# Availability Modelling of Fault Tolerant Cloud Computing System

**Deepa Mani[1]***      **Anand Mahendran[2]**

[1]*School of Information Technology and Engineering, VIT University, Vellore, India*
[2]*School of Computer Science and Engineering, VIT University, Vellore, India*
* Corresponding author's Email: mdeepa@vit.ac.in

**Abstract:** Cloud management organisation is an imperative part of cloud computing platform and serving as the resource manager for cloud platforms. The multifaceted nature of cloud-management base makes its high availability (HA), a standout amongst the most necessities. Different innovations have been produced to build a consistent quality and availability of cloud administration infrastructure. In any case, little work concentrated on quantitative examination of its accessibility. While this ability accomplishes a developed availability with small fault rates, corporate requests conveyed over the autonomous zones may encounter unique Quality of Service (QoS) because of various physical frameworks. The key target of this paper is to show how the Markov-based model can fulfil the client request. For this reason, a few scenarios of the failure rate of virtual machine's practices, for example, single system failure, multiple system failures, power outage are considered by applying the Markov model. The improved repair strategies of accessibility in various circumstances are also investigated. The Queuing models like Markovian and non-Markovian models are examined using phase type expansion and renewal theory keeping in mind the end goal to sufficiently speak to and to assess. The considered element unwavering quality perspectives if there should be an occurrence of the most part dispersed lifetimes and times to repair.

**Keywords:** Availability, Homogeneous continuous-time Markov chain (HCTMC), Quality of service, Mean time to failure (MTTF), Virtual machine.

## 1. Introduction

Cloud computing is the tactic to uplift the capacity and competencies without capitalising in licensing new software, building new infrastructure, and providing training to new workforces. It helps to develop the prevailing of IT industries [10]. One of the most prominent areas where cloud computing has grown faster in the business field. As more and more information on individuals and companies can be placed in the cloud, concerns are beginning to wind up about precisely how safe a situation it is [2].

Availability is one of the non - productive constraints which can sometimes express regarding a system or a service accessible by the customer. These terms identify the downtime of the administration for a given period [1]. One of the most stringent requirements is high availability, and

it provides the system downtime and incorporates the maintenance of the system which is, planned or unplanned. Although there is some reward for the customers to reserve data, some issues need to be addressed [3]. The software or hardware failure may happen due to the unavailability of the service provider, in that situation the clients could not recover the data. So, for this reason, the customers may not guarantee long-term availability of their cloud storage services [4].

One of the risks of protecting data from the failures is hardware, software failures, computer viruses and natural calamities, and the finest remedy is data backup [12]. With the help of the protecting techniques like data replication and mirroring, the data can be protected. Sometimes these techniques will be helpful to provide a high availability of data [5]. But sometimes these methods will fail to avoid

the downtime overhead, and it will lead to performance deprivation. So there is a requirement of operation up gradation by the fluctuations in technologies and the business [9].

From the availability system perspective, it is possible for the component to be contingent or simulate the other element. Some of the dynamic performances are a facility to repair the effect of ageing, causes of the method of let-downs and so on [7]. Specifically, ageing is one of the factors while assessing the availability of the virtual machine [11]. One of the major functionality of the cloud service providers is to measure their service significantly and make it available to the users also. For representing the component availability and the failure detections, various approaches used [3]. The model considered in this paper is a stochastic model, which is the best choice to model the dependability and the variability [15]. The types are classified into Markovian or non- Markovian process. Markov models are expedient when solving the decision problem [23].

The stochastic problem in Markov model can be solved by using the Markov property as follows, the future state process of the conditional probability depends on the current status only, and not the sequence of events indicated it [31]. One of the universally used stochastic models is Homogeneous Continuous-Time Markov chain (HCTMC) which applies the property of Markov to the events like service accomplishments, failures, repairs, etc.), with exponential repair time distributions [23]. In this paper, we consider the exponential distribution of Markov chain in the cloud environment and also cover the performance of the systems in cloud environment.

This paper develops the strategies for the practical implementation of non-homogeneous Markovian and non-Markovian state–space models in some specific consistency and availability scrutinises, addressing the diverse element aspects of the cloud environment. We outline the significant contributions of our paper:

- This paper discuss various heterogeneous availability models for a few failure structures, utilising measurements of limit arranged accessibility and framework steady state availability to look at those particular designs. The hardware failures are studied in the proposed explanatory models. The replicas are likewise used to acquire shut structure conditions, empowering proficient calculation of the pointed measurements

with the help of Continuous Time Markov Chain (CTMC) model.

- We have unified the fault tolerant and disaster sophisticated techniques, and also we have explained various system failures like a power outage, single system failure, multiple system failures, etc.,

The paper is organised as follows: We discuss problem issue and description methods of the counteracting problem in Section 1. We address the background in Section 2. Section 3 presents CTMC and numerical study and discussion. Section 4 shows the proposed system flow on failure analysis. In section 5 we analyse the models, and experimental results are given to approve the model arrangements. At last, we conclude the paper in section 6.

## 2. Background

This section deals with the ideas of clustering, virtualization and software rejuvenation.

