Dedication

To my mother, late Ms. Kamlesh Rani,
and my father Sh. SatPal Goyal
Contents

List of Figures ........................................................................................................ vi
List of Tables ........................................................................................................ viii
Certificate ............................................................................................................. ix
Acknowledgement ................................................................................................ x
Abstract .............................................................................................................. xii
1 Introduction ......................................................................................................... 1
  1.1 Cloud Computing: An Overview ................................................................. 2
    1.1.1 Evolution of Cloud Computing ......................................................... 3
    1.1.2 Types of Cloud .................................................................................. 4
    1.1.3 Benefits of Cloud Computing .......................................................... 7
  1.2 Need of Energy efficient Computing in Clouds ........................................... 8
  1.3 Energy Efficient Cloud Computing Techniques ......................................... 10
    1.3.1 VM Management and Scheduling .................................................... 11
      1.3.1.1 Thermal Aware Scheduling ...................................................... 11
      1.3.1.2 Energy efficient Resource Scheduling .................................... 12
      1.3.1.3 Energy Aware Data Storage and Placement Strategies ........ 12
    1.3.2 Energy Efficient Network ................................................................. 13
      1.3.2.1 Energy Efficient Network Devices .......................................... 13
      1.3.2.2 Traffic Engineering and Energy Efficient Routing ............... 13
      1.3.2.3 Energy Efficient Network Architecture ................................ 14
    1.3.3 Advance Data Center Design ............................................................. 14
      1.3.3.1 Energy Efficient Hardware ....................................................... 14
      1.3.3.2 Renewable Energy Sources .................................................... 15
      1.3.3.3 Data Center Locations ............................................................. 16
    1.3.4 Energy Aware User Contracts ........................................................... 16
      1.3.4.1 Energy Aware SLA ................................................................. 16
  1.4 Thesis Organization ....................................................................................... 16
  1.5 Thesis Contribution ...................................................................................... 18
2 Literature Review ................................................................. 20
  2.1 Energy Efficient Resource Scheduling in Cloud Computing ------- 21
    2.1.1 Consolidation based Heuristic Approaches --------------------- 21
    2.1.2 Greedy Heuristic Approaches ------------------------------- 23
    2.1.3 Bio-Inspired Heuristic Approaches -------------------------- 24
    2.1.4 Power-Aware Heuristic Approaches -------------------------- 25
    2.1.5 Miscellaneous Heuristic Approaches ------------------------ 26
  2.2 Comparative Analysis of Energy Efficient Resource Scheduling
    Algorithms-------------------------------------------------------- 30
  2.3 SLA Based Energy Aware Scheduling in Cloud Computing ---------- 32
    2.3.1 Power-Aware Heuristic Approaches -------------------------- 32
    2.3.2 Bio-Inspired Heuristic Approaches -------------------------- 33
    2.3.3 Miscellaneous Heuristic Approaches ------------------------- 33
  2.4 Comparative Analysis of Green SLA Scheduling Algorithms ------ 35
  2.5 Motivation ------------------------------------------------------ 38
  2.6 Problem Formulation and Objectives ----------------------------- 39
  2.7 Summary --------------------------------------------------------- 40
3 Proposed Energy-efficient Cloud Resource Scheduling (ECRS) Algorithm 41
  3.1 Formalization of the Proposed Energy-efficient Resources Scheduling
    Algorithm-------------------------------------------------------- 42
  3.2 Proposed Energy Efficient Cloud Service Framework: ACA-Cloud----- 43
    3.2.1 Energy Model ---------------------------------------------- 43
    3.2.2 Reservation Model ----------------------------------------- 44
    3.2.3 Prediction Model ------------------------------------------ 46
  3.3 Proposed Energy-efficient Cloud Resource Scheduling (ECRS) Algorithm 47
    3.3.1 Initial VM Placement --------------------------------------- 47
    3.3.2 Optimization the Current Workload ------------------------- 48
  3.4 Summary --------------------------------------------------------- 50
4 Proposed Green SLA aware Cloud Resource Reservation (GSLACRR)
  Algorithm------------------------------------------------------------- 51
  4.1 Introduction ------------------------------------------------------ 52
4.2 Proposed Green Service Level Agreement (GSLA) based Resource Management

