ; BILLHARDT, H. ; HERMOSO, R. ; OSSOWSKI, S. : Organising MAS: A Formal Model Based on Organisational Mechanisms. In: *Proceedings of the 2009 ACM Symposium on Applied Computing*, ACM, 2009 (SAC '09). – ISBN 978-1-60558-166-8, S. 740–746
- [53] CHABAREK, J. ; SOMMERS, J. ; BARFORD, P. ; ESTAN, C. ; TSIANG, D. ; WRIGHT, S. : Power Awareness in Network Design and Routing. In: *INFOCOM 2008. The 27th Conference on Computer Communications*. IEEE, 2008. – ISSN 0743-166X
- [54] CHANG, R.-S. ; CHANG, H.-P. ; WANG, Y.-T. : A dynamic weighted data replication strategy in data grids. In: *Computer Systems and Applications, 2008. AICCSA 2008. IEEE/ACS International Conference on*, 2008, S. 414–421
- [55] CHEN, G. ; HE, W. ; LIU, J. ; NATH, S. ; RIGAS, L. ; XIAO, L. ; ZHAO, F. : Energy-aware Server Provisioning and Load Dispatching for Connection-intensive Internet Services. In: *Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation*. USENIX Association (NSDI'08). – ISBN 111-999-5555-22-1, 337–350
- [56] CHEN, H. ; CARAMANIS, M. ; COSKUN, A. : The data center as a grid load stabilizer. In: *Design Automation Conference (ASP-DAC), 2014 19th Asia and South Pacific*, 2014, S. 105–112
- [57] CHEN, H. ; COSKUN, A. ; CARAMANIS, M. : Real-time power control of data centers for providing Regulation Service. In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, 2013. – ISSN 0743-1546, S. 4314–4321
- [58] CHEN, Y. ; WANG, T. ; KATZ, R. : Energy efficient Ethernet encodings. In: *Local Computer Networks, 2008. LCN 2008. 33rd IEEE Conference on*, 2008, S. 122–129
- [59] CHIARAVIGLIO, L. ; MELLIA, M. ; NERI, F. : Minimizing ISP Network Energy Cost: Formulation and Solutions. In: *IEEE/ACM Trans. Netw.* 20 (2012), Apr., Nr. 2, S. 463–476. – DOI 10.1109/TNET.2011.2161487. – ISSN 1063-6692
- [60] CHIHI, H. ; CHAINBI, W. ; GHÉDIRA, K. : A Multi-Agent based Architecture for Cloud Infrastructure Auto-adaptation. In: *CLOSER 2013 - Proceedings of the 3rd International Conference on Cloud Computing and Services Science, Aachen, Germany, 8-10 May, 2013*, 2013, S. 90–95
- [61] CHO, Y. ; CHANG, N. : Energy-Aware Clock-Frequency Assignment in Microprocessors and Memory Devices for Dynamic Voltage Scaling. In: *IEEE Trans. on CAD of Integrated Circuits and Systems* 26 (2007), Nr. 6, S. 1030–1040. – DOI 10.1109/TCAD.2006.885835

- [62] CHURCH, K. ; GREENBERG, A. ; HAMILTON, J. : On Delivering Embarrassingly Distributed Cloud Services. In: *HotNets 2008 Proceedings*, 55-60
- [63] CIDON, I. ; KUTTEN, S. ; SOFFER, R. : Optimal Allocation of Electronic Content. In: *Comput. Netw.* 40 (2002), Okt., Nr. 2, S. 205–218. – DOI 10.1016/S1389–1286(02)00251–7. – ISSN 1389–1286
- [64] CISCO: *Virtual Machine Mobility with VMware VMotion and Cisco Data Center Interconnect Technologies*. http://www.cisco.com/c/dam/en/us/solutions/collateral/data-center-virtualization/data-center-virtualization/white_paper_c11-557822.pdf. Version: 2009, Abruf: 25.11.2014
- [65] CISCO: Cisco Visual Networking Index: Forecast and Methodology, 2013-2018 / Cisco. 2014. – White Paper
- [66] CLARK, C. ; FRASER, K. ; HAND, S. ; HANSEN, J. G. ; JUL, E. ; LIMPACH, C. ; PRATT, I. ; WARFIELD, A. : Live Migration of Virtual Machines. In: *Proceedings of the 2nd Conference on Symposium on Networked Systems Design & Implementation - Volume 2*. USENIX Association (NSDI'05), 273–286
- [67] COROAMA, V. C. ; HILTY, L. ; HEIRI, E. ; HORN, F. M.: The direct energy demand of internet data flows. In: *Journal of Industrial Ecology* 17 (2013), Nr. 5, S. 680–688. – ISSN 1088–1980
- [68] CORPORATION, I. : An architectural blueprint for autonomic computing / IBM Corporation. 2003. – White Paper
- [69] DAS, R. ; KEPHART, J. O. ; LEFURGY, C. ; TESAURO, G. ; LEVINE, D. W. ; CHAN, H. : Autonomic Multi-agent Management of Power and Performance in Data Centers. In: *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems: Industrial Track*. International Foundation for Autonomous Agents and Multiagent Systems (AAMAS '08), 107–114
- [70] (DEHST), D. E.: *Kohlendioxidemissionen der emissionshandlungspflichtigen stationären Anlagen und im Luftverkehr in Deutschland im Jahr 2011*. Mai 2012. – VET-Bericht
- [71] DERHAB, A. ; BADACHE, N. : Data replication protocols for mobile ad-hoc networks: a survey and taxonomy. In: *Communications Surveys Tutorials, IEEE* 11 (2009), Second, Nr. 2, S. 33–51. – DOI 10.1109/SURV.2009.090204. – ISSN 1553–877X
- [72] DHIMAN, G. ; MIHIC, K. ; ROSING, T. : A system for online power prediction in virtualized environments using Gaussian mixture models. In: *Proceedings of the 47th Design Automation Conference, ACM, 2010 (DAC '10)*. – ISBN 978–1–4503–0002–5, S. 807–812
- [73] DÜMMEL, T. ; JANKIEWICZ, P. ; GRASSE, P. : *Prognosegüte von Wetterportalen*. http://www.geo.fu-berlin.de/met/ag/miks/medien/download/Guete_Wetterportale_2007_02.pdf. Version: 2007, Abruf: 08.09.2014
- [74] DOBBER, M. ; MEI, R. van d. ; KOOLE, G. : A prediction method for job runtimes on shared processors: Survey, statistical analysis and new avenues. In: *Perform. Eval.* 64 (2007), Aug., Nr. 7-8, S. 755–781. – DOI 10.1016/j.peva.2007.01.001. – ISSN 0166–5316

