FRAMEWORK TO IMPROVE DATA INTEGRITY IN SAAS CLOUD

Enrollment number -122217

Name of Student- Anandita Singh Thakur

Name of Supervisor- Dr. Pradeep Kumar Gupta

May- 2014

Submitted in partial fulfillment of the Degree of

Master of Technology

DEPARTMENT OF COMPUTER SCIENCE AND TECHNOLOGY

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY,

WAKNAGHAT

TABLE OF CONTENTS

CERTIFICATE

This is to certify that the work titled "**Framework to improve Data Integrity in SaaS Cloud"** submitted by "**Anandita Singh Thakur**" in partial fulfillment for the award of degree of M.Tech of Jaypee University of Information Technology, Waknaghat has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature of Supervisor

Name of Supervisor Dr Pradeep Kumar Gupta

Designation Associate Professor (Senior Grade 2)

Date

ACKNOWLEDGEMENT

No work is a single man's success. Apart from the hard work and personal efforts, which are key ingredients, it requires the knowledge, guidance, encouragement and support. A work on completion signifies the joint efforts of the persons involved in it.

Firstly, I take this opportunity to express a deep sense of gratitude towards my guide **Dr. Pradeep Kumar Gupta** for providing excellent guidance, encouragement and inspiration throughout the project work. Without his invaluable guidance, this work would never have been a successful one.

I would also like to thank **Mr. Punit Gupta** for his scientific advice and knowledge and many insightful discussions and suggestions.

I wish to express my gratitude and high regards to **Prof. Dr. Satya Prakash Ghrera** head of CSE department, JUIT.

I would like to express my profound sense of gratitude to all the faculty members for their cooperation and encouragement throughout my course.

Last but not the least; I would also like to thank my family and friends for their valuable suggestions and undue support which has been a constant source of encouragement in all my educational pursuits.

Signature of the student

Name of Student Anandita Singh Thakur

Date

ABSTRACT

Cloud computing is a technology that being widely adopted by many big companies like Google, Microsoft etc in order to make the resources available to multiple users at a time over the internet. Many issues are identified due to which cloud computing is not adopted by all users till now.

The aim of my thesis is to improve the data integrity in SaaS cloud. A framework is proposed and the data from different users is protected by encrypting it using cryptographic algorithms namely: RSA, Bcrypt and AES. The algorithm is selected by user based on the level of security needed to be applied to the user's data. Performance analysis of the given framework is done using Cloudsim in Netbeans. It is seen that the time taken for encryption using the proposed framework is less than the time taken for encryption of a file using the old technique i.e. for different file sizes, different level of security is provided by the proposed approach which in turn takes less time for encryption. In previous scenarios all files were provided with same level of security irrespective of the type of data in it. Thus time for encryption of files using old technique was more. When the data is not modified its original form is sent to the user. Hence verification of data is done.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

NO.	NAME	FULL FORM
1.	ACID	Atomicity, Consistency, Isolation, Durability
2.	AES	Advanced encryption standard
3.	API	Application Programming Interface
4.	AWS	Amazon Web Services
5.	CDIBA	Cloud data integrity backup agent
6.	CSP	Cloud service provider
7.	CSPA	Cloud service provider agent
8.	DOS	Denial of Service
9.	EDOS	Economic denial of service
10.	IaaS	Infrastructure as a service
11.	MAS	Multi agent system
12.	PaaS	Platform as a service
13.	PDI	Prove able data integrity
14.	PDP	Prove able data possession
15.	POR	Proof of retrieve ablity
16.	RSA	Rivest Shamir Adleman
17.	SaaS	Software as a service
18.	SLA	Service level agreement
19.	SSL	Secure socket layer
20.	TLS	Transport layer security
21.	TPA	Third party auditor
22.	TTP	Trusted third party
23.	VM	Virtual machine

CHAPTER -1

INTRODUCTION

Cloud computing is an emerging technology where the resources, information and software are shared and are provided to the users over the internet as per their demands. In cloud, multiple systems can interact with each other at a time and move computing tasks from their system to the cloud. This way their burden of storing and maintaining of data locally is released. The computing services are delivered over the internet. These services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations.

The cloud architecture is classified into 4 layers [1] shown in figure 1:

1) Fabric layer - The hardware level resources like network resources, computing resources and storage resources are all contained in this layer.

2) Unified resource layer - All the resources that have been virtualized lie in this layer. These resources can be exposed to upper layers and end users as integrated resources.

3) Platform layer - A collection of services, software tool, middleware, are contained in the platform layer that provides development and deployment platform.

4) Application layer - The applications that would run in the clouds in contained in the application layer.

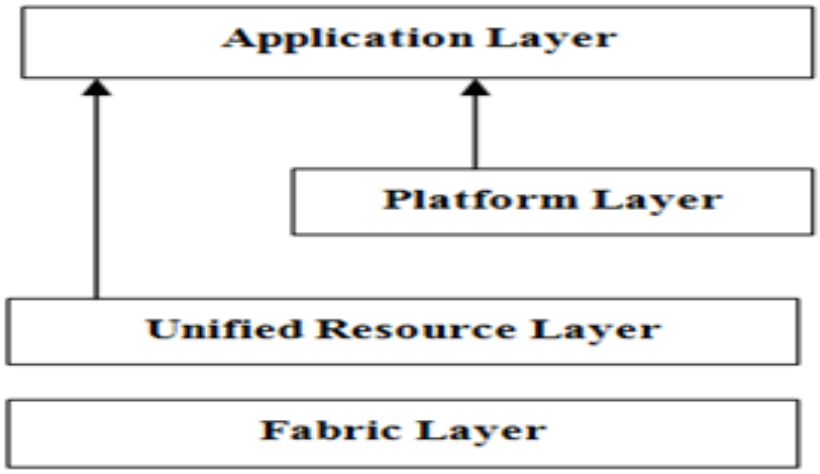

Fig.1 cloud architecture [1]

Cloud computing Service Models are classified into three categories [2] that are represented diagrammatically in figure 2.

1) *Infrastructure as a Service (IaaS)* – The IT infrastructures that are provided in the IaaS Cloud are directly used by the cloud Consumers. In order to meet the growing of resource demand from the consumers a concept called virtualization is used in IaaS cloud [2]. The infrastructure can comprise of virtualized storage, databases, servers and other items. Eg: Amazon Elastic Compute Cloud, a commercial platform offered as part of Amazon.com's Web Service platform.
2) *Platform as a Service (PaaS)* – It allows the cloud consumers to develop the applications directly onto the PaaS cloud. The deployment of applications directly into the virtual machine containers is minimized using PaaS [3]. Eg: Force.com and Amazon Web Services [AWS] offers services that allow developers to construct an application that is deployed using web-based tooling.
3) *Software as a Service (SaaS)* – The applications is released by the consumers in a hosting environment which can be accessed through the networks [4]. These are applications that offer an API to allow for greater application extensibility. Eg: Google Docs has been deployed solely within the Cloud and offers several APIs to promote use of the application.

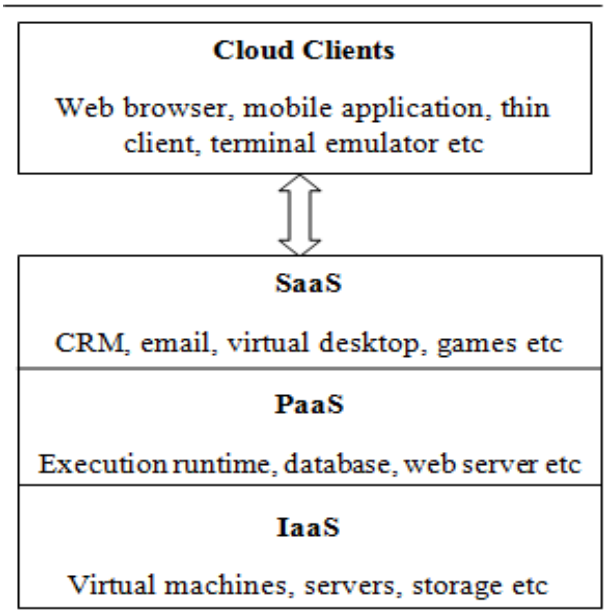

Fig.2 Cloud computing service models [4]

From research perspective focus of my work is on SaaS model of cloud computing.

There are different types of cloud each of which has their own benefits and drawbacks as shown in table 1.

1. *Public cloud* – On commercial basis the resources like applications and storage are made available to public by cloud service provider. The services may or may not be free. Using public cloud, the service provider does not require any initial capital investment on infrastructure. The only drawback is that it lacks fine control over data, network and security settings which reduces its effectiveness in various scenarios.[3]
2. *Private cloud* – The infrastructure of cloud is deployed and made available for specific group of people or organization. It provides highest degree of control over performance, reliability and security but it lacks providing benefit like no up- front capital costs.[3]
3. *Hybrid cloud* – It is a combination of various types of clouds. Some resources are made available for public use while some are to be used within a particular organization. It provides more flexibility than public or private cloud. Designing

the hybrid cloud is tricky as it requires determining the split between public and private cloud components very carefully.[3]

Table 1 Pros and cons for public, private, and hybrid cloud [5]

CLOUD TYPE	PROS	CONS
1. *PUBLIC CLOUD*	Implementation and usage is the simplest	Most expensive long term
	Upfront costs are minimal	Susceptible to prolonged service outages.
	Utilization efficiency gains through server virtualization	-
	Accessibility is widespread	-
	Requires no space dedicated for data center	-
	Can handle large spikes in workload	
2. *PRIVATE CLOUD*	Complete control of server software updates patches etc is allowed.	Upfront costs are large.
	Long term costs are minimal	Susceptible to prolonged service outages.
	Utilization efficiency gains through server virtualization	Access is limited.
	-	Large amount of space is required dedicated for data center.
	-	Cannot handle large spikes in workload
3. *HYBRID CLOUD*	Through utilization and flexibility of private and public cloud, this is most cost efficient.	Due to complex management schemes it is difficult to implement this cloud type.
	Less susceptible to prolonged service outages.	Moderate amount of space is required dedicated for data center.
	Utilization efficiency gains through server virtualization.	-
	Can handle large spikes in workload	-

Consumers, who wish to store their data in the cloud, either buy or lease storage capacity from them and use it for their storage needs. Cloud faithfully stores the data and return back to the owner whenever needed. But there is no guarantee that data stored in the cloud is secured and not altered by the cloud or TPA.

Cloud computing has various advantages which has made it one of the most exciting technologies and fastest growing parts of the IT industry.

1.1 Benefits of cloud computing [1, 4]

1. *Cost savings* - Companies can use operational expenditure and reduce their capital expenditure in order to increase their computing capabilities.
2. *Scalability and flexibility* - Companies can start with small deployment and with fair speed they can grow to large deployment and then can even scale back if required. The flexibility feature of cloud computing allows the companies to use extra resources at peak times so as to satisfy the customers.
3. *Reliability* - Services that make use of multiple redundant sites can support disaster recovery and business continuity.
4. *Maintenance* - The maintenance of system is done by the cloud service providers and access is through APIs which do not require application installations onto PCs, thus further reducing the maintenance requirements.
5. *User-centric interfaces* - cloud computing uses the concept of utility computing in which it becomes easy for consumers to obtain and employ the platforms in computing Clouds.
6. *On-demand service provisioning* - The resources and services are made available to the users according to their need.
7. *QoS guaranteed* - Cloud computing guarantees that quality of service would be rendered to its users.eg: size of memory, CPU speed etc.
8. *Autonomous System* - Cloud computing is an autonomous system and managed transparently to consumers.
9. *Pay-as-you-go* - This means that payment made by the user is according to the consumption of the resource.
10. *Energy efficiency* - The cloud reduces the consumption of unused resources thus making this technology an energy efficient technology.
11. *Multi-tenancy* - Services are owned by multiple providers in a cloud environment that are located in a single data center.

12. *Service oriented* - Cloud computing follows a service driven model where each PaaS, SaaS, IaaS providers provide service according to the Service Level Agreement (SLA) that is negotiated with its customers.

1.2 Issues in cloud computing

Though cloud computing has been adopted by the industries, it still has certain drawbacks. The foremost issue in cloud computing is security and privacy related to the data of the users. Since multiple users access the information in cloud at a time, the integrity and privacy of the information is at high risk. Cloud computing provides distribution of data over computers. When data is sent by the user to be processed in the cloud; the control of the data is given to a remote party that may not address security concerns of the user. As a user has no physical access to the data, he is unaware about the location of his data and is not sure whether the integrity of his data is maintained or compromised in cloud. It is important to ensure that the information being processed on cloud is secure and no tampering of information is done when previously unknown parties may be present [6].

1. *Security issues*: Security is termed as the prevention of any unauthorized access, unauthorized deletion or amendment of the information. The main dimensions of security that should be kept in mind for providing user satisfaction are availability, confidentiality and integrity [7].

    - Availability – The users can access the resources, information from any place and at any time. Denial of service attacks, equipment outages, and natural disasters are all threats to availability. According to Dimitrios Zissis and Dimitrios Lekkas [8], availability should not only be in terms of data, software but also hardware being available to authorized users upon demand.
    - Confidentiality – The aim is to keep the user's data secret. According to Xiao and Xiao [9], confidentiality is one of the major issues in cloud because the data that is outsourced by users on cloud servers is managed and controlled by untrustworthy cloud providers. In [10], Tianfield states

that threat to data increases because of increased number of applications, parties and devices which leads to increase in number of point of access. One way to achieve confidentiality is to encrypt the information sent by user before placing it in Cloud.

- Integrity – The aim is to preserve the integrity of the information. It should be checked that the information is not lost or modified by unauthorized user. One technique to maintain the integrity is usage of digital signature. In [10] the authors stated that by using SLA data is protected while it is on cloud, preventing intrusion or attack on data and responding swiftly to attacks such that damage is limited.

2. *Privacy issues*: The ability of individual or group to seclude themselves or information about themselves and selectively reveal them is termed as privacy. In term of organization, personally identifiable information is managed by providing privacy which involves the application of processes, standards, laws and mechanisms. Privacy issues vary according to different cloud scenario. The privacy issues are defined as follows [7]:
   a) Ensuring legal requirements for personal information is the responsibility of which party.
   b) How can the users maintain control over their data when it is stored and processed on cloud?
   c) How data replications can be guaranteed in a consistent state?

   In [11], Gharehchopogh and Hashemi stated that there are four forms of privacy:

a) Informational privacy
b) Environmental privacy
c) Relational privacy
d) Physical privacy

3. *Trust issues*: Trust is a measureable belief that is used to make trustworthy decisions based on experience. It has attributes like reliability, confidence, dependability, honest etc. The issues of trust in cloud computing are defined as [7]:

a) The attributes of cloud computing environment are unique, so the definition and evaluation of trust becomes difficult.
b) Based on the degree of trust, how to provide different security level of services.
c) The trust relationship in cloud computing is temporary and dynamic, so handling of malicious information is a tough task.

Apart from the issues stated above, there are other issues of cloud computing pertaining to its different service models i.e. IaaS, PaaS and SaaS.

- IaaS – It provides only basic level of security like load balancing, firewall etc. The applications moving to the cloud require high degree of security. One company may be hosting many other companies' workloads and data in a shared environment. In such cases, it may expose all parties to a higher risk of security or privacy related incidents [12]. The different issues in IaaS are stated in the table 2 below

Table 2 Issues in IaaS [13]

Area	**Security**	**Privacy**	**Trust**
DOS (denial of service)	Misconfiguration, vulnerabilities in system or OS	Access control compromised	Service not available
Robustness of virtual machine-level Isolation	Vulnerabilities in hypervisor	Internal network probing may occur	Compromised Virtual machines/ Hypervisors permit the loss of trust
EDOS (Economic denial of service)	AAA vulnerabilities	User provisioning, deprovisioning vulnerabilities	Access control compromised

- PaaS – The hacker can use the advantages of PaaS to influence the PaaS cloud infrastructure for malware command and control. One challenge that may be encountered while utilizing PaaS is compatibility. Since there is no list of features,

languages, APIs, software, database types, tools, or middleware which is common to all PaaS providers, it becomes difficult to choose the right one, or to switch. The various issues in PaaS are stated in table 3 below.

Table 3 Issues in PaaS [13]:

Area	**Security**	**Privacy**	**Trust**
Technical Immaturity	Compliance Challenges	Storage of data in multiple jurisdiction and lack of transparency about this.	Lack of information on jurisdictions
Lack of portability	Non availability of common authentication interface	"Data hostage" clause in supplier outsourcing contracts	Liquidate damage for lost business
Protecting API keys	Bad key management procedures	Service Information Leakage	Lack of sensitivity

- SaaS – The focus is not on application's portability but on migration of data and enhancement of security functionalities. For development and deployment of SaaS application certain security elements should be kept in mind that are seen below:
- Data security
- Network security
- Data locality
- Data integrity
- Data segregation
- Data access
- Authentication and authorization

The issues in SaaS are defined in table 4 below.

Table 4 Issues in SaaS [14]:

Area	**Security**	**Privacy**	**Trust**
Unauthorized Access	Data integrity and Confidentiality loss	Compromised communications secrecy	Loss of Trust in Service
Physical risks	Physically Destroyed data	-	-
Browser- Based Risks	Loss of data, Integrity and Confidentiality	Loss of user secret credentials	Loss of confidence upon channel
Network Dependence	Loss of availability	-	Trust on service reduces

In my research attention is paid on the security issues in SaaS which are defined as [14]:

- **Data security** – In SaaS model the data is stored at the vendor's end. The SaaS vendor should acquire additional security checks so as to ensure security of data and prevent data breach through unauthorized users. Strong encryption techniques should be involved for data security. Due to loophole in data security model, malicious users can gain access to the data.
- **Network security** – The sensitive data that is obtained from the users / organizations are stores at SaaS vendor end. All the data over the network must be protected to prevent leakage of data. This is achieved by using strong encryption techniques to manage network traffic like SSL and TLS.
- **Data locality –** The applications provided by SaaS are used by consumers and then data processing is done. The consumers are unaware of the fact as to where their data is getting stored. Example: In many European countries certain type of data

cannot leave the country because of the information being sensitive. The SaaS model should provide reliability to the customer in terms of location of data.

- **Data integrity** – It is easier to achieve data integrity in a single system with single database by making use of database constraints and transactions. Transactions follow ACID properties. The problem of data integrity gets magnified in case of cloud computing. The SaaS vendors unveil their web service APIs without any support for transactions. There are different levels of availability and SLA in each SaaS application which makes it difficult to manage the transactions and provide data integrity.
- **Data segregation** – Due to multi tenancy feature in cloud computing, multiple users can store their data on cloud. The data of various users will reside at same location. Intrusion in user's data by another user becomes easy in such environment. Intrusion can be done by hacking or by injecting client's code into SaaS system. Therefore, SaaS model should ensure boundary for each user's data not only at physical level but also at application level.
- **Data access –** Various security policies are provided by the organization to the users when accessing the data. Based on these policies each employee can access limited information. Cloud must stick to these security policies in order to avoid intrusion of data.
- **Data breaches –** Various users / organization put their data on cloud that would be at risk when there will be breaching into the cloud environment.
- **Backup –** The SaaS vendor should make sure that there is backup facility of the sensitive information so that it can be recovered in case of disasters. The backup data should be protected using strong encryption schemes that would prevent leakage of sensitive information.
- **Availability –** The services should be made available to the organization / users on anytime and anywhere basis. There will be architectural changes at the application and infrastructure levels to add scalability and high availability.

While reading various research papers on cloud computing and its service models, security was an important issue of concern apart from all the others issues. The users who move their data onto the cloud are unaware of the integrity of their data. They don't know as to where their data is getting stored. Due to these reasons some users are still adamant of making use of this technology.

There are many frameworks and algorithms that are proposed to resolve the issue of data integrity. The aim of my research is to provide a framework that would improve the integrity of data in SaaS model of cloud computing.

CHAPTER- 2
LITERATURE REVIEW

Cloud computing is an emerging trend in the field of technology that involves the development of parallel computing, distributed computing, grid computing and virtualization technologies [2]. There are various issues related to cloud computing, major ones being the security and integrity of data.

2.1 Security and privacy issues in cloud computing

In [15] Shaikh and Haider discovered that one of the reasons why cloud computing is not fully accepted by the users is, security. The users are always in fear of loss of their data as well as privacy. Some of the top security issues identified are: Data loss, Data Leakage, Client's trust, User's Authentication, Malicious users handling, Wrong usage of Cloud computing and its services, Hijacking of sessions while accessing data. To resolve these issues usage of Cloud Security Alliance (CSA) release of a new governance, risk management, and compliance stack for cloud computing was proposed. By using the cloud security tools, organizations create private and public clouds that comply with industry standards for accepted governance, risk, and compliance (GRC) best practices.

In [16] Dillon et.al discuss challenges and issues of Cloud computing. The relationships amongst Cloud computing, Service-Oriented computing, and Grid computing is articulated. Interoperability issue was highlighted which requires research and development. In cloud computing there are a number of levels that interoperability. Firstly, in order to optimize the computing resources, it is essential that an organization keeps in-house IT assets and capabilities associated with their core competencies while outsourcing marginal functions and activities on to the cloud. Secondly, a number of marginal functions must be outsourced by an organization to cloud services offered by different vendors. For example, it is highly likely that an SME may use Gmail for the email services and SalesForce.com for the HR service.

In [9] Xiao and Xiao, based on an attribute-driven methodology have systematically studied the security and privacy issues in cloud computing as shown below in figure 3. The security/privacy attributes like confidentiality, integrity, availability, accountability, and privacy-preservability, are identified and the vulnerabilities are discussed, which may be exploited by adversaries in order to perform various attacks.

