
Issa M. Khalil 1,*, Abdallah Khreishah 2 and Muhammad Azeem 3

1 Qatar Computing Research Institute (QCRI), Qatar Foundation, Doha, Qatar; E-Mail: ikhalil@uaeu.ac.ae
2 Department of Electrical and Computer Engineering, Newark College of Engineering, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA; E-Mail: abdallah@njit.edu
3 College of Information Technology, United Arab Emirates University, PO Box 15551, Al Ain, United Arab Emirates; E-Mail: azeemsamma@hotmail.com

* Author to whom correspondence should be addressed; E-Mail: issa.khalil@yahoo.com.

Received: 5 September 2013; in revised form: 14 November 2013 / Accepted: 27 January 2014 / Published: 3 February 2014

Abstract: Cloud computing is an emerging technology paradigm that migrates current technological and computing concepts into utility-like solutions similar to electricity and water systems. Clouds bring out a wide range of benefits including configurable computing resources, economic savings, and service flexibility. However, security and privacy concerns are shown to be the primary obstacles to a wide adoption of clouds. The new concepts that clouds introduce, such as multi-tenancy, resource sharing and outsourcing, create new challenges to the security community. Addressing these challenges requires, in addition to the ability to cultivate and tune the security measures developed for traditional computing systems, proposing new security policies, models, and protocols to address the unique cloud security challenges. In this work, we provide a comprehensive study of cloud computing security and privacy concerns. We identify cloud vulnerabilities, classify known security threats and attacks, and present the state-of-the-art practices to control the vulnerabilities, neutralize the threats, and calibrate the attacks. Additionally, we investigate and identify the limitations of the current solutions and provide insights of the future security perspectives. Finally, we provide a cloud security framework in which we present the various lines of defense and identify the dependency levels among them. We identify 28 cloud security threats which we classify into five categories. We also present nine general cloud attacks along with various attack incidents, and provide effectiveness analysis of the proposed countermeasures.
Keywords: cloud computing; cloud security; security vulnerabilities; threats; attacks; insider attackers

1. Introduction

Cloud computing provides a centralized pool of configurable computing resources and computing outsourcing mechanisms that enable different computing services to different people in a way similar to utility-based systems such as electricity, water, and sewage. In electricity, for example, people started to connect with central grids, supported by power utilities rather than relying on their own electricity production capabilities. This migration is beneficial in reducing the cost and time of production and in providing better performance and reliability [1]. Similarly, clouds provide their customers with high performance and more reliable computing services such as e-mail, instant messaging, and web services at a lower cost.

Cloud computing does not have a common accepted definition yet [2]. The National Institute of Standards and Technology (NIST) [3] defined five essential characteristics of cloud computing, namely: on-demand self-service, broad network access, resource pooling, rapid elasticity or expansion, and measured service. Also, cloud computing is described as a dynamic and often easily extended platform to provide transparent virtualized resources to users through the Internet [4]. Cloud computing architecture consists of three layers: (i) Software as a service (SaaS); (ii) Platform as a service (PaaS) and (iii) Infrastructure as a service (IaaS) [5]. The clouds are also viewed as five component architectures that comprise clients, applications, platforms, infrastructure and servers [6]. The current clouds are deployed in one of four deployment models: (a) public clouds in which the physical infrastructure is owned and managed by the service provider; (b) community clouds in which the physical infrastructure is owned and managed by a consortium of organizations; (c) private clouds in which the infrastructure is owned and managed by a specific organization and (d) hybrid clouds which include combinations of the previous three models [7]. Figure 1 shows cloud deployment models together with their internal infrastructure (IaaS, PaaS and SaaS). Cloud deployment models have similar internal infrastructure, but vary in their policies and user-access levels.

Clouds bring out tremendous benefits for both individuals and enterprises. Clouds support economic savings, outsourcing mechanisms, resource sharing, any-where any-time accessibility, on-demand scalability, and service flexibility. Clouds minimize the need for user involvement by masking technical details such as software upgrades, licenses, and maintenance from its customers. Clouds could also offer better security advantages over individual server deployments. Since a cloud aggregates resources, cloud providers charter expert security personnel while typical companies could be limited with a network administrator who might not be well versed in cyber security issues. Similarly, clouds are more resilient to Distributed Denial of Service (DDoS) attacks due to the availability of resources and the elasticity of the architecture. The clouds support mobile computations where Virtual Machines (VMs) migrate from one physical machine to another. In addition to alleviating dedicated DDoS attacks, mobile computations help to avoid settings in which a single administrator has exclusive control over the computation.
Figure 1. Cloud deployment models and infrastructure.

The new concepts introduced by the clouds, such as computation outsourcing, resource sharing, and external data warehousing, increase the security and privacy concerns and create new security challenges. Moreover, the large scale of the clouds, the proliferation of mobile access devices (e.g., smartphones and tablets), and the direct access to cloud infrastructure amplify cloud vulnerabilities and threats. As clouds become more and more popular, security concerns grow bigger and bigger as they become more attractive attack targets due to the concentration of digital assets.

Many researchers and practitioners work on identifying cloud threats, vulnerabilities, attacks, and other security and privacy issues, in addition to providing countermeasures in the form of frameworks, strategies, recommendations, and service oriented architectures (e.g., [2,8–13]). Additionally, efforts in other domains such as ad hoc networks have been tuned to address the emerging security problems in the clouds (e.g., [14-18]). Many researchers (e.g., [1,19–25]) have addressed single attributes of cloud computing security such as data integrity, authentication vulnerabilities, auditing, etc. Others (e.g., [3,22–27]) provide surveys that cover specific areas of cloud security concerns and proposed solutions. In [3,27], the authors briefly and broadly discuss cloud security issues involving data, applications and virtualization. The authors in [22] discuss similar cloud security issues but with deeper investigations. In [23,25], the authors present surveys on cloud security requirements such as confidentiality, integrity, transparency, availability, accountability, and assurance. In [24], the authors present a survey on the different security issues of the service delivery models of the clouds. In [26], the authors discuss the security challenges specific to the public clouds. In [28], Hashizume et al. classify the security issues in the cloud based on the SPI (SaaS, PaaS, IaaS) cloud infrastructure and services model. The authors provide deeper classification of the fourth category (C4) in our classification model (Section 2). Additionally, the authors explain fundamental security concepts including vulnerabilities, threats, and attacks and provide mapping among these concepts. Our work provides a higher classification level and assumes prior knowledge of the fundamental security concepts. In [29], Zissis et al. evaluate cloud security requirements. They propose a Trusted Third Party solution that calls upon cryptography to ensure the authentication, integrity and confidentiality of data and communications. In [30],
Whaiduzzaman et al. present key management and broad aspects of privacy and security issues in the domain of vehicular cloud computing.

To successfully address the cloud security issues, we need to understand the compound security challenges in a holistic way. Specifically, we need to: (i) investigate various cloud security attributes including vulnerabilities, threats, risks, and attack models; (ii) identify the security requirements including confidentiality, integrity, availability, transparency, etc.; (iii) identify the involved parties (clients, service providers, outsiders, insiders) and the role of each party in the attack-defense cycle; and (iv) understand the impact of security on various cloud deployment models (public, community, private, hybrid). The main contribution of this paper is that it provides a holistic study of the security issues in the clouds that cover almost all the cloud components (data centers, computing infrastructure, interfacing and networking, etc.), network layers (application, transportation, IP, etc.), and cloud stakeholders (providers, consumers, third party contractors, etc.). In this paper, we provide a comprehensive survey on the cloud security and privacy concerns that includes: (i) cloud computing security issues (vulnerabilities, threats, and attacks); (ii) attack classifications; (iii) relations and dependencies among attacks; (iv) known attacks; (v) comparative analysis of some of well-known countermeasures; (vi) insights from the current security solutions to identify and address unattended security challenges.

This paper is organized as follows. In Section 0, we describe cloud security categories, issues and dependencies. We present some of the well-known attacks and countermeasures in Section 0. In Section 0, we present the findings of comparative evaluation of some well-known general cloud computing solutions. Further discussion on cloud security issues is presented in Section 0. In Section 0, we conclude the paper.

2. Cloud Security Categories, Issues and Dependencies

As part of this work, we have conducted a survey on the current cloud security issues and the state-of-the-art security solutions. We identify 28 security issues (Table 2) that we categorize into five classes (Table 1). We have also provided a comparative analysis of the current security solutions and the state-of-the-art countermeasures.

2.1. Categories and Issues

We classify cloud computing security related issues into the following five categories, which are also summarized in Table 1. A similar approach to classify the issues is found in [19] but it is limited to small set of cloud security concerns and only partially covers four categories.

