



ENERGY EFFICIENT VIRTUAL MACHINE CONSOLIDATION IN CLOUD COMPUTING USING MULTIPLE FACTOR REGRESSION ALGORITHMS

By

Amany AbdElSamea Saeed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Computer Engineering

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

ENERGY EFFICIENT VIRTUAL MACHINE CONSOLIDATION IN CLOUD COMPUTING USING MULTIPLE FACTOR REGRESSION ALGORITHMS

By

Amany AbdelSamea Saeed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Computer Engineering

Under the Supervision of

Prof. ElSayed E. Hemayed

Professor

Computer Engineering Department
Faculty of Engineering, Cairo University

Prof. Hesham E. ElDeeb

Professor

Computer Engineering and System Department
Electronics Research Institute

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT

2017

ENERGY EFFICIENT VIRTUAL MACHINE CONSOLIDATION IN CLOUD COMPUTING USING MULTIPLE FACTOR REGRESSION ALGORITHMS

By

Amany AbdElSamea Saeed

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Computer Engineering

Approved by the Examining Committee:

Prof. ElSayed E. Hemayed, Thesis Main Advisor

Prof. Hesham E. ElDeeb, Advisor
Electronics Research Institute

Prof. Nevin M. Darwish, Internal Examiner

Prof. Salwa M. Nassar, External Examiner
Electronics Research Institute

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

Engineer's Name: Amany AbdElSamea Saeed
Date of Birth: 25/07/1976
Nationality: Egyptian
E-mail: amany@eri.sci.eg
Phone: 01273285455
Address: El dokki
Registration Date: 1/10/2010
Awarding Date: 2017
Degree: Doctor of Philosophy
Department: Computer Engineering

Supervisors:

Prof. ElSayed E. Hemayed
Prof. Hesham E. ElDeeb

Examiners:

Prof. ElSayed E. Hemayed (Thesis main advisor)
Prof. Hesham E. ElDeeb (Advisor) (Electronics Research Institute)
Prof. Nevin M. Darwish (Internal examiner)
Prof. Salwa M. Nassar (External examiner) (Electronics Research Institute)

Title of Thesis:

Energy Efficient Virtual Machine Consolidation In Cloud Computing Using Multiple Factor Regression Algorithms.

Key Words:

VM consolidation; Host overload detection; Multiple regression; Single factors; Multiple factors

Summary:

Green Computing is a recent trend towards designing, building, and operating computer systems to be energy efficient. Data centers are intensive consumers of energy both to power the computers and to provide the necessary cooling. Moreover, cloud providers have magnified the problem by building increased numbers of energy hungry data. Virtual Machine (VM) consolidation is an effective way to improve the utilization of resources and increase energy efficiency in cloud data centers. Most of current researches migrate VMs based on CPU utilization only for host overload detection. In this thesis we propose enhancement of VM consolidation using multiple factors Regression Host Overload Detection algorithm and normalization techniques.

Acknowledgements

The Whole Gratitude is due to ALLAH

I would like to express my deep and sincere gratitude to Dr Ali A. El-Moursy. Dr Ali has significant contribution in my work and during the revision of my thesis. Without Dr Ali's contribution in my work and thesis the work will never appear in that professional level. So I would like to thank Dr Ali for his support, guidance and patience.

I would like to express my deep and sincere gratitude to my supervisors Prof. ElSayed Essa Hemayed, Prof. Hesham Ezzet Eldeeb for their endless support, guidance and patience.

Appreciation is expressed to the examiners Prof. Salwa M. Nassar and Prof. Nevin M. Darwish for their valuable comments on the thesis.

I would like to express my deepest gratitude to my family for their support, encouragement, patience, and prayers.