### 2.1 Clustering Environment

Clustering has ended up being a powerful technique to scale larger systems for improved enactment, the number of clients, or different properties. It also provides more elevated amounts of availability and lower administration costs [14, 16]. A few clusters target to give zero downtime through joblessness; some intend to give high availability by load balancing responsibilities over the group, furthermore expect to deliver a superior performance using parallelism [28]. A target of the failure procedure is for it to happen smoothly, without an end client realising that a failure has happened. With a specific end goal to guarantee the capacity to play out a failover, adequate extra spares must be accessible to drifted workload [17].

### 2.2 Virtualization Environment

Virtualization is an innovation that provides various working situations like hardware and software dividing or collection, incomplete or complete machine reproduction, imitating, time-sharing, and others [19] that join or partitions are computing resources. Virtualization gives one PC a chance to carry out the occupation of various PCs, by sharing the resources of a single PC over different situations [30]. Virtual servers and virtual desktops have various working schemes, liberating from the confinements like physical and geographical. Notwithstanding vitality funds and

lower capital costs because of well-organized utilisation of hardware resources and upgraded catastrophe recovery forms when we construct a simulated base [18].

## 2.3 Virtualization of Clustering Environment

Virtualization of a cluster is the tremendously streamlined arrangement, administration of lacks of servers, finished through creating expensive gatherings of servers show up and performs like a solitary system, as simple to oversee as a solitary workplace [13].

Monetary and productivity advantages of this methodology are exceedingly convincing, the creation of virtualization of clustering are the most useful, knowledge systems. For diminishing intricacy, expense and general managerial weight of broad scale processing, empowering to get the most out of computing resources [20]. Virtualized clustering is additionally intrinsically intended for greatest information honesty and irrelevant recuperation time by location let-downs.

## 2.4 Renewal of software systems

Renewal of software is an active deficiency administration method intended for tidying up the systems internal state to avoid an event of more severe crash failure later on. By evaluating the collected fault conditions and operating system resources degradation, this strategy actively anticipates sudden framework blackouts [22]. Dissimilar to downtime brought on by unexpected let-down events, the lost time identified with software renewal can be planned at the attention of the client/manager. Software renewal is a cost-effective system for managing software shortcomings that incorporate insurance against hard failures, as well as against execution lack also. The software restoration method well suited with the arranged restart component, which can bring down the retrieval overhead to an incredible degree of a system.

## 3.   Related Works

Availability of the cloud is a pre-essential for the assessment and arrangement of at present carried and a key factor of research, connected components and advances for the cloud. Availability and high availability have characterised in a wide range of ways. Availability is one of the essential qualities of the cloud [24]. Their discussion shows that the failure rate and response time are not exponential

and the probability distribution for the high availability cannot be accomplished in the closed form. So,  a portion of exploration exertion is committed to ensuring the availability in the cloud. Considering an aspect in these endeavours, we can understand that the elucidation of availability may contrast from the suppliers. These days, storage providers offer another path for Cloud vendors to work together. This new pattern demonstrated by the quantity of distributed storage suppliers that are persistently showing up available. At this point, utilising cloud storage providers are turning into a typical practice for end-clients. In any case, the existing cloud storage suppliers do not offer any sureties on haul availability and protection.

Truth told, information put away in the cloud could be secured, lost, or damaged as far as protection [4]. They presented with several causes of achieving the long term availability. To claim the change of file dependability by bringing notice into an expansive storing framework abusing unique solutions,  such as eradication rectifying codes used in Redundancy Array of Independent Disks stages 5 and 6, introducing a few information arrangement, failure identification and exertion disciplines inside data centres [25]. They co- ordinated the reliability factor for QoS measures for the customers in cloud and cloud frameworks. Adjusted to that in [26] availability is characterised as "the extent to which a framework is working and is available to convey its administrations amid a given time interim". Their approach failed to achieve the steady state availability. It is the rate of time a structure is prepared to play out its tasks and reckoned as [29]. Availability is the coefficient of the predictable time to failure for a *non-repairable system* and the sum of the time required to repair a system. The failure time is calculated by using MTTF (the expected time to failure for a *non-repairable system*); MTBF (average time excludes the time spent waiting for repair) and represents the sum of MTTF and MTTR. (The amount of time required to repair a system) [26] Their approach failed to achieve the steady state availability. One can see the availability of a framework through the accessibility of its services. Service availability can characterise as:

Service Availability is the coefficient of service uptime and service outage. Where, service uptime is the span among which the framework conveys the given service, the amenity blackout (or additionally referred as stoppage) is the period of the administration not communicated [26]. Consolidate both sophisticated deficiency strategies, (for example, an excess of the swarm and Virtual Machine (VM), VM failure, and VM live migration)

and catastrophe understanding procedures, (for example, dispersed geological repetition, site fall flat over and reinforcement server establishment) into the framework [6]. The major objective of the proposed method disaster tolerant data center (DTDC) is to provide system availability, efficiency and downtime cost.

