# Improving Reliability in Cloud Computing Systems

Suma V
Dean
Research and Industry Incubation Centre (RIIC)
Department of Information Science and Engineering
Dayananda Sagar Institutions (DSI)
Bangalore, India

Chintureena Thingom
Post Graduate Programme
Computer Science and Engineering
Department of Information Science and Engineering
Dayananda Sagar Institutions (DSI)
Bangalore, India

## ABSTRACT
Cloud computing is one of today's most exciting technologies because of its capacity to reduce cost associated with computing while increasing flexibility and scalability for computer processes. During the past few years, cloud computing has grown from being a promising business idea to one of the fastest growing sectors of the IT industry. But there is a prime essence that IT organizations have expressed concerns about critical issues such as Security that accompany the widespread implementation of cloud computing. Reliability is one of the important factors for cloud computing resource for maintaining higher user satisfaction and business continuity. This paper brings out various challenges of cloud computing and some of influences factors to overcome disaster in cloud. Further, this paper analyzes the existing reliability assurance algorithm and proposes an enhanced algorithm using proactive filtering based redundancy scheduling technique. Enhanced Proactive Filtering based Redundancy Scheduling algorithm further reduces the impact of disaster in cloud computing due to reliability factors and hence improves business tactics.

## Index Words
Cloud Computing, IT Industry, Reliability, Proactive Filtering, Disaster Management

## 1. INTRODUCTION
The main value proposition of Cloud Computing is to provide the clients a cost-effective, convenient means to consume the amount of IT resources that is actually needed for the service provider and effective resource utilization of existing infrastructure can be achieved through a multi-tenant architecture [1]. From a business perspective, Cloud Computing is about improving organizational efficiency and reducing cost, often coupled with the objective of achieving a faster time-to-market. Centrally hosted services with self-service interfaces can help to reduce lead times between organizational units who use the cloud as a collaborative IT environment [2]. Low-cost access to data centres in different geographical regions may further reduce market entry barriers and enable new business models. From technology and engineering perspective, Cloud Computing can help to realize or improve scalability, availability, and other non-functional properties of application architectures. In this paper, we focus on the technology perspective, and in particular on challenges and opportunities of Cloud Computing research related to quality-driven architectures. It is worth to note that in cloud computing, the user tasks are run in a machine in a data centre , failure of machine should not affect the user in any way and the task must be get completed in the planned execution without affecting the user experience. Only in this case it can be said that the cloud computing system in reliable. However, there are algorithms to manage disaster through reliability and availability. This paper, therefore analyzes one of the popular reliability algorithm namely redundancy scheduling and further brings out an enhancement to this existing algorithm using proactive filtering based redundancy scheduling technique which in turn reduces level of disaster during cloud computing.

The structure of this paper are organized as follows, section I introduces the significance of Availability and Reliability in Cloud environment, in section II the Literature Survey is presented, in section III we define the problem identified, section IV presents the Research work contributed and in section V the Result analysis is done by comparing the two scheduling algorithms, finally section VI presents the Conclusion.

### 1.1 Benefits reaped due to Cloud computing
Cloud computing has put forward many vital benefits which includes optimized server utilisation, cost saving, Dynamic scalability, Reduced time for implementation, etc. As most organization typically underutilise their server, cloud resources and cloud computing will manage the server effective utilisation to the optimum level. IT infrastructure costs are always high and in many countries, it can lead for a tax advantage regarding income taxes. Also, cost can be saved during cloud computing using resource pooling approach. Many firms include an extensive computing requirement to make sure that capacity is in place to meet uncertain peak demand. Cloud computing provides low cost processing buffer as needed and without much capital investment or contingency fees to users. Cloud computing provides the processing power and data storage as needed at the capacity required. It can be mentioned that this can be obtained in real time batch processing instead of weeks or months that occur when a new business initiative is brought online in a traditional way.

Cloud computing has extremely appreciable benefit and it is one of best technologies which can reduce cost associated with computing along while best flexibility and scalability. However, despite the effective benefit of cloud computing still there prevails various challenges.

### 1.2 Challenges of Cloud Computing
- Data location - cloud computing technology allows cloud servers to locate anywhere across the globe, thus the organization need not know the physical location of the server used to store and process their data and applications. Despite the technology, location of the server is least relevant; this has become a critical issue for data governance requirements. It is essential to understand that many Cloud Service Providers (CSPs) can also specifically define where data is to be located.

