A Security Framework to Protect Data in Cloud Storage

Farashazillah Yahya1, Victor Chang2, Robert John Walters3 and Gary Brian Wills3

1Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia
2International Business School Suzhou, Xi’an Jiaotong-Liverpool University, Suzhou, China
3Electronics and Computer Science, University of Southampton, Southampton, U.K.

Abstract: With the success and widespread adoption of Cloud Computing Cloud storage has become the storage option of choice for many computer users wishing to keep their data online. This paper presents a framework to explore and evaluate security threats to data held in Cloud Storage. The Cloud Storage Security Framework (CSSF) has been developed both from consideration of established good practice as described in existing literature and the opinions of cloud storage managers and experts using a questionnaire and separate interviews. The purpose of the framework is to support researchers and managers of Cloud storage to understand the nine identified factors of security in Cloud storage and how to ensure security measures are successful. CSSF can also integrate with another framework to produce a greater impact and strengthens its research contributions.

Keywords: Availability, Authenticity, Cloud Storage, Confidentiality, Integrity, Non Repudiation, Reliability, Security Framework.

1 INTRODUCTION

With recent improvements to the availability, reliability and bandwidth of internet connections, Cloud computing has become a viable and highly attractive alternative to conventional computing for both individuals and business. A principal advantage is that using Cloud computing permits rapid adjustment of computing capabilities according to circumstances. This enables users to pay only for the facilities they need whilst retaining flexibility to buy extra resources when required. These advantages apply equally to computing processing power and data storage and the number of service providers offering both has increased significantly in recent years. It is estimated that in 2016 a typical Cloud storage user has at least 3.3 terabytes of data stored. There is also a significant increase in the use of commercial service providers; Dropbox, Box and Google Drive amongst others for personal storage (Gartner 2012). However, along with the cost and accessibility advantages, using Cloud based data storage services brings with it concerns about safety and security of valuable and potentially sensitive data (Zissis and Lekkas, 2012). There are regular reports of security breaches of poorly secured Cloud storage whereby outsiders have been able to discover passwords or exploit insecure APIs to gain access to unencrypted data (CSA, 2013a; GTISC and GTRI, 2013, Shaikh and Haider, 2011). There is significant variety in the techniques used from simple methods such as brute force discovery of passwords to more sophisticated approaches to read unencrypted files by compromising insecure service provider’s APIs or the Cloud service itself.

Efforts are being made by the service providers to address the problem, such as two-factor authentication and file encryption which make accessing data more difficult for outsiders. Unfortunately, these approaches can be unpopular with service providers and authorised users since they usually compromise usability and performance of Cloud storage services in what has become a highly competitive market (Honan, 2012; Zarandoon et al., 2012; Zhund and Chen, 2012; Zhao and Yue, 2014).

Whilst it is recognised that security is an issue in the provision of Cloud computing services, including the establishment of security frameworks and industry standards and best practices, previous work has addressed the security of generic Cloud services. There has been very little work directed at understanding issues and challenges specific to the provision of Cloud storage services.
2 CLOUD STORAGE

Cloud computing in general may be considered to be a style of computing which is delivered by internet technologies for users to access applications and other resources (Weiss 2007; Mell and Grance 2009; Vanquero et al. 2009). Cloud storage is a term used where the service provided in the Cloud is retention of data. Commercial Cloud storage provided by public Cloud service providers may also be referred to as Utility storage (Wu et al., 2010). Cloud storage provides the user with access to a flexible, scalable and provisioned virtual storage architecture, usually via an API (Ju et al., 2011). Services widely offered include storage protocols such as iSCSI (Satran et al. 2004), file storage (Miller, 2013) or databases or web servers. The data is stored on distributed servers which may be accessed from anywhere through the Internet. The service provider uses virtualization techniques to maintain, operate and manage the storage (Wu et al., 2010).

3 SECURITY GOALS

The adoption of Cloud storage can result in improvements to security of the data held, particularly where the storage service used is provided by a commercial Cloud service provider since, whilst management of the data is important to the owner, this activity will not normally be the owner’s principal activity. In contrast, managing data is central to the activities of a commercial Cloud storage service provider and it is consequently to be expected that such a provider will implement stronger and more recent security technologies than the data owner (Ryan, 2013). At the same time, adoption of Cloud storage brings some additional issues arising from data being stored on internet connected shared hardware.

