iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 6 **Issue:** XII **Month of publication:** December 2018

**DOI:**

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887*
*Volume 6 Issue XII, Dec 2018- Available at www.ijraset.com*

# Generating Fault Tolerance Mechanism in Cloud Computing System using EDP

B. Anand Kumar[1], A.Vinayababu[2], V. Malliah[3], K. Madhukar[4]
[1]*Research Scholar, Department of Computer Science, Rayalaseema University, Kurnool.*
[2]*Head of Department (IT), Stanley College of Engineering and technology, Hyderabad.*
[3]*Research Scholar, Department of Computer Science, Nagarjuna University, Guntur.*
[4]*Physics department, Osmania University, Hyderabad.*

***Abstract:*** ***Cloud computing is the technology that enables sharing resources to the end users. This technology integrates hardware, software and services. As a result, the dynamic changes in sharing the resources lead to unexpected faults at various service models of the cloud and makes services unavailable. In this paper we proposed a new algorithm that is known as EDP (Error Detection Program). The EDP algorithm proactively identifies faults, when reboot or update is required. To resolve the faults, EDP uses the rejuvenation mechanism of proactive fault tolerance technique. If errors still exist, it uses task resubmission mechanism of reactive fault tolerance technique.***

***Keywords: Cloud computing, SLA, Resource scheduling, Kernel, Zombie***

## I. INTRODUCTION

Cloud computing is a technology which provides services over the internet by using computational resources such as storage, operating system and middleware to end user [1].

This technology cuts the operating cost, increases the availability of computing resources; provide the reliable services round the clock throughout the globe, and offers portable and platform independent services to all types of users. IBM, Amazon, Microsoft and Google apps are the major cloud services providers available in the cloud market. This Technology offers on-demand services which involves huge amount of incoming data.

The Grid computing is the predecessor technology to cloud commuting. In the Grid computing all the servers offering services will be interconnected all over the world. Cloud computing is a collection of both physical and virtual data centres, as a result high usage of cloud services involves increased number of errors and faults such as server overload, crash, network congestion, and server re-booting etc. These common errors and faults will become very complex to handle at the levels of the operating system and network.

More efficient fault tolerant techniques are needed to handle this kind of errors and faults. These techniques are necessary for both system and applications to guarantee the availability and reliability of the critical services. When a fault is identified, different parameters like complexity of environment, whether user is facing issue only for this server or also for other servers in the datacentre should be verified.

When a fault is found in server or datacentre, following two strategies should be verified.

*1)* Weather production server connected to Disaster recovery (DR) server
*2)* Make sure when prod server is down, Disaster recovery (DR) server will replace prod server

As per National Institute of Standards and Technology (NIST) cloud computing reference architecture, the cloud architecture contains five characteristics, four deployment models and three service models [2][3],

Five characteristics of cloud architecture are

*a)* *On Demand Self-Services:* Cloud service providers can provide cloud resources (processor, RAM, power and storage) to the customers any time without any human interaction.
*b)* *Broad Network:* In cloud computing all the resources can be accessed through internet using different devices.
*c)* *Resources Pooling:* Cloud services can share multiple resources to multiple virtual machines (VM) on same physical server for different users.
*d)* *Rapid Elasticity:* User or customer can easily increase their number or quantity of resources and decrease their number or quantity of resources as per requirement.
*e)* *Measured Services:* A cloud is a pay and use service and, a customer usage will be measured based on processor time, storage space, internet bandwidth etc.

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887*
*Volume 6 Issue XII, Dec 2018- Available at www.ijraset.com*

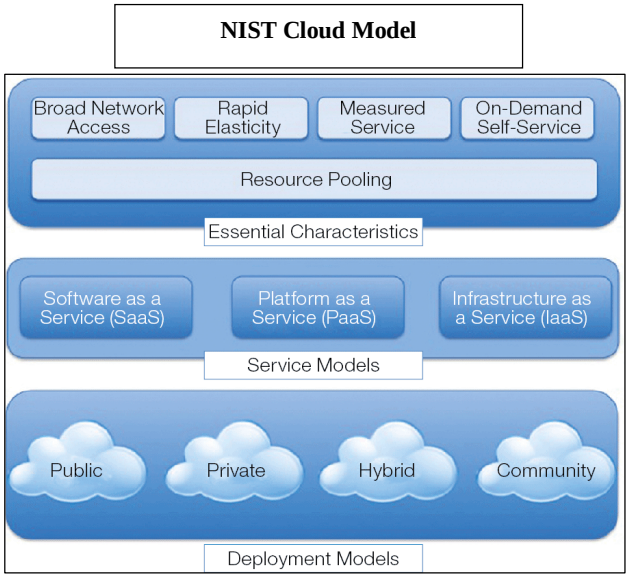



Fig.1 Cloud service models and deployed models as per NIST

*A. Cloud Deployment Models.*

There are four cloud deployment models available in the cloud computing market. These are offered by IBM, Amazon, Microsoft and other cloud service providers.

