

LB/TH/41/2025

TH6003

**COST-EFFECTIVE AND SECURE MULTI-CLOUD  
OPTIMIZATION STRATEGY FOR BLOCK  
STORAGE**

P.N. Sithara

229391N

Master's in Computer Science

Department of Computer Science and Engineering  
Faculty of Engineering

University of Moratuwa  
Sri Lanka

07 2025

**COST-EFFECTIVE AND SECURE MULTI-CLOUD  
OPTIMIZATION STRATEGY FOR BLOCK  
STORAGE**

P.N. Sithara

229391N

Thesis submitted in partial fulfillment of the requirements for the degree  
Master's in Computer Science

Department of Computer Science and Engineering  
Faculty of Engineering

University of Moratuwa  
Sri Lanka

07 2025

## **DECLARATION**

I declare that this is my own work and this Thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date: 2025-06-30

The supervisor should certify the Thesis with the following declaration.

The above candidate has carried out research for the Master's in Computer Science Thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. Indika Perera

Signature of the Supervisor:

Date:

## **ACKNOWLEDGEMENT**

I would like to express my deepest gratitude to everyone who supported me throughout the journey of completing this research thesis.

First and foremost, I extend my heartfelt thanks to my supervisor, Prof. Indika Perera, for his invaluable guidance, insightful feedback, and constant encouragement, which were instrumental in developing this work. His expertise and patience inspired me to push my boundaries and explore new dimensions in this research.

I am also grateful to my academic institution, Department of Computer Science, Faculty of Engineering, University of Moratuwa, for providing the resources and a conducive environment that supported me to carry out this study successfully.

Special thanks go to my colleagues and peers who offered their support and shared their perspectives, which enriched my understanding of the subject matter. I am equally thankful to my wife and family for their unwavering belief in my abilities, their emotional support, and their patience throughout this journey.

Thank you all for being a part of this journey.

## ABSTRACT

In the era of big data, the volume of information being generated and stored has grown rapidly from terabytes to petabytes. Cloud storage provides a practical solution by offering nearly unlimited storage capacity with flexibility and scalability. However, relying on a single cloud provider poses significant risks, such as vendor lock-in, data breaches, and service outages. Even having replication strategies puts users into series of complex technical and financial situations. Multi-cloud storage, which distributes data across different cloud providers is an emerging research area which helps reduce these risks by improving fault tolerance and ensuring higher data while maintaining security for different user requirements.

A major challenge in multi-cloud storage is determining how to store user's block data in cost-effective, high availability, high performance within the budget. This project addresses these challenges by developing COSMOS-B, an optimized approach to select the best combination of cloud providers for storing block data objects while balancing cost and availability, latency and other security requirements. It uses erasure coding for fault-tolerant data distribution and weighted sum strategies to evaluate cloud providers based on normalized cost and availability metrics while enhancing overall security. Finally, it demonstrates the model's adaptability to varying priorities, achieving significant cost savings and high reliability multi cloud environment.

**Keywords:** Cloud Computing, Multi-cloud, Cloud Platforms, Data Storage, Block Data, Erasure Coding, Cost Effectiveness, Latency, Availability

## TABLE OF CONTENTS

Declaration of the Candidate & Supervisor	i
Acknowledgement	ii
Abstract	iii
Table of Contents	iv
List of Figures	viii
List of Tables	ix
1 Introduction	1
1.1 Background and Motivation	1
1.1.1 Current Challenges in Cloud Storage	2
1.1.2 The Rise of Multi-Cloud Storage	3
1.2 Problem Statement	4
1.2.1 Core Challenges	4
1.2.2 Key Research Questions	6
1.2.3 Research Focus	7
1.3 Research Objective	8
1.3.1 Cost Optimization	8
1.3.2 Availability Enhancement	9
1.3.3 Security Implementation	9
1.4 Research Contribution	9
1.4.1 Mathematical Framework	10
1.5 Thesis Organization	10
2 Literature Review	12
2.1 Cloud Storage Technologies	12
2.2 Multi-Cloud Storage Systems	12
2.3 Erasure coding and multi-cloud environment	14
2.3.1 Reed-Solomon Coding	14
2.3.2 XOR-Based Coding	14

