Complying with Security Requirements in Cloud Storage Systems

Rodrigo Roman[1], Miguel Rodel Felipe[2*], Phua Eu Gene[2], and Jianying Zhou[1]
[1] Institute for Infocomm Research, Singapore.
[2] Data Storage Institute, Singapore.

* Corresponding author. Email: rodel_fm@dsi.a-star.edu.sg
Manuscript submitted October 7, 2014; accepted March 10, 2015.
doi: 10.17706/jcp.11.3.201-206

Abstract: One of the necessary steps to ensure the security of cloud storage systems is to adequately protect the infrastructure itself – the hardware and software that implements the storage services. Starting from an analysis of the security requirements that affect these storage systems, this paper studies the different strategies and approaches that are currently used to fulfill such requirements. The goal of this paper is twofold. Firstly, we aim to analyze the security components that should be used to **provide** a basic level of protection to storage systems, examining the actual technologies that are used to construct them. Secondly, we aim to identify gaps in the provisioning of security services, highlighting any areas that need of further research.

Keywords: Cloud computing, cloud storage, security.

1. Introduction

Cloud storage systems are one of the subsets of the cloud computing model, focused on the provisioning of on-demand storage services. Such services can be provided by external companies (public clouds), or even located within the enterprise infrastructure (private clouds). While these services provide several benefits (e.g. ubiquitous access, resource scaling), there are many challenges that must be considered. Some of those challenges are related to the security of the cloud infrastructure, i.e. the actual platform that implements and provides the cloud storage services.

The main goal of this paper is to study the security of these cloud storage infrastructures, identifying the security services that should be integrated into existing platforms and the open issues that need further research. We will identify a complete set of security requirements that influence over the design of these infrastructures (Section 2). We will then enumerate the most important security components that should be integrated into storage platforms (Section 3). Afterwards, we will evaluate the security of existing cloud storage infrastructures (Section 4). We will also highlight various open issues related to those components that need of further research (Section 5).

2. Security Requirements in Cloud Storage Systems

One of the first steps in the design of any secure system is to discover the security requirements. These requirements will help not only to identify which security components have to be developed (e.g. authentication), but also to describe various requisites that affect the design of those components (e.g. performance).

In order to obtain the security requirements, we have followed a methodology known as the Asset Table [1]. This methodology firstly identifies the different assets that belong to the infrastructure (in our case: devices, networks, protocols and services, users, information), and secondly creates a table describing how those assets should be protected and how they should be attacked by means of use cases.

Table 1. Security Requirements in Cloud Storage Systems

Requirements Category	Specific Requirements
General	Protocol Interoperability, Scalability, Performance / Availability, Extensibility, Updating
Management	Accountability (logging protection), Policy Management (usability), Simple Administration (constraints), Auditing (intrusion detection systems)
Data	Secure Storage (privacy), Migration, Redundancy, Retirement (zeroing), Authorization (minimal permissions, flexibility), Proof of Operations
Credential	Creation (randomness), Storage, Rekeying, Maintenance
Network	CIAA (Confidentiality, Integrity, Authentication, Availability), Robustness
User	Authentication (flexibility), Usability, Trust in the platform (perception)
External (Perimetral)	Perimetral Security (software, hardware, physical)

For the development of these use cases, we have used various threat modeling approaches, such as STRIDE by Microsoft [2]. As for the attacker model, we assumed not only that the cloud infrastructure is honest (albeit curious), but also that both external and internal entities can try to attack the infrastructure. Finally, the table (which is not included due to space restrictions) is used to extract the security requirements, as these requirements must cover all the use cases defined in the table. Note that one of the benefits of this methodology is that it can be easily adapted to perform a risk assessment (cf. Section 3), since such assessment largely depends on the definition of the 'Attack' use cases.

There are seven categories of requirements, and every particular requirement can also have various related sub-requirements. The final list of security requirements is introduced in Table 1.

3. Risk Analysis and Major Security Components

The list of requirements presented in Section 2 can help designers and architects to be aware of the main security issues that can affect a cloud storage infrastructure. To identify the most important security requirements, we can perform a risk management process, deriving the risk from the likelihood and the impact (cf. [3]). Firstly, we used the asset table (cf. Section 2) to analyze the likelihood (i.e. probability of the anomalous event to occur) and impact (i.e. effect on the system and its services) of every attack use case. Secondly, after assigning a score to every factor (from 'very low' to 'high'), we calculated the severity of the attacks following a risk combination table (cf. [4]). These values, once combined, provided a list of the most dangerous attacks – and in turn the most important security requirements.