The steps developed of open communication administrations, e-trade, and distributed computing based organisations have pushed an immeasurable advancement on server farm frameworks lately [21]. Their proposed method procedure used to optimise the performance with the already existing methods. Moreover, the vital of keeping up client's certainty and avoiding income misfortunes requires High Availability (HA) of cloud administrations facilitated on server farms. The distinction in minutes of inaccessibility, while apparently apparent, which expense the manager a large measure of cash because of interference of business exchanges [5]. Their discussion shows that process of improving the system data availability and storage system availability. The theme of information replication in geologically disseminated cloud computing server groups and proposes a novel replication arrangement which notwithstanding customary execution measurements, for example, accessibility of transmission system capacity, enhances vitality proficiency of the framework [27]. They co-ordinated the performance estimation metrics like availability and energy efficiency for QoS measures for the customers in cloud and cloud framework. What's more, streamlining of correspondence delays prompts the enhancements in nature of client experience of cloud applications. The studies aimed for surveying the determination of cloud frameworks don't address programming constancy or reflect the impact of inserting new bits of gear to give access to existing models [8]. They defined the model hierarchical heterogeneous model which is composed of reliability block diagrams (RBD) and markov reward model (MRM) which addresses the performance metrics like cost, energy and availability. Present studies on trustworthiness forecast for open source bunches of servers. They expect to upgrade the high availability (HA) highlight to Open Source Cluster Application Resources (OSCAR). The authors anticipate framework consistent quality and accessibility through Stochastic Reward Nets (SRN). The espoused Continuous Time Markov Chains (CTMC) to examine the accessibility of a group framework with different nodes [28]. Their examination considered in both continuous mode failures (CMF) and no CMF for the group of the node. CTMCs

additionally utilised as a part of for displaying HA clusters. Their model was intentionally devised to find the high availability of the cloud systems by using the analytical models.

## 4. System Organization

This study depicts our proposal to suggest accessibility component based on the failures generated in all possible components of the cloud system. By interfacing two servers, we can achieve high availability with the help of the dynamic/ reinforcement node structure. The reinforcement node screens the active node while the vibrant centre involves assets and provides administrations given the possessed assets. If failures are created in the dynamic hub, the reinforcement hub assumes control over the resources of the dynamic hub and ceaselessly gives a relating administration in light of the resources. The consistent administration has granted to an external client from side to side such a dynamic/reinforcement node construction. Virtualization technologies permit to bind various virtual servers on a solitary physical engine. Virtualization of cluster empowers expansive cluster of servers to performance and seem like a solitary, steady virtual machine. The subsequent scheme discussed in this paper,

- Single virtual machine failure
- Multiple virtual machine failure
- Hardware failure
- Power failure and Network failure

### 4.1 Proposed system flow

In this paper, we have referred some formulas [29]. Based on our proposed work, the variables/ parameters in the formulas are slightly modified.

Assume that the availability model of cloud computing has 'n' components in a virtual machine. Let us suppose that the components of the model ordered according to the critical parallel redundant system. Let $\mu_n$ be the repair rate for n components and $\lambda_n$ is the failure rate for n components. In cloud

Computing, there might be several types of failures

**Nomenclature**

Table 1. Input parameter values

| Notation | Meaning |
|---|---|
| N | Number of Virtual Machines |
| $\mu_n$ | Repair rate of 'n' components |
| $\lambda_n$ | Failure rate of 'n' components |
| $\rho$ | $\dfrac{\lambda}{\mu}$ |
| P | Probability |
| $\mu_1, \mu_2\ldots \ldots \mu_n$ | Repair time distribution for 'n' components |
| La$_f$ | Laplace transforms density function and f is underlying density f(t) |
| La$_w$ | Laplace transforms density variable and w is common probability density function w(t) |
| La$_g$ | Probability density function and g is probability density function g(t) |
| La$_m$ | Average number of repairs and m is replacements m(t) |
| MTTF | Mean Time To Failure |
| MTBF | Mean Time between Failure |
| MTTR | Mean Time To Repair |

will take place such as hardware and software failure, network failure, VM failure, power failure, application failure and soon. At the point, when a failure happens in any one of the components, there is a case of having backup storage. In cloud computing environment these elements frequently contain front end platforms that are a thick client, thin client, and mobile device. On the other hand the backend platform such as servers and storage and network like the internet, intranet, and intercloud.

## 4.2 Availability analysis for single virtual machine

The model parameters chosen for availability modelling are listed in Table 1.