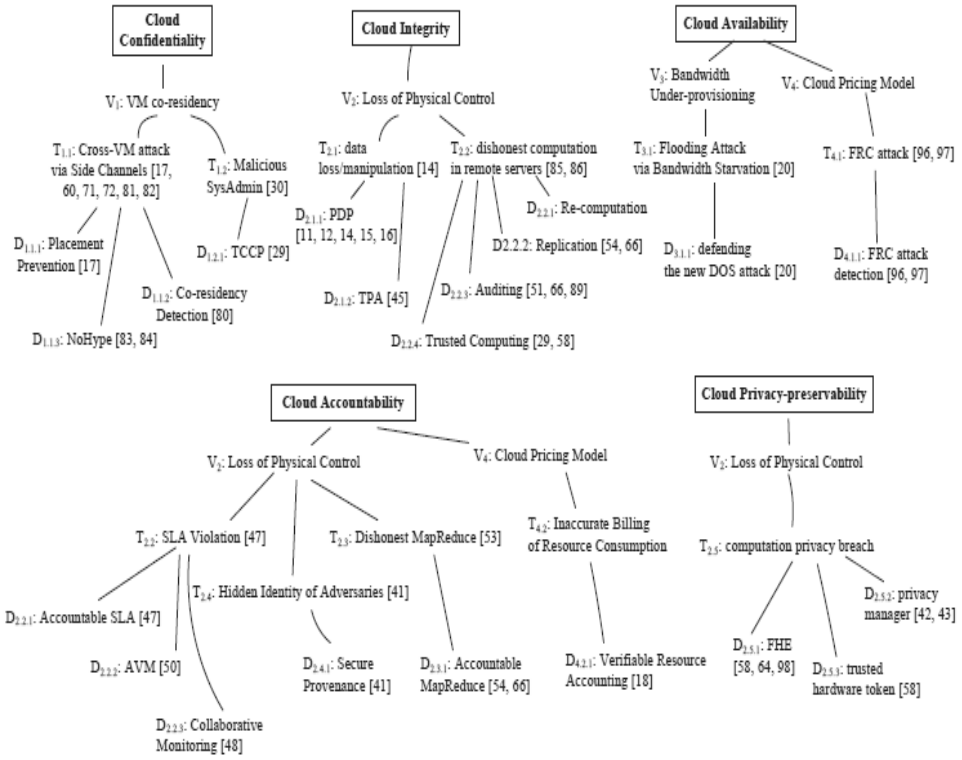

Fig.3 Summary of Research Advances in Cloud Security and Privacy [9]

Hu et.al in [17] present a survey on the architectures, concepts and challenges of cloud computing. A summary of challenges in cloud computing with respect to security, virtualization, and cost efficiency are discussed. Among the stated issues, security issues is the most important challenge in the cloud computing.

Chen and Zhao in [18] discussed about the analysis on data security and privacy protection issues associated with cloud computing across all stages of data life cycle along with some solutions. One concern is that what information to reveal and who can access that

information over the Internet. Another concern is whether web sites which are visited collect, store, and possibly share personal information about users. Privacy can be achieved by separating sensitive data from non-sensitive data followed by the encryption of sensitive elements.

Current security solutions for data security and privacy protection are discussed below.

- Roy I and Ramadan developed privacy protection system called airavat that can prevent privacy leakage without authorization in Map-Reduce computing process.
- A fully homomorphic encryption scheme was developed by IBM in June 2009. It allows data to be processed without being decrypted.
- A key problem for data encryption solutions is key management. On the one hand, the users have not enough expertise to manage their keys. On the other hand, the cloud service providers need to maintain a large number of user keys. The Organization for the Advancement of Structured Information Standards (OASIS) Key Management Interoperability Protocol (KMIP) is trying to solve such issues.
- NEC Labs's provable data integrity (PDI) solution can be used for public data integrity verification.
- Mowbray proposed a client-based privacy management tool for data storage and use stages. It provides a user centric trust model to help users to control the storage and use of their sensitive information in the cloud.
- A privacy protection framework was proposed by Randike Gajanayake based on information accountability (IA) components. The identification of users who are accessing information and the types of information they use is done by IA agent. In case of misuse being detected, the agent defines a set of methods to hold the users accountable for misuse.

Tianfield in [10] discusses about the various issues of security in cloud computing. In the paper cloud security requirements are analyzed in terms of fundamental issues like trust, availability, audit, integrity and confidentiality. As security is a major issue, it should be applied at different levels to ensure right implementation of cloud computing such as: security of host server, security of data storage, network security and security of application. Cloud security can be analyzed along three features:

1) *Identity security* - Key elements of cloud security is identity management at end-to-end level, authentication from third party and federated identity. The integrity and confidentiality of data and applications is preserved by identity security while providing access to appropriate users.

2) *Information security* - Information security is closely related to third-party data control. Concerns regarding information security include the way in which data is stored and accessed, compliance and audit requirements. All sensitive data including archive data, needs to be segregated properly on the cloud storage infrastructure. Encrypting and managing encryption keys of data in transit to the cloud or data at rest in the service provider's datacenter is critical to protecting data privacy and complying with legal and regulatory mandates.

3) *Infrastructure security* - The foundational infrastructure for a cloud must be inherently secure whether it is a private or public cloud or whether the service is SaaS, PaaS or IaaS. The cloud computing infrastructure, including servers, switches, routers, storage devices, power supplies, and other components that support operations and transaction of data and information, should be physically secure. Unauthorized user or employee should not be able to access any component.

- The various security issues in cloud computing were summarized diagrammatically in figure 4.

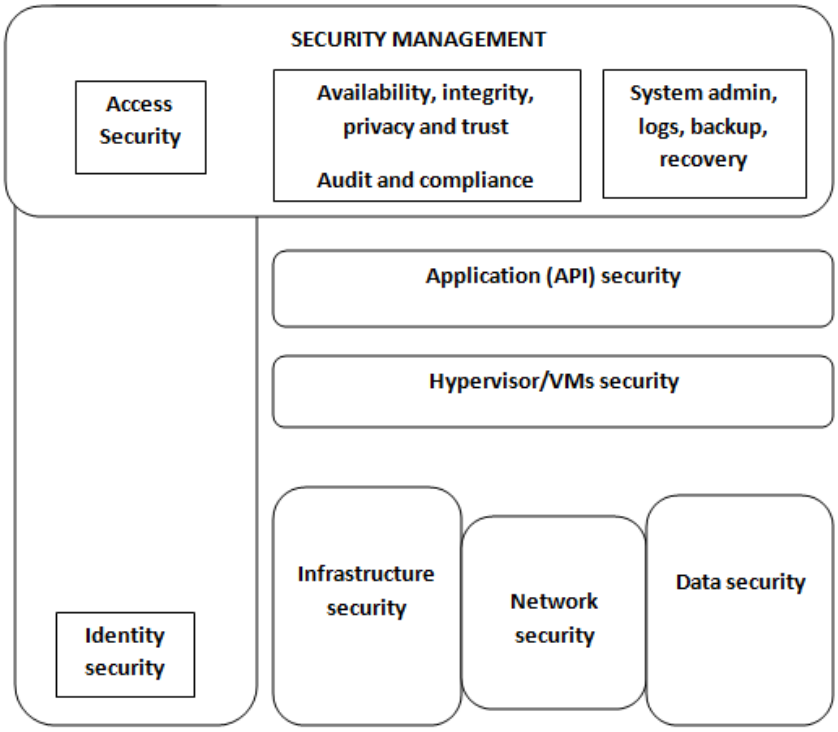

Fig.4 Cloud security architecture [10]

In [19] Bhadauria and Sanyal have stated that the entire data in a cloud computing environment lies over a set of networked resources and the access of data is provided through virtual machines. Since the location of data centers is not known and the users have no control of their data, there are many security and privacy challenges that have to be addressed. There are issues that are identified based on delivery models of cloud i.e. private cloud, public cloud. The various security threats are SQL injection, Cross Site Scripting (XSS), DoS and DDoS attacks, Google Hacking, and Forced Hacking. To resolve these threats, different methods are adopted by different service providers like to avoid the generated SQL usage in the code, identifying the meta-structures in the code, user entered parameters to be validated, etc.

Hamlen et.al in [20] discuss security issues for cloud computing and present a layered framework to secure clouds and then focus on two of the layers, i.e., the storage layer and the data layer. The issues include storage security, middleware security, data security,

network security and application security. The main goal is to store the data securely and manage the data not under control by owner of the data. A bottom up approach to security is proposed where work is done on small problems in the cloud that we hope will solve the larger problem of cloud security. Firstly, how documents can be secured is discussed so that they may be published in a third party environment. Next thing was that how security can be enhanced by using secure co-processors. It is found that due to complexity of cloud it is difficult to achieve end-to-end security. Even if some parts of the cloud fail the challenge is to ensure more secure operations. Building trust applications from untrusted components will be a major aspect with respect to cloud security.

To resolve the problem of privacy in the clouds P. Metri and G. Sarote [21] introduced a threat model which helps in analyzing a problem, designing appropriate strategies and evaluating the solutions. It has the following steps:

a) Attacks, threats are identified
b) Threats are prioritized according to the impact they have on the privacy of cloud user and cloud server. The prioritization of threats is done using STRIDE model.
    - Spoofing identity – means an attacker poses as another user or a machine poses as a valid machine
    - Tampering with data – means to maliciously modify the data.
    - Repudiation
    - Information disclosure – means to expose the information to the unauthorized users.
    - Denial of Service (DoS) – means to deny any services to valid users. Example: Web browser made temporarily unavailable.
    - Elevation of privilege – means the privileged access is gained by unprivileged users to destroy entire system.
c) Strategies are select for threats by considering the previous solutions along with the new strategies.
d) Based on these strategies solutions are built and applied.

Juels et al. [22] described a formal "proof of retrievability" (POR) model for ensuring the remote data integrity. It uses error correction code in order to check that the data is correct or not. In POR protocol the verifier stores only a single key for each file. The proposed scheme requires that only a small portion of file F is accessed by the prover in course of a POR. POR encrypts file F and randomly valued check blocks called *sentinels* are embedded. The prover is challenged by the verifier by specifying position of collection of sentinels and asking the prover to return associated sentinel values. If any modifications are made on file F by the prover then large number of sentinels is likely to be compressed. To protect corruption of file F by the prover, error correcting codes are employed.

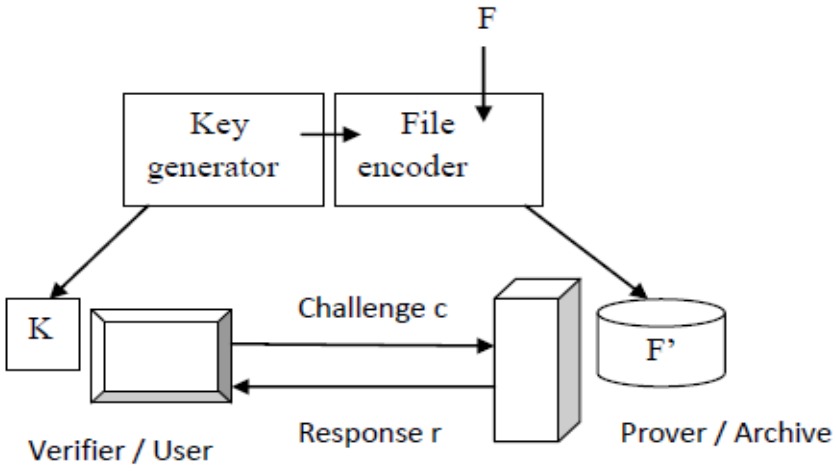

Fig.5 Schematic of a POR system [22]

In the figure 5, by using encoding algorithm raw file F is transformed into file F' and is stored with the prover / archive. A key K is produced by using the key generation algorithm and it is stored by the verifier. The key K is used in encoding mechanism. A challenge-response protocol is performed by the verifier with the prover to check that file F can be retrieved by the verifier.The drawback of the approach is that there is computational over head.

Ateniese et al. [23] considered public audit ability in their defined "provable data possession" model to verify if the client's data is stored at un trusted server. Homomorphic Verifiable Tags are used for data auditing. The model samples random sets of blocks from

the server and generates probabilistic proofs of possession which reduces I/O costs. The client verifies the proof by maintaining a constant amount of metadata. The transmission of a small, constant amount of data is done by the response protocol which minimizes network Communication. The PDP model for remote data checking supports large data sets in widely-distributed storage systems. There are certain performance parameters in PDP:

- Computation complexity: defines the cost to pre process a file, generate a proof of possession and verify such proof.
- Block access complexity: defines how many blocks are accessed to generate proof of possession.
- Communication complexity: defines the amount of data transferred.

To achieve scalable solution, the amount of computation and block accesses should be minimized. Experimentally it was shown that the previous schemes failed to give any verification for possession of large data sets which has been provided by proposed scheme.

In [24] Rana et al. proposed an architecture which combines IAAS and PAAS framework and remove the drawbacks of IAAS and PAAS. It describes how to simulate the cloud computing key techniques such as data storage technology (Google file system), data management technology Big Table as well as programming model and task scheduling framework using CLOUDSIM simulation tool. A machine contains the following four essential resources:

- CPU
- Memory (RAM)
- Disk
- Network connectivity

After examining these four parameters not enough information is provided by the server system metrics that would do meaningful capacity planning. Load cutting is a good parameter to be used under such conditions. This parameter is referred as resource cutting depending upon a server's configuration.

2.2 Frameworks to improve security

Mathew in [25] proposed a new framework that will help the cloud consumers and providers to safe guard the data. In the given framework, by using secured VPN the clients can access the provider's network. The providers check for user authentication. It should be checked that the clients that are approaching the providers are authorized and genuine. Once the providers are confident about the client's credentials their data will be encrypted and stored. The framework is shown below in figure 6.

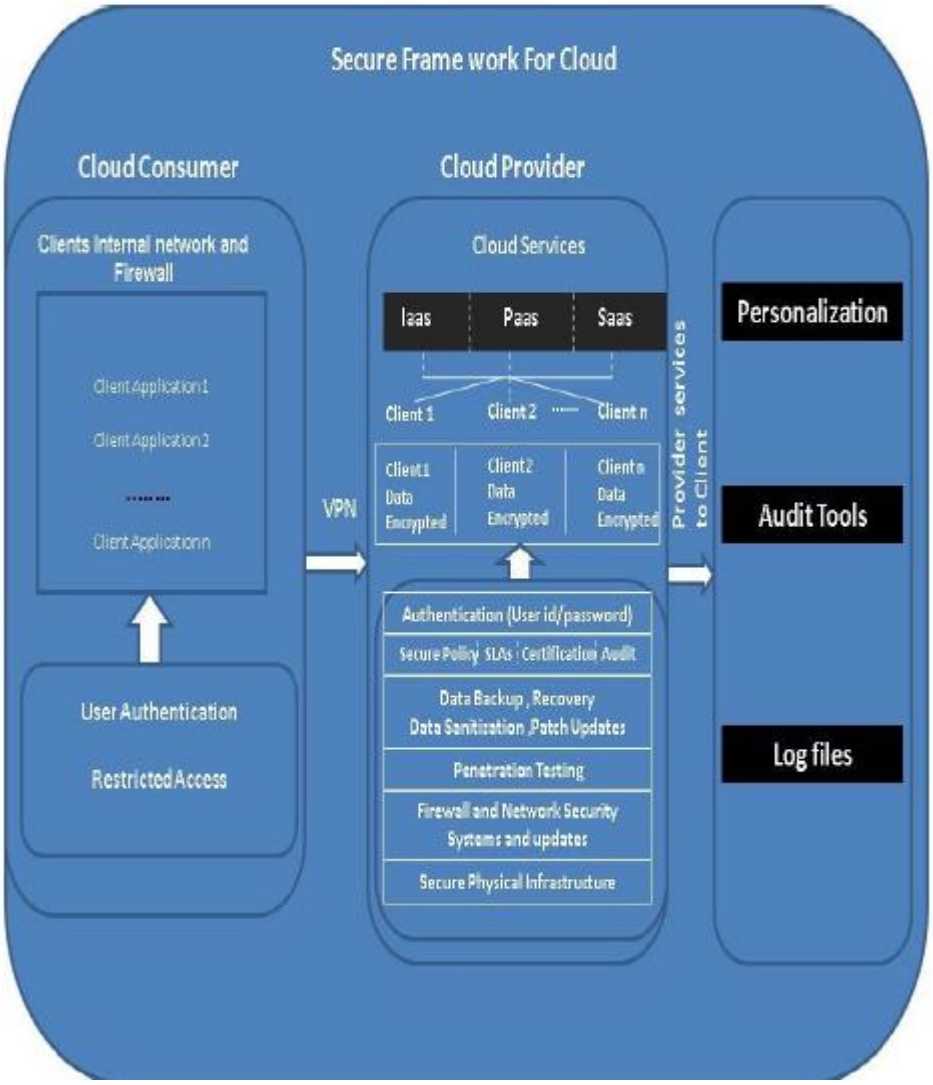

Fig.6 Cloud computing: A Secure Framework [25]

The steps to achieve security are defined as:

1) **POLICIES for SECURITY**
    - To ensure security a formal plan must be prepared by providers
    - Training should be given to employees on related technologies
    - Background check of employees.
    - Access restrictions / privilege setting to staff
    - Obligatory password change within stipulated time.
    - Supplier provided password of Server / hardware should not be used.
2) **DATA BACKUP, RECOVERY, SANITIZATION, PATCH UPDATES AND LOGS**
    - Backup of data to be carried out at regular intervals.
    - In order to meet unexpected disasters alternate plans should be ready.
    - Data recovery plans must be equipped with providers in all emergencies.
    - When the service or server is removed from the cloud, data should be deleted from servers and backup devices.
    - Updating system files and patches accurately.
    - System logs must be maintained with the following details that which users accessed the data, time spent and modifications made.

3) **PENETRATION TEST**
    - Testing has to be done at regular intervals to ensure providers system is not affected by any vulnerability.

4) **NETWORK SECURITY and FIREWALLS**
    - Firewalls must be installed and its policies, configurations, rules must be revised in habitual basis Antivirus updates must be done.

5) **PHYSICAL INFRASTRUCTURE SECURITY**
    - Physical Location of server is necessary; the storage devices should be kept in secured places by cloud providers with proper physical protection.

6) PROVIDER SERVICES TO THE CLOUD CONSUMERS

- Services like personalization, audit tools and log details are provided to consumers by the cloud providers.

It was concluded that if the providers and consumers follow the security measures discussed above cloud computing will be more secure.

Mohta et.al in [26] have discussed about a way to implement TPA who not only check the reliability of Cloud Service Provider (CSP) but also check the consistency and accountability of data. With the use of TPA, checking the risk in the cloud will be easy and less burdening to data owner but without encryption of data how data owner will ensure that his data are in a safe hand. A protocol is proposed that is used for supporting fully dynamic data operations, providing data privacy and integrity to end user. The architecture is stated below in figure 7.

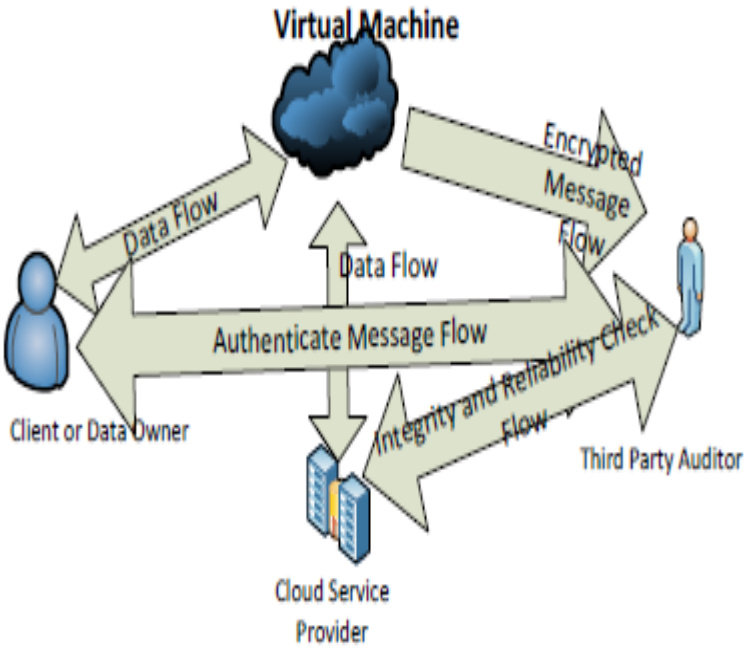

Fig.7 Architecture for Client, Third Party auditor and Cloud Service Provider [26]

The client asks the CSP to provide service where CSP authenticate the client and provide a virtual machine by means of Software as a service. In this Virtual Machine (VM), RSA algorithm are used where client encrypt and decrypt the file. In this VM, SHA-512 algorithms also defined which create the message digest. This message digest is a combination of client encrypted file, digital signature and mode of operation i.e. updating of records or insertion of records or deletion of records.

Experiments are performed on CloudSim, an open source simulator. The results demonstrate the efficiency and scalability the approach. All results were obtained after taking of 10 trials. From the results it is concluded that after getting digital signature of client and encrypted file the message digest take almost constant time that is shown in figure below in figure 8:

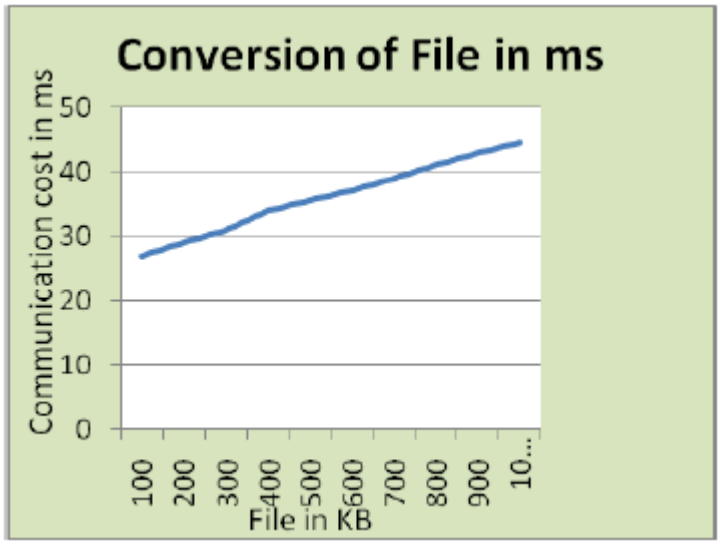

Fig.8 Communication cost verses File Conversion [26]

Bhosale et al. [27] provides a 3 dimensional framework along with digital signature and RSA algorithm where the user will upload the data over cloud based on the various security levels. Protection ring 1 will provide high level of security, ring 2 will provide less security and ring 3 will provide least level of security. Security of cloud is enhanced by using this framework with RSA and DSA algorithm combination. Availability of data is achieved by overcoming many existing problem like denial of services, data leakage. It also provides more flexibility and capability to meet the new demand of today's complex

and diverse network. Digital signature is a scheme that checks the authenticity of the document or a message. If a message is created by known user, then the digital signature will send a receipt to the sender stating that the message was not altered. An asymmetric type of cryptography is employed by the digital signature. When a digital signature is properly implemented then even if messages are sent through a non secure channel, the digital signature gives the receiver reason to believe that the message was sent by the claimed sender. Non-repudiation is provided by the digital signature, meaning that the signer cannot claim they did not sign a message. Even while claiming their private key remains secret. Some non-repudiation schemes offer a time stamp for the digital signature, so that even if the private key is exposed, the signature is valid nonetheless. Digitally signed messages may be anything that can be represented as a bit string. Examples: electronic mail, contracts, or a message sent via some other cryptographic protocol. RSA supports encryption and digital signatures. RSA gets its security by integer factorization problem. It is easier to understand and implement the RSA algorithm.

2.3 Frameworks to maintain data integrity

In [28] Chalse et.al provides a detailed analysis of the cloud security problem. The different problem in a cloud computing system and their effect upon the different cloud users are also analyzed. The architecture taken into consideration is shown in figure 9.