(C1) The Security Standards category deals with regulatory authorities and governing bodies that define cloud security policies to ensure secure working environment over the clouds. It includes service level agreements, auditing and other agreements among users, service provider and other stakeholders.

(C2) The Network category refers to the medium through which the users connect to cloud infrastructure to perform the desired computations. It includes browsers, network connections and information exchange through registration.
(C3) The Access Control category is a user-oriented category and includes identification, authentication and authorization issues.

(C4) The Cloud Infrastructure category includes security issues within SaaS, PaaS and IaaS and is particularly related with virtualization environment.

(C5) The Data category covers data integrity and confidentiality issues.

**Table 1. Cloud Security Categories.**

<table>
<thead>
<tr>
<th>No.</th>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Security Standards</td>
<td>Describes the standards required to take precaution measures in cloud computing in order to prevent attacks. It governs the policies of cloud computing for security without compromising reliability and performance.</td>
</tr>
<tr>
<td>C2</td>
<td>Network</td>
<td>Involves network attacks such as Connection Availability, Denial of Service (DoS), DDoS, flooding attack, internet protocol vulnerabilities, etc.</td>
</tr>
<tr>
<td>C3</td>
<td>Access Control</td>
<td>Covers authentication and access control. It captures issues that affect privacy of user information and data storage.</td>
</tr>
<tr>
<td>C4</td>
<td>Cloud Infrastructure</td>
<td>Covers attacks that are specific to the cloud infrastructure (IaaS, PaaS and SaaS) such tampered binaries and privileged insiders.</td>
</tr>
<tr>
<td>C5</td>
<td>Data</td>
<td>Covers data related security issues including data migration, integrity, confidentiality, and data warehousing.</td>
</tr>
</tbody>
</table>

In Table 2, we map the identified cloud security issues into the suitable categories defined earlier (Table 1). We have labeled these issues with “I1, I2, ..., In” where Ix refers to cloud security issue number x. Special attention is required towards mutual security standards such as Secure Sockets Layer (SSL)/Transport Layer Security (TLS), XML signature, XML Encryption Syntax and Processing, and Key Management Interoperability Protocols. Currently, cloud computing lacks appropriate security standards (I1) [1]. Even if security standards are defined properly, many security issues are still associated with compliance risks (I2) due to lack of governess for audits and assessment of corporate standards [1]. Cloud customers do not have enough knowledge of procedures, processes and practices of the provider, especially in the areas of identity management and segregation of duties. Organizations that seek to obtain certifications may be put on risk by denying an audit by cloud customers. One of the most important aspect of cloud computing security is auditability (I3); however, we do not have an audit net for cloud service providers [13,31]. If a service provider outsources a service to a third party where functionality is not transparent, users must be able to inspect the whole process [12]. Security standards (C1) and governing bodies are part of service level agreements (SLA) (I4) and legal aspects, respectively which have not been taken into practices for cloud computing [32,33]. SLA defines the relationship among parties (provider—recipient) and is extremely important for both parties [9]. It includes identifying/defining the customer’s needs, simplifying complex issues, encouraging dialog in the event of disputes, providing a framework for understanding, reducing/removing areas of conflict, eliminating unrealistic expectations. The user may suffer, in case of data loss, if the above factors are not taken into consideration as he may not be able to put claims on service providers. These interactions shape the Trust (I5) relationship between the users and the different cloud stakeholders which is required when users transfer data on cloud infrastructure [34]. Strong justifications are required to gain customers’ trust in that regard.
Table 2. Cloud Security Issues and Classifications.

<table>
<thead>
<tr>
<th>Category</th>
<th>Label</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Standards</td>
<td>I1</td>
<td>Lack of security standards</td>
</tr>
<tr>
<td></td>
<td>I2</td>
<td>Compliance risks</td>
</tr>
<tr>
<td></td>
<td>I3</td>
<td>Lack of auditing</td>
</tr>
<tr>
<td></td>
<td>I4</td>
<td>Lack of legal aspects (Service level agreement)</td>
</tr>
<tr>
<td></td>
<td>I5</td>
<td>Trust</td>
</tr>
<tr>
<td>Network</td>
<td>I6</td>
<td>Proper installation of network firewalls</td>
</tr>
<tr>
<td></td>
<td>I7</td>
<td>Network security configurations</td>
</tr>
<tr>
<td></td>
<td>I8</td>
<td>Internet protocol vulnerabilities</td>
</tr>
<tr>
<td></td>
<td>I9</td>
<td>Internet Dependence</td>
</tr>
<tr>
<td>Access</td>
<td>I10</td>
<td>Account and service hijacking</td>
</tr>
<tr>
<td></td>
<td>I11</td>
<td>Malicious insiders</td>
</tr>
<tr>
<td></td>
<td>I12</td>
<td>Authentication mechanism</td>
</tr>
<tr>
<td></td>
<td>I13</td>
<td>Privileged user access</td>
</tr>
<tr>
<td></td>
<td>I14</td>
<td>Browser Security</td>
</tr>
<tr>
<td>Cloud Infrastructure</td>
<td>I15</td>
<td>Insecure interface of API</td>
</tr>
<tr>
<td></td>
<td>I16</td>
<td>Quality of service</td>
</tr>
<tr>
<td></td>
<td>I17</td>
<td>Sharing technical flaws</td>
</tr>
<tr>
<td></td>
<td>I18</td>
<td>Reliability of Suppliers</td>
</tr>
<tr>
<td></td>
<td>I19</td>
<td>Security Misconfiguration</td>
</tr>
<tr>
<td></td>
<td>I20</td>
<td>Multi-tenancy</td>
</tr>
<tr>
<td></td>
<td>I21</td>
<td>Server Location and Backup</td>
</tr>
<tr>
<td>Data</td>
<td>I22</td>
<td>Data redundancy</td>
</tr>
<tr>
<td></td>
<td>I23</td>
<td>Data loss and leakage</td>
</tr>
<tr>
<td></td>
<td>I24</td>
<td>Data location</td>
</tr>
<tr>
<td></td>
<td>I25</td>
<td>Data recovery</td>
</tr>
<tr>
<td></td>
<td>I26</td>
<td>Data privacy</td>
</tr>
<tr>
<td></td>
<td>I27</td>
<td>Data protection</td>
</tr>
<tr>
<td></td>
<td>I28</td>
<td>Data availability</td>
</tr>
</tbody>
</table>

Network category (C2) related issues are deemed to be the biggest security challenges in clouds since cloud computing is more prone to network related attacks compared to the traditional computing paradigms [2]. In addition, cloud operations are tightly coupled and highly depend on networking. Therefore, cloud network security issues receive more attention in this work compared to the other security categories. The ratio of network attacks and fraud dramatically increases as people and organizations migrate their data into clouds. Security experts anticipate that clouds will be the focus of hackers in future due to the concentration of valuable “assets” (data and computation) within the clouds. The possible lack of proper installations of network firewalls (I6) and the overlooked security configurations (I7) within clouds and on networks, make it easier for hackers to access the cloud on behalf of legitimate users [35]. Hackers can occupy resources (hardware/application) by generating bogus data or they can run malicious code on the hijacked resources. Denial of service can be launched by first identifying vulnerabilities in Internet protocols (I8) such as SIP (Session Initiation Protocol) which could deem the Internet to be un-trusted [36]. Migrating to cloud will increase the Internet...
dependency (I9) as a main communication medium for cloud access. Therefore, if, due to some attacks, the Internet is disabled and the cloud services become unavailable, this may cause production to become severely crippled [37]. I9, therefore, implies all the network reliability issues.

Account and service hijacking (I10) involves phishing, fraud and software vulnerabilities where attackers steal credentials and gain unauthorized access to servers [1]. This unauthorized access is a threat to integrity, confidentiality and availability of data and services [1]. Unauthorized access can be launched from within or outside the organization. Malicious insiders (I11) such as dishonest administrators severely impact organizations’ security. Given their level of access, they infiltrate corporate and cause brand damage, financial and productivity losses. Therefore, it is critical for cloud customers to clearly determine the guarantees that the cloud providers use to detect and defend against insider threats. The current authentication mechanisms (I12) may not be applicable in cloud environments as customers no longer belong to or are able to access a single tightly controlled system [4]. A single customer may access data and compose services from multiple cloud providers using a mobile application or a browser. This kind of access brings in an inherent level of risk and this risk has been called privileged user access (I13) [6]. Unauthorized access becomes possible through browser vulnerabilities. Therefore, Internet browser (I14) is the first stage where security measures should be considered because vulnerabilities in the browser open the door for many follow-on attacks.

