A Study on Cloud Computing Security Challenges

Santosh Bulusu
Kalyan Sudia

School of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona
Sweden
This thesis is submitted to the School of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Software Engineering. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:
Author(s):
Santosh Bulusu 880216-5798
E-mail: santosh.bulusu@me.com

Kalyan Sudia 870418-0556
E-mail: kasu10@student.bth.se

University advisor(s):
Prof. Lars Lundberg
School of Computing

School of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona
Sweden
Internet : www.bth.se/com
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57
Abstract

Context: Scientific computing in the 21st century has evolved from fixed to distributed work environment. The current trend of Cloud Computing (CC) allows accessing business applications from anywhere just by connecting to the Internet. Evidence shows that, switching to CC organizations’ annual expenditure and maintenance are being reduced to a greater extent. However, there are several challenges that come along with various benefits of CC. Among these include security aspects.

Objectives: This thesis aims to identify security challenges for adapting cloud computing and their solutions from real world for the challenge that do not have any proper mitigation strategies identified through literature review. For this the objective is to identify existing cloud computing security challenges and their solutions. Identify the challenges that have no mitigation strategies and gather solutions/guidelines/practices from practitioners, for a challenge with more references but no mitigation strategies identified (in literature).

Methods: This study presents a literature review and a snowball sampling to identify CC security challenges and their solutions/mitigation strategies. The literature review is based on search in electronic databases and snowball sample is based on the primary studies searched and selected from electronic databases. Using the challenges and their solutions identified form literature review, challenges with no mitigation strategies are identified. From these identified challenges with no mitigation strategies, a challenge with more references is identified. The surveys are employed in the later stages to identify the mitigation strategies for this challenge. Finally the results from the survey are discussed in a narrative fashion.

Results: 43 challenges and 89 solutions are identified from literature review using snowball sampling. In addition to these mitigation strategies few guidelines are also identified. The challenge with more references (i.e., more articles mentioning the challenge) and no mitigation strategies identified is incompatibility. The responses identified for the three insecure areas of incompatibility (i.e., interoperability, migration and IDM integration with CC) in cloud computing security are mostly guidelines/practices opined by experienced practitioners.
Conclusions: This study identifies cloud computing security challenges and their solutions. Where these (challenges and solutions) are common to cloud computing applications and cannot be generalized to either service or deployment models (viz. SaaS, PaaS, IaaS, etc.). The study also identifies that there are methods (guidelines/practices identified from practitioners) to provide secure interoperability, migration and integration of on-premise authentication systems with cloud applications, but these methods are developed by individuals (practitioners/organization) specific to their context. The study also identifies the non-existence of global standards for any of these operations (providing interoperability/migration/IDM integration with cloud).

This identified non-existence of global standards and guidelines could be help academics to know the state of practice and formulate better methods/standards to provide secure interoperability. The identified cloud computing security challenges (43) and solutions (89), can be referred by practitioners to understand which areas of security need to be concentrated while adapting/migrating to a cloud computing environment.

Keywords: Cloud computing, Security, Challenges, Solutions, Incompatibility.
Acknowledgments

At the outset we thank Dr. Lars Lundberg and Mr. Hussain Aziz for their valuable guidance. Mr. Hussain had given full support almost day to day interaction for designing the thesis without which this could not be possible. We would finally want to thank friends and family for extending their support in successful completion of thesis.

“I would first want to thank my supervisor, its a dream come true to work with him. Every mail and every time I talked to him I learned new things. I would thank my friends and family for their support. I thank my mother as she gave me unstinted support all through the process. I also thank all my well wishes who have directly and indirectly helped me to frame this work. Last but not least I would thank the almighty for giving me this wonderful opportunity.”

-Santosh Bulusu

“I would also thank my partner, he has lots of patience and helped me a lot when I was confused. I never hesitated him for asking help. He also put his full effort on thesis and discussed about the problems regarding thesis, what is to be next. Its a great pleasure working with him.”

-Kalyan Sudia
# Contents

Abstract ................................. i

Acknowledgments ........................ iii

List of tables ............................ vi

List of figures ........................... vii

1 Introduction .......................... 3
  1.1 Aims and Objectives .................. 4
  1.1.1 Aim .................................. 4
  1.1.2 Objectives .......................... 4
  1.2 Research questions ................... 5
  1.3 Methods used .......................... 5
  1.4 Structure of report .................... 5

2 Background ........................... 7
  2.1 History of cloud computing .......... 7
  2.2 Defining the term cloud computing ... 7
  2.3 Benefits of cloud computing .......... 8
  2.4 Cloud Computing: Service models .... 8
  2.5 Cloud computing: Deployment models ... 10
  2.6 Importance of security in cloud computing ... 11
  2.7 Related work in cloud computing security ... 12

3 Research Design ....................... 17
  3.1 Aims and objectives of study ......... 17
  3.1.1 Aim .................................. 17
  3.1.2 Objectives .......................... 17
  3.2 Framing research questions .......... 17
  3.2.1 Research questions .................. 18
  3.3 Research methodology ................ 19
  3.3.1 Collecting data from literature ... 19
  3.3.2 Collecting data from real world ... 19
List of Tables

3.1 Research questions and Research methodologies ................................ 21
4.1 Business breach report blog ............................................... 31
4.2 Cloud computing challenges related to data .................................... 35
4.3 Cloud computing challenges related to network .............................. 41
4.4 Virtual machine related challenges in cloud computing .................... 45
4.5 Examples of cloud computing previous failure .............................. 47
4.6 Organizational security concerns that challenge cloud computing ........ 54
4.7 Scalability related concerns which challenge cloud computing ............. 56
4.8 Challenges faced by cloud computing in relation to confidentiality and privacy ................................................................. 59
4.9 Data backup and recovery related concerns cloud computing .......... 61
4.10 Other concerns that cater security of cloud computing ........................ 65
4.11 Data related challenges and solutions ...................................... 81
4.12 Network related challenges and solution .................................... 82
4.13 Virtualization related challenges and solutions ............................ 82
4.14 Organization related challenges and solutions .............................. 83
4.15 Scalability related challenges and solutions ................................ 83
4.16 Confidentiality and privacy related challenges and solutions .............. 84
4.17 Backup related challenges and solutions ................................... 84
4.18 Other concerns related challenges and solutions ........................... 84
5.1 Compatible on-premise authentication systems used with cloud ............. 99
A.1 Initial set of papers .............................................................. 124
C.1 Respondents notation ........................................................... 128
List of Figures

4.1 Process of selecting papers for study .......................... 23
5.1 Results of survey: Service providers used ....................... 89
5.2 Results of survey: Type of service used ............................ 90
5.3 Results of survey: Deployment models used ..................... 90
## Acronyms used in this document

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Acronym expanded</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>ACID properties</td>
<td>Atomicity, Consistency, Isolation and Durability</td>
</tr>
<tr>
<td>AWS</td>
<td>Amazon Web Services</td>
</tr>
<tr>
<td>Amazon EC2</td>
<td>Amazon Elastic Compute Cloud</td>
</tr>
<tr>
<td>Amazon S3</td>
<td>Amazon Simple Storage Service</td>
</tr>
<tr>
<td>ARP</td>
<td>Address Resolution Protocol</td>
</tr>
<tr>
<td>AUP</td>
<td>Acceptable Use Policy</td>
</tr>
<tr>
<td>AICPA</td>
<td>American Institute of Certified Public Accountants</td>
</tr>
<tr>
<td>BC</td>
<td>Business Continuity</td>
</tr>
<tr>
<td>CC</td>
<td>Cloud Computing</td>
</tr>
<tr>
<td>CSP</td>
<td>Cloud Service Provider</td>
</tr>
<tr>
<td>CSRF</td>
<td>Cross-site request forgery</td>
</tr>
<tr>
<td>CISO</td>
<td>Chief Information Security Officer</td>
</tr>
<tr>
<td>CICA</td>
<td>Canadian Institute Of Chartered Accountants</td>
</tr>
<tr>
<td>CV</td>
<td>Cumulative Voting</td>
</tr>
<tr>
<td>DR</td>
<td>Disaster Recovery</td>
</tr>
<tr>
<td>DoS</td>
<td>Denial of Service</td>
</tr>
<tr>
<td>DDoS</td>
<td>Distributed Denial of Service</td>
</tr>
<tr>
<td>DNS servers</td>
<td>Domain Name System</td>
</tr>
<tr>
<td>DHCP</td>
<td>Dynamic Host Configuration Protocol</td>
</tr>
<tr>
<td>DBMS</td>
<td>Database management system</td>
</tr>
<tr>
<td>ECPA</td>
<td>Electronic Communication privacy Act</td>
</tr>
<tr>
<td>ECV</td>
<td>Equality of Cumulative Votes</td>
</tr>
<tr>
<td>FTP</td>
<td>File Transfer Protocol</td>
</tr>
<tr>
<td>FISMA</td>
<td>Federal Information Security Management Act</td>
</tr>
<tr>
<td>FIPS</td>
<td>Federal Information Processing Standard</td>
</tr>
<tr>
<td>GAPP</td>
<td>Generally Accepted Privacy Principles</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hypertext Transfer Protocol</td>
</tr>
<tr>
<td>HMAC</td>
<td>Hash-based Message Authentication Code</td>
</tr>
<tr>
<td>IDM</td>
<td>Internet Download Manager</td>
</tr>
<tr>
<td>IP spoofing</td>
<td>Internet Protocol spoofing</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>LDAP</td>
<td>Light weight Access Protocol</td>
</tr>
<tr>
<td>MITM</td>
<td>Man-In-The-Middle</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>OAUTH</td>
<td>Open Authorization</td>
</tr>
<tr>
<td>Acronym</td>
<td>Acronym expanded</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------</td>
</tr>
<tr>
<td>PII</td>
<td>Personal Identifiable Information</td>
</tr>
<tr>
<td>PKI</td>
<td>Public Key Infrastructure</td>
</tr>
<tr>
<td>PCI DSS</td>
<td>Payment Card Industry Data Security Standard</td>
</tr>
<tr>
<td>PCI QA</td>
<td>Payment Card Industry Quality Assessors</td>
</tr>
<tr>
<td>REST</td>
<td>Representational state transfer</td>
</tr>
<tr>
<td>SP</td>
<td>Service Provider</td>
</tr>
<tr>
<td>SSO</td>
<td>Single Sign On</td>
</tr>
<tr>
<td>SLA</td>
<td>Service Level Agreement</td>
</tr>
<tr>
<td>SSL</td>
<td>Secure Socket Layer</td>
</tr>
<tr>
<td>SSH</td>
<td>Secure Shell</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>SMB</td>
<td>Small and Medium Business</td>
</tr>
<tr>
<td>SAML</td>
<td>Security Assessment Mark-up Language</td>
</tr>
<tr>
<td>SSO</td>
<td>Single sign-on</td>
</tr>
<tr>
<td>SPML</td>
<td>Service Provisioning Markup Language</td>
</tr>
<tr>
<td>SOAP</td>
<td>Simple Object Access Protocol</td>
</tr>
<tr>
<td>SAS 70</td>
<td>Statement on Auditing Standards</td>
</tr>
<tr>
<td>SOA</td>
<td>Service-Oriented Architecture</td>
</tr>
<tr>
<td>TTP</td>
<td>Trusted Third Party</td>
</tr>
<tr>
<td>TDB</td>
<td>Technology Development Board</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TPM</td>
<td>Trusted Platform Module</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>VMs</td>
<td>Virtual Machines</td>
</tr>
<tr>
<td>VMM’s</td>
<td>Virtual Machine Monitors</td>
</tr>
<tr>
<td>VPN</td>
<td>Virtual private networks</td>
</tr>
<tr>
<td>VSG</td>
<td>Virtual Security Gateway</td>
</tr>
<tr>
<td>VLAN’s</td>
<td>Virtual Local Area Network</td>
</tr>
<tr>
<td>XSS</td>
<td>Cross-site scripting</td>
</tr>
<tr>
<td>XACML</td>
<td>Extensible Access Control Markup Language</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible Markup Language</td>
</tr>
</tbody>
</table>
Cloud Computing (CC) is an emerging technology that has abstruse connection to Grid Computing (GC) paradigm and other relevant technologies such as utility computing, distributed computing and cluster computing [55]. The aim of both GC and CC is to achieve resource virtualization. In spite of the aim being similar, GC and CC have significant differences. The main emphasis of GC is to achieve maximum computing, while that of CC is to optimize the overall computing capacity [39]. CC also provides a way to handle wide range of organizational needs by providing dynamically scalable servers and application to work with [163]. Leading CC service providers such as Amazon, IBM, ‘Dropbox’, Apple’s ‘iCloud’, Google’s applications, Microsoft’s ‘Azure’, etc., are able to attract normal users through out the world. CC have introduced a new paradigm, which helps its users to store or develop applications dynamically and access them from anywhere and anytime just by connecting to an application using Internet [150]. Depending on customer’s requirement CC provides easy and customizable services to access or work with cloud applications. Based on the user requirement CC can be used to provide platform for designing applications, infrastructure to store and work on company’s data and also provide applications to do user’s routine tasks.

When a customer chooses to use cloud services, data stored in the local repositories will be sent to a remote data center [109]. This data in remote locations can be accessed or managed with the help of services provided by cloud service providers. This makes clear that for a user to store or process a piece of data in cloud, he/she needs to transmit the data to a remote server over a channel (internet) [75]. This data processing and storage needs to be done with utmost care to avoid data breaches.

If proper security measures are not implemented to the data transmitted and operated on cloud, the data is at higher risk than when stored or operated in local repositories [170]. Malicious user who wants to gain access to transmitted data in cloud, can do that by taping into the connection between user and remote location. He can also hack into users account and get access to sensitive information by creating another account (by using virtualized infrastructure that CC provides) in the same service provider with malicious intent [109]. Since cloud computing provides different services to chose from for diversified group of users
(nave, expert, malicious etc.), possibility of having data at risk when working in cloud computing systems is a huge.

From the above discussion, it is summarized that security challenges with cloud adoption and cloud interoperability need to be addressed first, before implementing CC in organizations [39]. A non-exhaustive search on CC challenges also reveal that most of the organizations consider security as an important challenge that needs to be addressed [39] [119] [90]. Even with the strongest security measures implemented there are always some other weakness found and exploited. Henceforth, identifying security challenges, improvising and updating solutions for handling these challenges is essential in implementing CC.

This study presents a combined effort of literature review, snowball sampling and a survey to identify list of solutions/guidelines/practices to handle a CC security challenge that have no proper mitigation strategies defined. It will also provide detailed challenge description and also the challenging situation faced by practitioners (identified in literature). Models, architectures, practices and solutions, which help to mitigate the challenges are also listed. This report has attempted to collect, gather and present a detailed study on cloud computing security challenges and solutions from literature. Then provide solution to the challenge which has no identified mitigation strategies.

1.1 Aims and Objectives

1.1.1 Aim

This thesis aims to identify security challenges for adopting cloud computing and solutions from real world for the challenge that do not have proper mitigation strategies identified through literature review.

1.1.2 Objectives

- Identify existing cloud computing security challenges and their solutions from literature.
- Identify the challenges that have no mitigation strategies defined.
- Collect solutions/guidelines/practices from organizations, for a challenge with more references but no mitigation strategies proposed (identified in literature).
- List out solutions/practices/guidelines to the cloud computing security challenge that has no mitigation strategies identified.
Chapter 1. Introduction

1.2 Research questions

1. What are the security related challenges in cloud?

2. What solutions are present to handle security related cloud challenges?

3. How does an organization handle the identified security challenge, which has no mitigation strategies?

4. What are the solutions or guidelines to enhance security in CC where no mitigation strategies are identified?

1.3 Methods used

Some solutions (practices, models, architectures etc.) used to handle CC security issues are presented in academia. This paper identifies CC security issues and solutions to handle these issues in software engineering context, with the help of literature review (LR) and snowball sampling. A snowball sampling is done on selected papers in the year 2009-2011 and LR is carried out between 2008-2012. The term cloud computing was popularly being used after 2008 we chose to perform a literature review on resources available after that year (popular CC Service Providers (SP) such as Salesforce.com, Google Apps Engine, Rackspace etc., came into popularity from 2008 ) [65]. Then based on results from literature, a survey is conducted to find out solutions/practices used in organizations to the CC security challenges with no mitigation strategies identified through literature. The idea behind doing so is to identify all the existing solutions and best practices/solutions used in real-time to mitigate CC security challenge and bridge the gap that stops enthusiastic users/organization from using the power of CC.

1.4 Structure of report

Chapter 2 Background, explains a brief review on terms and terminology required to understand the later sections.

Chapter 3 Research Design, explains an outline on methodologies used to infer the expected results.

Chapter 4 Literature Review, explains detailed process followed in selecting, sorting and describing extracted information.

Chapter 5 Survey: Design and results, explains survey design process and results obtained from survey.
Chapter 1. Introduction

Chapter 6 Discussion, explains a discussion on findings in this report (summarizes). There is also an explanation of validity threats in this section.

Chapter 7 Conclusion and Future Work, explains the results of this study and also the future work.
Chapter 2

Background

2.1 History of cloud computing

Cloud Computing (CC) is a new term given to a technological evolution of distributed computing and grid computing. CC has been evolving over a period of time and many companies are finding it interesting to use. Without the development of ARPANET (Advance Research Projects Agency Network) by J.C.R.Licklider in 1960’s and many other researchers who dreamt of improving the interconnection of systems, CC would never have come into existence. The advent of ARPANET, which helped to connect (for sharing, transferring, etc.) a group of computers [86], lead to the invention of Internet (where bridging the gap between systems became easy). This Internet helped to accelerate multifarious activities such as human interaction (social media, instant messaging, etc.), business needs of an organization (online shopping, financial services, etc.). Further advancement in this area of Internet resulted in development of Applications Service Provision (ASP), grid and utility computing and cloud computing [86]. CC introduced a new paradigm which changed the traditional interconnection of systems to a pool of shared resources that can be accessed through internet.

2.2 Defining the term cloud computing

NIST (National Institute of Standards and Technology) defines cloud computing as follows: “Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [106].

This definition clearly states that CC helps in minimizing an organization’s expenditure towards managing resources and also reduces the burden of maintaining software or hardware by its user. When burden of management, maintaining a software/hardware is reduced, the companies expenditure and time spent towards infrastructure management is reduced and time saved can be utilized in doing some creative work. This is a huge advantage for users/organizations, which
Chapter 2. Background

not only saves time but also boosts the performance of company by saving time spent on infrastructure.

2.3 Benefits of cloud computing

Some common benefits of CC are [119][130][3]:

- **Reduced Cost:** Since cloud technology is implemented incrementally (step-by-step), it saves organizations total expenditure.

- **Increased Storage:** When compared to private computer systems, huge amounts of data can be stored than usual.

- **Flexibility:** Compared to traditional computing methods, cloud computing allows an entire organizational segment or portion of it to be outsourced.

- **Greater mobility:** Accessing information, whenever and wherever needed unlike traditional systems (storing data in personal computers and accessing only when near it).

- **Shift of IT focus:** Organizations can focus on innovation (i.e., implementing new products strategies in organization) rather than worrying about maintenance issues such as software updates or computing issues.

These benefits of cloud computing draw lot of attention from Information and Technology Community (ITC). A survey by ITC in the year 2008, 2009 shows that many companies and individuals are noticing that CC is proving to be helpful when compared to traditional computing methods [123].

2.4 Cloud Computing: Service models

CC can be accessed through a set of services models. These services are designed to exhibit certain characteristics and to satisfy the organizational requirements. From this, a best suited service can be selected and customized for an organization’s use. Some of the common distinctions in cloud computing services are Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), Infrastructure-as-a-Service (IaaS), Hardware-as-a-Service (HaaS) and Data storage-as-a-Service (DaaS) [146][39][167][157]. Service model details are as follows:

- **Software as a Service (SaaS):** The service provider in this context provides capability to use one or more applications running on a cloud infrastructure. These applications can be accessed from various thin client interfaces such as web browsers. A user for this service need not maintain,
manage or control the underlying cloud infrastructure (i.e. network, operating systems, storage etc.) [106]. Examples for SaaS cloud’s are Salesforce, NetSuite [77].

- **Platform as a Service (PaaS):** The service provider in this context provides user resources to deploy onto cloud infrastructure, supported applications that are designed or acquired by user. A user using this service has control over deployed applications and application hosting environment, but has no control over infrastructure such as network, storage, servers, operating systems etc [106][130]. Examples for PaaS cloud’s are Google App Engine, Microsoft Azure, Heroku [77].

- **Infrastructure as a Service (IaaS):** The consumer is provided with power to control process, manage storage, network and other fundamental computing resources which are helpful to manage arbitrary software and this can include operating system and applications. By using this kind of service, user has control over operating system, storage, deployed applications and possible limited control over selected networking components [106]. Examples for IaaS cloud’s are Eucalyptus (The Eucalyptus Open-source Cloud-computing System), Amazon EC2, Rackspace, Nimbus [77].

- **Privacy and Anonymization as a Service (PAAS):** This service is proposed as a demonstration model to provide data privacy and protection in a particular organization. It also proposes a work-flow oriented approach to manage data in cloud [73][157].

- **Hardware as a Service (HaaS):** The idea of buying a hardware or an entire datacenter with a pay-as-you-use scheme which can scale up and down as per user requirements can be termed as Hardware as a Service (HaaS) [127]. Examples for HaaS cloud’s are Amazon EC2, IBM’s Blue Cloud Project, Nimbus, Eucalyptus, Enomalism [157].

- **Identity as a Service (IDaaS):** This service is targeted for third party service providers who provide Identity and access control functions (including users life cycle and sign-on process). This can be used in combination with various other services (software, platform or infrastructure services) and also for public and private clouds [135].

- **Data storage as a Service (DaaS):** This service allows user to pay for the amount of data storage he/she is using. With this service there is a separate cloud formed which provides storage as a service [39]. Examples of such kinds of users are Amazon S3, Google Bigtable, Apache Hbase, etc [157].
• **Security as a Service (SaaS):** This service allows users to create their own security policies and risk frameworks. In this kind of service cloud users must identify, assess, measure and prioritize system risks [88].

• **Anything as a Service (XaaS):** This is more general form of representing deployment of a service. These services could be of any type and ‘X’ in XaaS can be substituted by software, hardware, infrastructure, data, business, IT, Security, monitoring, etc. These days new service models are being developed [127][130]. Examples are: IT as a service [96], Cloud as a Service (CaaS) [130], Management as a Service (MaaS) [130], etc., are some other services that are identified in literature.

### 2.5 Cloud computing: Deployment models

Among the service models explained above, SaaS, PaaS and IaaS are popular among providers and users. These services can be deployed on one or more deployment models such as, public cloud, private cloud, community cloud and hybrid cloud to use features of cloud computing [123][9]. Each of these deployment models are explained as follows:

- **Public cloud:** This type of infrastructure is made available to large industrial groups or public. These are maintained and owned by organization selling cloud services [39].

- **Private cloud:** This type of cloud deployment is just kept accessible to the organization that designs it. Private clouds can be managed by third party or the organization itself. In this scenario, cloud servers may or may not exist in the same place where the organization is located [39].

- **Hybrid cloud:** With in this deployment model there can be two or more clouds like private, public or a community. These constituting clouds (combinations of clouds used, such as ‘private and public’, ‘public and community’, etc.) remain different but yet bound together by standardized or preparatory technology that enables application and data portability [39].

- **Community cloud:** This type of cloud infrastructure is shared by several organizations and supports a specific community with shared concerns. This can be managed by an organization or third party and can be deployed off or in the organizational premise [39].

Usage of deployments models and services modeled provided by CC changes how systems are connected and work is done in an organization. It adds up dynamically expandable nature to the applications, platforms, infrastructure or any other resource that is ordered and used in CC [130]. This means, when an individual/user uses more computation resources than normal, more money is spent
Chapter 2. Background

and when less resources are used than normal, individual/user needs to pay less (Pay-per-use policy) [136]. This reduces organizations/individuals expenditure towards managing resources (which includes buying computational resources, installing required software or applications to satisfy the daily computation needs and also maintaining them) [130].

Salesforce.com (an enterprise cloud computing company) in the year 1999 was the first to implement this idea. It provided a web based service model to satisfy enterprise needs. Later in 2002, Amazon web services launched cloud-based services, including storage, computation and even human intelligence through Amazon Mechanical Turk [87]. Now, there are various Service Providers (SP) in the market who offers various applications for cloud computing in various forms. The users are provided with feature rich applications, dynamically scalable storage services, application developing interfaces and many more by just signing into a web browser/ dedicated apps. In addition to this, since CC supports remote access feature and automatic updates (by cloud SP), any application once updated on a site gets updated to all its users [146][174].

