HIGH RELIABILITY PRIVATE CLOUD STORAGE WITH MINIMAL OPERATION COST USING MULTI-OBJECTIVE OPTIMIZATION

MUHD HANIS BIN HARUN

A thesis submitted in
fulfillment of the requirement for the award of the
Master of Computer Science (Software Engineering)

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

APRIL 2014

ABSTRACT

As far as internet connection brings people together to share idea and communicate through it, it continually invites service providers to share their resources with desired organizations in order to improve their business operations. Storage system used to store most of the organization's data needs to be accessed frequently. Thus, storage has become a vital part of an organization and needs to be available almost anytime. In order to maximize reliability and the availability of their data, many of organization converts their storage into a cloud storage platform. To set up a cloud server, a lot of consideration is needed. One of the main considerations is cost. Setting up a cloud storage without proper strategies will involve very high cost in future. In order to make sure the cloud storage server operates at maximum reliability in a certain amount of time, a high budget for servers is required. High server cost is required in order to obtain maximum reliability. Thus, setting up target duration will help organizations to control the cost without having to compromise the storage reliability. This research will focus on finding maximum server reliability based on target duration and cost constraint. In this thesis, maximum reliability is defined by using two main objectives which are budget and duration. The multi objective optimization technique is used in order to find the optimum point from the constraint. The optimal result will be the minimum number of servers required to achieve maximum reliability according to target duration.

ABSTRAK

Saban hari talian internet membangun secara meluas membawa orang ramai untuk berkongsi idea dan berkomunikasi melaluinya, seterusnya menjemput pembekal perkhidmatan untuk berkongsi sumber mereka dengan organisasi ini untuk meningkatkan operasi perniagaan mereka. Sistem penyimpanan yang digunakan untuk menyimpan hampir keseluruhan maklumat organisasi dan perlu kepada akses yang lebih kerap. Oleh itu, pusat penyimpanan data telah menjadi bahagian yang penting dalam sesebuah organisasi dan perlu beroperasi hampir setiap masa. Bagi memaksimumkan kebolehpercayaan dan ketersediaan data mereka, kebanyakan organisasi menukar pusat penyimpanan data mereka ke platform Cloud. Dalam usaha untuk menubuhkan Cloud Server, banyak pertimbangan yang diperlukan. Salah satu pertimbangan utama adalah kos. Membina Cloud Server sendiri tanpa strategi yang betul akan melibatkan kos yang sangat tinggi pada masa hadapan. Bajet yang tinggi juga diperlukan bagi memastikan Cloud Server beroperasi pada kebolehpercayaan maksimum dalam jumlah masa tertentu. Kos server yang tinggi diperlukan untuk mendapatkan kebolehpercayaan maksimum. Oleh itu , menetapkan jangka masa sasaran akan membantu organisasi untuk mengawal kos tanpa perlu berkompromi dengan keupayaan Cloud Server. Dalam tesis ini, kebolehpercayaan maksimum ditentukan dengan menggunakan dua objektif utama iaitu bajet dan tempoh masa. Multi-objective optimization digunakan untuk mencari titik optimum dari kekangan yang ditetapkan. Hasil optimum akan bilangan minimum server yang diperlukan untuk mencapai kebolehpercayaan maksimum dalam jangka masa yang ditetapkan.

CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background

Cloud Computing has turned to become one of the most popular and latest technology now days. The need of increasing processing power and efficient storage capacity influenced businesses to take part in cloud computing technology. As a general concept, we can imagine all business data are being stored in some place different from the organization. The data is replicated in a few forms as an assurance of 99.9% maximum uptime if anything happens to the server. This is in opposition to the distributed computing paradigm, where replication of data is the main issue to be solved. The term cloud computing is becoming quite synonyms with computing technology now days. The idea of cloud computing was coined by Leonard Kleinrock back then in 1969 where computer usage was likened to utilities. Even though, there are certain people working hard toward the idea, some computer users still do not have a clear view of what it is.

