Cloud Audits and Privacy Risks

Thomas Rübsamen and Christoph Reich

Cloud Research Lab,
Furtwangen University of Applied Science,
D-78120 Furtwangen, Germany
{thomas.ruebsamen,christoph.reich}@hs-furtwangen.de

Abstract. In cloud computing users are giving up control over resources such as storage. Lacking transparency of cloud services (e.g. data access and data lifecycle reports) is an important trust issue, that hinders a more wide-spread adoption of cloud computing. Giving the customer of cloud services more information about data usage, compliance test reports and accordance to best-practices make the cloud more transparent. Reporting about such verifications is the main objective of cloud audits and is performed by third party auditors (TPAs). However, public auditing by TPAs can introduce new privacy problems. In this paper, a survey of current cloud audit privacy problems is given and techniques are shown how they can be addressed. Also, requirements for a privacy-aware public audit system are discussed.

Keywords: Audit, Privacy, Cloud Computing, Transparency.

1 Introduction

Cloud computing as a paradigm is becoming more important for today's information technology. The shift from using traditional in-house datacenters towards on-demand provision of resources for delivering services has provided significantly more flexibility. However, a major aspect of adopting the cloud is giving up control of resources such as servers and network infrastructure. This also implies giving up control of data, which is stored in the cloud as well as services running in the cloud. The benefits provided are: increased flexibility, on-demand scalability and sometimes cost advantages.

Large companies like Amazon, Google and Microsoft have recognized the potential business opportunities and are providing well-established cloud services. Depending on the cloud service model, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), or Software as a Service (SaaS) [1] cloud customers have different levels of control over their resource stack for providing and using their services. For instance, a customer does have more control of their resources in a IaaS compared to a SaaS resource. Nevertheless some parts of the cloud infrastructure always remain hidden from the customer. This leads to a lack of transparency of cloud services, which is in turn the most important reason for a lack of trust in cloud computing. Highly visible security and privacy incidents of

R. Meersman et al. (Eds.): OTM 2013, LNCS 8185, pp. 403–413, 2013.

major cloud providers also make this problem worse, e.g., as reported by CNN: "Google fires engineer for privacy breach" [2].

The Cloud Security Alliance (CSA) launched a new initiative to encourage transparency of security practices within cloud providers, called Security, Trust & Assurance Registry (STAR) [3]. Providers are encouraged to provide self assessment reports using STAR.

Some cloud providers address the transparency problem by providing tools such as Amazon's dashboard [4], but to increase trust cloud providers need to be able to prove compliance with external and internal regulations. Cloud audits are one instrument to provide cloud customers with more insight on cloud service delivery. They increase trust in cloud providers by proving the provider's compliance to regulatory requirements (SAS70, HIPAA, PCI) and industry best-practices (ISO27001, BSI Grundschutz, ITIL) regarding the implementation of privacy controls. Cloud audits may be characterized from different perspectives. They can be performed internally by the cloud provider or externally by a trusted third-party auditor. Additionally, audits may also be performed by the customer (e.g., public audit), by providing him with the necessary interface to conduct them.

However, cloud audit capabilities themselves can be problematic in multi-tenant environments such as the cloud. By providing an auditor access to cloud infrastructures, the customer's data might be at risk. In case of an untrustworthy auditor (third-party or internal) sensitive information is potentially at risk. The Privacy Rights Clearinghouse published such an incident, where an untrustworthy auditor used his access privileges to sell customer's credit card information obtained from the Amsterdam Hospitality Group [5]. Such an incident may also happen in cloud environments. Information about other tenants might also be leaked, when a customer is given the capability to audit the provider's services himself. Additional protection mechanisms beyond contractual terms are needed in cloud audit scenarios. Another problem arises from making the cloud more transparent in terms of data locality and audits of policies, which require certain types of data to be stored at distinct locations. However, tracking of data in the cloud is not the focus of this paper.

In this paper, we focus on the privacy implications of providing publicly available audit capabilities to cloud customers. In Section 2 we define transparency, privacy and cloud audit concepts, followed by Section 3 where we provide a survey on privacy issues, which stem from public cloud audits as well as some ways to address these issues. After that, requirements for a privacy-aware public audit system are discussed in Section 4. We end this paper with a conclusion in Section 6.

2 Transparency, Privacy and Cloud Audit

The dilemma is to provide the highest levels of transparency while maintaining strong privacy and being auditable by a third party. In this section, we describe the concepts of transparency, privacy and cloud audit.

2.1 Transparency

Transparency in the sense of visibility is an important issue in cloud computing. According to the NIST definition of cloud computing one essential characteristic of cloud computing is measured service [1]. Resources are monitored, controlled and reported on continuously. This provides some degree of transparency for the user and provider. However, this information might not be enough for showing *what* happens in the cloud. Additionally, technical (e.g. cloud architecture, configurations etc.) and operational information (information about security processes, incidence management, etc.) needs to be provided to make cloud services transparent. Also, the different cloud service delivery scenarios (IaaS, PaaS, SaaS) need to be considered. In IaaS the cloud user usually needs less information from the cloud provider, due to having full control about virtual machines at the operating system level. In PaaS and SaaS however, additional information needs to be provided by the provider because of the administration shift away from the customer.