1) *Private Cloud:* In which all the services and infrastructure is accessed with in the organization. An organization can host their own private cloud internally or externally as per their requirements. [4] Private cloud is the best choice for those organizations whose maintain their own applications and databases for their customers. Once the private cloud is designed and implemented, it will not be much difference to public cloud. Following are some of the advantage of private cloud

*a)* More security
*b)* More accessible and reliable
*c)* Maximum usage of all resources
*d)* Designed for enterprise level and can be used as public cloud.
*e)* Greater control on infrastructure.

2) *Public Cloud*: In which all the services and infrastructure is accessed by the public over the internet. [4] Public cloud is less securable because of various threats that can be attacked through the internet. In public cloud, service provider is responsibility of IT maintenance and support. Public cloud is the best choice for start-up companies, because no need to invest on infrastructure and services. Service provider will provide all the resources on pay and use process. Following are some of the advantage of public cloud.

*a)* Low cost
*b)* Pay and use method
*c)* Reliable and security
*d)* Easy to adapt

3) *Hybrid:* Hybrid cloud is combination of both public and private cloud. In this cloud critical activities are performed by the private cloud and non-critical activities are performed by the public cloud. [4] In this cloud the service provider ensures availability of all the resources. Some of the advantages of hybrid cloud are more safety, scalability, reliability and performance. IBM is one of organization offering the best hybrid cloud services in cloud computing market.

4) *Community:* Community cloud is like a public cloud, in which systems and services are accessible only to a specific community. [4] This cloud can be maintained by on their own or by a service provider. In this cloud environment, there is a limited access to the cloud. It is not compulsory that entire members in the Community only access the IT resources, and outside members of Community can access the IT resources.

*B. Cloud Services Models*

Three cloud service models are

1) *Infrastructure-as-a-Service [IaaS]:* IaaS provides fundamental physical hardware resources like Physical machines, virtual machans, storage, virtual storage, networking etc. In IaaS service model, pre-installed and configured network, hardware resources along with operating system are provided to the end user. End user can install his own software as per his own requirements. Service provider provides complete access on virtual machine to the customer to perform some common operations such as start, reboot, installing applications, and granting permissions. Service provider also supports the previous infrastructure of the customer. [5]
2) *Platform-as-a-Service [PaaS]:* PaaS provides development tools and runtime environment applications. Software development or start-up companies will get benefited with this service model. PaaS will publish many services needed for many applications, and all these services are up-to-date. Service providers will test all these services and apps in different test environments so that it will get benefit to the customer. Customer need not to test their software in their environment which will save their money and valuable time [5].
3) *Software-as-a-Service [SaaS]:* SaaS model allows to customer or end user use applications as a service. In this model most commonly used licensed software such as SAAP, Sales Force, E-mail, and Office 365 are provided to the customers over the internet based on the type of subscription. SaaS provides all these features over internet by using browser. Customer may not know which operating system is used in the cloud background because it is only pay and use. [5]

## II. LITERATURE SURVEY

In the recent years various research work done in cloud computing. Most of the research studies were focused on cloud security and fault tolerance.

Youssef M. Essa, a Big Data Consultant, implemented survey paper on Fault tolerance techniques in cloud computing. This paper explores the best and low-cost techniques of cloud fault tolerance. The paper identified one of the best algorithms which distributes workload across multiple computers, computer clusters, network links, CPU, disk drives or other resources, to archive optimal resources utilization. They also discovered various techniques of maximizing throughput, minimizing response time, and avoiding overload. It concluded that every mechanism has its own issues and a new technique must be identified to overcome the issues. [4]

Virendra Singh Kushwaha, Sandip Kumar Goyal & Priusha Narwariya, paper investigated a method for fault-tolerance in load balancing schemes in the cloud environment. This paper mainly focused on various fault tolerance implementations in cloud computing during the load balancing and they concluded one of the best and low-cost technique in the cloud computing. [6]