Case 1:

In this section, we discuss single virtual machine failures in the cloud system, N=5 where N is the number of the virtual machines, among that if the single virtual machine fails. Then for the users, another virtual machine will be available, and each component has its repair facility with repair rate $\mu_{n,}$ then the availability of individual components and $\lambda_n$ is the failure rate for n components. Assume that the failure of a segment it is restored and reestablished to be on a par with new one. Let Time$_j$ be the span of the j$^{th}$ working period, and Do$_j$ is the system downtime for the j$^{th}$ repair or replacement. Let us assume that the sequence of random variables $\{X_j=Time_j+Do_j\}$ Where $(j=1,2,\ldots,n)$ is mutually independent. Then X$_j$ are also distributed, $\{X_j/ j =1,2,\ldots,n\}$ is a renewal process. Let the fundamental density f(t) of the reestablishment procedure is the intricacy of w and g. Thus, the Laplace transform of the function, we have,

$$La_f(s) = La_w(s)\, La_g(s) \tag{1}$$

By using the convolution property of transforms, we get,

$$La_m(s) = \frac{La_f(s)}{1 - La_f(s)} \tag{2}$$

Using this equation we have,

$$La_M(s) = \frac{La_w(s)\, La_g(s)}{1 - La_w(s)\, La_g(s)} \tag{3}$$

the average number of repairs are M (t) in the time interval (0,t), and the Laplace transforms as follows,

$$La_M(s) = \frac{La_w(s)\, La_g(s)}{S\left[1 - La_w(s)\, La_g(s)\right]} \tag{4}$$

The instantaneous availability of the component is Avail (t), and t is the system functioning time. Let us consider that during the absence of repair, the availability is merely equal to the reliability $R(t) = 1-W(t)$ of the component. After substituting, we get,

$$LaAvail(s) = \frac{1 - La_w(s)}{S\left[1 - La_w(s)\, La_g(s)\right]} \tag{5}$$

Based on the factors like expected time to failure for a *non-repairable system* and the amount of time required repairing a system, we have restrictive

accessibility as the final- value theorem of Laplace transforms, Avail=$\lim 1-[1-\frac{s}{\lambda}]$ (6)

Where, $s \rightarrow 0$.

Let us assume that each component has its own repair facility.

Steady state availability = $\dfrac{1}{1+\dfrac{\lambda}{\mu}}$ = $\dfrac{1}{1+\rho}$ (7)

Where, $\rho = \dfrac{\lambda}{\mu}$

Then system availability avail$_I$=$1-\left(\dfrac{\rho}{1+\rho}\right)^n$ (8)

## 4.3 Availability analysis on cost

Case 2:

We need to conserve on the repair facility and offer a separate repair facility of rate μ among all the "n" virtual machines. When the virtual machines performance is down, then only the components will undergo for repair.

Then virtual machine availability *avail$_{II}$ = 1- P$_n$*
Let us consider k systems,

$P_k = P_0 \left(\dfrac{\lambda}{\mu}\right)^k \dfrac{n!}{(n-k)!} = P_0\, \rho^{\,k}\, \dfrac{n!}{(n-k)!}$ (9)

Hence, $P_0 = \dfrac{1}{\displaystyle\sum_{k=0}^{n} \rho^k \dfrac{n!}{(n-k)!}}$

avail$_{II}$=$1-P_n$= $1-\dfrac{\rho^n n!}{\displaystyle\sum_{k=0}^{n}\rho^k \dfrac{n!}{(n-k)!}}$ (10)

## 4.4 Availability analysis on the performance

Case 3:

If avail$_{II}$ is low, at that point we have to accelerate the procedure of repair rate facility by μ$_n$,

We get, avail$_{III}$ = $1-\dfrac{\left(\dfrac{\lambda}{\mu}\right)^n n!}{\displaystyle\sum_{k=0}^{n}\left(\dfrac{\lambda}{n\mu}\right)^n \dfrac{n!}{(n-k)!}}$ (11)

## 4.5 Availability analysis on multiple virtual machines

In cloud system, if multiple virtual machines fail, then the system is unavailable for the user, and each component has its repair facility with repair rate μ$_n$, then the availability of individual components,

Case 1:

Steady state availability = $\dfrac{1}{1+\dfrac{\lambda_n}{\mu_n}}$ = $\dfrac{1}{1+\rho_n}$ (12)

Where, $\rho_n = \dfrac{\lambda_n}{\mu_n}$

Then system availability avail$_I$= $1-\left(\dfrac{\rho_n}{1+\rho_n}\right)^n$ (13)

## 4.6 Availability analysis on cost

Case 2:

The reduced cost of the repair facility and share a multiple repair facility of rate μ among all n components. When the all the systems performance is down then only all the components are undergoing for repair.

Then system availability *avail$_{II}$ = 1- P$_n$* (14)
Let us consider k systems,

$P_k = P_0 \left(\dfrac{\lambda_n}{\mu_n}\right)^k \dfrac{n!}{(n-k)!} = P_0 \rho_n^{\,k} \dfrac{n!}{(n-k)!}$ (15)

Hence, $P_0 = \dfrac{1}{\sum_{k=0}^{n} \rho_n^{\,k} \dfrac{n!}{(n-k)!}}$        (16)

$\text{avail}_{\text{II}} = 1 - P_n = 1 - \dfrac{\rho^n n!}{\sum_{k=0}^{n} \rho_n^{\,k} \dfrac{n!}{(n-k)!}}$    (17)

## 4.7 Availability analysis on performance

Case 3:

If $\text{avail}_{\text{II}}$ is low, then we need to speed up the process of repair rate facility by $n_\mu$, We get,

$$\text{avail}_{\text{III}} = 1 - \frac{\left(\dfrac{\lambda_n}{\mu_n}\right)^n n!}{\sum_{k=0}^{n} \left(\dfrac{\lambda_n}{\mu_n}\right)^n \dfrac{n!}{(n-k)!}} \quad (18)$$

## 5. Modelling and Analysis

The procedure can be automated by utilising the following three states of the system,

1. Check for the functionality of all the components of the cloud system
2. Check the component located in detection and location phase
3. Check the component is ready to repair the failed device.