The widely accepted security goals of confidentiality, integrity and availability, commonly referred to as CIA have been extended to include non repudiation, authenticity and reliability (ISECT, 2014a; CSA, 2013b; ISO/IEC, 2016). The implementation of a secure service requires policies, procedures and control in addition to technical measures to keep user data safe (Brock and Goscinski, 2010; Firesmith, 2004; Takabi et al., 2010; Zissis and Lekkas, 2012; Mapp et al., 2014) to meet the following goals:

• Ensuring confidentiality. Data must be handled correctly to prevent unauthorised exposure and ensure only those intended are able to gain access to the data through the application of access, authentication and authorisation controls to ensure access to data is only permitted to verified users with the necessary permissions (Vrable et al., 2012; Mapp et al., 2014; El Booz et al., 2016).

• Integrity checks. It is important that when users retrieve data which has been stored, they receive their data back unchanged from when it was stored. It is therefore necessary for storage service providers to implement integrity tests to ensure they are able to detect when data have been altered. A common approach is to use a hash or checksum from the contents of units of data. Generation of these checksums is often combined with the encryption of data where this is used in connection with ensuring confidentiality (Bowers et al., 2009; Yao et al., 2010).

• Maintaining availability. Although users need assurance that their data is safely stored and protected from unauthorised access, they also need timely access to their data. Hence, it is important for the service provider to ensure (server side) availability by ensuring the service is protected from both physical failures such as power failures and network disruption, and logical issues such as denial of service attacks as a result of malicious attacks. Storage providers have a variety of measures available to ensure continued availability, including mirroring of data in multiple clouds (Firesmith, 2004; Mapp et al., 2014; Takabi et al. 2010).

• Non repudiation of data. Attribution or provenance of data enables owners and other users to be confident that data cannot be disputed, including preventing a recipient from denying data have been received and is necessary in transactional interactions (Firesmith, 2004).

• Preserving authenticity. Data authenticity is concerned with its original creation by its author and ensuring that it has not suffered subsequent alteration. For example, when a document includes a digital signature, anyone using the document can use the signature to establish that the content and form of the document they have is unchanged from that created by the original writer (Brock and Goscinski, 2010; Zissis and Lekkas, 2012).

• Reliability of service provider. Service provider reliability is related to maintenance of availability but is also concerned with wider issues of the ability of the service provider to provide the offered (or intended) service consistently. In addition to

308
provision of satisfactory availability, the service provider needs to have proper procedures and mechanisms such as logging, monitoring and version controls to ensure smooth running of their services and contingencies to handle exceptional events.

There are International and Industry Standards and Best Practice guidelines which have been widely adopted by major Cloud service providers including Amazon, Oracle, RedHat and Salesforce (CSA, 2013; NIST, 2013a; ISECT, 2014b; ISO/IEC, 2016; ENISA, 2009; CPNI, 2014a; ASD, 2014b). A summary of these is given in Table 1.

Table 1: Summary of International and Industry Security Standards, Guidelines and Best Practice.

<table>
<thead>
<tr>
<th>Goals</th>
<th>CSA</th>
<th>NIST</th>
<th>ISECT</th>
<th>ENISA</th>
<th>CPNI</th>
<th>ASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>√</td>
<td></td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Av</td>
<td>√</td>
<td>√</td>
<td></td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>N</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At</td>
<td></td>
<td>√</td>
<td>√</td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>R</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

* C – Confidentiality, I – Integrity, Av – Availability, N – Non-repudiation, At – Authenticity, R – Reliability

4 CLASSIFICATION OF THREATS

In the context of computer system security, a vulnerability is a weakness of the system which may be exploited to achieve unauthorised access or other harm to the system. Alone a vulnerability is not a problem so long as it is undetected or remains unknown to potential attackers. However, vulnerabilities need to be addressed as, once potential attackers are able to identify them and discover how to exploit them, possibly in combination with others, they become a threat. Vulnerabilities can include all weaknesses of system which permit disruption of the system by misleading users, phishing, triggering incorrect system behaviour by introduction of false information or denial of service attacks as well as more “traditional” weaknesses such as weak or ineffective passwords, session hijacking and interception or alteration of data in transit (Chang and Ramachandran, 2016; Wang et al., 2010; Sabahi, 2011; CSA, 2013a). A threat model enables designers to estimate risks from attackers. Recognised threat analysis techniques include DREAD and STRIDE (Swiderski and Snyder, 2004) but these only provide industry best practice or standards which generally need adjustment or enhancement to fit particular systems (Myagmar et al., 2005).