Siddharth Shyamsunder1, Kunal Varkhede2, Shirish Bhuruk3, Rasika Vibhute4, Prof R.P. Karande5, proposed a service layer. This layer provides fault tolerance techniques on the servers. The responsibility of service layer is to detect the faults that occur at server's end and to accurately find a solution keeping the internal details hidden from the user. They used FTP mechanism for overall system using three blocks. 1) Client sends request to the server when it detected the fault. 2) FTP will control all the server requests subject to many processes to handle the fault. 3) Once the fault is properly handled then response sent to the server, which in turn sends the client's response. [7]

Mehdi Effatparva, Seyedeh Solmaz Madani, proposed a model in which, results are analysed by CPN tools and demonstrated the degree of reliability. This proposal concluded that, when there is an increase of requests, the FT is reduced, and same way reliability is also reduced and vice versa. The proposed method was implemented by using byzantine fault tolerance by using colored Petri-nets simulator in interconnected network clouds, which is used to evaluate the reliability. This method is used because it has strong mathematical support. Results which are obtained from simulator they are compared with optional resources and they and they found same method is used in both end. [8]

Mehdi Effatparva proposed model "System created based on FT-FC framework" Dr. Chandralekha1, Satish Kumar Suman2, Anupriya, proposed a model which has relationship between fault tolerance and Chandralekha,Satish Kumar proposed model Image developer matrices, to measure the system performance.

The proposed model implemented in mobile node network. This proposed new model which is optimized using genetic algorithm and effective fault tolerance shown by the system.

The prework has great evaluation metric to measure the extent of fault-tolerance towards system improvement over period. They concluded that this proposal will identify the best way to find system efficiency and robustness in fault tolerance. [9]

## III.FAULT TOLUENE IN CLOUD COMPUTING

Cloud computing is one of the complicated and challenging technology where 'N" numbers of physical and virtual servers are connected to a network. Various faults occur in cloud computing are very critical in real time and could lead to SLA (Service level agreement). There are four different types of fault layers 1) Physical hardware 2) virtual machine 3) network 4) application. Faults can happen essentially on two occasions one is on operating system, and another one is on VM or hardware level. [10]. Following are some of the faults that can occur at cloud infrastructure level.

1) *Network Fault:* It occurs in a network due to problems in one of the network hardware such as Ethernet, physical cables, switch/router and network firewall. Presence of this fault leads to Packet Loss, corruption, link failure and destination unreachable.
2) *Physical Faults:* This Fault occurs due to the problems in computer hardware such as CPU, memory, storage devices, storage adapter and power supply.
3) *Storage Faults:* It occurs in two different ways, one is due to hard disk crashes, and the other is due to the disk is not connecting to network or not accessible. When this fault occurs communication from physical server to storage server will be failed and VM will be crashed.
4) *Processor Faults:* When this fault occurs, operating system will be crashed.
5) *Process Faults:* This fault occurs when there is a bug in software and operating system (kernel) or shortage of resources, etc. This fault leads to service unavailability or server reboot.
6) *Service Expiry Fault:* This fault occurs when a process is killed due to hung state. The situation is also known as Zombie process.
7) *Permanent Fault:* This type of faults occurs when a device is at fault or a software issue. Permanent faults will be resolved when the device is replaced by a new device or replaced/reinstalled with software.
8) *File System Faults:* This fault occurs when a files system crashes. When it occurs, all the applications stored in the file system are crashed.
9) *Unknown Faults:* This fault occurs and suddenly disappears. This fault may occur due to network connection fails and will recover automatically.

### *A. Techniques Used in Fault Tolerance*

Different types of fault tolerance techniques are used in cloud computing to resolve the faults either at task level or workflow level. The faults can be resolved using either proactive technique or reactive technique.

1) *Proactive Fault Tolerance Technique:* It will identify the errors and faults before it occurs [or] suspected to occur. [11] Following three techniques will come under the proactive fault tolerance techniques.

*a)* Software rejuvenation: In this technique the cloud servers will be rebooted and updated periodically (approximately after every six months), and the server will start in clean state. If any unknown or zombie processor is running, that will be killed automatically. [11], [12]

*b)* Proactive fault tolerance – self healing: This technique is used when an application is running in multiple virtual severs at a time. This technique is also known as self healing as it resolves all the failure automatically.