Case 1:

The steady state probabilities are,

$$p_n = \frac{1}{1 + \dfrac{\lambda}{\mu_1} + \dfrac{\lambda}{\mu_2}, ..., \dfrac{\lambda_n}{\mu_n}} \quad (19)$$

The steady state probability assigned as S,

$$S = p_n = \frac{1}{1 + \lambda_n\left(\dfrac{1}{\mu_1} + \dfrac{1}{\mu_2}, ..., \dfrac{1}{\mu_n}\right)} \quad (20)$$

Case 2:

By developing the above condition for k-stage hypo exponential repair time dissemination with the parameters like $\mu_1, \mu_2, ..., \mu_k$.

$$\text{Now, S} = \frac{1}{1 + \lambda_n\left(\dfrac{1}{\mu_1} + \dfrac{1}{\mu_2}, ..., \dfrac{1}{\mu_k}\right)} \quad (21)$$

Case 3:

Occasionally, it is hard to recognise the fundamental reason for the failure,

Let us assume that $E_H$, $E_S$, $E_N$ and so on as a disjoint or mutually exclusive events, a set

$$E = (E_H \cup E_S \cup E_N) \quad (22)$$

Where $E_H$, $E_S$, $E_N$ are the events causing the hardware, software and network failures.
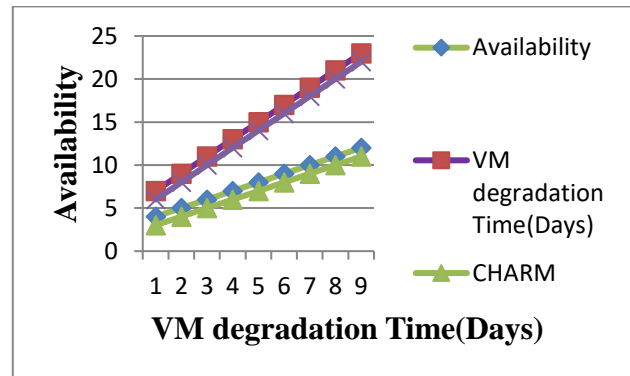


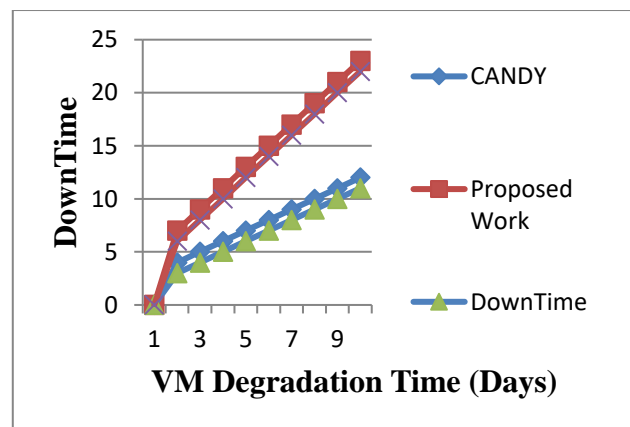Figure. 1 Availability vs. Different VM Degradation Rate



Figure. 2 Downtime vs. Different VM Degradation Rate

Figure 1 illustrates the availability changes for the proposed model with 3 physical hosts system. The

influence of VM degradation rates and rejuvenation trigger rates on availability of the CHARM technique is shown. The rejuvenation transition firing rates $\lambda_n$ are assumed 1 time/3 days and 1 time/4 days. It can be observed that the rejuvenation trigger rate increases for VM, the higher availability can be achieved in the proposed model.

Figure 2 plotted the downtime as a function of the VM degradation rates and rejuvenation trigger rates. For the system with higher VM degradation rate, it can be shown that the rejuvenation trigger rate increase for VM, the lower downtime can be achieved and compared with the CANDY method. The differences in downtime with different VM degradation time and different rejuvenation time are shown in Figure 2 from the result, our proposed model is apparent that the quicker rejuvenation time for VM can enhance the availability and reduce the downtime.

We analyse the availability on different physical hosts as a function of different migration rates. The change in the availability of system with the different numbers of physical hosts and different migration rates is plotted in Fig 3. The more physical hosts in the system by using RBD technique, the more chance the operational VM to be migrated. We also observe that the amount of availability increment depends on migration rates.