The insecure interface of Application Programming Interface (API) (I15) issue covers the vulnerabilities in the set of APIs in the cloud portal (customers use these APIs to connect to a cloud) which can expose an organization to several threats such as unauthorized access, content transmission, reusable token and logging capabilities [1]. Quality of service (QoS) (I16) is an unattended issue [32] because many cloud service providers focus only on fast performance and low cost [31]. In this work, we consider QoS in the domain of any function or activity that directly or indirectly affect security. A simple error in the configuration of one or more of the cloud components may cause severe consequences because cloud configurations could be shared by many services [4]. Technical flaws (I17), also known as reputation fate sharing [38], in which errors transfer from a corrupted server to each virtual machine created on that server becomes worse when the corruption transfers through the infected mobile VM to other servers. Therefore, it is extremely important to identify and fix fate sharing incidents and implement the best practices to prevent them from reoccurring. Reliability of suppliers (I18) is an important factor that requires a background check on the staff to control data and hardware access [4]. It is highly recommended that companies should evaluate its staff in order to protect its assets and data and provide this information to public to gain customer’s trust. Servers in the cloud are the backbone to its infrastructure that provides numerous services such as directory service, data storage and mail. Intruders can access the system if the security attributes of the servers are not configured properly (security misconfiguration—I19) [39]. This misconfiguration could happen in the application stack, the framework, the web server, the custom code as well as the platform. Note that this is different from inadequate network security configuration (I7), which includes network level security misconfigurations. Cloud serves serve multiple simultaneous users through virtualization, which allows the sharing of the same software and hardware resources by different users. This multi-tenancy (I20) capability could lead to information leakage from one tenant to other server mates [40]. Attacks such as VM-to-VM and compromised VM are becoming hub for future attacks. In terms of server location (I21) precautions, it is important to keep in mind that the floor should be anti-static, should have no window
for security reasons, should have a rack with seismic bracings and should be properly grounded [41]. Cloud infrastructure cannot be completely trusted at this stage and it is critical to maintain backup offline.

Data redundancy (I22) [6], data loss and leakage (I23) [40], data location (I24) [6], data recovery (I25) [4], data privacy (I26) [42], data protection (I27) [43] and data availability (I28) [43] have been marked as major and important issues in different case studies which require data to be properly encrypted, transmitted, protected, controlled and available in the time of need.

Figure 2 shows the cloud components where security issues may be raised. Each component, such as policies, clients, cloud infrastructure, and network, is prone to certain security attacks and requires attack prevention/detection/response strategies.

Figure 2. Cloud components that are prone to security threats.

2.2. Dependencies among Cloud Security Categories and Issues

In addition to identifying cloud security issues and classifying them into several categories, we have identified dependencies among these categories and the security issues they encompass. If one of the categories is prone to certain attacks, other categories may also become prone to these attacks. For example, issues I1–I4 (Table 2) fall under category C1 (Table 1). Security related issues under this category might be entry doors through which other threats infiltrate into the cloud. Suitable management and security precautions taken in one category (e.g., C1) may greatly minimize or even eliminate security issues in the other categories (C2, C3, etc.). If proper policies are implemented at one category, then fewer issues will arise in other categories (only those issues that are genuine to a particular category rather than issues that take advantage of vulnerabilities in other categories). We have covered many threats/vulnerabilities in this survey and will now investigate the dependency relationships among them. We follow thematic analysis ([31,43]) to extract the dependency relationships among the cloud security issues we surveyed. Initially, we extracted the appropriate text from the literature we surveyed (column 1 in Table 3). We then identified the security issues in the selected text and performed manual coding to identify the particular features of that selected text. This coding process is essential for organizing the data into meaningful groups. Table 3 presents an example, which contains textual description (text
extraction), identified code, identified rule and the rule description. The symbol used in describing the rules is represented as “→”. It simply means that the security issue on the left of the symbol leads to or increases in the probability of occurrence of the issue(s) on the right of symbol.

Table 3. Data Extraction Example.

<table>
<thead>
<tr>
<th>Text Extraction (from Primary Studies)</th>
<th>Problems Coded</th>
<th>Identified Rule</th>
<th>Rule Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Account and service hijacking involves phishing, fraud and software vulnerabilities where attackers steal credentials and gain an unauthorized access to server [1].</td>
<td>• Account and service hijacking (I10)</td>
<td>Account and service hijacking</td>
<td>Account and service hijacking increase the possibility of data loss and leakage</td>
</tr>
<tr>
<td></td>
<td>• Data Loss and Leakage (I23)</td>
<td>→ Data Loss and Leakage</td>
<td></td>
</tr>
<tr>
<td>Insecure interface of API's (customers use these APIs to connect to cloud) which can expose organization to several threats such as unauthorized access, content transmission [1]</td>
<td>• Insecure interface of API’s (I15)</td>
<td>Insecure interface of API’s → authentication mechanism</td>
<td>Insecure interface of API’s leads to increase in the probability of information leakage and the weakness of the authentication mechanism.</td>
</tr>
<tr>
<td></td>
<td>• Weak authentication mechanisms (I13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Data Loss and Leakage (I23)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The same procedure has been followed in order to determine dependencies among other cloud security issues that we have identified in this survey. A sample of these dependencies has been presented in the form of rules in Table 4. These rules aim at providing practitioners with deep insights about cloud computing security issues to take suitable precaution measures. Practitioners can also trace the source of any security issue through these rules. For example, the main cause of wrapping attacks can be tailored to weak browser security as stated by Rule 7, whereas weak browser security can be linked to security misconfiguration as stated by Rule 11. These rules could be helpful in protecting a cloud computing environment against various attacks by providing a comprehensive view of the attacks’ map rather than considering each attack in an isolated setup.

Figure 3 shows a directed graph representation of the dependencies among the surveyed cloud security issues. In the figure, nodes represent security issues, a directed edge from node X to node Y indicates that the occurrence of issue Y is facilitated by the occurrence of issue X. Finally, codes on edges refer to the rules in Table 4, where Rn refers to rule number n.
Table 4. Rules that Capture Dependencies among Cloud Security Issues.

<table>
<thead>
<tr>
<th>No.</th>
<th>Rules</th>
<th>Sample References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lack of security standard (I1) → compliance risks (I2), account and service hijacking (I10), internet protocol vulnerabilities (I8), malicious insider (I11), authentication mechanisms weakness (I12)</td>
<td>[1,4,7,41,42]</td>
</tr>
<tr>
<td>2</td>
<td>Lack of auditing (I3) → malicious insider (I11)</td>
<td>[7,37,39,41,44,45]</td>
</tr>
<tr>
<td>3</td>
<td>Lack of legal aspects (Service level agreement) (I4) → authentication mechanism weakness (I12)</td>
<td>[1,2,4,19,32,46]</td>
</tr>
<tr>
<td>4</td>
<td>Insecure interface of API (I15) → account and service hijacking (I10)</td>
<td>[1,4]</td>
</tr>
<tr>
<td>5</td>
<td>Compliance risk (I2) → account and service hijacking (I10)</td>
<td>[41]</td>
</tr>
<tr>
<td>6</td>
<td>Weak Browser security (I14) → account and service hijacking (I10)</td>
<td>[4,41]</td>
</tr>
<tr>
<td>7</td>
<td>Internet protocol vulnerabilities (I8) → authentication mechanism weakness (I12)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Security misconfiguration (I19) → insecure interfaces of API (I15), account and service hijacking (I10)</td>
<td>[4,41–43]</td>
</tr>
<tr>
<td>9</td>
<td>Malicious insider (I11) → Data loss and Leakage (123)</td>
<td>[6,33]</td>
</tr>
<tr>
<td>10</td>
<td>Lack of quality of service (I16) → insecure interfaces of API (I15), Data loss and Leakage (123)</td>
<td>[4,6,33,41,43]</td>
</tr>
<tr>
<td>11</td>
<td>Authentication mechanism weakness (I12) → Data loss and Leakage (123)</td>
<td>[6,33]</td>
</tr>
</tbody>
</table>

Figure 3. Dependencies among cloud security issues.

3. Known Attacks and Countermeasures: An Evaluation

“All the vulnerabilities and security issues that on-premise, non-virtualized and non-cloud deployments have still remain in the cloud”, Lawrence Pingree, analyst for Gartner, said. “All that cloud and virtualization does is enhancing the potential risks by introducing virtualization software and potentially mass data breach issues, if an entire cloud provider’s infrastructure is breached [105]104.” In this work, we classify cloud applicable attacks into nine groups; provide sample attack incidents from
each group, and present comparative analysis of some of the famous security mitigation techniques. Table 5 presents a summary of attack group names, known attack incidents in each group, attack consequences, attack category (from Table 1), exploited vulnerabilities (from Table 2), and references for further readings about each attack group. A description of each attack group along with evaluation of the state-of-the-art countermeasures are presented in the following subsections.