2.6 Importance of security in cloud computing

The power, flexibility and ease of use of CC comes with lot of security challenges. Even though CC is a new intuitive way to access applications and make work simple, there are a number of challenges/issues that can effect its adoption. A non-exhaustive search in this field reveals some issues. They are: Service Level Agreements (SLA), what to migrate, security, etc. [39]. CC has a feature of automatic updates, which means a single change by an administrator to an application would reflect on all its users. This advertently also leads to the conclusion that any faults in the software are visible to a large number of users immediately, which is a major risk for any organization with little security. It is also agreed up on by many researchers that security is a huge concern for adoption of cloud computing. A survey by IDC on 263 executives also shows that security is ranked first among challenges in CC [85][7][100].

Even though a company boasts to have top class security and does not update its security policies from time to time, it will be prone to security breaches in near future. In this regard, through this detailed study, we propose to update the readers with different distinctions (types of) in security challenges and their solutions. We also include real-time practices to mitigate challenges, include improved solutions proposed by researchers to show which areas of cloud computing need more attention.
2.7 Related work in cloud computing security

Some methods have been proposed in literature for handling security issues in organizations implementing cloud computing. A brief discussion on these methods is given below:

- Popovi and Hocenski, discussed security issues, requirements and challenges that cloud service providers (CSP) need to address in cloud engineering [119]:
  1. Security issues describe the problems encountered during implementation of cloud computing (CC).
  2. Security standards provide some security templates, which are mandatory for cloud service providers. The Open Visualization Format (OVF) is a standard for creating new business models that help the company to sell a product on premises, on demand, or in a hybrid deployment model.
  3. Security management models are designed based on the security standards and best practices.

- Maggi and Zanero, addressed countermeasures (anti-viruses, intrusion detection systems) developed to mitigate well-known security threats. The focus is mainly on anomaly-based approaches which are mostly suited for modern protection tools and not for intrusion detectors. The pattern-based changes (example: from thin client connected to the main frame or powerful workstations connecting to thin clients) are observed, which cause some simultaneous changes in work environment and new problems to security of CC [101].

- Ertaul et al., mentioned CC’s features like reduced total cost of ownership, scalability and competitive differentiation. They claim CC also minimizes complexity and provides faster and easier acquisition of services to customers. Virtualization is the technique used to deal with quality of service (QOS). Usage of CC is considered to be unsafe in an organization. For dealing with this type of situation, they investigated a few major security issues with CC and also existing countermeasures to those security challenges. Advantages for implementing CC from a different point of view are also discussed. They also stated that some standards are required in CC for security [46].

- Subashini and Kavitha, dealt with the security risks faced in the CC. They provided empirical evidence on security risks and issues encountered during deployment of service delivery models in an organization. The service models are placed in cloud and the empirical validation was made in order
to justify the safety of the environment. Security was the main issue while
there were also complications with data protection and data privacy in a
continuous manner that affected the market [144].

- Md. Tanzim Khorshed et al [90] boast that cloud computing helps reduces
cost of services and improves business outcomes. But to market this and
popularize its use by IT user community, there are many security risks to
be solved. They also mentioned that the cloud services pose an attrac-
tive target to cyber attacks and criminal activities as these services have
information from many organizations and individuals stored in their repos-
itories. The author performs a survey in cloud computing to find out gaps
and security concerns and mentions 5 common types of attacks:

a. Denial of service: In this type of attack the attacker prevents the
   legitimate user from accessing his resources,

b. Malicious insider attacks: This type of attack the attacker is an insider.
   This person can easily gain access to sensitive user information namely:
   passwords, cryptographic keys, etc.

c. Cross virtual machine side channel attacks: Is the type of attack in
   which attacker resides in the same physical hardware as that of the
   target virtual machine and gains access to his sensitive information ,

d. Phishing attacks: In this type of attacks the attacker sends links to
   the target user through email or instant messages. These links look as
   if they were sent by a trusted party but through this links the attacker
   can gain access to user sensitive information,

e. Attacks targeting shared memory: The shared memory between the
   user and the attacker is used to perform unwanted, unauthorized ac-
   tions.

They proposed a method for automatic identification of these attacks, tested
its effectiveness by simulating attacks in a real, actual cloud setup. The
design of this model, they say, is based on machine learning models. Various
models are considered and a support vector machine (SVM) is able to detect
maximum attacks.

In the end, the authors conclude that their results are outcomes of simula-
tion and express that there can be difference in depth, volume and intensity
of attacks in simulation as compared to an actual environment. This leaves
us an impression that the method proposed still needs real world experi-
mentation and hence the attacks mentioned in this paper are still a threat
in CC environment.

- A study by Farhan Bashir Sheikh et al in [137], includes information regard-
ing vulnerable security threats from 11 articles. The authors tabulated their
findings i.e., problem discussed and technique used to solve the problem in their paper. But in the end, they conclude expressing that cloud computing from user perspective is suffering from numerous security threats. This, they say, is the only worth mentioning disadvantage in CC. They also list out the following as key concerns in their point of view:

a. Users authentication: User authentication process must be improvised to ensure that malicious users do not get access to powerful computing systems in CC [1].

b. Leakage of data or Data loss: Data can be at risk if an unauthorized person gains access to shared pool of resources and deletes or modifies data. This risk can increase further if there exists no backup for that data [1].

c. Clients trust: There must be strong authentication practices implemented to ensure that the clients data is being protected from unauthorized access [1].

d. Malicious users handling: Malicious users can be attackers using cloud services with a malicious intent or an insider who has gained the trust of company but works to gain access to sensitive information stored in cloud [1].

e. Hijacking of sessions: These kind of attacks happen when a legitimate user is prone to phishing or insecure application interfaces that can be exploited by attackers. Through this kind of attacks, attackers gain user credentials and hijack legitimate users sessions [1] and

f. Wrong usage of CC and its services: Cloud computing service providers give access to try their cloud services for a limited period of time for free. Some users utilize this trial period to misuse the resources obtained through CC service provider [1].

Also, in their ‘future work’ section, the authors mention that CC is not fully mature and is a treat to both vendors and users. They also expressed concerns of not having any strict security standards for CC. They conclude that even though multiple tools are developed and models proposed, CC is not secure and nothing fruitful is found.

• Iliana Iankoulova et al in [75] have performed a systematic review to identify which security requirements need to be further researched. To find that, the authors used an existing model with 9 sub-factors namely: access control, attack/harm detection, non-repudiation, integrity, security auditing, physical protection, privacy and confidentiality, recovery and prosecution to categorize their finding from 55 papers. From this review they found that nonrepudiation, physical protection, recovery and prosecution are the
least researched in security areas. Integrity, access control and security auditing are the most popular areas. A surprising finding in their review is that privacy and confidentiality had been observed only in 7% publications. In addition to security requirements, solutions to these identified challenges were also mentioned.

A limitation to this study is usage of an elaborated categorization scheme. Due to this, they say, they missed including some papers (and requirements), which is due to selection of databases or availability. They propose that further study in CC is required to understand why some of these requirements are least researched. They also said further study should follow another structure to describe CC security requirements (which might help in identifying requirements missed in their study).

- Eystein Mathisen in [105] discusses some key security issues of cloud computing (policy, software and hardware security) and techniques implemented to reduce the risk. The author expresses that usage of CC will increase in near future and more companies will share their information to cloud servers, which could attract large groups of hackers. He also says that in future there are possibilities for interoperability and data lock-in issues, which can be reduced by using open standards from the time of CC adoption.

The author concluded by saying that security is always addressed late while adopting CC and also mentioned that security standards are still missing for CC. If an organization wishes to shift to CC but is reluctant due to lack of proper measures or standards, it can refer to Open Cloud Manifesto which is the largest initiative surrounding open standards. These standards are restrictive and so most companies do no wish to follow the Open Cloud Manifesto standards.

The proposed methods address security challenges in CC and solutions to overcome these challenges. The following points can be observed from above related work:

1. In the study performed by Ertaul et al., he mentions that CC is considered unsafe to be used by organizations and he also stated CC requires some standards [9][46]. This provides a need for further research to ensure security for all those who are using CC applications [80].

2. Eystein Mathisen concluded in their article that security is always addressed late while adopting CC. He also say that no proper security standards for CC exist [105].

3. Md. Tanzim Khorshed et al and Farhan Bashir Sheikh et al both advocate that security challenges are still a major hindrance for adopting CC [90][137].
4. Md. Tanzim Khorshed et al have identified some threats to CC and proposed a method for automatic detection of network attacks, but it is still not used in real world [90].

5. Iliana Iankoulova et al identified few security areas of CC to be less researched and also suggest to use another way of categorization in further studies [75].

From these studies it can be clearly understood that there are no security standards defined, even after a few researchers trying to formulate them. It can also be understood that even though few organizations and researchers tried to formulate strategies to handle security issues in cloud, there are still many companies that are reluctant to join the group of CC users. Their major concern is still security in cloud computing. This research tries to identify every possible challenge cloud faces and their practice/solution from literature and then pick a challenge that has no proper solutions/practices/models proposed and ask the people working in organizations to suggest a set of guidelines/practice to mitigate/control the challenge. This study will help both organizations and academics to identify the extent of research. It also will help to identify a set of solutions/practices/guidelines to harness the power of CC securely. This study will also include benefits of using a specific technique listed out, which can help organizations to choose a solution that fit their requirements.
Chapter 3

Research Design

3.1 Aims and objectives of study

This thesis aims to find cloud computing security challenges and provide a set of guidelines/solution/best practices to handle CC security challenges based on the aims and objectives of this thesis.

3.1.1 Aim

To identify security challenges for adopting cloud computing and solutions from real world for the challenge that do not have proper mitigation strategies identified through literature review.

3.1.2 Objectives

- Identify existing cloud computing security challenges and their solutions from literature.

- Identify the challenges that have no mitigation strategies defined.

- Collect solutions/guidelines/practices from organizations, for a challenge with more references but no mitigation strategies proposed (identified in literature).

- List out solutions/practices/guidelines to the cloud computing security challenge that has no mitigation strategies identified.

3.2 Framing research questions

From the previous section i.e., related work (discussed in section 2.7) it is not hard to accept the fact that in the area of CC, ‘security’ is still a challenge. As long as this new paradigm does not evolve into a more secure computing platform, which users (organizations/individual) can trust, depend on and utilize in everyday work, harnessing the benefit of CC is not possible. Preliminary studies show
Chapter 3. Research Design

that most of the organizations are concerned about security when considering CC applications [9][46][51]. It is also clear from previous discussion that CC has no specific security standards. If there are no security standards, how can cloud be safe? Supporting this argument, another researcher mentions that CC is not safe to be used in organizations, which means that existing solutions or practices are not sufficient [46]. If the existing methods are not satisfactory, there could be better ways (solutions/guidelines/practices) developed to improve security in CC applications. Before developing better methods, existing challenges and their solutions need to be identified (first and second research question are framed to know challenges and mitigation strategies, specific to CC RQ 1 & RQ 2). From the above discussion its clear that CC does not have specific security standards and security is the major concern for adopting cloud computing [137]. To handle the security challenges in CC, researchers have proposed, designed and experimented different mitigation strategies (some of them discussed as related work in section 2.7). They have also framed a set of guidelines to handle security issues in a better way. Still, there are a few security challenges unattended/insecure, which need improved methods (observed from related work chapter 2, section 2.7). These challenges (challenges which are still insecure/unattended and need improved methods) can be identified from the literature review used to answer ‘RQ 1’ and ‘RQ 2’. As the next step to promote the usability and trust in security of CC, finding out which security measures are being used by CC security professionals in practice (for the unattended/insecure challenges) is required to be known (The third research question is framed to investigate which security practices are used in practice for the challenges which have no solutions RQ 3). The last section is to list out solutions/practices/guidelines for handling CC security issues. (the last research question helps to find out which are popular methods to enhance the security in CC RQ 4).

3.2.1 Research questions

To find answers for issues discussed above, the following research questions are framed:

1. What are the security related challenges in cloud?

2. What solutions are present to handle security related cloud challenges?

3. How does an organization handle the identified security challenge, which has no mitigation strategies?

4. What are the solutions or guidelines to enhance security in CC where no mitigation strategies are identified?
3.3 Research methodology

3.3.1 Collecting data from literature

To identify which areas of cloud computing security need more research, initially CC challenges are found (this is done by searching the literature). Available methods for achieving this are literature review (LR) and systematic literature review (SLR). SLR is used to find all available data relevant to a particular research area [30]. Since the topic of cloud computing is a novel one, literature review and snowball sampling are employed [60].

The article written by the researchers are based on scientific papers, online sources, journals etc. Snowball sampling revisits all these references used in writing an article. This in turn increases the final set of papers used to understand the problem (selected for study). In this context, since this study had threat for finding fewer references, authors choose snowball sampling. Literature review is also used addition to snowball sampling for collecting information to satisfy ‘RQ 1’ and ‘RQ 2’, this method helped to gather references that are missed through snowball sampling.

Based on the information gathered from literature review, an analysis was employed to develop general explanations. This helps researcher to identify the key concepts, terms and also resources used by other researchers. This data is used to develop alternative designs or find out need for further research. Literature review using online databases involves a series of steps [52]:

1. Identifying the keywords for the topic.
2. Creating a list of possible search terms.
3. Using search engines, electronic databases to find information.
4. Modify the list of terms and repeat step 3.

Based on similarity in the findings (inferred from respective articles), identified challenges are grouped (explained in detain under data analysis section 4.2). Literature review in this context gathers required information for answering ‘RQ 3’.

3.3.2 Collecting data from real world

Based on results of literature review and empirical study, mechanisms that are actually used in organizations to handle the security challenges identified from literature are needed to be found. Other methods available under empirical studies are interviews, case studies, experiment and post-mortem. The interview is a process of gathering in-depth information to the interviewer’s question from responder. Here the interviewer asks a question and expects an answer to that
from the responder. The interviewer can also change the questions based on responses to get required information. On the other hand surveys are referred to research-in-the-large, their execution or measurement cannot be controlled. The surveys generally start with a set of defined questions and target a group of people. Case study on the other hand is an observational method where researchers understand the ongoing project or activity. Experiment is a formal rigorous and controlled investigation where normally the objective is to distinguish between two situations for example a situation under investigation and a control situation. Post-mortem is a mix of both case study and survey. This post-mortem contains survey features, but concentrates on a specific case [161]. The objective here is to identify solutions/practices used in organizations. Surveys are chosen over other methods of empirical studies, as these can be simple and also easy to answer. The surveys are used to provide a quantitative or numeric description of trends opinions or attitudes [36].

The surveys are used to gather the solutions/practices/guidelines to the challenges which have no mitigation strategies proposed (and have been repeatedly cited by many papers as a challenge). The process of survey is chosen and used to answer RQ 3, as the objective is to gather solutions/practices to a security challenge identified through literature review. For doing this we perform a survey on a selected group of people (who are essentially professionals working in cloud computing and are in various designations, experience levels and companies) by providing each of them with an online survey link which has same set of questions. Since most respondents wish to complete answering questionnaire quickly, the process of survey will suite best in saving time and effort.

### 3.3.3 Presenting results

The results of literature review and surveys are going to be guidelines/practices/solutions. These are going to be responses from diversified group of users and with different levels of experience. Since the observations from the literature review are not helpful to handle the challenge identified, solutions/practices/guidelines observed from surveys are going to be new to the reader. Hence a narrative analysis is employed to explain the various responses observed from the surveys. This narrative analysis will help users (individual/organizations) to consider the possible ways to mitigate the challenge and for the researchers giving an idea as on what are the practitioners expecting as a solution to the challenge (to answer RQ 4).
Table 3.1: Research questions and Research methodologies

<table>
<thead>
<tr>
<th>Sl no</th>
<th>Research question</th>
<th>Methodology employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What are the security related challenges in cloud?</td>
<td>LR, Snowball sampling</td>
</tr>
<tr>
<td>2</td>
<td>What solutions are present to handle security related cloud challenges?</td>
<td>Data analysis of LR and snowball sampling</td>
</tr>
<tr>
<td>3</td>
<td>How does an organization handle the identified security challenge, which has no mitigation strategies?</td>
<td>Surveys</td>
</tr>
<tr>
<td>4</td>
<td>What are the solutions or guidelines to enhance security in CC where no mitigation strategies are identified?</td>
<td>Narrative analysis</td>
</tr>
</tbody>
</table>

Table 3.1 summarizes the discussion of research questions and research methodologies. The research questions of this study and the corresponding research methodologies are laid out in a tabular form. As we can see, LR and Snowball sampling are used to identify the security related challenges in cloud. Data analysis on LR and Snowball sampling helps to answer the question ‘What solutions are present to handle security related cloud challenges?’ Surveys are used to answer ‘How does an organization handle the identified security challenge, which has no mitigation strategies?’ Finally, the survey will be analysis to find the solutions/guidelines to enhance security in CC.
Chapter 4

Literature Review

The study starts by identifying cloud computing security challenges and their mitigation strategies from the literature. To identify cloud computing security challenges and their solutions, grey literature, systematic literature review, snowball sampling etc., could be used, but this report uses snowball sampling and literature review (to answer ‘RQ 1’ and ‘RQ 2’). Literature Review (LR) helps to identify state of art in a study and snowball sampling helps to revisit into references used in the article and find information related to the current study. The rationale for selecting snowball sampling and literature review is as follows:

- The topic of choice (cloud computing) is new and using other techniques might result in a few papers.
- The snowball sampling technique considers most relevant papers as initial set and then traverses through all the references in them.
- The LR helps to identify articles that are relevant to the study but are missed to be identified by snowball sampling.

4.1 Process of selecting papers for study

Snowball sampling starts with a set of 13 initial (will be discussed in Table A.1) and 19 authors suggested papers (total 32 articles as initial set for snowball sampling). A two level snowballing on this initial set of scientific papers (32 articles), resulted in a total of 88 articles (32 initial + 56 articles resulting from snowball sampling). The 13 initial papers were selected after an exhaustive search on different databases (i.e., ACM, IEEE, Inspec, Springerlink, Scopus, ISI) using a common search string (cloud computing security). Papers with more number of references, more challenges, more citation and papers published in the period 2009-2011 were selected (as initial articles) for performing the snowball sampling process. The articles with irrelevant abstract (not related to cloud computing security), title, unavailable in full text, not related to software engineering and published in other languages are excluded. The process is diagrammatically
In addition to the 88 papers selected through snowball sampling, 81 additional papers were selected through a literature review (from ACM, IEEE, Scopus, Evillage and ISI). The search string used was ‘(cloud AND computing) AND (security AND challenges)’. The exclusion criteria used were same as that of the snowball sampling (excluded papers related to electronic commerce, health care, telecommunication network, mobile computing, industrial etc.) and period of search was from 2008-2012. The idea behind exclusion criteria was to focus on software engineering.

The objective of performing literature review is to identify challenges and solutions proposed till date (answer for ‘RQ 1’ and ‘RQ 2’), data analysis is applied to the final set of papers obtained through literature review. Data analysis is “The process of systematically searching and arranging the interview transcripts, field-notes, and other materials that you accumulate to increase your understanding of them and to enable you to present what you have discovered to others” and it helps in [115]:

a. Reducing the data by filtering, selecting and simplifying to transform the raw data from different sources.
b. Reorganizing data in such a way that conclusions or a logical evidence can be built on these results.

Data analysis can be done in various methods (qualitative, quantitative, or a combination of qualitative and quantitative). A qualitative analysis technique ‘coding’ was employed for data analysis. This coding was done based on Perspectives held by subjects, because we have coded similar information, which have shared rules and norms as well as general points of views. Other possible qualitative data analysis methods are narrative summary, thematic analysis, grounded theory, meta-ethnography, meta-study, realistic synthesis, Miles and Huberman’s data analysis techniques, content analysis, case survey, qualitative comparative analysis and Bayesian meta-analysis [40]. Meta-analysis is a method for summarizing empirical studies that has been advocated in many areas of behavioral research. Narrative summary is a traditional method of summarizing data from many areas of research [37]. Grounded theory is an inductive, theory discovery methodology that allows the researcher to develop a theory based on general features of the topic [103]. Case survey bridges the gap between surveys and case-studies to combine their respective benefits of generalizable procedures [93]. Meta-ethnography refers to the synthesis of interpretive research [114]. Qualitative comparative analysis is a combination of both qualitative and quantitative data [40]. Since the objective here is to identify CC security challenges and solutions coding technique is employed. This technique refers to a collection of data and categorizing it based on codes designed by the author. First, regularities and patterns from articles are searched, selected and grouped in sections, this helps to easily analyze complex data [115]. Each of these grouped challenges are explained in section 4.2. For example, in table 4.2 that describes data related challenges describes the challenges identified and grouped in ‘data related’ challenges. From this table it can be understood that ‘data security’ a challenge mentioned in data related challenges is given code ‘D1’. Similarly a challenge ‘data locality’ mentioned in ‘data related’ challenges is given code ‘D2’.


4.2 Categories of Security Challenges in Cloud Computing

Cloud computing security is the major concern and has various challenges that need attention [165][17]. From the recent surveys on IT executives and CIO’s conducted by IDC, it was clear that security was the highly cited (74%) challenge
in the cloud computing field [3][85]. A comparison with grid computing systems also proves that for cloud computing security the measures are simpler and less secure [55]. Security in cloud computing is totally based on the cloud service provider, who is responsible for storing data and providing security [63].

Based on the data analysis process explained in the previous section, terms in literature with similar understanding (terms such as data security, data locality etc., are categorized in data related challenges) are grouped under 8 sections and each section is explained in detail below. These sections are grouped based on how they are explained and based on categorizations proposed by some authors in their discussions.

4.2.1 Data

Data security (D1)

Information from articles that discuss about data security and data protection are considered.

Security provided by cloud SP’s might not be highly cost effective when implemented in small companies. But when two or more organizations share a common resource there is a risk of data misuse. In such situation it is required to secure data repositories [105]. Not only the data repositories but also data should be secured in any stage such as storage, transit or process [92]. Since this kind of sharing resources is prevalent in the CC scenario, protection of data is important and is the most important challenge among other CC challenges [170][134][109]. In shared areas to keep data secure is challenging than protecting in a personal computer [160][57]. This problem has begun due to the introduction of new paradigm CC [107]. The author of article [164] explains how data security effects in various service models namely SaaS, PaaS and IaaS and in the article [70] author advocates that data security is the primary challenge for cloud acceptance and author for [107] expresses that cloud data security is an issue to be taken care of. For enhanced security on data repositories it is important to provide better authentication, authorization and access control for data stored on CC in addition to on-demand computing capability [63][28][133].

Given below are three key areas in Data security that CC refers to [169];

1. Confidentiality: When enterprise data is stored outside organizational boundaries it needs to be protected from vulnerabilities. To protect data from vulnerabilities, employees must adopt security checks to ensure that their data stays protected from malicious attacks [10][2][58]. Few test are used to help organizations to assess and validate, to which extent data is protected from malicious user and they are as follows [144][26][178][154]:

   (a) Cross-site scripting [XSS]

   (b) Access control weaknesses
Chapter 4. Literature Review

(c) OS and SQL injection flaws
(d) Cross-site request forgery [CSRF]
(e) Cookie manipulation
(f) Hidden field manipulation
(g) Insecure storage
(h) Insecure configuration

Example: With the help of Payment Card Industry Data Security Standard (PCI DSS) the data is not allowed to go outside the European Union [144][26][154]. This can also enforce encryption on certain areas of data and by encrypting data in this way permission is given only to specific users to access specific areas of data [119].

2. Integrity: There is no common policy that exists for data exchange. To maintain security on client data, thin clients are used where only few resources are possible. Since only few resources are given access user are not suggested to store any personal data such as passwords. Since passwords are not stored on desktops, passwords cannot be stolen by anyone. Integrity of data can be further assured by [10]:

- Using some extra features which are like unpublished API’s for securing a particular section of data.
- Using DHCP and FTP for long time has been rendered as insecure.

3. Availability: Availability is the most problematic issue, where several companies face downtime (i.e., denial of service attack) as a major issue. The availability of a service generally depends on contract signed between client and vendor.

