# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ABSTRACT</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF SYMBOLS AND ABBREVIATIONS</td>
<td>xii</td>
</tr>
</tbody>
</table>

1. INTRODUCTION 1

1.1 BACKGROUND 1

1.1.1 Cloud Service Models 2

1.1.2 Types of Cloud 2

1.2 VIRTUAL MACHINE 3

1.3 LOAD BALANCING AMONG VM 4

1.3.1 Challenges in Load Balancing 4

1.3.2 Goals of Load Balancing 6

1.3.3 Classification of Load Balancing Algorithms 6

1.3.4 Load Balancing Algorithms 6

1.4 DYNAMIC RESOURCE ALLOCATION 9

1.4.1 Dynamic Resource Allocation Strategies 10

1.5 SCHEDULING IN VM 10

1.5.1 Content-based Scheduling 11

1.5.2 Dynamic Adaptive Scheduling 11

1.5.3 Task Scheduling 12

1.5.3.1 First Come First Served (FCFS) 12

1.5.3.2 Round Robin Scheduling (RR) 12

1.5.3.3 Generalized Priority Algorithm (GPA) 13

1.5.4 Firefly Algorithm (FA) 14

1.5.5 Min-Min Scheduling Algorithm 14

1.5.6 Max-Min Scheduling Algorithm 15

1.5.7 Shortest Job First (SJF) Algorithm 15

1.6 MOTIVATION 15

1.7 OBJECTIVES 16
1.8 ORGANIZATION OF THESIS 16

2. LITERATURE SURVEY 18
   2.1 INTRODUCTION 18
   2.2 OPTIMAL VIRTUAL MACHINE SELECTION 18
   2.3 LOAD BALANCING 21
   2.4 RESOURCE SHARING 26
   2.5 RESOURCE ALLOCATION 27
   2.6 PARALLEL JOB SCHEDULING 32
   2.7 SUMMARY 36

3. REVIEW OF RELATED RESEARCH WORKS 37
   3.1 RELATED WORKS 37
   3.2 PROBLEM STATEMENT 45
   3.3 OBJECTIVES 46
   3.4 SPECIFICATION, REQUIREMENTS AND TOOLS 46
   3.5 PERFORMANCE METRICS 47
      3.5.1 Response Time 47
      3.5.2 Processing Delay 47
      3.5.3 Processing Speed 47
      3.5.4 Migration 47
      3.5.5 Makespan 47
      3.5.6 Waiting Time 47
      3.5.7 Idle Time 48
      3.5.8 Throughput 48
      3.5.9 Latency 48
      3.5.10 Cost and Accuracy 48
      3.5.11 Execution Time 49
      3.5.12 Access Storage Capacity 49
      3.5.13 Power Consumption 49
      3.5.14 Data Transfer Rate 49
      3.5.15 Response Time 49
      3.5.16 Allotted Jobs 49
      3.5.17 Data Storage 50
3.5.18 Efficiency Rate 50
3.5.19 Time Consumption 50
3.5.20 Resource Utilization 50

3.6 SUMMARY 51

4. AN EFFICIENT FRAMEWORK TO HANDLE INTEGRATED VM WORKLOADS IN CLOUD 52

4.1 INTRODUCTION 52
4.1.1 Load Balancing Techniques for Resolving Load Imbalance Issues 53
4.1.2 Low Resource Utilization 53

4.2 DISTRIBUTED VIRTUAL ENVIRONMENTS (DVE) 54
4.2.1 Data Model 54
4.2.2 Communication Model 54
4.2.3 View Consistency 55
4.2.4 Network Traffic Reduction 55

4.3 OPTIMAL VIRTUAL MACHINE CONFIGURATION (OVMC) STRATEGY 57
4.3.1 Generation of VM nodes 59
4.3.2 Cost Factors of Virtual Machines 60
4.3.2.1 Strategic Decision 61
4.3.2.2 Selection of service provider 61
4.3.2.3 Cloud Types Service Charge IaaS 61
4.3.3 Fuzzy Decision-Making Approach 62
4.3.4 Integer Linear Programming Constraints 64

4.4 RESULTS AND DISCUSSIONS 68
4.4.1 Time Consumption 69
4.4.2 Efficiency Rate 70
4.4.3 Throughput 71
4.4.4 Latency 72
4.4.5 Cost and Accuracy 73
4.4.6 Execution Time 74
4.4.7 Access Storage Capacity 76
5. AN INTEGRATED STP-DRS ALGORITHM FOR ON-DEMAND RESOURCE SHARING IN CLOUD

5.1 INTRODUCTION

5.2 VM ASSESSMENT

5.3 LOAD BALANCING
- 5.3.1 Mathematical Model
- 5.3.2 Load Balancing Decision
- 5.3.3 Virtual Machine (VM) Grouping
- 5.3.4 Job Transfer

5.4 DYNAMIC RESOURCE ALLOCATION
- 5.4.1 On-demand Sharing
- 5.4.2 Queue Size Estimation
- 5.4.3 Load Computation
- 5.4.4 Job Demand Estimation
- 5.4.5 Function for Load Check
- 5.4.6 Function for Termination Criteria

5.5 RESULTS AND DISCUSSION
- 5.5.1 Throughput
- 5.5.2 Response Time
- 5.5.3 Allotted Jobs
- 5.5.4 Data Storage
- 5.5.5 Resource Utilization
- 5.5.6 Processing Delay
- 5.5.7 Processing Speed
- 5.5.8 Migration
- 5.5.9 Makespan

5.6 SUMMARY

6. A NOVEL APPROACH FOR SCHEDULING PARALLEL JOBS BASED ON PRIORITY BASED CONSOLIDATION METHOD
6.1 INTRODUCTION 108
6.1.1 Full Virtualization 109
6.1.2 Para Virtualization 109
6.2 LOAD BALANCING 110
6.2.1 Static Load Balancing Approaches 111
6.2.2 Dynamic Load Balancing Approaches 111
6.3 JOB SCHEDULING 112
6.4 TYPES OF SCHEDULING 115
6.4.1 Static and Dynamic Scheduling 115
6.4.2 Centralized, Hierarchical and Distributed Scheduling 116
6.4.3 Preemptive and Non-preemptive scheduling 116
6.4.4 Online and Batch mode scheduling 116
6.5 SCHEDULING ALGORITHMS 117
6.6 PARALLEL JOB SCHEDULING 118
6.7 PRIORITY-BASED JOB SCHEDULING 118
6.8 BACKFILLING ALGORITHM 119
6.9 BEE BACKFILL WITH MIGRATION 120
6.9.1 System Architecture 120
6.10 RESULTS AND DISCUSSION 123
6.10.1 Response Time 123
6.10.2 Migration 124
6.10.3 Makespan 125
6.10.4 Waiting Time 127
6.10.5 Idle Time 128
6.11 SUMMARY 129
7. RESEARCH FINDINGS 131
8. CONCLUSION AND FUTURE ENHANCEMENT 133
8.1 CONCLUSION 133
8.2 FUTURE ENHANCEMENTS 134
REFERENCES 135
LIST OF PUBLICATIONS 157