*c)* Pre-emptive migration: It relies on feed back loop control mechanism where application is continuously monitored and analysed. [11]

2) *Reactive Fault Tolerance:* This technique can't fulfil the requirements in real-world, because it identifies errors after they occurred, and these errors will be resolved later. [13] But this technique will help to reduce the faults on software and hardware. Following Fig. 2 shows reactive fault tolerant real image on server.

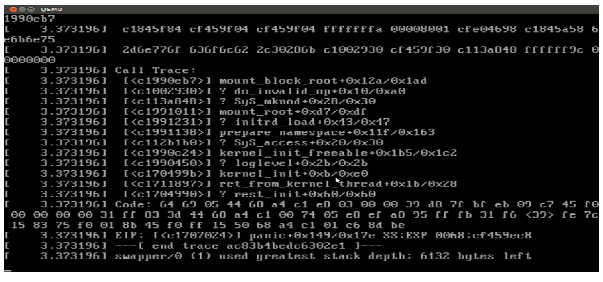



Fig.2 Kernel Panic error (Reactive Fault Tolerance)

a) *Replication:* Means copying, data can be copied from one server to another to avoid faults at further. For example, we can set up disaster recovery server for the production server using the replication. Disaster recovery server copies all the data from production server and the same data can be used to process the tasks as in production server.
b) *Job Migration:* Job migration moves the data from primary node to secondary node when primary node gets fails, the jobs in the batch also move to another secondary node automatically. This is also known as cluster.
c) *Task Resubmission:* Sometimes a fault may occur due to heavy load on server or network, task resubmission algorithm resubmits the task until the fault in the task is solved. [14]
d) *Checkpointing:* When a task fails, instead of initiating from beginning, it will restart from the recently check pointed state. Check pointing will be carried out periodically, and the process is executed from the recent check point, once system progress the feet. [14]
e) *Resource Allocation:* this method allocates resources from physical machine to virtual machine as per requirements. Some algorithms allocate the resources dynamically to the virtual machines in a cluster, when a faulty occurs.

## IV. PROPOSED METHOD – ERROR DETECTION PROGRAM(EDP)

In real-time environment, all the cloud servers (both Physical and virtual) are available 24/7 to provide continues services to the end users. It is also required that servers should reboot, update periodically to avoid unnecessary errors and faults. But after the update and reboot, the server may not be 100% in previous state.

The proposed project EDP (Error detection program) focuses on detection of errors or faults in operating system level (both physical and virtual) while the server is up and running in the cloud environment. EDP supports Linux flavours such as Red hat, Fedora, Centos, Mandrake and SUSE.

EDP uses software rejuvenation and task resubmission techniques of pro-active and reactive fault tolerance. This paper focus only on proactive rejuvenation technique which is presents below.

### A. *EDP Algorithm*

1) Run the EDP on up and running cloud servers
2) EDP implements proactive rejuvenation fault tolerance technique to resolve the faults before the server reboot or update. And execute several operating system commands at kernel level
a) Server info: Host name, IP Address of the server, Data center, Connection status.
b) Operating system information: Server type (physical or virtual), architecture, uptime status, Date & time zone info, OS version, kernel version.
c) Physical resources information: CPU cores, memory, out of memory(oom), dimm from edac error, Storage related errors.
d) File system information: Mounted file system, read only file system information.
e) Logical volume Information: Volume Group, physical volume, logical volume information.
f) Yum repos: yum repo list, latest installed packages, duplicate package, conflict packages, rpm db problems, repo name, yum summary.
g) NFS: Nfs version, NFS mount error, NFS status
h) Service info: VMware tool status, server monitoring status, last backup date, disaster recovery status, NFS service at boot uptime.
i) Network: Network routes, network ethernets, network gateway access,
j) Errors in logs: /var/log/messages
k) NTP Service: NTP connection, Date and time.
l) Error comments
3) After executing above commands, It generates output values in a text file. The text file contains unresolved errors which are in alphabetical, numerical, and alpha-numerical or some other format along with an error code for each error. The text file is later mapped into a database.
4) If still errors not resolved, It will implement reactive task resubmission fault tolerance technique and further also if any unresolved errors, those will be sent to the configuration management tool to resolve manually. These aspects are focused on next paper.