3.1. Theft of Service Attacks

The Theft of Service attack [37] utilizes vulnerabilities in the scheduler of some hypervisors. The attack is realized when the hypervisor uses a scheduling mechanism, which fails to detect and account of Central Processing Unit (CPU) usage by poorly behaved virtual machines. This failure may further allow malicious customers to obtain cloud services at the expense of others. This attack is more relevant in the public clouds where customers are charged by the amount of time their VM is running rather than by the amount of CPU time used. Since the Virtual Machine Manager (hypervisor) schedules and manages virtual machines, vulnerabilities in the hypervisor scheduler may result in inaccurate and unfair scheduling. These vulnerabilities mainly result from the use of periodic sampling or low-precision clock to measure CPU usage: like a train passenger hiding whenever ticket checkers come for tickets. In the Theft of Service attack, the hacker ensures that its process is never scheduled when a scheduling tick occurs. The common incidents of this attack include: (1) using cloud computing services (e.g., Human Resource, HR, systems) for long period of time while keeping it hidden from the vendor and (2) using cloud computing resources (e.g., storage system or OS platform) for a long period without representing it in a billing cycle.

A countermeasure to this attack has been provided by Zhou et al. in [37] by modifying the scheduler to prevent the attack without sacrificing efficiency, fairness or I/O responsiveness. These modifications do not affect the basic credit and priority boosting mechanisms. The modified schedulers are: (1) exact scheduler; (2) uniform scheduler; (3) passion scheduler and (4) Bernoulli scheduler. The main differences among these schedulers are in the scheduling and monitoring policies and in time-interval calculations. The experiment conducted by authors with the modified schedulers provides accurate and fair scheduling. The modifications in hypervisor are shown to be beneficial, as compared to Xen hypervisor (currently running in Amazon Elastic Compute Cloud—EC2).

Another theoretical countermeasure has been provided by Gruschka et al. in [44]. They suggest using a new instance of cloud-to-user surface in victim machine to monitor the scheduling of parallel instances. Then, the outputs of both the attacker and the legitimate instances are compared. A significant difference in results is reported to the responsible authorities as an attack. This solution has not been validated or verified by authors and does not provide any guarantee for a beneficial result. There are other solutions provided for hypervisor scheduling such as [45,47,48] but they are only limited to improving other aspects of virtualized I/O performance and VM security such as CPU-bound issues. These studies do not examine scheduling fairness and accuracy in presence of attackers, which is the backbone for the Theft-of-Service attack.
### Table 5. Known Attacks against Clouds.

<table>
<thead>
<tr>
<th>Sr. #</th>
<th>Attack Name</th>
<th>Attack Incidents</th>
<th>Consequences</th>
<th>Category</th>
<th>Vulnerability/ Caused by</th>
<th>Sample References</th>
</tr>
</thead>
</table>
| 1     | Theft-of-service |                  | • Cloud service usage without billing  
• Cloud resource stealing with less/no cost | Cloud Infrastructure | I1, I3, I6, I8, I11, I14, I19 | [37,44] |
| 2     | Denial of service | DDoS, Http-Based DDoS, Xml-Based DDoS, REST-Based-DDoS, Shrew attack (light traffic) DoS | • Service/hardware unavailability  
• Wrapping a malicious code in Xml signature to gain unauthorized access to information  
• Accessing a browser history or any other private information through unsecure Http browsing | Network, Cloud Infrastructure | I1, I3, I10, I14, I19 | [34,49] |
| 3     | Cloud malware injection |                  | • Credential information leakage  
• User data leakage  
• Cloud machine abnormal behavior | Cloud Infrastructure | I7, I11, I13, I15 | [44,50–52] |
| 4     | Cross VM side channels | Timing side channels, Energy-consumption side channels | • User data/information leakage  
• Cloud resources/infrastructure information leakage | Cloud Infrastructure | I15, I19 | [53–55] |
| 5     | Targeted shared memory |                  | • Cloud resource’s information leakage  
• User information/data leakage  
• Provides open window for other attacks such as side channels and cloud malware injection | Cloud Infrastructure | I1, I3, I10, I15, I19 | [53,56,57] |
### Table 5. Cont.

<table>
<thead>
<tr>
<th>Sr. #</th>
<th>Attack Name</th>
<th>Attack Incidents</th>
<th>Consequences</th>
<th>Category</th>
<th>Vulnerability/ Caused by</th>
<th>Sample References</th>
</tr>
</thead>
</table>
| 6     | Phishing    |                  | • Unauthorized access to personal information  
        • Installing a malicious code into user computer  
        • Force cloud computing structure to behave abnormally  
        • Make server unavailable for end-user | Cloud Infrastructure, Network, Access | I1, I6, I8, I10, I12, I14 | [58,59] |
| 7     | Botnets     | Stepping stone attack | • Unauthorized access to cloud resources  
        • Make cloud system work abnormally  
        • Stealing sensitive information  
        • Stealing user data | Network, Cloud Infrastructure, Access | I1, I6, I10, I12, I14 | [60,61] |
| 8     | Audio Steganography |                  | • Unavailability of cloud storage system  
        • Accessing user data  
        • User data deletion | Cloud infrastructure, Access | I1, I3, I6, I10, I14, I19 | [62,63] |
| 9     | VM rollback attack |                  | • Launch brute force attack  
        • Damage cloud infrastructure  
        • Leakage of sensitive information | Cloud Infrastructure, Access | I1, I3, I6, I10, I14, I19 | [64,65] |

### 3.2. Denial of Service Attacks

Most of the serious attacks in cloud computing come from denial of service (DoS), particularly HTTP, XML and Representational State Transfer (REST)-based DoS attacks. The cloud users initiate requests in XML, then send requests over HTTP protocol and usually build their system-interface through REST protocols such as those used in Microsoft Azure and Amazon EC2. Due to vulnerabilities in the system-interface, DoS attacks are easier to implement and very difficult for security experts to countermeasure [49]. XML-based distributed denial of service (DDoS) and HTTP-based DDoS attacks are more destructive than traditional DDoS because these protocols are widely used in cloud computing with no strong deterrence mechanisms available to avoid them. HTTP and XML are critical and important elements of cloud computing, so security over these protocols becomes crucial to providing healthy development of a cloud platform.
Karnwal et al. in [49] provide a framework called “cloud defender” that is based on five stages: (1) sensor filter; (2) hop count filter; (3) IP frequency divergence filter; (4) puzzle resolver filter and (5) double signature filter. The first four filters detect HTTP-based DDoS attacks and the fifth filter detects XML-based attacks. REST-based attacks are mentioned without providing a framework to prevent it. One of the reasons could be that REST-based attacks are closely related with the user interface, which may vary from user-level to system-level applications. These applications are different in nature, based on different requirements, and there is no single hard and fast rule to implement the security measurements at the interface level. The solution provided in this article consists of the following modules:

- **Sensor**: It monitors the incoming request messages. If it identifies that there is a hypothetical increase in the number of messages coming from the same or particular consumer, it marks it as suspicious.
- **HOP Count filter**: It will count the hop count value (how many nodes, does message traverse from source to destination) and compare it with pre-defined HOP count. If a difference is found, it means that the header or the message has been modified on the hacker machine and thus is marked suspicious.
- **IP Frequency Divergence**: Marks a message suspicious, if there is the same frequency of IP messages.
- **Double Signature**: It doubles the XML signature: one in the header and one in the bottom. In case of attack, both XML signatures need to be verified.
- **Puzzle Solver**: It deals with some intelligent puzzles, where results should be imbedded in some Simple Object Access Protocol (SOAP) header. In case of attack (HTTP DDoS), the cloud defender will send back the puzzle to the IP, from which it is receiving messages. If the cloud defender received back the solved puzzle then the request is deemed legitimate, otherwise it is marked as HTTP DDoS attack.

The problem in this framework is that it lacks practical validation and is based on the assumption that the number of modules in the system is directly proportional to the number of attacks expected. Moreover, exhaustive monitoring of messages on each node would considerably slow the network traffic. Finally, the framework lacks the proper mechanisms for node coordination in case of attack incidence detection.