- [75] DOBSON, S. ; STERRITT, R. ; NIXON, P. ; HINCHEY, M. : Fulfilling the Vision of Automatic Computing. In: *Computer* 43 (2010), Nr. 1, S. 35–41. – DOI 10.1109/MC.2010.14. – ISSN 0018–9162
- [76] ECKART, M. : Cloud Computing / TecChannel. 2010. – Forschungsbericht. – 164 S.
- [77] EILENBERGER, G. J. ; BUNSE, S. ; DEMBECK, L. ; GEBHARD, U. ; ILCHMANN, F. ; LAUTENSCHLAEGER, W. ; MILBRANDT, J. : Energy-efficient transport for the future internet. In: *Bell Labs Technical Journal* 15 (2010), Nr. 2, S. 147–167. – DOI 10.1002/bltj.20446. – ISSN 1538–7305
- [78] ELGHIRANI, A. ; SUBRATA, R. ; ZOMAYA, A. : A Proactive Non-Cooperative Game-Theoretic Framework for Data Replication in Data Grids. In: *Cluster Computing and the Grid, 2008. CCGRID '08. 8th IEEE International Symposium on*, 2008, S. 433–440
- [79] ELMROTH, E. ; MARQUEZ, F. G. ; HENRIKSSON, D. ; FERRERA, D. P.: Accounting and Billing for Federated Cloud Infrastructures. In: *Proceedings of the 2009 Eighth International Conference on Grid and Cooperative Computing*. Washington, DC, USA : IEEE Computer Society, 2009 (GCC '09). – ISBN 978–0–7695–3766–5, S. 268–275
- [80] ERBES, J. ; MOTAHARI-NEZHAD, H. R. ; GRAUPNER, S. : The Future of Enterprise IT in the Cloud. In: *Computer* 45 (2012), S. 66–72. – DOI 10.1109/MC.2012.73. – ISSN 0018–9162
- [81] FAN, X. ; WEBER, W.-D. ; BARROSO, L. A.: Power Provisioning for a Warehouse-sized Computer. In: *Proceedings of the 34th Annual International Symposium on Computer Architecture*, ACM, 2007 (ISCA '07). – ISBN 978–1–59593–706–3, S. 13–23
- [82] FAUCHEUX, S. ; NICOLAÏ, I. : {IT} for green and green IT: A proposed typology of eco-innovation. In: *Ecological Economics* 70 (2011), Nr. 11, S. 2020 – 2027. – DOI 10.1016/j.ecolecon.2011.05.019. – ISSN 0921–8009. – Special Section - Earth System Governance: Accountability and Legitimacy
- [83] FEO, T. A. ; RESENDE, M. G. C.: A Probabilistic Heuristic for a Computationally Difficult Set Covering Problem. In: *Oper. Res. Lett.* 8 (1989), Apr., Nr. 2, S. 67–71. – DOI 10.1016/0167–6377(89)90002–3. – ISSN 0167–6377
- [84] FESTA, P. : Greedy Randomized Adaptive Search Procedures. In: *AIROneWS* 7 (2002), Nr. 4, S. 7–11
- [85] FICHTER, K. ; CLAUSEN, J. ; EIMERTENBRINK, M. : Energieeffiziente Rechenzentren – Best-Practice-Beispiele aus Europa, USA und Asien / BMU, Borderstep. 2009. – White Paper. – 40 S.
- [86] FRANKLIN, B. : Data Center Interconnect, Implementation Guide for Virtualized Workload Mobility with Cisco, EMC and VMware / Cisco. 2011. – Forschungsbericht. – 96 S.
- [87] FREUND, Y. ; SCHAPIRE, R. E.: A decision-theoretic generalization of on-line learning and an application to boosting. In: *Proceedings of the Second European Conference on Computational Learning Theory*. Springer-Verlag (EuroCOLT '95). – ISBN 3–540–59119–2, 23–37

- [88] GHAMKHARI, M. ; MOHSENIAN-RAD, H. : Optimal integration of renewable energy resources in data centers with behind-the-meter renewable generator. In: *Communications (ICC), 2012 IEEE International Conference on*, 2012. – ISSN 1550–3607, S. 3340–3344
- [89] GHATIKAR, G. ; PIETTE, M. A. ; FUJITA, K. S. ; MCKANE, A. T. ; DUDLEY, J. H. ; JR., A. R. ; MARES, K. ; SHROYER, D. : Demand Response and Open Automated Demand Response Opportunities for Data Centers / Lawrence Berkeley National Laboratory. Berkeley, Jan 2010. – Report. – 56 S. – California Energy Commission, PIER Program and Pacific Gas and Electric Company (PG&E).
- [90] GMACH, D. ; ROLIA, J. ; BASH, C. ; CHEN, Y. ; CHRISTIAN, T. ; SHAH, A. ; SHARMA, R. K. ; WANG, Z. : Capacity planning and power management to exploit sustainable energy. In: *CNSM*, 2010, S. 96–103
- [91] GOIRI, I. n. ; LE, K. ; HAQUE, M. E. ; BEAUCHEA, R. ; NGUYEN, T. D. ; GUITART, J. ; TORRES, J. ; BIANCHINI, R. : GreenSlot: Scheduling Energy Consumption in Green Datacenters. In: *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, ACM, 2011 (SC '11). – ISBN 978–1–4503–0771–0, S. 20:1–20:11
- [92] GREENBERG, A. ; HAMILTON, J. ; MALTZ, D. A. ; PATEL, P. : The Cost of a Cloud: Research Problems in Data Center Networks. In: *SIGCOMM Comput. Commun. Rev.* 39 (2008), Dez., Nr. 1, S. 68–73. – DOI 10.1145/1496091.1496103. – ISSN 0146–4833
- [93] GREENBERG, A. ; HAMILTON, J. R. ; JAIN, N. ; KANDULA, S. ; KIM, C. ; LAHIRI, P. ; MALTZ, D. A. ; PATEL, P. ; SENGUPTA, S. : VL2: A Scalable and Flexible Data Center Network. In: *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication*, ACM, 2009 (SIGCOMM '09). – ISBN 978–1–60558–594–9, S. 51–62
- [94] GREENPEACE: Make IT green: Cloud computing and its contribution to climate change / Greenpeace International. 2010. – Forschungsbericht
- [95] GUEYE, M. ; SARR, I. ; NDIAYE, S. : Database Replication in Large Scale Systems: Optimizing the Number of Replicas. In: *Proceedings of the 2009 EDBT/ICDT Workshops*, ACM, 2009 (EDBT/ICDT '09). – ISBN 978–1–60558–650–2, S. 3–9
- [96] GUPTA, A. : The 24-Hour Knowledge Factory: Can It Replace the Graveyard Shift? In: *Computer* 42 (2009), Nr. 1, S. 66–73. – DOI 10.1109/MC.2009.25. – ISSN 0018–9162
- [97] GURUMURTHI, S. ; SIVASUBRAMANIAM, A. ; KANDEMIR, M. ; FRANKE, H. : DRPM: Dynamic Speed Control for Power Management in Server Class Disks. In: *Computer Architecture, 2003. Proceedings. 30th Annual International Symposium on*, 2003. – ISSN 1063–6897, S. 169–179
- [98] GUTTMANN, C. : Towards a Taxonomy of Decision Making Problems in Multi-Agent Systems. Version: 2009. In: BRAUBACH, L. (Hrsg.) ; HOEK, W. van d. (Hrsg.) ; PETTA, P. (Hrsg.) ; POKAHR, A. (Hrsg.): *Multiagent System Technologies* Bd. 5774. Springer Berlin Heidelberg, 2009. – DOI 10.1007/978–3–642–04143–3_19. – ISBN 978–3–642–04142–6, S. 195–201
- [99] HÄRDER, T. ; RAHM, E. : *Datenbanksysteme: Konzepte und Techniken der Implementierung ; mit 14 Tabellen*. Springer Berlin Heidelberg https://books.google.de/books?id=tjEAejth3_AC. – ISBN 9783540421337

- [100] HARNEY, E. ; GOASGUEN, S. ; MARTIN, J. ; MURPHY, M. ; WESTALL, M. : The efficacy of live virtual machine migrations over the internet. In: *Proceedings of the 2nd international workshop on Virtualization technology in distributed computing*. New York, NY, USA : ACM, 2007 (VTDC '07). – ISBN 978-1-59593-897-8, S. 8:1-8:7
- [101] HATAMI, S. ; PEDRAM, M. : Minimizing the Electricity Bill of Cooperative Users under a Quasi-Dynamic Pricing Model. In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, 421-426
- [102] HERMENIER, F. ; LORCA, X. ; MENAUD, J.-M. ; MULLER, G. ; LAWALL, J. : Entropy: a consolidation manager for clusters. In: *Proceedings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments*, ACM, 2009 (VEE '09). – ISBN 978-1-60558-375-4, S. 41-50
- [103] HINTEMANN, R. ; FICHTER, K. : Energieverbrauch und Energiekosten von Servern und Rechenzentren in Deutschland / Borderstep. 2012. – Kurzstudie
- [104] HINTEMANN, R. ; PFAHL, S. : Energieeffizienz im Rechenzentrum / BITKOM. 2010. – Leitfaden. – 50 S.
- [105] HIROFUCHI, T. ; OGAWA, H. ; NAKADA, H. ; ITOH, S. ; SEKIGUCHI, S. : A Live Storage Migration Mechanism over WAN for Relocatable Virtual Machine Services on Clouds. In: *Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on*, 2009, S. 460-465
- [106] HOLT, A. ; WEISS, K. ; HUBERTY, K. ; GELBLUM, E. ; FLANNERY, S. ; DEVGAN, S. ; MALIK, A. ; ROZOF, N. ; WOOD, A. ; STANDAERT, P. ; MEUNIER, F. ; LU, J. ; CHEN, G. ; LU, B. ; HAN, K. ; KHARE, V. ; MIYACHI, M. : Cloud Computing Takes Off – Market Set to Boom as Migration Accelerates / Morgan Stanley Research. 2011. – Blue Paper
- [107] HORENKAMP, W. ; HUBE, W. ; JÄGER, J. e. a.: VDE-Studie Dezentrale Energieversorgung 2020 / Energietechnische Gesellschaft in der VDE (ETG). 2007. – Studie
- [108] HOYER, M. : *Resource Management in Virtualized Data Centers Regarding Performance and Energy Aspects*, C.v.O. Universität Oldenburg, Diss., 2011
- [109] HOYER, M. ; SCHRÖDER, K. ; NEBEL, W. : Statistical static capacity management in virtualized data centers supporting fine grained QoS specification. In: *Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking*, ACM, 2010 (e-Energy '10). – ISBN 978-1-4503-0042-1, S. 51-60
- [110] HOYER, M. ; SCHRÖDER, K. ; SCHLITT, D. ; NEBEL, W. : Proactive dynamic resource management in virtualized data centers. In: *Proceedings of the 2nd International Conference on Energy-Efficient Computing and Networking*, ACM, 2011 (e-Energy '11). – ISBN 978-1-4503-1313-1, S. 11-20
- [111] HUANG, H. ; HUNG, W. ; SHIN, K. G.: FS2: Dynamic Data Replication in Free Disk Space for Improving Disk Performance and Energy Consumption. In: *Proceedings of the Twentieth ACM Symposium on Operating Systems Principles*, ACM, 2005 (SOSP '05). – ISBN 1-59593-079-5, S. 263-276
- [112] HUANG, W. ; LIU, J. ; ABALI, B. ; PANDA, D. K.: A Case for High Performance Computing with Virtual Machines. In: *Proceedings of the 20th Annual International Conference on Supercomputing*, ACM, 2006 (ICS '06). – ISBN 1-59593-282-8, S. 125-134

- [113] HUTTENLOHNER, R. : Stromverteilung über Schienen entlastet das Budget. In: *Rechenzentren und Infrastruktur* 1/2009 (2009), S. 16–19
- [114] IBM: *IBM Energieeffizienzanalyse für Rechenzentren*. 2007. – Report
- [115] IBM: *Wirtschaftlichere x86 Infrastrukturen mit IBM System x & IBM Site and Facilities Services*. Juli 2010. – Vortrag
- [116] IDC: *Transforming the Datacenter: Consolidation, Pervasive Virtualization and Energy Optimization*. März 2009. – IDC, Report, Doc #DIR2009-T4_MB
- [117] INOUE, T. ; IKEDA, M. ; ENOKIDO, T. ; AIKEBAIER, A. ; TAKIZAWA, M. : A Power Consumption Model for Storage-based Applications. In: *Complex, Intelligent and Software Intensive Systems (CISIS), 2011 International Conference on*, 2011, S. 612–617
- [118] ISCI, C. ; BUYUKTOSUNOGLU, A. ; CHER, C.-Y. ; BOSE, P. ; MARTONOSI, M. : An Analysis of Efficient Multi-Core Global Power Management Policies: Maximizing Performance for a Given Power Budget. In: *Proceedings of the 39th Annual IEEE/ACM International Symposium on Microarchitecture*. Washington, DC, USA : IEEE Computer Society, 2006 (MICRO 39). – ISBN 0–7695–2732–9, S. 347–358
- [119] ITU-T: *NGNs and Energy Efficiency / ITU-T*. 2008 (7). – Technology Watch Report
- [120] JIANG, J. ; YANG, G. ; WANG, D. : Modeling Replication Strategies in Data Grid Systems with Arbitrary Clustered Demands. In: *Proceedings of the 3rd International Conference on Scalable Information Systems*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (InfoScale '08). – ISBN 978–963–9799–28–8, 8:1–8:9
- [121] JOO, Y. ; CHOI, Y. ; SHIM, H. ; LEE, H. G. ; KIM, K. ; CHANG, N. : Energy exploration and reduction of SDRAM memory systems. In: *DAC '02: Proceedings of the 39th annual Design Automation Conference*, ACM, 2002. – ISBN 1–58113–461–4, S. 892–897
- [122] KANSAL, A. ; ZHAO, F. ; LIU, J. ; KOTHARI, N. ; BHATTACHARYA, A. A.: Virtual Machine Power Metering and Provisioning. In: *Proceedings of the 1st ACM Symposium on Cloud Computing*, ACM, 2010 (SoCC '10). – ISBN 978–1–4503–0036–0, S. 39–50
- [123] KHAN, K. M. ; MALLUHI, Q. : Establishing Trust in Cloud Computing. In: *IT Professional* 12 (2010), Nr. 5, S. 20–27. – DOI 10.1109/MITP.2010.128. – ISSN 1520–9202
- [124] KHAN, K. M. ; MALLUHI, Q. : Trust in Cloud Services: Providing More Controls to Clients. In: *Computer* 46 (2013), Nr. 7, S. 94–96. – DOI 10.1109/MC.2013.254. – ISSN 0018–9162
- [125] KHANNA, G. ; BEATY, K. ; KAR, G. ; KOCHUT, A. : Application Performance Management in Virtualized Server Environments. In: *Network Operations and Management Symposium, 2006. NOMS 2006. 10th IEEE/IFIP*, 2006, S. 373–381
- [126] KILPER, D. ; ATKINSON, G. ; KOROTKY, S. ; GOYAL, S. ; VETTER, P. ; SUVAKOVIC, D. ; BLUME, O. : Power Trends in Communication Networks. In: *Selected Topics in Quantum Electronics, IEEE Journal of* 17 (2011), March, Nr. 2, S. 275–284. – DOI 10.1109/JSTQE.2010.2074187. – ISSN 1077–260X

- [127] KLADNIK, B. ; ARTAC, G. ; GUBINA, A. : An assessment of the effects of demand response in electricity markets. In: *International Transactions on Electrical Energy Systems* 23 (2013), Nr. 3, S. 380–391. – DOI 10.1002/etep.666. – ISSN 2050–7038
- [128] KLOBASA, M. ; SENFUS, D. F. ; RAGWITZ, D. M.: *CO₂-Minderung im Stromsektor durch den Einsatz erneuerbarer Energien im Jahr 2006 und 2007*. Version: Februar 2009. http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-1178853.pdf, Abruf: 04.08.2015
- [129] KOOMEY, J. G.: *Estimating Total Power Consumption by Servers in the U.S. and the World*. http://www.hpcadvisorycouncil.com/pdf/vendor_content/amd_svrprwusecompletefinal.pdf. Version: Februar 2007, Abruf: 04.08.2015
- [130] KOOMEY, J. G. ; BERARD, S. ; SANCHEZ, M. ; WONG, H. : *Assessing Trends in the Electrical Efficiency of Computation over Time*. August 2009. – Report
- [131] KUSIC, D. ; KEPHART, J. O. ; HANSON, J. E. ; KANDASAMY, N. ; JIANG, G. : Power and performance management of virtualized computing environments via lookahead control. In: *Cluster Computing* 12 (2009), März, Nr. 1, S. 1–15. – DOI 10.1007/s10586-008-0070-y. – ISSN 1386–7857
- [132] LAMEHAMEDI, H. ; SHENTU, Z. ; SZYMANSKI, B. ; DEELMAN, E. : Simulation of dynamic data replication strategies in Data Grids. In: *Parallel and Distributed Processing Symposium, 2003.*, 2003. – ISSN 1530–2075, S. 10 pp.
- [133] LEAVITT, N. : Is Cloud Computing Really Ready for Prime Time? In: *Computer* 42 (2009), Jan., Nr. 1, S. 15–20. – DOI 10.1109/MC.2009.20. – ISSN 0018–9162
- [134] LEFURGY, C. ; RAJAMANI, K. ; RAWSON, F. ; FELTER, W. ; KISTLER, M. ; KELLER, T. W.: Energy Management for Commercial Servers. In: *Computer* 36 (2003), Nr. 12, S. 39–48. – DOI 10.1109/MC.2003.1250880. – ISSN 0018–9162
- [135] LEI, M. ; VRBSKY, S. V. ; HONG, X. : An on-line replication strategy to increase availability in Data Grids . In: *Future Generation Computer Systems* 24 (2008), Nr. 2, S. 85 – 98. – DOI 10.1016/j.future.2007.04.009. – ISSN 0167–739X
- [136] LEMAY, M. ; NGUYEN, K.-K. ; ARNAUD, B. S. ; CHERIET, M. : Toward a Zero-Carbon Network: Converging Cloud Computing and Network Virtualization. In: *IEEE Internet Computing* 16 (2012), Nr. 6, S. 51–59. – ISSN 1089–7801
- [137] LEVERICH, J. ; MONCHIERO, M. ; TALWAR, V. ; RANGANATHAN, P. ; KOZYRAKIS, C. : Power Management of Datacenter Workloads Using Per-Core Power Gating. In: *IEEE Computer Architecture Letters* 99 (2009), Nr. 2. – DOI 10.1109/L-CA.2009.46. – ISSN 1556–6056
- [138] LI, J. ; LI, Z. ; REN, K. ; LIU, X. : Towards Optimal Electric Demand Management for Internet Data Centers. In: *Smart Grid, IEEE Transactions on* 3 (2012), March, Nr. 1, S. 183–192. – DOI 10.1109/TSG.2011.2165567. – ISSN 1949–3053
- [139] LIU, Q. ; THEODOROPOULOS, G. ; SILVA, D. da ; LIU, E. : Towards an agent-based symbiotic architecture for autonomic management of virtualized data centers. In: *Simulation Conference (WSC), Proceedings of the 2012 Winter*, 2012. – ISSN 0891–7736, S. 1–13

- [140] LIU, Z. ; LIN, M. ; WIERMAN, A. ; LOW, S. H. ; ANDREW, L. L.: Geographical Load Balancing with Renewables. In: *SIGMETRICS Perform. Eval. Rev.* 39 (2011), Dez., Nr. 3, S. 62–66. – DOI 10.1145/2160803.2160862. – ISSN 0163–5999
- [141] LOUKOPOULOS, T. ; AHMAD, I. ; PAPADIAS, D. : An overview of data replication on the Internet. In: *Parallel Architectures, Algorithms and Networks, 2002. I-SPAN '02. Proceedings. International Symposium on*, 2002. – ISSN 1087–4089, S. 27–32
- [142] MA, J. ; LIU, W. ; GLATARD, T. : A Classification of File Placement and Replication Methods on Grids. In: *Future Gener. Comput. Syst.* 29 (2013), Aug., Nr. 6, S. 1395–1406. – DOI 10.1016/j.future.2013.02.006. – ISSN 0167–739X
- [143] MARTELLO, S. ; TOTH, P. : *Knapsack Problems: Algorithms and Computer Implementations*. John Wiley & Sons, Inc., 1990. – ISBN 0–471–92420–2
- [144] MARTIN, W. : Cloud Transformation Services – Großes Potential. In: *IT-Management* (2013), S. 24–26
- [145] MEISNER, D. ; GOLD, B. T. ; WENISCH, T. F.: PowerNap: eliminating server idle power. In: *Proceedings of the 14th international conference on Architectural support for programming languages and operating systems*, ACM, 2009 (ASPLOS '09). – ISBN 978–1–60558–406–5, S. 205–216
- [146] MELL, P. ; GRANCE, T. : The NIST Definition of Cloud Computing / National Institute of Standards and Technology (NIST). 2011. – Forschungsbericht. – 7 S.
- [147] MEZA, J. ; SHAH, M. A. ; RANGANATHAN, P. ; FITZNER, M. ; VEAZEY, J. : Tracking the Power in an Enterprise Decision Support System. In: *Proceedings of the 2009 ACM/IEEE International Symposium on Low Power Electronics and Design*, ACM, 2009 (ISLPED '09). – ISBN 978–1–60558–684–7, S. 261–266
- [148] MILANO, J. ; LEMBKE, P. ; REDBOOKS, I. : *IBM System Blue Gene Solution: Blue Gene/Q Hardware Overview and Installation Planning*. IBM Redbooks (IBM System Blue Gene Solution). <http://books.google.de/books?id=8BvCAGAAQBAJ>. – ISBN 9780738438221
- [149] MINES, C. : *Making Corporate IT a Greener Place*. März 2008. – Forrester Research, Vortrag
- [150] MÜNZL, G. ; PRZYWARA, B. ; RETI, M. ; SCHÄFER, J. ; SONDERMANN, K. ; WEBER, M. ; WILKER, A. : *Cloud Computing – Evolution in der Technik, Revolution im Business / BITKOM*. 2009. – Leitfaden. – 85 S.
- [151] MOHR, W. ; PATZAK, E. ; WALF, G. ; WOLFF, I. : Aspekte der Reduzierung des Energieverbrauchs und der Verbesserung der Energieeffizienz in Kommunikationsnetzwerken VDE-Studie / VDE. 2009. – Forschungsbericht. – 23 S.
- [152] MOHSENIAN-RAD, A. ; LEON-GARCIA, A. : Coordination of Cloud Computing and Smart Power Grids. In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, 2010, S. 368 –372

- [153] MOHSENIAN-RAD, A.-H. ; LEON-GARCIA, A. : Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments. In: *Smart Grid, IEEE Transactions on* 1 (2010), Sept, Nr. 2, S. 120–133. – DOI 10.1109/TSG.2010.2055903. – ISSN 1949–3053
- [154] MOIZ, S. A. ; SAILAJA, P. ; VENKATASWAMY, G. ; PAL, S. N.: Database Replication: A Survey of Open Source and Commercial Tools. In: *International Journal of Computer Applications* 13 (2011), January, Nr. 6, S. 1–8
- [155] MOLFETAS, A. ; MEGINO, F. B. ; TYKHONOV, A. ; LASSNIG, M. ; GARONNE, V. ; BARISITS, M. ; CAMPANA, S. ; DIMITROV, G. ; JEZEQUEL, S. ; UEDA, I. ; VIEGAS, F. T. A.: Popularity framework to process dataset traces and its application on dynamic replica reduction in the ATLAS experiment. In: *Journal of Physics: Conference Series* 331 (2011), Nr. 6, 062018. <http://stacks.iop.org/1742-6596/331/i=6/a=062018>
- [156] MOORE, J. ; CHASE, J. ; RANGANATHAN, P. : Weatherman: Automated, Online and Predictive Thermal Mapping and Management for Data Centers. In: *Autonomic Computing, 2006. ICAC '06. IEEE International Conference on*, 2006, S. 155–164
- [157] MOORE, J. ; CHASE, J. ; RANGANATHAN, P. ; SHARMA, R. : Making Scheduling “Cool”: Temperature-aware Workload Placement in Data Centers. In: *Proceedings of the Annual Conference on USENIX Annual Technical Conference*. USENIX Association (ATEC '05), 61–74
- [158] MORENO-VOZMEDIANO, R. ; MONTERO, R. S. ; LLORENTE, I. M.: IaaS Cloud Architecture: From Virtualized Datacenters to Federated Cloud Infrastructures. In: *Computer* 45 (2012), Nr. 12, S. 65–72. – DOI 10.1109/MC.2012.76. – ISSN 0018–9162
- [159] MUKHERJEE, T. ; BANERJEE, A. ; VARSAMOPOULOS, G. ; GUPTA, S. K. S. ; RUNGTA, S. : Spatio-temporal thermal-aware job scheduling to minimize energy consumption in virtualized heterogeneous data centers. In: *Comput. Netw.* 53 (2009), Dez., Nr. 17, S. 2888–2904. – DOI 10.1016/j.comnet.2009.06.008. – ISSN 1389–1286
- [160] MÜLLER-SCHLOER, C. ; SCHMECK, H. ; UNGERER, T. : Organic Computing. In: *Informatik Spektrum* 35 (2012), Nr. 2, S. 71–73. – DOI 10.1007/s00287-012-0599-2
- [161] NEBEL, W. : *Green IT ?- Energy Efficiency and ICT* Datum: 23.09.2008
- [162] NEBEL, W. ; HOYER, M. ; SCHRÖDER, K. ; SCHLITT, D. : Untersuchung des Potentials von rechenzentrenübergreifendem Lastmanagement zur Reduzierung des Energieverbrauchs in der IKT / OFFIS. Version: Dezember 2009. <http://www.lastmanagement.offis.de/>, Abruf: 04.08.2015. – Studie
- [163] NEVES, L. ; CAMPINO, I. ; KUBINK, W. ; HOLD, A. ; LING, H. ; BEHAMMER, J. ; GEMMA, M. Paolo nad C. Paolo nad Chebbo ; FRIEDMANN, P. ; SCHWARZLÄNDER, F. ; SCHÖNFELDT, A. ; BOCKEMÜHL, M. ; RUBEL, H. : Smart 2020 Addendum Deutschland: Die IKT-Industrie als treibende Kraft auf dem Weg zu nachhaltigem Klimaschutz / Deutsche Telekom, Huawei, SAP, Siemens, The Boston Consulting Group. 2009. – Studie
- [164] NGUYEN, K. ; CHERIET, M. ; LEMAY, M. ; ARNAUD, B. ; REIJS, V. ; MACKAREL, A. ; MINOVES, P. ; PASTRAMA, A. ; VAN HEDDEGHEM, W. : Renewable Energy Provisioning for ICT Services in a Future Internet. Version: 2011. In: DOMINGUE, J. (Hrsg.) ; GALIS,

- A. (Hrsg.) ; GAVRAS, A. (Hrsg.) ; ZAHARIADIS, T. (Hrsg.) ; LAMBERT, D. (Hrsg.) ; CLEARY, F. (Hrsg.) ; DARAS, P. (Hrsg.) ; KRCO, S. (Hrsg.) ; MÜLLER, H. (Hrsg.) ; LI, M.-S. (Hrsg.) ; SCHAFFERS, H. (Hrsg.) ; LOTZ, V. (Hrsg.) ; ALVAREZ, F. (Hrsg.) ; STILLER, B. (Hrsg.) ; KARNOUSKOS, S. (Hrsg.) ; AVESSTA, S. (Hrsg.) ; NILSSON, M. (Hrsg.): *The Future Internet* Bd. 6656. Springer Berlin Heidelberg, 2011. – DOI 10.1007/978-3-642-20898-0_30. – ISBN 978-3-642-20897-3, S. 419–429
- [165] NGUYEN, X.-N. ; SAUCEZ, D. ; BARAKAT, C. ; TURLETTI, T. : Optimizing Rules Placement in OpenFlow Networks: Trading Routing for Better Efficiency. In: *Proceedings of the Third Workshop on Hot Topics in Software Defined Networking*, ACM, 2014 (HotSDN '14). – ISBN 978-1-4503-2989-7, S. 127–132
- [166] NICOLAI, M. : IT-Klimatisierung / RITTAL. 2005 (2). – White Paper
- [167] NIEMANN, J. : Hot Aisle vs. Cold Aisle Containment / APC. 2008 (135). – White Paper
- [168] *Open Cloud Manifesto*. <http://www.opencloudmanifesto.org>. Version: 2009, Abruf: 05.10.2009
- [169] ORTIZ, S. : The Problem with Cloud-Computing Standardization. In: *Computer* 44 (2011), July, Nr. 7, S. 13–16. – DOI 10.1109/MC.2011.220. – ISSN 0018-9162
- [170] OSTERBURG, S. ; PINNOW, A. ; RAUTENSTRAUCH, C. ; WINTER, M. : Neue Computing-Grundlagen für das Rechenzentrum. In: *Informatik-Spektrum* 32 (2009), Nr. 2, S. 118–126. – DOI 10.1007/s00287-008-0308-3. – ISSN 0170-6012
- [171] PADMANABHAN, P. ; GRUENWALD, L. ; VALLUR, A. ; ATIQUZZAMAN, M. : A survey of data replication techniques for mobile ad hoc network databases. In: *The VLDB Journal* 17 (2008), Nr. 5, S. 1143–1164. – DOI 10.1007/s00778-007-0055-0. – ISSN 1066-8888
- [172] PAKBAZNIA, E. ; GHASEMAZAR, M. ; PEDRAM, M. : Temperature-aware Dynamic Resource Provisioning in a Power-optimized Datacenter. In: *Proceedings of the Conference on Design, Automation and Test in Europe*. European Design and Automation Association (DATE '10). – ISBN 978-3-9810801-6-2, 124–129
- [173] PATEL, C. ; SHARMA, R. ; BASH, C. ; GRAUPNER, S. : Energy aware grid: Global workload placement based on energy efficiency. In: *ASME 2003 International Mechanical Engineering Congress and Exposition* American Society of Mechanical Engineers, 2003, S. 267–275
- [174] PEDRAM, M. ; HWANG, I. : Power and Performance Modeling in a Virtualized Server System. In: *Proceedings of the 2010 39th International Conference on Parallel Processing Workshops*. Washington, DC, USA : IEEE Computer Society, 2010 (ICPPW '10). – ISBN 978-0-7695-4157-0, S. 520–526
- [175] PELLE, S. ; MEISNER, D. ; WENISCH, T. ; VANGILDER, J. W.: Understanding and Abstracting Total Data Center Power. In: *Workshop on Energy Efficient Design (WEED)*, 2009
- [176] PIERRE, G. ; STEEN, M. van ; TANENBAUM, A. : Dynamically selecting optimal distribution strategies for Web documents. In: *Computers, IEEE Transactions on* 51 (2002), Jun, Nr. 6, S. 637–651. – DOI 10.1109/TC.2002.1009149. – ISSN 0018-9340

- [177] PIETTE, M. A. ; GHATIKAR, G. ; KILICCOTE, S. ; KOCH, E. ; HENNAGE, D. ; PALENSKY, P. ; MCPARLAND, C. : Open Automated Demand Response Communications Specification (Version 1.0) / Public Interest Energy Research (PIER), California Energy Commission. 2009. – Report
- [178] PRICE, M. : The Paradox of Security in Virtual Environments. In: *Computer* 41 (2008), Nr. 11, S. 22–28. – DOI 10.1109/MC.2008.472. – ISSN 0018–9162
- [179] PRIMEENERGYIT: Best Practice Beispiele für energieeffiziente Rechenzentren / Intelligent Energy, Europe. 2012. – White Paper. – 42 S.
- [180] PU, X. ; LIU, L. ; MEI, Y. ; SIVATHANU, S. ; KOH, Y. ; PU, C. : Understanding Performance Interference of I/O Workload in Virtualized Cloud Environments. In: *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, 2010, S. 51–58
- [181] PULVERMÜLLER, P. : *Bestandsaufnahme effiziente Rechenzentren in Deutschland*. 2008. – eco – Verband der deutschen Internetwirtschaft e.V., Report
- [182] QURESHI, A. ; WEBER, R. ; BALAKRISHNAN, H. ; GUTTAG, J. ; MAGGS, B. : Cutting the Electric Bill for Internet-scale Systems. In: *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication*, ACM, 2009 (SIGCOMM '09). – ISBN 978–1–60558–594–9, S. 123–134
- [183] RAHMAN, R. ; BARKER, K. ; ALHAJJ, R. : Replica Placement Strategies in Data Grid. In: *Journal of Grid Computing* 6 (2008), Nr. 1, S. 103–123. – DOI 10.1007/s10723–007–9090–8. – ISSN 1570–7873
- [184] RAMAKRISHNAN, K. K. ; SHENOY, P. ; MERWE, J. Van d.: Live Data Center Migration Across WANs: A Robust Cooperative Context Aware Approach. In: *INM '07*. New York, NY, USA : ACM, 2007. – ISBN 978–1–59593–788–9, S. 262–267
- [185] RAMCHURN, S. D. ; VYTELINGUM, P. ; ROGERS, A. ; JENNINGS, N. R.: Putting the “Smarts” into the Smart Grid: A Grand Challenge for Artificial Intelligence. In: *Commun. ACM* 55 (2012), Apr., Nr. 4, S. 86–97. – DOI 10.1145/2133806.2133825. – ISSN 0001–0782
- [186] RANGANATHAN, K. ; FOSTER, I. : Computation Scheduling and Data Replication Algorithms for Data Grids. Version: 2004. In: NABRZYSKI, J. (Hrsg.) ; SCHOPF, J. M. (Hrsg.) ; WĘGLARZ, J. (Hrsg.): *Grid Resource Management* Bd. 64. Springer US, 2004. – DOI 10.1007/978–1–4615–0509–9_22. – ISBN 978–1–4613–5112–2, S. 359–373
- [187] RAO, L. ; LIU, X. ; XIE, L. ; LIU, W. : Minimizing Electricity Cost: Optimization of Distributed Internet Data Centers in a Multi-Electricity-Market Environment. In: *INFOCOM, 2010 Proceedings IEEE*, 2010. – ISSN 0743–166X, S. 1–9
- [188] RASMUSSEN, N. : AC vs. DC Power Distribution for Data Centers / Schneider Electric. 2011. – White Paper
- [189] RAWSON, A. ; PFLUEGER, J. ; SADER, T. ; BELADY, C. (Hrsg.): GREEN GRID DATA CENTER POWER EFFICIENCY METRICS: PUE AND DCiE / the green grid. Version: 2008. <http://www.thegreengrid.org/Global/Content/white-papers/The-Green-Grid-Data-Center-Power-Efficiency-Metrics-PUE-and-DCiE>, Abruf: 04.08.2015. – White Paper

- [190] RITTAL: Whitepaper Rittal Liquid Cooling Packages / Rittal. Version:2015. <http://whitepaper.computerwoche.de/whitepaper/das-liquid-cooling-package-lcp-massstab-fuer-effizienz>, Abruf: 04.08.2015. – White Paper
- [191] RIVOIRE, S. ; SHAH, M. ; RANGANATHAN, P. ; KOZYRAKIS, C. ; MEZA, J. : Models and Metrics to Enable Energy-Efficiency Optimizations. In: *Computer* 40 (2007), Dec, Nr. 12, S. 39–48. – DOI 10.1109/MC.2007.436. – ISSN 0018–9162
- [192] RIVOIRE, S. ; RANGANATHAN, P. ; KOZYRAKIS, C. : A Comparison of High-level Full-system Power Models. In: *Proceedings of the 2008 Conference on Power Aware Computing and Systems*. USENIX Association (HotPower’08)
- [193] ROLIA, J. ; ANDRZEJAK, A. ; ARLITT, M. F.: Automating Enterprise Application Placement in Resource Utilities. In: BRUNNER, M. (Hrsg.) ; KELLER, A. (Hrsg.): *Self-Managing Distributed Systems* Bd. 2867, Springer Berlin Heidelberg, 2003 (Lecture Notes in Computer Science). – ISBN 978–3–540–20314–8, S. 118–129
- [194] RYCKBOSCH, F. ; POLFLIET, S. ; EECKHOUT, L. : Trends in Server Energy Proportionality. In: *Computer* 44 (2011), Sept, Nr. 9, S. 69–72. – DOI 10.1109/MC.2011.130. – ISSN 0018–9162
- [195] SACHVERSTÄNDIGENRAT FÜR UMWELTFRAGEN: *Wege zur 100 % erneuerbaren Stromversorgung: Sondergutachten*. Erich Schmidt Verlag, 2011. – ISBN 9783503136063
- [196] SANKARANARAYANAN, A. N. ; SHARANGI, S. ; FEDOROVA, A. : Global Cost Diversity Aware Dispatch Algorithm for Heterogeneous Data Centers. In: *Proceedings of the 2Nd ACM/SPEC International Conference on Performance Engineering*, ACM, 2011 (ICPE ’11). – ISBN 978–1–4503–0519–8, S. 289–294
- [197] SATO, H. ; MATSUOKA, S. ; ENDO, T. ; MARUYAMA, N. : Access-pattern and Bandwidth Aware File Replication Algorithm in a Grid Environment. In: *Proceedings of the 2008 9th IEEE/ACM International Conference on Grid Computing*. Washington, DC, USA : IEEE Computer Society, 2008 (GRID ’08). – ISBN 978–1–4244–2578–5, S. 250–257
- [198] SAUNDERS, H. D.: The Khazzoom-Brookes postulate and neoclassical growth. In: *The Energy Journal* (1992), S. 131–148
- [199] SCHLARB, A. : *Potentiale der IT-gestützten Energieeffizienz in Gebäuden*. Oktober 2011. – Siemens, Vortrag
- [200] SCHLITT, D. ; NEBEL, W. : Load dependent data center energy efficiency metric based on component models. In: *Proceedings of 2012 International Conference on Energy Aware Computing (ICEAC 2012)* IEEE, 2012, S. 20–25
- [201] SERRANO, D. ; PATIÑO-MARTÍNEZ, M. ; JIMÉNEZ-PERIS, R. ; KEMME, B. : Boosting Database Replication Scalability through Partial Replication and 1-Copy-Snapshot-Isolation. In: *PRDC*, 2007, S. 290–297
- [202] SHARMA, R. ; BASH, C. ; PATEL, C. ; FRIEDRICH, R. ; CHASE, J. : Balance of power: dynamic thermal management for Internet data centers. In: *Internet Computing, IEEE* 9 (2005), Jan, Nr. 1, S. 42–49. – DOI 10.1109/MIC.2005.10. – ISSN 1089–7801

- [203] SIMOES, M. ; ROCHE, R. ; KYRIAKIDES, E. ; SURYANARAYANAN, S. ; BLUNIER, B. ; MCBEE, K. ; NGUYEN, P. ; RIBEIRO, P. ; MIRAOU, A. : A Comparison of Smart Grid Technologies and Progresses in Europe and the U.S. In: *Industry Applications, IEEE Transactions on* 48 (2012), July, Nr. 4, S. 1154–1162. – DOI 10.1109/TIA.2012.2199730. – ISSN 0093–9994
- [204] SIMON, G. J.: *Data Center Challenges for Webhosters*. März 2009. – Equinix, Vortrag
- [205] SIRIWARDANA, J. ; SCHOTT, W. ; HALGAMUGE, S. : The power grabbers. In: *Power and Energy Magazine, IEEE* 8 (2010), January, Nr. 1, S. 46–53. – DOI 10.1109/MEPE.2009.934881. – ISSN 1540–7977
- [206] SIVASUBRAMANIAN, S. ; ALONSO, G. ; PIERRE, G. ; STEEN, M. van: GlobeDB: Autonomic Data Replication for Web Applications. In: *Proceedings of the 14th International Conference on World Wide Web, ACM, 2005 (WWW '05)*. – ISBN 1–59593–046–9, S. 33–42
- [207] SIVASUBRAMANIAN, S. ; PIERRE, G. ; STEEN, M. van: Web Content Caching and Distribution. Version: 2004. <http://dl.acm.org/citation.cfm?id=1061831.1061851>. Kluwer Academic Publishers. – ISBN 1–4020–2257–3, Kapitel A Case for Dynamic Selection of Replication and Caching Strategies, 275–282
- [208] SOLOMON, B. ; IONESCU, D. ; LITOIU, M. ; ISZLAI, G. : Designing autonomic management systems for cloud computing. In: *Computational Cybernetics and Technical Informatics (ICCC-CONTI), 2010 International Joint Conference on*, 2010, S. 631–636
- [209] SOTOMAYOR, B. ; MONTERO, R. S. ; LLORENTE, I. M. ; FOSTER, I. : Virtual Infrastructure Management in Private and Hybrid Clouds. In: *IEEE Internet Computing* 13 (2009), S. 14–22. – DOI 10.1109/MIC.2009.119. – ISSN 1089–7801
- [210] SPEES, K. ; LAVE, L. : Demand Response and Electricity Market Efficiency. In: *The Electricity Journal* 20 (2007), Nr. 3, 69–85. <http://EconPapers.repec.org/RePEc:eee:jelect:v:20:y:2007:i:3:p:69-85>, Abruf: 04.08.2015
- [211] SRIKANTAIHAH, S. ; KANSAL, A. ; ZHAO, F. : Energy Aware Consolidation for Cloud Computing. In: *Proceedings of the 2008 Conference on Power Aware Computing and Systems*. USENIX Association (HotPower'08)
- [212] STIELER, W. : Von der Börse an die Steckdose. In: *Technology Review* (2008), 09, S. 27–32
- [213] STOBBE, L. ; NISSEN, D.-I. N. F. ; PROSKE, M. ; MIDDENDORF, A. ; SCHLOMANN, B. ; FRIEDEWALD, D. M. ; GEORGIEFF, P. ; LEIMBACH, T. : Abschätzung des Energiebedarfs der weiteren Entwicklung der Informationsgesellschaft / Fraunhofer IZM and ISI. Version: März 2009. <http://www.bmwi.de/DE/Mediathek/publikationen,did=305306.html>, Abruf: 04.08.2015. – Abschlussbericht
- [214] STOCKINGER, H. ; SAMAR, A. ; ALLCOCK, B. ; FOSTER, I. ; HOLTMAN, K. ; TIERNEY, B. : File and object replication in data grids. In: *High Performance Distributed Computing, 2001. Proceedings. 10th IEEE International Symposium on*, 2001. – ISSN 1082–8907, S. 76–86
- [215] SURY, U. : Cloud Computing und Recht. In: *Informatik-Spektrum* 32 (2009), S. 183–184

- [216] SVÄRD, P. ; HUDZIA, B. ; TORDSSON, J. ; ELMROTH, E. : Evaluation of delta compression techniques for efficient live migration of large virtual machines. In: *Proceedings of the 7th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments*, ACM, 2011 (VEE '11). – ISBN 978-1-4503-0687-4, S. 111–120
- [217] TANG, M. ; LEE, B.-S. ; TANG, X. ; YEO, C.-K. : The impact of data replication on job scheduling performance in the Data Grid. In: *Future Generation Computer Systems* 22 (2006), Nr. 3, S. 254 – 268. – DOI 10.1016/j.future.2005.08.004. – ISSN 0167-739X
- [218] TANG, Q. ; GUPTA, S. ; VARSAMOPOULOS, G. : Energy-Efficient Thermal-Aware Task Scheduling for Homogeneous High-Performance Computing Data Centers: A Cyber-Physical Approach. In: *Parallel and Distributed Systems, IEEE Transactions on* 19 (2008), Nov, Nr. 11, S. 1458–1472. – DOI 10.1109/TPDS.2008.111. – ISSN 1045-9219
- [219] TANG, X. ; XU, J. : QoS-aware replica placement for content distribution. In: *Parallel and Distributed Systems, IEEE Transactions on* 16 (2005), Oct, Nr. 10, S. 921–932. – DOI 10.1109/TPDS.2005.126. – ISSN 1045-9219
- [220] TAYLOR, C. ; KOOMEY, J. : Estimating Energy Use and Greenhouse Gas Emissions of Internet Advertising: Working Paper Prepared for IMC2 / IMC2. Version: Febr. 2008. <http://www.imc2.com/documents/carbonemissions.pdf>, Abruf: 12.09.2012. – Forschungsbericht
- [221] THOME, H. : Univariate Box/Jenkins-Modelle in der Zeitreihenanalyse. In: *Historical Social Research* 19 (1994), Nr. 3, 5-77. <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-32646>, Abruf: 04.08.2015
- [222] TIETZ, V. ; BLICHMANN, G. ; HÜBSCH, G. : Cloud-Entwicklungsmethoden. In: *Informatik-Spektrum* 34 (2011), Nr. 4, S. 345–354. – DOI 10.1007/s00287-011-0531-1. – ISSN 0170-6012
- [223] TRAVOSTINO, F. ; DASPIT, P. ; GOMMANS, L. ; JOG, C. ; LAAT, C. de ; MAMBRETTI, J. ; MONGA, I. ; OUDENAARDE, B. van ; RAGHUNATH, S. ; WANG, P. Y.: Seamless live migration of virtual machines over the MAN/WAN. In: *SC '06*. New York, NY, USA : ACM, 2006. – ISBN 0-7695-2700-0, S. 290
- [224] TRAVOSTINO, F. ; DASPIT, P. ; GOMMANS, L. ; JOG, C. ; LAAT, C. de ; MAMBRETTI, J. ; MONGA, I. ; OUDENAARDE, B. van ; RAGHUNATH, S. ; WANG, P. Y.: Seamless live migration of virtual machines over the MAN/WAN. In: *Future Gener. Comput. Syst.* 22 (2006), Okt., Nr. 8, S. 901–907. – DOI 10.1016/j.future.2006.03.007. – ISSN 0167-739X
- [225] TSIROGIANNIS, D. ; HARIZOPOULOS, S. ; SHAH, M. A.: Analyzing the Energy Efficiency of a Database Server. In: *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, ACM, 2010 (SIGMOD '10). – ISBN 978-1-4503-0032-2, S. 231–242
- [226] UPTIME INSTITUTE, L. : Data Center Site Infrastructure – Tier Standard: Topology / Uptime Institute. 2010. – Report
- [227] VAUGHAN-NICHOLS, S. : OpenFlow: The Next Generation of the Network? In: *Computer* 44 (2011), Aug, Nr. 8, S. 13–15. – DOI 10.1109/MC.2011.250. – ISSN 0018-9162