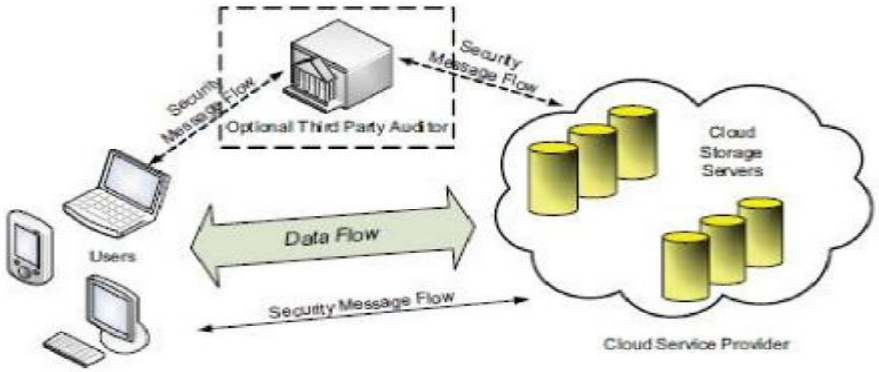

Fig.9 Cloud data storage architecture [28]

In the architecture, there is

1) A client who wants to store large amount of data in multiple clouds and have the permissions to access and manipulate stored data.
2) Cloud service providers (CSP) who work together to provide data storage services and have enough storage and computation resources.
3) Trusted Third Party (TTP), who is trusted to store verification parameters and offer public query services for these parameters.

In order to store and maintain client's data, multiple CSPs are taken into consideration. All data of client is verified using PDP.

The verification procedure is described as follows:

- A client uses the secret key to pro-process a file which consists of a collection of n blocks, generates a set of public verification information that is stored in TTP, transmit the file and some verification tags to CSPs and may delete its local copy
- By using a verification protocol, the clients can issue a challenge for one CSP to check the integrity and availability of outsourced data with respect to public information stored in TTP.
- Analysis of security is performed. The data is encrypted to ensure that the file will not be intercepted by an unauthorized person. Because encryption and decryption SecretKeyGen and VeriTagGen cryptosystem uses public key and private key, security is based on calculating private key. If private key is not known, the file cannot be decrypted.
- Integrity is checked in the verification phase. The owner would like to do verification of cipher text M stored on the server. The server will calculate the value of z to prove that he has completely stored cipher text file M. If the server calculated value z calculated with the owner of the verification value is equal to V, it means the server does have the correct storage cipher text file M. The results are shown in the snapshots stated in figure 10 and figure 11.

Fig.10 Cloud server login [28]

Fig.11 Client server side [28]

The proposed solution support dynamic outsourcing of information making it a more realistic application of cloud computing.

In [29] R and Saxena provided a scheme that gave a proof of data integrity in which the correctness of the user data can be done by the user. This proof can be incorporated in SLA and is agreed upon by both the cloud and the customer. The proposed scheme does not involve the encryption of the whole data. Few bits of data per data block are encrypted; hence the computational overhead on the clients is reduced. The overhead of client storage is also minimized as it does not store any data with it. In the given protocol irrespective of the size of data file, the verifier needs to store only a single cryptographic key and two functions which generate a random sequence. No data is stored with it by the verifier. The verifier before storing the file at the archive preprocesses the file and appends some metadata to the file and stores at the archive. In order to verify the integrity of the data, the verifier uses the metadata that is stored. The protocol just checks the integrity of data i.e. whether the data has been tampered or deleted. The modification of the data by archive is not prevented. The scheme applies only to static storage of data. It cannot handle to case when the data need to be dynamically changed.

M.V and Santhanalakshmi in [30] proposes an effective and flexible Batch Audit scheme with dynamic data support to reduce the computation overheads. To ensure the correctness of user's data a third party auditor (TPA) is used, to verify the integrity of the data stored in the cloud.

Symmetric encryption is considered for effective utilization of outsourced cloud data under the model, it achieve the storage security in multi cloud data storage. The new scheme further supports secure and efficient dynamic operations on data blocks, including data insertion, update, delete and replacement. Extensive performance analysis shows that the proposed scheme is highly efficient and resilient against Byzantine failure, malicious data modification attack, and even server colliding attacks. The protocol supports public audit ability and data dynamics. By processing the integrity of data using data reading protocol and data management algorithm after and before the entering of data into the cloud, user

can assure that all data in cloud must be in protected condition for its trustworthiness. The actual size of stored data before and after in cloud is maintained even though the user himself has done any modification, deletion, and update for his own purpose by using proposed scheme. These processes are carefully done using proposed scheme. So here user takes full control and process on the data stored in cloud apart from TPA and give strong assurance and protection to the data stored in multiple cloud server environments. To avoid server failure and any unexpected error one server restore point is put in cloud server database for efficient data back up or restore using multi server data comparison method. It is major advantage of proposed system. This process is done with the help of CSP for cloud database process since we have physical data possession in cloud server. The model is shown in figure 12.

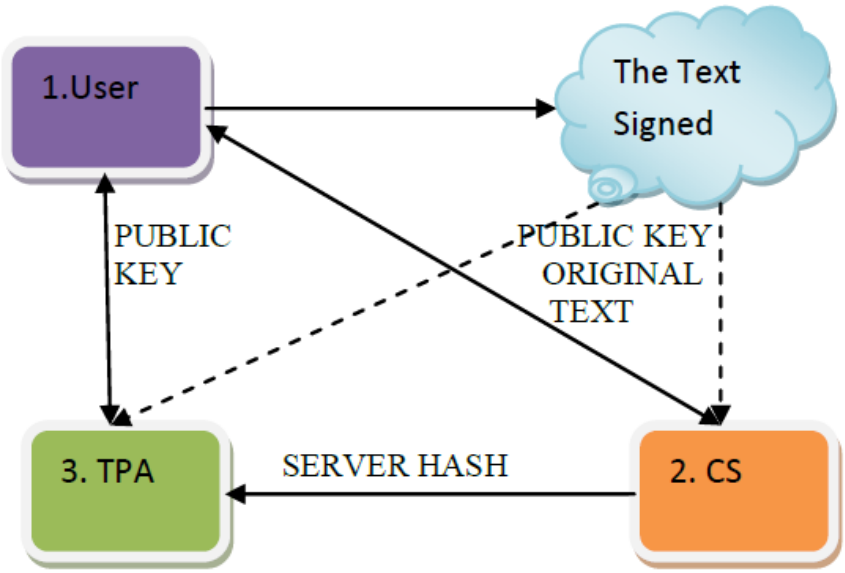

Fig.12 Proposed Architecture [30]

1)User: First a random parameter is chosen by user to construct the public and the private keys then he\she will sign the data using the private key to be uploaded to the cloud, and then he\she will send the signed data to the cloud server and deletes its local copy.

2)CS: CS will compute a hash value from the original data to send it to the TPA, and then takes this hash value along with the data signed in the cloud for verification using the public key.

3) TPA: After the cloud server finishes its role, the TPA will be initiated to verify over the cloud server work by taking the hash value from the cloud server. TPA will take the data signed form the cloud and decrypt it with the public key.
From this protection for cloud data, user can be strong belief for his uploaded data for any future purpose or his any other related process without worry.

Raju et al. [31] introduced a protocol for integrity checking of cloud storage that would provide integrity protection of user information. This protocol supports public verifiability and is evidenced to be secure against associate un-trusted server. It's additionally non-public against third-party verifiers. The prediciton of information consistency was missing in the existing systems and with the help of the proposed protocol this problem is resolved.A concept of Voucher was introduced that arbitarily checked the knowledge blocks by causing a challenge request and challenge response from the packet. If the challenge request and challenge response matched, then the block is valid otherwise the block is affected. The protocol supports insertion, modification and deletion of information at block level.

Attas and Batrafi [32] proposed an integrity checking model over cloud with help of TPA using DSA algorithm. User can examine and verify the data from unauthorized people who manipulate with the cloud or extract data. Upon uploading large files on the remote cloud servers, the users cannot verify the data in remote servers.TPA does this verifying task in order to help the users as TPA is reliable and independent. The TPA has access to cloud provider environment and understands the service level agreements between the user and the provider. A random parameter is chosen by the user for construction of public and private keys. The data to be uploaded on cloud is signed by the user using private key. This signed data is sent to the cloud server and its local copy is deleted. Hash value is computed from the original data by the cloud server and sent to TPA. Using the public key, verification of the hashed value with the data signed in the cloud is done by cloud server. The user is informed by cloud server whether there is any intrusion in data on cloud or not.
The TPA verifies the work done by cloud server by taking hash value from cloud server. TPA will decrypt the data signed from the cloud using public key. The result of decryption

will be a hash value which will be compared with the hash value that the cloud server computed in its part. After verification, the users will be informed by TPA whether the cloud server was trustworthy or not. Evaluation of the model is done using Windows Azure project that involves digital signature coding. The results show that the proposed model worked according to what was claimed.

In [33] Kumar et al. assess that how cloud providers can gain trust of their customers and provide them with security, privacy and reliability on the data when processing of sensitive data is done by the third party in remote machines located in various countries. Various services are made available to the user by using the concept of utility cloud. The advantage of proposed approach was to incorporate the trusted computing technology into the cloud computing environment to achieve trusted computing requirements for the cloud computing and then fulfill the trusted cloud computing. The importance of trust varies from one organization to another depending upon the data values.

Talib et al. [34] proposed layered architecture based on MAS architecture having two main layers: (a) cloud resource layer [cloud server side] (b) MAS architecture layer [cloud client side]. The MAS architecture has two agents namely Cloud service provider agent [CSPA] and Cloud data integrity backup agent [CDIBA]. This layered architecture collectively was called "Cloud Zone". The cloud data travels as a series of messages, queries between the cloud server and user that are sent using cloud protocols. CDIBA provides backup of cloud data by defining some set of rules using logical grouping of cloud components. Specific activities pertaining to certain user or part of CDS is monitored by CSPA. There are certain requirements of Cloud Zone that are kept in mind:

- Only MS SQL databases are backed up in "Cloud Zone."
- Component based backup is not supported by "Cloud Zone."
- Backup and recovery of Windows Oracle 11i is supported.
- There is no usage of Visual SourceSafe (VSS) for backup and recovery.

Nirmala et al. [35] proposed a new scheme to resolve integrity problem by introducing user authenticator to audit and check the integrity of data. Their research focused on providing

solutions to all issues of cloud computing and to develop a model that would provide secure cloud infrastructure which would help to adopt the cloud as and when required. There are four stages in the proposed scheme:

a) System model: Homomorphic authenticator technique is opted to avoid bringing of data blocks from blocks and support audit ability as depicted in figure 13.

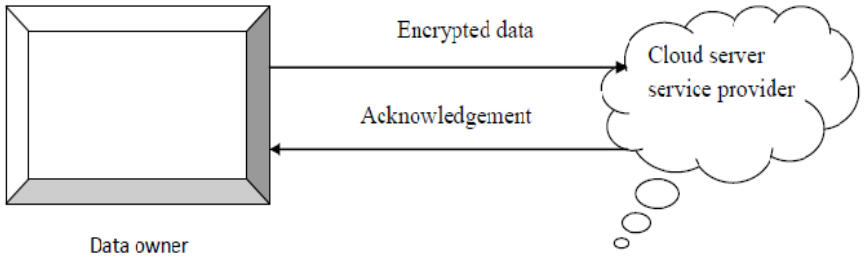

Fig.13 System model [35]

b) System startup: the generation of public key and private key is done by invoking KeyGen( ) and the data file is preprocessed and production of meta data along with homomorphic authenticators is achieved by running SigGen( ).

c) Integrity verification of static data: To verify the integrity of data, the owner sends a request to send hash code for any block as shown in figure 14

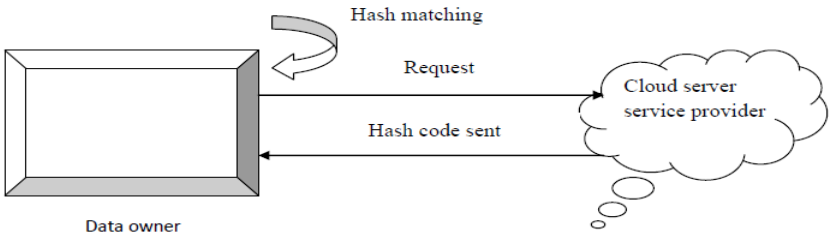

Fig.14 Integrity verification of static data [35]

d) Integrity assurance for vibrant data: The proposed scheme can handle the dynamic data operations (insert, update, delete) efficiently.

- Data update: in figure 15 update operation performs the changes in the given data set

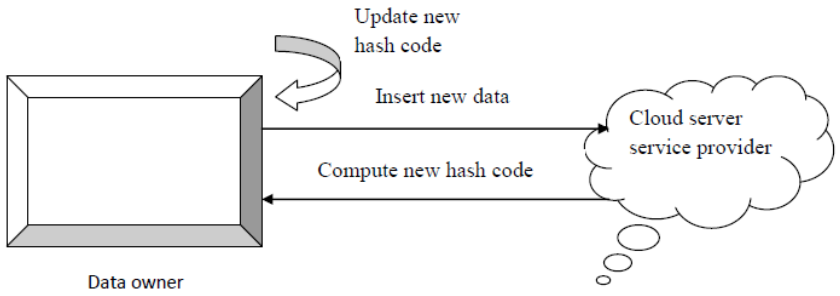

Fig.15 updating vibrant data [35]

- Data insertion: This mechanism changes the logical structure of the file. New blocks are inserted in file F after some specified positions. Insertion can be done by data owner or data sharer.
- Data deletion: After deletion of a specified block, all the latter blocks are moved one block forward.

In [12] Eswaran et al. proposed an approach to secure the data and ensure the integrity of data in cloud using cryptographic keys. Proof of retrieve (POR) ability is used in order to verify the integrity of data. The schematic view of POR is as:

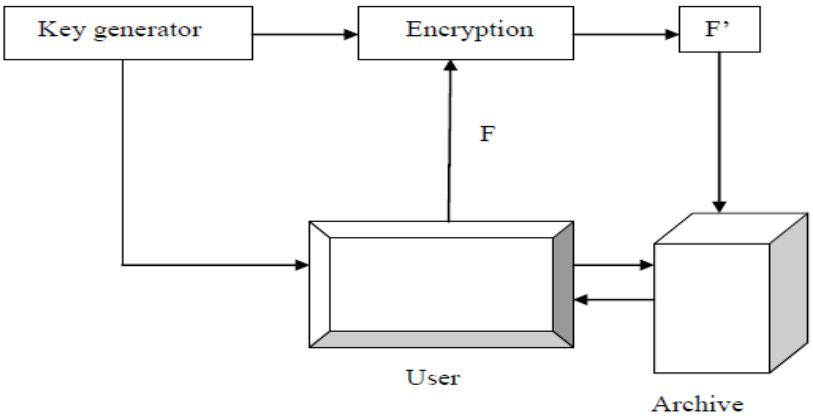

Fig.16 schematic view of POR

From figure 16, it is clear that the owner of file must encrypt the data in file before storing it in cloud in order to prevent unauthorized access to the file. The proposed system makes sure that unauthorized users are not permitted to login. Only authorized client can upload the file into cloud. While uploading the files into the cloud, key generator of the proposed system generates an encryption key and sends to the owner. Every file which is uploaded in the cloud is verified by TPA stating whether it is secured or not. The verification process can be done in two ways:

1) *Direct verification* - In direct verification, the file is verified by TPA without any need for cryptographic key.

2) *Download verification* - In download verification, cryptographic key is required. Request of key is sent by TPA to the owner. Key is sent by the owner to the TPA so that the integrity of the file is checked by TPA. Once verification of file is done, TPA sends a report to the owner. If TPA modifies the file and again uploads it, this means that cloud has sent alert the owner. Through this process integrity of file is verified.

CHAPTER-3

PROPOSED FRAMEWORK AND ALGORITHM

3.1 Proposed framework

This section defines the proposed framework to provide data integrity in multi cloud system. Proposed framework shown in figure 17 has three main roles [36]:

1*) Users* - who will store the data by selecting appropriate layer depending on the level of security needed for the data stored on cloud.

2) *Cloud service provider (CSP)* - provides the storage of data service with flexible resources to keep the user data. The CSP manages cloud server (CS) which informs the user about the intrusion of data on cloud.

3*) Third party auditor (TPA)* - verifies the cloud server and checks whether there is any manipulation of user data by the cloud server. It then sends a report to the user stating that the cloud server (CS) was trusted or not.

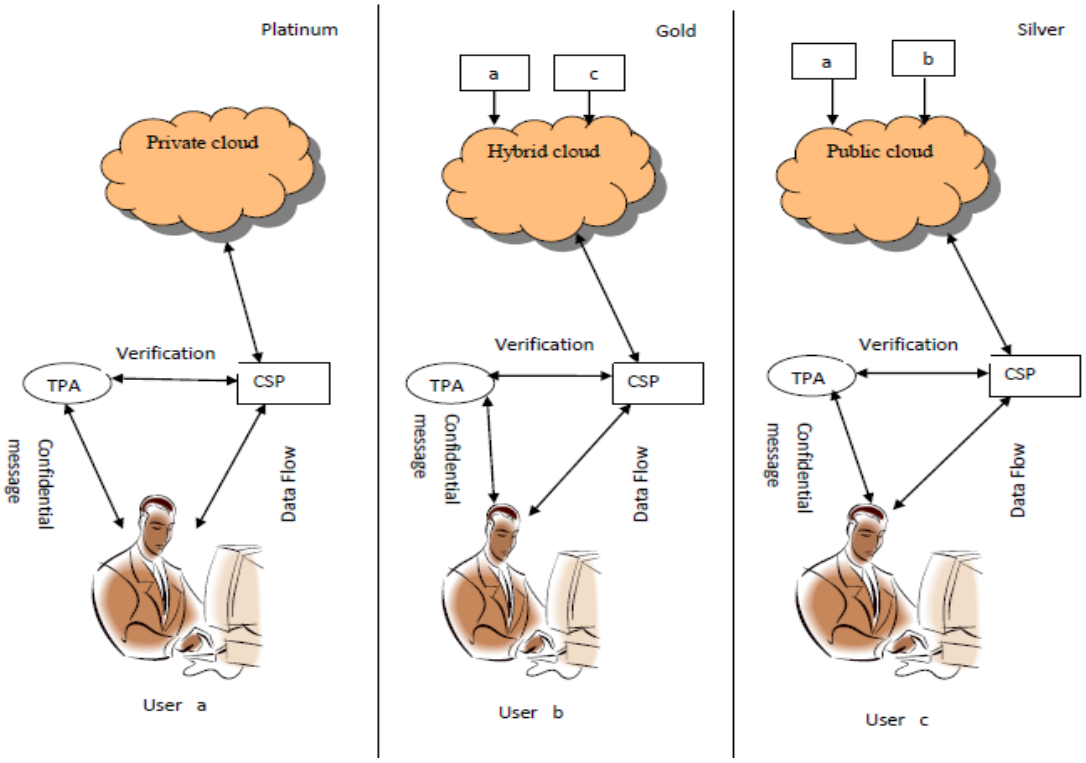

Fig.17 The proposed framework [36]

There are many cloud service providers and each of them provides different storage plan along with different QoS parameters so it becomes a tough task for users to keep moving their data from one cloud to another based on QoS and cost optimization [37]. In the proposed model concept of multi cloud is used to provide best cost optimization for various requirements of user. To give a clear design of the model, use of connectors is made labeled as a, b and c.

The useage of connectors has been made in order to provide better view of the model. It can clearly be seen that depending on type of data the user can move from one service provider to another.

Eg: the user "a" can put his data over hybrid or public cloud depending on the security needed for the data stored.The same thing can be applied by other two users.

Depending on the type of data to be stored on various clouds, there are three main platforms in the model namely:

- *Platinum*- sensitive data will be stored here like data related to transactions of atm, bank account information along with high level of security on the data. The data will be stored on private cloud.

- *Gold*- data related to simple login on any page like facebook, ebooking and email login is stored. The level of security needed is not that high. Security only on password is required.

- *Silver*- data related to only simple browsing of sites, uploading of images, downloading of files like downloading of music files or images is stored. The level of security needed is the least.

Based on these levels the user will decide on which platform to store the data.

3.2 Proposed algorithm

In this section we have implemented various algorithms like RSA algorithm, Bcrypt algorithm, and AES algorithms to implement the proposed framework.

/* Variables used for:

User => u,

Platinum => p,

Gold => g

Silver => s

*/

Begin

If *u chooses p*

Then *call module m1*

If *u chooses g*

Then *call module m2*

If *u chooses s*

Then *call module m3*

End

Module m1:

Begin

1. *User's data is encrypted using RSA and sent to CSP.*
2. *Data is verified by CSP*
3. *If data is valid*

   success message
4. *Go To Module T*

End

Module m2:

Begin

1. *Data stored is encrypted using Bcrypt algorithm*
2. *Data is verified by CSP*
3. *If data is valid*

success message

4. Go to module T

End

Module m3:

Begin

1. *Data is encrypted using AES algorithm*
2. *Verification of data is done by CSP*
3. *If data is valid*

   success message
4. *Go To Module T*

End

Module T:

Begin

1. *Check the data stored in file.*

If *user's data == cloud data then*

data valid

Else

corrupted data

End

In module m1 RSA algorithm [38] is used to provide integrity of data because for storing sensitive information on cloud, hashing algorithms are used. RSA is based on the difficulty of factoring large numbers. There are various advantages of RSA due to which it is preferred over DSA.

- DSA can only be used for authentication while RSA can be used for both authentication and to encrypt a message.
- A bad random number generator will leak DSA key bits.
- Faster at encrypting than DSA.

In module m2 Bcrypt algorithm is used for hashing the passwords. A password hashing algorithm should preferably be slow in order to prevent brute force attacks; it should have

features which actually decrease the feasibility of a distributed brute force attack on the hashes. The following hashing algorithms are not considered for the purpose:

- MD-5
- SHA-1
- SHA-2
- SHA-3

Bcrypt algorithm is derived from the Blowfish block cipher which uses look up tables that are initiated in memory to generate the hash.

In module m3 AES algorithm is used to provide security on the data stored. AES is asymmetric encryption algorithm in which to encrypt the message sender uses public key of receiver and its private key is used by receiver to decrypt the message.

a) AES is preferred over DES algorithm as it is more secure.
b) AES data encryption is mathematically more efficient and elegant cryptographic algorithm. Key length option is the main strength of the algorithm. Time required to crack an encryption algorithm is directly related to the length of the key used to secure the communication. AES gives an option to choose a 128-bit, 192-bit or 256-bit key, making it exponentially stronger as compared to the 56-bit key of DES.
c) Block size of DES is small compared to AES
d) A balanced Feistel structure is used by DES while substitution-permutation is used by AES.

CHAPTER- 4
PERFORMANCE ANALYSIS

A framework has been proposed in order to improve the integrity of SaaS service model of cloud using the concept of multi cloud. Encryption algorithms are used to achieve the results. An algorithm is proposed to implement the cryptographic algorithm for encryption. The analysis of the different files with varied sizes is done using Netbeans tool.