Riquet et al. [34] claim that “there is no strong solution available to prevent the DDoS attacks”. To validate the claim, the authors conduct the experiment to evaluate the effectiveness of the actual security solutions against distributed attacks. The security solutions involved in the experiment are SNORT and commercial firewall. The authors conclude that the failure of security systems lies within two aspects: either the security solution can be obsolete because it is not updated, or the solution can rely on unsuitable methods. They did not propose any solution that can prevent distributed security attacks. Other widely used DDoS countermeasures are firewalls. However, due to firewall location (at the border of a network), it would not be able to detect distributed attacks once they are in the network [50].
3.3. Malware Injection Attacks

Cloud malware injection attack refers to a manipulated copy of the victim’s service instance, uploaded by attacker to cloud, so that some service requests to the victim’s service are processed within that malicious instance. An attacker can get access to user data through this attack. The attacker actually exploits its privileged access capabilities in order to attack that service security domain. The incidents of this attack include credential information leakage, user private-data leakage and unauthorized access to cloud resources. The challenge does not only lie in the failure to detect the malware injection attack but also in the inability to determine the particular node on which the attacker has uploaded the malicious instance [44]. Retrospective detection (examination of hard-drive and memory) has been a widely used technique to detect the host of malware instances. Liu et al. in [50] propose a new retrospective detection approach based on portable executable (PE) format file relationship. This approach has been implemented and validated in HADOOP platform. This approach proves higher detection rate as well as lower false positive rate. The main drawback of this approach is that its success is based on three assumptions (pre-requisites): (1) most legitimate programs and malware files are in PE format and lie within a windows platform; (2) the number of legitimate files is greater than that of malware files in user’s computer; and (3) creating/writing/reading PE format files seldom happen in a user’s computer. However, an attacker could exploit any vulnerability in cloud to attack without following any of these pre-requisites. The authors fail to discuss the consequences of the absence of these pre-requisites such as (1) how efficient this approach would be if one or more of the assumptions are not fulfilled; (2) how much damage and attacker could cause to system or data in absence of these assumptions.

Another countermeasure to the attack called “CloudAV” is provided by Oberheide et al. in [51]. CloudAV provides two main features that make it more efficient, accurate and fast as a malware detection system:

- Antivirus as a network service: the detection capabilities by host-based antivirus can be more efficiently and effectively provided as cloud-network-service. Each host runs a light weight process to detect new files and then sends them to network service for quarantine and for further analysis rather than running complex analysis software on each end-host.
- N-version protection: malicious software identification is determined by multiple heterogenous detection engines in parallel similar to the idea of N-version programming. The notion of N-version protection has been provided in this solution so that the malware detection system should leverage detection capabilities of multiple heterogeneous detection engines to determine malicious and unwanted files more effectively. However, the number of false positives encountered during normal operations increase compared to 1-version engines. To manage the false positives, the administrator has to set a trade-off between coverage (a single detector is enough to mark a file as malicious) and false positives (a consensus of a number of detectors is required to mark a file as malicious).

The authors prove the efficiency of CloudAV through validation in a cloud environment. CloudAV also provides better detection of malicious software, enhanced forensics capabilities, new threat detection through retrospective detection approach and improved deployability and management. The validation experiment proves that CloudAV provides 35% better detection coverage against threats
compared to single antivirus engines and 98% detection coverage of an entire data-set of a cloud. However, cloud-based security solutions generally suffer from three problems, namely—security coverage, scalability, and privacy. As malware can be embedded in a large number of file types, attackers may be able to bypass cloud solutions as they are limited to few file types and hence degrade the detection coverage. Additionally, exporting all binaries or PDF files to the cloud for investigation does not scale and may create a single point of failure by flooding the cloud with benign binaries. Finally, exporting the binaries and files into the cloud for inspection creates privacy issues since there is always a risk that a sensitive file could be exported to the cloud as well.

In [52], the authors provide a framework based on behavior analysis to detect suspicious programs in clouds. It allows end users to delegate security labs, the execution and the analysis of a potential malicious program and force the program to behave as if it were executed directly in real end-user environment. There are two advantages for this: (1) it allows security lab to monitor the execution of potential malicious programs in realistic end user environment and (2) it allows end-users to raise their level of protection by leveraging better computational resources [52].

### 3.4. Cross VM Side-Channel Attacks

VM side channel attack is an access-driven attack in which an attacker VM alternates execution with the victim VM and leverages the processor caches to infer the behavior of the victim. It requires that the attacker resides on a different VM on the same physical hardware as that of the victim’s VM. Ristenpart et al. in [57] discuss a comprehensive example on how to collect information from a target VM through cross VM side channel attack. One incident of side channel attacks is the timing side channel attack [54] which is based on measuring how much time various computations take to perform. Successful modulation of this measured time may lead to leakage of sensitive information about the owner of the computation or even the cloud provider. Timing channels are especially hard to control and pervasive on clouds due to massive parallelism. Moreover, timing side channel attacks are hard to detect since they do not leave trails or raise any alerts. Cloud customers may not have the authorization to check for possible side channels from other cloud mates obviously due to privacy concerns. On the other hand, cloud providers can thoroughly check and detect timing attack incidents but may not be willing to report such breaches due to many considerations such as protecting company reputation. Another incident of side channel attacks is the energy-consumption side channel attack [55]. Instead of directly attacking the software stack (virtualization layer), attackers can indirectly collect sensitive information about the cloud using energy consumption logs. This type of data (energy consumption log) is maintained to monitor the infrastructure status and to provide computer energy efficient workload mapping. In [55], authors investigate the potential of extracting valuable information from raw energy consumption logs which may affect user’s privacy and security. Dozens of hypervisors could exist in a cloud computing environment, each of which could be the host of the targeted VM. Therefore, it may take a while for the attacker to determine which hypervisor is hosting the targeted VM. The more time it takes an attacker to determine the host machine, the higher is the probability of attack detection. However, if the attacker can somehow get the power consumption data, it may become possible for him to narrow the possible set of servers that could be running the targeted VM. This gives attackers better
chances to determine the correct server before being detected. Currently, no efficient solutions have been provided against timing-side channel attacks and energy consumption side channel attacks.

3.5. Targeted Shared Memory Attacks

In this attack, attackers take advantage of shared memory (cache or main memory) of both physical and virtual machines. It is an initial level attack in cloud computing that can lead up to several different types of attacks such as side channel attacks and malware injection attacks [53]. For example, authors in [57] perform cross-virtual-machine-side-channels attack on Amazon EC2 and measure the cache activity of other users, which provides an example of activity-information leakage in cloud computing. Attackers can get unauthorized access to information that reveals the internal structure of the cloud such as the number of processes running, the number of users logged-in in a specific time and the temporary cookies residing in memory. Another example of targeted shared memory attack is explored by Rochsa and Correia in [56]. The goal is to access the memory dumps in virtual machines through malicious insider attack. This access has led to the extraction of the current running processes in the system and users’ private information.

Thus far, in the literature, no one has claimed to solve or prevent targeted shared memory attacks. Researchers and practitioners are working to get more information about the attack and no strong solution is available to prevent it except current anti-viruses or firewalls that limit users’ access to the shared memory.

3.6. Phishing Attacks

Phishing is an attempt to access personal information from unsuspecting user through social engineering techniques. It is commonly achieved by sending links of webpages in emails or through instant messages. These links appear to be correct, leading to a legitimate site such as bank account login or credit card information verification but they practically take users to fake locations. Through this deception, the attacker can obtain sensitive information such as passwords and credit card information. Phishing attacks can be classified into two categories: (1) an abusive behavior in which an attacker hosts a phishing attack site on cloud by using one of the cloud services and (2) hijack accounts and services in the cloud through traditional social engineering techniques [58].

Cloud security alliances (CSA) mentioned that cloud service providers do not maintain sufficient control over systems in order to avoid being hacked or spammed. To prevent such attacks, CSA proposes a few precaution measurements such as strict registration process, secure identity check procedure and enhanced monitoring skills [58]. Privacy laws in cloud computing do not allow cloud service providers to look at what customers are doing, so if a malicious individual or organization is performing something nefarious (phishing attack or uploading malicious code) by using cloud services, it cannot be detected until or unless notified by some security software [59]. Researchers in [58,59] discuss the fact that the present cloud privacy laws restrict cloud providers to become the first to know about nefarious activities in their clouds, regardless of the enhanced monitoring and comprehensive inspection of network traffic.
3.7. Botnet Attacks (Stepping-Stone Attack)

In Stepping Stone, attackers try to achieve their goals (such as spying, DoS, damaging, etc.) while avoiding revealing their identities and locations to minimize the possibility of detection and trace-back. This is achieved by indirectly attacking the targeted victim through a sequence of other hosts (called stepping stones). Stepping stone hosts can be recruited through illegal botnets. A bot-master, through botnet attack, can setup command and control server and stepping-stones into clouds in order to steal sensitive information and to gain unauthorized access to cloud resources in a bid to make it behave abnormally. In recent years, botnet attacks have been reported in Amazon EC2, Google AppEngine and Raytheon UK. A Zeus command and control was hosted on Amazon EC2 cloud [66]. A computer was infected by using relay commands through Google AppEngine [67], which allowed attackers to steal sensitive information from the Raytheon cloud [68]. Cloud computing is an ideal environment for botnet attacks [60]. A cloud has rich and elastic computing resources (bandwidth, processing and storage), which are easy to access. The attacker can either compromise a cloud-based server as their command and control server or they can lease a high performance virtual machine with the help of fake or stolen credit cards.