As a result of the previous analyses, we can provide an ordered list of the most important security components that should be included in the design of cloud storage systems. The list is shown below:

1) **Logging System and Auditing System**. These two components can help administrators to understand the actual (and past) state of the system.
2) **User Authentication and User Authorization.** As a cloud storage platform deals with user data, it is essential that only those who are authorized can access it.
3) **Device Authentication/Authorization and Secure Communications.** Not only the elements of the system must prove that they belong to the same infrastructure, but also all communications must be protected as well.
4) **Data Protection.** The system should protect the data at all times, even when it is stored in the storage nodes.

5) **Credentials Storage.** The credentials stored within all devices must be managed securely.
6) **Extended Services.** The storage system should provide an open interface where diverse mechanisms (e.g. proof of storage) can be integrated.
7) **Policy Administration.** The management of all the policies of the storage system, as well as other management tasks such as user administration, should be simple and usable.

4. Analysis of Existing Cloud Storage Platforms

The major security components described in the previous section can be used as a foundation for analyzing the security of existing enterprise-centric cloud storage platforms, such as the open source platforms Walrus (Eucalyptus cloud), Swift (OpenStack cloud) and Cumulus (Nimbus cloud). Note that closed source platforms like Amazon S3 (Amazon WS cloud) cannot be studied at the same level of detail. A summary of the instantiations of the security components in every platform is shown in Table 1, and will be explained in detail in the following paragraphs.

Table 2. Security Components and Existing Cloud Storage Platforms

	Logging/ Auditing	User Auth[2]	Device Auth[2] Secure Comm	Data Protection	Credential Storage	Extended Services	Policy Admin
Eucalyptus	Logs Health System	ID: Keys, X509… Resources: ACL	Cloud: WS Storage: None	None	Stored in folder	No support	Command line Web Platform
OpenStack	Logs Health System	ID: Keys, X509… Resources: ACL	Cloud: Keystone Storage: None	None	None	Authentication/Authorization	Command line Web Platform
Nimbus	Logs	ID: Keys, X509… Resources: ACL	Cloud: SSH Storage: None	None	Stored in folder	"Plugin" support	Command line
Amazon S3	Logs, Mgmt. & Health System	ID: Keys, X509… Resources: ACL		Server side Client side		No support	Command line Web Platform

Some of the components (logging / auditing and policy administration) are partially supported: logs are stored in known locations, and diverse tools (health monitoring subsystems, administration interfaces) are available. Still, the usability of these mechanisms should be improved, either by implementing extensions or by using third-party components. Besides, there are specific improvements in these areas that should be considered. For example, the outputs of the auditing subsystem should be connected to internal intrusion detection systems, so as to detect anomalous situations inside the cloud storage system.

Other components (user auth2 – authentication and authorization) provide a satisfactory level of security. Users can make use of different authentication mechanisms (e.g. passwords, certificates) to access the data. Also, most platforms provide an Access Control List (ACL) mechanism based on the Amazon S3 specification, where owners can assign policies to specific users, groups and buckets (i.e. data containers). Note, however, that there is still room for improvement in these components. For example, only OpenStack provides a simple interface to extend the authentication and authorization mechanisms without recompiling the whole platform. Also, as ACLs are in some cases inadequate to fully capture the complexity of enterprise environments, other approaches such as Role-based Access Control (RBAC) might be integrated [5].

Finally, there are various components that are not supported in most platforms. For instance, existing device authentication/authorization and secure communications components are designed to protect the communications between cloud entities, but they are not used to protect the communications inside the storage subsystem. Also, there is no explicit support for storing the credentials in secure and tamper-resistant containers. As for data protection, no platform provides mechanisms that implement data-at-rest encryption. Lastly, with the exception of the Nimbus platform and (partially) the OpenStack platform, it is not possible to implement specific plugins that provide additional extended services such as proof of storage services.

We should note here that the Amazon S3 platform actually provides or extends some of the previously

mentioned services. For example, Amazon provides support for transparent server-side data encryption, although the actual location of the encryption keys and their physical and logical security is not known. There is also explicit support for client-side encryption through specific APIs. As for logging and auditing, while the actual internal mechanisms are not known, users can be able to access detailed server logs that indicate which files are being accessed and who is accessing them

5. Applicability of Existing Protocols and Research Solutions

As most existing cloud storage platforms do not provide complete implementations of all security components, it is necessary to check what solutions (e.g. academic research, industrial standards) could be used to improve this situation.