- Commingled data - application sharing and multi-tenancy of data is one of the characteristics associated with cloud computing. Data encryption is another control that can assist data confidentiality.
- Cloud security policy / procedures transparency - some CSPs may have less transparency than others about their information security policy. As a result, it may create conflict with the enterprise's compliance of information requirement. The firm needs to have detailed layout for the effective understanding of the service level agreements (SLAs).
- Cloud date ownership - in the contract agreements it may state that the CP owns the data stored in the cloud computing environment. The Cloud Service Providers are likely to demand for significant cost for the data to be returned to the enterprise when the cloud computing SLAs terminates.
- Lock-in with CSP's proprietary application programming interfaces (APIs) - currently many Cloud Service Providers implement their application by adopting the interface. As in turn it results in cloud services transition from one service provider to another service provider of Cloud, which has become very vital and complicated, intensive requirement of labour and time-taking.
- Compliance requirements - today's cloud computing services does challenge many important compliance audit issues which are currently in place. To mention a few, location of the data inputs, security of cloud computing policy are all major issues in compliance auditing efforts.
- Disaster recovery - it is a concern of firms about the effective utilisation of cloud computing, since data may be complicated and scattered through many servers across various geographical areas. It may be possible that the data for a specific point of time cannot be identified. But in this current era of dynamic server hosting, almost all the firms knows exactly where the location is of their data, to be rapidly retrieved in the event of disaster recovery.

In addition to the above challenges of cloud computing, one of the most important challenges is the disaster management for cloud computing which has to be addressed to understand and determine the solution for cloud disaster management.

## 1.3 Disaster Management

Cloud is an environment where tasks pop up constantly; these tasks are generally of different size demanding for different types of resources for computation. When the computer system fails to provide service to its clients, the system collapses and leads to Disaster in Cloud. Various factors that would prevent Disaster in Cloud are Availability and Reliability.

## 1.4 Availability and Reliability in Cloud

Cloud computing is based on virtualization technology, in which each user uses a virtual machine. Virtualization technology includes two levels of virtual machines, which are VMs (virtual machine) and hypervisors. The hypervisor has administrative rights to control VMs. But virtualization has some issues that could endanger system performance. From a cloud viewpoint, there are many important dimensions of virtualization technology to consider, but the hypervisor's Reliability, Availability and Serviceability (RAS) is an important aspect of virtualization technology and requires special attention. One such problem involves overflows of system due to excessive combination of VM to a physical server that affects availability and reliability. Because of these issues, cloud systems are vulnerable to traditional attacks as well as new attacks that some of them have migrated from virtualization. Privacy is another issue which can decrease virtualization and cloud's overall performance, because the VMs are located practical in a multitenant environment, thus making it possible for a user to access a past tenant's information in the same space. This can make the cloud vulnerable to some attacks, like the DoS family (Denial of Service) which aims to make the target server inaccessible to legitimate users [3]. Therefore, applying countermeasures to deal with security problems in the cloud is critical, whereas one of the main countermeasures is controlling access control in the cloud. Besides security, cloud providers are also responsible for reliability and availability, because all users expect the highest level of QoS (Quality of Service). The cloud providers use some solutions such as partitioning to achieve maximum performance. From investigations, a system availability and reliability can be enhanced through efficient scheduling techniques

# 2. LITERATURE SURVEY

Research has progressed ever since the inception of cloud technology in IT industries. Disaster management related research too has progressed in order to minimize Disaster in Cloud. Some of this research includes techniques that focus on aspects such as load balancing and Scheduling to enhance the resource availability to minimize Disaster in Cloud.

Mousumi Paul et al, have proposed a scheduling mechanism which follows the Lexi –search approach to find an optimal feasible assignment. Task scheduling has been treated as general assignment problem to find the minimal cost. Here, the cost matrix is generated based on some most vital condition of efficient task scheduling such as arrival rate, waiting time and the most important task processing time in a resource [4]. Gang Yao et al, have done analysis on pre-emptive and non pre-emptive scheduling techniques, they say both approaches have advantages and disadvantages, and no one dominates the other when both predictability and efficiency have to be taken into account in the system design. In particular, limiting pre-emption's allows timing analysis more predictable with respect to the fully pre-emptive case. Three methods (with different complexity and performance) are presented to calculate the longest non-pre-emptive interval that can be executed by each task, under fixed condition, without reducing the schedules of the task set, with respect to the fully pre-emptive case. The methods are also compared by simulations to evaluate their effectiveness in reducing the number of pre-emption's [5]. Mahitha O, et al, have focused on for disaster management through efficient scheduling mechanism. This work presents a Priority Pre-emptive scheduling (PPS) with aging of the low priority jobs in Cloud for disaster management. The implementation results proved that the jobs at any event of time are provided with the resources and henceforth preventing them to enter the starvation, which is the most important factor for cloud computing disaster [6]. Vignesh V, et. al, the process of execution of the cloud computing requires Resource Management due to the high process to the resource ratio. Resource Scheduling is a complex task in cloud computing environment because there are many alternative computers with varying capacities. The goal of their work is to propose a

model for job-oriented resource scheduling in a cloud computing environment. Resource allocation task is scheduled for the Process which gives the available resources and user preferences. The computing resources can be allocated according to the rank of job .Their work constructs the analysis of resource scheduling algorithms. The time parameters of three algorithms, viz. Round Robin, Pre-emptive Priority and Shortest Remaining Time First have been taken into consideration. From this, they have computed that SRTF has the lowest time parameters in all respects and is the most efficient algorithm for resource scheduling [7].