Many researches worked to countermeasure botnets and stepping stone attacks by identifying whether a particular host is a stepping stone or not. Most of the detection work is based on the hypothesis of strong correlation between the incoming and outgoing traffic of a possible stepping stone host. Such correlation can be based on packet content, login behavior, frequency of network activity, timing properties, and periodicity of network traffic. However, many of these techniques are easy to fool by attackers using encrypted traffic and authentication forging or by introducing random delays (jitter), while others are shown to be inefficient due the huge traffic that need to be monitored and analyzed. One of the most notable detection techniques is presented by Lin et al. in [60]. They introduce what they call a “pebble” trace scheme to trace-back the bot master. It first identifies the cryptographic keys of the botnet communication in order to configure the botnet operations and then it traces back the bot master. It involves the design and implementation of a new key identification scheme and an approach for tracing-back bot master across stepping-stones beyond multiple clouds. The solution only considers symmetric key cryptography with no discussion about asymmetric cryptography.

A different mechanism, based on self-protection, has been present by Kourai et al. in [61]. The mechanism, dubbed xFilter, is a packet filter that runs in a virtual machine that monitors the underneath VM and achieves pinpoint active response by using VM introspection. xFilter inspects the memory of the VM being monitored, using VM introspection, and obtains information about guest operating systems without interacting with them. xFilter can deny packets, by using sender process’s information, from particular sender or process. When xFilter detects an outgoing attack, it automatically identifies the attack source and generates a new filtering rule in order to stop the stepping-stone attack. This mechanism is proved to be efficient because even if cloud server is compromised, this mechanism will continue to provide other services as much as possible. For example, when the apache server is compromised, only the privileges of user www-data are taken over at worst, while other applications such as Postfix mail server will continue running legitimately. This solution has a limitation that can be dangerous for the business of cloud providers. For example, the attacker may intentionally use local SMTP server to mount SPAM attacks because other legitimate applications are also using SMTP server.
to send emails. If xFilter detects the SPAM attack, it will update the rule repository to deny all the traffic coming from the SMTP server including that from legitimate applications or users.

Srivastava et al. in [69] provide a detection mechanism that inspects the Internet traffic. This solution, named VMwall, is based on application-level firewall using VM introspection. One of the main differences of this solution compared to xFilter is that it degrades the network performance considerably when large number of packets and nodes are to be inspected. On the other hand, since xFilter inspects the server memory, the degradation would be minimized [69].

3.8. Audio Steganography Attacks

Audio Steganography attack has been regarded as one of the most serious attack to cloud storage systems. Audio Steganography helps users to hide their secret data within regular audio files. The steganography user can transmit secret information through sending media files, which appear to be normal sound files. Hackers utilize this feature to deceive the current security mechanisms or traditional countermeasures (e.g., steg-analysis) for protecting cloud storage systems by hiding their malicious code in sound files and sending it to victim servers [70]. Very little research has considered proposals to thwart Audio Steganography attacks against cloud storage systems, which make it an open area that requires practical solutions [71].

Liu et al. in [71] perform a careful analysis of Audio Steganography attack on cloud storage systems. They design and implement a scheme called StegAD (steganography Active defense) to tackle the threat of data leakage by using Audio Steganography attacks. The first step, in this solution, is to scan the hiding place of audio files under cloud storage system through famous RS image gray scale steg-analysis algorithm. After acquiring suspicious files, authors use SADI (Steganography Audio Dynamical Interference) technique to interfere in all the possible places in those suspicious files. Authors try to avoid damaging innocent files (which have been marked as suspicious during the scanning process) though adopting an approach where an interference has to be in multiple hiding places or most significant places. However, this solution does not provide any information about how to decide the significant hiding places. Authors first use random noise to replace information including the most significant one and then compare the previous unchanged information with the changed one. By doing so, steganography and innocent files will have different consequences. This difference will determine whether the audio file has appropriate content or some malicious code that can damage the cloud storage system. There are many questions that have not been answered in this solution such as: (1) How is the comparison between steganography and innocent file carried out? (2) What are the specific types of audio files considered in this experiment?

3.9. VM Rollback Attacks

The virtualization environment in cloud computing is the most vulnerable area to attack. The hypervisor can suspend a VM at any time during execution, take a snapshot of current CPU states, disk and memory and resume a snapshot later without guest VM awareness. This feature has been widely used for fault tolerance and VM maintenance; however, it also provides an open window to an attacker to launch VM rollback attacks. In a rollback attack, a user can take advantage of previous snapshots and run it without the user’s awareness and then clean the history and again run the same or different
snapshot. By cleaning the history, the attacker will not be caught for his suspicious activities. For example, an attacker can launch a brute force attack to guess a login password for VM, even if the guest OS has a restriction on the number of attempts such as blocking the user after three failed attempts or erasing all data after 10 times, the attacker can still rollback the VM to its initial state after each try. The attacker will clear the counter inside the VM and bypass the restriction and run the brute-force attack again [64].

Szefer et al. in [71] provide an architecture named “Hyperwall” in order to manage hypervisor vulnerabilities. The solution to prevent VM rollback attack is based on disabling the suspend/resume functionalities of the hypervisor. The suspend/resume feature is powerful for virtualization and disabling it will not provide a better solution. Another limitation of this solution is the excessive user interaction with the cloud system. It requires end users to get involved during VM booting, suspending and resuming. This means that the system needs to ask for permission every time it reboots, migrates or suspends a VM which makes it inconvenient and impractical [71]. Xia et al. in [64] provide a solution that works without disabling any basic functionalities of the hypervisor as compared to the Hyperwall. In this solution, only the end user can tell whether a rollback is malicious or not by auditing the log of VM activities. Although this solution has minimized the user involvement compared to the Hyperwall, the changing infrastructure of cloud computing still demands autonomous working of VM operation with some user involvement.


The European Union Agency for Network and Information Security (ENISA) has done significant work in addressing many security issues related to the cloud. It provides stakeholders with information that helps them to understand, assess, and manage the risks when migrating into the clouds. It also provides advisory services on setting and monitoring SLAs to optimize security gains. ENISA also provides cooperative studies with various stakeholders to identify the critical cloud services and analyze the impact of the cloud service failure in such circumstances. In the following subsections, we present the state-of-the-art general tools that are individually and collectively used to countermeasure cloud security attacks.

4.1. Intrusion Detection Systems (IDS)

Intruders, through impersonating legitimate users, can access cloud infrastructures causing it to be unavailable for legitimate users. It has been shown that attackers can easily get information regarding victim machines in the IaaS component of the cloud [70]. This information can help in attacking cloud users by co-locating the malicious virtual machine with the victim’s virtual machine. These attacks include denial of service (DoS) and distributed denial of service (DDoS) attacks that mainly target data confidentiality, integrity, and availability [72]. Such attacks can be avoided by implementing IDS which offers additional security measures by investigating network traffic, log files and user behavior [73]. IDS is defined as a system that collects and analyzes information, from different key points, for security auditing and monitoring in order to check whether this is a violation of network policies or not. Intrusion detection can be classified into two categories (1) misuse detection (MD) and (2) anomaly detection (AD) [74]. MD deals with information characteristics of users input and making a comparison with
database results (previous inputs by the same user). On the other hand, anomaly detection stores user behavior in feature database, which can be compared with current behavior. If there is a high rate of difference in comparison, then the invasion occurred. IDS have two types (1) host based IDS (HIDS) which monitors the behavior on a single host and (2) network based IDS (NIDS) which analyses traffic flowing through a network [75]. Roschke et al. in [76] discuss a third type called hybrid IDS also known as distributed IDS, which combines the functions of both NIDS and HIDS.

The authors in [73] propose an IDS system named Grid and Cloud Computing Intrusion Detection System (GCCIDS) which is based on NIDS and HIDS. It consists of an audit system that detects and covers attacks that have not been covered, previously, by other NIDS and HIDS systems. It works by integrating knowledge and behavior analysis in order to detect intrusions. The main components of GCCIDS include node, service, event auditor and storage system. The main limitations in GCCIDS include the high communication overhead and redundancy. Each node identifies the local event and alerts all other connected nodes. The overhead increases dramatically when the number of nodes increases due to the massive computations and communications involved. Additionally, GCCIDS does not provide any information as to whether a node should immediately alert other nodes as intrusion occurs or at a certain predefined periodic time interval. Let us consider the case of immediate sharing with a grid, for example, of 1000 nodes. Each time intrusion occurs, the detecting node will alert all the other 999 nodes generating high overhead of communication exchange. Attackers can utilize this ‘alert sharing’ as an open window to disrupt the network. Attackers behave as intruders in different intervals of time, triggering certain nodes to detect these intrusions and hence notifying other nodes, which considerably increases the communication overhead. If a node selects not to immediately alert other nodes, inconsistency arises. A node that identifies an attack contains the information about the attack, while other nodes are not aware of it. Each node, in this system, consists of a local database that has information related to intrusions occurred previously. The local repository leads to redundancy.

GCCIDS uses both knowledge-based and behavior-based techniques for attack incident detection. Knowledge-based techniques cannot detect new attacks because detection depends on pre-defined rules. However, this limitation can be alleviated through regularly updating local repositories with information about new attacks. For behavior-based technique, GCCIDS uses feed forward artificial neural network (FFANN). However, FFANN is not useful at the starting phase due to little or no data availability. However, as the time passes and more computations are done by FFANN, the results improve.

The redundancy of repository in GCCIDS has been removed in an IDS system called Distributed, Collaborative and Data driven Intrusion Detection and Prevention framework (DCDIDP) [77]. DCDIDP creates global database to be used for detection tasks by the ID prevention module. DCDIDP consists of three level architectures, namely: network, host and global infrastructure. Network and host architecture maintain local database of policy and rules and contribute to global database. The global database shares information regarding intrusions among different clouds. The main features of DCDIDP include being distributed (policies are distributed among hosts), collaborative (hosts collaborate with each other to stay synchronized for information sharing) and data driven (dynamic evaluation of rules and access list). DCDIDP can be implemented in IaaS, PaaS and SaaS and provides effective intrusion prevention. However, the collaboration among different clouds requires an extensive trust management, which does not exist in the current DCDIDP framework. It does not provide approaches to promote trust among cloud users beyond what is stated in SLA. Information sharing among different clouds is also
dependent on the structure of each cloud. Finally, DCDIDP has not been evaluated and verified through practical implementations.

Dastjerdi et al. [78] propose a new IDS system by combining and extending the peer-to-peer IDS based on mobile agents [79] and the distributed intrusion detection using mobile agent (DIDMA) [80]. It consists of four main components namely IDS control center (IDS CC), agency, application specific static agent detector and specialized investigative mobile agent. IDS CC is central part of IDS component administration. Dastjerdi et al. claim that their proposal reduces network load and provides better trust management.

Figure 4 represents the overall view of IDS structure that consists of NIDS and HIDS. Each IDS (regardless of being deployed on distributed networks, grid or cloud) requires some basic components such as alert sharing, analyzer, technique to detect the suspicious behavior, etc. Researchers and practitioners are providing solutions, in the form of IDS, DIDS or HIDS, by reducing the number of components and by increasing the performance of IDS. Cloud infrastructure is changing so rapidly that current intrusion detection systems are not flexible enough to cope with these changes. One solution can be autonomous IDS that can update its policy as soon as cloud infrastructure changes. An unattended area here is to develop mechanisms for distributed IDS (Network or host based), which involve both inter and intra clouds considerations. Infrastructure heterogeneity, among clouds, and use of efficient communication protocols poses another challenge in cloud security.

4.2. Autonomous Systems

An autonomous system is an IDS that works with pre-specified basic rules. These rules configure, heal, optimize and protect themselves automatically, thereby reducing human efforts and involvement [81]. These rules are specified either by management or through artificial intelligence. It is impossible, without autonomic computing, to manage next generation distributed-systems such as clouds and grids effectively. Deolitzshcer et al. [82] present an intelligent autonomous agent for incident detection named Security audit as a Service (SaaS). This agent is assumed to be aware of the underlying business flow of
instances of the deployed clouds. SaaS collects data directly at the source, analyses it, aggregates information and then distributes it. The central part of this system is based on data analysis that is performed through security service level agreement (SSLA). SaaS addresses three main problems of cloud computing namely: (1) abusing cloud resources; (2) missing security monitoring in cloud infrastructure and (3) defective isolation of shared resources. SaaS involves the use of large number of sensors that can capture many events. These sensors receive security policies from SSLA. Even though SaaS is based on autonomous agent, its security policy is still based on pre-defined rules, which limits the detection capabilities only to those attacks that are already known. Moreover, SaaS involves the use of large number of agents and functions such as initiating agent, moving agent, killing agent, etc., which creates high communication among agents and increases the computational overhead.

Balen et al. [83] provide a pure concept of autonomous system, as compared to SaaS [82], particularly autonomic manager in grid and cloud computing. It provides a feature of self-configuration, self-healing, self-optimization and self-protection. The autonomous system is one which includes the autonomic manager, which is able to build and execute plans for implementation, based on sent and received information. A pure autonomous system architecture is presented by Sodhi et al. in [84]. This architecture is based on IaaS, where control of cluster nodes is fully autonomous. This architecture uses real time information from cluster nodes and decentralizes the policy management from the master node to other working nodes. It has several main components, namely: (1) cloud controller—a gateway for clients into cloud which determines the suitable node to run VM that satisfies client needs; (2) cloud agent—an intelligent software component that responds to the queries of the cloud controller regarding the availability of VM configuration for a specific lease duration; (3) VM foundry—VM image repository interface dedicated to answer queries for particular VM configurations and it creates the one-time-URLs for the VM image. The cloud agent is further based on several components including request handler, VM manager, policy manager, capability manager and data store. The key point which differentiates this system from other IaaS based systems is that it consists of decentralized policy management rather than based on master–slave relationship architecture that provides a bottleneck issue. If a problem occurs in the master, it may cause the system to function abnormally or even to shut down. This issue has been avoided through decentralization of policies to other nodes. Workload distribution mechanisms for IaaS are static. Decentralization needs to be an autonomous, due to rapid change in cloud infrastructure. Decentralization of policy management among different nodes will increase reliability and security (e.g., if one node is compromised, information leakage from other nodes can still be controlled).

Non-functional requirements with respect to autonomous IaaS node structure is still not much explored. Yazir et al. in [85] discuss an approach for dynamic and autonomous resource management in cloud computing with respect to scalability, feasibility and flexibility. Similar to [84], Yazir et al. present an approach that deals with decentralization of autonomous tasks, carried out by independent agents. The authors’ contributions are based on two steps (1) resource management is decomposed into independent tasks and each task is executed by autonomous node agent by adopting distributed architecture and (2) configuration is carried out by node agent through multiple criteria decision analysis (MCDA) using PROMETHEE method [86]. The simulation of this system proves to achieve scalability through distributed approach which reduces computational complexity. The approach used in the system is potentially feasible in large data centers as compared to centralized approaches. Dynamic resource
allocation provides higher flexibility due to its ability to change/add configurations. The main purpose of introducing an autonomous system in a cloud is to avoid wasting time and to utilize resources efficiently. The authors conclude that the PROMETHEE approach is promising with respect to non-functional attributes such as scalability, feasibility and flexibility although there is no function for proper resource management among virtual machines [87]. For example, HDD volume and memory volume setting is done manually or by cost estimation. There is no flexible management system by number of user access, at a given time or by load averages.

4.3. Federated Identity Management System

The backbone of cloud computing security is tightly coupled with identities used to access cloud infrastructure. Management of identities (IDM) is about maintaining the integrity of identities, throughout their life cycle, to make it and its related data (e.g., authentication and authorization results) available to different services in secure and privacy-protected manner [88]. The concept of federated identity management (FIM) is about managing identities by allowing an identity subject to establish links between his/her identities, each of which can be used for a different service, across geographical and organizational borders [88]. Establishing a logical link between identities is called identity federation [88]. The federation is a group of organizations that establish trust among themselves in order to cooperate safely in business [89]. The process to repeat authentication of user (Single Sign-On) can be an example of federated identity [89]. The main issue with single Sign-On lies in the wider damage that it causes in case of compromise. If a user identity is compromised, the illegitimate user will not be verified again, which could create higher level of information leakage. Another issue is lacking dynamic federation and agile mechanism in FIM systems [90]. It is an architectural concern and requires further investigation.