Our proposed method achieve faster migration rate for the VM can enhance the availability. When we compare our proposed technique with the existing techniques like CHARM and RBD our method attains the QOS metrics like cost, energy, and downtime very prominently.
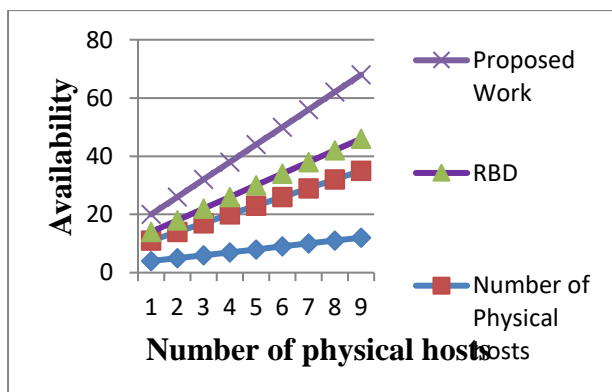


Figure .3 Availability vs. Number of physical hosts

## Influence of Cloud Computing on Healthcare:

Access to medical services shifts crosswise over nations, gatherings, and people, generally affected by social and monetary conditions. Nations and jurisdictions have distinctive approaches and plans in connection to the individual and population based human services objectives inside their social orders. Healthcare systems are associations built up to meet the wellbeing needs of target populations. Their current setup changes amongst national and subnational substances. In few nations and jurisdictions, human services arranging is disseminated among market members. Though in others, arranging happens all the more halfway among governments or other planning bodies. Patient centricity turned into the key pattern in human services provisioning like electronic medicinal records (EMR), electronic health records (EHR), personal health records (PHR).

Markov simulation is another random procedure displaying device that frequently is utilized for monetary assessments when looking at changed results of complex medical intercessions. It has been utilized to explore ICU clinical basic leadership by uncovering proof for sex-based hazard distinction in ICU patients [32, 33]. The element time is clearly connected with the probability of a patient taking certain states in a progression of discrete timeframes. In Markov models, these timeframes are called cycles. In other words, a disease is divided in distinct cycles, and probabilities are attributed to the transition between these states. At the point, when patient data is put away in the cloud, medical suppliers can get to lab and radiology results. In addition some other related test outcomes during whichever time outline and at any position also collected. Cloud computing offers data which is available to the entire suppliers and the clients. So that the Emergency unit (ER), intensive care unit (ICU), different therapeutic nursing units, auxiliary divisions, for example, lab and radiology, and different offices. For example, a joined nursing home, for instance, has entry to similar sort and measure of data.

In medical informatics, cloud resources can be arranged into compute cloud, storing cloud, and information cloud. Cloud computing comprises of resources like hardware, software, central processing unit (CPU), and operating framework. Besides, human services information should be shared across different settings and geographies which promote load the medical services supplier and the patient bringing about significant delay in treatment and loss of time. Cloud takes into account every one of these necessities along these lines

giving the social insurance associations. An unbelievable chance to enhance administrations to their clients, the patients, to share data more effectively than any other time in recent memory, and enhance operational proficiency in the meantime.

The above analysis ascertains the frequently failed components furthermore, the reason for component repair. This investigation is useful in ascertaining the cost spent in repairing hardware and repair time, also approximating accessibility by figuring the related downtime and so forth. The measures making resourcefailures are perceived, and the methodology for measuring the risk of these activities is displayed. The examination focused on the quantifiable belonging of the let-down information, including the cause of failure, average time spent waiting for repair, and the time required to repair a system.

Because of their high convolution, organization segments are liable to an expansive number of failures that may keep the system from satisfying its proposed usefulness. Cases of key perceptions from these studies are as per the following:

The yearly failure rate for servers is around 8%. The normal number of repairs is 2 for each machine. (e.g., 20 repair or substitution occasions in 9 machines were recognized over a 14-month time span.). Hard disks are the most failure inclined hardware segments and the most noteworthy purpose for server failures (around 78% of aggregate deficiencies or substitutions influenced hard disks).

Our proposed strategy separates the framework for protecting cloud-based applications and organizations from every one of the levels of disappointments and distinctive estimations to scale and evaluate accessibility. In our proposed work, we have demonstrated that at whatever point the failure rate is builds an ideal opportunity to repair is less when it is contrasted and alternate frameworks. Distinctive healthcare services associations and divisions will have their individual differing sensitivities with regards to the uptime availability, system responsiveness, latency, and versatility prerequisites for their healthcare services IT applications and work process.
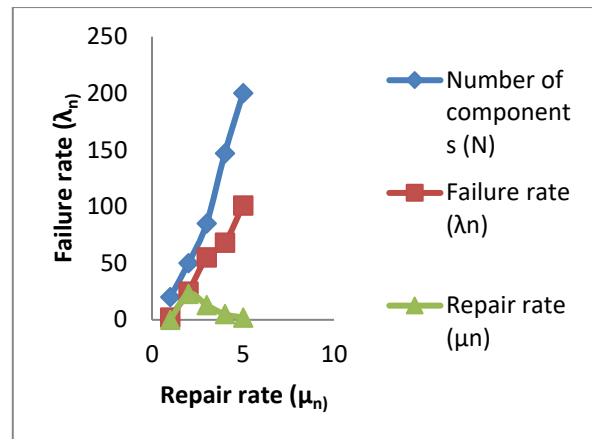


Figure.4 the comparison outcomes of failure rate

Vs. Repair rate

## 6. Comparative Analysis

The simulation of the availability modelling is done using SHARPE tool. The availability analysis of analytical model is investigated. We have investigated the proposed model, using SHARPE tool, and it is obvious that the proposed availability model with Markov chain process model predicts the steady state availability of the virtual machines. The simulation of this model is investigated to envision the Quality of Service metrics such as cost, failure rate, repair rate and utilisation of the system. The proposed model is compared with conventionally proven techniques like OSCAR (Open Source Cluster Application Resources), and SRN (Stochastic Reward Nets).