4.2.1 Energy Based Resource Provisioning Policy (EBRPP): A Negotiation Approach

4.2.2 GSLA Workflow

4.2 Proposed GSLA aware Cloud Resource Reservation (GSLACRR) Algorithm

4.3 Advantages of the Proposed (GSLACRR) Algorithm

4.4 Summary

5 Design and Implementation of Proposed Algorithms

5.1 Architecture of ACA-Cloud

5.1.1 Interface Services

5.1.2 Account Manager

5.1.3 Cloud Cluster Controller

5.1.4 Host Controller

5.2 Design Details of ACA-Cloud

5.2.1 Use Case Diagrams

5.2.2 Class Diagram

5.2.3 Activity Diagram

5.2.3.1 ECRS Algorithm’s VM Handler for Initial VM Placement

5.2.3.2 ECRS Algorithm’s Load Balancer for Optimizing Current Workload

5.2.4 Sequence Diagram

5.2.4.1 GSLACRR Successful User Negotiation

5.2.4.2 GSLACRR Renegotiation for User’s Requirement

5.2.5 State Diagram

5.3 Implementation Details

5.3.1 Experiment Setup

5.3.2 User Interface

5.4 Summary
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Simple view of Cloud Computing</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Convergence of various technologies advancement leading to the advent of cloud computing</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Cloud computing comprised a collection of technologies, solutions and models</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>The cloud computing stack</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>Energy distribution in a data centre</td>
<td>8</td>
</tr>
<tr>
<td>1.6</td>
<td>Average CPU utilization of more than 5,000 servers for a six-month period</td>
<td>9</td>
</tr>
<tr>
<td>1.7</td>
<td>A taxonomy of Green Cloud Framework</td>
<td>11</td>
</tr>
<tr>
<td>2.1</td>
<td>Energy efficient resource scheduling taxonomy</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>High level architecture of proposed energy efficient cloud service framework</td>
<td>45</td>
</tr>
<tr>
<td>3.2</td>
<td>VM migration band for a host</td>
<td>49</td>
</tr>
<tr>
<td>4.1</td>
<td>Proposed GSLA-based resource management</td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>Proposed Green Service Level Agreement</td>
<td>54</td>
</tr>
<tr>
<td>5.1</td>
<td>Architecture of proposed ACA-Cloud framework</td>
<td>62</td>
</tr>
<tr>
<td>5.2</td>
<td>Communication protocol VMH-GE and LB-GE</td>
<td>64</td>
</tr>
<tr>
<td>5.3</td>
<td>Use case diagram for user authentication and management</td>
<td>66</td>
</tr>
<tr>
<td>5.4</td>
<td>Use Case Diagram for resource provision</td>
<td>66</td>
</tr>
<tr>
<td>5.5</td>
<td>Class diagram for ACA-Cloud</td>
<td>67</td>
</tr>
<tr>
<td>5.6</td>
<td>GlobalExecutor class of ACA-Cloud</td>
<td>67</td>
</tr>
<tr>
<td>5.7</td>
<td>Host class of ACA-Cloud</td>
<td>68</td>
</tr>
<tr>
<td>5.8</td>
<td>Activity diagram of VMHandler module of ECRS algorithm</td>
<td>69</td>
</tr>
<tr>
<td>5.9</td>
<td>Activity diagram of LoadBalancer module of ECRS algorithm</td>
<td>70</td>
</tr>
<tr>
<td>5.10</td>
<td>Sequence Diagram of successful negotiation process of GSLACRR</td>
<td>71</td>
</tr>
<tr>
<td>5.11</td>
<td>Sequence diagram of GSLACRR renegotiation for user's requirement</td>
<td>72</td>
</tr>
<tr>
<td>5.12</td>
<td>State diagram of user resource requests</td>
<td>73</td>
</tr>
<tr>
<td>5.13</td>
<td>Authenticated user resource request screen</td>
<td>75</td>
</tr>
<tr>
<td>5.14</td>
<td>GSLA negotiation screen</td>
<td>76</td>
</tr>
<tr>
<td>5.15</td>
<td>ACA-Cloud administration panel screen</td>
<td>77</td>
</tr>
</tbody>
</table>
5.16 Development environment and output generated by GlobalExecutor for ACA-Cloud ................................................................. 77
5.17 Snapshot of HostSupervisorGlobal for ACA-Cloud ---------------- 78
6.1 Accepted VM requests with low resource heterogeneity ............. 84
6.2 Average power consumption of a centre with low resource heterogeneity 85
6.3 Accepted VM requests with high resource heterogeneity ............ 86
6.4 Average power consumption of a centre with high resource heterogeneity .............................................................. 86
6.5 Accepted VM requests ............................................................... 89
6.6 Rejected VM requests ............................................................... 90
6.7 Average power consumption of a centre------------------------------ 90
6.8 Number of PMs used with respect to VMs ................................. 91
6.9 (a) The experimental results with different types of workload (a) Test 1 with low number of VMs request ........................................ 91
6.9 (b) The experimental results with different types of workload (b) Test 2 with high number of VMs request.......................... 92
6.10 Effect of number of resources/servers on the power consumption in a data center ....................................................... 94
List of Tables