# 3. PROBLEM DEFINITION

The cloud model performance is influenced by availability and reliability of the system, which could be achieved through efficient scheduling and resource allocation techniques. Insufficient scheduling in the cloud reduces the effective performance of the cloud computing. Resource Scheduling is a complex task in cloud computing environment because there are many alternative computers with varying capacities. The goal of their work is to focus on the techniques to ensure reliability for the user task using effective scheduling methods.

# 4. RESEARCH WORK

Reliability means the ability of the system to return the expected results to the users in spite of machine failure. Machines can fail due to lot of reasons either hardware or software problems. Once the failure happens, the user should not get affected by the failures in the system. A good reliable system will abstract the failures from the user as much as possible with little or no delay in user's task execution time.

## 4.1 Reliability Scheduling Algorithms

### *4.1.1 Redundancy Scheduling (RS)*

In this method the user task is executed in two different machines in the data center. Both the machines run in hot standby mode. Both run the user task in same time but only the hot machine will respond to the user. In case the hot machine failed, the standby machine assumes the role of hot node and responds to the user. Due to this user observer no changes and the cloud is reliable from the user perspective. This is explained in terms of an algorithm mentioned in section 4.2

**Algorithm**

The pseudo code of redundancy based scheduling is given below:

The scheduler algorithm will wait for user request and the user request is processed as follows.

```
While (shutdown)
{
    ReqWait for User Request( );
   hostList  sort host List based on decreasing order of load();
   Primary hostList(1);
   BackUphostList(2);
   RunTask(Req,Primary,Hot);
RunTask(Req,Backup,standby);
}
```

### *4.1.2 Enhance Redundancy scheduling using Proactive Filtering based Redundancy Scheduling (PFRS)*

In the data center not all machines fail in the same way. Some machine has high probability of failure than other machines. The reasons can anything from aging to a presence of malicious code in the machine which cause it to fail often. In this algorithm we will calculate the reliability of each host in the datacenter in form of score. The score is calculated based on number of failures occurring in duration of time, duration can be any value tuned by the cloud administrator. For our work we will take it as 1 hour. The number of failures of the system in one hour duration is taken and assigned as score.

The failure score is calculated in form of weighted average means

FS = A*FS + (1-A)*num_of_time_failed

A is constant from 0 to 1.

When machines fail more often their score is a very high value. A very high value indicates the machine is not reliable.

While choosing the primary & backup machines for running the task , we always choose the machines with least FS score , by this way we ensure that the chances for switchover is avoided totally improving the user experience and user never see a switch over delay which occurs in case of previous algorithm.

**Algorithm:**

The pseudo code of the algorithm is given below:

```
While (shutdown)
{
    ReqWait for User Request( );
   for host in hostList
       If FS(host)<LowThres
          filterhostList.add(host);
       end
   end
   filterhostList  sort filterhostList  based on decreasing order of load();
   Primary filterhostList 1);
   BackUp filterhostList (2);
   RunTask(Req,Primary,Hot);
RunTask(Req,Backup,standby);
}
```

# 5. RESULT ANALYSIS

We conducted the simulation of both the algorithms on cloud sim for different number of user request per second and measured the switch over delay between the two algorithms. From this we find that switch over hardly happens in case of Proactive Filtering based redundancy scheduling. Failure for host is introduced as poison random variable.

**Table.1. Performance of User Task Vs Switch Delays**

| No of User Task | Average Switch delay in RS | Average Switch delay in PFS |
|---|---|---|
| **50** | **100 ms** | **0ms** |
| **100** | **200 ms** | **10ms** |
| **150** | **250 ms** | **12 ms** |
| **200** | **300 ms** | **12ms** |