What is cloud computing? Cloud computing refers to both the application delivered as a service over the internet and the hardware and systems software in the datacenters that provide those services (Fox *et al.*, 2009). By using the cloud, an organization does not have to deal with the technical issues on the infrastructure and does not have to invest the budget on buying and maintaining their vault of data. They will have to outsource their data to the cloud provider's server and need to pay for the size of data stored. They can also enlarge the storage size every time they need to. That is why the term "pay-per-use" was very synonymous with the cloud service. It is as if we have to pay for what we have used only. In general, there are three common service models offered in a cloud environment, namely: Application as a Service (AaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) (Sourabh, 2012). As part of cloud computing, IaaS plays important features of its thriving compared to the other two services. This is because infrastructure serves as a third party provider for organization's hardware and operations. All of the company's confidential data will be outsourced to the provider and will reside outside their premises. The capability of a service provider to manage and handle huge amount of data is very crucial. At the same time, the providers have to manage their storage server and maintain the data to make sure it is available anytime when being requested by the organization. To maintain the storage reliability and the data availability, cloud offers the best framework with massive scalability, 99.99% reliability, high performance, and specific configuration by using multiple redundancy solution (Wu, 2010).

1.2 Problem Statement

In terms of data availability, cloud offers a very good data reliability and performance as it offers 99.99% service up-time. In order to maintain the service up-time, it requires huge amount of storage capacity. Current cloud systems apply three replication strategies in order to achieve data availability. Such a 3-replica strategy means to store one gigabyte of data would need three gigabytes of data space and incur two times additional cost, which significantly affects the cost-effectiveness of the cloud system. In addition, such huge resource allocation will indirectly increase the energy consumption to power up the storage server for 24/7 a week that will indirectly affect the operational cost.

Organizations today tend to keep all of their business activity and data on an online server, which enables them and their customer to access the resource anytime, anywhere. Organization's size often affects the storage amount for their data. Most cloud storage providers nowadays provide various storage packages for consideration, but the main consideration is: in order for a certain amount of data to be stored for a particular period of time, how much storage would be required? Was the amount of storage able to maintain the reliability of storage in a certain amount of time? If it does, how long will it take before it changes to a failed state?

Having additional servers to maintain storage reliability will increase the electricity consumption to power up and cooling down the server. According to Barett (2011) in his article on eHow.com, a server consumes electricity from 500 to 1200watt per hour. A server powered up by using natural gas, produced approximate 0.553kg carbon dioxide per hour (Carbon Fund, 2012). If an organization needs five servers to operate twenty-four hours a year to serve their customers for 10 years, the operational server itself will produce at least 242,214kg carbon dioxide. Even worst, twenty-four hours operated servers require a good cooling system that might double up the electricity usage. Therefore, reducing the number of servers without compromising the system reliability will save the service provider's operational cost and at the same time save the earth.

1.3 Objectives Of Study

In order to achieve the research aim, few objectives has been set which are:

I. To obtain highest reliability based on target and minimum storage based on the budget allocated.
II. To solve the multi-objective problem using the linear programming method.

1.4 Scope Of Study

Nowadays, large-scale cloud-based applications have put forward a higher demand for storage ability of data centers. Data in the cloud storage system needs to be stored with high efficiency and cost-effectiveness while meeting the requirement of reliability. In order to ensure data reliability in private cloud, cloud storage environment offers replication that sometimes is not needed by organizations. To achieve the cost-effectiveness in storage infrastructure, strategies to minimize a storage system without compromising its reliability is the focus of this research. This will be the key factors of effective storage strategies in order to save high cost on large storage centers. The focus of this research is on the storage environment itself without considering the virtualization of the storage structure.

1.5 Thesis Outline

Chapter 1 of this thesis explains on the research background, problem statements, the objectives and scope of research. Detail explanations on the problem statements will be discussed in Chapter 2 as well as the revision of related works, research gap, storage reliability, failure rate, power consumption and carbon dioxide. Research methodology in Chapter 3 will discuss in details on how to calculate system reliability by using failure rate and produce solutions for minimizing cost, but maximizing reliability, and apply effective solutions based on organization's constraint. Chapter 4 will discuss detail implementation of program, flow chart, codes as long as reliability calculation and expected output. Chapter 5 is discussing the details of testing results, and optimization in order to achieve minimal operation cost with maximum system reliability. Finally, Chapter 6 will conclude the entire thesis and provide recommendation for future work.