Appreciation is expressed to my colleagues at Electronics Research Institute for their support guidance and encouragement

Dedication

To My Family

Table of Contents

ACKNOWLEDGEMENT.....	i
TABLE OF CONTENTS.....	iii
LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
LIST OF SYMBOLS AND ABBREVIATIONS	viii
ABSTRACT.....	x
1 INTRODUCTION.....	1
1.1 Motivation	1
1.2 Contribution	3
1.3 Organization of the Thesis	3
2 CLOUD COMPUTING.....	5
2.1 Definition	5
2.2 Evolution of Cloud Computing	5
2.3 Layered Architecture	8
2.4 Cloud Computing Model	9
2.4.1 Essential Characteristics	9
2.4.2 Service Models	10
2.4.3 Deployment Models	12
2.5 Challenges	12
2.6 Summary	13
3 POWER MANAGEMENT.....	15
3.1 Green Computing	15
3.2 Power Consumption and Energy Consumption	16
3.3 Energy Consumption Modeling	17
3.3.1 Linear Power Models	17
3.3.2 Empirical Power Models	20
3.4 Power Management Categories	21
3.4.1 Node Level Optimization	22
3.4.2 Dynamic Voltage and Frequency Scaling (DVFS)	23
3.4.3 Thermal-aware Scheduling Techniques	24
3.4.4 Dynamic Virtual Machine Consolidation	25
3.5 Summary	26
4 VIRTUAL MACHINE CONSOLIDATION.....	27
4.1 Challenges and Opportunities of VM Consolidation	27
4.2 Virtual Machine Consolidation Steps	29
4.3 Host Overload Detection Algorithms	30
4.3.1 Static CPU Utilization Threshold Algorithms	30

4.3.2 Adaptive Utilization Threshold based Algorithms	31
4.3.3 Regression based Algorithms	33
4.4 VM Selection Algorithms	35
4.4.1 Fixed-Criterion Techniques	35
4.4.2 Dynamic Criteria Techniques	35
4.5 VM Live Migration Algorithms.....	36
4.6 Virtual Machine Placement Techniques	38
4.6.1 Taxonomy of Virtual Machine Placement Techniques	39
4.6.2 Comparison of Various Virtual Machine Placement Strategies . .	39
4.7 Summary	41
5 HYBRID HOST OVERLOAD DETECTION ALGORITHMS	42
5.1 Multiple Regression	44
5.1.1 Multiple Regression Assumptions	44
5.1.2 Applications that use Multiple Regression	44
5.1.3 Multiple Regression Model	45
5.2 Utilization Modeling	46
5.2.1 Geometric Relation	46
5.2.2 Euclidean Distance	47
5.2.3 Absolute Summation (AS)	48
5.3 Multiple Factor Regression	48
5.3.1 Hybrid Local Regression Host Overload Detection Method (HLRHOD)	48
5.3.2 Multiple Regression Host Overload Detection Algorithm (MRHOD)	51
5.4 Summary	53
6 EVALUATION METHODOLOGY AND RESULTS	54
6.1 Simulation Methodology	54
6.1.1 Modeling of Power	54
6.1.2 Performance metrics	55
6.1.3 Experimental setup	56
6.2 Simulation Results and Analysis	57
6.2.1 VM Selection Policy Evaluation	57
6.2.2 Host Overload Detection using Random Workload	60
6.2.2.1 Sensitivity Analysis	60
6.2.2.2 Algorithms Comparative Analysis	63
6.2.3 Host Overload Detection using Planetlab Workload	64
6.2.3.1 Sensitivity Analysis	64
6.2.3.2 Algorithms Comparative Analysis	66
6.2.3.3 Multiple regression significance test	67
6.2.3.4 The Usage of CPU and RAM only in MRHOD algorithm	67
6.3 Summary	68
7 CONCLUSION AND FUTURE WORK	69
References	72