Some other points that need to be highlight when it comes to data security [28]:

- Who has rights over data (i.e., does data still belong to company?)
- If there is any other company or organization being involved (i.e., is there involvement of any third party organization) [10][78].
- Customers using CC applications need to check, if the data provided by cloud service providers is carried out in a lawful way or not.
- If data protection fails while data is being processed, it could result in administrative, criminal sanctions or civil type of issues (which depends on country controlling data). These issues may occur due to multi transfers of data log between federated cloud providers.
Chapter 4. Literature Review

- Cryptographic algorithm should be maintained well and updated regularly, failing to do so could lead in disclosing personal data [169].

- Data is not completely protected when it is encrypted and stored. When searching for a piece of information again in CC servers care should be taken to retrieve information in a secured process. Traditional searches can disclose data to other companies/individuals [156]. Not only this but also using complex ways to encrypt can also raise issues while retrieving data from storage [126].

Data locality (D2)

Information from articles that discuss about data locality, jurisdictional issues, risk of seizure and loss of governance are considered.

Using CC applications or storages services questions such as “does CSP allow to control the data location?” arise and reason for asking this question is explained in this section [123]. We know that in CC the data can be hosted anywhere and in most cases the customer does not know the location of his data i.e., the data is generally distributed over number of regions [110][85][100]. It is also known that when the geographical location of data changes the laws governing on that data also change. This clarifies that the user’s data (information, applications, etc..) that is stored in cloud computing (distributed over number of regions) is affected by the compliance and data privacy laws of that country (which ever country user’s data is located). So it is necessary that the customer should be informed about the location his data stored in cloud [144]. SP can provide the location of data whenever there is a change or if the SP provide a mechanism to track the location of data it can be very helpful for customer [110]. If the customer shows any concerns towards the location of data they should be dealt immediately [142]. This is because if the customer is found violating laws of certain territory his/her data can be seized by the government. Since all the information stored in cloud computing architecture is in a same data repository, there is always a chance for government seizing or compromising data of another company [46]. Hence, before storing data on the cloud, users must ensure providers whether data are stored keeping jurisdiction constraints into consideration or not. They must also verify existing contractual commitment which can shows agreement to local privacy requirements [86][22][119][72]. For example: Some countries like Europe, South Africa do not let their data leave country borders, as information is potential and sensitive in nature [144] [109].

Because of all these problems some customers also concerned that their data should stay in the same geographical locations in which they are [7]. In [55], the author mentions that clouds will face significant challenges to handle Cloud applications while managing the data locality of applications and their access patterns. In addition to this, author in [105] also concerned regarding the physical
Data integrity (D3)

If a system maintains integrity, its assets can be only be modified by authorized parties or in authorized ways. This modification could be on software or hardware entities of system [179]. Data integrity in any isolated system (with a single database) can be maintained via database constraints and transaction. But in a distributed environment, where databases are spread out in multiple locations data integrity must be maintained correctly, to avoid loss of data [99][100]. For example when the premises application is trying to access or change data on a cloud the transaction should be complete and data integrity should be maintained and failing to do so can cause data loss [75]. In general every transaction has to follow ACID properties (Atomicity, Consistency, Isolation and Durability) to preserve data integrity [179][46][130][19]. This data integrity verification is one of the key issues in cloud data storage especially in case of an untrusted server [120][81].

Web services face problems with transaction management frequently as it still uses HTTP services. This HTTP service, does not support transaction or guarantee delivery. The only way to handle this issue is by implementing transaction management at Application Programing Interface (API) level. There are some standards (such as WS-Transaction and WS-Reliability) to manage data integrity with web services exists. But since these standards are not matured they are not implemented. The majority of vendors who deal with SaaS, expose their web service API’s without any support for transactions. Additionally, each SaaS application may have multiple levels of availability and SLA’s (Service Level Agreement), complicating it further with management of transactions and data integrity across multiple SaaS applications.

Lack of data level integrity controls could result in some profound problems. Architects and developers need to handle this carefully, to make sure that database’s integrity is not compromised when shifting to cloud computing applications [144]. Failing to check data integrity may lead to data fabrication or in some cases even if data is removed by CSP as its rarely accessed, the user won’t be known until he attempts to access [153]. In article [98] the author compares protocols that are used for remote data integrity and expresses that the protocols in comparison were either focused on single server scenario or multi-server but not in dynamic situation such as CC. In the article [126], author expresses concern regarding maintaining a local copy to check the data integrity by each user. CSP’s are supposed to have some user’s metadata to grant access or identify a user. To manage the integrity of data in cloud storage metadata should be managed correctly [31].

Author in [120] has performed a comparison of different protocols (entire data dependent tag, data block dependent tag, data independent tag based, data
replication-based protocol) to check data integrity that are discussed and compared. The comparison shows most proposed methods are having data integrity as a primary objective and also support dynamic operations in cloud storage. But there is room for improving these methods.

**Data segregation (D4)**

Another issue in cloud computing is multi-tenancy. Since multi-tenancy allows multiple users to store data on cloud servers using different built-in applications at a time, various user’s data resides in a common place. This kind of storage shows a possibility for data intrusion. Data can be intruded (malicious user retrieving or hacking into others data) by using some application or injecting a client code [144][72]. The user should ensure that data stored in the cloud should be separated from other customer’s data [55][85][123]. Article [142] suggests that an encryption scheme used should be assessed and certified that they are safe and cloud provider should use only standardized encryption algorithms and protocols.

Vulnerabilities with data segregation can be detected or found out using the following test [144]:

1. SQL injection flaws
2. Data validation
3. Insecure storage

**Data access (D5)**

Information from articles that discuss about data access, access rights, privileged user access, access control, administrative access are considered. This issue mainly relates to security policies. Policies are described as “Conditions necessary to obtain trust, and can also prescribe actions and outcomes if certain conditions are met” [121]. Every organization has their own security policies. Based on these policies employee will be given access to a section of data and in some cases employees might not be given a complete access. While giving access it is necessary to know which piece of data is accessed by which user [11][109]. And for this various interfaces or encryption techniques are used and keys are shared with only authorized parties. Wrong management of keys can also cause difficulty in providing security. To prevent wrong management of keys access control list might be used, but with increase in the number of keys, the complexity of managing keys also increases [95]. Even in the case of interfaces used to manage security, if the number of interfaces increase management of access can also become complicated [166].

In any case, to ensure that data stays away from unauthorized users the security policies must be strictly followed [144][72]. The unauthorized access
could be from an insider or by any user trying to access CC [136]. The data access control is a specific issue and various standalone approaches used in access control of data in CC are mentioned in article [45].

Privileged user access: Since access is given through the Internet, giving access to privileged users is an increasing security risk in cloud computing. When sensitive data is transferred through the Internet there is a possibility for an unauthorized user to gain access and control data. To avoid this, the user must use data encryption and additional protection mechanisms like one time password protection or multi-factor authentication, that can be used to provide strong authentication and encrypted protection for all administrative traffic [3][86][22].

Data confidentiality issue (D6)

Cloud computing allows users to store their own information on remote servers, which means content such as user data, videos, tax preparation chart etc., can be stored with a single cloud provider or multiple cloud providers. When users store their data in such servers, data confidentiality is a requirement [170]. Storing of data in remote servers also arises some privacy and confidentiality issues among individual, business, government agency, etc., some of these issues are mentioned below [144][76]:

1. Privacy of personal information and confidentiality of business and government information implications is significant in CC.

2. The terms of service and privacy policy established by cloud provider are the key factors that vary user’s privacy and confidentiality risks significantly.

3. Privacy and confidentiality rights, obligations and status may change when a user discloses information to cloud provider based on information type and category of CC user.

4. The legal status of protections for personal or business information may be greatly affected by disclosure and remote storage.

5. Location of information may have considerable effects on the privacy and confidentiality information protection and also on privacy obligations for those who process or store information.

6. Information in cloud may have multiple legal locations at the same time but with differing legal consequences.

7. A cloud provider can examine user records for criminal activity and other matters according to law.

8. Access to the status of information on cloud along with privacy and confidentiality protections available to users would be difficult due to legal uncertainties.
Chapter 4. Literature Review

In addition to these to maintain confidentiality understanding data and its classification, users being aware of which data is stored in the cloud and what levels of accessibility govern that piece of data should also be known [160].

Data Breaches (D7)

Since data from various users and organizations is stored in a cloud environment, if user with malicious intent enter the cloud environment, the entire cloud environment is prone to a high value target [144][166]. A breach can occur due to accidental transmission issues (such breaches did happen in Amazon, Google CC’s) or due to an insider attack [100][143]. In any case of breach data is compromised and is always a security risk which is also a top threat mentioned by CSA [143]. There is a high requirement for breach notification process available in the cloud. It is because if breaches are not notified the cloud might not be able to notify serious attacks [90].

Table 4.1: Business breach report blog

<table>
<thead>
<tr>
<th>Threat</th>
<th>Impact</th>
<th>Resulting in Pseudo Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Criminals Pose</td>
<td>Greatest (73%)</td>
<td>Least (30,000 compromised records)</td>
</tr>
<tr>
<td>Insiders Pose</td>
<td>Least (18%)</td>
<td>Greatest (375,000 compromised records)</td>
</tr>
<tr>
<td>Partners are middle</td>
<td>73.39%</td>
<td>73.39%</td>
</tr>
</tbody>
</table>

The business breach report blog gives information on the impact of breaches [144], as shown in table 4.1. We can see that the threat of external criminals is greatest by 73% but with least compromised records. On the other hand threat of insiders is least with 18% but the impact they cause is greatest.

Reliability of data storage (D8)

As long as there exists no problems with virtualization manager, developer will have better control over security. The virtual machines have many issues within it, but it still a good solution for providing secure operation in CC context. With growing virtualization in every aspect of cloud computing, there is an issue with reliability of data storage and owner retaining control over data, regardless of its physical location [144]. The users also think that storage mechanisms are not reliable in CC [33]. In CC, reliability of data storage is a general issue [99]. This
issue constitutes to every data entity stored in the cloud and when Infrastructure as a service (IaaS) is considered the service provider is expected to assure that an organization’s data is kept secured along its life cycle (even after the user has removed his account) [38]. Another issue to be considered here is that even a virtual machine needs to be stored in a physical infrastructure. This can also cause security risks, which needs to be protected [125][164]. In addition to these problems article [126] explains various storage concerns and show to what extent cloud can be dependable with respect to reliability of data storage.

**Data center operations (D9)**

Information from articles that discuss about losing control over data, incident response, data center operational management, data center operations, data management, disaster recovery, data transfer bottlenecks, incident response are grouped her.

Organizations using cloud computing applications are concerned about protecting data while it’s being transferred between the cloud and the business. The concern is about what will happen to user’s data if something happens to cloud storage? [162]. If data is not managed properly, data storage and data access can become an issue [15]. In article [54] author expresses his concern that though there is growing interest in cloud governance issues such as data integration, data consistency, policy management etc., are not given required attention. Adding to this [178] mention that cloud is not secure unless mechanisms to debug, diagnose distributed queries and analyze exists for the cloud providers.

Explanation on issues related to data center operation are discussed below:

Recovery: In case of disaster, the cloud providers should be answerable to users questions, such as what happened to data stored in the cloud? If the cloud service provider does not replicate data across multiple sites, such system could result in failure under certain circumstances (if disaster’s occur)[119][10][162]. Even though there is a disaster in cloud it should be able to recover or provide some other means to avoid halting in the business needs of user [65]. Therefore, service providers can be questioned “is there any option for complete restoration if exists some way and how long does it take?” [119]. In case of PaaS, disaster and recovery are issues that need to be paid more attention and in IaaS data center construction is a key point [164].

Losing control over data: Data is outsourced, control of data will be lost gradually. To avoid major loss there should be transparency in how data is managed and how data is accessed [142]. As a solution to this, Amazon Simple Storage Service S3 API’s, provide both object level access control and bucket level control. Each authenticated user under this kind of security in a system is authorized to perform certain actions, specific to each object which he/she needs access to perform his/her task. In case of bucket level control the authorization is granted at the level of bucket (is a container for objects stored in Amazon S3).
This security can be accessed at both object and bucket level [46].

Data transfer bottleneck: If potential consumer transfers data across the cloud boundaries, it might lead to data transfer cost. To reduce cost, when CC applications are used cloud user and cloud providers need to focus on the implication of data placement and data transfer at every level. During amazon development, ship disk was used to overcome this issue with Data transfer bottleneck [13][12].

Data sanitization (D10)

Information from articles that deal about data sanitization and insecure destruction or incomplete data deletion are considered.

It is the process of removing unwanted or outdated sensitive data from the storage device. When a user updates data in the cloud, he/she can secure the data by encrypting while storing data on the cloud. Users are very much concerned about what will happen to the data after it passes its user’s “use by date” date, will it be deleted after the contract is completed? [69][46][78][7]. Even if the data has to be deleted or no longer needed should be deleted in a secure way such a way that unauthorized access is not possible [126][100][142]. It is also a benefit if the user is kept informed how his data is deleted (if asked for deletion) and also helps user keep informed if the service provider is keeping the data even if asked for deletion [119]. Amazon Web Services (AWS) procedure includes a decommissioning process when the storage device reached to the end of useful life. It also means that the user data is not exposed to unauthorized users. Sanitization is also applied to backup the data for recovery and restoration of service [78]. While performing an anti-malware scan identifying useful piece of data and deleting unwanted information can become complicated [138].

Example: Researchers are obtaining data from online auction and other sources. They also extract data by retrieving large amounts of data from them (with proper equipment, we can recover the data from the failed drives which are not disposed by cloud providers) [78].

Data storage and lock-in (D11)

Information from articles that deal about data storage, where data is stored?, problems with DBMS and data lock-in are considered.

Data storage is a concern in CC [54]. There are a number of data storage concerns expressed by different authors and as follows:

- Is the data secure? Will data be available when requested? [136][92].
- Authors of [90][5][142][162][17][110] have mentioned data loss or leakage as a challenge, concern or issue in their articles.
- Many doubts such as how data is stored, where is it stored (is it distributed in various places?), what will happen if the cloud provider is
Chapter 4. Literature Review

taken over, what security measures are taken to protect the user’s data etc., [162][119][109][45].

- Since a large amount of data is stored in cloud it can attract attention from malicious users [166]. Due to this user might not be interested to store mission critical data for processing or storing into CC [55]. Some others feel that data stored in CC as insecure and not reliable [33][108].

- Customer lock-in might be attractive for service providers, but it is an issue (viz., vulnerable to increase in price, reliability problem, etc..) for customers [13][12][92][110]. As an example customers of ‘Coghead’ were forced to re-write their applications on another platform when its cloud computing platform got shut down [35].

- It’s a known fact that different organization’s data is stored in a place and this increases the possibility that an organization data is sold to another organization for money. The author wants to bring it to notice that since all the data is stored in the same place there are increased chances of data misuse, data can be intentionally leaked and if this happens the customer is at loss [100][133].

- The author of [142] mentions that data loss/leakage issue affects only public cloud.

The issues with data storage in CC can arise when proper sanitization or segregation of data is not implemented, which could result in users not able to extract their data from repositories when necessary or when company willing to shift data to another location.

Data storage: With cloud computing users can utilize a wide variety of flexible online storage mechanisms to store their information, which have been known as computing and storage “clouds”. Examples are Amazon S3, Nirvanix CloudNAS and Microsoft SkyDrive [26]. The architecture of storage mechanism also depends on cloud type e.g., internal or external cloud computing and types of services i.e., SaaS, PaaS or IaaS. This variation exists because in internal cloud computing organization keeps all data within its own data center but in case of external cloud computing data is outsourced to CSP [132]. In any of these cases the data is not under the control (physical or logical state) of user and traditional cryptography can not be used by user [68][95][164]. Since in CC data is stored in a remote location and even traditional cryptographic algorithms cannot be applied the security of data stored in remote locations is a huge concern in cloud computing [153].

Users while accessing flexible storage mechanisms can maintain a local trusted memory, use a trusted cryptographic mechanism and upload data into the cloud (by doing this the user need not have to trust the cloud storage provider); to
verify the integrity of data user can have a short hash of in local memory and authenticate server responses by re-calculating the hash of received data [26].

To create trust in cloud storage, data storage systems need to fulfill different requirements such as maintaining user’s data, high availability, reliability, performance, replication and data consistency; but since these requirements are all interrelated and conflicting no system has implemented all of these at once. DaaS providers facilitate storage as a service by implementing one feature over other, which is mentioned by its customers in Service Level Agreements (SLA) [168]. To assure security of information, CSPs should take care to protect the data not only when it’s stored but when its under transmission [31][123]. In article [11], author mentions storage devices should support different storage patterns. In addition to this various other concerns in relation to data storage specific to IaaS are explained in [38]. To have a better security in cloud storage long term storage correctness and remote detection of hardware failure is suggested [126].

Table 4.2: Cloud computing challenges related to data

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data related</td>
<td>Data Security</td>
<td>D1</td>
</tr>
<tr>
<td></td>
<td>data locality</td>
<td>D2</td>
</tr>
<tr>
<td></td>
<td>Data integrity</td>
<td>D3</td>
</tr>
<tr>
<td></td>
<td>Data segregation</td>
<td>D4</td>
</tr>
<tr>
<td></td>
<td>Data access</td>
<td>D5</td>
</tr>
<tr>
<td></td>
<td>Data confidentiality issue</td>
<td>D6</td>
</tr>
<tr>
<td></td>
<td>Data breaches</td>
<td>D7</td>
</tr>
<tr>
<td></td>
<td>Reliability of data storage</td>
<td>D8</td>
</tr>
<tr>
<td></td>
<td>Data center operations</td>
<td>D9</td>
</tr>
<tr>
<td></td>
<td>Data sanitization</td>
<td>D10</td>
</tr>
<tr>
<td></td>
<td>Data storage</td>
<td>D11</td>
</tr>
</tbody>
</table>

4.2.2 Networking

Network security (N1)

Information from articles that discuss about network security and VPN network are considered.

Data should not be leaked while transmission and it is one of the requirements in information security [136]. To prevent leakage of sensitive information while transferring, a strong traffic encryption technique such as Secure Socket Layer (SSL) and the Transport Layer Security (TLS) are required. Sensitive data are obtained from enterprises, processed by any service application and stored at the service vendor end. Amazon Web Services (AWS), provide more protection to its users from traditional network based attacks like MITM (Man-In-The-Middle)
attacks, IP spoofing, port scanning, packet sniffing, etc. The assessment tests to find vulnerabilities in security are based on the following [144]:

1. Network penetration and packet analysis
2. Session management weaknesses
3. Insecure SSL trust configuration

VPN network: If an organization is distributed globally and employees a single vendor, then such organization might experience lower transfer rates when sending a file from one side to another side. A solution to this is usage of Virtual Security Gateway and maintaining multiple vendors, for implementing this usage of some commercial solutions that give customer-controlled security in a cloud is necessary. This helps to establish a bridge over private infrastructure, where control over cloud lies within the organization. It enables confidentially leverage over the cloud for redundancy, scalability and failover during critical transitions, which may lead to scale up grow or scale down to the organization or business [3].

Given below are some of the network attack types, concerns and suggestions that various authors mention that can be aggregated under network security:

- The author in [14] mentions how challenging it is to identify the intrusion severity in CC.

- Author of [90] mentions attack methods such as phishing, fraud, Denial of Services (DoS) and account hijacking that are used to steal user credentials.

- The authors of [51] mentions that using traditional network based methods such as IPSec proxies encryption and digital signature, key exchange through SSL proxy which are still being used in the cloud are insecure. The authors mention vulnerabilities of using such methods in the article by considering Amazon Web Services (AWS), Azure and Google App Engine (GAE) as examples.

- Attack types mentioned in [169] launching dynamic attack points, key and password cracking, hosting malicious data, botnet command and control, DDOS, building rainbow tables and also CAPTCHA solving farms.

- The author in [134] mentions, backdoors, TCP hijacking, replay, social engineering (where the attacker tries to gain private information from user’s social behavior), dumpster diving (information from trash can be helpful), password guessing, trojan horses and malware are some of the network attacks mentioned.
• Metadata spoofing is another kind of attack where a new system similar to cloud system can be built by analyzing and re engineering from metadata [85].

• Account or service hijacking [17][5][110]. In addition to these Dos, IP address modifying helps malicious users to hack into accounts [15].

• Some attack types that are specific to IaaS are DDOS, MITM, port scanning and IP spoofing [38]. The author of [164] points out network security, transport security as key security issues for IaaS in CC.

• Side channel attacks and incident handling [110]. SQL injection and phishing by service provider [7].

• The problems account or service hijacking mentioned above affects public and not private cloud [142].

In addition to all these kinds of attacks, the author in [75] mentions that even the cloud firewalls are still under chaotic state. From all the attack types and point put before us there is a strong requirement for security measures, anyone’s credentials can be stolen easily with all these different types of security risks pointed out. If an attacker can gain access to someone’s credentials, he or she can eavesdrop on a customer’s activity, transactions and also become a huge threat to user’s data [51]. A server side protection which includes application security and virtual server should also be provided to strengthen network security [92].

Application security (N2)

Information from articles that discuss about web application security, API security, application vulnerabilities and application security are considered.

With new advantages cloud also bring to the developers novel vulnerabilities and threats related to APIs [5]. It is known that in the cloud, any application or software that is used lies in cloud but not with the actual user and if this software/application has vulnerabilities then it can have a decremental impact on all the customers using the cloud [142][17][164]. If the insecure APIs are not secured the vulnerabilities in that can lead to compromising security, can also lead to men in the middle of attacks and affect the availability of CC [166][169][139]. Security surveys bring to notice that application level attacks are more bandwidth efficient compared to network level attacks [65]. In [110] the author mentions insecure APIs as one of the top threats to CC. Traditional security mechanisms such as network firewalls, network intrusion detection and prevention mechanisms do not adequately satisfy being used as a solution [70]. The typical issues arising with applications technology are: Session riding, hijacking and injecting vulnerabilities. Other web-application-specific vulnerabilities are browser’s front-end components in which, data sent from the user component to server component is
Chapter 4. Literature Review

manipulated [144]. The author of [31] also mentions that modern browsers compatible with AJAX is not secure though it’s good for I/O and data encrypted by XML needs to be also encrypted by using transport layer security. XML signature attacks, browser based attacks for cloud authentication are few other insecurities in application security that can be used for cloud authentication [80]. One easy way of access cloud for an intruder is to search for saved passwords in history which can be a huge risk and using insecure browsers [123]. Hence securing the browser is necessary to secure transaction between users and his application that are stored in the cloud [92]. The application security is the main threat to SaaS platform [164]. In article [142] author mentions that both public and private clouds are affected by insecure applications.

Host and Network intrusion (N3)

This problem arises in PaaS, where control might be given to user by SP to some extent. Service providers should keep in mind that control below the application level, such as host and network intrusion shall not be given to the user and the provider should maintain inaccessibility between applications [144]. By using programs such as trojan horses and malware which leak sensitive data can help intruders to gain access to sensitive information [134][80][67].

Denial of service (N4)

Denial of service attacks are possible in cloud which can be a threat to data under transmission [32][136][169][105]. Unlike bypassing preventive and security measures the attacker uses methods such as packet splitting, payload mutation, shell-code mutation and duplicate insertion [65]. The three proposed distinctions [159][24][46] are:

- **Direct denial of service**: If the workload becomes high in cloud computing operating system, the environment will automatically allocate more computational power (more service instance, more virtual machines) to overcome this additional workload. But since resources are limited with service provider even providing additional resources for an extended amount of time is not quite possible. In such case (flooding scenario) the cloud system is not working against the attackers (providing more computational power) but to some extent system is giving attackers rights to do damage on service’s availability, starting from a single flooding attack. The attacker to cause disturbance does not need to attack ‘n’ servers, an attacker can target one of the cloud based address in order to attack ‘n’ servers and perform full loss of availability [80][67].

- **Indirect denial of service**: In-direct denial of service, attackers’ damage availability of a service. If an instance in cloud service is flooded with too
much work load, other services running on the same cloud server might face issues with availability. If resources are completely utilized by a single instance, other instance has to face consequence and wait for resources to be allocated.

Depending on level of sophistication, if cloud service provider notices lack of availability and tries to “evacuate”, affected instances on to other servers the complexity rises even further. The reason for over complication is, flooding attack which existed in previous server will now get extended to the server where it is transferred and this situation is called “jump over” [79][80]. In the worst case, to manage flooding attacks utilization of another cloud computing system is suggested. In that case, cloud should provide more computational resources, respectively fending, until it reaches the full loss of availability [80][67].