In order to provide assurance of characteristics in cloud systems like reliability, security, fault-tolerance, sustainability, and scalability computational services timely, repeatable, and controllable methodologies are required for evaluation of new cloud applications and policies before actual development of cloud products [39]. **Simulation** is a flexible methodology that is used for analysis of behavior of a present or proposed business activity, new product, manufacturing line etc. **Performing simulations** and **analyzing the results**, helps to know the functioning of the present system, and what would happen if changes are made to it – or estimation of behavior of the proposed new system is done. IT companies have various benefits of using simulation based approaches by allowing:

(i) Testing of their services in repeatable and controllable environment

(ii) Before deploying on real clouds tuning of system bottleneck should be done

(iii) For developing and testing adaptive application techniques experimentation with different workload mix and resource performance scenarios on simulated infrastructures should be done [40].

In order to evaluate the proposed framework simulation is done using a cloud simulation tool named “CloudSim”.

4.1A brief introduction to CloudSim simulation tool

A simulation framework that enables modeling, simulation, and experimentation of emerging Cloud computing infrastructures and management services. Main features of CloudSim include [41]:

- To model and simulate large scale Cloud computing data centers
- To model energy-aware computational resources and simulate it
- Data center network topologies and message-passing applications are modeled and simulated
- Simulation elements inserted dynamically, simulation stop and resume feature.
- User-defined policies for allocation of hosts to virtual machines and policies for allocation of host resources to virtual machines.

Its advantages are:

- Time effectiveness
- Flexibility and applicability
- Test policies in repeatable and controllable environment
- Tune system bottlenecks before deploying on real clouds

4.2 Results

4.2.1Analysis of RSA and AES algorithm

To estimate the output of the code, matlab tool is used.

- RSA ALGORITHM ANALYSIS

RSA algorithm is tested for integer numbers ranging from a single digit message length to 16-digit message length. The execution time t is in seconds. The execution time depends on the values of p and q which are prime numbers.Different values of p and q are taken and depending on these values graph between message length and time is plotted as shown in figure 18 and figure 19.

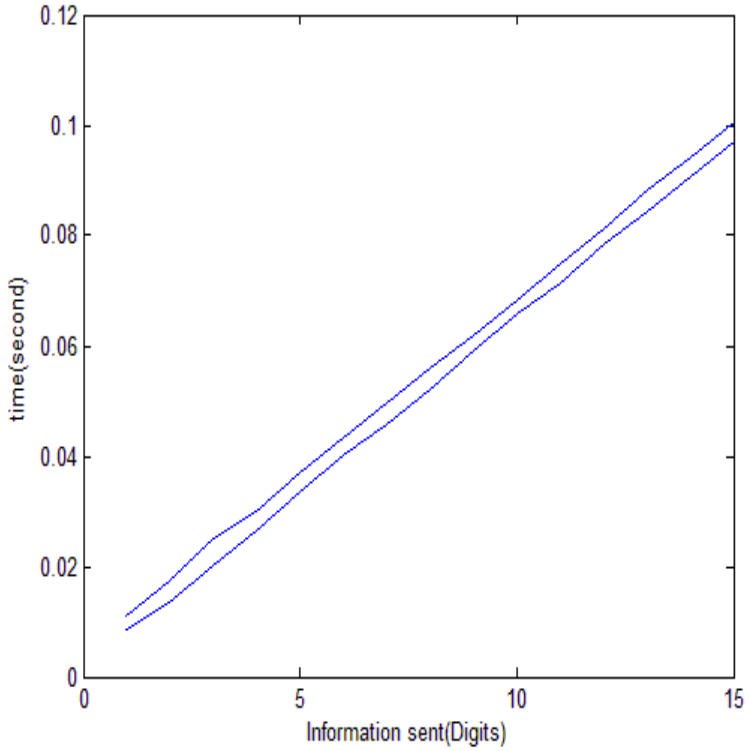

Fig.18 Message length vs. Time for p=3 and q=7

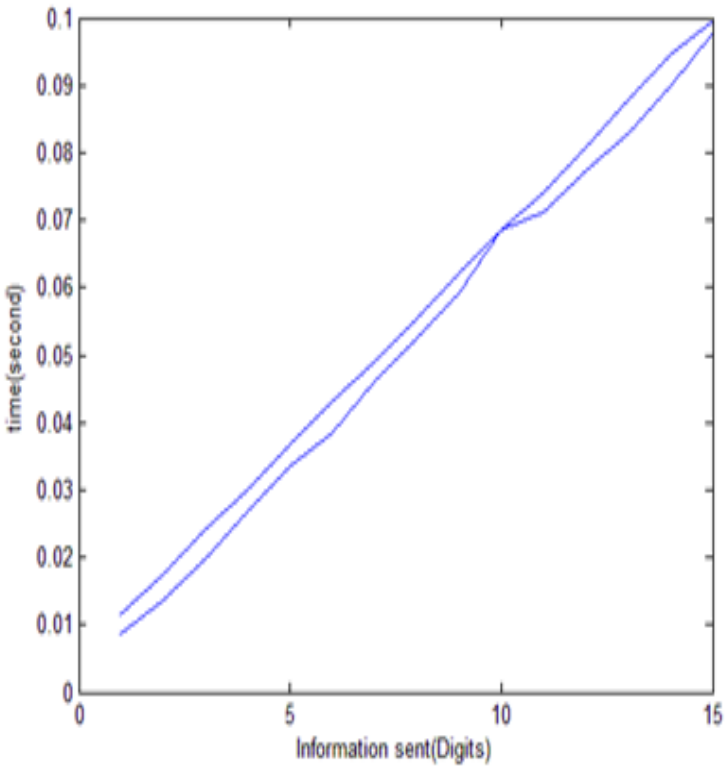

Fig.19 Message length vs. Time for p=23 and q=17

- AES ALGORITHM ANALYSIS

  AES is used here to provide integrity to data while simple browsing of internet. Plain text is encrypted to hexa decimal format. The change in graph depends on the value of plain text. The time taken increases if there is a use of combination of text and digits. A graph between information sent and time is plotted which can be seen in the figure 20.

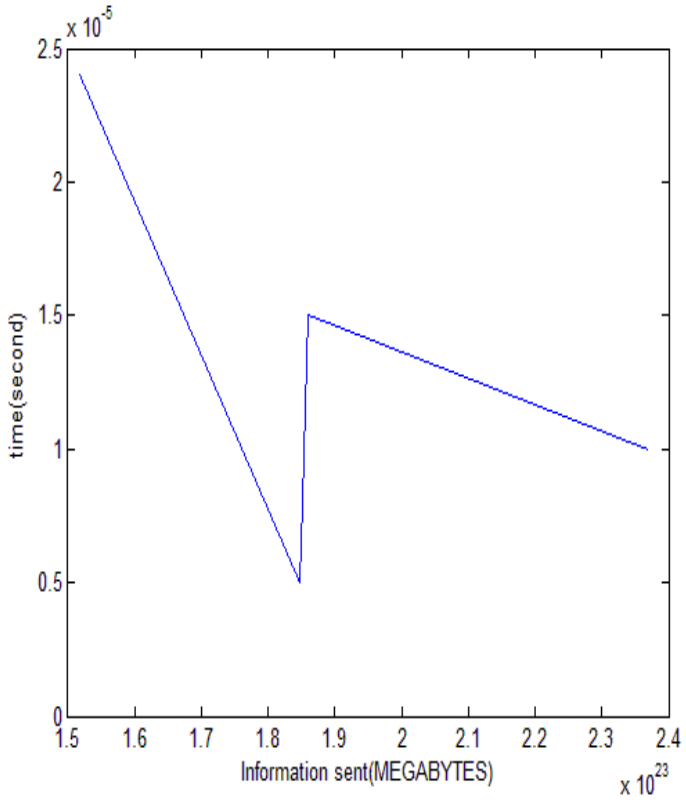

Fig.20 AES algorithm

4.2.2 Analysis of framework

Cloudsim is installed on netbeans and a GUI is built as shown in figure 21 that is used to create different datacenters, cloudlets and virtual machines. The snapshots of the creation of all the above mentioned components is shown below in figure 22, figure 23 and figure 24.

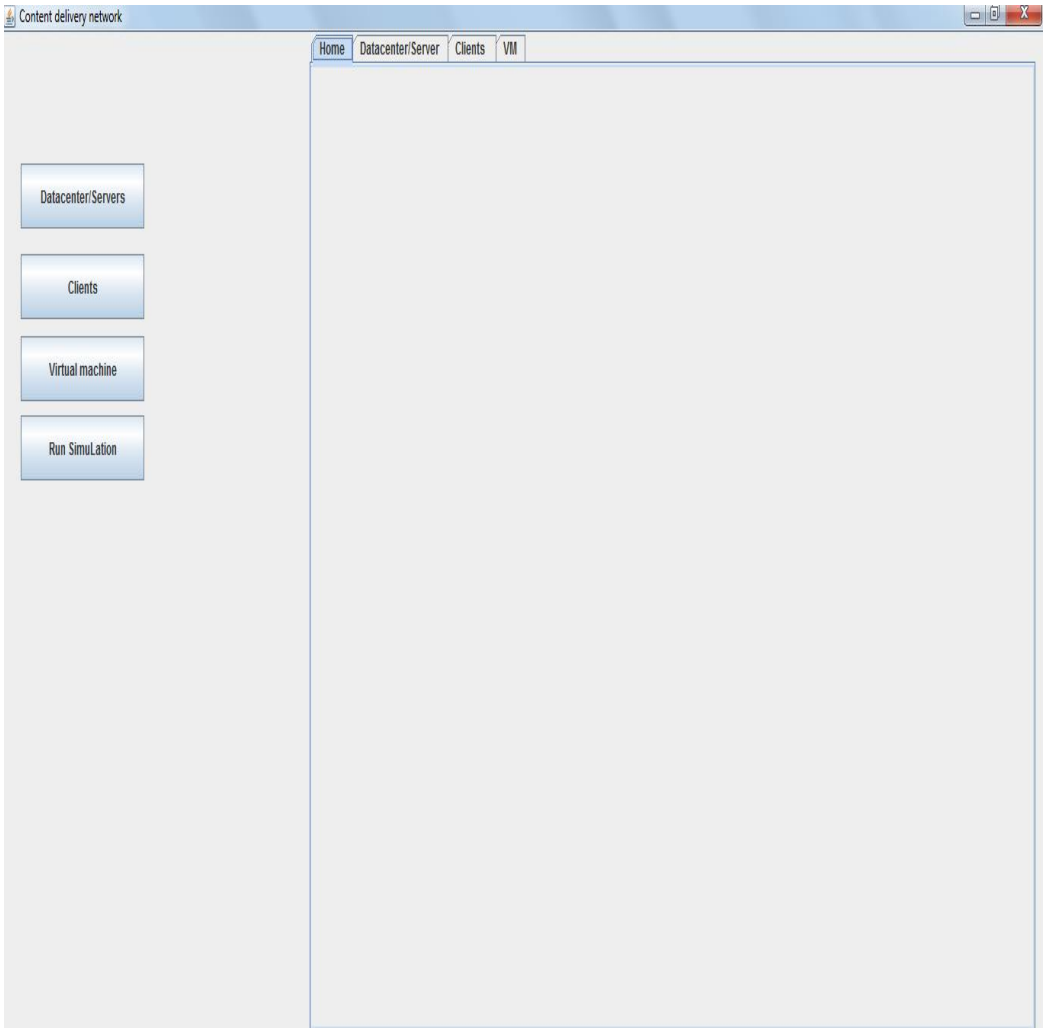
Content delivery network
Home
Datacenter/Server
Clients
VM
Datacenter/Servers
Clients
Virtual machine
Run SimuLation

Fig.21 GUI

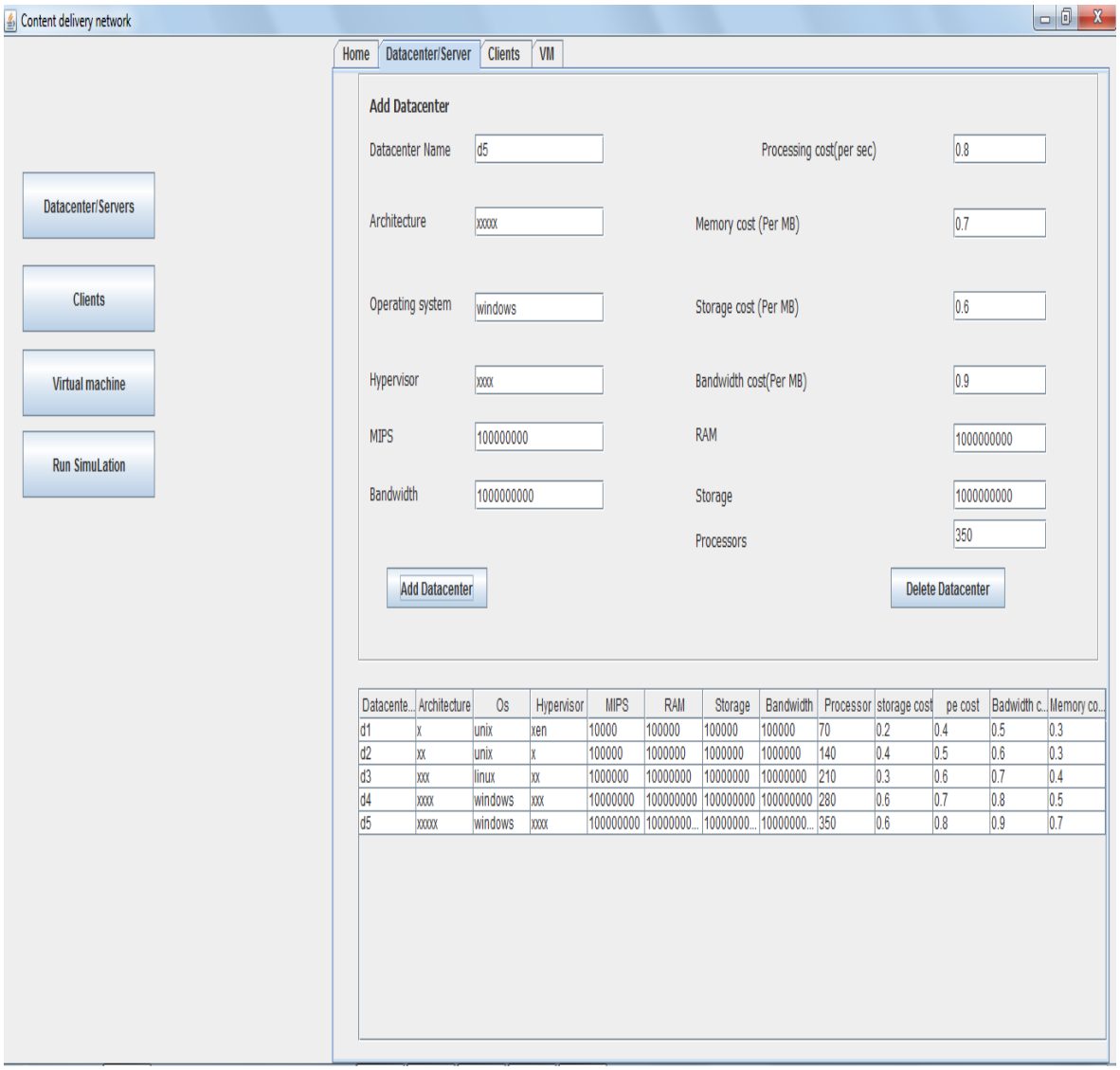

Fig.22 Creating datacenters

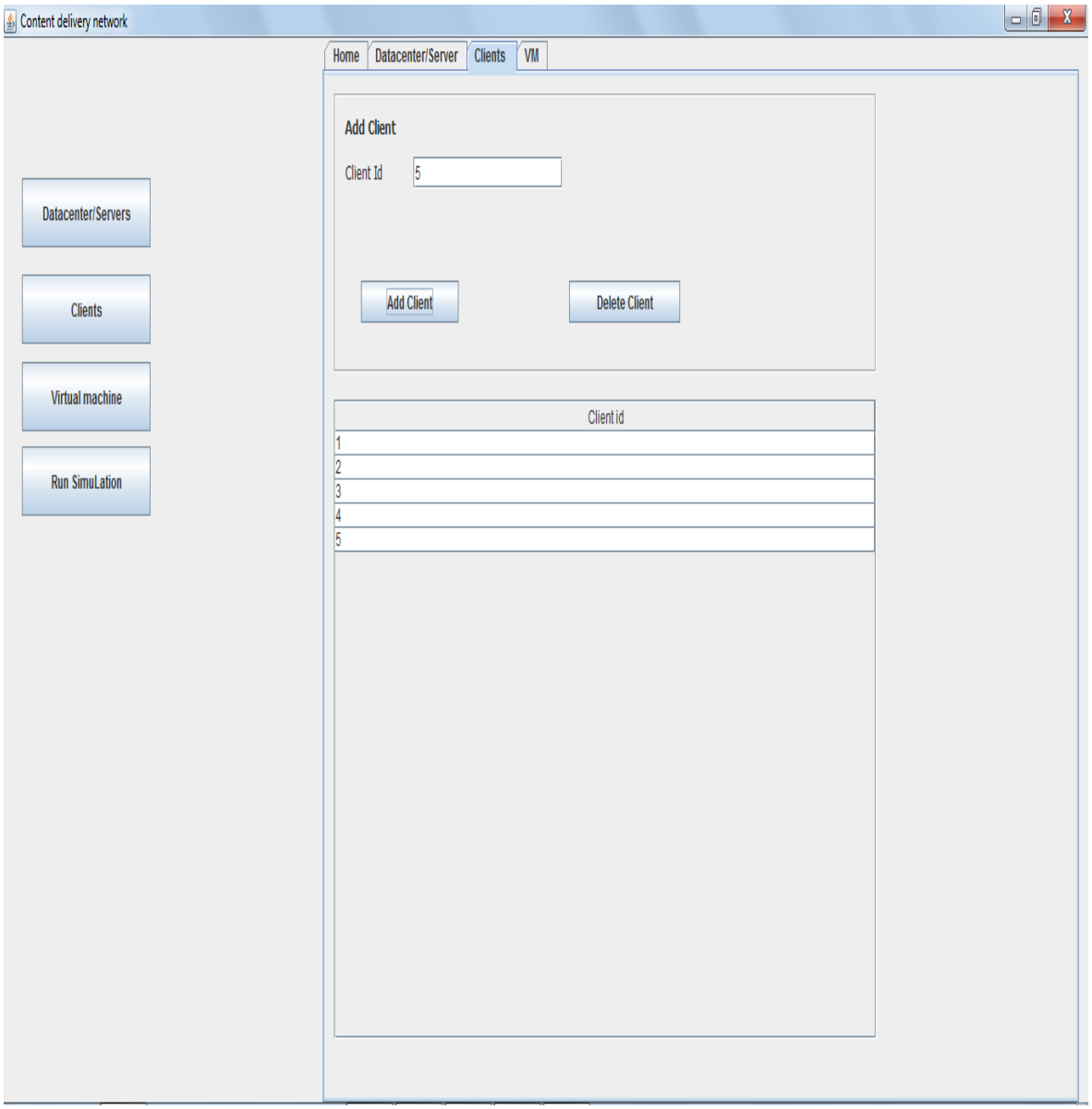

Fig.23 Creating different clients

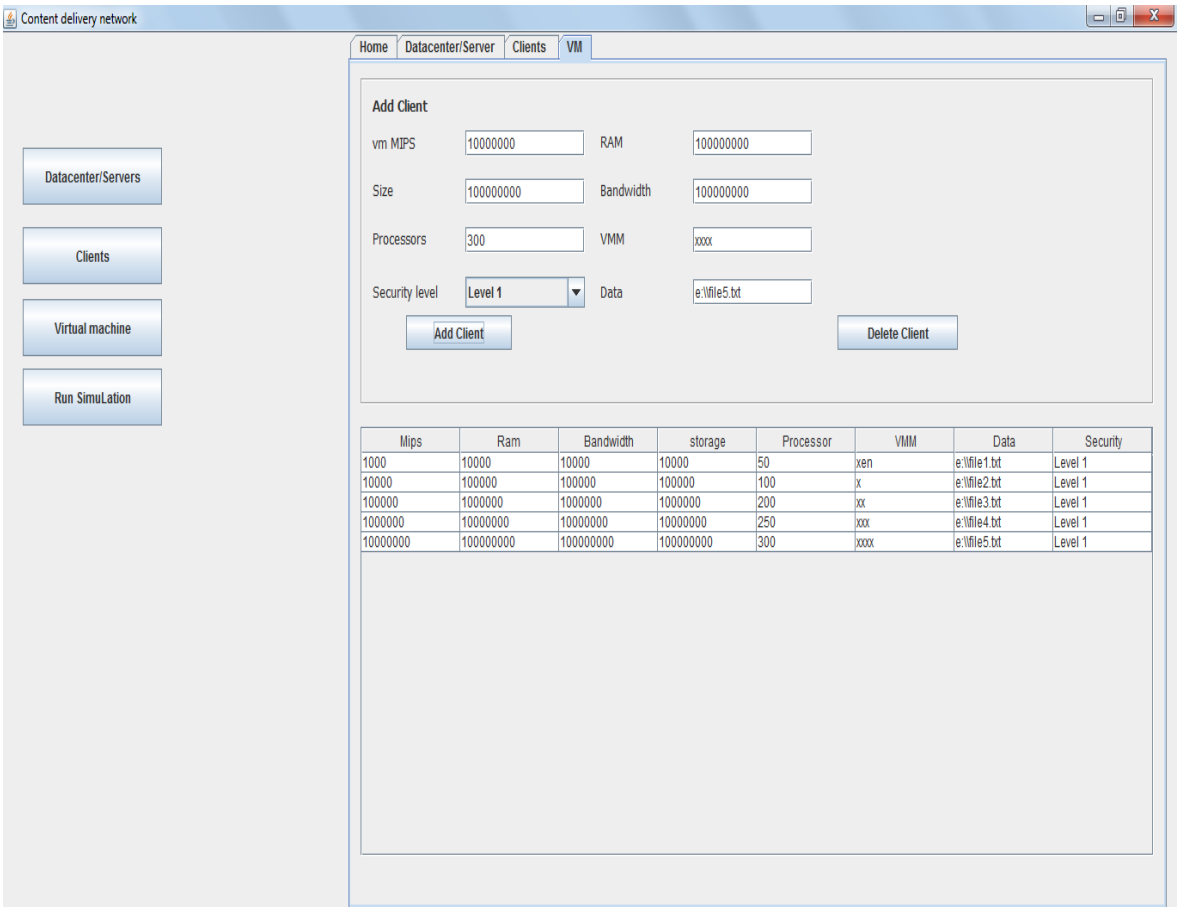

Fig.24 Virtual machines created

On clicking on Run Similation button, simulation is started. The datacenters are created and resources/ virtual machines are allocated to the datacenters. Figure 25 and figure 26 depicts the simulation.