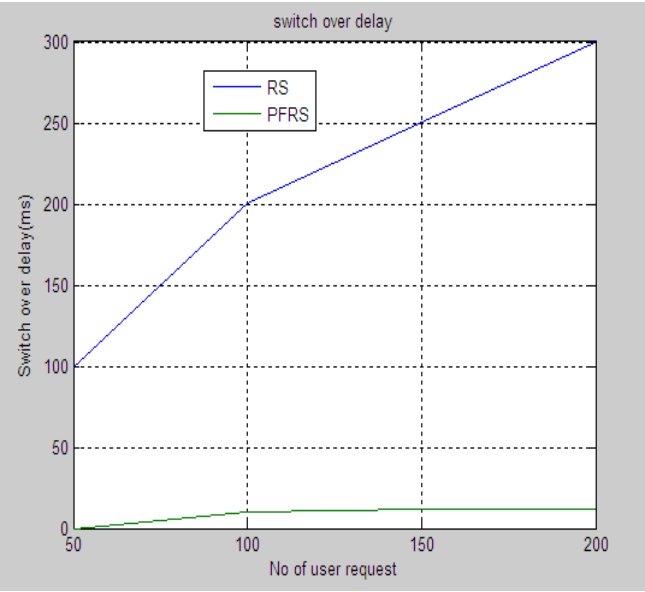


**Fig.1. Graphical Representation of the Performance**

From this above graph show that the average switch delay is very less in PFS because, we always choose highly reliable nodes to execute the task in primary hot mode, so switchover hardly happens.

## 6. CONCLUSION

Cloud computing has become today's one of the highly business accelerating domain in IT field. The main crux of Cloud Computing is to enhance a system which can provide the clients a cost-effective, convenient means to consume the amount of IT resources that is actually needed for the service provider, better resource utilization of existing infrastructure is achieved through a multi-tenant architecture. As one of the most influencing factor for the barrier in the performance of cloud computing is the Disaster. The cloud model performance is influenced by availability and reliability of the system, which could be achieved through efficient scheduling and resource allocation techniques. In this work, we focus on the techniques to ensure reliability for the user task using effective scheduling methods. An efficient reliable model is tremendously required to avoid the Disaster in the cloud. To do so, this paper shows the ways of improving the Cloud performance by having an effective redundancy based scheduling. We have enhanced the redundancy scheduling by reducing the average switch delay by filtering nodes based on failure scores which will lead to reduction in the disaster and can effectively manage the performance of cloud computing.

## 7. REFERENCES

[1] Andrea Bianco, Luca Giraudo, David Hay, "Optimal Resource Allocation for Disaster Recovery", Global Telecommunications Conference, 2010, pp.1-5

[2] Jorge Finochietto, Luca Giraudo, Marco Modesti, Fabio Neri ,"Network Planning for Disaster Recovery", proceedings of the 2008, 16th IEEE workshop on Local and Metropolitan Area Networks, 2008, pp.43-48

[3] Manish Pokharel, Seulki Lee, Jong Sou Park, "Disaster Recovery for System Architecture using Cloud Computing", 2010 10th Annual International Symposium on Applications and the Internet, 2010, pp.304-307.

[4] Mousumi Paul, Debabrata Samanta, Goutam Sanyal, "Dynamic job Scheduling in Cloud Computing based on horizontal load balancing", 13th International Journal of Computer Technology and Applications, Oct 2011

[5] Gang Yao, Giorgio Buttazzo and Marko Bertogna, "Bounding the Maximum Length of Non-Preemptive Regions Under Fixed Priority Scheduling", Embedded and Real-Time Computing Systems and Applications, 2009. RTCSA '09. 15th IEEE International Conference on 24-26 Aug. 2009

[6] Mahitha.O, Suma.V, "Pre-emptive Priority Scheduling with Aging Technique for Effective Disaster Management in Cloud", Graduate Research in Engineering and Technology (GRET): An International Journal , ISSN 2320 – 6632, volume 1, Issue-2, 2013.

[7] Vignesh V, Sendhil Kumar KS, Jaisankar N, "Resource management and scheduling in cloud environment", International Journal of Scientific and Research Publications, Volume 3, Issue 6, June 2013 1ISSN 2250-3153.

[8] Manish Pokharel, Seulki Lee, Jong Sou Park, ―Disaster Recovery for System Architecture using Cloud Computing‖, 2010 10th Annual International Symposium on Applications and the Internet, 2010, pp.304-307.

[9] Mahitha.O, Suma.V, Vaidehi.M, "Disaster Recovery in the Cloud Environment – An Analysis", International Conference on Advanced Computer Science and Information Technology (ICACSIT), 10th March 2013, Pune, India.

[10] Mahitha.O, Suma.V, Vaidehi.M, "A Comparative Study of Tools for Disaster Management in Cloud", 2nd International Conference on Recent Trends in Engineering Technology(ICRTET), 30th April, 2013, Bangalore, India.

[11] Chintureena, Suma V, "Ensured Availability of resources in a highly reliable mode through Enhanced approaches for Effective Disaster Management in Cloud", International Conference on Electronics and Communication System (ICECS), IEEE, 13th – 14th Feb 2014, Coimbatore, India.

[12] Chintureena, Suma V, "Effective Scheduling Approach to Enhance Availability of Resources to Manage Disaster In Cloud", 10th International Conference on Systemics, Cybernetics and Informatics (ICSCI), 3rd – 6th Feb 2014, Hyderabad, India.