• **Distributed Denial of Service (DDoS):** This is even more dangerous and difficult to identify or control. In this type of attack the attacker floods the target servers by controlling a large group of geographically distributed hosts. Since Cloud is shared by a large group of users, DDoS could have greater impact on customers [100][83].

**Men in middle of attack (N5)**

Attackers create an independent connection which relay on the messages between user and provider. The attacker makes the user and provider believe that this connection is secured and makes them talk directly, but behind the scene attacker controls the whole connection and receives every message which is sent between them (user and provider) [159][46][100]. The attacker can also possibly modify the message before sending to the respondent [134]. XML signature wrapping is one such attack where the attacker without changing the header modify the data in it and sends to the receiver [31]. This makes clear that in cloud computing, while data is in transfer state denial of service is a threat [32].

**IP Spoofing(N6)**

Using somebody’s IP address and creating TCP/IP packets is called IP spoofing. In this scenario intruder gains access to trusted system and sends messages as if trusted host is sending these messages. Here the malicious user impersonates the actual user with the help of IP address [169][134]. Amazon EC2 cannot send instance spoofed network traffic. Firewall designed in amazon web server prohibits an instance sending messages with any other IP address or MAC address other than its own [46].
Port scanning (N7)

A port is a place from which data goes in or out of a system. When a system’s security fields are configured to send or receive data through a port. Then that specific port is vulnerable to port scan. When a network is scanned for vulnerabilities port scanning shows these vulnerable ports as open doors. When a computer tries to access internet, a port is opened by default and it is not possible to stop port scanning [130][46]. This can cause security breaches in cloud computing platforms.

Packet Sniffing (N8)

Packet sniffing is listen to network devices and capture raw packets [169]. If a software finds a packet suitting a particular criteria, it logs into a file (login and password may be mandatory and significant ones). A virtual instance running in promiscuous mode can’t receive or “sniff” traffic which is intended for a different virtual instance. The hypervisor will not deliver any traffic when customers place their interfaces into promiscuous mode. If two virtual instances are determined on the same physical host and are owned by same customers then they cannot listen to each other’s traffic. For Amazon EC2 attacks such as Address Resolution Protocol (ARP) cache poisoning do not work at all. Protection is provided by Amazon EC2 against customer who are attempting (maliciously) to view other’s data, sensitive traffic should be encrypted by customers as a standard practice [46].

Sharing computing resources (N9)

Sharing technology is a top threat to cloud computing because it inherits all the problems that are possible with sharing resources in stand alone systems [5][17][110][71]. In the world of cloud computing information is stored in data servers which are globally distributed. This cloud computing architecture is supported by virtual machines that run on hypervisors [90]. Due to this the user will lose control of physically securing data and this might result in security risks because this data is stored in a location where resources (storage, computational resources, etc.,) are shared with some other organizations [119][34]. Sharing resources between different projects and products and remote storage and processing of data can be beneficial but there are also some risks (such as how data is handled and misuse in command) and can complicate computation (i.e., monitoring, analysis and reporting for company needs) [50][169][123]. In addition to the complications, there are various challenges that author mentions need immediate attention when sharing computing resources, such as for tenant-isolation, customer-specific application customization etc., that can effect customers if not using proper measures [50]. The issues that arise with shared technologies effect both public and private cloud [142].
Table 4.3: Cloud computing challenges related to network

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network related</td>
<td>Network security</td>
<td>N1</td>
</tr>
<tr>
<td></td>
<td>Application vulnerabilities</td>
<td>N2</td>
</tr>
<tr>
<td></td>
<td>Host and network intrusion</td>
<td>N3</td>
</tr>
<tr>
<td></td>
<td>Denial of service</td>
<td>N4</td>
</tr>
<tr>
<td></td>
<td>Men in middle of attack</td>
<td>N5</td>
</tr>
<tr>
<td></td>
<td>Ip spoofing</td>
<td>N6</td>
</tr>
<tr>
<td></td>
<td>Port scanning</td>
<td>N7</td>
</tr>
<tr>
<td></td>
<td>Packet sniffing</td>
<td>N8</td>
</tr>
<tr>
<td></td>
<td>Sharing computing resources</td>
<td>N9</td>
</tr>
</tbody>
</table>

4.2.3 Virtualization

Vulnerability in Virtualization (V1)

Information from articles that discuss about vulnerability in virtualization and cross-vm information leakage are considered.

The most commonly used multiple way to create multiple virtual machines on a single physical machine is done using Virtual Machine Monitor (VMM) approach. Hypervisors are also used to manage multiple VMs and any flaws in the hypervisors can allow attacker to gain access in an inappropriate way, even when tools such as Xen access are used many security risks can be found, which allow admin to see through the user level process while the customer is running his VM and attacker can easily install a malicious code [90][38][105]. In article [144][50] the author mentions that major risk is to ensure different instances running on the same physical machine are isolated and this requirement is not yet met completely. If virtual machines have vulnerabilities then sharing such hardware can result in many vulnerabilities [15]. To control and administer host, guest operating system is another issue which needs attention. The author also argues that most of the current existing systems do not satisfy perfect isolation and many bugs found here are from Virtual Machine Monitors (VMM’s). These bugs can be exploited by malicious local users and bypass some security restrictions or gain privileges. Ex., as been explained by the author “vulnerability in Xen caused due to an input validation error in tools/pygrub/scr/GrubConf.py”, which can be exploited by ‘root’ users of a guest domain to execute arbitrary commands in a domain using specially designed entities in grub.conf (when guest is booted) [144][84][44][82]. If an unauthorized user hacks into CC, he can access all the virtual machines running in CC and control them [166]. Since CC is relatively cheap and provides resources to users easily and also since virtual machines (which is a main part of CC) are easy to copy or clone, the VMs can be easily used to launch new attacks [109][67].
Chapter 4. Literature Review

The following are some of the vulnerabilities by different authors mention in their work:

- Author for [134] mentions that trusted hypervisor is more important than having a trusted VM, because an untrusted hypervisor can affect to a greater extent.

- The author in [123] expresses that increased usage of virtualization is a threat.

- In [50] the author mentions insecure areas in virtual machine which need to be secured and they are version management, customer specific application customization, migration between service providers, etc.

- An attacker can also cause damage to a cloud that is connected to a virtual machine by installing a bug in the source machine [126].

The above discussion makes clear that, if VM are not constantly monitored there can be service breakage [136].

Virtual machine protection (V2)

Information from articles that discuss about virtual machine protection and securing virtual machine boundaries are considered.

Multiple VM’s can be instantiated or halted in a single server (machine) to satisfy list of services accepted. These services can also run multiple applications which are based on different operating system environments [25]. In relation to this author in [6] expressed his concern towards securing the boundaries of virtual machines. Since the virtual machines created in cloud server have virtual boundaries (unlike general isolation where multiple hard disk drives), it’s the responsibility of CSP to ensure VM’s that use common resources on the same physical server (i.e., CPU, memory, I/O, NIC and others) are separated [6]. Also the host needs to be secure, because if the host is not secure protecting offline/online VMs is going to become a challenging task [38].

Virtualization in the cloud: The author expressed views on virtualization has been vulnerable to breaches, be it on traditional infrastructure or a cloud architecture. This virtualization is a process of creating multiple Virtual Machines (VM’s), on a physical server to use a shared pool of resources, which also raises a concern on possibilities of spreading malicious activity. To prevent this kind of malicious activity in physical environment segmentation can be used, but in case of virtualized environment there is no possibility of segmenting. But without segmentation cloud could possibly allow hacker in web systems jump over to financial systems or databases. The solution to such kind of problem in virtual environments is to use Virtual Security Gateway, this enables users to apply some critical rules, log-in and access privileges, similar to what is done with systems.
placed in-house [3]. Supporting this author in [74] mentions that possibility that dynamically assigned virtual spaces in CC can already have some security threats.

**Networks in virtual networks (V3)**

Information from articles that discuss about networking in virtual networks, networking problems, virtual management, virtual network security and hypervisor security are considered.

Virtual machine instances interconnectivity (i.e., communication) is a huge concern in CC [75][38]. Traditional mechanisms such as VLAN’s (Virtual Local Area Network) and firewalls are proving to be inefficient when used in a virtualized environment [73]. The security of a computer depends on the quality of underlying security mechanisms or kernels which control the execution of process [78]. Also when the devices are virtualized user might lose the visibility of how the VM’s are connected and insecure information transfer can be possible (information of one customer being disclosed to another customer) [163][138]. In addition to losing visibility, when considering security at the network level in virtual machines there arise two issues as discussed in article [6]. One of them is virtual network security, when sharing network infrastructure between different users within same server or physical network can increase possibility to exploit various vulnerabilities (DNS servers, DHCP, IP protocol, etc.) [6]. The other issue mentioned by [6], is on hypervisor security, this is the ‘virtualizer’ that links physical resources to virtualized resources and virtual machines in a virtual environment are given access to physical resources by this hypervisor. Hypervisor provides isolation between the different guest and virtual machines. Since these hypervisors are virtual machines that run on physical machines and also smaller in comparison to physical systems, this helps in controlling and isolating processes. But now a days hypervisors or virtual machines are large, complex and comparable to an operating system [78]. Hence understanding hypervisor security is critical, any compromise on these machines can effect security of all users connected through it [6]. Examples of hypervisor software’s are VMware or Xen. Also the author in [56] mentions that the auto arrangement of virtual resources can lead to problems and suggests that SPs need to perform security checks and constantly verify if user’s security requirements are met or not. One way to see that VMs are protected and not controlled by malicious users, administrators since they are the key persons in managing VM, the CSP should regulate the administrators from time to time [166].

Within most of virtualization platform existing there lies inbuilt ability, to create software based swatches and network configurations as part of the virtual environment and provide communication between virtual machines which is more efficient and direct. Traffic over virtual networks may not be visible to security protection on physical networks, so there is a necessity to duplicate such services in virtual networks [78].
Some common attacks that can effect virtual machine’s network are “buffer overflow, DDoS, zero day attacks, viruses, covert channels, trojans, etc” mentioned, which can infect the hypervisor as it is the major VM controlling entity [5].

**VM isolation (V4)**

Inter positioning is an inherent feature of VMM, isolation of VM is a key concern as it is not that possible CC and active inspection mechanisms are yet to be designed [73][75]. Multi-tenancy is a process where multiple cloud customers have their virtual instances running on the same physical server. In such environment, there might be a possibility for malicious user allocating an instance in same the server and possibility is that he/she (malicious user) might penetrate the isolation between the VM’s and exploit them [128]. This could be possible by creating a side channel between two VM’s on a same physical machine and enabling a SSH keystroke timing outlined attack [34][141]. There is also a possibility that data accidentally crosses the virtual boundary [136]. The author in [126] mentions that even in an isolated environment if an attacker interferes with the VM placing strategy and position their instance on a physical machine of the victim, all the private information of target was visible.

To secure VMs, the VM instances need to be isolated, but at what level isolation should be implemented i.e., is it at VM, physical machine, local area network, VMM or at data centers [73]. The author in [105] mentions that logical separation need to be validated. Will that be sufficient? But there are more problems to handle.

Author in [38] mentions about protection from the interconnection of the VM, which connections are going into and which are coming out. The author explains various issues that can arise with existing technologies. The author [38] also mentioned there could be issues arising in the following areas of VM if isolation of VMM’s is not done properly by:

1. Monitoring VMs from the host
2. Communication between VM and host
3. Monitoring VM’s from other VM’s and virtual machines in mobility.

**VM Securitization (V5)**

Information from articles that discuss about VM securitization and VM security are considered.

VM introspection or behavioral monitoring in virtual machines requires a lot of computational power. It is easy to clone and distribute VM’s between physical servers but this distribution could result in propagation of configuration errors.
Maintaining a hypervisor level, Media Access Control (MAC) and trusted computing techniques are suggested mechanisms used to build future secure cloud systems. In IaaS virtual machine security lies with the cloud consumer, a consumer can use their own security controls based on situation or security management process and expected risk level assessment [6].

\textbf{Table 4.4:} Virtual machine related challenges in cloud computing

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtualization Related</td>
<td>Vulnerability in virtualization</td>
<td>V1</td>
</tr>
<tr>
<td></td>
<td>Virtual machine protection</td>
<td>V2</td>
</tr>
<tr>
<td></td>
<td>Networking in virtual networks</td>
<td>V3</td>
</tr>
<tr>
<td></td>
<td>VM isolation</td>
<td>V4</td>
</tr>
<tr>
<td></td>
<td>VM securitization</td>
<td>V5</td>
</tr>
</tbody>
</table>

\subsection{4.2.4 Organization}

Organizational security management (O1)

When adapting to cloud computing, some changes are introduced to the security management, information security lifecycle models, even the corporate IT standards and policies need to be changed [160]. There are issues such as less coordination among different communities of interest within client organizations. The customer also has to face new risks introduced by perimeter-less environment, such as data leakage due to multi-tenancy, issues like local disasters and provider’s economic instability. But since the cloud computing environment is distributed in nature, re-evaluate best practices and adoption of secure cloud computing applications become extremely complex as they require to have a well structured cyber insurance [146]. Another way is to adjust to the new features provided by the cloud computing else gaining complete benefit from CC would not be possible [50]. Some point needed to be considered and as follows:

- Will proprietary interfaces used to administer services provided be accessible? [65]
- Security models must be assessed in detail [142].
- Gather and manage security related information on non native interfaces such as CC is important [91].
- Some organization policy and legislations are non compliant with CC, example enterprise network which uses NAT (Network Address Translation) and company specific firewalls and is not easy to integrate with CC [97][49].
Trust management and Policy integration (O2)

Trust is the key problem in the CC environment as CC is not completely trustable for users [164][136][67]. An organization using cloud computing application has to give some of its critical information to use certain services provided by the cloud service provider. For this organization needs to develop trust in remote execution/storage and placing sensitive data in the cloud is tough as the company feels that its losing control over data and there is much less transparency in cloud [89][77][4][32]. Lack of this control is leading to distrust in CC [123]. The remote execution in CC is carried out based on some policies and principles. There are evidences that even though the policies at individual domains remain verified, the violation is still possible to occur at integration that needs to be taken care of or have some attention. Hence, there is a need for a trust based secure interoperation which helps in managing trust and support policy integration [146]. Insider access, data stored or processed outside organizational boundaries and security controls are some features which when used with cloud computing could have some inherent risks. Threats that arise with the insider access are well known for various organizations. Its influence over cloud computing has not at all changed [78][84]. These threats could be raised by a user or a third party organization that was given access to a part of application in the cloud. These threats could result in theft of critical information or various types of frauds [78]. The CSA mentions that violation in cloud can be caused by a third party provider when it tries to access users data stored in the cloud [7]. The author for [4] mentions that trust in customers can be increased when SP develops trust worthy tools, automated processes for working in cloud and methods that establish healthy connections between user and provider. The author also mentions other attributes of CC such as, compositional chain of trust, trust re-evaluation and transparency vs trust evaluation [4]. The author of this article illustrates multiple examples of possible threats and possible security measures [179]. Since there are no extensible trust building models available in CC, SLAs should be monitored and enforced properly as they are one way to build trust between user and provider [38][64]. Mutual auditing users/providers in the CC will help in increasing trust and assurance on others in the cloud [110]. In article [148] author suggests a need for trust building approach to negotiate, establish and maintain trust. He mentions that this should adaptively support policy integration. The author poses some questions that need to be answered [148]:

1. How to manage and maintain dynamically changing trust values

2. How to adapt to the access requests as trust keeps evolving

3. The most important one, how to establish trust and determine assess mapping which satisfies inter-domain access requirements
Failure in providing security (O3)

Information from different articles that discuss about fault tolerance and failure in providing security are considered. Failure in providing security to the infrastructure under control of cloud provider can result in compromising subscriber’s security. Even a single weak link in cloud computing can cause a security threat to multiple entities connected in it. For customer to secure his data, he/she should believe in service provider’s security [154][87][46]. It is evident that customer must also have trust in providers security. For a cloud to be trusted and considered strong simple but yet important features such as logging, security policies, incident response, etc., should also be strong [90]. A third party poor management can lead to damaging the provider and also the user indirectly [138].

In addition to the above discussion, article [127] gives us a few examples that are tabulated in table 4.5, which exemplify different cloud failures.

<table>
<thead>
<tr>
<th>Service and outage</th>
<th>Duration</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft Azure; Malfunction in windows Azure</td>
<td>22 hours</td>
<td>March 13-14, 2008</td>
</tr>
<tr>
<td>Gmail and google apps engine</td>
<td>2.5 hours</td>
<td>Feb 24,2008</td>
</tr>
<tr>
<td>Google search outage:program error</td>
<td>40 minutes</td>
<td>Jan 31,2008</td>
</tr>
<tr>
<td>Google site unavailable due to outage in contacts system</td>
<td>1.5 hours</td>
<td>Aug 11,2008</td>
</tr>
<tr>
<td>Google AppEngine partial outage:programming error</td>
<td>5 hours</td>
<td>Jun 17,2008</td>
</tr>
<tr>
<td>S3 outage: authentication service leading to unavailability</td>
<td>2 hours</td>
<td>Feb 15,2008</td>
</tr>
<tr>
<td>S3 outage: Single bit error leading to gossip protocol backup</td>
<td>6-8 hours</td>
<td>Jun 20,2008</td>
</tr>
<tr>
<td>Flexi Scale core network failure</td>
<td>18 hours</td>
<td>Oct 31,2008</td>
</tr>
</tbody>
</table>

Identity and access management (O4)

Information from articles that discuss about identity management system, identification, identity and access management, identity provisioning (Independent IDM stack, credential synchronization and federated IDM) are considered.

Identity management is an administrative process area, where all users of a system are identified (entities), and by enforcing some restrictions on these entities access to resources in that system are controlled [144]. Since this plays a key role in securing Cloud Computing (CC) environment access should be given only to privileged users, which should also be blackened by physical monitoring and background checking [92]. The major concern in this is how complex it
is to manage a diverse population of users, and provide access to internal and external services in a constantly changing environment (changes can happen due to change in business needs/processes) [104]. It is also true that need for fine grained access control on the information stored on CC is increasing and since CC is vast satisfying these needs and meet growing requirements has become complicated [170]. For this there are various mechanisms but not all are up to the satisfaction of users [170]. The browser based authentication protocols are not secure as they cannot issue XML based security tokens [75]. In addition to this, organizations may contain complex web of user identities, access rights and procedures.

Author of [126][105] says access control in CC is an issue and it is necessary that a user is able to view only the portion of data he is given access. Access control models should also be able to a pick out relevant areas of SLA and change access rights accordingly [148]. Generally organizations give access to its employees based on principle of least access, where access is given to only those services to which he/ she needs basic access [119]. Identity is the core of any system and to have a manageable control on users and identify breaches, there are various areas need to be considered: “Identity Provisioning and deprovisioning, identity information privacy, identity linking, identity mapping, identity federation, identity attributes federation, single sign on, authentication and authorization” [6]. Identity management can be done in three different ways [144]:

- To just create, manage and delete entities without regard to access or knowledge regarding the access rights of the entity. This is known as a pure identity paradigm.

- Access can be provided to users with the help of some smart cards or a traditional scheme of authenticating. This enables user to log on to a system for accessing a service or services, which is other way, called as the user access (log-on) paradigm.

- Providing access to users on various devices based on different distinctions such as Role based, online, on-demand, multimedia content and presence based, is termed as a service based paradigm.

A traditional way of having control over an application was to keep track of collection of users and manage them and they are still not sufficient in CC [170][112]. This system is no longer easy as sharing of Personal Identifiable Information (PII) is not suggested because doing so can cause a breach to privacy. The PII is the information that a user sends to Service Provider (SP) to authenticate it. PII differs from user to user [124][18]. Some suggested methods for managing identities are [144]:

- Independent IDM stack: In this system, all information related to user accounts, passwords, etc., is stored with the Cloud Service Provider (CSP) and it provides management and sign on services.
• Credential synchronization: In this system, user account information is replicated between enterprise and service provider. The user account information is done at each guest who pays for the services provided by the CSP. This replication is done at each user based on their company regulations. Replication also considers relevant portions of user credentials at the CSP to provide sign on and other access control features for services used in CC.

• Federated IDM: Within this kind of system, user credentials are stored and managed at the enterprise itself. Authentication is done in enterprise boundary itself and if necessary, certain attributes are communicated to the CSP [144][165].

Another issue here is that companies having a strong IDM system might not be able to extend its functionalities on CC or change the existing framework to support this new architecture of cloud [148]. To separately maintain a different authentication for internal (organization authentication) and external (authentication at SP’s end) is a complex solution to this problem [78]. To nullify the effect of these challenges on CC, systems need to find possible breaching zones/areas before they cause some impact. Given below are some assessment test to validate the security in IDM [144]:

• Authentication weakness analysis

• Insecure trust configuration

The author in [90] explains a special case, by taking an example of an employee with malicious intent, working in a company which provides CC services. This employee can easily gain access to secure areas and perform malicious activities which could be difficult for outsiders. The impact of this employee could be severe if this employee is not identified by the service provider. Supporting the previous example, article [142] mentions that more than 67% of respondents in a survey by CERT insider threat center 2010 cyber security watch survey admitted that insider incidents are the costliest. The author also suggests to use a insider security policy to avoid insider attacks [142]. The author for [166] expresses that insider attacks are alarming and this might happen when security standards provided by SP do not meet the standards used by customer’s standards. The author of [109] mentions that one third of breaches happen form lost or stolen devices and from employees sharing their data accidentally on the internet and nearly 16% due to insider threat. This issue of malicious insider has also been mentioned in following articles [17][5][105]. Not only there is an insider threat but even a malicious user might gain access to CC and deploy malware if he can produce a stolen credit card from legitimate user [138]. These insider threats or entry for a hacker or malicious insider is easy because SPs don’t give customers transparency into security in CC. Inefficient access and identity management can
Chapter 4. Literature Review

lead to illegal service process, fabricated services process and malicious service interruption [32]. The author for [43] also mentions various types of insider and the impact and also mentions that it can be dangerous if go unattended. Article [91] mentions Identity management, credential management, attribute management, privilege management, cryptographic key management as security management areas in CC.

Authentication and authorization: Most companies store their employee information in some kind of Lightweight Directory Access Protocol (LDAP). Small and Medium Business (SMB) companies which have SaaS as their highest adoption rate seem to use ‘Active Directory’ as the most popular tool for managing users. In case of SaaS, the user credentials get stored multiple times in provider’s databases, because SaaS requires the user data to be stored in SP’s data repositories. To avoid misuse SaaS customers should remember to remove/disable accounts or create/enable as employees leave or join the company [144].

Authentication and identity management: Identity management and authentication is being realized as a critical requirement in CC [66][163]. Users can be authenticated based on service based credentials and characteristics using identity management (IDM) mechanism. This system is supposed to protect private and sensitive information related to users and processes [162]. In multi-tenant cloud environments customer’s identity and authentication information should be segregated. These components need to be easily integrated table with other security component. It is not clear yet on how multi-tenant cloud environments can effect privacy of information [146].

Most of the service providers use Security Assessment Mark-up Language (SAML) in administering and authenticating user before giving access to resources. This language provides an interface where information related asserting subjects or authentication information is exchanged between cooperating domains. SAML request and response messages are mapped over Simple Object Access Protocol (SOAP) which in turn depends on extensible Mark-up Language (XML) for formatting. This SOAP message security validation can become complicated and threat prone. As this SAML depends on XML, there are some evidences where wrapping attacks on XML were successful and in turn that can be effecting SOAP message. This in turn can become a threat to the SAML authentication procedure [78]. The detailed description on challenges with implementation of other modes of authentication such as, single sign-on, OpenID, RBAC, etc., are given in [66].

Given below are some cloud specific issues mentioned in [62] article:

- Denial of service by account lockout: The most general way of authenticating a user is by username and password. To avoid misusing an account, after a certain numbers of failed attempts an account is locked. This feature can be used by an attacker to launch a Denial of Service (DoS) attack against an authorized user [144][38][20][143].
Chapter 4. Literature Review

- Weak credential-rest mechanism: Password recovery mechanism prevailing is proven to be weak, if a Cloud computing (CC) provider manages user credentials manually than using a federated authentication, there should be a method provided for resetting credentials in case of lost or forgotten credentials.

- Insufficient or faulty authorization checks: Web applications and service cloud are vulnerable to faulty or insufficient techniques of verification that can make users access unauthorized information. URL guessing attacks are the outcome of missing some authorization checks, the user can modify URL’s to display other user’s information.