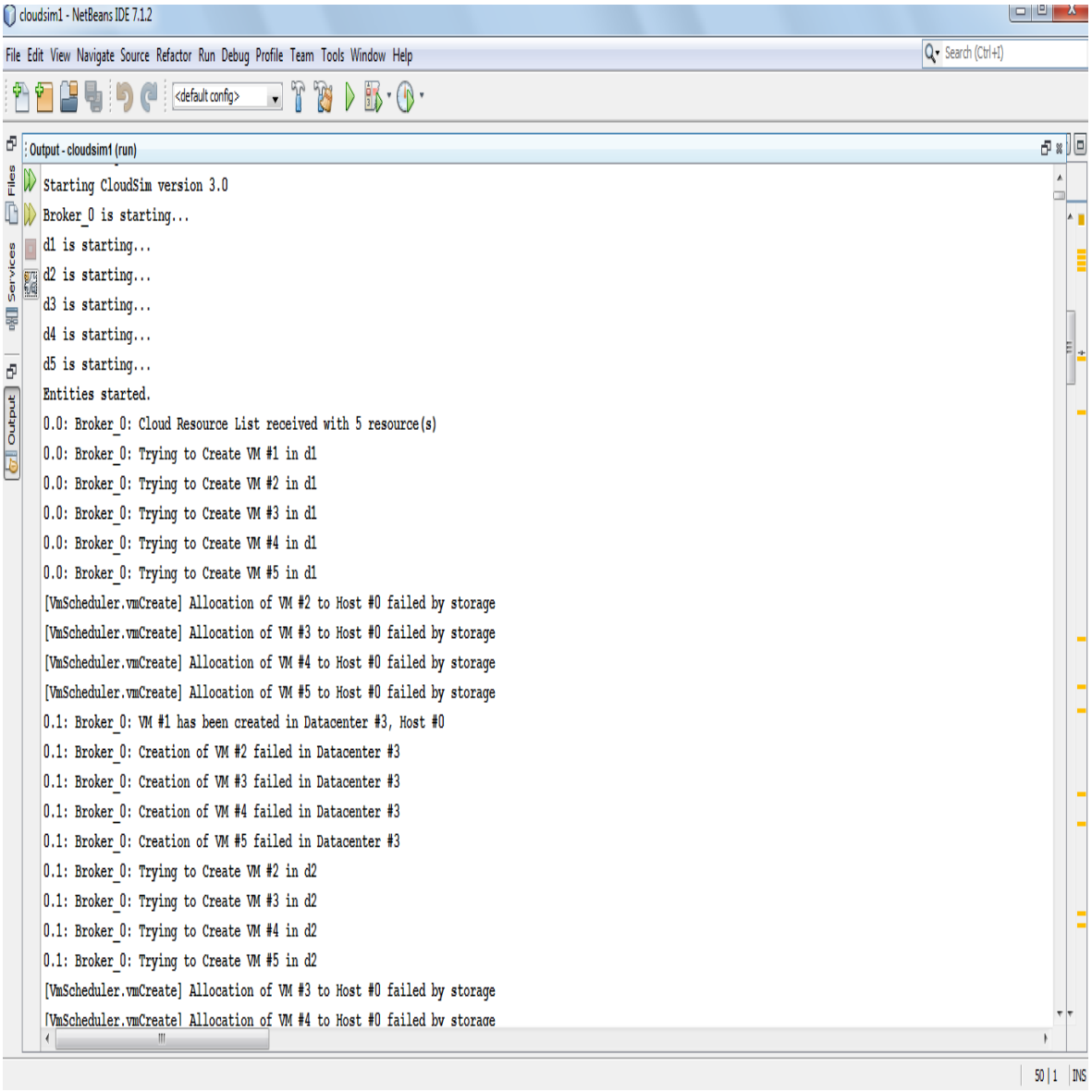

Fig.25 Simulation started

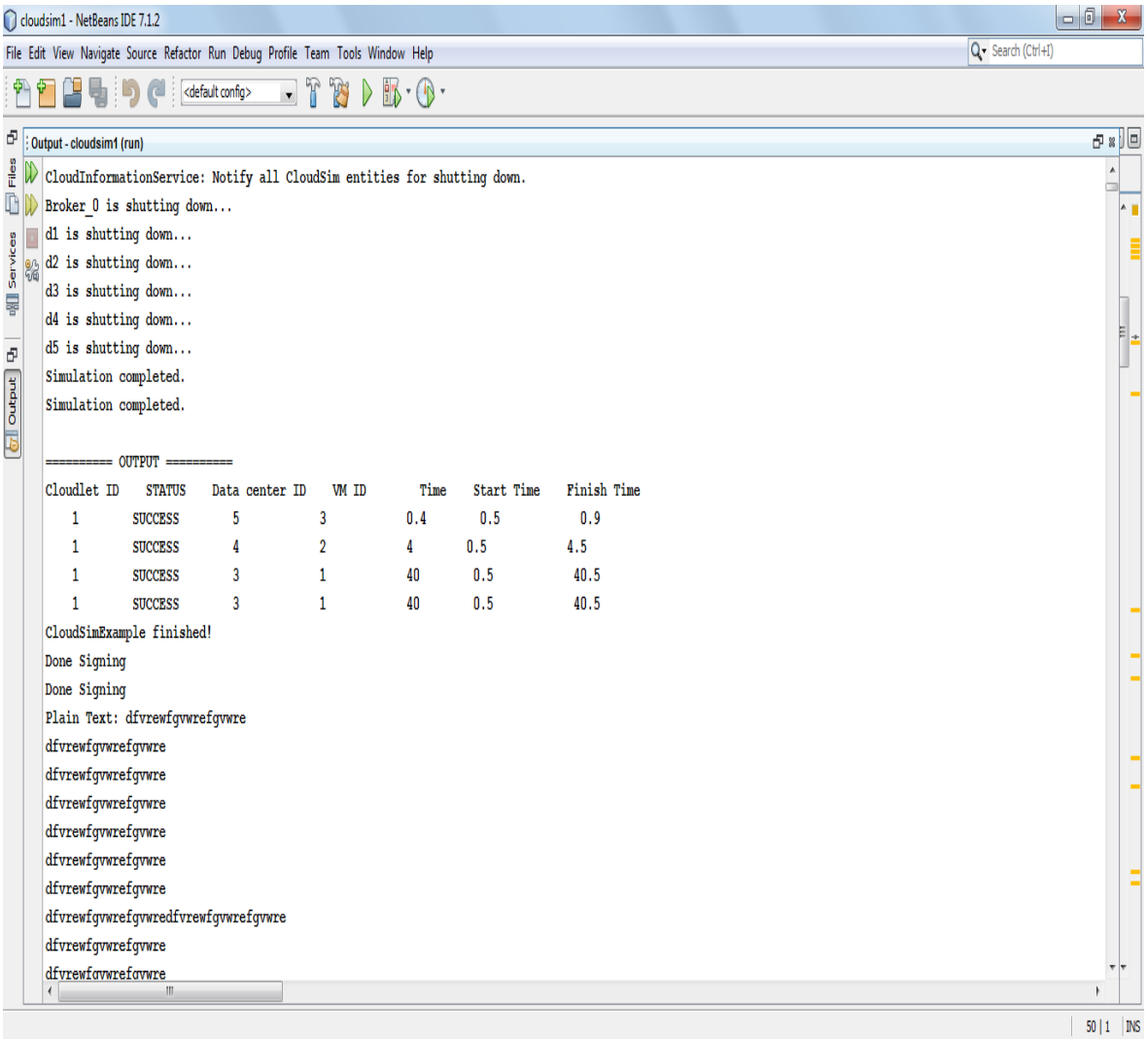

Fig.26 Simulation completed

The data in all three files is encrypted using RSA algorithm by selecting security level 1. Integrity of the files is checked by using hashing algorithm, where firstly the files are signed and then they are verified by using the secret keys. Since the files are not modified data is decrypted and the execution time for both encryption and decryption process is calculated as shown in figure 27.

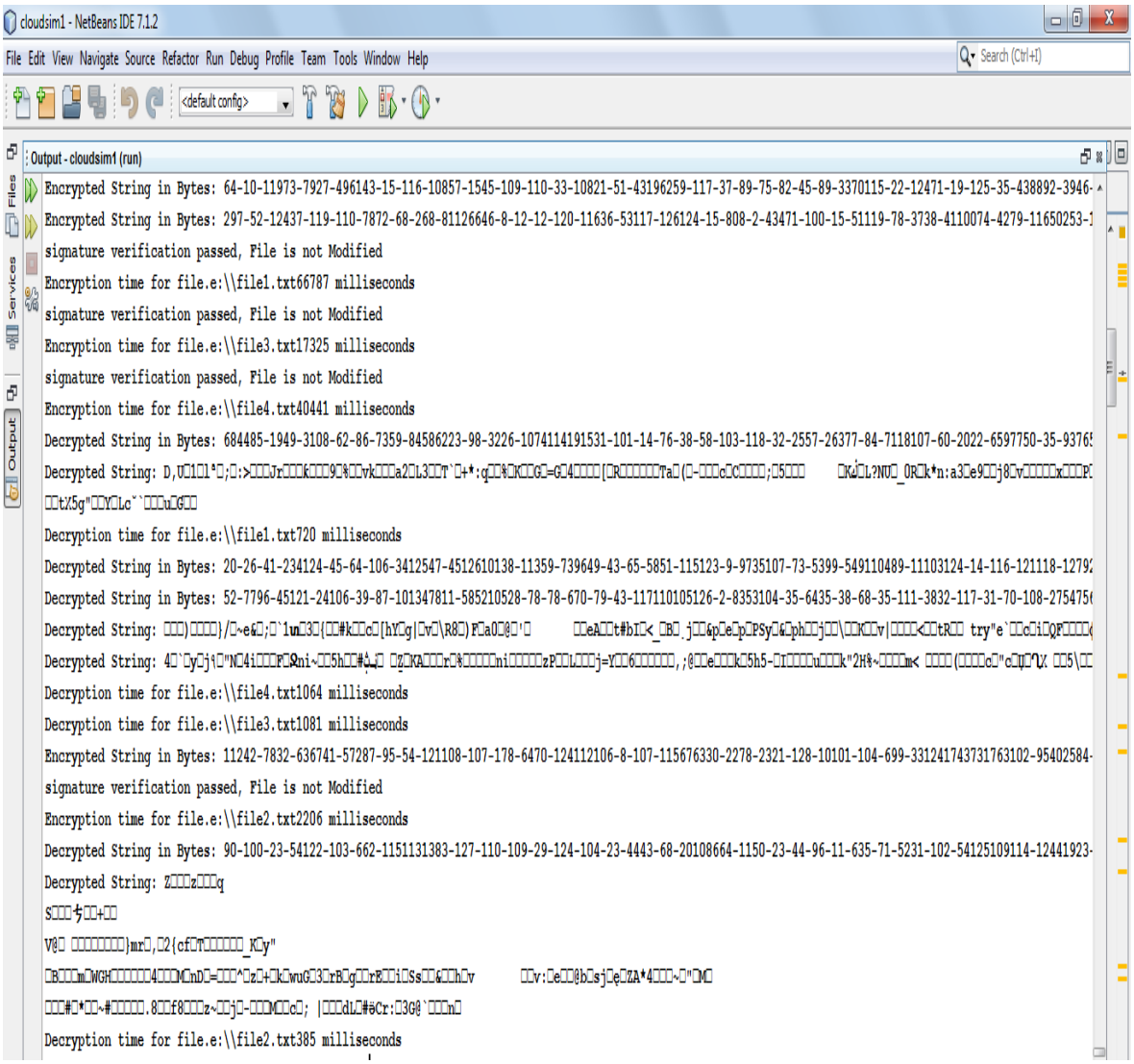

Fig.27 The encryption and decryption time of file 1, file 2, file 3 and file 4.

Now the same process is followed on the files but security level 2 is applied where the algorithm used is Bcrypt. The results generated after simulation are depicted in figure 28, figure 29 and figure 30.