For this study, threats were modelled using a three step process (Myagmar, 2005), characterising the system, identifying system assets and identifying system concerns as described in the following sections.

4.1 Characterising the System

A cloud storage system has three participants: the users, the clouds and the storage provider(s) (Gruschka and Jensen, 2010). Interactions between these participants take place at the two interfaces between them. It involves with three steps as follows. Firstly, users are authenticated and get access to their respective Cloud storage. Secondly, they can send commands and messages to Cloud Storage and execute commands, such as uploading files. Thirdly, service provider can receive commands and messages from users and store files. These three steps can be done seamlessly within seconds. Examples of interactions include a user requesting a service or a service requesting additional storage from the service provider. Attacks may be analysed by examining the interactions between the elements of this model.

4.2 Identifying System Assets

Each of the system participants offers a particular interface to the others which may all be subject to attacks as follows, see Figure 1.

- **Service to user.** Provided by the server, this layer is subject to different types of attacks, including account hijacking, SQL injection and privilege escalation (CSA, 2013b).
- **User to service.** Threats to this interface involve user programmes requesting services, including browser malware or phishing based attacks (Marlinspike, 2009; CSA, 2013b).
Cloud to service. Although the separation between the service instance and the cloud provider may not always be clear, threats to this interface include exhaustion attacks directed at triggering the provider to provide excess resources or attacks on the system hypervisor (CSA, 2013b).

Service to cloud. All kinds of attack against Cloud systems by services, including privacy related attacks and tampering with data.

User to cloud and Cloud to user. These include all kinds of attack which may be made by the user against the cloud provider, which may include inducing users to delete data, phishing based attacks or presentation of falsified usage accounts.

4.3 Identifying System Concerns

Many concerns may arise from within or outside the system from authorised users, others who masquerade as genuine users or have managed to bypass security mechanisms (Myagmar et al., 2005). Others arise from innocent errors or external factors such as severe weather or other natural disasters. The goal is to establish risks to the system based on the information gathered.

Each threat is the damage an adversary may inflict on the system (Swiderski and Snyder, 2004) which may be described as the capability of the adversary. A reasonable start to threat modelling is known system concerns and vulnerabilities. Each of these then needs to be considered carefully to in the context of the quality and suitability to the system. There are fourteen security concerns and factors worth to be investigated (Gruschka and Jensen, 2010; Shaikh and Haider, 2011; CSA, 2013a; GTISC and GTRI, 2013). They are presented in Table 2 and Table 3, so that threat classification can be compared in details.

Each security factor may be mapped to security concerns showing the potential effects of each on the system. This shows security factors and concerns using the CSA Control Matrix (CSA, 2013 a, 2013 b), to which reliability has been added as an additional factor.

In this work, STRIDE has been used it matches the results of the threat identification process. In general, threats fall into six types or classes depending on their effect (Swiderski and Snyder, 2004). They are:

1. Spoofing – Using false credentials to gain access to protected assets.
2. Tampering – Mounting an attack by changing data.
3. Repudiation – Denying having taken actions in circumstances where the target is unable to prove otherwise.
4. Information Disclosure – Unauthorised release of data.
5. Denial of Service – Actions which reduce the ability of legitimate users to access resources.
6. Elevation of Privilege – Exploitation by users able to gain access to system features reserved for others.

In defining a threat model, the modeller defines attacks and prioritises them. Using a risk assessment, each threat is given a priority and mapped to a mitigation mechanism. With the concerns identified, the security factors have been identified as shown in Table 2 and Table 3.

<table>
<thead>
<tr>
<th>Threat Classification</th>
<th>STRIDE Threat Modelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Breach</td>
<td>S √</td>
</tr>
<tr>
<td>Data Leakage and loss</td>
<td>T √</td>
</tr>
<tr>
<td>Insecure APIs</td>
<td>R √</td>
</tr>
<tr>
<td>Account hijacking</td>
<td>I √</td>
</tr>
<tr>
<td>Denial of Service</td>
<td>D √</td>
</tr>
<tr>
<td>Malicious insiders</td>
<td>E √</td>
</tr>
<tr>
<td>Inadequate cloud planning</td>
<td>S √</td>
</tr>
<tr>
<td>Cloud related malware</td>
<td>T √</td>
</tr>
<tr>
<td>Closure of cloud service</td>
<td>R √</td>
</tr>
<tr>
<td>Shared technology vulnerabilities</td>
<td>I √</td>
</tr>
<tr>
<td>Insufficient due diligence</td>
<td>D √</td>
</tr>
</tbody>
</table>

Table 3: Threat classification with security factors.