Figure 4 illustrate the availability model efficiently decreases the repair time of the virtual machine compared to the OSCAR and SRN techniques. The proposed availability modelling attains less repair time for more number of failure rates compared to OSCAR and SRN. Our proposed technique has derived a prominent result. Our main aim to attain 100% steady state availability of the system. We have analysed the QOS metrics such as cost, energy, repair rate, failure rate with the already existing system bur still some of the proposed techniques failed to satisfy all the QOS metrics.

Our proposed approach stabilises the availability to reduce the repair rate accordingly. Thus, the comparison result proves that our fault tolerant system is more efficient and performs well, when the number of failures in the virtual machine increases. Table2. Describes the comparison of proposed technique with the already existing skills.

Table 2. Comparison of proposed technique with the existing techniques

| Work carried out | Tools used | Analytical model/ Numerical model | Quality of service factors | Metrics of QOS achieved |
|---|---|---|---|---|
| Availability modelling of fault tolerant system using Markov chain model(proposed work). | SHARPE | Analytical model | Steady state availability | Cost, energy, failure rate, repair rate, degradation rate, rejuvenation rate. |
| CHARM A Cost-Efficient Multi-Cloud Data hosting Scheme. | Apache Libcloud | Analytical model | Guaranteed availability | Cost factor, Degradation Rate |
| Eucalyptus-based private clouds: availability modelling and comparison to the cost of a public cloud | Eucalyptus platform | Analytical model | Capacity oriented availability and system steady state availability. | Down time and the acquisition cost |
| Availability Modelling and Analysis for Data Backup and Restore Operations | CASSI (Computer Aided System model-based System Integration environment | Analytical model | Storage availability, system availability, and user-perceived availability | Cost factor |

## 7. Conclusion

Various experts and cloud suppliers discourse the availability of the cloud and suggest elucidations taking account into their particular meaning of availability. These definitions impact the projected methods. Our proposed method differentiate the system for shielding cloud-based applications and is recognised, and the strategy for measuring the threat of these actions is exhibited. The investigation concentrated on the measurable possessions of the let-down data, including the origin of failure, average time spent waiting for repair, and the time required to repair a system. The best model for average time excludes the time spent waiting for repair is the CTMC, with reducing risk capacity rate.

Overhaul periods are greatly improved and demonstrated by a lognormal appropriation than an exponential conveyance. We verified our test results with the assessment results through SHARPE. By using the same method, we can identify the fault in hardware, software and network as a future work.

## References

[1] M. Nabi, M.Toeroe, F. Khendek, "Availability in the cloud: The state of art", *Journal of Network and Computer Applications*, No. 6, pp. 54- 67, 2016.
[2] Z. Zhang, Y. Wang, H. Chen M. Kim, J.M. Xu, H. Lei, "A cloud queuing service with strong consistency and high availability", *IBM Journal*

*of Research and Development*, Vol. 55, No. 6, pp. 10:1- 10:12, 2011.

[3] G. Radhakrishnan, "Adaptive application scaling for improving fault- tolerance and availability in the cloud", *Bell Labs Technical Journal*, pp. 5-14, 2012.

[4] A. Celesti, M. Fazio, M. Villari, A. Puliafito, "Adding long term- availability, obfuscation, and encryption to multicloud storage systems", *Journal of network and computer Applications*, vol. 59, pp. 208- 218, 2016.

[5] X. Yin, J. Alonso, F. Machida, "Availability modelling and analysis for data backup and restore operations", *31$^{st}$ International Symposium on Reliable Distributed systems*, pp. 208-218, 2012.

[6] N.T. Anh, D.S. Kim, and J.S. Park, "Availability modelling and analysis of a data center for disaster tolerance", *Future Generation Computer System*s, Vol. 56, pp. 27- 50, 2016.

[7] Q. Zhang, S. Li, Z. Li, Y. Xing, Z. Yang, and Y. Dai, "CHARM: A Cost- Efficient Multi- Cloud Data Hosting Scheme with High Availability", *IEEE Transactions on cloud computing*, Vol. 3, No. 3, pp. 372- 386, 2015.

[8] J. Dantas, R. Matos, J. Araujo, P. Maciel, "Eucalyptus-based private clouds: Availability modelling and comparison to the cost of a public cloud", *Springer Verlag*, Vol. 97, pp. 1121- 1140, 2015.

[9] U. Franke, M. Buschle, " Experimental Evidence on Decision – Making in availability service level agreements", *IEEE Transactions on network and service management*, Vol. 13, No. 1, pp. 58- 70, March 2016.

[10] Q. Zhang, L. Cheng, R. Boutaba. "Cloud computing: state-of-the-art and research challenges", *J Internet Serv Appl*, vol.1, pp. 7–18, 2010.