2.3.3	Locally Repairable Codes (LRC)	15
2.4	Optimization Techniques for Cost and Availability	15
2.4.1	Availability Optimization	16
2.4.2	Performance Optimization	17
2.5	Security in Multi-Cloud Storage	18
2.6	Data Consistency Models	18
2.6.1	Traditional Consistency Models	19
2.6.2	Multi-Cloud Specific Models	19
2.6.3	Implementation Approaches	20
2.6.4	Performance Impact	21
2.6.5	Emerging Technologies	21
2.7	Related Work Analysis	22
2.7.1	Storage Technology Evolution	22
2.7.2	Data Protection Approaches	23
2.7.3	Optimization Strategies	24
2.7.4	Research Opportunities	25
2.7.5	Research Direction Alignment	25
3	System Architecture and Design	27
3.1	Research Objective	27
3.1.1	User Scenario	27
3.2	System Overview	29
3.2.1	Proposed Framework components	30
3.2.2	Supporting Systems	32
3.3	Data Flow Architecture	32
3.3.1	Upload Data Flow	33
3.3.2	Download Data Flow	34
4	Security Framework	37
4.1	Threat Model	37
4.2	Security Requirements	38
4.3	Security Tier Model	39
4.3.1	Tier 1: Single-Server (Low Security)	40

4.3.2	Tier 2: Multi-Cloud Solution (Medium Security)	40
4.3.3	Tier 3: Multi-Cloud Solution (High Security)	41
4.3.4	Integration with the Multi-Cloud Model	42
5	Data Management Framework	43
5.1	Data Placement Strategies	43
5.1.1	Security Tiers and Their Impact on Data Placement	43
5.1.2	Erasure Coding in Data Placement	44
5.2	Consistency Management	46
5.2.1	Recovery Mechanisms in Consistency Management	47
5.3	Data Retrieval Strategy	49
5.3.1	Single Server Data Retrieval	49
5.3.2	Multi-Cloud Retrieval (Erasure Coding)	50
6	Methodology	52
6.1	Security Tier 1: Finding the Best Single Storage Node	53
6.2	Security Tier 2: Finding Optimum Storage Nodes in Single Cloud Provider	54
6.2.1	Problem Description	54
6.2.2	Optimization Approach	55
6.3	Security Tier 3: Finding the Optimum Cloud Storage Nodes in any Cloud Providers	56
6.3.1	Optimization Approach	56
6.4	Optimization for security levels	57
6.4.1	Pareto Optimal Solutions	58
6.4.2	Entropy-Based Weightage Mechanism	58
6.5	Metadata Table for Data Tracking	60
6.5.1	Metadata table work flow	60
7	Implementation	63
7.0.1	Entropy Weights Calculation	65
7.0.2	Pareto Optimal Solutions	66
7.0.3	Availability Calculation	67

8	Evaluation and Results	70
8.1	Experimental Setup	71
8.2	Impact on Number of Storage Nodes	71
8.3	Cost Analysis over Amount of Data Transferred	72
8.3.1	Latency analysis	73
8.3.2	Cost and Number of Requests Analysis	74
8.4	Data Retrieval Analysis	75
8.4.1	Retrieval time Performance	75
8.4.2	Recovery Success Rate	76
8.4.3	Geographic Latency Analysis	76
8.5	Cost and Availability Comparison over Popular Cloud Providers	77
8.6	Security Evaluation	79
9	Conclusion	80
9.0.1	Future Improvements	82
	References	83

## LIST OF FIGURES

<b>Figure</b>	<b>Description</b>	<b>Page</b>
Figure 1.1	Erasure coding data management	8
Figure 3.1	Proposed Multi-cloud Framework	30
Figure 3.2	Upload data flow	33
Figure 3.3	Data download flow	35
Figure 5.1	Consistency Workflow	48
Figure 8.1	Cost comparison over the number of data nodes	71
Figure 8.2	Cost and Amount of Data transferred	73
Figure 8.3	Latency Analysis	74
Figure 8.4	Cost and No of Read requests	75
Figure 8.5	Data Retrieval Time	76
Figure 8.6	Recovery Success Rate	76
Figure 8.7	Recovery Success Rate	77

## LIST OF TABLES

<b>Table</b>	<b>Description</b>	<b>Page</b>
Table 3.1	Cloud Service Pricing Comparison	29
Table 5.1	Data Placement example	45
Table 5.2	Data Placement example	48
Table 5.3	Comparison of Single Server and Multi-Cloud Retrieval	51
Table 8.1	cloud provider data used in evaluation	70
Table 8.2	Parameter Setup	71
Table 8.3	Tier1 optimization output	72
Table 8.4	Availability comparison	72
Table 8.5	Parameter Setup	78
Table 8.6	cloud provider pricing and availability Comparison	78