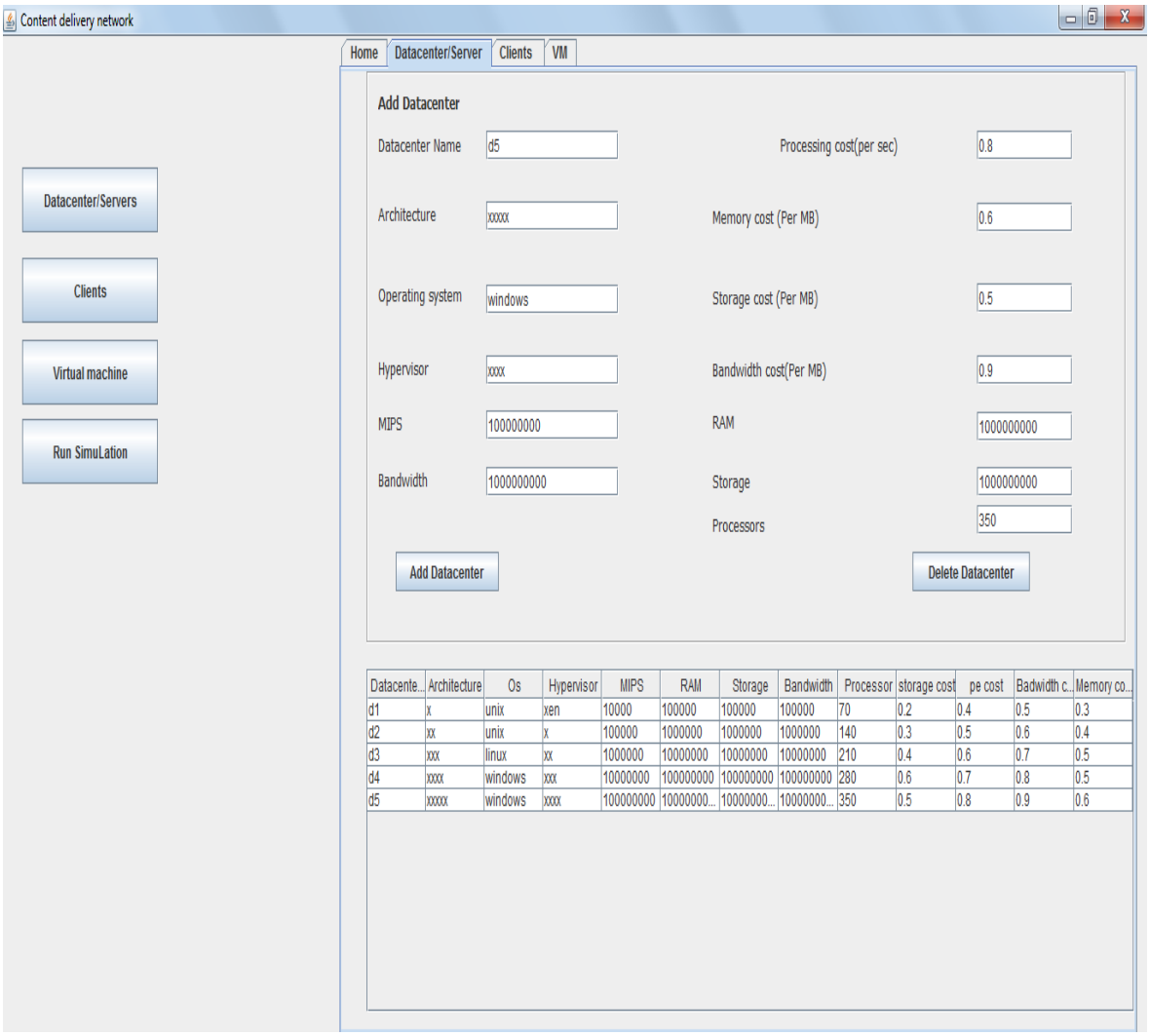

Fig.28 Creating datacenters for bcrypt method

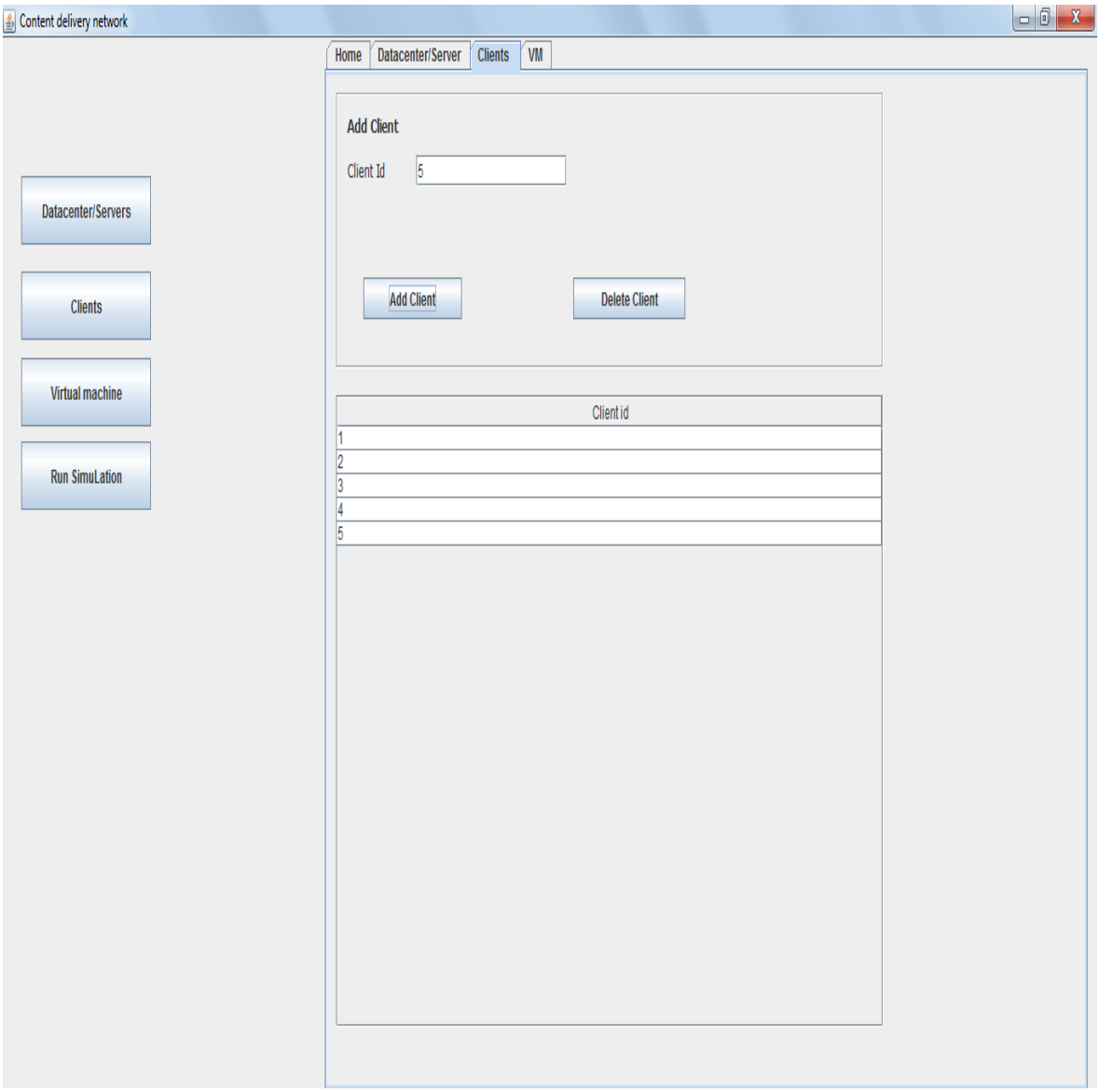

Fig.29 Clients are created

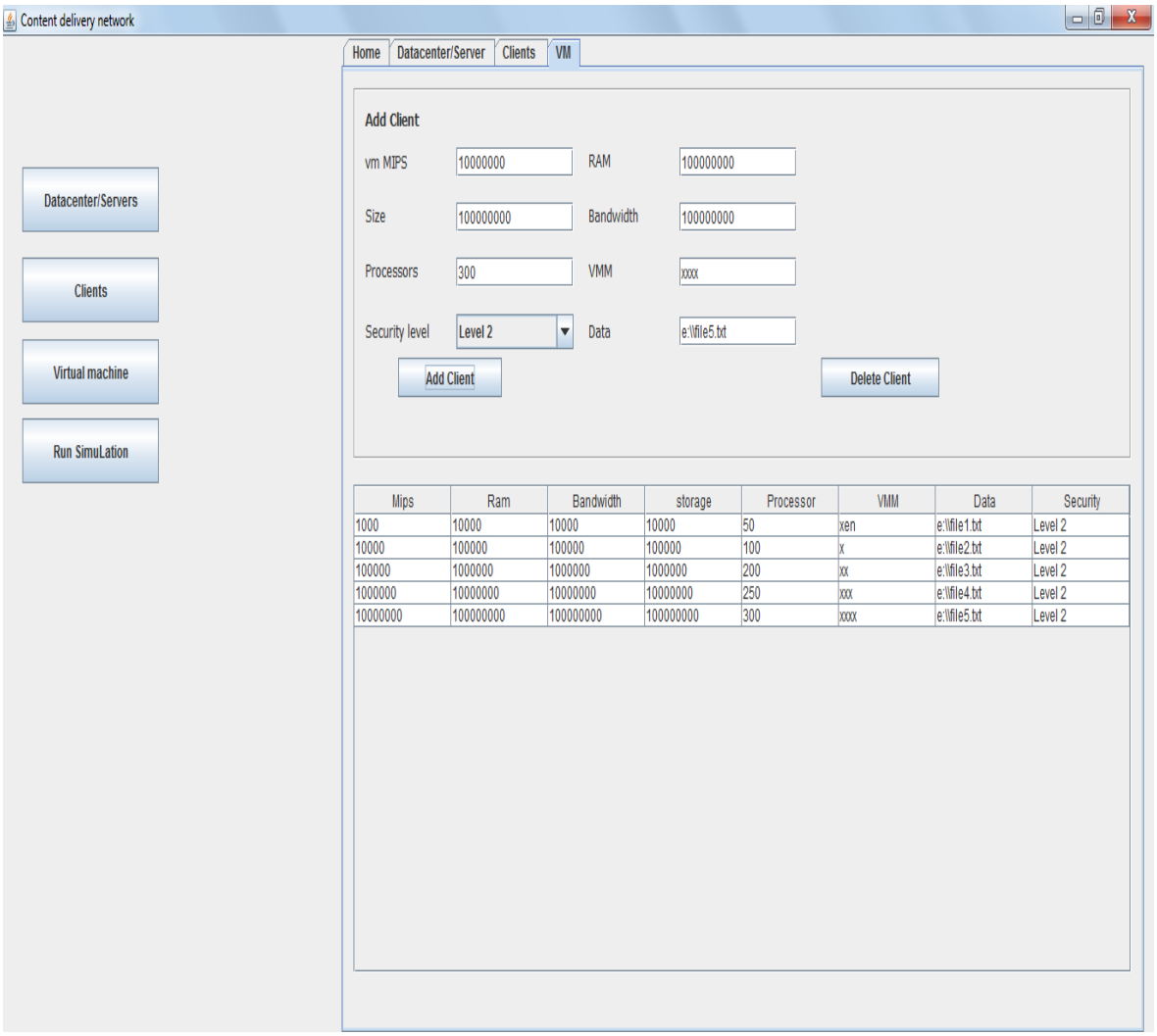

Fig.30 Virtual machine created for bcrypt method

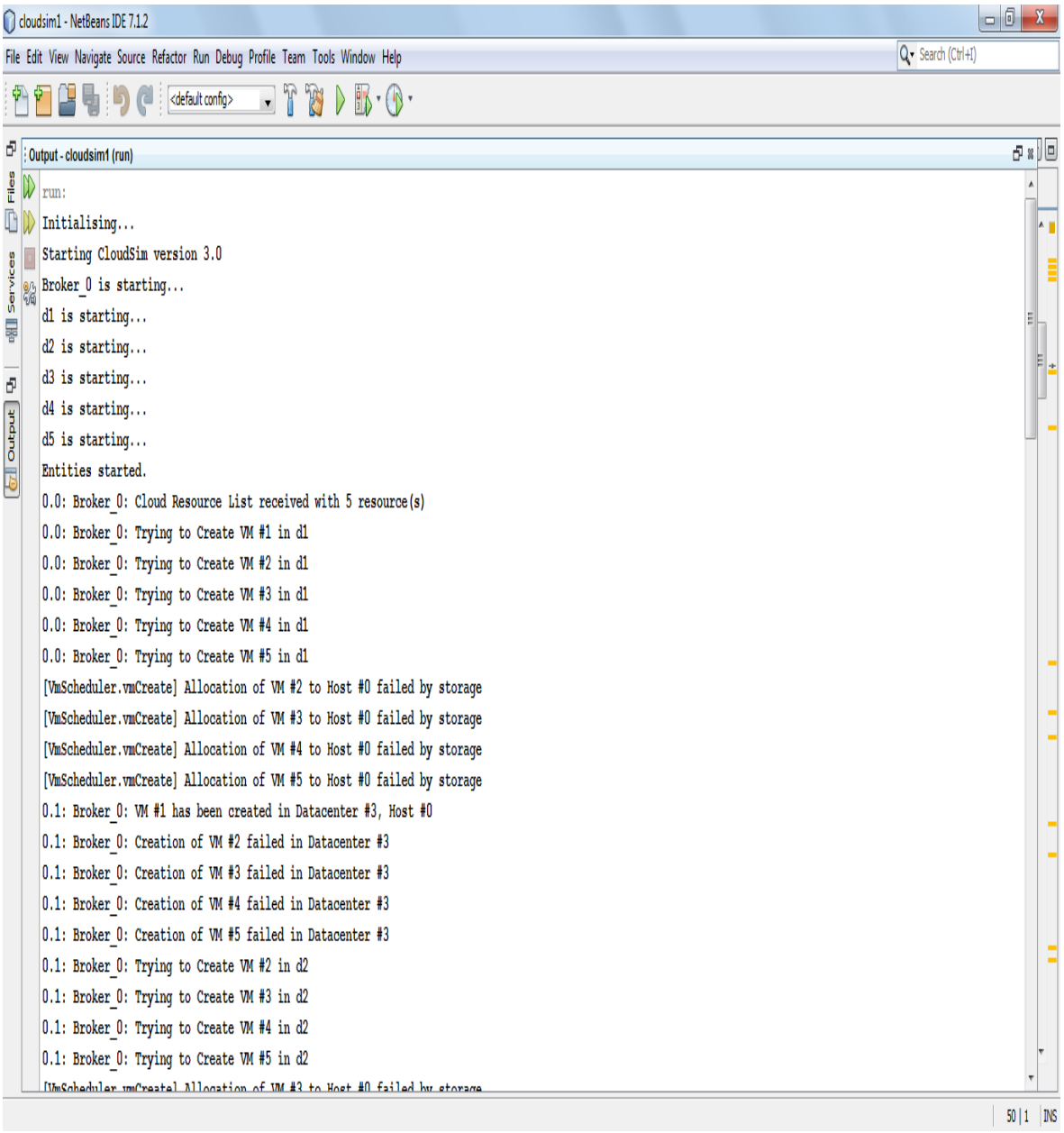

Fig.31 Simulation started

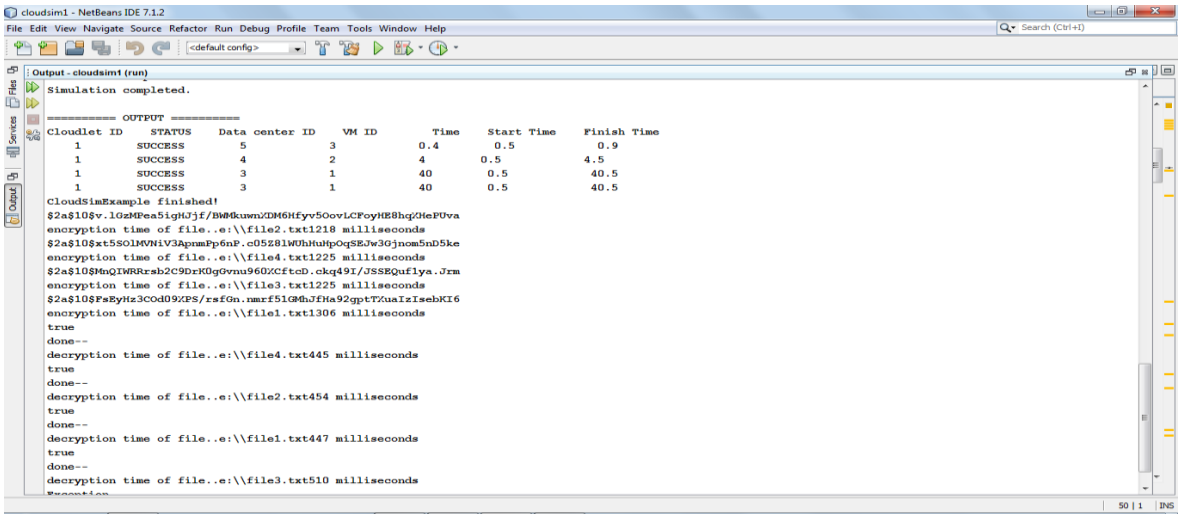

Fig.32 Simulation completed

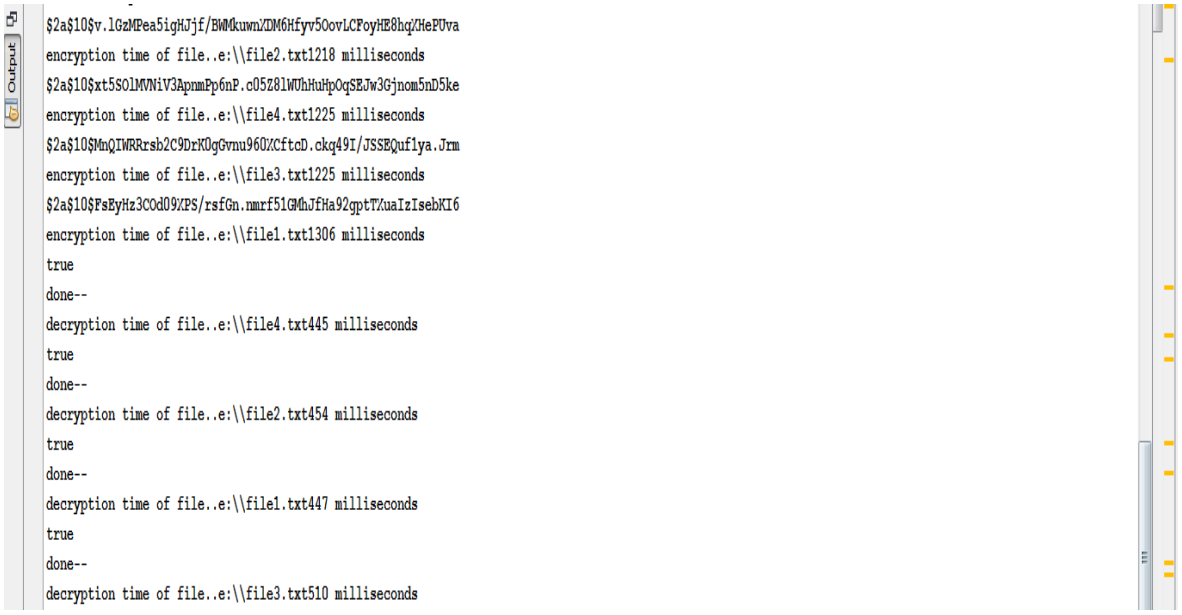

Fig.33 Encryption and decryption of all three files along with time taken to perform both these actions

Now the data in files is encrypted and decrypted using AES cryptographic algortihm, the result of which are shown in figure 34, figure 35, figure 36, figure 37 and figure 38.