1.6 Summary

Reliability is the main criteria for the cloud storage system. There are several ways to achieve system reliability, and some of them require high resource usage in term of operational cost and cost of server itself. Understanding system reliability is the main concern in order to be able to provide the solution to this problem. Next chapter will explain in detail about reliability and some of the previous research carried out by others.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

As more and more cloud-based applications tend to be data intensive, huge amount of data are stored in cloud infrastructure. The problem of managing the huge amount of data will require high amount of storage capacity and cost. Few researches are conducted to manage these issues. Most of them manage to reduce the redundancy by proposing a new method in term of backup and fault tolerance. Instead of providing fault tolerant backup, cloud storage could be optimized to reduce cost by managing data storage which requires huge storage size. This chapter will review some of techniques used by researchers to optimize reliability in cloud storage. At the end of this chapter, a strategy to manage storage reliability based on cost optimization will be proposed.

2.2 Storage Reliability

A simple definition of storage is a place or locations where users save and manage data. In order to do such operations, users need to frequently access the storage device. In the internet era, the storage enables user to share its content through a network as a server. While the evolution of internet plays an important part in managing and improving businesses online, the server capability has become a great concern. As we know that the server stores vital organizational data which are being accessed frequently, hence the reliability of a server becomes a major concern.

As we know that all electronic components in one way or another will exhaust its lifetime and therefore tend to crash by any factor. This keeps researchers working hard to find a solution to the problem. By implementing backups of current server with additional servers, they help organizations maintain the reliability of the cloud storage system. At the same time, the growths of data each day will also affect the cost of storage itself. Therefore, we found that, adding multiple backup servers without appropriate strategies is not a good solution to overcome storage reliability issue. In order to handle massive amounts of data with minimal cost, the appropriate storage reliability model needs to be implemented. Since the performance changes during the product's storage period, it has nothing to do with stress of storage servers. The main reasons responsible for the failure are the release of stress, which is created during the processing and assembling stages, the decay of materials, corrosion of electronic products, open and short circuit, and the material aging (Hong, 2012)

2.3 Previous Study

Few frameworks and building blocks of cloud server architecture in the research paper by Bakshi (2011) serves as a guidance for enterprises and service providers before considering cloud data center. Major motivation of cloud storage implementation is to reduce operating cost and increase reliability of the system. In order to adopt cloud architecture, the developer needs detailed explanations on how exactly cloud storage system is able to reduce the operational cost. The researcher groups the adoption steps into three phases which are: (1) Architecture development, (2) Migration and transition, and (3) Operation and optimization (Bakshi, 2011). The architecture development phase creates an accurate inventory of the application and data center related to the organization. The second phase applies to new data centers as well as redesign data centers and the third is focused on the operation and performance of the cloud system.

The cloud's storage major consideration is about its reliability. Most of the cloud storage providers are likely to offer a 99.9 % uptime guarantee. This means, the system will be online and available for access twenty-four days a week without fail. In order to understand reliability in detail, few research papers were used as references. In a paper authored by Huang et al., (2011), redundancy scheme was classified into replication and erasure codes. Replication scheme is simple and intuitive which replicates the original data into the numbers n of that, then distribute them to the different network nodes. Erasure codes split the original data into k blocks and encode them into (n-k) blocks with n blocks in total, then distributed these n blocks into different nodes. The study on erasure codes implementation towards redundancy consumes smaller storage space with the same reliability as compared to replication technique (Huang et al., 2011). During the implementation process, this method takes several steps to complete. At the first stage, storage allocation will be built using generic function. Then it will be transformed into the problem to calculate the multiplication on several polynomials. After that an erasure codes, algorithm is applied to find the minimum redundancy but satisfying reliability. In order to evaluate the method, there are a few parameters which will be used. These are reliability for all nodes ($\bar{r}$), standard deviation of node's reliability (δ), number of stored block's (l), number of blocks to reconstruct data (k), total number of blocks (n), power of the combinations (p), and cumulative data reliability for the power not more than p (CDR(p)). Based on the parameters, the impact of $\bar{r}$, δ, and l toward data reliability and redundancy optimality is easily identified.