2.1 Energy efficient scheduling algorithms for cloud computing-------- 27
2.2 Comparative analysis of energy efficient scheduling algorithms and policies ----------------------------------------------- 31
2.3 Comparative analysis of green SLA Scheduling algorithms-------- 36
3.1 Acronyms used in the algorithms-------------------------------- 47
4.1 Acronyms used in the GSLACRR algorithm ------------------------ 57
5.1 Implementation Technology used by the ACA-Cloud --------------- 74
5.2 Experimental setup --------------------------------------------- 75
6.1 VMs deployed --------------------------------------------------- 81
6.2 T-test: paired two sample for means------------------------------ 87
6.3 T-test: paired two sample for means with respect to MEC metric-- 93
6.4 T-test: paired two sample for means with respect to AVM metric--- 93
CERTIFICATE

I hereby certify that the work which is being presented in this thesis entitled "Energy Efficient Resource Scheduling Algorithms for Cloud Computing", for the award of degree of "Doctor of Philosophy" submitted to the Department of Computer Science and Engineering, Thapar University, Patiala, is an authentic record of my own work, carried out under the supervision of Dr. Seema Bawa and Dr. Bhupinder Singh. It refers to the work done by other researchers which are duly listed in the reference section.

The matter presented in this thesis has not been submitted in part or full to any other institute or university, for the award of any other degree.

(Sudhir Goyal)
Reg. No. 951003012

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. Seema Bawa)
Professor,
Computer Science & Engineering Department
Thapar University, Patiala 147004, Punjab, INDIA

(Dr. Bhupinder Singh)
Senior Program Manager,
Microsoft India (R&D) Ltd.,
Hyderabad, INDIA
Acknowledgement

First of all, I express my gratitude to that Almighty, Who blessed me with the zeal and enthusiasm to complete this work successfully.

My endeavors in this thesis have been made possible by several remarkable people who helped and supported me, and most importantly prayed for me. I would like to express my sincere appreciation and gratitude to them all in my humble acknowledgment.

Firstly, I would like to express my sincere thanks to my supervisors Prof. Seema Bawa, and Dr. Bhupinder Singh for their supervision, timely advice, and comments from the very early stage of this research which kept me stayed on the right track. It has been a great pleasure and a privilege to have the opportunity to work under their supervisions.

I also wish to extend my gratitude to the members of the PhD committee: Prof. Rajesh Khana, Dr. Inderveer Chana, and Dr. V.P Singh for their encouragement and insightful comments in relation to my research.I would like to many thanks all faculty and staff members of Thapar University, who have been kind enough to advise and help in their respective roles.