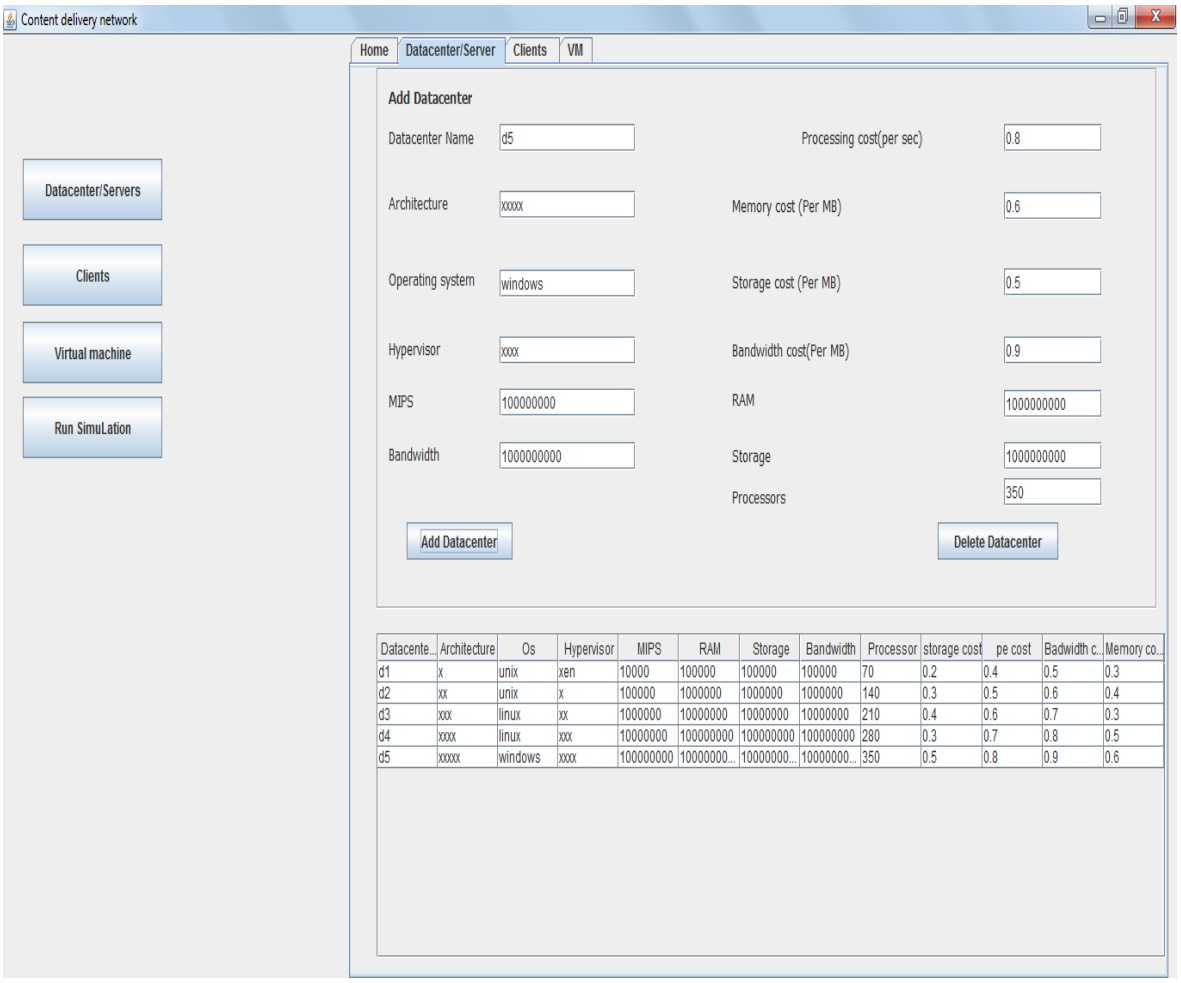

Fig.34 Datacenters created for AES method

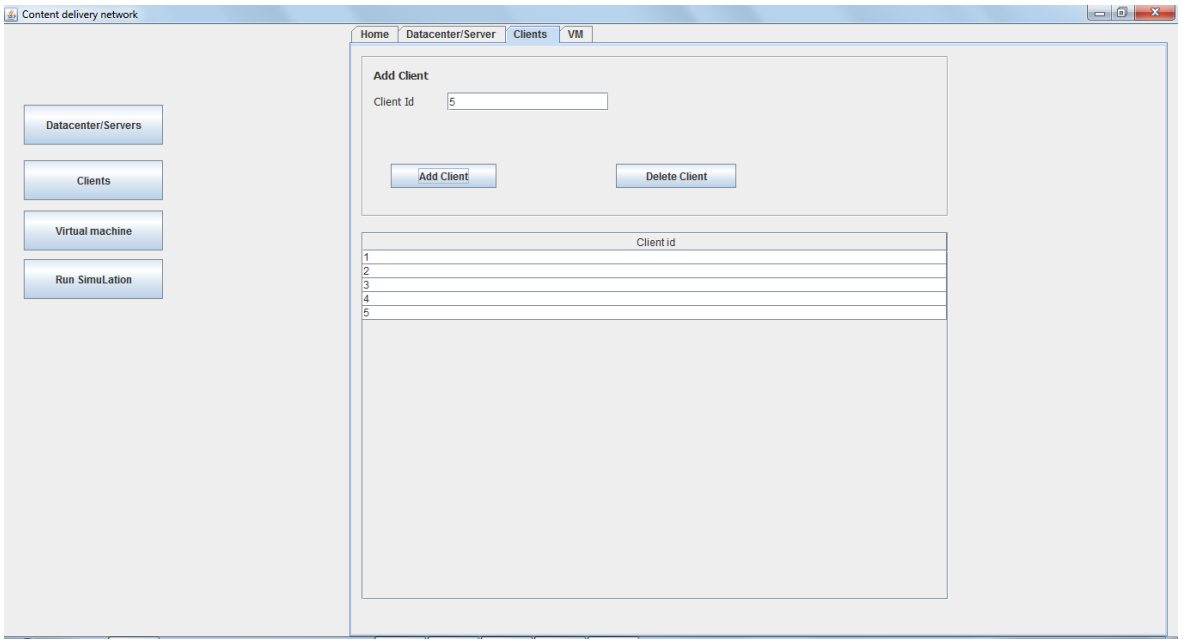

Fig.35 Creating cloudlets

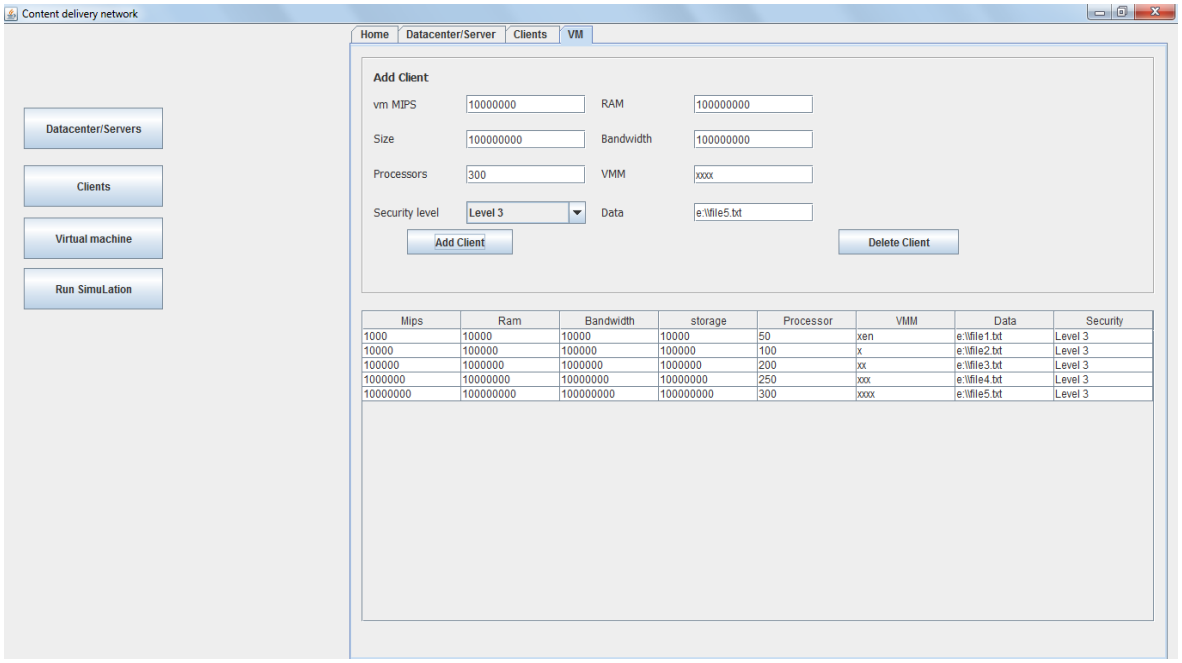

Fig.36 Creating virtual machines for AES method

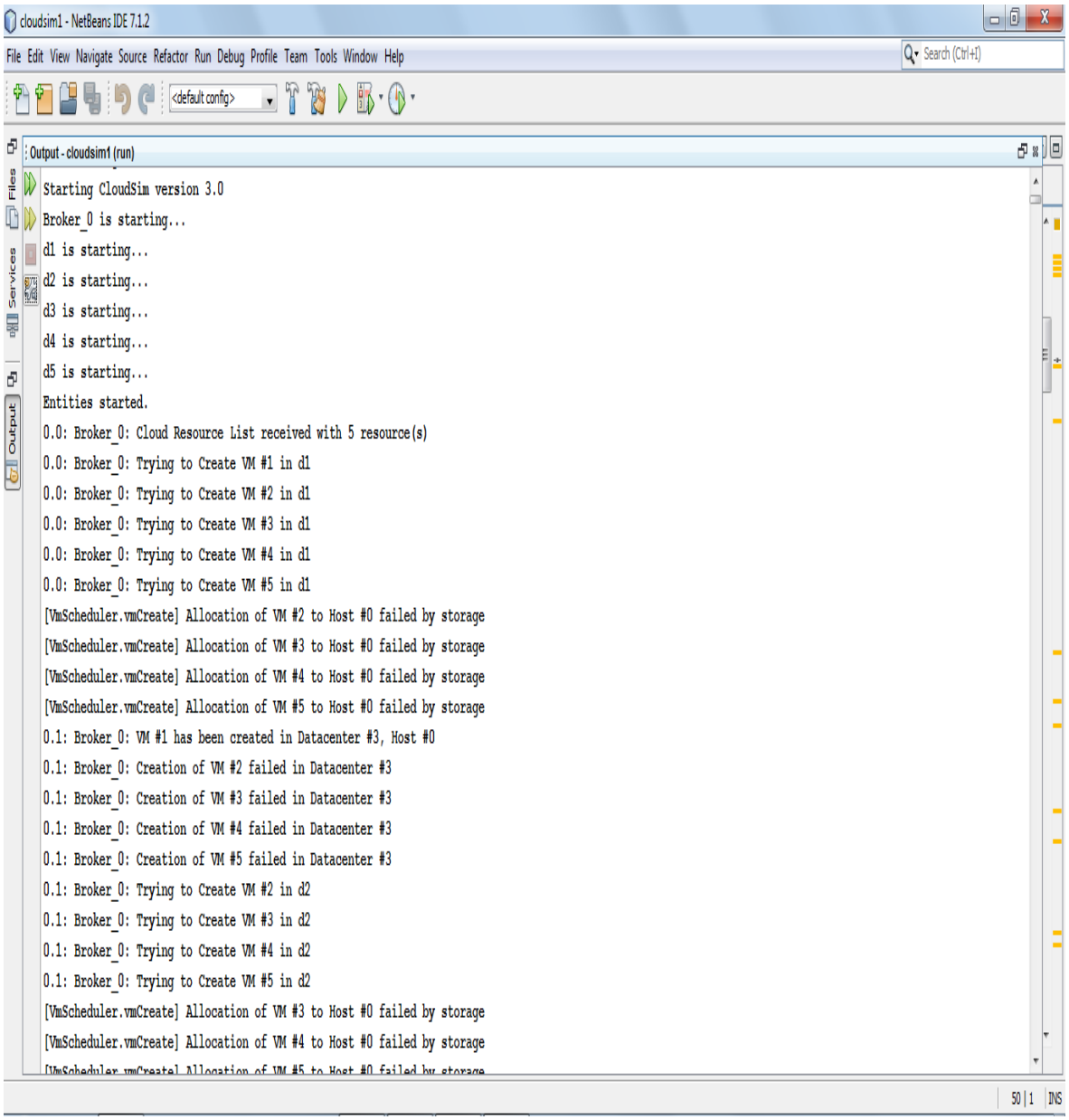

Fig.37 Simulation started

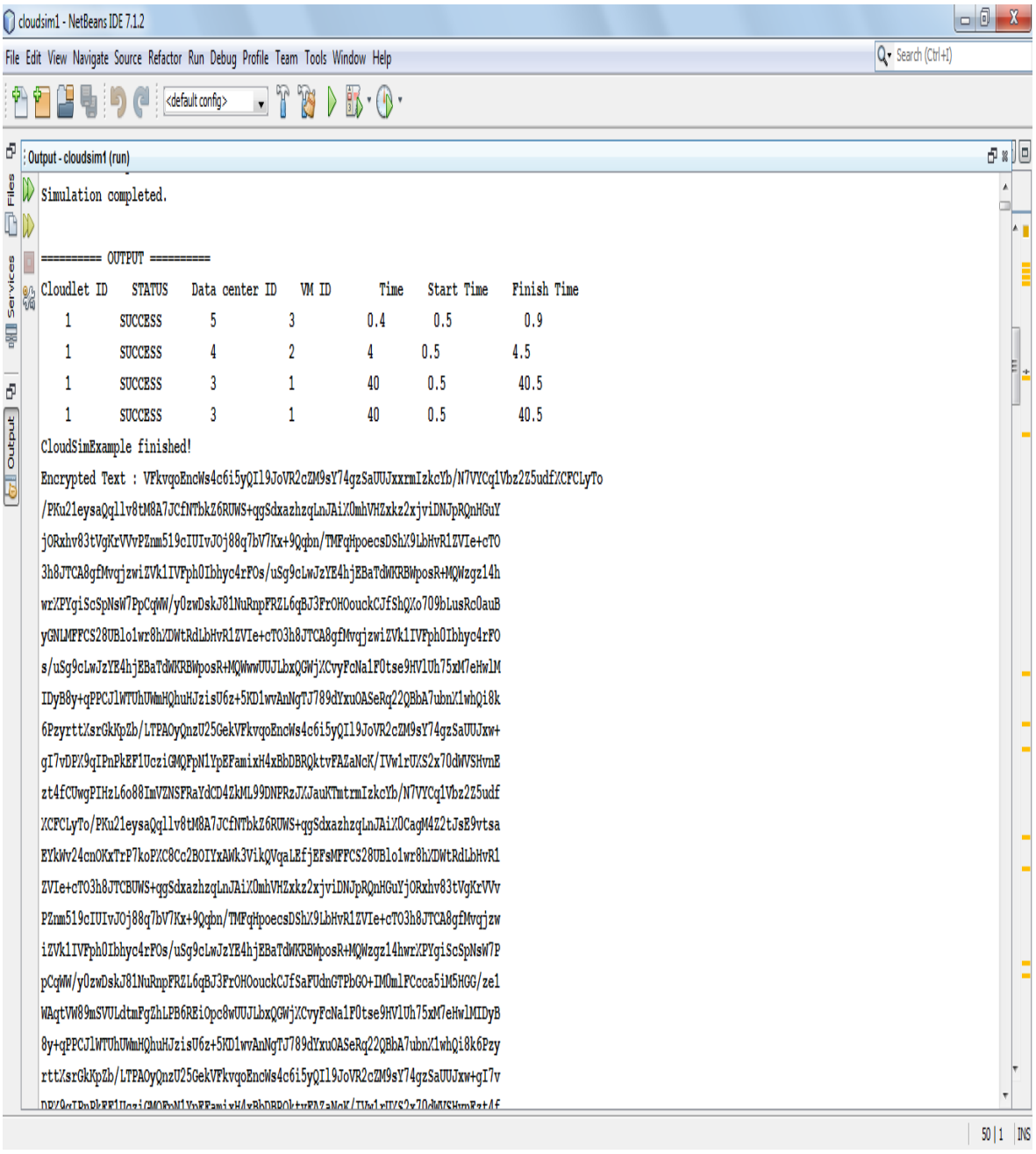

Fig.38 Simulation finished

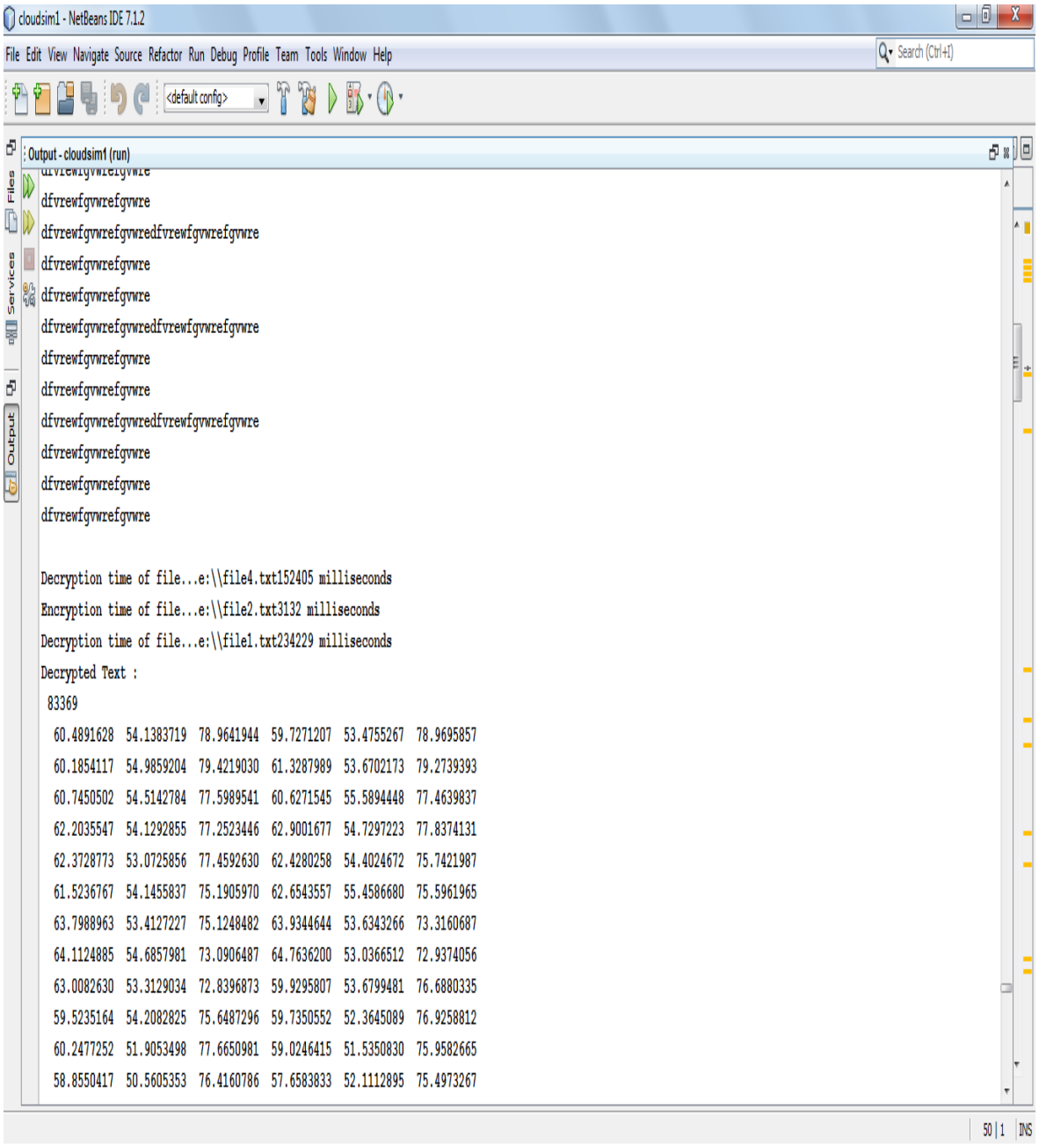

Fig.39 Encryption and decryption of files along with time of execuion of both methods

Now by using the proposed framework multi level security is provided on different files based on the type of information that has to be stored.

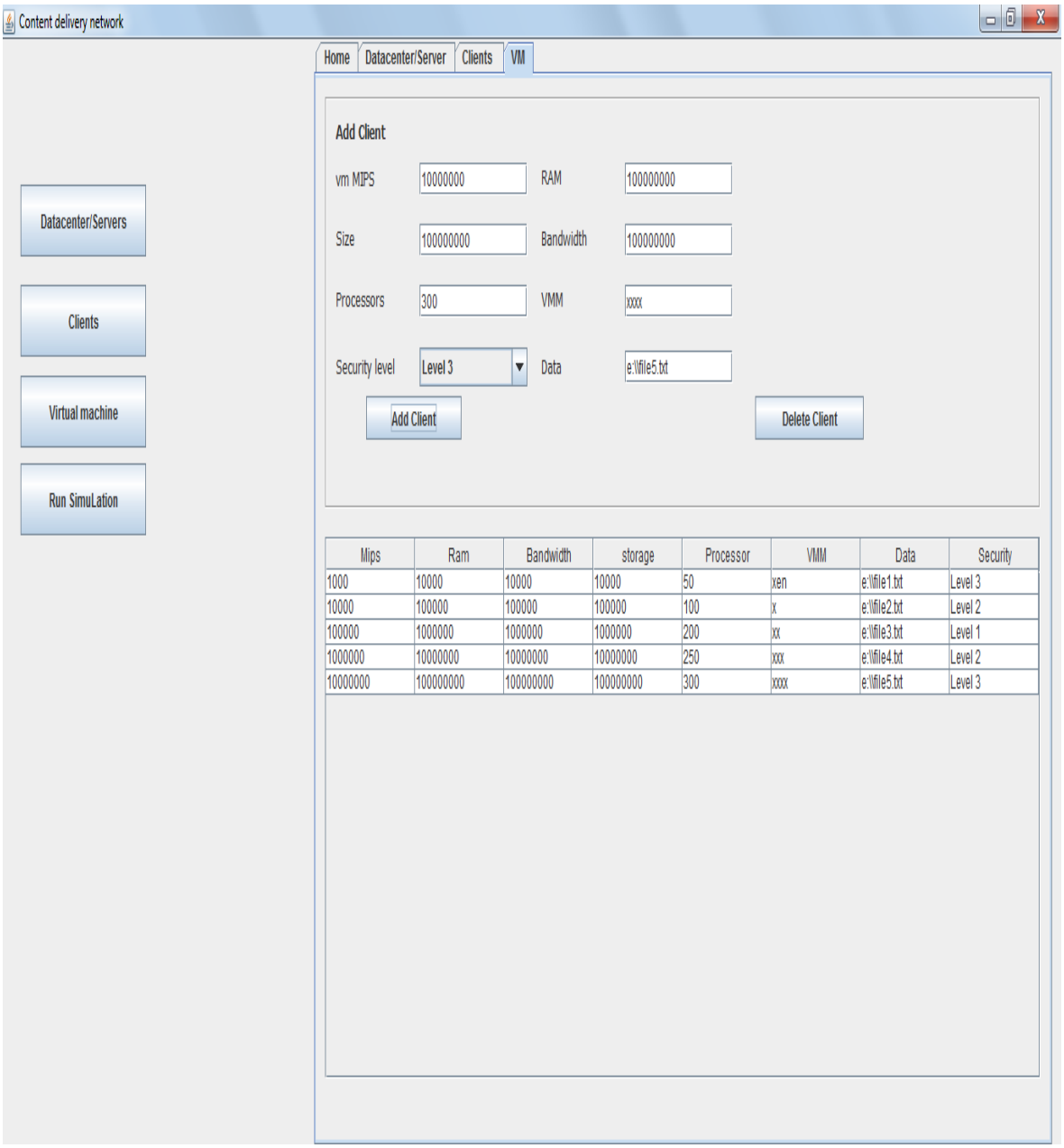

Fig.40 Virtual machines created with different level of security

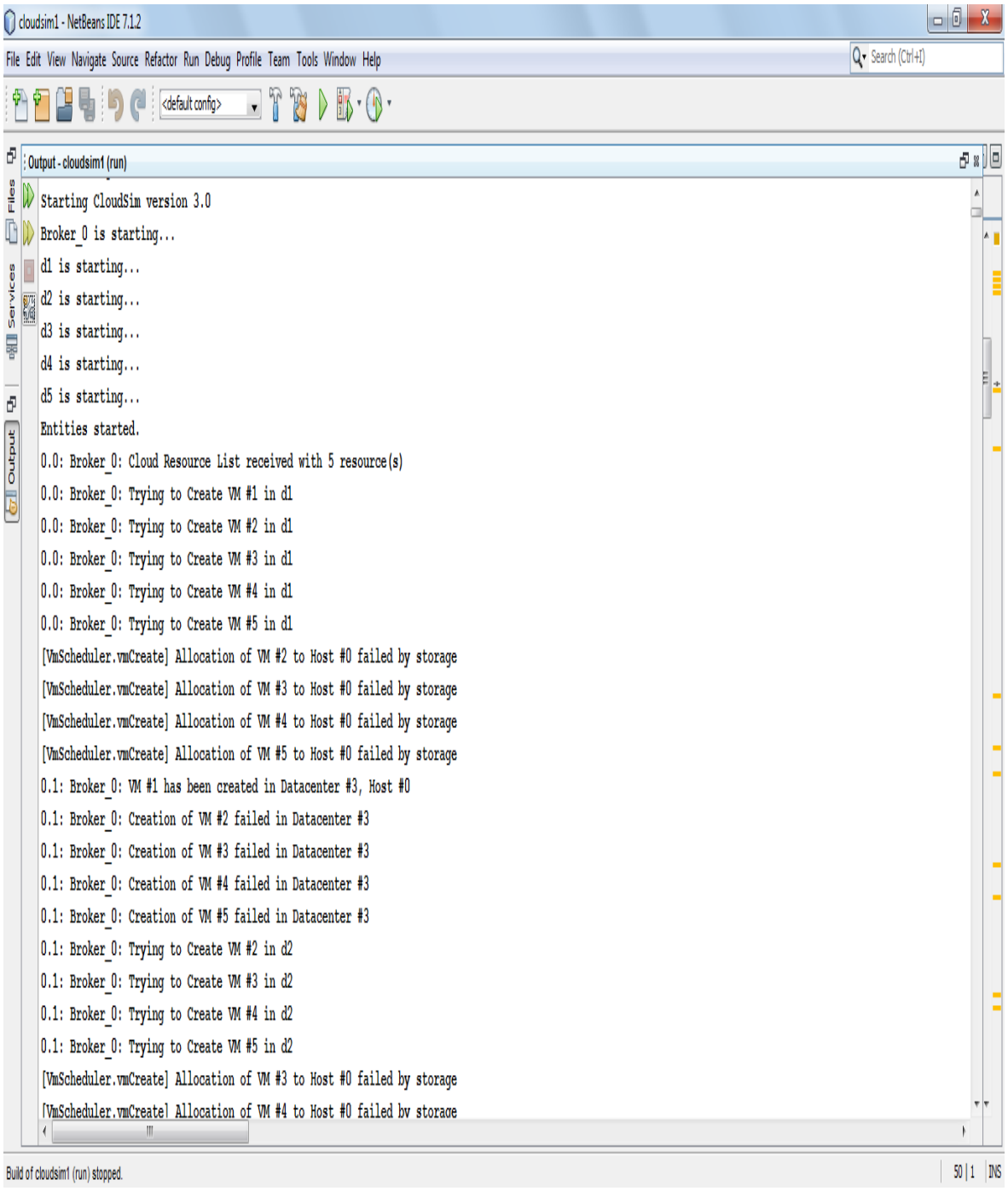

Fig.41 Simulation started

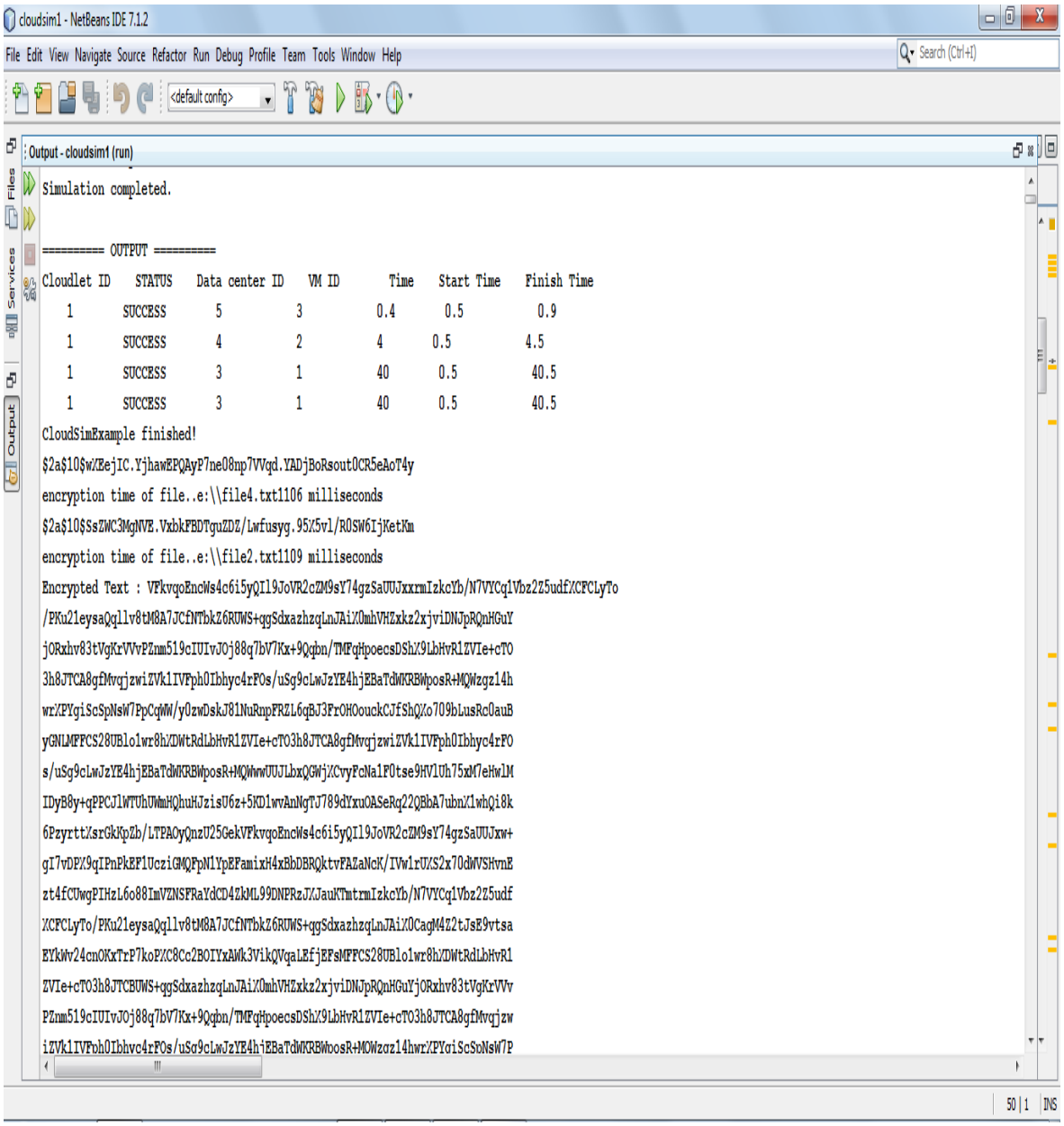

Fig.42 Simulation completed

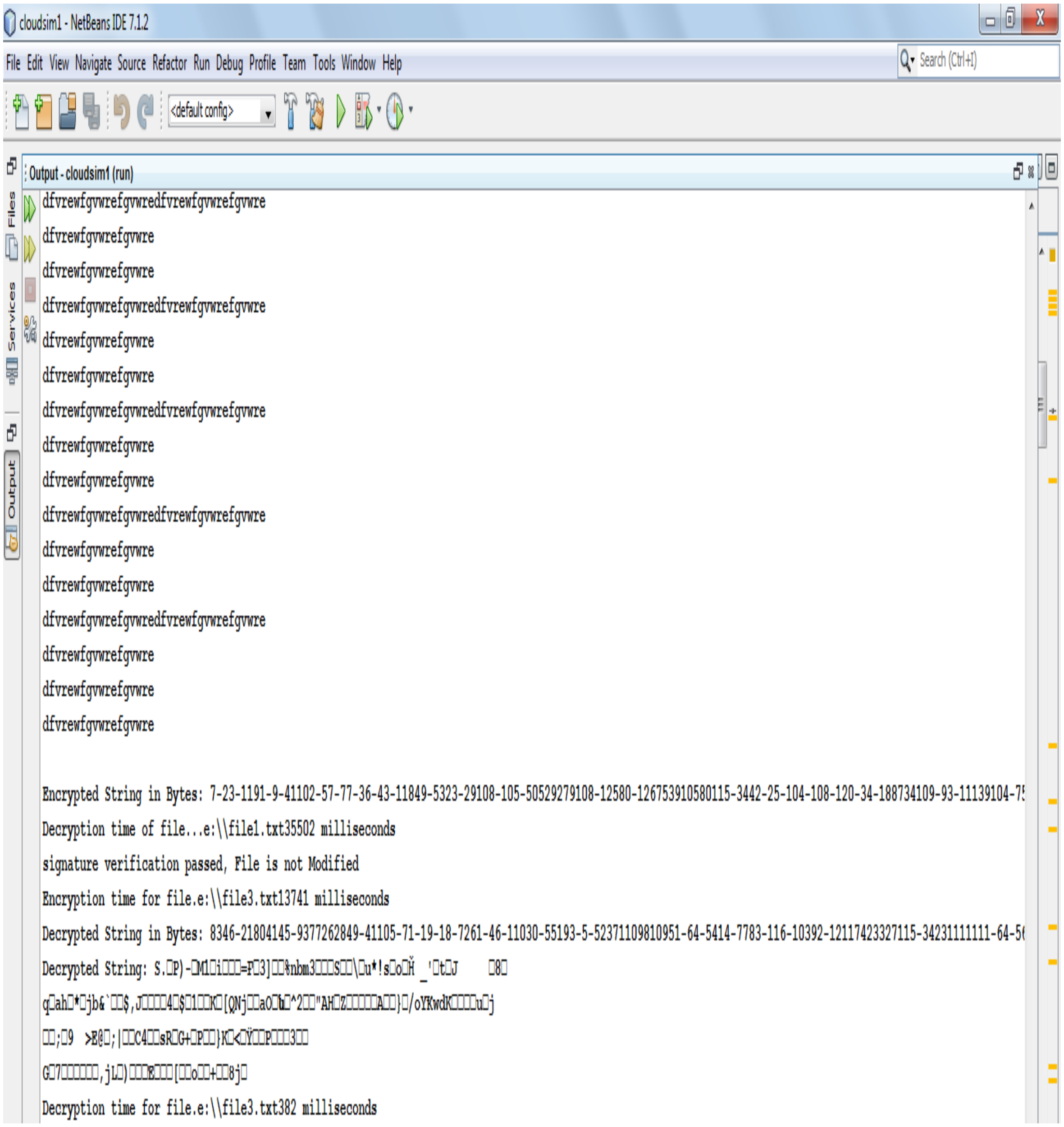

Fig.43 Encryption and decryption of files with different level of security applied

An evaluation of old scheme and proposed one is done by calculating the encryption time for different files. The analysis is shown in figure 44, 45, 46, 47.