## V. IMPLEMENTATION

Install EDP on stand alone or network Linux servers in cloud environment. Run the EDP before reboot or update of cloud servers. It collects all the data as per the above commands and output the results in a text file. Text file is later mapped to a database as shown in fig. 3. Part I of EDP is written in bash shell using the ruby GTK and ganety APIs, which helps in finding duplicate server names or host IP addresses.

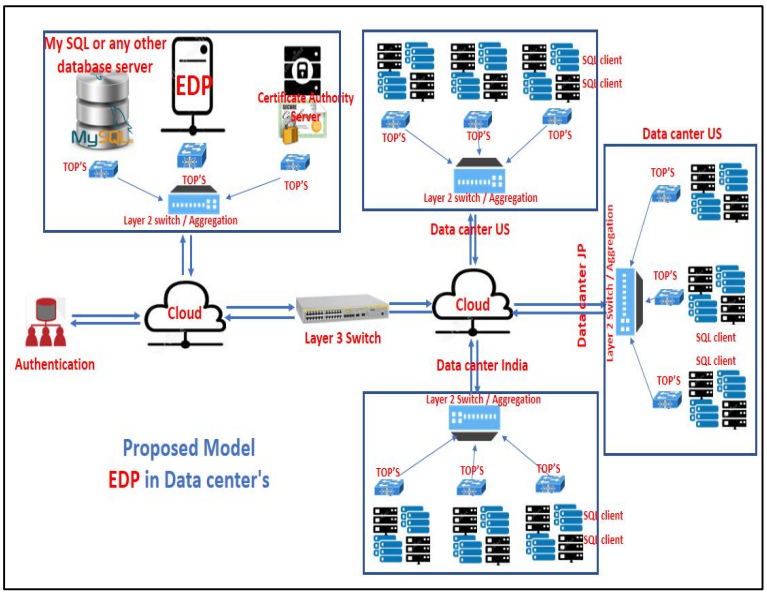



Fig.3 Proposed EDP implementation diagram

Part I of EDP is as follows: Store part I of EDP program in a folder and specify each server name in a separate line in the text file. When part I start running, it reads all the server names from the text file. Ganety program retrieves the server names from asset inventory for each datacentre. Ganety program pops-up when it finds a duplicate server name and asks for the correct server name. EDP uses an authentication service such as LDAP, which verifies user name and password and connect to the secure VPN which inturn connect to the data centre. By using SSH a secure communication will be established from EDP to servers in the data centre. By using SSL certificate authority, the EDP will establish end to end encryption to collect the error information from the servers. The experimental results of EDP shown in the following fig. 4.

### *A. Sample Script*

```
function check_pkg {
################ get our dc information
……..   *)
dc="xxxxxxxxxxxx"
    ;;
  esac
 fi
function update issue {  ……..
     done ;;    *) return ;;   esac }
##################function for yum update
function (){
check server info…….
      if [……. preferred authentications
update info / repo function redhat release
    return
  fi
 fi
```

```
finalize ()
{
secure connection authentication
        report _output = and stored in database
         ${out_report}
        database
        echo "final output report : ${out_report}" }
pbar for status ()
{}
######################execution of commands t kernel level
echo -e "\"${os_ver}\",\"${sys_date……,\"$status\",\"$comment\""
')
 echo -e "$enduser…….result}"> check_${pid}_${ip}.txt
                    fi      else
  echo "$....." > check_${pid}_${ip}.txt
       fi }
pbar $((number++)) ${_end}
echo
echo "creating report" finalize
```

```
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
duplicate packages :
-------------
**************************************************


====================================
conflict packages :
-----------------
Unsatisfied dependencies for libibnetdisc5-xxxxxxx.x86_64:
        libad.so.5()(64bit) is needed by (installed) libibnetdiscX-xxxxxxx.x86_64
        libad.so.5(AD_1.3)(64bit) is needed by (installed) libibnetdisc5-xxxxxxxxxxx.x86_64
        libad.so.3()(64bit) is needed by (installed) libibnetdisc5-xxxxxxxxxxxxx.x86_64
        libibumad.so.3(IBUMAD_1.0)(64bit) is needed by (installed) libibnetdisc5-xxxxxxxxxxxxx.x86_64
        libosmcomp.so.3()(64bit) is needed by (installed) libibnetdisc5-xxxxxxxxxxx.x86_64
        libosmcomp.so.3(OSMCOMP_2.3)(64bit) is needed by (installed) libibnetdisc5-xxxxxxxxxxxxx86_64
Unsatisfied dependencies for kernel-syms-9999999999999999.x86_64:
        pesign-obs-integration is needed by (installed) kernel-syms-9999999999.x86_64
++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
rpm DB issue :
----------------

  /run/lvm/lvmetad.socket: connect failed: No such file or directory
  WARNING: Failed to connect to lvmetad. Falling back to internal scanning.
  /run/lvm/lvmetad.socket: connect failed: No such file or directory
  WARNING: Failed to connect to lvmetad. Falling back to internal scanning.
  /run/lvm/lvmetad.socket: connect failed: No such file or directory
  WARNING: Failed to connect to lvmetad. Falling back to internal scanning.
```