<table>
<thead>
<tr>
<th>Threat Classification</th>
<th>Security Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Breach</td>
<td>C I Av N At R</td>
</tr>
<tr>
<td>Data Leakage and loss</td>
<td></td>
</tr>
<tr>
<td>Insecure APIs</td>
<td>√ √</td>
</tr>
<tr>
<td>Account hijacking</td>
<td>√ √ √</td>
</tr>
<tr>
<td>Denial of Service</td>
<td></td>
</tr>
<tr>
<td>Abuse of cloud planning</td>
<td>√</td>
</tr>
<tr>
<td>Inadequate cloud planning</td>
<td></td>
</tr>
<tr>
<td>Cloud related malware</td>
<td>√</td>
</tr>
<tr>
<td>Closure of cloud service</td>
<td>√ √</td>
</tr>
<tr>
<td>Natural disaster</td>
<td>√</td>
</tr>
<tr>
<td>Hardware failure</td>
<td></td>
</tr>
<tr>
<td>Shared technology vulnerabilities</td>
<td>√</td>
</tr>
</tbody>
</table>

* C – Confidentiality, I – Integrity, Av – Availability, N – Non-repudiation, At – Authenticity, R – Reliability

5 (DEVELOPMENT OF) CLOUD STORAGE SECURITY FRAMEWORK (CSSF)

This section proposes a security framework for Cloud storage. As already indicated, there is little existing work concerned specifically with factors affecting security or to investigate appropriate frameworks for Cloud storage. This section proposes an appropriate framework together which the process used in its development. An outline of the process is given in Figure 2.

5.1 Framework Development Process

An initial review revealed that, despite the great growth in use of Cloud based storage, there is very little existing work concerned specifically with Cloud Storage security. Hence the development of the framework proposed here which seeks to identify security factors and concerns in Cloud storage.

5.2 Listing Unique Factors

Factors were identified from a review of existing work on security frameworks for Cloud computing in general and Cloud storage in particular as described above. The factors identified from Cloud Security research are presented as follows (Catteddu and Hogben, 2009; CSA, 2013b; NIST, 2013a; ASD, 2014b; ISO/IEC, 2016). Firstly, confidentiality, it refers to whether data is unavailable to the unauthorised entities. Secondly, integrity, it measures whether data is complete and accurate. Thirdly, availability, it is focused on whether data can be accessed and used on demand for the authorised entity. Fourthly, non-repudiation, it deals with the ability to validate the occurrence of a claimed event and its originating entities. Fifthly, authenticity, it means if data is original as it claims. Lastly, reliability, it refers to the ability to demonstrate consistent results. To accommodate these six factors, Table 4 lists a further seven factors identified from secure Cloud Storage research. In all, a total of 13 goal driven factors and 20 potential characteristics were identified.

Figure 2: Development of the Cloud Storage Security Framework (CSSF).

5.3 Grouping Factors with Similar Meanings

Following completion of Section 5.3 and Table 4, each of the factors revealed were considered for inclusion in the final framework. This involved consideration of whether a factor is unique to its area or whether it could be mapped to a factor, or a combination of factors in the other area as well as consideration of the extent to which the factor is relevant and important in the context of Cloud
storage. With this comparison completed, there remain seven distinct factors which are included in the final framework.

5.4 The Proposed Cloud Storage Security Framework

The final Cloud Security Storage Framework (CSSF) is presented as follows. Firstly, Cloud Storage Security, since its policies and security procedures are important. Secondly, confidentiality, it deals with users’ identification and authorisation for accessing data. Thirdly, integrity, it specialises in accurate ownership and encryption of data. Fourthly, availability, it deals with up-to-date available and accessibility of data. Fifthly, non-repudiation, it is focused on accurate time stamping of accessed data with right user signature. Sixthly, authenticity, it checks data upon authentication and synchronised data in the storage. Lastly, reliability, it presents consistency and validity of Cloud service.