Appendix A HLRHOD Java Code.....	8
Appendix B MRHOD Java Code.....	85

List of Tables

2.1	Cluster, Grid and Cloud computing Comparison [1]	7
3.1	Comparative study on linear power models [2]	19
3.2	Comparative study on empirical power models [2]	21
4.1	Comparison between Host Overload Detection techniques	34
6.1	Power consumed by the chosen hosts at various load levels in Watts . . .	54
6.2	Evaluation of Maximum Correlation (MC)	59
6.3	Safety parameter for HLRHOD	60
6.4	Safety parameter for MRHOD-GR	61
6.5	Energy consumption vs. Number of VMs	61
6.6	ESV metric vs. Number of VMs	62
6.7	ESV vs Scheduling Interval	63
6.8	Simulation result of host overload detection algorithms	63
6.9	Planetlab Workload data [3]	64
6.10	Energy Consumption vs. Number of VMs	64
6.11	ESV metric vs. Number of VMs	65
6.12	Energy consumption vs. scheduling interval	65
6.13	Multiple Regression significance test using Anova	6

List of Figures

2.1	IT Evolution of Cloud Computing [4]	6
2.2	Cloud computing layered architecture [5]	8
2.3	Cloud computing model [4]	9
2.4	Cloud computing Service model [6]	11
3.1	Green Computing [7]	16
3.2	Data center power distribution [8]	17
3.3	Green Computing Power Management Strategies [7]	21
4.1	Virtual Machine Consolidation Steps	29
4.2	Classification of VM Placement Algorithms [9]	40
5.1	CPU , RAM, BW and Geometric Relation (GR) and predicted host utilization	43
6.1	Minimum Migration Time (MMT) algorithms comparison for single and multiple factors for VM selection	58
6.2	Evaluation of Random choice (RC) algorithms	58
6.3	ESV comparison for RC vs. MMT vs. MC	59
6.4	Algorithms comparative comparison using Planet Lab workload	66
6.5	Energy consumption of MRHOD with and without the usage of BW . .	68
6.6	ESV of MRHOD with and without the usage of BW	68

List of Symbols and Abbreviations

AS	Absolute Summation
DBMS	Data Base Management Systems
DVFS	Dynamic Voltage and Frequency Scaling
E	Total energy consumption
EL	Euclidean distance
ESV	Energy and SLA Violations
FQL	Fuzzy Q-Learning
GA	Genetic Algorithm
GR	Geometric Relation
HLRHOD	Hybrid Local Regression Host Overload Detection algorithm
IaaS	Infrastructure as a Service
IPMI	Intelligent Platform Management Interface
IQR	Interquartile Range
LR	Local Regression
LRR	Local Regression Robust
MAD	Median Absolute Deviation
MC	Maximum Correlation
MMT	Minimum Migration Time
MRHOD	Multiple Regression Host Overload Detection algorithm
NIST	National Institute of Standards and Technology
OLS	Ordinary Least Squares
PaaS	Platform as a Service
PDAs	Personal Digital Assistants
PDM	Performance Degradation due to Migrations
QoS	Quality of Service
RC	Random Choice

ROI	Return On Investment
SaaS	Software as a Service
SLAs	Service Level Agreements
SLATAH	SLA violation Time per Active Host
SLAV	SLA violation
SOA	Service Oriented Architecture
SPOF	single point of failure
TCA	Total Cost of Acquisition
THR	Averaging Threshold-based algorithm
VM	Virtual Machine
VoIP	Voice over IP
XaaS	Everything as a Service

Abstract

Cloud computing is an internet based computing in which a huge number of computing virtualized resources, different infrastructures, valuable softwares and distinct development platforms are provided in a pay as you go fashion as a service to customers. Fast development request to computational power due to scientific, business and web-applications mandates the production of extensive-scale data centers hence consumes a large amount of electrical energy. Virtual Machine (VM) consolidation is a compelling approach to enhance the resources utilization and increases the cloud data centers energy efficiency. VM consolidation includes live migration of VMs henceforth the capacity of exchanging a VM between physical servers with a near zero down time. VM consolidation consists of host overload/underload detection, VM selection and VM placement. Most of the recent VM consolidation approaches rely on CPU utilization only for host overload detection. However, for many applications, the performance does not rely on CPU utilization only. For applications that require communication among services, the communication cost can also influence the overall performance. Furthermore, there are applications require a huge amount of memory hence; memory utilization can also influence the overall performance.

This thesis proposes enhancement of VM consolidation using multiple factors host overload detection. Our newly developed multiple factor algorithms are based on three parameters (CPU, Memory, Bandwidth) utilizations instead of one parameter only (CPU utilization). First, we developed Hybrid Local Regression Host Overload Detection algorithm (HLRHOD) that is based on local regression using hybrid factors. It outperforms the single factor algorithms. Then we developed a multiple regression algorithm for host overload detection. The proposed algorithm, Multiple Regression Host Overload Detection (MRHOD), significantly reduces the energy consumption while guaranteeing Service Level Agreements (SLA) since it gives a real indication of host utilization. Our newly developed multiple factor algorithm is based on three different models for the VM utilization to combine the multiple factors namely Geometric Relation (GR), Euclidean distance (EL) and Absolute Summation (AS). Through simulations we show that our approach reduces power consumption by 6 times compared to single factor algorithms for random workload using MRHOD-GR. Using Planet Lab workload traces we show that MRHOD-AS improves the ESV metric by about 24% better than other single factor regression algorithms (LR and LRR). MRHOD-AS is the best algorithm for enhancement of VM consolidation then MRHOD-GR then HLRHOD.

Chapter 1: INTRODUCTION

Cloud computing causes a revolution in the IT industry by providing computing resources taking into account user requests and pay-as-you-go premise which are refined through virtualization, Service Oriented Architecture (SOA) and utility computing. To manage the issue of large energy consumption, it is essential to remove inefficiencies and waste in the manner electricity is conveyed to computing resources and in the manner these resources are utilized to fulfill the running workloads since most of the time servers work at 10-50% of their full capacity only so by enhancing the datacenters physical infrastructure in addition to resource management and allocation algorithms the energy consumption can be enhanced. The energy that is consumed due to cooling cannot be minimized since cooling is very essential for decreasing the temperature around servers. If the temperature and heat increases this will degrade the performance of CPU and also the applications that run on the cloud computing environment.

1.1 Motivation

Cloud computing is a new computing paradigm with numerous exciting features like on-demand computing resources, disposal of beginning capital and operational costs, elastic scaling and building up a pay-as-you-go plan of action for provided services [10]. On account of this, cloud computing has gained a significant consideration amid the recent decade. Furthermore, cloud suppliers have reacted by building expanded numbers of energy hungry data centers so as to fulfill the growing customer's (e.g. capacity, processing power) needs [11]. Such datacenters do not just force autonomy and scalability challenges on the frameworks that they control, but additionally they bring up issues with regards to energy efficiency [12]. They expend gigantic measures of electrical energy bringing about large amount of CO₂ emissions and high operational cost. The energy which a datacenter expends is divided into two portions [13]:

1. A static (or fixed) portion that relies on the size of the system and the type of the component (computing, network devices and storage elements); the amount consumed is acquired by leakage currents found in any powered system.
2. A dynamic (or variable) portion that outcomes from computing, storage and network resources; brought on by the activity of the system and variations in clock rates.

The decrease in power consumption does not generally diminish the energy consumed [14] since power is the rate that a system will consume electric energy while energy is the overall amount of electric energy utilized in a particular time frame. The minimization in power consumed causes a decrease in the expenses of the infrastructure provisioning. Enhancing the efficiency of energy by minimizing the fixed portion and conveying more performance proportional to the variable portion has become an extremely attractive development and research area. It is also a superior challenge in cloud computing.

The purpose behind huge energy consumption is not only the power inefficiency of hardware and the quantity of computing nodes, but also the inefficient usage of these hosts.