In the research proposed by Feng *et al.,* (2012) he was more focused on finding a solution to make a trade-off between high reliability and low space overhead for cloud storage systems. High reliability and low space cost are always contradictory in existing storage systems. This research paper introduces Magicube, a storage architecture with high reliability and low space overhead for cloud computing. To reduce the space overhead of file storage, Magicube keeps only one copy of each file in HDFS (Feng *et al.*, 2012), and to achieve high reliability, it uses a special encoding algorithm for fault-tolerance. Once the system parameters (n,k) have been set, the source file will be encoded and the chipper text will be split into n splits, then stores the corresponding file splits in the cluster (Feng et al., 2012). If the source file is lost, Magicube can repair the lost file from any k (k < n) of splits of that encoded file. It also allows users to adjust the system the system's fault-tolerance level and space cost by changing the setting of (n,k). Based on the research paper, to achieve high reliability and low space coast, there are several methods that can be used, which are Triplication at the beginning, File splitting and distribution, Extra replication deletion and File repair. All experiments were carried out on a cluster of nine nodes, where a node served as a master and the others as client and slaves for MapReduces and HDFS (Feng et al., 2012). Each node has two quad-core 2.40GHz Xeon processors, 11GB of memory and two disks. The experiment result shows that the (n,k) coding introduced negligible influence on the performance of Magicube, and the file repairing performance was also acceptable.

Meanwhile, LI et al., (2011) proposed a replication strategy that works especially for data that are used temporarily and/or have a relatively low reliability requirement. Based on the observation, the researcher found that the typical 3-replicas data replication strategy or any other fixed replica number replication strategy might not be the best solution for data with uncertain reliability duration. Most cloud systems contain a large number of tasks. During the execution, large volumes of intermediate data are generated and the amount could be much larger than the original data, but most of them are only aimed for temporary use. All the intermediate data are deleted after use or some of these will be stored for later use, but for an uncertain period of time (LI *et al.,* 2011). The average failure rate of a single storage unit in most cloud infrastructure without replication is 0.01%. Such reliability assurance and storage duration is sufficient to meet the requirement of most intermediate data in scientific applications without extra data replication. Thus, in order to reduce the storage cost and fully utilize the storage resource in current cloud systems, researchers propose a strategy named Cost-effective Incremental Replication (CIR). It is an incremental replication approach that works by calculating the replica creation time based on prediction, which indicates the storage duration to meet the reliability. To evaluate the performance of the proposed model, there are a few parameters which are being used, namely: number of data instance, storage unit number, reliability assurance, storage duration and cost. Based on the evaluation, the CIR strategy can reduce the data storage cost substantially, especially when the data is only stored for short duration or have a lower reliability requirement.

2.4 Research Gap

The storage reliability of components is especially important to the electronic products which store for a long-term and are used at one time (Yang, 2011). Based on the previous studies, the main issues in the research are reliability and cost. Most of the researches focus on finding minimal redundancy using erasure code, increasing the storage reliability from time to time by creating an additional replica to current nodes. By doing these, redundancy issues are yet to be resolved because the existing storage size will keep on increasing over time. What will happen to the existing storage nodes if it failed to operate in a certain amount of time? Moving replicas from the existing server to another is a good solution as a backup, but for how long will the recovery process take to ensure the reliability of the storage.

From the previous study, the reliability of the storage system is $Rs = 1 - qc$, where $q=1-p$ and p are the reliable server, c is the number of servers existing in the system. By using this formula, we can assume the reliability of a storage system is set to $1 - qc$, where q is the unreliability of the storage which is the system failure rate. In cloud storage system features, each of the data is stored using minimum 3-replication strategies, which mean the reliability of the storage is equal to *1 – q3* reliability. The basic storage rules provide good data access to the organization, but it will become insufficient if the data keep increasing from time to time. If a certain amount of data is to be stored in a certain amount of time, the exact number of storage system needs to minimize the server cost.

Most applicable methods to improve storage reliability and availability is by using redundancy scheme. Redundancy scheme is divided two classes: Replication and Erasure codes. Replication works by replicating original data into n copies and distribute to different nodes in or outside the network while erasure codes split the original data into k blocks, and encode them into *(n-k)* blocks with *n* blocks in total, then distribute these n blocks into different network nodes. From this solution, reliability equation *Rs=1-qc* will be used to

i. determine how many storage will be used,

ii. predict how long it will be operated until it gets to the next failure state,

iii. how long it will take to recover and

iv. how many storage nodes are required to ensure reliability.

2.5 Failure Rate

Failure rate (λ) is a number of failures per unit depending on operating hours. Each of storage devices consists of a failure rate during its operation, which is the number of units failing per time. Therefore, the failure rate is used to measure the reliability of a storage based on the probability of the storage itself. In the meantime, lower failure rates indicate higher reliability of storage. According to Rooney (1989), failure rates of storage are estimated at 2.64 and up to 5.90 failures per million hours.

Failure rates determine the reliability of a storage framework to which it displays the frequency of failure a device undergoes within a specific period (Torell, 2004). In a server, system failure rate is being identified according to the position of the storage, whether they are stacked in serial or parallel order. Another factor that directly affects the system reliability is the failure rate of each of the storage contained in the system. Each of the storage device contained in the system might have different failure rate or might not i.e. the system might contain identical or different storage failure rate.

2.5.1 Failure Rate for Serial System

Serial systems work with two or more storage combination operating in series. The failure of either storage will result in failure of the combination. The failure rate of combined storage in the series is shown as follows:

$$Rs = FR_A * FR_B \tag{2.1}$$

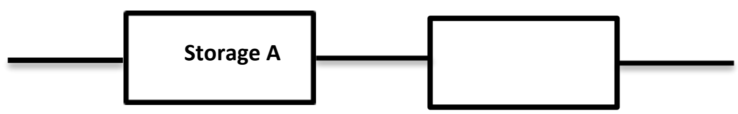

Figure 2.1: Serial Storage Connection

2.5.2 Failure Rate for Parallel System

Parallel systems work with two or more storage combination operating in side by side in parallel. The system remains operable unless all storages fail. The failure rate of parallel system shows as follows:

$$Rs = 1 - (1 - FR_A)(1 - FR_B) \quad (2.2)$$

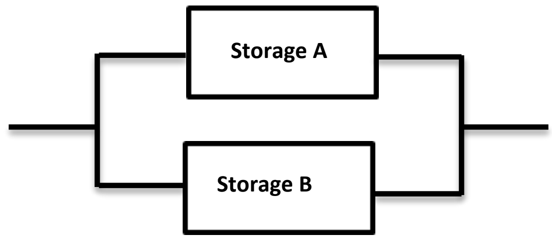

Figure 2.2: Parallel Storage Connection

2.5.3 Identical and Different Storage Failure Rate

Cloud servers consist of multiple storage units, and in order to identify its reliability, the failure rate of the system is calculated by finding the failure rate of each storage unit. A set of storage units in a server that have the same failure rate is called identical failure rate, and a set of storage that have varying failure rates is called different failure rate.

2.6 BATHTUB CURVE

Every single device comes with a lifespan. Failure rates and time are used to determine the reliability of a device during its life span. Normal life span for electronic device can be divided into three distinct periods (Klutke, 2003). Bathtub curve is used to show how the failure rates vs. time for each of an electronic device to help users predict the actual failure before it is happening. This helps a lot in managing user access and retrieval to the device.

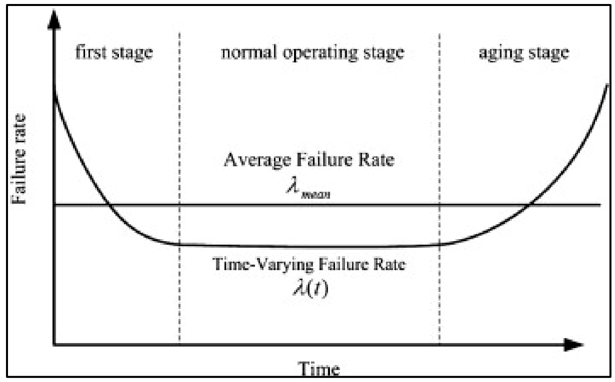

Figure 2.3: Bathtub Curve (Klutke, 2003)

The bathtub curve consists of three periodic stages which are infant mortality period (first stage) with a decreasing failure rate followed by normal operating stage period with a low, relatively constant failure rate and concluding with aging stage that exhibits an increasing failure rate. Bathtub curve does not depict the failure rate of a single device, but describes the relative failure rate of an entire population of devices over time. Some individual units will fail relatively early, the other will last until wear-out and some will fail during normal operating state.

2.7 Power Consumption

High-end servers use a lot of electricity to operate, and about the same is used to cool it down. Using electricity without proper control will affect global warming since the emission of carbon dioxide (CO2) from electric device is high especially a computer server with non-stop operation. Before developing a storage system, an organization needs to manage and study the system reliability prediction in order to minimize the power consumption.

2.8 Carbon Dioxide Emission

Every electric device emits carbon dioxide (CO_2) for every kilowatts of usage. According to Barett (2011) in his article on eHow.com, a server system consumes electricity from 500 to 1200 watts per hour. According to Carbon Fund (2012), on average electricity usage emits 1.222lbs per kilowatts (kWh) or approximate 0.5533827kg per kWh. If on the average, a server used 850 watts per hour, and about 7,446 kilowatts per year, that means more than 4120kg carbon dioxide (CO_2) is released into the atmosphere. In order to reduce the carbon dioxide emissions, less storage system is required without compromising the system reliability.

.

2.9 Summary

In next chapter, we will discuss about the strategies to minimize storage cost by reducing the number of server but maintaining the reliability of data. We will be focusing on the mean time between failure as well as mean time to repair and suggest a tradeoff between costs and performance based on the needs of the organization.

CHAPTER 3

METHODOLOGY

3.1 Introduction

In order to determine storage reliability and minimal operation cost, several factors need to be considered. This research focuses on how organizations can obtain storage reliability using the allocated budget. Instead of having budget as the key point to the system reliability and optimization, few variables such as failure rate, number of server and duration of the storage system are used. The research methodology is used to identify the reliability formula based on previous study and analyzes the appropriateness to organizational requirements based on current cloud storage pricing structure. The research also focuses on identifying the minimum number of servers required based on the budget allocated.

3.2 Operational Framework

The aim of this research is to get high reliability for cloud storage framework subject to minimal cost allocated. In order to achieve the objective, multi-objective optimization will be used to solve multi criteria decision making involving more than one objective function to be optimized simultaneously. At the end of this research, a table consisting of comparison from the selected variable will be published as analysis result and determine the best storage implementation depending on the number of servers required to maximize the system reliability.

In order to maintain its reliability, cloud storage uses replication strategies to ensure that all of its servers are operating using the latest updated data whenever one or more of its server fails. To avoid wasting storage resource during replication strategies, an exact calculation on estimation of when and how many times the server will fail in certain period is needed. To achieve these objectives, storage failure rates are used.

Figure 3.1shows the relation between servers in cloud storage:

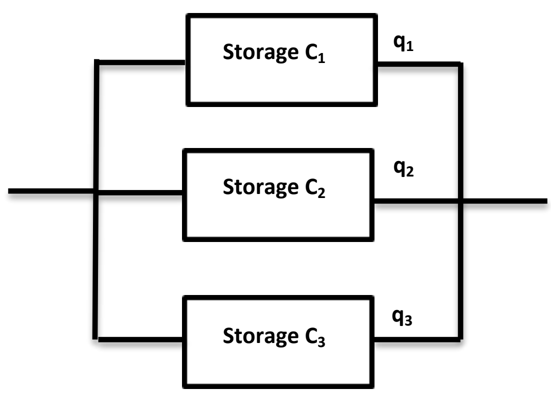

Figure 3.1: Relation between Servers in Cloud Storage

Figure 3.1 indicates that in a cloud storage framework consist of three servers' *c*. Each of the servers has a reliability of *q*. The overall storage framework reliability is represented as $Rs = 1 - q^3$, that is achieved from the equation of $R = 1 - q_1 q_2 q_3$ or $R = 1 - \prod_{i=1}^{3} q_i$ (3.1)

In order to determine the reliability of the storage, unreliability of each of the storage will have to be identified. The unreliability of each of the storage is calculated using failure rate (λ) or *q*.

3.3 Failure Rate Calculation

The failure rate of a storage device can be found from the hardware itself, which is provided by the device manufacturer. Failure rate enable user to identify how reliable the device is based on number of predicted fail per certain of times. For example, a device having 0.2%/1000 hours, which means the device will fail to operate two times within one million hours. This can be expressed on the following equation:

$$q = \frac{0.2}{100} * \frac{1}{1000} = \frac{2}{10^6} \qquad (3.2)$$

From the failure based on failure rate per 1000 hours, the probability of failures per year could be calculated as:

$$q = \frac{0.2}{100} * \frac{1}{1000} * 24 * 365 = 0.01752$$

3.4 Reliability Calculation

Based on the failure rate, the reliability of a device with 0.2% of failure rate per 1000 hours can be express as:

$$Rs = 1 - q$$
$$Rs = 1 - 0.01752 \qquad (3.3)$$
$$Rs = 0.98248$$

That means, if a device with such failure rate operates 24 hours a day for one year, the probability of a device will be reliable is 98%. In order to increase the reliability to 99%, one additional server with such failure rate is required.

$$Rs = 1 - q^2$$
$$Rs = 1 - (0.01752)^2$$
$$Rs = 0.999693049$$

The reliability will decrease by 1% for every server over a year and will require additional server to maintain the reliability. With minimum two servers, the storage reliability, maintained at 99% at least for five years.

If an organization need to set up a storage system using three servers for the duration of three years, using the previous failure rate for each server, the following calculation applies.

Year 1:

Failure rate: $q = \frac{0.2}{100} * \frac{1}{1000} * 24 * 365 = 0.01752$

Reliability:

$$Rs = 1 - q^3$$
$$Rs = 1 - (0.01752)^3$$
$$Rs = 0.999994622$$

Year 2:

Failure rate: $q = \frac{0.2}{100} * \frac{1}{1000} * 24 * 365 * 2 = 0.03504$

Reliability:

$$Rs = 1 - q^3$$
$$Rs = 1 - (0.03504)^3$$
$$Rs = 0.999956977$$

Year 3:

Failure rate: $q = \frac{0.2}{100} * \frac{1}{1000} * 24 * 365 * 3 = 0.05256$

Reliability:

$$Rs = 1 - q^3$$
$$Rs = 1 - (0.05256)^3$$
$$Rs = 0.9998548$$

3.4.1 Variable and Calculation Formula

In order to identify the reliabilities of a storage system for certain duration, there are few variable are used. The variable listed as follows:

Table 3.1: Variable for Reliability Calculation

Variable	Description
n	number of server
h	constant number of hours server operate (24 hours)
d	Constant number of the day server operates (365 days)
y	total number of years to calculate the server reliability

mh	million time is set to 10^{-6}
nF	number of failed per million hours
b	budget of the organization
Sc	cost per server
Mc	Minimum cost
Tr	Target reliability
Rs	System reliability

Number of possible server (*n*) will be calculated from the organization budget (*b*) and cost per server (*Sc*) based on market price.

$$n = b/Sc \qquad (3.4)$$

Reliability calculation formula

$$Rs = 1-[(nF*mh)*(h*d*y)]^{n} \qquad (3.5)$$

Reliability of the system will be calculate by using the formula to identify what is the reliability of the storage based on variables that were listed previously. The result will be displayed is the system reliability (%) based on year and the number of storage used.

3.5 Multi Objective and Linear Programming

In order to find system reliability, there are several objectives used as the constraint. Thus, the multi-objective optimization is used. Multi-objective optimization involves more than one objective function in order to optimize the result (Deb, 2014). Multi objective optimization is used to find minimum and maximum value from multi variable constraint. In this thesis, reliability is calculated using the constraint from cost and duration to find maximum reliability with lowest cost allocated.

The main objective of this research is to provide reliability solutions for organizations if they want to create cloud storage environment. For this purpose, several constraints are count for calculation. The focus is to find maximum server reliability based on target reliability and the minimum numbers of server involve in order of achieving maximum reliability in target duration. Number of server involve is relate to budget allocated.

In order to achieve result base on several objective, optimization technique is required. In this problem, linear programming is use to find the maximum and minimum point in each calculation The following formula shows the relation of multi objective variable based on constraint below:

Let,

i is the number of server,

$q = 1 - p$ is the probability of failure,

The maximum reliability is given as,

$$\text{Maximum } Rs = (1 - \prod_{i=1}^{n} q_i) \quad (3.6)$$

Subject to:

$$\sum_{i=1}^{n} i \leq budget$$

and

$$\bigcup_{i=1}^{n} i \geq duration$$

3.6 Application Development

Based on reliability calculation formula, we were able to calculate the system reliability for a given number of servers and duration. In order to find the system reliability based on certain duration, an application is needed to automate the calculation.