- Coarse authorization control: Cloud service management interface needs to offer the authorization control models, which are still not refined. This makes standard security mechanism (duty separation) not so easy to be implemented in a cloud environment.

- Insufficient logging and monitoring: Present day cloud has no standards or mechanisms that can help customers to monitor and log data within cloud resources. This can cause problems when the log file needs to be searched for a single tenant and also there is not all possibilities from CSP’s side to monitor user’s data.

Authorization and authentication is an important information security requirement on which cloud architecture designer must concentrate and build a secure and trusted cloud computing infrastructure [123][8]. Inefficient authorization mechanism can cause side-channel attacks in the cloud [8]. The author in [171] gives an example of a special case where two parties using the cloud as a medium for transferring applications/code/information, only the target user is supposed to get access to the piece of program but with the help of an insider (an employer in cloud) the target user can create another copy of the same code and sell it. This is a copyright issue which can cost the developing company. Traditional encryption, hashes or access-control labels can not help to avoid this. Which gives an example that shows if access management is not given enough importance a user who is not fit for having access to a part of the transaction can misuse his rights and cause damage to customer. To avoid this proper monitoring should be done for employees in cloud.

Some suggestions to prevent illegitimate access are as follows:

- Mentioning human resource requirements as part of SLAs, transparency in overall information security and management practices, compliance, reporting and determining a security breach notification method and the most important enforcing strict supply chain management are some suggestions given by CSA [90].
• Before committing to a SP the customer must make a note about the privileged user access policies, know who will be gaining access to the data, how can it be accessed etc., [55][85][123].

• An effective and continuous monitoring system which detects malicious activity will be better than depending on firewall and security mechanism [105].

Availability (O5)

Information from articles that discuss about availability, cloud provider goes down, job starvation due to virtualization, business continuity and service availability are considered.

Availability is a primary concern for CC and is one of the key issues in information assurance [169][65][166][164]. Not being able to access services can become a serious issue [169]. It is evident in literature that the availability of a system is important and many organizations/researchers have their questions such as “if downtime occurs can the business still operate?” [162].

Availability of a system can be defined as, the extent to which an organization provides its user with round the clock service (resources accessible and usable) is availability. There are chances of having a complete or permanent impact on the organizations when the availability of a system is affected. Denial of Service (DoS) attacks, natural disasters and equipment outages are all threats to availability [78][134][31][100]. The cloud services also have experience in outages and performance slowdowns; if not planned and designed properly customers using cloud computing services will be at loss. There are various examples illustrated by [78], where the effect of shortcomings were borne by well-established organizations such as Twitter (in 2009), lighting storm caused some IaaS customers go down (in July 2009), gmail (in 2008 for one day) etc., [78][35][7], these outages can also occur because of non-technical issues [13][12]. Service interruptions could also be due to malicious users/use [32]. If at all there is an outage or server crash all the VM, services running on that server become unresponsive, in addition to that high availability or automatic fail over or high availability is not possible in public cloud [138]. If an organization wishes to migrate its servers which is so vast that it has multiple data centers distributed over different geographical locations using a variety of network providers but with some common software service to connect with a cloud, in such situation the organization might not want to migrate to a new system without a good business continuity plan modeled. In this case, it is suggested to have a multiple service providers providing different stacks, such that if one stack fails there will be another stack (i.e., service provider) supporting as a substitute [13][12][72][116].

Generally CSP has to provide required application level scalability and availability, by making necessary changes to the applications at the architecture level
and satisfy the customer requests. CSP can design specialized hardware and software to deliver higher reliability and earn an additional penny out of this [13][12]. CC should be constantly available with all the access control mechanisms and data access working [70]. To make this possible, a multi-tier architecture is deployed with a load balancing architecture running on a variable number of servers. To enhance the availability of a system built in CC, it is also suggested to secure system for hardware/software failures and DoS attacks. For safeguarding data from unplanned emergencies or minimizing enterprise downtime there is a need to have an appropriate plan for Business Continuity (BC) and Disaster Recovery (DR). For example the amazon API’s use a Standard Distributed Denial of Service (DDoS) mitigation techniques such as synchronous cookies and connection limiting. Amazon also maintains a higher level internal bandwidth than supplied to its customer, to further ensure mitigation of DDoS attacks. To validate the vendor’s availability, given below are few tests [144]:

- Authentication weakness and
- Session management weakness

Another solution for maintaining availability is, securing the system by usage of a Trusted Third Party (TTP). TTP provides end to end security services which are scalable across different domains and usable with different geographical areas. TTP’s are impartial organizations delivering business confidence through some technical security features and securing an electronic transaction. These TTP’s are generally connected through a chain of trust to provide a web of trust using the Public Key Infrastructure (PKI). PKI in a distributed information system becomes effective when combined with a directory (it is a set of similar attributes organized in a hierarchical and logical way). The Light weight Access Protocol is internet’s (LDAP) general way of accessing a directory service. In CC environments PKI used in combination with Single Sign On (SSO) mechanism is ideal as they provide smooth and transparent authentication across various physical resources. The reasons for depending on TTP’s are as follows [179]:

- Low and High level confidentiality.
- Server and Client Authentication.
- Creation of Security Domains.
- Cryptographic Separation of Data.
- Certificate-Based Authorization

Usage of TTP or any other method used for ensuring availability is generally mentioned through the Service Level Agreements (SLA). To ensure greater extent of availability it is important to identify the right deployment and delivery model.
set out by the International Organization for Standardization (ISO) [123]. Usage of HTTPS+WS security can cause heavy load on the server leading to service break down (reasons for outages in google and amazon in 2008), there should be perfect balance between availability and security measures to keep services active [123].

Table 4.6: Organizational security concerns that challenge cloud computing

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization related</td>
<td>Organizational security management</td>
<td>O1</td>
</tr>
<tr>
<td></td>
<td>Trust</td>
<td>O2</td>
</tr>
<tr>
<td></td>
<td>Failure in providing security</td>
<td>O3</td>
</tr>
<tr>
<td></td>
<td>Identity and access management</td>
<td>O4</td>
</tr>
<tr>
<td></td>
<td>Availability</td>
<td>O5</td>
</tr>
</tbody>
</table>

4.2.5 Scalability

Incompatibility issue (SC1)

Information from articles that discuss about portability, interoperability, scaling quickly, smart scaling, automatic scalability, security management function, collaboration, integrity of security, general requirement cloud security management model are considered.

For instance there might be a situation arising where a user wishes to change his Cloud Service Provider (CSP), in such situation the new CSP should be capable in moving data from an old service provider. This transfer includes data, components and also security policies. In most cases user won’t be provisioned to take physical firewall but users should make sure to have a copy of its configuration on virtual machines [3]. There are chances of customer goes locked to a SP if the customer chooses a wrong service provider [109][105]. Cross data center operations would find this as a serious issue [55]. In article [91] author mentions that controlling defining and ordering in relation to security are one of the ten things to be noted while shifting to cloud computing.

CSPs are creating the hosting world using “sticky services”, these services are reasons for causing difficulty in moving between different service providers. Example, the Amazon’s simple storage service S3 is incompatible with IBM’s blue cloud or Google or dell [24][130][46]. Open cloud manifesto which was newly published was declined by both Microsoft and Amazon and are pursuing interoperability on their own terms [46][16][24][173].

Things to be noted before/after moving from one CSP to another are [3]:

- In infrastructural service, backing up of data is easy compared to other services.
• In case of web application, there should be a perfect plan and process on how and when to move data. If all of your data is visible then it is not necessary to move it completely, but instead necessary data can be selected and copied.

• Application migration and reconsidering application infrastructure are areas to be considered by infrastructure architects. This is because some companies such as Microsoft are provider-specific and are not built to be easily reused [50].

When data is stored in the cloud, there are some virtual perimeters and security models with some shared responsibilities, among CSP and customer. These shared responsibility models will induce some changes into organizations IT staff. To maintain stability in the organization, the Chief Information Security Officer (CISO) is required to know if the cloud service provider allows its user to control and manage implementation of security policies and get assurance that business data in cloud stays protected. To assure data stays protected there are two considerations, one is that there should be an analysis made to find out what additional security controls need to be implemented on existing controls provided by CSP. The other point to be noted is how enterprise security management tools and process adaptation manage security in cloud [104]. Policy integration task in the cloud should also be able to address secure interoperability, semantic heterogeneity and policy evolution management [148].

Scalability of storage: One advantage with cloud computing is that it can provide infinite capacity based on demand of the user. It’s not yet clear on how persistent storage systems could be when it is practically implemented. The attempts to provide these scalable conditions was done by trying to vary richness of query and storage API’s. There is still a challenge in this area for the researchers i.e., not only to meet the expectations of programmers in providing durability, high availability and ability to manage data quickly but also to combine these qualities with cloud where advantages of scaling up and down can be used efficiently [13][12].

Interoperability: Interoperability is an important to maintain global stabilization to reduce rework and management [41]. Organizations willing to adopt cloud computing wish to have an Identity Management (IdM) solutions that can interoperate with all the existing IT systems and solutions, with or without small changes. Some of the most commonly used authentication mechanisms are Microsoft Windows authentication, SSO, LDAP, SAML, OPENID and OAUTH, OpenSocial, FaceBookConnect, etc., and cloud users expect easy integration with these types of services. To conclude CSP would have to provide an authentication module where each and every type of authentication system that user wishes, to use are easily comparable with cloud system [61]. Organizations are interested in CSP’s, who provide a framework which can enable interoperability with different other Service Providers (SP), Amazon API’s standards, are the most commonly
used standards to design other on demand instances. CSP’s which use these API standards are easily comparable with Amazon (Amazon EC2 can run easily in Eucalyptus). In certain situations cloud service providers which assure to provide interoperability with other SP is also problematic, example Hadoop which provides on demand capacity can’t for sure provide compatibility to run on another system developed using C++. Several organizations such as Cloud Computing Interoperability Forum and Open Cloud Consortium are still trying to provide firm and stable standards for CC. There is a framework named as ‘Thrift’ which relies on code-generation engine, to provide scalable cross-language services development. To provide interoperability and compatibility with various programs there could be a common language, which can be used by different service providers. As a solution to this several people attempted but there is no single language that is up to the requirement [63]. A panel report on cloud computing standards discussed about proposing “open standards and Predictive Markup Language (PMML)” and expressed views that implementing these could be extremely helpful not only for software vendors and data mining community in general [172].

### Constant feature addition (SC2)

In cloud computing scenario users need to be updated with new security features. These updates need to be done by the service provider and are required to be quick. The rate at which updates are installed can effect security as well as software development lifecycle [46].

### Scalability (SC3)

CC requires the ability to scale multiple transactions managed over multiple connections. Scaling the resources manually is still persistent with IaaS, where SP will define how the service has to be scaled based on personal experience and different factors influencing it [109]. There should be a reduction in management tasks and automatic scaling with-in the CC applications should be enhanced [131]. With software today it takes approximately 6 months for a single SAM-L/SSO connection, which does not address the compliance and access control issues [151]. Open cloud manifesto states that, to keep out of problems with allocation of resources during peak hours cloud services have to dynamically scale up and down [61].

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability related</td>
<td>Incompatibility issue</td>
<td>SC1</td>
</tr>
<tr>
<td></td>
<td>Constant feature addition</td>
<td>SC2</td>
</tr>
<tr>
<td></td>
<td>Scalability</td>
<td>SC3</td>
</tr>
</tbody>
</table>
Chapter 4. Literature Review

4.2.6 Confidentiality and Privacy

Confidentially (C1)

Confidentiality plays a major role in securing organizational content stored across different databases [65]. It’s a key issue and since most data is virtually accessed, protecting and maintaining confidentiality of user profiles is of utmost consideration [123][136][177]. Viruses, trojans, malwares, etc., are some unauthorized ways to exploit customer’s information [67]. In some cases it is also important for an organization to handle data remains, this is to protect the confidentiality of an employee’s information even after his data is removed or erased. Remanence is an issue that can lead to the disclosure of private data [179]. In article [48] author mentions that even though there are methods to provide confidentiality they are not widely used by SPs. As a special case in [142] author mentions that because of confidentiality agreements made with user, preventing/identifying malicious attacks is becoming difficult (top threat mentioned by CSA). This is because a SP can not monitor or look what’s happening in the user’s space due to confidentiality agreements, which can be exploited by malicious users for unauthorized activity [142][90].

To prevent data being disclosed or misused in cloud and protect customer’s data, every customer should be educated as to how data is stored in the cloud. By doing this the user will be careful while storing sensitive data in the cloud [160].

Privacy (C2)

Privacy is one of the cloud computing security requirement [92][65][74][109][160][54][126]. Keeping data private in a distributed system is challenging when compared to personal possession and in CC it is risky [149][23]. Privacy or obligation is related to the collection, use, storage, disclosure and destruction of data that is personal to someone. The rules and the concept of privacy varies with countries, cultures and jurisdictions [104][111][74]. The author in [179] mentions privacy as a desire to control disclosure to his personal information and presents that there are a number of legal challenges to cloud. Privacy is being accountable to an organization’s data subjects and also be transparent towards organizations practice around personal information, there is also a little knowledge on how privacy laws that govern within an organization [111]. There is no universal agreement towards defining what constitutes personal data. “The rights and obligations of individuals and organizations with respect to the collection, use, retention, and disclosure of personal information.”, is one way of defining privacy and this is gaining popularity among American Institute of Certified Public Accountants (AICPA) and the Canadian Institute Of Chartered Accountants (CICA) in the Generally Accepted Privacy Principles (GAPP) [104]. As a special case in [142] author mentions that because of privacy agreements made by the user, prevent-
ing/identifying malicious attacks is becoming difficult (top threat mentioned by CSA). This is because a SP can not monitor or look what’s happening in the user’s space due to privacy agreements, which can be exploited by malicious user for unauthorized activity [142][90]. And this case preserving privacy i.e., protecting the private data from unauthorized users in cloud computing while maintaining sharing of resources is a security issue [177][17][140][75].

Services can keep varying among customers and also the service providers in CC, since these services keep changing private data and personal data moves within an organization or could also cross organizational boundaries and protecting such information is important. Fields such as financial and health, are concerned about safety of data. Following are most important privacy risks, which need to be covered [117]:

- In case of a user, he might be forced or persuaded to give his personal information against his will [49].
- From organizations perspective, compliance with origination policies, legislation, creditability and loss of reputation are some other issue.
- With implementers of cloud platform, possibility of exposure of sensitive information that will be stored on platform, legal liability, loss of reputation and credibility, lack of user trust and take-up.
- For providers of applications designed over the cloud platforms, loss of reputation, legal noncompliance ‘function creep’ which uses the personal information stored on the cloud (i.e., it might later use for purpose other than that for which it was originally intended).
- In case of data privacy, exposure of personal information. To prevent loss of personal information there is a need for a special committee, which keeps track and makes decisions related to data privacy. If a security compliance team already exists within an organization it won’t be having formalized training on data privacy, a possible solution to this is to hire a privacy expert or train an already existing member well [119].
- Organizations can be held liable even if the subcontractor causes security breaches and CSP is legally considered same as that of a subcontractor. This confirms that organizations should ensure that a CSP is/are compliant to respective privacy legislations. Various governments have posed laws, which make them accessible to data stored in their jurisdiction for electronic discovery or anti-terrorism purposes. To gain access to data stored in the cloud i.e., search for data in cloud government should in most cases issue a search warrant and this can differ from service provider to a service provider. In some cases CSP fails to provide required computing resources, in such situations CSP may be forced to outsource data for a different CSP. The
subcontracting CSP’s don’t inquire about compliance with privacy regulations when establishing a relationship. The organizations are not aware of these privacy regulations or they think these are not important to comply with. Raising awareness about both issues and existing regulations seems like a good step to solve this issue [132].

The author of [166] suggests that privacy of user data and laws regulation and policies need to be framed taking CC into context, which will prevent security and user involvement in using the CC application. Methods like homomorphic encryption technique, which helps to protect data will complicate handling huge amounts of data and in an environment where data keeps growing is tough [23]. Providing individual user control, anonymous services for individual use, limiting identity information, and requiring authentication for high level transactions are some of the features that safeguard privacy of users. For this SPs have to encrypt the user information, isolated data processing and storage, managing privacy and security requirements are some of the issues to be dealt while working in cloud computing [122]. Also transaction histories, identity information, policy components during integration, etc., need to be included and protected [146].

Table 4.8: Challenges faced by cloud computing in relation to confidentiality and privacy

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidentiality &amp; privacy</td>
<td>Confidentiality</td>
<td>C1</td>
</tr>
<tr>
<td></td>
<td>Privacy</td>
<td>C2</td>
</tr>
</tbody>
</table>

4.2.7 Backup and Recovery issues

Backup (B1)

CC servers are place where users store all the sensitive enterprise data and regular backup of the user data needs to be done as a fault tolerant mechanism and recover case of disasters where original data is destroyed [144][166]. But the author of [105] is concerned what will happen to the data backup if the company switches? Or company goes down? He also mentions relying on CSPs backup could be foolish. There is also another concern from customer point of view, which mentions that will data stored in the cloud will still be valid even though the cloud provider go broke? Will the data stay intact, accessible, without any logistical problem even when there are merged and acquisitions made by the service provider (long-term viability) [85][142][123][109][55]. The main aspect in all this discussion is to verify whether the client data has high probability in server side. Malicious vendors try to make it fake and collect the data from the server. For example: the server claims that it is storing five copies of data but actually it is storing three copies of data and shows only 5 copies of data occupation [42].
Chapter 4. Literature Review

1. Insecure storage
2. Insecure organization

For this there are assumptions such as:

1. Trusted Platform Module (TPM) is installed for each data backup.
2. Private key is certified by the third party.
3. Assume that the server cannot launch any sophisticated hardware to the stored data.

Article [116] exemplifies a situation of issue when a backup is not properly managed.

Data retention and recovery (B2)

Disaster recovery is another important issue [55]. To recover data service provider needs to have business continuity and disaster recovery planning policies [142]. Even if the customer do not know where his/her data is, cloud provider should be able to tell what will happen to it in the event of a disaster and how long will it take to recover? [85][123]. Industry pundits warn that if any offering is made which does not replicate the data and application of infrastructure across multiple sites is ‘vulnerable to total failure’. Data replication policies should be established along with the proof that the vendor can enact a complete restoration and indicate them how long will it take [22][72]. The author from [164] mentions that disaster and recovery are paid more attention in PaaS.

Investigative support: Investigating data in certain cases is necessary, and data stored in cloud provides some complexities [55]. Author in [71] mentions that digital forensic investigation of information, which include seizing system for investigation in cloud is complicated. In a report on security issues of cloud computing, Gartner pointed out that investigating inappropriate or illegal activity in the cloud is impossible because data of multiple users could be co-located or could be spread out in an ever changing set of hosts and data centers, which was also agreed in the articles [86][3][142]. It also says that the only way to safeguard your data on cloud is to ask the provider for any previous evidences of any such investigations successfully performed, which means there is no investigation possible when cloud services are used [3].

Risk management: Risk management is a process of identifying and assessing risks and plan accordingly to mitigate or the reduce impact of risk. In CC services some components, subsystems or complete system could be distributed and
may not be under the control of the organization using them. Most likely organizations have a better risk management when there is control over process and equipment. With traditional information systems, risks are managed through systems lifecycle and in case of CC services, assessing and managing becomes challenging. Since the organization does not get full control, organization should ensure that security controls are implemented correctly and they are operated as expected. Organization’s trust towards cloud service varies based on the extent of control provided to the organization for its data, applications and also on evidence provided about the effectiveness of those controls. Performing all these tasks to estimate the functionality of service provider is difficult, so solutions to this is using third party audits and establish trusts based result. Finally if the level of trust falls below the level of expectation and no compensatory controls can be employed then the organization has to reject the service or accept with a greater risk.

There is utmost need to have control over physical host security and also virtual machine. If the physical security is compromised all virtual machines residing on that specific host get compromised [78].

Table 4.9: Data backup and recovery related concerns cloud computing

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention or backup</td>
<td>Backup</td>
<td>B1</td>
</tr>
<tr>
<td></td>
<td>Data retention or Recovery</td>
<td>B2</td>
</tr>
</tbody>
</table>

4.2.8 Other concerns

Auditability (A)

Information about auditability, audit and monitoring, audit and compliance, auditability and data, confidentiality, compliance and auditability from different articles are considered.

Difficulty to audit is another side effect of cloud computing which shows, as there is lack of control in cloud [51]. Current auditing scheme in CC goes on with the help of documentation or manual audits. Audit is an internal, external entity that organization implements to identify requirements that organization must abide with and helps to put into practice those policies, procedures and process which are needed to satisfy such requirements. Audit is also used to constantly keep checking if policies and principles are followed within organization or not. For organizations to use cloud computing platforms, cloud service providers (CSP) has to maintain, monitor and demonstrate on going compliance with customers business and regulatory requirements. Sustainability is not possible while maintaining separate efforts on different regulations or standards. Combination of internal policy, regulatory compliance and external auditing should be used
When data is stored in remote locations auditing can become challenging or cost effective when doing it on site [7][126]. Also managing auditing is given as one of the 10 security management areas of CC in [91].

Data confidentiality/auditability: Security is the most common area for critics by CC analysts, who keep asking “who would trust their essential data to be stored somewhere?”. Many cloud computing security threats are similar to those faced by data centers. This responsibility is divided among many parties such as vendor, cloud user and third party vendors that CC users depend in securing sensitive information. In cloud computing architecture user is responsible for his securing at application level; cloud service provider is responsible at physical level and enforcing firewall security. These responsibilities can also be outsourced and given to third party service providers who sell especially security services. Additional features such as firewall rule analysis can be provided by standardized interfaces of platforms such as Amazon EC2. With cloud computing, internal security threats are more when compared to external threats.

Virtualization is key ingredient in cloud computing, with many benefits and at the same time it also brings numerous threats. Incorrect virtualization code might allow user to access sensitive portion of information of other user or provide access to service provider’s infrastructure. This all happens because virtualization software contains some bugs, which might allow virtualized code go loose to some extent. The service provider by default controls the software stack bottom layer which effectively circumvents most of the known security techniques.

A common problem that exists not only in case of cloud computing is, data lost into public and reason for this could be disposing a hard disk without being wiped or a bug within the program that makes data visible to unauthorized users partially.

Similarly, auditability could be added as an additional layer, which is kept out of reach for virtualized guest os, this provides arguably more security than those built into application. Such new features reinforce the cloud computing perspective of changing focus form specific hardware to virtualized capabilities [13][12]. In addition to all these mutual auditing should also be supported to cross check stakeholders in CC [110].

It might be easy, or too easy to start using CC services but hard to govern cloud related activity.

Compliance (COM)

All the compliance related terms (such as compliance, regulatory compliance, audit and compliance) are identified from different articles are considered.

Countries have their own security, privacy and regulatory laws at different levels (i.e., national, state and local), which makes compliance a complicated issue for cloud computing. Compliance requires conformance with the local established specifications, standards, regulations, or laws, which is hard for CC to
demonstrate [135][78][92]. The customer must check if provider allows to timely audits, since customer is the one who is ultimately going to be responsible of data even when stored at SP [85]. The provider must allow customer to check if these through third party audits and also use preventive measures need to be employed for avoiding such violations [65][142][55][123]. Also the compliance can help the providers to restrict the customer from violating regulations that are agreed upon [109]. The author in [50] suggests to provide approaches for license management which ensure a compliant deployment of cloud resources.

There are various issues discussed related, such as:

1. Data location [78] (as discussed in subsection 4.2.1).

2. Are service providers for cloud computing willing to support external audits and security certifications, similar to those of the traditional service providers [3].

3. Payment Card Industry Data Security Standard (PCI DSS), is a set of comprehensive requirements for enhancing payment account data security and this was developed by PCI Security Standards Council to adopt consistent data security measures on a global basis. This PCI DSS is a multifaceted security standard, which includes requirements for management of security, policies, procedures, network architecture, software design and other critical protective measures [3]. This is designed to protect customers account data.

4. Traditional service providers are allowed to external audits and security certifications. So, providers who refuse this are suggested to be used for only trivial functions [178][22].

5. Virtualization and cloud computing have many issues that PCI Quality Assessors (PCI QA), have concerns about and it is necessary to show compliance to these rules, while deploying virtualization technology in PCI environment. Some of the important points to be noted are [3][28]:

   a. Segregation of systems with one primary function per server and network with isolation of all management and control networks

   b. Virtual media that contains cardholder data needs to be protected.

   c. Support auditing on system components.

   d. PCI DSS may require additional processes or technology to ensure patching and change control compliance in virtual environment.

   e. Intrusion protection.
Chapter 4. Literature Review

Guest operating system (OC1)

Generally an each customer in cloud is given a virtual private server (VPS), which can run any operating system requested by user in a physical machine. This means multiple operating systems can be running in a physical machine and for a hacker find and trespass through any one of these operating systems running on VPS is not hard. Its the user responsibility to keep constantly patching up their own VM’s [105].

While interacting with CC services in some cases (where user is given complete virtual machine such as in IaaS) virtual instances are completely under control of customer. Amazon Web Services (AWS) does not give access to instances of customers and so cannot log into the guest OS [146].

Host operating system (OC2)

Administration hosts (build specifically to protect the management plane cloud) built to administer business needs in organization should be using multi-factor authentication before giving access. When an employee quits from job or has no longer access to the management plane, then his privileges should be revoked [46]. This is because if a malicious user gets access to the host he/she can effect the entire guest OS running in it which is a huge risk [105].

Instance isolation(OC3)

Instance isolation, Software isolation are different terms selected from primary selection and discussed here.

When user access a CC service there is possibility for another user to be accessing same part of cloud or other part of it by creating its instance. Isolation is to ensure that different instances running on same physical machine remain isolated from each other. Since the administration of instances here is through instances but not direct, it increases the risk and possibility of threat to the security of CC users. Hence, there needs to be efficient system control and access control restriction and a strict monitoring is needed to track changes. This isolation is efficient in Xen hypervisor and is being used by Amazon [46].

Cost (OC4)

Cloud computing applications can be easily used, but securing these applications requires management to spend additional resources (in terms of cash) [65]. Article [44] discusses multiple challenges and inter-relates their success of implementation to amount of cash an organization can spend.
Table 4.10: Other concerns that cater security of cloud computing

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditability</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Compliance</td>
<td>Com</td>
<td></td>
</tr>
<tr>
<td>Guest os</td>
<td>OC1</td>
<td></td>
</tr>
<tr>
<td>Host operating system</td>
<td>OC2</td>
<td></td>
</tr>
<tr>
<td>Instance isolation</td>
<td>OC3</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>OC4</td>
<td></td>
</tr>
</tbody>
</table>

4.3 Identified solutions

These identified solutions are suggestions from different authors and could vary from person to person or organization to organization. These solutions can be used individually or in combination to give better solution for identified challenges.

Tables mentioned in section 4.4 (i.e., Tables 4.11-4.18), show the solutions that are suitable for the challenges respectively (identified in section 4.2).

(S1) Encryption: Encryption is suggested as a solution to secure information which is being transferred, stored (at rest) or under any other operation. This section maps to the challenges identified in the previous section and explains how these solutions can minimize the effect of challenges while using cloud computing.

- Proposed solution for D1 (mentioned in 4.2.1) As data is at rest before uploading into a cloud server to protect data from unauthorized parties, data is suggested to be encrypted [144][146][132][155]. If data is encrypted in IaaS, data-at-rest is possible and strongly suggested for S3 [104], where as IaaS and SaaS cloud base application use Google apps, or salesforce.com etc., which is feasible. Generally encryption can’t be used for cloud based application because it prevents indexing or data searching [104][78][44].

- If someone is sharing computing resources with other companies in a public cloud, government may seize data with reasonable cause and might result in exposure of data. To avoid this and protect information stored in public cloud data encryption is done, this avoids even cloud providers from having access to data or decryption keys. If the government or someone wishes to access data, a person needs to get down to the user. This helps in maintain user’s data in private yet in the same level for data access to cloud [46].
• To avoid access of data from other users Gatner states “applying encryption on data that makes data totally unusable and normal encryption can complicate availability” [22][46][86].

• To avoid this problem (explained in section 4.2.2), it is necessary to maintain multiple vendors and have Virtual Security Gateway empower. This is considered as the best proposed solution till date, which will help organization get their own VPN network with full encryption between all virtual resources around the world. By doing this they can communicate and transfer the data by using encryption and security [3].

• Data in SaaS vendors such as Amazon S3 is not encrypted by default (to prevent problems explained in section 4.2.7, there is a requirement to have the data encrypted before backup so that unauthorized parties don’t get to access to such (backup) data [144].

• User-level encryption: Proposed as a solution to challenge specified in 4.2.8, in this problem to prevent data from being visible to unauthorized user, most commonly used solution is to have a user-level encryption, which is effective in cloud and for providing user-level encryption there are tools and expertise already available [13][12].

• Encryption while transmission, the data should not only be transferred to right person but should be transferred in from an authentic party with total integrity maintained [92].

• Encrypting before uploading sensitive information into cloud storages. Digitally signing and using RSA algorithm to encrypt is suggested [31].

• Cryptographic methods can prevent from losing data and protect data integrity while transmission [100]. Symmetric and asymmetric cryptography [75].

(S2) **Access rights:** Data owner should give permission to a particular party so that they can access the data easily. To provide this data access control, a standard based heterogeneous data centric security is used to give data protection to application for preventing problems specified in sections 4.2.1 [146].

(S3) **Use of central global transaction manager:** Usage of central global transaction manager refers to 2-phase commit protocol as per XA standards. As a mix of on-premise and SaaS application can arise data integrity problems in the world of SOA and cloud computing (as mentioned in section 4.2.1. SaaS application’s functionality usually gets exposed to XML based API’s. In SOA, many on-premise applications expose functionality based on SOAP and REST web services protocols. There are standards avail-
able for data integrity in HTTP but are yet immature (WS-transaction and WS-reliability) [144][123].

(S4) **Ensuring boundaries:** SaaS ensure that there must be clear boundary for individual user data. This boundary must be ensured not only on physical level, but also at application level to segregate the data from different users (solution to data segregation and data access mentioned in section 4.2.1) [144].

(S5) **Following Electronic Communication Privacy Act (ECPA):** Protection from government accessing electronic mails and other records done by enforcing Electronic Communication privacy Act of 1986 (ECPA). When this ECPA is applied in cloud computing activities, privacy or protection is still unpredictable [144]. This is a suggestion for challenge on data confidentiality issue 4.2.1.

(S6) **Payment Card Industry-Data Security Standards (PCIDSS):** SaaS states it can provide better security (to customer’s data) when compared to conventional methods. But even in such secure conditions the possibility for an insider to still have an effect on customer’s data in a different way. Insiders wont have direct access to database but still risk to security is not reduced. To avoid this, SaaS vendor must be compliant with Payment Card Industry-Data Security Standards (PCIDSS) where host merchants must comply with PCIDSS. Which can be given as a solution to 4.2.1 [144].

(S7) **Encryption using HMAC-SHA1:** Before accessing data, first step is authenticated using HMAC-SHA1 signature request with help of private key [46][152][22]. By doing this the user can have full control over the data to access which is a solution to problems in 4.2.1 [46].

(S8) **Ship disks:** For the problems mentioned in section (4.2.1, data transfer bottlenecks), Jim Gray found the cheapest way is to send using low cost ship disk. This will handle the case of large delay-tolerant point-to-point transfer [13][12].

Example: if we want to transfer data of 10TB from U.C Berkely to Amazon in Seattle, W.A. Garfinkel measured the bandwidth from S3 to three sites where he found an average of 5 Mbits/sec to 18 Mbits/sec. for WAN link 20 Mbits/sec then

\[
10TB = \frac{10 \times 1012 \text{Bytes}}{20 \times 10^6 \text{bits/second}} = \frac{8 \times 1013 \text{seconds}}{2 \times 10^7 \text{seconds}} = 4,000,000 \text{sec - approximately 45 days}
\]

Where ship disks can transfer 10TB data in one day that means like 1500 MBits/sec. AWS recently offering such a service called import/export [13][12]. This could be used as a solution to data transfer bottlenecks.
(S9) **Decommissioning and SSL encryption while transfer:** AWS Procedure decommissioning when data reaches useful life as stated in 4.2.1 AWS uses a decommissioning process ensures that data is not exposed to unauthorized individuals.

Usage of SSL encrypted endpoints can provide more security to Amazon S3. Using this method lets users access cloud features from Internet and also from with in the Amazon EC2. There are certain test used to validate level of security in SaaS, if there exist any vulnerabilities found out they need to be given additional security 4.2.2 [144].

(S10) **Usage of AWS and APL:** To reduce possibility of denial of service (explained in 4.2.2) amazon AWS and APL are hosted large endpoints in internet scale, world which can possibly reduce the chance of service downtime [46].

For problem explained in 4.2.2, Amazon uses SSL protected end points to authenticate AWS API’S and also generates a SSH host certificate, logs them in instance’s console. This enables customers to use secure API’S before logging into instance for first time and to make transactions more secured customers are forced to use SSL for interaction with AWS [46].

(S11) **Acceptable Use Policy (AUP):** Solves the problem of port scanning mentioned in 4.2.2, it is required to maintain a policy; every action of customer need to be monitored and any violations by customer should be reported to the service provider. Amazon EC2 maintains a ‘Amazon EC2 Acceptable Use Policy (AUP)’ which checks for any policy violators and warns service provider in early stages [46].

(S12) **Re-evaluating models and using enhanced security mechanisms:**

To get viable solution for set of issues mentioned in section 4.2.4 and for utilizing potential benefits of cloud, existing life cycle models, management processes, service attestation, penetration testing and risk analysis must be re-evaluated. Security mechanisms must be upgraded to use Security Assertion Markup Language (SAML), Extensible Access Control Markup Language (XACML), and Web services standards [6]. For policy engineering, there are several approaches proposed in cloud adoption, which may cause disruptions to the organization [175][146]. To avoid this disruption set of roles should be proposed for existing and optimal sets. Another possible solution is StateMiner approach that helps to get heuristic based solutions on RBAC state with least structural complexity [146][145][6].

(S13) **Usage of standards:** This is a solution/suggestion given to identity and access management system (described in section 4.2.4) and using this system (cloud computing system, maintaining identity and access management)
cloud service providers are suggested to adopt standards such as SPML, SAML [78], open id standard [78], OAuth, and XACML to have a secure interaction between entities (users of a system) in different cloud computing applications [6]. Entities in cloud platform could also be identified and granted access rights based upon delivery models (SaaS, PaaS, IaaS, etc.), and type of cloud (hybrid, public, private etc.), to ensure better security [123]. A comparison between different ways of identity management and standards for identity management are discussed in [66]. The authors for [85] say SAML is better than OAuth.

(S14) **Disaster recovery and Business continuity:** For safeguarding data from unplanned emergencies or minimizing enterprise downtime (the effects of these are mentioned in 4.2.4), appropriate plan for Business Continuity (BC) and Disaster Recovery (DR) is suggested [144]. For the problem mentioned under failure in providing security (4.2.4), is to have a system which restores data from backups silently, without any interruption and this system called failover [127].

(S15) **Maintaining multiple stacks:** is a solution for maintaining availability (challenge description in 4.2.4), is to maintain multiple service providers providing different stacks, such that if one stack fails there will be another stack (i.e., service provider) supporting as a substitute [13][12].

(S16) **Trusted Third Party:** Another solution to maintain availability (challenge discussed in 4.2.4), is by securing the system by using Trusted Third Party (TTP) [179], when a trusted third party is used it provides the strength that all organizations involved in its activities are trustworthy [94]. To ensure policy management and trust management (for problems mentioned in section 4.2.4) is to have a Trusted Third Party (TTP) within the cloud to ensure confidentiality integrity and communication [179].

(S17) **Keeping track of users:** Managing multiple SaaS products may also increase IT management overhead and might also complicate the user management system and in section 4.2.4 explains a scope of misusing user information if not managed properly and to avoid this issue, customers should keep a track of persons leaving or joining the organization, correspondingly add or remove their information [144].

(S18) **Privacy preserving protocols:** Solution to the problem of authentication and identity management (mentioned in section 4.2.4), is to have privacy preserving protocols, which verifies various attributes. An example of such technique is zero-knowledge, proof-based technique, these techniques use pseudonyms and accommodate multiple identities to protect privacy of
users and can further help built a user centric federated IDM for clouds [146].

(S19) **Dynamical scaling:** As a suggestion to the problem mentioned in section 4.2.5, open cloud manifesto states that stay away from problems with allocation of resources during peak hours cloud should be dynamically scalable (scale up and down) [61].

(S20) **Virtual Security Gateway (VSG):** Virtual Security Gateway (VSG) which is proposed as a solution to the challenge mentioned in section 4.2.3, is a firewall that runs inside virtual infrastructure and ensures all communications inside virtual environment stick to defined security policies. In this infrastructure no two virtual machines are allowed to talk to each other without passing through the security gateway. This gateway uses VPN encryption to ensure secure interconnection between VM’s, since VSG is placed inside virtual infrastructure auditing can also be achieved, regulatory compliance requirements can also be satisfied. As the cloud computing platform is scalable user can extend his/her services easily, which means this VSG can be easily deployed. VSG also allows creating lab/test environments and thereby improving overall security [3].

(S21) **Alter this to software:** This is given as a solution to a problem mentioned in 4.2.3, this can be done by maintaining a hypervisor level, MAC and trusted computing techniques are suggested mechanisms used to build future secure cloud systems [73].

(S22) **Including auditability as a layer:** As a solution to the problem explained in auditability (section 4.2.8), this could be added as an additional layer that stays out of virtualized guest operating system’s reach, which provides arguably more security than those built into the application. Such new features reinforce the cloud computing perspective of changing focus from specific hardware to virtualized capabilities [13][12].

There are some well known auditing guidelines such as in SAS 70, SOX and HIPAA. US government follows FISMA, NIST and FIPS [35].

(S23) **Multi factor authentication:** The solution recommends for challenge in this section 4.2.8, is to remove traditional password based encryption and use some multi-factor (two ways of checking before providing access) to provide access at least use a certificate-based SSH Version 2 authentication to access their instances [147][145][21][146]. The author explains with an example such as if someone using Linux as his/her guest OS, hardens the instance they should use certificate based SSHv2 to access the virtual instance, disable root login, use command line logging and ‘sudo’ command to change privilege [146]. General methods of access control are use of roles
and password, but using a multi-factor authentication is more secure than single password protection [105].

(S24) **Permission revoke:** Problem mentioned in 4.2.8 gets solved if permissions are revoked as soon as an employee quits from his job or no longer has access to management (which he had before) [146].

(S25) **Xen hypervisor:** This solution is suggested for a challenge mentioned in 4.2.8, which mentions that to effectively isolate instances there is needs for an efficient system control, access control restriction and a strict monitoring that tracks changes within virtual machines. It is also stated that isolation is efficient in Xen hypervisor, which is currently being used by Amazon [146].

(S26) **Calculate hash:** This is a solution proposed for backup related issues (section 4.2.7). Before uploading a file, the client is suggested to verify whether the file is stored on backup drives or not during the intermediate transaction and whether keywords in file remain unchanged. Calculate the hash of the file before uploading, which also divides file into equal size randomly and stores it in the locality [42].

A general solution in providing security to the cloud computing system is by having a service which acts as a watchman and secures all the applications designed in it [88].

For storing data on the cloud there might be some issues, as mentioned in 4.2.1. Storing hash in trusted local memory is a method to avoid issues, it helps user to check if the data stored in cloud stays unaltered. If there are huge sets of data is being stored, individual can maintain his own hash table/hash tree. This mechanism is used in many storage prototypes such as TDB [26][102], SiRiUS [26][59] and is commercially being used in Solaris ZFS filesystem1 [26].

This can also be used for data integrity but maintaing hash can become complicated [81].

(S27) **Credential or attribute based policies:** Using credential or attribute-based policies is a suggestion to identify unauthorized users [146].

(S28) **Standardize API’s:** Suggestion to the problem with data lock-in (4.2.1) is to standardize API’s, such that customers can deploy services and data across multiple cloud computing providers. The advantage of doing this even if one service provider fails it would not corrupt all other copies stored on other places (data or services). If this format is introduced, it could create a new usage model where software structures can be used in an internal data center and in public cloud [13][12].
Chapter 4. Literature Review

(S29) **Isolate personal information and software:** For problems mentioned in privacy (i.e., section 4.2.6) author in [29] proposes that to maintain privacy of an individual’s data a solution could be to isolate users private data in databases and distribute the software around the internet. Author in [65] suggests using software isolation where each domain in the cloud environment has a trust agent.

(S30) **Middleware for configuring storage infrastructure:** Cloud computing providers often use Storage Area Networks (SAN) to gain storage scalability, which have a vast parameter range and hence makes manage cloud storage, components to configure. To solve this problem a solution was proposed which introduces a SAN configuration middleware that manages task of updating and debugging heterogeneous SAN developments [47].

(S31) **SaaS protection:** Homomorphic token with distributed verification of erasure-coded data towards ensuring data storage security which supports dynamic update, delete and append data without loss or corruption. This method is efficient over byzantine failures and server colluding attacks, data modification [65].

(S32) **Border Gateway Protocol (BGP):** To identify malicious systems that announce themselves as the righteous destination for all the data transferred in a network. This architecture is vulnerable to DDoS attacks [65].

(S33) **Quantitative risk assessment framework (QUIRC):** Author in [136] mentions that this framework can be used to assess the level of security in the enterprise. The usage of this framework is suggested by [17]. In addition to this wide-band delphi model is also suggested [17].

(S34) **Distributed access control:** Author [8] suggests distributed access control architecture for access management in CC.

(S35) **Intrusion severity analysis:** Author [14] identified intrusion severity analysis in CC is a problem and for that he proposed a new method to detect intrusion, which is based on machine learning.

(S36) **Performance oriented architecture:** This is a reactive but not proactive architecture, which analyses the CC intruders before they can cause any damage [113].

(S37) **Trusted virtual data center:** These data centers have different VMs and associated hardware resources which will help to identify, which VM will access which resource. This can also separate customer workload to different associated virtual machines [74].
Chapter 4. Literature Review

(S38) **Attack detection model:** Article [90] proposes a proactive attack detection model which can detect an attack (at time of attack or before), alert the system/security administrator and helps the customer to understand the kind of attack that happened by looking at pattern of attack. Solution to attacks from fraud pissing or denial of service etc., is to keep login/account credentials personal and never disclose them.

(S39) **Access control based on group signatures:** By this method user can access CC using a group account, instead of personal account. This will allow the customer to preserve his/her privacy [176].

(S40) **Proactive approach for CC security:** A proactive approach to prevent customer falling into wrong hands is being proposed. This model proposes a security cloud which can perform tasks such as: Monitoring public blacklists for ones own network blocks, Vulnerability Assessment, Penetration Testing, Log Analysis, Host Based Intrusion Prevention System [142].

(S41) **Efficient remote data possession checking (RDPC):** The author for article [33] mentions that this scheme is better in terms of communication and computation, verification without need to be compared with the original data and mentions that user needs to store only two secret keys and several random numbers.

(S42) **Non-repudiation protocol:** A nonrepudiation protocol is proposed. This protocol is a two-party non-repudiation protocol that is like TCP/IP 3 phase handshaking protocol. This protocol used for both integrity checks and non-repudiation [51].

(S43) **Two stage API access control:** The author of [139] proposes a two stage API access controlled mechanism that developed based on a Rule Based Access Control mechanism (RBAC).

(S44) **CISCO secure data framework:** This framework provides multiple security layers with different existing security technologies that can be used to enhance security [169].

(S45) **Security measuring framework:** This framework is suitable for SaaS and can be used to determine the status of user’s applications that have been running in the VM for some time [75].

(S46) **Community watch service:** This service is supposed monitor, analyze cloud users and detect newly injected malware attacks [75].

(S47) **Network based intrusion prevention system (NIPS):** The author of [75] gives this a multi-technology based approach. This NIPS has 4 different technologies clubbed together such as, hardware acceleration, active defense
technology and linkage with a firewall and a synthesis detection method that are used in combination to detect block visits when threats are detected in real-time.

(S48) **Data fragmentation**: This solution is suggested to prevent intrusion into user’s privacy. The suggestions are to use a cloud based malware scanners and use data fragmentation technology which can be used to hide the connection between data attributes and/or separate customer’s data from the software [75].

(S49) **Virtual network framework**: Using a virtual network framework which consists of routing, firewall, shared network and ability to control intercommunication of VMs deployed, can provide secure access at VMs level [75].

(S50) **SaaS Role Based Access Control (S-RBAC) and Intercloud identity management infrastructure**: Are two methods that the author of [75] mention which can avoid the conflicts with access control while working with a multi-tenant environments.

(S51) **Data protection as a service**: This service model is proposed to ensure privacy and data security, also offers evidence of privacy to owners [140].

(S52) **Confidentiality as a service (CaaS)**: Proposes to protect the visibility of even clear text data from unauthorized users by having multiple cumulative layers of encryption [48].

(S53) **Proposes a data security model**: This model consists of 3 layers which perform their own task. First layer OTP authentication, second (data encryption, data integrity, private user protection), third fast data recovery. Test this model and implement software to improve work of data security model in CC [15][164].

(S54) **Third party auditing scheme**: This is to protect the data storage security and proposal here is to move the auditing process into cloud computing and integrate with the architecture of cloud provider [68][32][81]. Public and private auditing [153].

(S55) **Remote data position with public verifiability**: Proposes Remote data possession checking protocol with public verifiability. This uses HLAs and RSA construction [98].

(S56) **Fine grained access control**: Proposes a fine grained access control mechanism for data [95].
Chapter 4. Literature Review

(S57) **Fog computing:** This method uses behavior profiling of user and a decoy which can prevent malicious users from damaging the CC environment [143].

(S58) **RSA based Storage security (RSASS):** This method is supposed to address remote data security, which is based on RSA for storing files in remote servers. This method can also compute large files with different sizes [153].

(S59) **Chinese wall policy:** This method is used to address the insecure information flow at by using information flow policies at IaaS layer [163].

(S60) **Privacy aware inter-cloud:** This method takes charge of data protection standards, processing the privacy requirements and user requirements [149].

(S61) **Multi cloud model:** An OPNET tool is suggested here which has separate security as a service and tested simulation of all traffic [5].

(S62) **sTiles:** This is to preserve the privacy from both inside computational nodes and intruders [23].

(S63) **Trust evaluation model:** This ETCT architecture proposed includes time variant and space variant trust evaluation methods for calculating direct trust and recommends trust [64].

(S64) **Multi-replica PDP model:** This model is proposed to preserve integrity and data availability in the cloud. The authors of [99] proposed an improvement to this model, which can support third party auditing also.

(S65) **Mutual protection for cloud computing architecture (MPCC):** Is designed on a reverse access control concept which gives the user the control for authorization and authentication; also allows to check if any CC security violation are present [7].

(S66) **Network protection methods:** Using SSL/TLS, IPsec, Traffic cleaning and network based intrusion detection [32].

(S67) **Honeypot:** Used to detect and counter attack or deflect unauthorized access attempts [32]. This can be loaded into the user’s logical instance running on SaaS, PaaS or IaaS and help them to effectively handle viral infections and service hijacking in the cloud.

(S68) **Preserving cloud computing privacy:** In [122] authors of this paper propose a PccP (preserving cloud computing privacy), which prevents unauthorized users from accessing user information or derive such information from IP addresses.
(S69) **Permission as a service (PaaS):** This solution is proposed in par with other solutions, which allow users to tell what part of data can be accessed in which way, but the problem here is when there are large chunks of data to be managed. So to provide a solution the authors designed a new service which separates access control panel and allows user to set access to all the data from a single location. This PaaS uses Attribute Based Encryption (ABE) in their framework [45].

(S70) **Tree based key management:** Existing tree based encryption methods have some weaknesses mentioned in [177]. So to avoid this weakness and propose a practical solution for private data management, the authors develop a new tree based key management system and prove that the above problem can be solved [177].

(S71) **Byzantine fault-tolerant protocol:** This byzantine fault-tolerant replication protocol is used to manage collaborative data storage. Zetta which is similar to RAID systems can also be used as a solution [100].

(S72) **Denial of service prevention methods:** Network access control (NAC), Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) [100]. Intrusion detection system (IDS) or firewall [67].

(S73) **Computational Intelligence (CI):** CI is a mathematical model for CC, and authors of [57] propose an improvement that can enhance the performance and automated service provisioning in data security.

(S74) **Ubiquity and integration of services (UBIS):** The new architecture, UBIS, discussed in article [67] is proposed to satisfy cloud user requirements and cloud security challenges. This architecture is based on QoS management and includes security aspect based on an event driven architecture and ubiquitous services in the cloud.

(S75) **Risk perception model:** Author in [49] proposes a model useful in defining risk perception before moving on to CC. This model was redesigned to reduce the 9 variables (such as, voluntaries, immediacy of effect, knowledge about risk, knowledge of science, etc. that were used by Baruch Fischoff [53] to determine when a product is safe to be accepted) into two areas (understanding and consequences). The perceived risk in this model is a function of these two terms understanding and consequences.

(S76) **Proxy based firewall/NAT traversal solution (PASS):** This solution is proposed for SaaS integration, which allows on-premise applications to work securely on CC [97]. Other solutions are VPN, which is difficult in SaaS but easy on IaaS, Microsofts AppFabric (specific to Microsoft platforms) and TARGET.
(S77) **RSA based Assumption data integrity check:** This method is proposed by [81] since other methods such as hash, 3rd party auditing etc., he found are complicated or insecure. This new method combines both identity based cryptography and RSA signature [81].

(S78) **Capability based approach:** Use of Access Control List (ACL’s) or file groups lack scalability, fine grain access control in cloud computing. The scalable method (proposed in article [170]) that is based on using the combination of ABE, lazy re-encryption and proxy re-encryption is supposed to become complicated in an extremely dynamic environment such as CC. The re-encryption scheme and framework for access control on published XML documents by using different cryptographic keys on XML documents are also not suitable. Since XML scheme becomes complex to manage keys and XML documents and in case of re-encryption, if one user is malicious the data is exposed. So the author for [133] proposes a scalable, efficient and secure data access control mechanism using Capability based access control.

(S79) **Improved Capability based approach:** The author suggests to use a capability based approach which is already used in distributed file systems and suggest some extensions to them. The extensions constitute integration with external identity or access management components used by enterprises, capability based access control architecture for data centers, user to user and user to application delegation and fine-grained access, dynamic access right’s scope (a survey on the existing methods by the authors show that the features they proclaim to have in the proposed architecture are missing within the existing ones) [70].

(S80) **Multi tenancy for databases:** Proposes a multi tenant database design for ad-hoc clouds. The authors boast that their method allows multiple organizations to collaborate and benefit without security leaks [118].

(S81) **Secure ranked keyword search:** Searching through encrypted data in CC using traditional systems could accidentally reveal data, this proposed system avoids leaking others data [156]. This proposed method is experimentally tested.

(S82) **Security framework:** The author in [148] proposes a comprehensive security framework for CC, which has separate modules for trust issues of key CC components and also a module to handle security.

(S83) **Automatic virtual infrastructure evaluation:** The author proposes an approach to automatically check if the virtual infrastructure is following the security requirements imposed by the user [56].
Multi-tenant authorization model: This model is suitable for middleware services in PaaS layer. This model is said to support RBCA, Multi-tenancy, path based object hierarchies, hRBCA, and federation. Some other methods of authorization are also described in related work. But this model is based on abstraction of path concept and the authors mentions that it is better suited to CC authorization [27].

Data-driven framework: The author for [178] proposes a data-driven framework which provides secure data processing and sharing between cloud users. This framework is designed in such a way that secure query processing exists in multi-user scenario, integration of declarative access control with data processing, efficient end to end data verification and system analysis and forensics by capturing accurate historical records exists. This model is ready to use and is also available for download.

Clout Trace Back (CTB): This model is being proposed to trace back the DDoS attacks in CC [83]. It helps to find the real attacker among the number of requests that server receives under DDoS attack state [83].

Software watermarking: This paper designs a robust and new cloud based watermarking mechanism that can report and alert the required personal [171].

Tag: Tag entities which should not be moved and prevent CSP from moving sensitive data to another jurisdiction [71].

Fine-grained access with delegation of computational intensive tasks: This model achieves fine grainedness, scalability and data confidentiality for data access control in cloud computing. This proposed scheme enables the owner to delegate most of computation intensive tasks to cloud servers without disclosing data contents or user access privilege information. The proposed scheme is provably secure under the standard security model [170].

Apart from the above mentioned solutions some of the authors have mentioned general guidelines to provide security:

- Article [15], Good and well formulated SLAs are to some extent helpful points for ensuring availability. To avoid risk of insider malicious attacks the user must be informed about the information security and management practices. To prevent outsider attacks there can be a network access control system built up to at OS level. VMs should be isolated for each customer. Stopping to share account credentials and two factor authentication proactive monitoring (detect unauthorized activity) will probably help service disruptions. A strict registration and validation can reduce a lot of damage
from attacks. Defense-in-depth approach at different layers using different protection mechanisms can protect the VM’s better. Being aware about threats and preventing sensitive data into cc can be helpful in case of user and providers should allow the users to port their security protocols so that the user’s data stays protected.

- Article [105], Settlement on open standards can help reduce vendor lock-in problems and incompatibility issues. While transferring data user needs to use a SSH-tunneling or VPN protocol. The user can look for CSP with continuous data protection where every version of file is saved by the CSP in their server. Data portability should be provided to avoid Vendor lock-in. To avoid DoS the simplest way is to check if the source ip-address is invalid and another is to implement a strong and robust firewall. A reverse firewall can help to protect attacker from accessing an external server from a VPS.

- Article [89], The author highlight some issues that can build up a trust in cloud computing platform:

  1) Notify the customer when any entity accesses data,
  2) Without informing the customer the SP won’t save any backups and
  3) When a customer withdraws from Cloud all his data, direct and indirect access to that data will be destroyed.

In addition to this the software provided should be reliable, transparency at service level security properties and where the data is stored (its physical location). Some emerging technologies such as remote access control, reflective property, certification and private enclaves should be maintained so that the trust enhances.

- Using CSA suggested security solutions [137].

- Article [54], Providing tools for security management and auditing, disclosing security policies, compliance and practices can attract people towards cloud computing. Cloud providers need to have strong internal monitoring and auditing scheme.
4.4 Results of literature review

Literature review on Cloud Computing (CC) research area resulted in numerous challenges related to security (43 challenges), which were categorized and explained. Following tables summarize the findings of literature review. The tables 4.11-4.18 show grouping of challenges, their respective code and number of solutions and the number of research papers that discuss the challenges.

As an example, from table 4.11 which shows ‘data related’ challenges their codes, solutions identified and number of references; we can identify that the challenge ‘data security’ which is given the code ‘D1’ (mentioned in ‘data related’ challenges) has ‘(S1), (S2), (S79), (S73), (S53), (S51), (S81)’ as solutions with ‘27’ articles mentioning this challenge in their discussion.

Similarly other tables ‘4.12’, ‘4.13’, ‘4.14’, ‘4.15’, ‘4.16’, ‘4.17’ and ‘4.18’ shown below mentioned about network, virtualization, organization, scalability, confidentiality and privacy, backup and other concerns related challenges respectively. The summary includes corresponding codes (used to mention the challenge), solutions (identified from literature which is a mitigation strategy for the identified challenges) and number of references (number of articles mentioning the challenge).


<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data related</td>
<td>Data Security</td>
<td>D1</td>
<td>(S1), (S2), (S79), (S73), (S53), (S51), (S81)</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Data locality</td>
<td>D2</td>
<td>(S1), (S88)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Data integrity</td>
<td>D3</td>
<td>(S3), (S54), (S1), (S87), (S26), (S77)</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Data segregation</td>
<td>D4</td>
<td>(S1), (S4), (S34)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Data access</td>
<td>D5</td>
<td>(S4), (S27), (S89), (S1), (S78), (S56), (S69), (S79)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Data confidentiality issue</td>
<td>D6</td>
<td>(S5), (S85)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Data breaches</td>
<td>D7</td>
<td>(S6), (S44)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Reliability of data storage</td>
<td>D8</td>
<td>(S41), (S55)</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Data center operations</td>
<td>D9</td>
<td>(S7), (S8), (S71), (S1), (S53)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Data sanitization</td>
<td>D10</td>
<td>(S7), (S8)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Data storage</td>
<td>D11</td>
<td>(S28), (S26), (S54), (S42), (S31), (S58), (S80)</td>
<td>35</td>
</tr>
</tbody>
</table>
### Table 4.12: Network related challenges and solution

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network related</td>
<td>Network security</td>
<td>N1</td>
<td>(S1), (S9), (S35), (S36), (S40), (S46), (S47), (S46)</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Application vulnerabilities</td>
<td>N2</td>
<td>(S13), (S43)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Host and network intrusion</td>
<td>N3</td>
<td>(S36), (S72)</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Denial of service</td>
<td>N4</td>
<td>(S10), (S86), (S72), (S38)</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Men in middle of attack</td>
<td>N5</td>
<td>(S10)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>IP spoofing</td>
<td>N6</td>
<td>(S32)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Port scanning</td>
<td>N7</td>
<td>(S11)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Packet sniffing</td>
<td>N8</td>
<td>No solution</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Sharing computing resources</td>
<td>N9</td>
<td>No solution</td>
<td>11</td>
</tr>
</tbody>
</table>

### Table 4.13: Virtualization related challenges and solutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtualization Related</td>
<td>Vulnerability in virtualization</td>
<td>V1</td>
<td>(S37), (S83)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Virtual machine protection</td>
<td>V2</td>
<td>(S20)</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Networking in virtual networks</td>
<td>V3</td>
<td>(S59), (S49)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>VM isolation</td>
<td>V4</td>
<td>(S25), (S29)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>VM securitization</td>
<td>V5</td>
<td>(S21)</td>
<td>1</td>
</tr>
</tbody>
</table>
### Table 4.14: Organization related challenges and solutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization related</td>
<td>Organizational security management</td>
<td>O1</td>
<td>(S12), (S75), (S12), (S76), (S33), (S45), (S74), (S82)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Trust</td>
<td>O2</td>
<td>(S16), (S63)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Failure in providing security</td>
<td>O3</td>
<td>No solution</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Identity and access management</td>
<td>O4</td>
<td>(S13), (S17), (S18), (S34), (S79), (S57), (S84), (S59), (S50), (S61), (S65), (S67), (S84)</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Availability</td>
<td>O5</td>
<td>(S14), (S15), (S16)</td>
<td>21</td>
</tr>
</tbody>
</table>

### Table 4.15: Scalability related challenges and solutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability related</td>
<td>Incompatibility issue</td>
<td>SC1</td>
<td>No solution</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Constant feature addition</td>
<td>SC2</td>
<td>No solution</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Scalability</td>
<td>SC3</td>
<td>(S19)</td>
<td>4</td>
</tr>
</tbody>
</table>
Table 4.16: Confidentiality and privacy related challenges and solutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidentiality &amp; privacy</td>
<td>Confidentiality</td>
<td>C1</td>
<td>(S52)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Privacy</td>
<td>C2</td>
<td>(S29), (S62), (S48), (S39), (S60), (S68), (S70), (S64), (S81)</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 4.17: Backup related challenges and solutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retention or backup</td>
<td>Backup</td>
<td>B1</td>
<td>(S1), (S26)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Data retention or Recovery</td>
<td>B2</td>
<td>No solution</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 4.18: Other concerns related challenges and solutions

<table>
<thead>
<tr>
<th>Relation</th>
<th>Description</th>
<th>Code</th>
<th>Solutions</th>
<th>No of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other concerns</td>
<td>Auditability</td>
<td>A</td>
<td>(S1), (S22)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Compliance</td>
<td>Com</td>
<td>No solution</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Guest os</td>
<td>OC1</td>
<td>(S1), (S23)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Host operating system</td>
<td>OC2</td>
<td>(S24)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Instance isolation</td>
<td>OC3</td>
<td>(S25)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Cost</td>
<td>OC4</td>
<td>No solution</td>
<td>2</td>
</tr>
</tbody>
</table>

From Tables 4.11-4.18, it can be understood that in most cases identified challenges have at-least a single paper describing issue, but every challenges identified does not have a solution or practice proposed. From tables, it’s clear that there are no solutions found for some challenges. Given below are list of those challenges:

1. In Network related challenge N8 (‘Packet sniffing’), N9 (‘Sharing computing resources’).
2. In Organization related challenge O3 (‘Failure in providing security’).

3. In Scalability related challenge SC1 (‘Incompatibility issue’), SC2 (‘Constant feature addition’).

4. In Backup and recovery B2 (‘Recovery’).

5. In Other concerns COM (‘Compliance’), OC4 (‘Cost’).

Out of these challenges with no solutions, ‘Incompatibility’ under scalability challenges with ‘SC1’ as code is a challenge that has 19 articles mentioning the challenge and does not have any proper mitigation strategy mentioned.
Chapter 5

Surveys: Design and results

The results that are obtained from chapter 4 conclude that there are some challenges with no mitigation strategies proposed. These are ‘packet sniffing’, ‘sharing computing resources’, ‘failure in providing security’, ‘incompatibility’, ‘constant feature addition’, ‘data retention or recovery’, ‘compliance’ and ‘cost’. Out of these (challenges with no mitigation strategies) incompatibility is a challenge mentioned by a number of articles but no mitigation strategies identified. To identify mitigation strategies for the challenges, the surveys are employed.

This chapter describes the steps for designing the surveys, identify the survey respondents and implement the surveys.

5.1 Survey design

Surveys are described as the best method to collect the original data in a population that is too large to observe directly [158]. Surveys generally help the researcher to understand and generalize the results on a population by asking questions to a sample from that sample [36]. Surveys are chosen over other methods because these can be simple and in this case since the objective is just to identify solutions/practices/guidelines for the challenge that has no mitigation strategies, surveys stand a better chance in reaching large population. Three types of surveys are reported [158] and as follow:

- Needs assessment survey: to ask for problems and gather possible solutions.
- Marketing survey: surveys to understand the demand and the nature of specific programs.
- Evaluation surveys: to evaluate the outcome of newly implemented programs and policies.

These surveys can be conducted in different ways, by:

- Conducting surveys through mails i.e., by sending a set of questionnaires to the respondents with a covering letter which describes the purpose of the survey.
• Dialing a random number from a sample selected from a telephone directory and registering their responses to the questions either by taking a note on paper or by entering them into a survey form.

• Face-to-Face surveys are another way of gathering responses by asking the respondent a question. These kinds of surveys are expensive and time taking.

• Drop-off survey where the questionnaire is given in hand to the respondent.

• Electronic surveys are the kind of surveys where electronic media are used to contact the respondents. The respondents are sent a set of questions through email and requested to respond.

The needs assessment type is considered in this work to gather possible solutions/practices/guidelines. The medium used to communicate with the respondents is using e-mail and the responses are collected using an online survey tool.

5.1.1 Steps for conducting an online survey

For conducting an online survey the steps similar to other methods such as planning, data collection, data analysis, reporting and application need to be followed. Each of this general procedure has few specific tasks to be followed [129].

• Identify the evaluator: The evaluators are the master students of their master thesis report under the guidance of the supervisor.

• Identify and engage the stakeholder: The stakeholders are those who have an interest in the program (results of the thesis). In this context the stakeholders are the company professionals who could use the benefit of working with multiple CSP’s provided issue with incompatibility will get solved. These can be solutions architects, senior practice managers, chief architects, people who have rigorously worked with cloud computing application deployment, founders and CEO’s. The survey process used in this study identifies these stakeholders from different companies and ask them to respond to the survey questions.

• Determine resources: The resources is based on the respondent’s time spent on the survey link, which is generated from a free online survey tool. The results can also be analyzed with the help of an online survey tool. This online survey tool generates a report, which summarizes the responses for each question presented in the survey questionnaire.

• Writing goals and objectives: The survey aims to collect expert opinion and experiences on how to mitigate the issues identified under ‘incompatibility’. Since there were no dependable mitigation strategies (which can be used as
Chapter 5. Surveys: Design and results

a mitigation strategy by companies implementing CC) identified from the literature review. The incompatibility issue identified in the preliminary study was broken down into 3 issues for easy interpretation in survey questionnaires. These 3 issues can be stated as objectives of performing surveys, which were clearly stated in the beginning of the survey.

- Evaluation or results: The evaluation of results will be done in such a way that, the results mentioned through this study can help new CC users to identify mitigation strategies for the incompatibility issue. For the researchers this evaluation will help to identify which areas must be concentrated to improve security for CC users.

- Using a software to implement surveys: The survey link is created at (www.quicksurveys.com). The survey link is kept active to gather as many responses as possible that can suggest mitigation strategies. The survey link is also a one time response link which assures that there won’t be multiple responses from the same person (this avoids duplication) and this can improve the quality of surveys.

5.1.2 Designing the survey questionnaire

To see that the results of the study are helpful for the research field, there are 5 questions which were mentioned in [158] and have been taken into consideration while designing questions. The questions are as follows:

- What do we want to know?
- About whom do we want to know?
- How do we word the questions?
- How do we elicit appropriate and adequate responses?
- How do we interpret the results?

The above mentioned 5 questions were kept into consideration and survey questionnaire were designed. These were designed to focus on the challenge that has no mitigation strategies identified from the literature. The questions in the surveys were framed taking suggestions from our supervisor and we constantly simplified (checked and corrected) to avoid grammatical errors and help better understanding for the respondents. The final questionnaire used for surveys are presented in the appendix B. The survey questionnaire created following the above mentioned procedure was open for answers through the link (http://tolu.na/17cFxZz). The respondents were contacted through e-mail and requested to fill out the surveys. The respondents were also sent reminders at regular intervals to make sure that the responses were received without delay.
It was also necessary that the survey responses are strong (number of responses gathered from experienced practitioners) because the sample size considered was random. The survey link was kept open till 15 filled responses (with responses from experienced personnel) were received.

5.2 Observations from survey responses

Given below is the description of responses observed from surveys. There were 27 survey visitors and 15 responses. Among these, there were responses from CEOs, Founders, Chief architect, solution architect, senior technology practice manager, etc. These respondents are Information Technology (IT) professionals having 8, 7, 5, 3 years etc., experience in cloud computing and from different companies. It is also possible that some of these respondents were working with cloud computing from the year it was being used. Given below is the discussion of responses received for each question.

5.2.1 Results: Question related to service providers, service models and cloud types used

There were few basic questions asked to the survey respondent such as, who is your service provider, which cloud service model and deployment model they use. These questions were asked to know which service provider, which service model and deployment model are popular in practice.

Q1: Who is your service provider for cloud computing applications?

The answers to this question are shown graphically in figure 5.1. From this figure its understood that Amazon is the leading service provider with 60% votes, followed by Google with 33.33%.

![Figure 5.1: Results of survey: Service providers used](image-url)
Chapter 5. Surveys: Design and results

Q2: Types of services used
In figure 5.2, it can be understood that SaaS is being predominantly used (with 66.67% votes from respondents), followed by IaaS (with 46.67% votes from respondents) and in the last comes PaaS (with 40% votes from respondents).

Figure 5.2: Results of survey: Type of service used

Q3: Types of cloud used
In figure 5.3 shows the types of cloud are being used. From this it can be understood that public clouds are mainly being used (with 66.67% votes from respondents), followed by private cloud (with 46.67% votes from respondents), the next position is hybrid cloud (with 26.67% votes from respondents) and in the last comes community cloud (with 13.33% votes from respondents).

Figure 5.3: Results of survey: Deployment models used

5.2.2 Results: Questions in relation to cloud computing incompatibility issue
For the question related to ‘incompatibility’ in cloud computing systems that are presented in this questionnaire, the responses were guidelines and precautions.
Some of these suggestions were to use global standards to provide interoperability between service providers and different ways of encryption while interfacing between multiple cloud providers.

It is also observed that there have been maximum responses for the first 3-4 questions and for the last question related to on premise authentication systems there were few responses. For easy understanding each respondent of the survey will be given a notation such as R1, R2, R3 so on till R15 to describe respondent 1, respondent 2, respondent 3 so on till respondent 15 respectively. Given below are the responses of respondents to each question.

First question:

a) What steps can be followed to provide interoperability between two or more service providers?

Responses: Total Respondents to this question are 13. Given below are the solution/practice/guideline mentioned against each respondent.

**R14**
1. Gateway functions.
2. Identity aware applications
3. Federated authentication and Pass through mechanisms.

**R8**
1. Check compatibility of software
2. Check upgrade path of both the softwares
3. Have schemas and architecture very clear as how do you want to use data flow between them and how do you want to see it as a single solution rather than 2 softwares.

**R12**
1. Avoiding synchronous communication between clouds as much as possible by engaging an acquire-store-resend model.
2. Monitoring connections in the integration hub at all available levels by reserving a mechanism for an automated acquiring of lost connections.
3. Putting the maximum attention on semantics and ontologies of operations and data involved in the interactions between clouds.
4. Minimizing the number of interactions between clouds.

**R4** Cloud computing vendors have formed a common platform cloud computing interoperability forum (CCIF) to address the problem of cloud interoperability and standardization. Unified cloud computing is trying to unify various cloud APIs and abstract it behind an open and standardized cloud interface. Thus a key driver of the unified cloud interface (UCI) is to create an API about other APIs

**R2** Common Standards

**R5** DK
R6 Flexible licensing options

R7 Follow the proper regulations, security, standardization

R9 It depends upon the specific scenario and would be driven by the applications which have to be interoperable. As of now the IaaS and PaaS providers are still evolving rapidly so thoughts around integration are premature. Bigger challenge is for a market facing service provider who has to offer a bouquet of services to the end customer and these services are hosted on different platforms.

R11 Standard interfaces and security Inter-operable software Cross-platform communication software This would be best achieved by following a global standard in implementing cloud platforms. But I doubt if such a thing exists.

R3 Two or more systems or components are used for exchanging information to provide effective and efficient exchange between more service provider components like SAS, SSIS.

R13 We deploy a platform/interface for management/orchestration between different service providers.

Observations:
From the answers that are given by the responder to question one, it can be understood that the responses to this question are mostly guidelines. 3 out of 13 respondents for this question say that using standardization will be better option and one of these 3 say that “global standards is a best option. But I doubt if such thing exists” (response from Founder and CEO, i.e., R11). From the remaining responses, R5 says DK (don’t know).

a) What precautions can be taken to protect client data while interfacing between multiple cloud service providers?

Responses: Total Respondents to this question are 14. Given below are the solution/practice/guideline mentioned against each respondent.

R2 Make sure that Vendor follow compliance best practices like ISO27k, SOX, etc

R3 Encryption between Data and Other service provider . Network Security and Firewall Blockage at and within Data center.

R4 Data Integration Data Encryption and security Data Authentication Usage Policy

R5 DK
R6 DATA SECURITY IS CRITICAL. TYPICALLY I HAVE NOT SEEN FOOL PROOF DATA SECURITY BETWEEN MULTIPLE SERVICE PROVIDERS. IT IS BEST TO RESTRICT DATA IN AN EXTENDED PRIVATE NETWORK OVER A SINGLE CLOUD PROVIDER.

R7 Use proper authentication, security for each service provider.

R8 Have public cloud and check for sql injections and security measures that each cloud vendor offers. Have continuous monitoring of data slippage.

R9 Can’t comment.. have not thought about it.

R10 Again, if well-defined standards (similar to HIPAA for transfer of healthcare information across service providers) exist, the client data will be well-protected and strict rules will be defined for transfer between multiple service providers.

R11 High encryption Perimeter defense Access control Controlled ports Isolation

R12 Do not trust any declarations of SaaS providers regarding security. Protect your channel to SaaS from the integration hub (this is one of the major roles of having such hub) with all security means your corporate policies specify. If your applications are deployed in another cloud, the communication channel with this one has to be equally protected.

R13 Secure connections, multi-layer access controls and user education

R14 (1) TDE implementations are already in industry.
(2) 256 bit encryption
(3) Three phase encryption with hardware acceleration
(4) DLP mechanisms with secure replication

R15 Scalability, security and performance

Observations:
Most of the responses (6 out of 14) given to this question consist encryption as the solution, such as: 256-bit encryption, data encryption, encryption between data and other service provider, etc. Answers that need to be noted are:

- “Data security is critical. Typically I have not seen fool proof data security between multiple service providers. It is best to restrict data in an extended private network over a single cloud provider” (R6, 2 years experience).
“Do not trust any declarations of SaaS providers regarding security. Protect your channel to SaaS from the integration hub (this is one of the major roles of having such hub) with all security means your corporate policies specify. If your applications are deployed in another cloud, the communication channel with this one has to be equally protected.” (R12, 7 years experience).

As an answer to this question most of the respondents say that, there are security measures such as encryption techniques or secure connections with multi layered access control to be implements. From the above-mentioned points, it makes clear that cloud computing is not secure and interconnection between cloud is to be taken extreme care.

Second question:

a) If service provider allows migration, what steps (methods used, precautions taken or guidelines) can be followed to provide easy migration without compromising security?

Responses:

Total Respondents to this question are 12. The responses of each respondent in detail are as follows

R2 Not sure. No experience on this

R3 Data Center and Other service provider should be used the steps like Network Security and Firewall Blockage at and within Data center.

R4 There are some set of rules to be followed prior to migration. both the parties must agree on then, then only migration is possible

R5 DK

R7 Nothing have any thought

R8 (1) Have a sound schema - which should be in your full control.
    (2) Don’t rely too much on services offered by vendor in core areas.
    (3) Check the new cloud vendor offering with integrating your software, with security offerings

R9 Cant comment as we are not using PaaS

R10 Transfer of security policies across multiple service providers will not be possible until there is a common underlying framework that all service providers implement. Since this does not exist, we are left to individual expert consultants to advice on migration.

R11 Standard interfaces and data formats
R12 A workload that executes in one cloud provider can be uploaded to another cloud provider. Some standardization efforts that support this use case are Amazon Machine Image (AMI), Open Virtualization Framework (OVF), and Virtual Hard Disk (VHD). Data that resides in one cloud provider can be moved to another cloud provider. A standardization effort that supports this use case is Cloud Data Management Interface (CDMI). In addition, even though SOAP and REST are not data-specific standards, multiple cloud-storage providers support data- and storage-management interfaces that use SOAP and REST. A user who has established an identity with a cloud provider can use the same identity with another cloud provider. Standardization efforts that support this use case are Amazon Web Services Identity Access Management (AWS IAM), OAuth, OpenID, and WS-Security.

R13 (1) Create robust clauses within the contracts - to ensure timely and accurate extractions of data, followed by data erasal, upon termination of services
(2) Create maps between the source and target service providers - to ensure seamless data migration
(3) Always keep periodic backups of essential/critical data, onsite. Needless to say, most confidential information is better stored on-premise

R14 This is already in place. MSFT to G to Amazon (tools exists for import) Security for transfer of data/meta-data is done with encryption. Problems occur if domains do not allow rights keys.

Observations:
Three responder didn’t answer this question i.e., people with 2.5 years, 2 years and 1-year experience had no answer. Those who have answered have given unique steps to be followed to handle the secure migration in a cloud computing environment. Interesting responses to this question are as follows:

• “Transfer of security policies across multiple service providers will not be possible until there is a common underlying framework that all service providers implement. Since this does not exist, we are left to individual expert consultants to advice on migration.” Which means having a common set of rules followed by different service providers could benefit migration of data (R10, 5 years experience, founder and CEO).

• “This is already in place. MSFT to G to Amazon (tools exists for import) Security for transfer of data/meta-data is done with encryption. Problems occur if domains do not allow rights keys” Which means that
there a migration assistance provided by service providers (response by R14).

• The respondent R12, has listed out the frameworks, protocols and standards that are helpful to look at when a user wishes to migrate his system from one CSP to another. In this case his responses are more specific to the Amazon service provider.

a) If service provider does not allow migration, what precautions need to be followed to avoid data lock-in?

Responses:
The total respondents to this question are 10 and their answers as follows:

R2 Need to ensure that contracts are written in a way that we don’t get into this sort of mess

R3 Smooth network traffic within and out of data center. You can use agreement , RAA (Risk Analysis Acceptance) between vendor and service provider during migration activity

R4 I am not able to comment on this topic

R5 DK

R7 Nothing have any thought

R8 Regular back up of data in your personal system space.

R10 We need to ensure that the application inherently provides the ability to export data in some form that will make it easy to migrate to a different provider.

R11 Store data internally, and use service provider only for processing.

R12 According to me the below measures should be taken to avoid vendor lock-in in case a migration path does not exist :

(1) Data movement and encryption, both in transit and when it reaches the target environment.

(2) Setting up networking to maintain certain relationships in the source environment and preparing to connect into different network options provided by the target environment.

(3) The application itself, which lives in an ecosystem surrounded by tools and processes. When the application is moved to a target cloud, you may have to re-architect it based on the components/resources that the target cloud provides.

R13 (1) Find ways to replicate data that’s stored in the cloud

(2) Avoid storing sensitive data off-premise
R14  (1) read the fine print.
    (2) Most service providers will give back the data as is
    (3) Meta data management is usually a hurdle.

Observations:
To avoid this data lock-in due to no migration can be solved by following precautions suggested by respondents:

(a) Ensure that the contract are written well
(b) Store data internally and use the cloud for processing
(c) Replicate data stored in the cloud and also avoid storing sensitive info
(d) Design an application on cloud interface that could give export user data. This will solve the problem of exporting or fetching data manually.

Third Question:

a) What steps can be employed to provide compatibility between on-premise authentication systems and cloud computing systems?

Responses: Total Respondents observed to this question in survey are 10. Following are the responses mentioned against the respondents notation.

R2  For Cloud to be successfully vendors need to figure out how to integrate all of these prevailing identity management systems. Can't keep reinventing stuff.
R3  Restricted Zone , Limited access in production zone
R5  DK
R7  Nothing have any thought
R8  Have an interface between your application and social media login’s , populate your DB with the same on regular basis and then allow users to use the req. uid/passwd accordingly.
R10 Not sure.
R11 Similar software/hardware stacks Same patch levels on both Pre-production validation environments
R12 I have not experienced this scenarios, but my suggestion would be to adopt some in-house development of modules to help smooth interaction between on-premise authentication systems and cloud computing systems. Since organization confidential and user credential data flow is involved via such interfaces, encryption would be recommended for such communication channels.
(1) Create an on-premise layer of authentication

(2) Explore security brokers to help integrate on-premise and off-premise systems.

(3) Setup a local web server (with on-premise security) to serve up data to the cloud users directly from the the local web server, so there is a physical disconnect between the data in cloud and the data on the local web server but to the user it looked like all the data was in cloud.

Firstly IDM is not enough. Identity is to be tagged across application, users and hardware.. e.g. Mr. X can run Program Y on server Z and W only... this is achieved with compatible signed hardware software combos.

Observations:
There are 4 respondents who had no experience in this situation or had no answer (7 years, 5 years, 2.5 and 2 years). A respondent (person who is part of 6 of the largest implementations in the industry. Across platforms, engineering and hosting) says “Firstly IDM is not enough. Identity is to be tagged across application, users and hardware.. e.g. Mr. X can run Program Y on server Z and W only... This is achieved with compatible signed hardware software combos.”

Another respondent says, “For Cloud to be successfully vendors need to figure out how to integrate all of these prevailing identify management systems. Can’t keep reinventing stuff.”

There were few opinions, which suggested how to create compatibility between on-premise and cloud computing systems.

a) What on-premise authentication system compatible with cloud computing system do you use?

Responses: Total Respondents to this question are 10. Following are the responses:

R2 LDAP Integration

R3 It is very useful as every employee who is a part of CC they have to follow the rules by cc as limited permission within and out of production zone , never share your password with others.

R5 DK

R7 Nothing have any thought

R8 It’s always preferable to use SSO login, which is in your control.

R10 Haven’t used anything yet.
**R11** o-auth

**R12** SSO, LDAP, MS Windows authentication etc

**R13** Oracle Identity and Access Mgmt solution

**R14** Many.

**Observations:**

From the responses it can be identified that there are a few respondents unaware of the IDM used. Others mention all most all the IDMs identified through literature review. Following table 5.1 shows which on-premise authentication system compatible with which service provider in a cloud computing system are used. For example, from the table 5.1, we can understand that Director for IT/SW Services who has experience with cloud applications for a year says LDAP integration is compatible with google cloud. The other respondent in this survey who is an associate director and has been using Amazon cloud for 3 years says its preferable to use SSO (single sign-on) as it is compatible with cloud applications. Similarly the other responses have been tabulated to show which on-premise authentication system is compatible with and from which service provider.

**Table 5.1:** Compatible on-premise authentication systems used with cloud

<table>
<thead>
<tr>
<th>Position (experience with cloud in years)</th>
<th>Type of cloud used</th>
<th>Type of compatible on-premise authentication system used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Director IT/SW Services (1 year)</td>
<td>Google cloud</td>
<td>LDAP integration</td>
</tr>
<tr>
<td>Associate director (3 years)</td>
<td>Amazon cloud</td>
<td>Preferable to use SSO</td>
</tr>
<tr>
<td>Senior technology practice manager (5 years)</td>
<td>Amazon cloud</td>
<td>O-auth</td>
</tr>
<tr>
<td>Solution architect (7 years)</td>
<td>Microsoft and Amazon cloud</td>
<td>SSO LDAP MS windows authentication etc.</td>
</tr>
<tr>
<td>Strategy and governance leader (8 years)</td>
<td>Amazon and salesforce.com</td>
<td>Oracle identity and access management solution</td>
</tr>
<tr>
<td>Chief architect (part of 6 largest implementations in the industry)</td>
<td>All the service providers</td>
<td>Says many!</td>
</tr>
</tbody>
</table>

**5.3 Summary**

This section summarizes the observations of the survey based on the following.
1. Through this survey we have answers (guidelines) to the problems identified in literature and presented in the survey questions:
   a. Avoid customer lock-in.
   b. Secured migration between two or more service providers (an example is identified in the survey that shows Google to Amazon tools exist).
   c. We have also a list of on-premise IDMs that are compatible with cloud computing service providers.

2. Answers given to every question in the survey by respondents is different and is a guideline or precautionary measure that users can follow to overcome/avoid the problem mentioned.

3. Few users stress the use of global standards or common standards (to provide interoperability) and standard interfaces (for migration), as this did not quite exist in the cloud computing scenarios.

4. Some of the answers mention that there is utmost care required when data is migrated, interfaced or interconnected in cloud computing scenario. The respondents mention different ways of encryption and security measures that are to be implemented while transferring information from one cloud to another.

5. From these surveys it can also be understood that framing strict standards and having open source ways of designing CC applications is also stressed for better security and compatibility of different cloud service providers.
Chapter 6

Discussion

Cloud computing is the most interesting technology among organizations and most of these are organizations are concerned about its security. To identify CC security concerns a literature review and snowball sampling were employed, which revealed 43 challenges and 89 mitigation strategies. The identified challenges are categorized and are explained in chapter 4. In addition to the challenges and the mitigation strategies there are few guidelines and suggestions observed from literature. The identified mitigation strategies and guidelines are discussed in section 4.3. The identified challenges which can be mitigated using the observed strategies are tabulated in section 4.4. This helps the reader to identify the mitigation strategies that can be applied to effectively handle a challenge when encountered during CC implementation. All this done to answer the research question 1 and 2, which is to identify the security challenges and the mitigations. The categorization of challenges into groups used in this report is done to simplify explaining challenges. Grouping of challenges in this way was also used by other authors (as been highlight in section 2.7), it was also mentioned that using another type of categorization is suggested in further studies to avoid the missing of security requirements.

Observations from the literature clearly state that most of the challenges identified through literature have at least one solution/guideline/practice/architecture/model specified that can effectively handle the challenge. From the tables and results mentioned in ‘section 4.4’ it can be observed that there are some challenges which have no mitigation strategies. The challenges with no mitigation strategies are, ‘Packet sniffing’, ‘Sharing computing resources’, ‘Failure in providing security’, ‘Incompatibility issue’, ‘Constant feature addition’, ‘Recovery’, ‘Compliance’ and ‘Cost’. This study focuses on one challenge that has more article mentioning the challenge and no mitigation strategies identified through literature. The challenge chosen to be focused in this report is ‘Incompatibility’. This is because there are 19 articles mentioning the challenge but no dependable mitigation strategies identified.

Later, for this challenge identified (i.e., incompatibility), a survey method was employed to gather expert opinion on mitigation strategies from practitioners. To do this author needs to be aware of areas where this challenge can affect the
practitioners and then questionnaire for survey need to be framed. The insecure areas of this challenge were identified from the literature review and are as follows:

1. Interoperation between different service providers,
2. Migration of data, application from one service provider to another,
3. Compatibility of traditional (on-premise) authentication methods with cloud security systems.

From literature review there were very vague practices/suggestion/mitigation strategies identified, which suggest using of open cloud manifesto standards. These were not mentioned to be used by any company (observed from chapter 4). The survey questionnaire in this study was designed by taking into account the above mentioned points and focused on gathering expert opinion from practitioners. This answers the 3rd research question. The survey questionnaire is presented in appendix B and the list of respondents is presented in table C.1 (due to confidentiality reasons detailed information of respondents can not be presented).

The results of the surveys show that there is a need for security. The respondents give guidelines, some precautions and their experiences as to how the challenge situations mentioned in the questionnaire can be handled. From the above, 3 insecure areas in cloud, there have been satisfactory responses from surveys (observed from experienced people in cloud computing). The respondents stress if global standards are available the incompatibility issue with CC (problems of interoperability between different service provider) can be avoided. In addition to this it is also observed that, while interoperability being facilitated between multiple service providers, utmost care is required to be taken and channels through which data is being communicated need to be secured. Failing to do so can be risky and also information leakage (number of preventive measures are given to facilitate secure interoperability between cloud providers). From the surveys list of all compatible on-premise authentication systems that can be used with cloud computing and what steps should be followed to provide secure migration between two or more service providers (to avoid data lock-in situations) are observed. Detailed survey responses to these questions are discussed in chapter 5.
6.1 Validity threats

The following are the validity threats that can be identified in this thesis.

- There might be a chance of missing papers when selecting initial papers for snowball sampling. To avoid this authors of initial papers were contacted and requested to suggest papers that might be helpful when included in our study. In addition to this the author performed a literature review in addition to snowball sampling to strengthen the results of this study.

- Since there are numerous challenges identified from literature review and if they are not properly recorded there are chances of missing some important challenges or solutions. To prevent this from happening there was a data extraction form maintained which helped in tracking challenges identified.

- Too many challenges listed in an unorganized way, can confuse the reader. To avoid this, challenges are categorized under 8 sections, some of them might find this categorization unclear and hard to understand what they constitute. To avoid this confusion, each categorization is explained clearly in terms, which were collectively grouped under them.

- The challenges were categorized into different sections and few challenges with similar threat and risk were grouped under a challenge. Based on this categorization the identified mitigation strategies were also mentioned. Possibility of grouping irrelevant challenges and mitigation strategies can happen, which is avoided by giving a detailed explanation of each challenge and their mitigation strategies. Doing this will remove confusion among the readers.

- Collecting information from a large group of practitioners could have biased results. There might be chances for people replying even when they are not related to cloud computing. To avoid these practitioners who have high experience in this field were contacted after searching in different forums and companies that implement cloud.

- Since the surveys were conducted through a survey link, there is possibility that respondents find it hard to understand the purpose of the survey. To avoid this issue each respondent sent a startup mail in addition to the introduction of purpose mentioned before presenting questions in the survey link.

- There is a possibility for responders finding it difficult to understand the terms used in interviews. To avoid this issue, interviews were conducted only with those persons who have worked in this area. They were also provided with an e-mail id, through which respondents could contact the authors and clear their doubts.
There is a threat that readers might consider these results are generalized, to avoid this problem it is stated in the report that these generalized results are obtained from expert opinion or from literature and can differ with the change from organization to organization.
Chapter 7

Conclusions and Future Work

7.1 Conclusions

This study collectively describes cloud computing security challenges in general and describes the mitigation practices that have been proposed to handle the identified challenges. But there are still some challenges with no mitigation strategies, which might stand as a risk and a concern for some enthusiastic CC lovers. Through this study the author tried to focus on one such challenge ‘incompatibility’ and find mitigation practices from CC practitioners.

The conclusion shall be divided into 3 subsections and inferences obtained from each research question is explained below:

RQ1 & RQ2: After an exhaustive search (using literature review and snowball sampling) on electronic databases, 43 challenges and 89 mitigation strategies were identified in cloud computing security. In addition to these mitigation strategies there are also some guidelines been identified.

RQ3: The guidelines and mitigation strategies identified through initial study are not available for all the challenges identified. There are some challenges with no mitigation strategies and from them a challenge that is mentioned by multiple articles is selected. For this challenge (‘incompatibility’) identified from literature that has no mitigation strategies, solutions/practices/mitigation strategies are found from cloud computing practitioners.

RQ4: Based on the survey method applied to identify the practices to mitigate this ‘incompatibility’ challenge there are some conclusions made and as follow:

1. There are ways to provide interoperability between multiple cloud service providers (data/work load/instances/application), but there is no proper standard or standard interface observed by cloud users. Each of them have designed a series of steps through which this is interconnectivity is established. It can also be observed that this individually designed steps by users could have flaws as they are not standardized.
Hence, the suggested methods need to be deployed with extreme caution to prevent security risks (can be observed by the responses from the survey that suggest the usage of multiple encryption techniques at various levels to prevent data leakage).

2. For migration of data, an issue pointed in survey, the responses state that there are some standards user can look at while making a decision to migrate data/application to another CSP (but these are given in specific to Amazon services). These standards not all cloud users are familiar, which is because there is no common framework that is existing with the CSP to provide migration. In addition to this, in case migration is not allowed by the service provider, the best way to stay secure and safe from data lock-in is by not uploading sensitive information to the cloud and make sure that the contract is well written.

3. For the on-premise authentication (IDM’s) being compatible with cloud security, the conclusions can be that there are no proper seamless integration strategies existent and most of the cloud user have to depend on vendors. In addition to this, there are some set of on-premise authentication strategies need to be identified which can be integrated with cloud service providers that are mentioned.

7.2 Future work

There are very few studies that discuss all these 43 challenges and 89 solutions. Among these challenges, some do not have solutions or clear description. These need to be evaluated further to understand their impact on cloud computing and then discarded or practices should be designed.

However, most important future work identifies here is that there are concrete standards for cloud computing security still missing. There are some open cloud manifesto standards and few efforts made by the cloud security alliance to standardize the process in the cloud. The cloud vendors and users don’t encourage the usage of these standards as they are restrictive. In addition to this the cloud computing with such great offering such as storage, infrastructure and application designing capabilities on the go to the IT industry still fail to have proper standards for interoperability with other cloud service providers. This failure to provide concrete security standards, common underlying framework for data migration and global standards for cloud interoperability, make the leading technology “the cloud computing” still a vulnerable option for aspiring users.
References


important issues to highlight. In our research we review these concept ... 10, 15, 18


References


References


[107] Parastoo Mohagheghi and Thor Sæther. Software engineering challenges for migration to the service cloud paradigm. 25


References


References


References


References


### Initial Set of papers

*Table A.1:* Initial set of papers

<table>
<thead>
<tr>
<th>Sl.no</th>
<th>Papers</th>
<th>Authors responses (yes/no)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A survey on security issues in service delivery models of cloud computing</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Addressing cloud computing security issues</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>An Investigation into Cloud Configuration and Security</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Can a Trusted Environment Provide Security?</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Cloud Computing Implementation, Management, and Security</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Cloud Computing: Issues and Challenges</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Cloud computing security issues and challenges</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>Infrastructure as a Service Security: Challenges and Solutions</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>Locking the sky: a survey on IaaS cloud security</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>Privacy and Anonymization as a Service: PASS</td>
<td>No</td>
</tr>
<tr>
<td>11</td>
<td>Security and Privacy Challenges in Cloud Computing</td>
<td>Yes</td>
</tr>
<tr>
<td>12</td>
<td>Security Challenges in Cloud Computing</td>
<td>No</td>
</tr>
<tr>
<td>13</td>
<td>Storm Clouds Rising: Security Challenges for IaaS Cloud Computing</td>
<td>No</td>
</tr>
</tbody>
</table>
Appendix B

Survey questionnaire

The survey questions are distributed through an online survey link. This link is prepared with 2 pages. The first page is to gather basic information and the second page to gather solution/practices/guidelines.

**Page One:**

General information (OPTIONAL):
Name: ________
Designation: ________
Organization: ________
Total work experience in cloud: ________
Email-id: ________

**Question 1:** Who is your service provider for cloud computing applications?

a) Microsoft
b) Google
c) Amazon
d) Salesforce.com
e) Rackspace

Other, please specify: ________

**Question 2:** Types of services used

a) SaaS
b) PaaS
c) IaaS

Other, please specify: ________

**Question 3:** Types of cloud used

125
a) Public cloud
b) Private cloud
c) Hybrid cloud
d) Community cloud
Other, please specify: ________

Page 2

Thank you for sharing your basic information.

We are Master students in Software Engineering from Blekinge Institute of Technology (BTH), Sweden. Currently working on our Master thesis in Cloud Computing security under the guidance of Prof. Lars Lundberg, BTH, Sweden.

This survey aims to collect expert opinion and experiences on how to mitigate the issues identified under ‘incompatibility’. We have 3 major concerns identified under this challenge and they are:

1. Interoperation between different service providers,
2. Migration of data, application from one service provider to another,
3. Compatibility of traditional (on-premise) authentication methods with cloud security systems.

Following are few questions related to these issues and we request you share your opinion and experiences to mitigate them.

Please note that these questions in no way require any business information from you. We just need your OPINION and EXPERT advice.

1. Using resources from different cloud service providers collectively could help a company achieve improved scalability and performance. For example an application designed on Microsofts Azure platform using Oracle or Amazon storage services or vice versa. This interoperation between different service providers is a challenge in cloud computing.

   a) What steps can be followed to provide interoperability between two or more service providers?
   Character Count 1000

   b) What precautions can be taken to protect client data while interfacing between multiple cloud service providers?
   Character Count 1000
2. Company might want to migrate its data/applications from one service provider to another. This migration from one cloud to another is also a challenge, as while transferring the new service provider must not only transfer data but also the security policies enforced on client data. If migration is not allowed data can be locked-in with the service provider and might lead to conflicts.

(a) If service provider allows migration, what steps (methods used, precautions taken or guidelines) can be followed to provide easy migration without compromising security?

(b) If service provider does not allow migration, what precautions need to be followed to avoid data lock-in?

3. For any user to access the company data/applications he/she must be authenticated by an Identity management (IDM) system. Commonly used authentication ways are Microsoft Windows authentication, Single Sign-On (SSO), Light weight Directory Access Protocol (LDAP), Security Assertion Mark up Language (SAML), OPENID and OAUTH, OpenSocial, Facebook Connect, etc. Some of these methods are not compatible with cloud computing systems. If the enterprise IDM is not compatible with cloud platforms, providing access to employees for different cloud applications can become a challenge. Hence delaying in the adoption of cloud.

a) What steps can be employed to provide compatibility between on-premise authentication systems and cloud computing systems?

b) What on-premise authentication system compatible with cloud computing system do you use?
### Appendix C

#### Respondents

*Table C.1:* Respondents notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Position</th>
<th>Years of experience in cloud computing</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Asst. Manager</td>
<td>0</td>
</tr>
<tr>
<td>R2</td>
<td>Director IT/SW Services</td>
<td>1</td>
</tr>
<tr>
<td>R3</td>
<td>Associate Consultant</td>
<td>18 Months In Atos and 7+ yrs In IT</td>
</tr>
<tr>
<td>R4</td>
<td>Senior Software Engineer</td>
<td>2</td>
</tr>
<tr>
<td>R5</td>
<td>Jr. Programmer Datamatics</td>
<td>2 years</td>
</tr>
<tr>
<td>R6</td>
<td>BPM HEAD</td>
<td>2 YEARS</td>
</tr>
<tr>
<td>R7</td>
<td>Survey Programmer</td>
<td>2.5</td>
</tr>
<tr>
<td>R8</td>
<td>Associate Director</td>
<td>3 years</td>
</tr>
<tr>
<td>R9</td>
<td>CEO</td>
<td>3 years</td>
</tr>
<tr>
<td>R10</td>
<td>Founder and CEO</td>
<td>5</td>
</tr>
<tr>
<td>R11</td>
<td>Senior Technology Practice Manager</td>
<td>5 years</td>
</tr>
<tr>
<td>R12</td>
<td>Solution Architect</td>
<td>7 years</td>
</tr>
<tr>
<td>R13</td>
<td>Strategy &amp; Governance Leader</td>
<td>8 years</td>
</tr>
<tr>
<td>R14</td>
<td>Chief Arch</td>
<td>part of 6 of largest implementations in the industry. Across platforms, engineering and hosting.</td>
</tr>
<tr>
<td>R15</td>
<td>Sr.OTM Consultant</td>
<td>very less</td>
</tr>
</tbody>
</table>