CSSF framework can also work together with other framework such as Cloud Computing Adoption Framework (CCAF) to ensure all the services are valid and data can be protected in real-time. Chang et al. (2016) present CCAF, which designs and develop Cloud services. The security is based on multi-layered security as follows. Firstly, firewall and authentication to execute all the seven factors in CSSF. The second layer is based on the intrusion detection and prevention system to minimise the impacts of security. The third layer is on encryption and its main purpose is to identify any comprised and disguised data mixed with real and validated data. Petabytes of data were in the data centre. Real-time largescale tests were conducted. Results showed that CCAF can protect more than 99% of data under the intensive penetration tests. In this case, CSSF provides guidelines and important foundation to build up the first layer of defence to make Cloud Storage more robust and resilient. Case studies from organisations that adopt our recommendations can be presented to illustrate the positive impacts and research contributions offered by the blended approach. Therefore, getting two frameworks to work together can provide better security services for users.

<table>
<thead>
<tr>
<th>Security Factors</th>
<th>Characteristics/Elements</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Storage Security</td>
<td>Security policy, security procedure</td>
<td>Firesmith, 2004; Brock and Goscinski, 2010; Takabi et al., 2010; Zissis and Lekkas, 2012; Mapp et al., 2014</td>
</tr>
<tr>
<td>Ensuring Confidentiality</td>
<td>Identification of Cloud storage user, authorisation to access data</td>
<td>Bessani et al., 2011; Kamara et al., 2011; Popa et al., 2011; Stefanov and Dijik, 2012; Zhou et al., 2013; Yao et al., 2013; Mapp et al., 2014; Tawalbeh et al., 2015; Vu ât al., 2015; El-Booz et al., 2016</td>
</tr>
<tr>
<td>Integrity Checks on Remote Data</td>
<td>Accurate ownership of data, encryption of data</td>
<td>AlZain et al., 2011; Bessani et al., 2011; Kamara et al., 2011; Mahajan et al., 2011; Napal et al., 2011; Vrable et al., 2012; Mu et al., 2012; Stefanov and Dijik, 2012; Wang et al., 2013; Ahou et al., 2013; Tawalbeh et al., 2015; El-Booz et al., 2016</td>
</tr>
<tr>
<td>Maintaining Availability</td>
<td>Access to data, up-to-data available data</td>
<td>AlZain, et al., 2011; Bessani et al., 2011; Mahajan et al., 2011; Popa et al., 2011; Mu et al., 2012; Stefanov and Dijik, 2012; Mapp et al., 2014; Vu et al., 2015</td>
</tr>
<tr>
<td>Guaranteeing Non-repudiation to Data</td>
<td>Accurate time-stamping of accessed data, assurance with user signature</td>
<td>Wang et al., 2013; Tawalbeh et al., 2015; El-Booz et al., 2016.</td>
</tr>
<tr>
<td>Preserving Authenticity</td>
<td>Verified data based on authentication, synchronised data in storage</td>
<td>Yao et al., 2013; Mapp et al., 2014; Vu et al., 2015; El-Booz et al., 2016.</td>
</tr>
<tr>
<td>Reliability of Service Provider</td>
<td>Consistency of Cloud service, valid service</td>
<td>Bessani et al., 2011; Mahajan et al., 2011; Stefanov and Dijik, 2012; Mapp et al., 2015, Vu et al., 2015.</td>
</tr>
</tbody>
</table>
6 CONCLUSIONS

With recent improvements to the availability, reliability and bandwidth of internet connections, Cloud computing has become a viable and highly attractive alternative to conventional computing for both individuals and business enabling users to pay only for the facilities they need whilst retaining flexibility to buy extra resources when required. However, along with the cost and accessibility advantages, using Cloud based data storage services brings with it concerns about safety and security of valuable and potentially sensitive data (Zissis and Lekkas, 2012) and there are regular reports of security breaches of poorly secured Cloud. Efforts are being made by the service providers to address the problem. However, whilst there exist frameworks and industry standards and best practices for generic Cloud services, there has been very little work directed at understanding issues and challenges specific to the provision of Cloud storage services.

This paper has described work to identify factors relevant to security of Cloud storage by considering existing work on security, Cloud security in general and Cloud storage in particular. Having analysed the resulting factors and items for distinct elements, common themes and duplication, the results presented as the framework given in Section 5. The joint work with CCAF can produce a greater impact and more validated research contributions, so that users and data can be protected in real-time. Future work may also include integration with other frameworks to demonstrate that our work can provide not only recommendation and guidelines, but also the real examples and case studies of building secure services.

REFERENCES