Figure 5 represents the basic functionality of IDM that consists of eight steps: login, requesting application or data, requesting token for ID verification, generating token, verifying it and sending/accessing the data/application. The user logs in IDM through his login ID (step 1) and requests to access application/data from the cloud (step 2) simultaneously. The cloud asks the user for a suitable token to be generated by the IDM for further authentication (step 3). The user requests the token from the IDM (step 4). The IDM generates the token and shares it with the user and the cloud (step 5). The user sends the token to the cloud in order to complete the step for final authentication (step 6). The cloud compares the token sent by the user and that sent by the IDM (step 7). Finally, if the verification succeeds, the cloud grants access to the user. There are three major standards to generate tokens in IDM namely—SAML, OpenID and Information Card. SAML consists of different sets of technical standards in order to implement the FIM. The notable service designed by SAML is single sign-on. It has been deployed widely and follows the strict security/privacy requirements, e.g., enterprise, governments and telecom [91]. OpenID is similar to SAML except it provides a smaller set of functions together with simple expressions of identity related data [91]. In Information Card standard, identities are managed as a set of cards. In the real word, a person usually has a set of cards such as national ID and driving license in order to represent her and all these cards are used for different purposes. The same mechanism has been applied in Information Card standards.
Leandro et al. [89] propose multi-tenancy authorization system for cloud computing using SAML standards (Shibboleth [92]). They promote the use of Shibboleth as access control system without using a trusted third party [89]. Shibboleth consists of four components (1) handle service—designated to authenticate user and issue handle token to the user; (2) attribute authority—designated to handle requests for software product attributes and apply corresponding privacy policy; (3) directory service—attribute storage of local user and (4) authentication mechanism—designated for user authentication with central service through login/password. It provides strong authorization but not authentication. In other words, it is difficult to steal identity during the running session once the user is authenticated with Shibboleth. However, Shibboleth does not provide a mechanism to ensure the legitimacy of the person connected with the system. Illegitimate person with valid username and password can avail the services without being double-checked. With this argument, the use of shibboleth does not guarantee complete secure transactions. Moreover, the multi-tenancy authorization system generates separate identity provider for each user/organization where each identity provider is using same policies. This particular architecture design helps users to modify security policies according to their need. However, outsourcing identity providers few concerns [93] including: (1) how to make sure that there will be no shift of data and processes to another location; (2) how to delete data after expiration of contract and (3) how to avoid both accidental and deliberate interference across the domain of clouds. In principle, there is no scheme to match the access control requirements with those that are provided by cloud providers. Organizations have no control over their data even after signing a contract with a cloud provider because it is not an easy task to enforce that the cloud provider will always meet the enterprise’s requirements.

Another concept of central approach in IDM systems is discussed in [94,95]. Angin et al. in [94] propose entity-centric approach called Active Bundle Scheme for IDM with comparison of application-
centric approach. This approach allows the entity to (1) create/manage its digital identities in order to authenticate in such a way that it does not reveal its actual identity and relationship between identities to vendor/service providers and (2) to protect personally identifiable information from unauthorized access. The same concept has been discussed in [89] except that Angin et al. do not implement or validate the solution. Identity-centric IDM is rationale response for the need of trust and security, generated by business explosion and social transaction, facilitating the exploit of an ever-increasing amount of personal data. The work in [74] refers to a particular type of Internet where trust and security are native, by design. This concept can be achieved through (1) total separation of storage of personal data and their exploitation by organizations and (2) through disaggregating organization centric data systems to re-aggregation around identity-centric systems.

5. Discussion

Cloud computing is an emerging paradigm that involves all the basic components of computing such as end-user machines (PCs), communication networks, access management systems and cloud infrastructures. To achieve comprehensive cloud security, the data and cloud infrastructure must be protected against known/unknown attacks across all cloud components. The number of browser based attacks in 2011 increased from 580,371,937 to 946,393,693 [96]. This notable increase is mainly due to the wide adoption of cloud computing which makes the platform very attractive for attackers due to the growing value of the data assets and resources available on the clouds. Unfortunately, we cannot protect the cloud-computing infrastructure from all the known/unknown attacks because it requires additional computational overhead and resources. Current security solutions like IDS, DIDS, firewallsing, outsourcing the identity management systems, installing antivirus, etc., [97-102] are expensive and degrade performance. Therefore, the major cloud security research challenge lies not only in providing high level security measures but also in doing so with minimum resources and reduced performance degradation.

Most of cloud vendors falsely claim to provide secure data and computational environments for cloud users. However, for such claims to be realistic, collective efforts are required at higher levels (e.g., governing bodies) instead of leaving it to individual organizations. Our current work helps in achieve that goal by providing a comprehensive study of the attacks against clouds, establishing dependencies among various attacks, and correlating attacks to vulnerabilities across various cloud components. This study can support the endeavors to provide preventive measures as well as proactive tools in defending the clouds. Using this study, we found that data and system security should be embedded in the design of cloud architecture to achieve better security. Moreover, security measures should be dynamic and autonomous. Cloud computing infrastructure is changing fast requiring security measures and policies to be updated regularly at the same pace to match the changing behavior of the clouds. Furthermore, licensing is crucial to the security of clouds. Standard policies should be strictly implemented in clouds and organizational/governing bodies should visit clouds’ staff and infrastructure on regular bases to evaluate the efficiency of the security precautions adopted by the vendors. The statistics for attacks, occurring in any cloud, should be publically available to determine the reliability of cloud vendors. This type of sharing helps other cloud’s security experts to guard against new attacks. Also, it is extremely important to holistically investigate the various cloud security related parameters
including risks, threats, challenges, vulnerabilities, and attacks. The possibility of being attacked can be reduced by deeply understanding the dependencies among these parameters. Finally, note that virtualization is a backbone of cloud computing. However, the concept of using virtualization in cloud computing is not yet mature as there are numerous number of attacks that target the virtualization environment. Examples of these attacks include information leakage during VM migration, service theft by manipulating VMs, uploading malicious VMs on cloud server, and rolling back VMs. Therefore, it is extremely important to develop reliable schedulers that, by design, contain sufficient security mechanisms.

We have identified a few areas that are still unattended in cloud computing security such as auditing, and migration of data from one cloud to another. Emphasis in current research has always been on fast performance and low cost services, but the quality (e.g., availability, scalability, reliability, etc.) of service has not been seriously considered. Non-functional requirements with respect to autonomous IaaS node structure is still not much explored. Another big concern is related with data. Data migration from one cloud to another is not achievable, at this moment, because of the heterogeneous nature of clouds. Moreover, clouds lack the tools that ensure that user data has been deleted from the cloud if the contract has expired. These data privacy threats require researchers’ attention to provide some standards for permanent data deletion. We plan in the future to explore the possibility of providing suitable frameworks for DIDS for heterogeneous clouds together with scheduling algorithms for Green IT.

Other related research issues in the area of cloud infrastructure include mobile computing challenges [103,104]. In mobile platforms, limited memory, low processor speed and higher computational requirements create hurdles in the efforts to provide best performance on these platforms. In addition to this, mobile applications provide higher level of threat, as there is a weak or even no security “check and balance” on application’s development. It has been claimed in [91], that 20% of the 48,000 applications, in android market, allow third party applications to access your sensitive data as well as allowing it to make calls and send texts without user consents. Apple Inc. has implemented a security framework in which each application should be submitted to Apple security team to determine the threat to user data or system before including it in the App Store.

6. Conclusions

The adoption of cloud computing paradigm is continuously growing. In 2010, the IT spending in America to migrate to cloud computing solutions was estimated at $20 billion. Analysts believe that the cost reduction factor in cloud computing will further accelerate the adoption of cloud computing in the public sectors. With the massive growth in cloud computing adoption, the security attracted the attention of researchers and practitioners but still has not received enough attention.

In this work, we conduct a survey on the current cloud security issues and the state-of-the-art security solutions. We identify 28 cloud security issues such as firewall misconfigurations, malicious insiders, tampered binaries, multi-tenancy, side channels, weak browser security, and mobility. Then, we classify these issues into five security categories, namely: security standards, network, access, cloud infrastructure, and data. We also identify nine attack classes that target the clouds and present variable incidents of each attack such phishing, fate sharing, botnet, and malware injection. For each attack class, we present the state-of-the-art countermeasures and provide a comparative analysis of the effectiveness
and the shortcomings of the proposed solutions. Finally, we present and evaluate the effectiveness of the state-of-the-art general countermeasures for cloud security attacks including intrusion detection systems, autonomous systems, and federated identity management systems. We also highlight the shortcomings of these systems that include the high communication and computation overhead and the detection efficiency and coverage.

Conflicts of Interest

The authors declare no conflict of interest.

References


58. Top Threats to Cloud Computing VI.0; Cloud Security Alliance: March 2010.


© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).