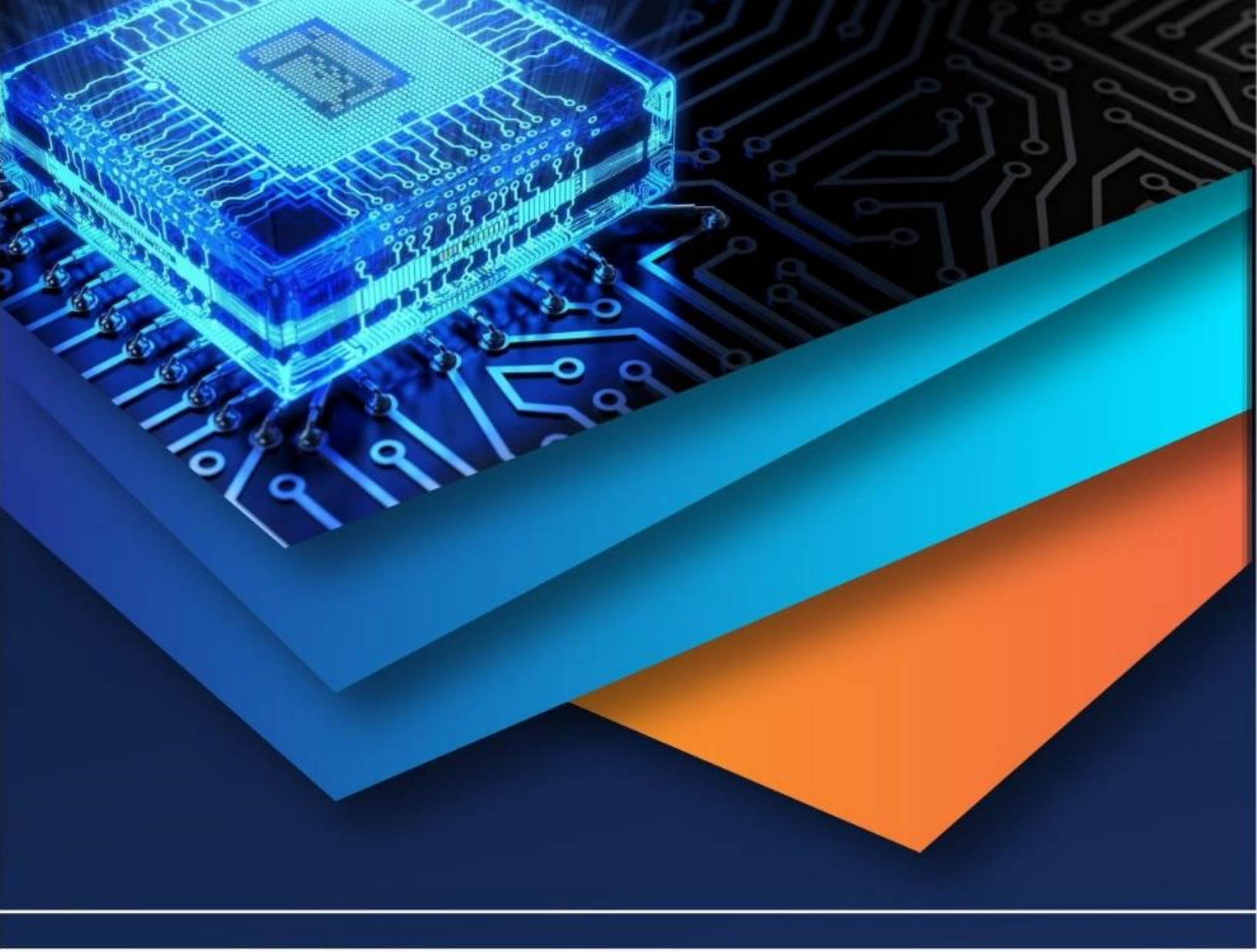
Fig.4 EDP Test Experimental results

## VI. CONCLUSION

In this paper we have discussed different types of faults or errors that lead to service and server unavailability in real-time cloud computing. We proposed a solution known as Error Detection Program (EDP) to resolve the faults. The EDP runs on all Linux flavors such as Redhat, Suse, Centos, Mandrake and other few flavors. EDP proactively identifies the errors and stores them in a database for further analysis. Proactive fault tolerance technique may not resolve all the errors and some of the errors remain unresolved. Further we are working on reactive fault tolerance technique which uses task re-submission strategy to resolve errors which are stored in the database.