This research has profited from the friendship, advice, encouragement and support of a fine set of people. I am indebted to my dear friend Naveen Gupta, Panjab University, for continual counsel, willingness to support and helped me, with a smiling face, as and when required. I would also like to special thanks to Khushneet Jindal, Sunil Singla, Manmohan Chhibber, Ajay Kakkar, Ashish Aggarwal, Anuj Gupta, Arun Bansal for their friendship and help during my PhD. I have had a great time with them.

Last but not least, I am heartily thankful to all my family members for their love, encouragement and endless support at all times. I can never repay the greatness of my father’s love for me. He has sacrificed everything for his sons, my younger brother and me. This thesis is a result of his endless love. I remain indebted to my divine mother, late Ms. Kamlesh Rani,who always believed in my capabilities. Thanks to my loving brother, Sandeep Goyal, and my sisters-in-law Meenu Goyal, their daughters for being supportive and caring. Finally, I thank to my wife Ritu and daughter Avani, for their support, love, inspiration,
patience, and for making my life filled with joy and happiness. They never give up on me, but love me from the bottom of their heart.

Sudhir Goyal
Abstract

Cloud Computing presents an exciting new horizon for the Information Technology (IT) industry. It provides a cost-effective solution, as it allows hosting of storage, computational and supported network services on a shared infrastructure of physical servers. It offers utility oriented IT services to the users worldwide. Cloud computing empowers companies to host engineering, scientific, and business applications, and makes them accessible across the world. However, to accommodate the increasing trend of online computing applications and the ever growing massive amount of data, the data centers are also continually expanding in size. This means a huge consumption of electrical energy that ultimately results in high operational costs and emission of green house gases into the environment. Therefore, to curb this unsustainable increase in energy consumption, research on "energy efficient computing" becomes a critical need and a roadblock of great magnitude with respect to energy efficient utilization of resources while preserving the desired users’ Quality of Service (QoS) standards.

This thesis focuses on cloud resource management with special attention to energy efficient resource scheduling and Green Service Level Agreements (GSLAs). Initially, it compares, analyses and reports on the existing energy aware resource scheduling frameworks and heuristics on the basis of various aspects. From the literature survey, it is apparent that there is a great energy saving potential with respect to the system operations and workload specificities. This in particular holds for small and medium sized datacenters which can’t afford expensive hardware and renewable energy sources to save energy. To address this challenge, this thesis presents novel techniques, models, and algorithms for the cloud environment.

To achieve the goal of improving resource utilization and reducing energy consumption, the Energy efficient Cloud Resource Scheduling (ECRS) algorithm has been designed. It gathers information such as host utilization level, power consumption, number of VMs and their state etc. regarding available resources. Using this information along with an estimate of resources required for future requests, the energy efficiency of cloud compute cluster is assessed. This assessment is used to manage resources energy efficiently. Further, to involve the users in eco-system cloud services, a Green Service Level Agreement (GSLA) aware
Cloud Resource Reservation (GSLACRR) algorithm has been proposed. It is an endeavor to incline the users towards sustainable computing through user negotiation strategies. It ultimately results in cost benefits for users as well as service providers and helps to minimize the energy consumption. To address the various cloud resource management challenges such as performance, energy efficiency etc., an energy efficient cloud framework, named ACA-Cloud has been proposed, designed, developed and tested, and further used to demonstrate the applicability of the proposed algorithms.

The proposed energy efficient cloud framework, ACA-Cloud, has four layers architecture, with Host Controller (HC) as the base layer, providing actual resources to a user. The second layer is Cloud Cluster Controller (CCC) which is responsible for monitoring the status of all the physical machines/hosts and making appropriate resource management decisions in response to the current workload and incoming user requests. The next layer is Account Manager, which renders user authentication management to the framework to handle different account types and authentication. The topmost layer is the Web Portal Interface, through which users can submit their requests for cloud services.

This proposed framework has been installed at Thapar University, Patiala. The experimental results reveal the competitive performance and usage of the proposed algorithms implemented on this framework.

Finally, the proposed algorithms are compared with the existing one to validate the outcome. The results show that the scheduling algorithms successfully and collectively address the issues of energy efficiency and performance to establish an efficient Cloud.