[11] S. Distefano, F. Longob, K.S. Trivedi, "Investigating dynamic reliability and availability through state–space Models", *Computers and Mathematics with Application,* vol. 64, pp. 3701–3716, 2012.

[12] E. Hyytia, S. Bayhanb, J. Otta, J. Kangasharjub, "On search and content availability in opportunistic networks", *Computer Communications,* Vol. 73, pp. 118–131, 2016.

[13] R.D. Matos, J. P. Maciel, F. Machida, D.S. Kim, K.S. Trivedi, "Sensitivity Analysis of Server Virtualised System Availability", *IEEE Transactions on Reliability*, Vol. 61, No. 4, pp. 994- 1006, 2012.

[14] R. Ghosh, F. Longo, V.K. Naik, K.S. Trivedi, "Modelling and performance analysis of large-scale IaaS Clouds", *Future Generation Computer Systems*, Vol. 29,pp. 1216-1234, 2013.

[15] F. Machida, E. Andrade, D.S. Kim, K. S.Trivedi, "Candy: Component- based Availability Modelling Framework for Cloud Service Management using SysML", *30$^{th}$ IEEE International Symposium on Reliable Distributed Systems*, Vol. 47, pp. 209-218, 2011.

[16] R. Ghosh, F. Longo, F. Frattini, S. Russo, K.S.Trivedi, " Scalable Analytics for IaaS Cloud Availability", *IEEE Transactions on cloud computing*, Vol. 2, No.1, pp. 57- 70, 2014.

[17] M. Unuvar, S. Tosi, Y.N. Doganata, M. Steinder, A.N. Tantawi, " Selecting optimum cloud availability zones by learning user satisfaction levels", *IEEE Transactions on services computing*, Vol. 8, No.2, 2015.

[18] M.L. Yin,John , E. Angus, K.S. Trivedi, " Optimal preventive maintenance rate for best availability with Hypo- Exponential failure distribution", *IEEE Transactions on Reliability*, Vol. 62, No. 2, pp. 351- 361, 2013.

[19] I.A. TargioHashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S. Ullahkhan, " The rise of big data on cloud computing: Review and open research issues", *Information systems*, Vol. 47, pp. 98- 115, 2015.

[20] M. Ali, S.U. Khan, A.V. Vasilakos, "Security in cloud computing: Opportunities and challenges", *Information Sciences*, Vol. 305, pp. 357- 383, 2015.

[21] M. Ramachandran, V. Chang, "Towards Validating Cloud service providers Using Business Process Modelling and Simulation", *International Journal of Information*, Vol. 36, No. 6, pp. 618- 625, 2016.

[22] T. Thenin , J. SouPaark , "Availability analysis of application server using software rejuvenation and virtualization", *J compSci Technol,* Vol. 24, pp. 339- 346, 2009.

[23] J. Jassen, R. Mance, " Semi Markov Risk model Finance, Insurance, and reliability", *Springer* , pp. 1-41, 2007.

[24] M. Armbrust , A. Fox, R. Griffith , A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, " A view of cloud computing", *ACM* , Vol. 52, No. 4, pp. 50-58, 2010.

[25] D. Bhagwat, K. Pollack, D.D. Lon, T. Schwarz, E.L. Miller, J.F. Paris, "Providing high

reliability in a minimum redundancy archival storage system", *In: proceedings of the 14th IEEE International Symposium on Modeling, analysis, and simulation, MASCOTS'06. Washington, DCC, USA: IEEE Computer Soceity*, pp. 413- 442, 2006.

[26] M. Toeroa, F. Tam, "Service availability principles practice", *John Wiley and Sons Ltd publication*, pp. 476- 486, 2015.

[27] D. Boru, J. Dejene, "Energy- efficient data replication in cloud computing data centers", *Cluster Computing*, Vol. 18, pp. 385- 402, 2015.

[28] Z. Hong, Y. Wang, M. Shi, "Ctmc- based availability analysis of cluster system with multiple nodes", *In Advances in future computer and control systems, Springer*, pp. 121- 125, 2012.

[29] K.S. Trivedi , " Probability, and Statistics with Reliability, Queuing and Computer Science Applications", second ed., *Wiley*, 2001.

[30] S. Nanda, T. Chiueh, "A survey on virtualization technologies", *Stony Brook University, Tech. Rep. TR*, Vol. 179, pp. 1-42, 2005.

[31] M. Afshari, A. Ghaffaripour, "Modelling of Imperfect Data in Medical Sciences by Markov chain with Numerical Computation", *Advances in Bioscience and Biotechnology,* Vol. 5, pp. 1003- 1008, 2014.

[32] A.S.M. Mosa , L. sheets, "A systematic review of Healthcare application for smart phones", *Medical informatics and decision making*, Vol. 12, pp.1-31, 2012.

[33] N. Rolim, C. Oberdan, "A cloud computing solution for patients data collection in health care institutions", *eHEalth, Telemedicine, and Social Medicine,* Second International conference on IEEE, Vol. 49, pp. 95- 99, 2010.