Table 5,6,7 and 8 display the execution time of different files using both schemes when the number of client requests have been increased

Table 5 Performance analysis of old scheme v/s proposed framework

REQUESTS	PREVIOUS SCHEME (TIME in ms)	PROPOSED SCHEME (TIME in ms)
1	3074	3074
2	3074	2917
3	3075	796
4	3009	1108
5	2995	1108
6	2905	796
7	3079	921
8	3420	827

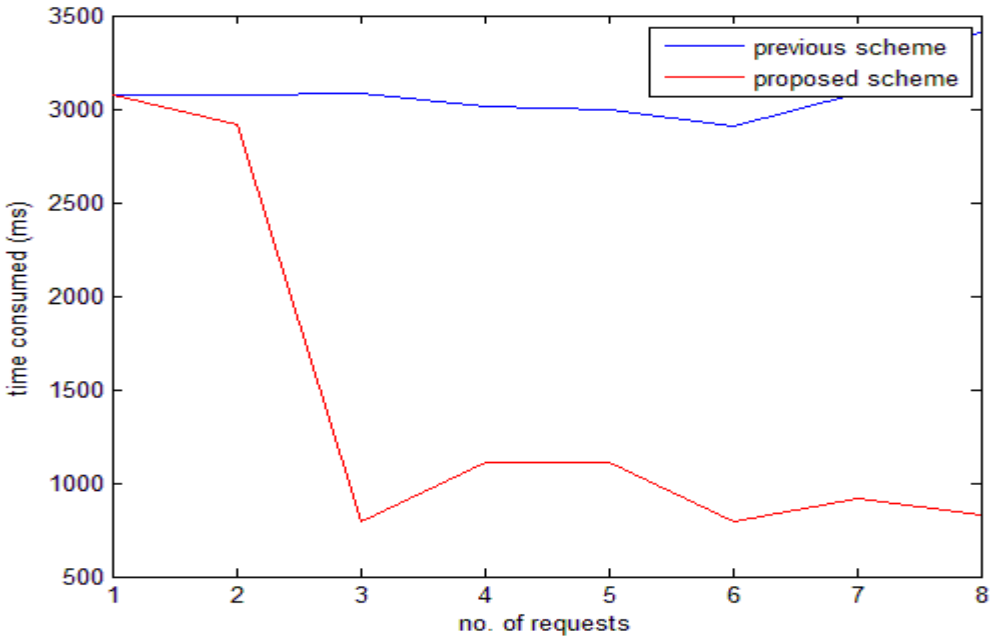

Fig.44 The encryption time of previous and proposed scheme when number of requests are 8

Table 6 Performance analysis of previous scheme v/s proposed scheme when number of requests increased to 16

REQUESTS	PREVIOUS SCHEME (TIME in ms)	PROPOSED SCHEME (TIME in ms)
1	3248	3006
2	3291	3110
3	3106	900
4	3341	1249
5	3570	1289
6	3320	1160
7	3334	1276
8	3456	900
9	3418	1000
10	3534	958
11	4018	2987
12	3987	3540
13	3765	3491
14	4211	4010
15	4300	2667
16	4696	3800

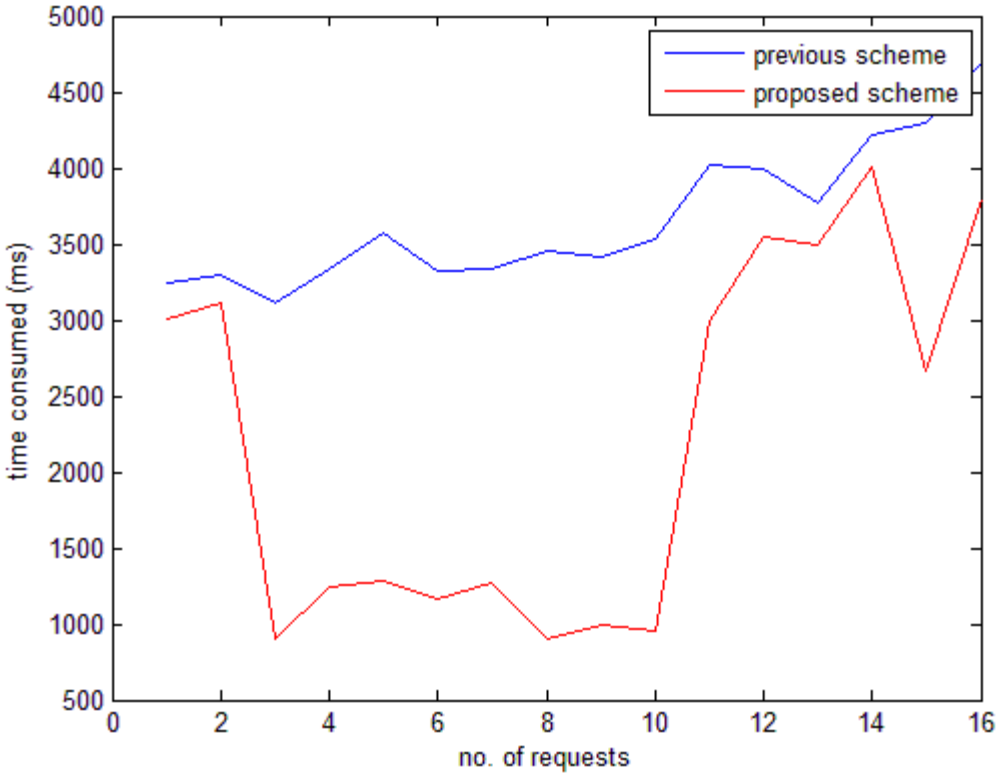

Fig.45 Encryption time of previous and proposed scheme when number of requests are 16

Table 7 Performance analysis of previous scheme v/s proposed scheme when number of requests increased to 32

REQUESTS	PREVIOUS SCHEME (TIME in ms)	PROPOSED SCHEME (TIME in ms)
1	3896	3694
2	3424	3341
3	3308	3208
4	3410	3400
5	3540	3459
6	3691	3576
7	3330	3200
8	3120	3110

9	3600	3567
10	3498	1000
11	3900	3741
12	4009	4003
13	3774	3669
14	4219	4198
15	4166	4047
16	4333	4216
17	3681	3538
18	3908	3497
19	3724	3623
20	3805	3551
21	3500	3348
22	4460	4234
23	4000	3811
24	4790	4626
25	4546	4377
26	3980	3694
27	2196	987
28	2897	1248
29	3071	2381
30	4298	3106
31	4571	3963
32	5168	4685

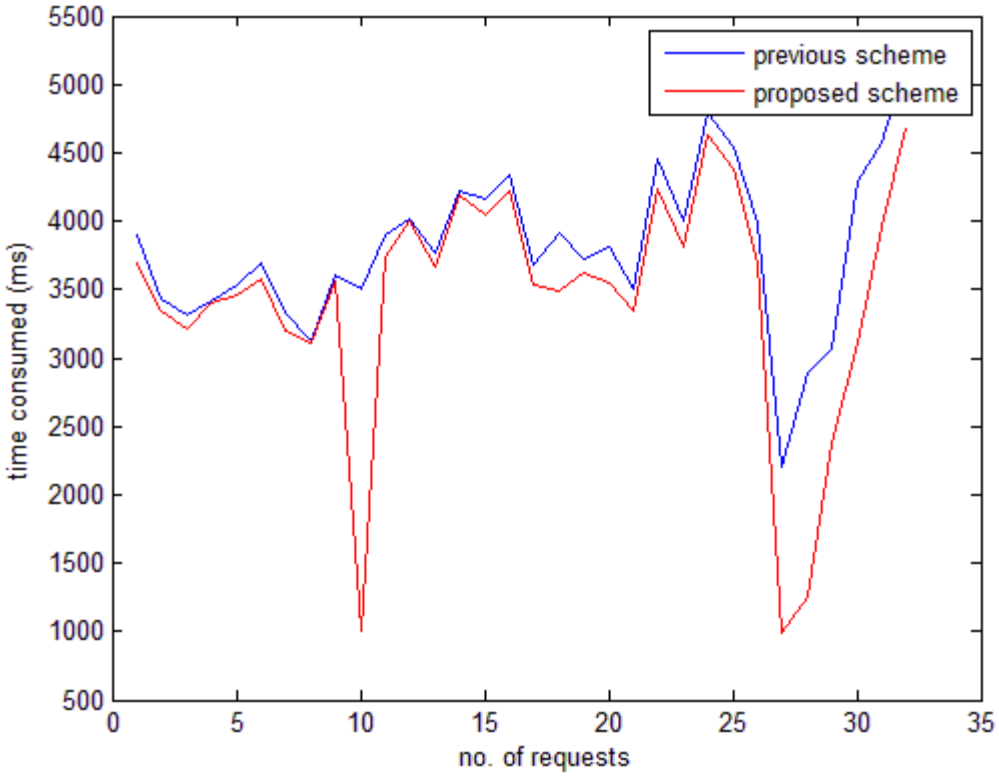

Fig.46 Encryption time of previous and proposed scheme when number of requests are 32

The average time taken in all three case is noted and performance is evaluated as shown in table 9 and figure

Table 8 Average time taken by both schemes when requests increased from 8 to 32

REQUESTS	PREVIOUS SCHEME AVG TIME (ms)	PROPOSED SCHEME AVG TIME (ms)
8	3105.91	1086.3
16	3870.19	2519.88
32	3899.72	3149.19

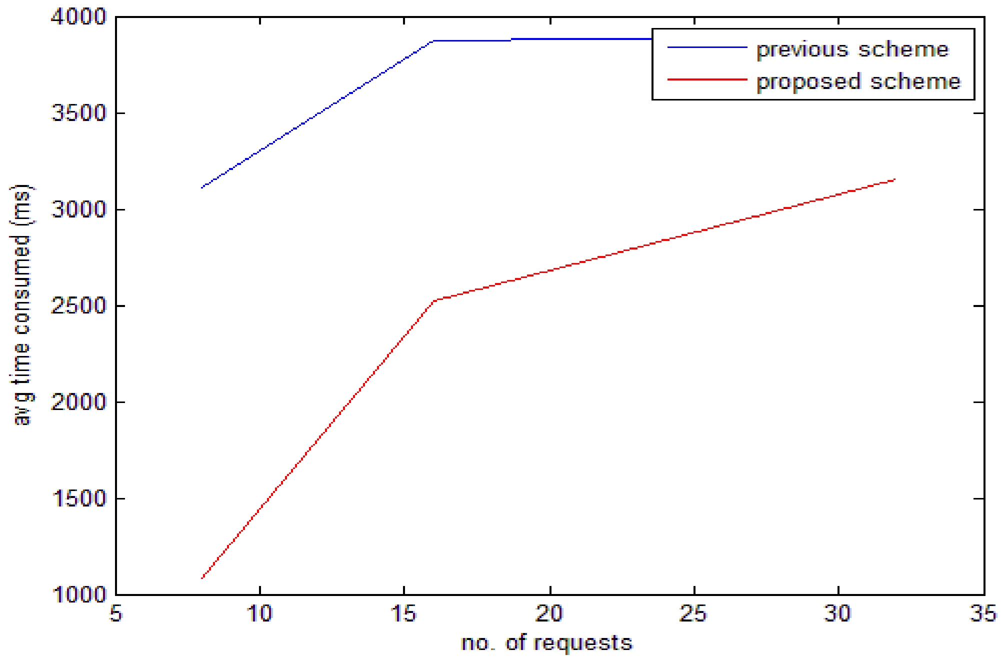

Fig.47 Average time consumed

If the file is modified then the data in file will get encrypted using RSA algorithm but it will not get decrypted as shown in figure 48 below

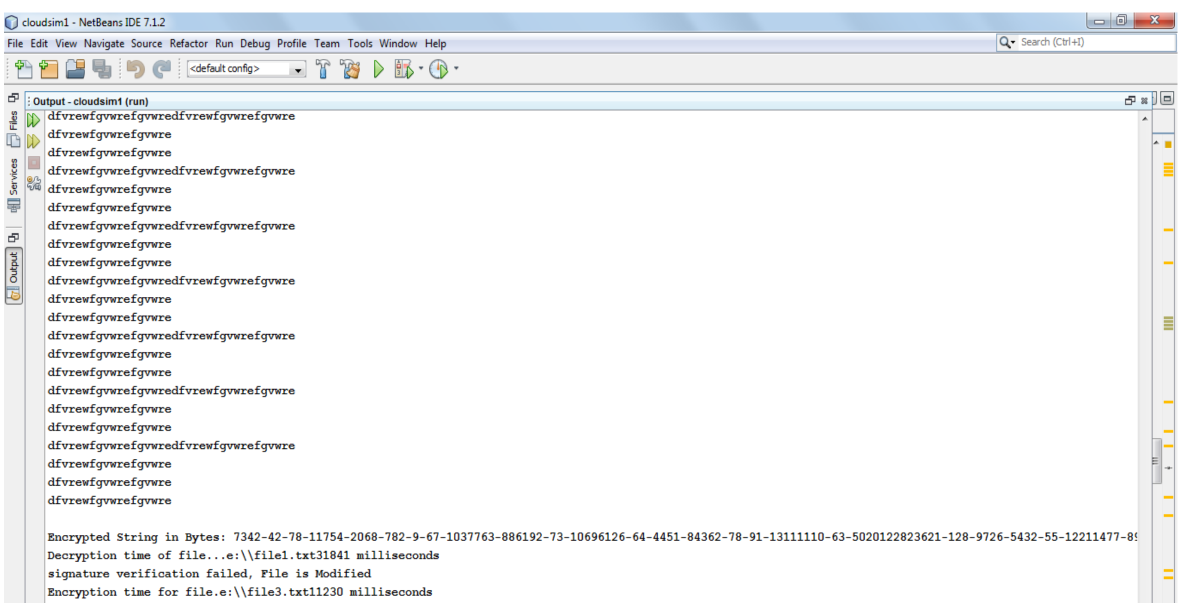

Fig.48 Output when file is modified

The encryption and decryption time in such case is shown in table 9 below

Table 9 Encryption and decryption time when file is modified

FILE SIZE (MB)	ENCRYPTION (ms)	DECRYPTION (ms)
1 (AES)	1037	31841
2.90 (Bcrypt)	698	17868
3.98 (RSA)	11230	-
4.65 (Bcrypt)	695	17819

CHAPTER- 5
CONCLUSION

Cloud computing is an emerging technology that is considered as a boon for the IT sector. It follows mulitenancy where several users can share the resources at a given time over the internet. Along with the benefits of cloud computing ,comes its drawbacks which cause hinderance in full acceptance of this technology.

Security is major issue in cloud computing.The data of consumers is not safe as data of different users is collected and stored at Cloud servers. To maintain the integrity of data a framework is proposed in which data is stored to different cloud based on the type of information. Validation is done on only the sensitive information.Since the file is not modified the data is decrypted and hence the integrity of file is maintained.Performance analysis of the framework is done by calculating the encryption time of different files and increasing the number of requests.It is noted that the time taken for encryption of files is less if the proposed scheme is used, compared to the time taken for encryption of file if old scheme is used where at a time only single level of security is provided to files irrespective of the data stored in them.

In future, validation will be performed on AES and Bcrypt algorithm as well.

REFERENCES

1) Frederic Magoules, Jie Pan, and Fei Teng, "*Cloud computing Data-Intensive computing and Scheduling*," CRC Press, pp. 231, 2012.
2) Santosh Kumar, R.H. Goudar, "*Cloud Computing – Research Issues, Challenges, Architecture, Platforms and Applications: A Survey*," International Journal of Future Computer and Communication, Vol .1, No. 4, December 2012.
3) Qi Zhang, Lu Cheng, Raouf Boutaba, "*Cloud computing: state-of-the-art and research challenges*," Springer 2010.
4) Heena I. Syed, Naghma A. Baig, "Survey on Cloud Computing," International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 4, April 2013
5) Fei Hu, Meikang Qiu, Jiayin Li, Travis Grant, Draw Tylor, Seth McCaleb, Lee Butler and Richard Hamner, "*A Review on Cloud Computing: Design Challenges in Architecture and Security*," Journal of Computing and Information Technology vol. 19, 2011, pp. 25–55.
6) Jeffrey Naruchitparames, Mehmet Hadi Giines, "*Enhancing Data Privacy and Integrity in the Cloud*," IEEE 2011.
7) Dawei Sun, Guiran Chang, Lina Sun and Xingwei Wang, "*Surveying and Analyzing Security, Privacy and Trust Issues in Cloud Computing Environments*," Published by Elsevier Ltd, 2011.
8) Dimitrios Zissis and Dimitrios Lekkas, "*Addressing cloud computing security issues*," Future Generation Computer Systems, vol. 28, no. 3, 2012, pp. 583–592.
9) Zhifeng Xiao and Yang Xiao, "*Security and Privacy in Cloud Computing*," IEEE Communications Surveys & Tutorials, vol. 15, no. 2, 2013, pp. 843-859.
10) Huaglory Tianfield, "*Security Issues in Cloud Computing*," IEEE International Conference on Systems, Man, and Cybernetics (SMC'12), Seoul, Korea, pp. 1082-1089, 2012.
11) Farhad Soleimanian Gharehchopogh, and Sajjad Hashemi, "*Security Challenges in Cloud Computing with More Emphasis on Trust and Privacy*," International Journal of Scientific & Technology Research, vol. 1, no. 6, July 2012, pp. 49-54.

12) Saranya Eswaran and Dr.Sunitha Abburu, "*Identifying Data Integrity in the Cloud Storage*," IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.

13) Unnati S. Shah, Mitali N. Sonar, and Heta K. Desai, "*A Concise Study on Issues Related To Security, Privacy and Trust in Cloud Services*," 2nd International Conference on Mobility for Life: Technology, Telecommunication and Problem Based Learning, 2013, pp.1-6.

14) S. Subashini, and V. Kavitha, "*A survey on security issues in service delivery models of cloud computing*," Journal of Network and Computer Applications, vol. 34, no. 1, pp. 1-11, 2011.

15) Farhan Bashir Shaikh and Sajjad Haider, "*Security Threats in Cloud Computing*," 6th International Conference on Internet Technology and Secured Transactions, 11-14 December 2011, Abu Dhabi, United Arab Emirates.

16) Tharam Dillon, Chen Wu and Elizabeth Chang, "*Cloud Computing: Issues and Challenges*," 2010 24th IEEE International Conference on Advanced Information Networking and Applications.

17) Fei Hu, Meikang Qiu, Jiayin Li, Travis Grant, Draw Tylor,Seth McCaleb, Lee Butler and Richard Hamner, "*A Review on Cloud Computing:Design Challenges in Architecture and Security*," Journal of Computing and Information Technology - CIT 19, 2011, 1, 25–55 doi:10.2498/cit.1001864.

18) Deyan Chen and Hong Zhao, "*Data Security and Privacy Protection Issues in Cloud Computing*," 2012 International Conference on Computer Science and Electronics Engineering.

19) Bhadauria, Rohit, and Sugata Sanyal, "Survey on Security Issues in Cloud Computing and Associated Mitigation Techniques." International Journal of computer applications 47 (2012).

20) Kevin Hamlen, Murat Kantarcioglu, Latifur Khan, Bhavani Thuraisingham, "*Security Issues for Cloud Computing*," International Journal of Information Security and Privacy, 4(2), 39-51, April-June 2010 39.

21) Priya Metri and Geeta Sarote, "*Privacy Issues and Challenges in Cloud computing*," INTERNATIONAL JOURNAL OF ADVANCED ENGINEERING SCIENCES AND TECHNOLOGIES Vol No. 5, Issue No. 1, 001 – 006.

22) Juels, Ari, and Burton S. Kaliski Jr, "*PORs: Proofs of retrievability for large files*," In Proceedings of the 14th ACM conference on Computer and communications security, pp. 584-597. ACM, 2007.

23) Giuseppe Ateniese,Randal Burns, Reza Curtmola,Joseph Herring, Lea Kissner , Zachary Peterson, Dawn Song, "*Provable Data Possession at Untrusted Stores*," Proceedings of the 14th ACM Conference on Computer and Communications Security(CCS 2007).

24) Poonam Rana, P.K. Gupta, Rajesh Siddavatam, "*Combined and Improved Framework of Infrastructure as a Service (IAAS) and Platform as a Service (PAAS) in Cloud Computing,"* Springer, 2013.

25) Prof: Asha Mathew, "*Security and Privacy Issues of Cloud Computing; Solutions and Secure Framework*," ZENITH International Journal of Multidisciplinary Research Vol.2 Issue 4, April 2012, ISSN 2231 5780.

26) Abhishek Mohta, Ravi Kant Sahu, Lalit Kumar Awasthi, "*Robust Data Security for Cloud while using Third Party Auditor*," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 2, February 2012

27) Pradeep Bhosale Priyanka Deshmukh Girish Dimbar Ashwini Deshpande, "*A Review Paper on Enhancing Data Security in Cloud Computing Using 3D Framework & Digital Signature with Encryption*," International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 8, October – 2012.

28) Rajkumar Chalse, Ashwin Selokara and Arun Katara, "*A New Technique of Data Integrity for Analysis of the Cloud Computing Security*," 2013 5th International Conference on Computational Intelligence and Communication Networks.

29) Sravan Kumar, R., and Ashutosh Saxena, "*Data integrity proofs in cloud storage*," In Communication Systems and Networks (COMSNETS), 2011 Third International Conference on, pp. 1-4. IEEE, 2011.

30) Khaba M.V and M.Santhanalakshmi, "*Remote Data Integrity Checking in Cloud Computing*," International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 1 Issue: 6 553 – 557.

31) Regil V Raju,M.Vasanth, Udaykumar P, " *DATA INTEGRITY USING ENCRYPTION IN CLOUD COMPUTING*," Journal of Global Research in Computer Science, Volume 4, No. 5, May 2013.

32) Dalia Attas and Omar Batrafi, "*Efficient integrity checking technique for securing client data in cloud computing,*" International Journal of Electrical & Computer Sciences IJECS-IJENS Vol: 11 No: 05, 2011.

33) Pardeep Kumar, Vivek Kumar Sehgal , Durg Singh Chauhan, P. K. Gupta and Manoj Diwakar, "*Effective Ways of Secure, Private and Trusted Cloud Computing,*" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 2, May 2011.

34) Amir Mohamed Talib Rodziah Atan, Rusli Abdullah & Masrah Azrifah Azmi Murad, "*CloudZone: Towards an Integrity Layer of Cloud Data Storage Based on Multi Agent System Architecture*," 2011 IEEE Conference on Open Systems (ICOS2011), September 25 - 28, 2011.

35) V.Nirmala, R.K.Sivanandhan, Dr. R.Shanmuga lakshmi, "*Data Confidentiality and Integrity Verification using User Authenticator scheme in cloud*," Proceedings of 2013 International Conference on Green High Performance Computing March, 2013.

36) Anandita Singh Thakur and P. K. Gupta, "*Framework to Improve Data Integrity in Multi Cloud Environment*," International Journal of Computer Applications, vol. 87, no.10, February, 2014.

37) S. B. Shivakumar, Ramesh B. E., Kavitha G. M., Mala M. "*Multi Cloud Architecture for Improved User Experience*," International Journal of Inventive Engineering and Sciences (IJIES) Volume-1, Issue-7, June 2013.

38) Sombir Singh, Sunil K Maaka and Dr. Sudesh Kumar, "*A Performance Analysis of DES and RSA Cryptography*," International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 2, Issue 3, May – June 2013

39) Code.google.com/p/cloudsim/

40) Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, C´esar A. F. De Rose and Rajkumar Buyya, "*CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms*," SOFTWARE – PRACTICE AND EXPERIENCE Softw. Pract. Exper. 2011.
41) Cloudbus.org/cloudsim

LIST OF PUBLICATIONS

- Anandita Singh Thakur and P. K. Gupta, "*Framework to Improve Data Integrity in Multi Cloud Environment*," International Journal of Computer Applications, vol. 87, no.10, February, 2014.
- Anandita Singh Thakur and P. K. Gupta, "SECURITY, PRIVACY AND TRUST IN CLOUD COMPUTING: A TECHNICAL REVIEW," 9th National Conference Smarter Approaches in Computing Technologies & Applications (SACTA-2014).