When creating *secure communication channels* inside the storage infrastructure, it is necessary to consider how the underlying transport layer is implemented. For example, the distributed file system standard NFSv4.1 can use Remote Direct Memory Access (RDMA)-based transport protocols (e.g. RoCE, InfiniBand [6]) or non-RDMA-based transport protocols (e.g. TCP, SCTP). If RDMA is used, security protocols like TLS and IPsec cannot be integrated, as solutions such as InfiniBand or RoCE (RDMA over Converged Ethernet) implement their own protocol stack. Note, however, that RoCE provides support for Ethernet at the network layer, so layer 2-based technologies such as IEEE 802.1ae might be applied.

Regarding *policy administration*, while most storage platforms provide mechanisms that perform this task (e.g. access permissions management, users / groups / roles / domains management), various factors highlighted by the research community should be taken into account. For example, securing the extremely vulnerable administration tasks done over web interface [7]. Data security developers should take note about usability [8] and feedback mechanisms [9]. Policies can be enforced at various layers of the architectures of cloud storage and cloud computing systems [10].

As for existing *logging systems*, their usability is normally low, as they are in most cases simple text files stored within the machines' file systems. In order to improve the accessibility of this information, not only these logs can be sent to the administration interfaces, but also it is possible to aggregate them into existing distributed monitoring systems (e.g. Nagios and Ganglia).

In the integration of *data protection* mechanisms, the implementation of data-at-rest encryption can be pushed onto the user (users send the data already encrypted) or onto the storage infrastructure (data is encrypted and decrypted in the storage nodes). If the cloud storage infrastructure is in charge of protecting the data, there are various issues to consider. First, as storage nodes can perform thousands of transactions per second, performance becomes a core requirement, thus it should be necessary to apply diverse optimizations such as HW acceleration (e.g. Intel's AES extensions, dedicated HW acceleration cards). Second, all credentials and keys should be closely guarded, preferably on trusted software modules and appliances [11]. On the other hand, if users apply data protection mechanisms, there are other issues that must be addressed, such as the management of the keys. If users retain the keys, it is possible to lose all the data if these keys are lost. Also, in order to share the data, users must entrust these keys to other users. There are some basic solutions that aim to manage these issues, such as storing the keys in Trusted Third Parties that are independent from the original cloud storage [12].

However, some researchers have devised more advanced procedures, which effectively intermix data protection with *access control*. The core concept of this idea is to integrate the actual authorization policies into the data itself. In other words, any user can access the data, but only those who are authorized can actually decrypt it. This vision can be accomplished in different ways: from embedding policies into the session keys by making use of a key derivation structures [13] to integrating the policies into the actual ciphertext by using ciphertext-policy attribute-based encryption (CP-ABE) and hierarchical identity-based

encryption (HIBE) [14]. Note that in some of these approaches a trusted third party is needed. Also, the computational cost of these mechanisms is high in comparison with more traditional encryption mechanisms.

Finally, for *user authentication*, while some researchers have provided additional authentication mechanisms, such as USIM-based authentication [15], other researchers have focused on the integration of cloud services with federated identity systems. The technologies that are used to implement this idea are numerous: from SAML assertions [16] to Shibboleth-enabled applications [17]. Precisely, this diversity is the source of syntactic (differences between protocols) and semantic (different names and meanings for identity attributes) problems that need to be carefully considered [18]. Moreover, many technologies such as OpenID and SAML provide their services using HTTP as a transport mechanism, which might not be suitable for certain deployments.

6. Conclusions

In this paper, we have studied the security of cloud storage systems from the point of view of their security requirements and components. Most of the existing cloud storage platforms do not provide a complete implementation of all security components. Moreover, there are various research issues that must be carefully considered. Some research areas (policy administration, logging, and auditing) are slightly underdeveloped in comparison with other areas (entity authentication, authorization, and data protection). Also, there are various architectural issues that are not explicitly studied (e.g. location of security services, modular and extensible components).

References

[1] Jaatun, M., & Tøndel, I. (2008). Covering your assets in software engineering. *Proceedings of the 3rd International Conference on Availability, Reliability and Security (ARES)* (pp. 1172-1179).

[2] Microsoft Corporation. *The STRIDE Threat Model*. Retrieved July 2012, from http://msdn.microsoft.com/library/ms954176.aspx

[3] Ahmed, A., Kayis, B., & Amornsawadwatana, S. (2007). A review of techniques for risk management in projects. *Benchmarking: An International Journal*, *14(1)*.

[4] The Open Web Application Security Project (OWASP). *OWASP Risk Rating Methodology*. Retrieved July 2012, from https://www.owasp.org/index.php/OWASP_Risk_Rating_Methodology

[5] Takabi, H. (2010, Nov.-Dec.). Security and privacy challenges in cloud computing environments. *IEEE Security & Privacy*, *8(6)*.

[6] Shah, H., Pinkerton, J., Recio, R., & Culley, P. (2007, Oct.). Direct data placement over reliable transports. *RFC 5041*.

[7] Johnston, J., Eloff, J. H. P., & Labuschagne, L. (2003, Dec.). Security and human computer interfaces. *Computers & Security*, *22(8),* 675-684.

[8] Rode, J., Johansson, C., DiGioia, P., Silva, F. R., Nies, K., Nguyen, D. H., Ren, J., Dourish, P., & Redmiles, D. (2006, Jul.). Seeing further: extending visualization as a basis for usable security. *Proceedings of 2nd Symposium on Usable Privacy and Security* (pp. 145-155).

[9] Waller, A., Sandy, I., Power, E., Aivaloglou, E., Skianis, C., Muñoz, A., & Maña, A. (2011, Jun.). Policy based management for security in cloud computing. *Proceedings of the 1st International Workshop on Security & Trust for Applications in Virtualised Environments* (pp. 130-137).

[10] Ficco, M., & Rak, M. (2012, Jul.). Intrusion tolerance in cloud applications: The mosaic approach. *Proceedings of the 6th International Conference on Complex, Intelligent and Software Intensive Systems* (pp. 170-176).

[11] Yao, J., Chen, S., Nepal, S., Levy, D., & Zic, J. (2010, May). Trust store: Making Amazon s3 trustworthy with services composition. *Proceedings of the 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing* (pp. 600-605).

[12] Vimercati, De C. di S., Foresti, S., Jajodia, S., & Livraga, G. (2012, Jul.). Enforcing subscription-based authorization policies in cloud scenarios. *Proceedings of the 26th Annual WG 11.3 Conference on Data and Applications Security and Privacy* (pp. 314-329). LNCS 7371, Springer Verlag.

[13] Wang, G., Liu, Q., & Wu, J. (2011, Aug.). Achieving fine-grained access control for secure data sharing on cloud servers. *Concurrency and Computation: Practice and Experience*, *23(12),* 1443-1464.

[14] Bernal, B. J., Marin, P. J. M., Alcaraz, C. J. M., Garcia, C. F. J., Martinez, P. G., & Gomez, S. A. F. Semantic-aware multi-tenancy authorization system for cloud architectures. *Future Generation Computer Systems*.

[15] Celesti, A., Tusa, F., Villari, M., & Puliafito, A. (2010, Jun.). Security and cloud computing: intercloud identity management infrastructure. *Proceedings of 19th IEEE International Workshop on Enabling Technologies: Infrastructures for Collaborative Enterprises* (pp. 263-265).

[16] Rieger, S., Richter, H., & Xiang, Y. (2011, Sep.). Introducing federated webdav access to cloud storage providers. *Proceedings of 2nd International Conference on Cloud Computing, GRIDs, and Virtualization* (pp. 46-51).

[17] Nuñez, D., Agudo, I., Drogkaris, P., & Gritzalis, S. (2011, Jun.). Identity management challenges for intercloud applications. *Proceedings of Secure and Trust Computing, Data Management, and Applications* (pp. 198-204).

[18] Ada P. R., Lorch, D., Molnar, J. R., Wang, H. J., & Zhuang, L. (2011, Jun.) Enabling security in cloud storage slas with cloudproof. *Proceedings of the USENIX Annual Technical Conference*.

Rodrigo Roman received his Ph.D. degree in computer science from Universidad de Málaga in 2008. His areas of research/interests include identifying security issues, applying security mechanisms, internet of things, wireless sensor networks, cloud storage services, SCADA systems, and digital homes.

Rodel Felipe Miguel hails from Laguna, Philippines. He received his bachelor of science degree in computer engineering (cum laude) from AMA Computer College, Makati City, Philippines.

He has more than 12 years of software engineering experience focused on embedded systems, device drivers, and network applications. He led teams from the Philippines and Singapore and has worked for industry leaders like Agilent Technologies and JDSU. He is currently a senior research engineer in Data Storage Institute — A*STAR, Singapore. His research interests include storage and network security.