3.7 Summary

Research methodology is the most important part in which we will provide the guideline and procedure on how to get the data and verify what to do in the next chapters. With all variables used, input, expected result and appropriate steps were being explained within this chapter. This will make the implementation phase come in handy. Details of application to calculate the system reliability is discussed in the next chapter.

CHAPTER 4

IMPLEMENTATION AND TESTING

4.1 Introduction

With the detailed explanation about the formula in previous chapter, this chapter will conduct an implementation and testing in order to find the storage reliability based on few variables selected. The objective of this chapter is to apply what have been described in previous chapter which is the formula to find the system reliability and execute it in a program in order to find the reliability based on variables set by user.

4.2 Program Implementation

In order to find the system reliability, a custom program was used to test all variables which have been mentioned in the previous chapter. HTML and JavaScript were used as the implementation platform since they were the most compatible program and can be run in many browsers. The implementation process requires input from the user in order

REFERENCES

Bakshi, K. (2011, March). Considerations for cloud data centers: Framework, architecture and adoption. In *Aerospace Conference, 2011 IEEE* (pp. 1-7). IEEE.

Borthakur, D. (2007). The hadoop distributed file system: Architecture and design.

Carbon Fund (2012), *"How We Calculate"*, Retrieve October 13, 2013, from URL: http://www.carbonfund.org/how-we-calculate

Deb, K. (2014). Multi-objective optimization. In Search methodologies (pp. 403-449). Springer US.

Feng, Q., Han, J., Gao, Y., & Meng, D. (2012, June). Magicube: High Reliability and Low Redundancy Storage Architecture for Cloud Computing. In *Networking, Architecture and Storage (NAS), 2012 IEEE 7th International Conference on*(pp. 89-93). IEEE.

Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., ... & Stoica, I. (2009). Above the clouds: A Berkeley view of cloud computing. *Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, 28.*

Hong, W., Wang, S., Wang, X., & Wang, W. (2012, July). Overview of storage reliability for high reliability products. In Industrial Informatics (INDIN), 2012 10th IEEE International Conference on (pp. 682-687). IEEE.

Huang, Z., Yuan, Y., & Peng, Y. (2011, May). Storage allocation for redundancy scheme in reliability-aware cloud systems. In *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on* (pp. 275-279). IEEE.

J.T Barett (2011), *"How Much Electricity Does a Computer Use Per Hour?"*, Retrieve October 13, 2013, from URL: http://www.ehow.com/info_8763694_much-computer-use-per-hour.html

Klutke, G. A., Kiessler, P. C., & Wortman, M. A. (2003). A critical look at the bathtub curve. *Reliability, IEEE Transactions on*, *52*(1), 125-129.

Li, W., Yang, Y., & Yuan, D. (2011, December). A novel cost-effective dynamic data replication strategy for reliability in cloud data centres. In *Dependable, Autonomic and Secure Computing (DASC), 2011 IEEE Ninth International Conference on* (pp. 496-502). IEEE.

Rooney, J. P. (1989, January). Storage reliability. In *Reliability and Maintainability Symposium, 1989. Proceedings., Annual* (pp. 178-182). IEEE.

Sourabh (2012), *"What is Cloud Computing"*, Retrieve March 10, 2013 from http://sourcedigit.com/41-what-is-cloud-computing.

Torell, W., & Avelar, V. (2004). Mean time between failure: Explanation and standards. White Paper, 78.

Wu, J., Ping, L., Ge, X., Wang, Y., & Fu, J. (2010, June). Cloud storage as the infrastructure of cloud computing. In Intelligent Computing and Cognitive Informatics (ICICCI), 2010 International Conference on (pp. 380-383). IEEE.

Yang, Y., Fu, G., Wan, B., & Gu, H. (2011, May). Component storage reliability evaluation method based on multi-performance parameters degradation analysis. In Prognostics and System Health Management Conference (PHM-Shenzhen), 2011 (pp. 1-6). IEEE.