## REFERENCES

[1] Sun Microsystems “Introduction to computer Archstractures “white paper first edition, June 2009.
[2] Anju Bala, indeer chnna, “Fault tolerance challenges, techniques and implementation in cloud computing”. IJCSI international Journals of computer science issue, vol. 9, Issue. 1, January 2012.
[3] Youssef M. Essa Bigdata Consultant, EMC, Cairo, Egypt, "A Survey of Cloud Computing Fault Tolerance: Techniques and Implementation" IJCA (0975 – 8887) Volume 138 – No.13, March 2016
[4] R. Kabilan, Dr. N. Jayaveeean, “Data Mining in Cloud Computing techniques: a survey paper” IJSEAS: International Journal of Scientific Engineering and Applied Science, Issue-8 NOV – 2015
[5] https://www.ibm.com/cloud
[6] Virendra Singh Kushwah1, Sandip Kumar Goyal2 & Priusha Narwariya3 "A Survey On Various Fault Tolerant Approaches For Cloud Environment During Load Balancing" (IJCNWMC) ISSN(P): 2250-1568; ISSN(E): 2278-9448 Vol. 4, Issue 6, Dec 2014, 25-34
[7] Siddharth Shyamsunder1, Kunal Varkhede2, Shirish Bhuruk3, Rasika Vibhute4, Prof R.P. Karande5 "Efficient Fault Tolerant Strategies in Cloud Computing Systems”, IARJSET 10.17148/IARJSET.2015.2208
[8] Mehdi Effatparva, Seyedeh Solmaz Madani, "Evaluation of Fault Tolerance in Cloud Computing using Colored Petri Nets " IJACSA, Vol. 7, No. 7, 2016.
[9] Dr. Chandralekha1, Satish Kumar Suman2, Anupriya, "Effective Fault Tolerance Management in Cloud Computing " IJETAE, Volume 6, Issue 3, March 2016
[10] Nikit Jain, Vishal srinivas “Data Mining techniques: a survey paper” ijret: International journal of research in engineering and technology, volume : 02 issue : 11 NOV – 2013
[11] Novel steps of authentication Technique for privacy presentation in cloud computing infrastructure.
[12] Y. Huang, C. Kintala , N Kiletties, and N. Fultoon, “Software rejuvenation: Analysis, Module and application”, In Proceeding of fault tolerant Computing Synmposium FTCS-25, June 1995.
[13] Dr. Laxmi Prasad Saiki, Yumana Langlen Devi, “Fault Tolerance Techniques and Algorithms in Cloud Computing” International Journals of computer science & communication Networks Vol(1), 01-18.
[14] Virendra singh kushwah1, sandip kumar goyal2 & priusha narwariya3 "a survey on various fault tolerant approaches for cloud environment during load balancing" International Journal of Computer Networking, Wireless and Mobile Communications (IJCNWMC) ISSN(P): 2250-1568; ISSN(E): 2278-9448 Vol. 4, Issue 6, Dec 2014, 25-34
[15] Dr. M.H Dunham, “ Data Mining, Introductory and Advanced Topics”, Prentic Hall, 2002.
[16] Imad M.Abbadi “self-management services, conceptual model in trustworthy clouds infrastructures”.
[17] Juna Li Pallvi Roy Samee u: Khan Lizhe Wang Yan Bhi” Data Mining Using Cloud; An Experimental Implementation of Apriori over mapreduce”
[18] A.Mahendiran, N. Sarvanan N. Venkata Subramanaian, N Sriram Implementation of K-Means clustering in cloud computing environment research journals of applied science, engineering and technology 4(10) ISSN 2040-7467.
[19] A Survey of Cloud Computing Fault Tolerance: Techniques and implementation, IJCA Volume 138, March 2016



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY