



**Ressourcen-Provisionierungsverfahren zur Erweiterung
energetischer Freiheitsgrade von Rechenzentren unter
Veränderung der IT-Last**

Von der Fakultät II für Informatik, Wirtschafts- und
Rechtswissenschaften der Carl von Ossietzky Universität Oldenburg zur
Erlangung des Grades und Titels eines

Doktors der Naturwissenschaften (Dr. rer. nat)

angenommene Dissertation

von Herrn Alexander Borgerding

Gutachter: Prof. Dr.-Ing. Wolfgang Nebel
Weiterer Gutachter: Prof. i.R. Dr. Michael Sonnenschein
Tag der Disputation: 15.02.2021

Zusammenfassung

Durch die voranschreitende Digitalisierung und Etablierung neuer IT-Services ist ein steigender Bedarf an Rechenzentrumsdienstleistungen zu verzeichnen, folglich steigt auch der Energiebedarf in diesem Bereich an und stellt ein relevantes Volumen im Strommarkt dar. Im Energiesektor rückt der Fokus verstärkt auf die erneuerbaren und klimaneutralen Energien. Hierbei ergibt sich neben der Bereitstellung der benötigten Kapazitäten auch die Herausforderung, eine kontinuierliche Stromversorgung sicherzustellen. Je höher der Anteil an volatilen Energielieferanten steigt, umso aufwendiger wird auch die entsprechende Kompensationsleistung.

Ein Ansatz, den Kompensationsbedarf zu reduzieren, ist es, den Energiebedarf entsprechend der aktuellen Verfügbarkeit zu steuern. Der Energiebedarf von Rechenzentren stellt dabei ein attraktives, wachsendes Volumen für diese Aufgabe dar. Allerdings lässt sich die Leistungsaufnahme eines nativen Rechenzentrums nicht oder nur unzureichend steuern. Maßgeblich wird die Leistungsaufnahme von der Rechenleistung bestimmt, die durch die jeweils betriebenen Applikationen abgerufen wird. In dieser Arbeit wird ein Verfahren vorgestellt, das es ermöglicht, die abgerufene Rechenleistung dennoch zu beeinflussen und die energetischen Freiheitsgrade von Rechenzentren zu erweitern.

Auf Basis einer Ressourcen-Überprovisionierung wird ein heuristisches Verfahren vorgestellt, um die Performance der betriebenen Applikationen zu beeinflussen und damit auch die Leistungsaufnahme der Server zu steuern. Mit Server-Virtualisierung als Basistechnologie werden virtuelle Maschinen als Vehikel zum Verschieben von Rechenlast verwendet und so eine Serverauslastung erzeugt, die einer geforderten Leistungsaufnahme entspricht. Mit diesem Ansatz kann eine Steuerung der Server-Leistungsaufnahme erreicht werden, ohne dass Zugriff auf die betriebenen Applikationen benötigt wird oder die fachliche Aufgabe der Applikationen durchdrungen werden muss.

Das in dieser Arbeit entwickelte Online Heuristic Machine Allocation (OHMA)-Verfahren wird als Automaten-Modell in der Simulationsumgebung Uppaal modelliert, um eine transparente Prozessdarstellung zu ermöglichen, und wird anhand von Datensätzen aus der Industrie evaluiert. Es wird gezeigt, dass das heuristische Verfahren vergleichbare Ergebnisse liefert wie bekannte Ansätze, die jedoch auf Scheduling-Technologie basieren und einer fachlichen Implementierung in den betriebenen Applikationen bedürfen.

Das OHMA-Verfahren kann mit linearer Komplexität beschrieben werden und zeigte in synthetischen Experimenten vergleichbare Ergebnisse wie eine bekannte Optimal-Lösung. Es wird eine Leistungsanpassung von bis zu 50% mit der Technik einer Ressourcen-Überprovisionierung in Kombination mit der Migration virtueller Maschinen erreicht.

Abstract

Due to the advancing digitalization and establishment of new IT services, there is a growing demand for data center services, and at the same time the energy demand in this area is also increasing, thus representing a relevant volume in the electricity market. In the energy sector, the focus is now increasingly shifting to renewable and climate-neutral energies. In addition to providing the required capacities, this also poses the challenge of ensuring a continuous supply of electricity. The availability of solar and wind energy can vary greatly. In order to balance out these fluctuations, several types of power plants are always combined for stable grid operation. The higher the proportion of volatile energy suppliers increases, the more expensive the corresponding compensation line becomes.

One approach to reducing the need for compensation is to manage energy demand according to current availability. The energy requirements of data centers represent an attractive, growing volume for this task. However, the power consumption of a native data center cannot be controlled or only insufficiently controlled. The power consumption is largely determined by the data processing line, which is accessed by the applications running in each case. In this thesis, a method is presented which allows to influence the retrieved data line nevertheless and to extend the energetic degrees of freedom of data centers.

Based on resource over-provisioning, a heuristic method is presented for affecting the performance of the applications being operated and thus also controlling the power consumption of the servers. With server virtualization as the base technology, virtual machines are used as a vehicle to shift computational load, creating a server workload that corresponds to a required power consumption. With this approach, control of server power consumption can be achieved without requiring access to the applications being operated or having to penetrate the business task of the applications.

The Online Heuristic Machine Allocation (OHMA) approach developed in this thesis is modeled as an automaton model in the simulation environment Uppaal to enable a transparent process representation, and is evaluated using data sets from industry. It is shown that the heuristic method provides comparable results to known approaches, which, however, are based on scheduling technology and require a specialized implementation in the operated applications.

The OHMA approach can be described with linear complexity and showed comparable results to a known optimal solution in synthetic experiments. A performance adjustment of up to 50% is achieved with the technique of resource overprovisioning in combination with virtual machine migration.