- [228] VAZHKUDAI, S. ; SCHOPF, J. M.: Using Regression Techniques to Predict Large Data Transfers. In: *Int. J. High Perform. Comput. Appl.* 17 (2003), Aug., Nr. 3, S. 249–268. – DOI 10.1177/1094342003173004. – ISSN 1094–3420
- [229] VENUGOPAL, S. ; BUYYA, R. ; RAMAMOCHANARAO, K. : A Taxonomy of Data Grids for Distributed Data Sharing, Management, and Processing. In: *ACM Comput. Surv.* 38 (2006), Jun., Nr. 1. – DOI 10.1145/1132952.1132955. – ISSN 0360–0300
- [230] VERMA, A. ; AHUJA, P. ; NEOGI, A. : pMapper: power and migration cost aware application placement in virtualized systems. In: *Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware*. Springer-Verlag New York, Inc. (Middleware '08). – ISBN 3–540–89855–7, 243–264
- [231] VERMA, A. ; DASGUPTA, G. ; NAYAK, T. K. ; DE, P. ; KOTHARI, R. : Server workload analysis for power minimization using consolidation. In: *Proceedings of the 2009 conference on USENIX Annual technical conference*. USENIX Association (USENIX'09), 28–28
- [232] VINAY KULKARNI, K. S.: Evaluating Microsoft Hyper-V Live Migration Performance Using IBM System x3650 M3 and IBM Systems Storage DS3400 / IBM. 2010. – White Paper
- [233] VISHWANATH, A. ; ZHU, J. ; HINTON, K. ; AYRE, R. ; TUCKER, R. : Estimating the energy consumption for packet processing, storage and switching in optical-IP routers. In: *Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC), 2013*, 2013, S. 1–3
- [234] VMWARE: VMware vCloud Architecting a vCloud / VMware. Version:2010. <http://www.vmware.com/files/pdf/VMware-Architecting-vCloud-WP.pdf>, Abruf: 04.08.2015. – White Paper
- [235] VOORSLUYS, W. ; BROBERG, J. ; VENUGOPAL, S. ; BUYYA, R. : Cost of Virtual Machine Live Migration in Clouds: A Performance Evaluation. In: *Proceedings of the 1st International Conference on Cloud Computing*. Springer-Verlag (CloudCom '09). – ISBN 978–3–642–10664–4, 254–265
- [236] WALLNER, A. : Das Rechenzentrum der Zukunft: Shared IT Infrastructure. In: *ntz* 1 (2011), S. 54–55
- [237] WANG, H. ; HUANG, J. ; LIN, X. ; MOHSENIAN-RAD, H. : Exploring Smart Grid and Data Center Interactions for Electric Power Load Balancing. In: *SIGMETRICS Perform. Eval. Rev.* 41 (2014), Jan., Nr. 3, S. 89–94. – DOI 10.1145/2567529.2567556. – ISSN 0163–5999
- [238] WANG, R. ; KANDASAMY, N. ; NWANKPA, C. ; KAELI, D. R.: Datacenters as Controllable Load Resources in the Electricity Market. In: *ICDCS, IEEE*, 2013, S. 176–185
- [239] WEI, Q. ; VEERAVALLI, B. ; GONG, B. ; ZENG, L. ; FENG, D. : CDRM: A Cost-Effective Dynamic Replication Management Scheme for Cloud Storage Cluster. In: *Cluster Computing (CLUSTER), 2010 IEEE International Conference on*, 2010, S. 188–196

- [240] WEINHARDT, C. ; ANANDASIVAM, A. ; BLAU, B. ; BORISSOV, N. ; MEINL, T. ; MICHALK, W. ; STÖSSER, J. : Cloud Computing ? A Classification, Business Models, and Research Directions. In: *Business & Information Systems Engineering* 1 (2009), Nr. 5, S. 391–399. – DOI 10.1007/s12599-009-0071-2
- [241] WENGORZ, J. : *Transparenz im Data Center*. 2007. – SAS Institute, Vortrag
- [242] WINTER, M. : Data Center Consolidation: A Step Towards Infrastructure Clouds. In: *Proceedings of the 1st International Conference on Cloud Computing*. Berlin, Heidelberg : Springer-Verlag, 2009 (CloudCom '09). – ISBN 978-3-642-10664-4, S. 190–199
- [243] WÖGINGER, G. J.: Exact Algorithms for NP-Hard Problems: A Survey. Version: 2003. In: JÜNGER, M. (Hrsg.) ; REINELT, G. (Hrsg.) ; RINALDI, G. (Hrsg.): *Combinatorial Optimization ? Eureka, You Shrink!* Bd. 2570. Springer Berlin Heidelberg, 2003. – DOI 10.1007/3-540-36478-1_17. – ISBN 978-3-540-00580-3, S. 185–207
- [244] WOLSKI, R. : Forecasting network performance to support dynamic scheduling using the network weather service. In: *High Performance Distributed Computing, 1997. Proceedings. The Sixth IEEE International Symposium on*, 1997. – ISSN 1082-8907, S. 316–325
- [245] WOLSKI, R. ; SPRING, N. T. ; HAYES, J. : The Network Weather Service: A Distributed Resource Performance Forecasting Service for Metacomputing. In: *Future Gener. Comput. Syst.* 15 (1999), Okt., Nr. 5-6, S. 757–768. – DOI 10.1016/S0167-739X(99)00025-4. – ISSN 0167-739X
- [246] WOOD, T. ; RAMAKRISHNAN, K. K. ; SHENOY, P. ; MERWE, J. van d.: CloudNet: dynamic pooling of cloud resources by live WAN migration of virtual machines. In: *SIGPLAN Not.* 46 (2011), März, Nr. 7, S. 121–132. – DOI 10.1145/2007477.1952699. – ISSN 0362-1340
- [247] WOOD, T. ; SHENOY, P. ; VENKATARAMANI, A. ; YOUSIF, M. : Black-box and gray-box strategies for virtual machine migration. In: *Proceedings of the 4th USENIX conference on Networked systems design & implementation*. USENIX Association (NSDI'07), 229–242
- [248] WU, J.-J. ; LIN, Y.-F. ; LIU, P. : Optimal Replica Placement in Hierarchical Data Grids with Locality Assurance. In: *J. Parallel Distrib. Comput.* 68 (2008), Dez., Nr. 12, S. 1517–1538. – DOI 10.1016/j.jpdc.2008.08.002. – ISSN 0743-7315
- [249] XU, J. ; ZHAO, M. ; FORTES, J. ; CARPENTER, R. ; YOUSIF, M. : Autonomic Resource Management in Virtualized Data Centers Using Fuzzy Logic-based Approaches. In: *Cluster Computing* 11 (2008), Sept., Nr. 3, S. 213–227. – DOI 10.1007/s10586-008-0060-0. – ISSN 1386-7857
- [250] ZHANG, Y. ; WANG, Y. ; WANG, X. : GreenWare: Greening Cloud-scale Data Centers to Maximize the Use of Renewable Energy. In: *Proceedings of the 12th ACM/IFIP/USENIX International Conference on Middleware*. Berlin, Heidelberg : Springer-Verlag, 2011 (Middleware'11). – ISBN 978-3-642-25820-6, S. 143–164
- [251] ZYBELL, G. ; WAGNER, E. : *Wie setzen sich die Strompreise zusammen (Stand 2006)?* <http://www.energie-fakten.de/pdf/strompreise-2006.pdf>. Version: Juli 2006, Abruf: 13.09.2014

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Oldenburg, den 10. Juli 2016