2.2 Privacy

Privacy can be looked at from different angles. From a consumer's perspective it comprises the protection and appropriate use (according to the customer's expectations) of the customer's personal data. From an organization's point of view, privacy includes the application of laws, policies, standards and processes by which personally identifiable information (PII) is managed [6]. These aspects have to be considered carefully in cloud computing scenarios, where PII is processed and transferred.

Confidentiality is a concept, which focuses on the protection of data from unauthorized access. Usually, cryptographic mechanisms are used to achieve confidentiality. Confidentiality can be linked to privacy in the sense of protecting PII by using encryption techniques. However, in current cloud computing services, data is usually not encrypted to enable processing of data by the cloud provider. Therefore, other ways have to be found to protect customer's privacy.

2.3 Cloud Audit

Common cloud computing scenarios usually include at least one cloud service provider and a customer in the service provision chain. The introduction of a third party auditor (TPA) offloads time-consuming and cost-intensive tasks, as well as required knowledge to conduct audits from the customer. The entity, which performs audits is called cloud auditor. Generally, audits are used to "verify conformance to standards through a review of objective evidence" [7]. In cloud computing this means cloud services are examined regarding their performance, privacy, and security controls, which shall assure the cloud customer, that appropriate measures are in place. To enable the TPA to fulfill his role, public auditability becomes a requirement. The cloud provider has to provide auditors

with interfaces, techniques and necessary documentation as well as any additionally needed information to conduct audits. The auditor can than leverage his expertise and conduct audits on behalf of the provider (for external audits) or the customer.

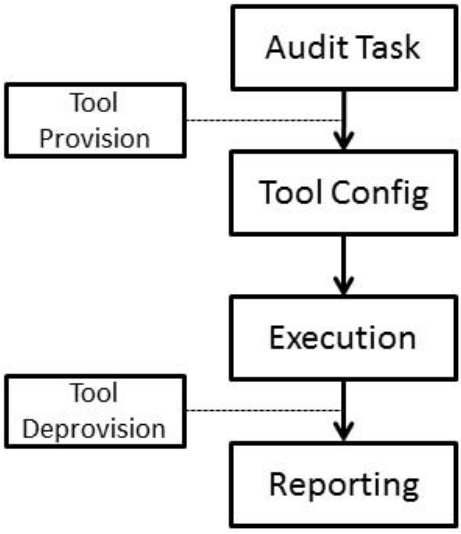

Fig. 1. Cloud Audit Process

Figure 1 depicts an automatic, tool-driven cloud audit process. Starting from a specific audit task, which describes a distinct objective, a tool is provisioned to perform the task (e.g., checking specific configuration files or parsing logs for certain events). The configuration of the tool specifies what has to be checked during execution. On the basis of the evidence collected by the tool, a report is generated and the tool is deprovisioned.

Auditors can be internal or external. External auditors are members of independent organizations specialized on performing audits. Internal auditors are usually contracted to the organization, whose entities are to be audited. They generally adhere to the same standards as external auditors with respect to performing an independent analysis.

3 Audit and Privacy

Cloud auditing needs to consider privacy of cloud customers. There can be several views on public cloud auditing scenarios, which will be shown in the following:

– Disclosure of customer's data
  Cloud customers store information in the cloud. Examples for this are cloud storage services, which are used to store files, database as a service, which

stores more structured data and data stored in SaaS services (e.g., customer relationship information).
- Disclosure of metadata about customer's data
  This kind of metadata includes data access and data usage logs, location data and data tracking information.
- Disclosure of metadata about customer
  This kind of metadata includes information about customer's usage profiles of cloud services.
- Disclosure of security relevant information
  Cloud customers may build virtual infrastructures. Auditing capabilities for such infrastructures may be provided by the cloud service provider as well as a third party. It is important to design them in a way, that no potentially harmful information is disclosed.

In this section, we provide an overview about recent research conducted in the area of privacy with respect to audits. We also give a short overview of how audits increase trust in cloud computing infrastructures.

3.1 Operational and Reliability Audit

By moving to the cloud, customer's (also owners of data) are placing control of their data into the hands of the cloud service provider. Two big arising issues are reliability and availability. Threats to these goals are for instance (silent) data corruption and (unintentional) deletion. Data corruption can have multiple causes like hardware faults in the networking or storage subsystems, or software bugs during processing of data. Deletion of data can be caused by careless provider personnel as well as software bugs. Risks evolving around these problems are for instance high costs caused by downtimes. On the provider side there is a risk of potential reputation loss. To mitigate these risks and provide the customer with an adequate level of assurance, that his data are safe with the cloud provider and stored correctly, audits can be used.

However, actually checking the correctness and retrievability of data is not a trivial task. Simply downloading data from the cloud and checking whether or not its integrity was harmed or parts are missing is not a feasible solution in cloud environments. Having to download terabytes of data and calculating checksums on that data would place a heavy burden on the customer or auditor, financially as well as in terms of resources needed.

Therefore, new ways of reliability auditing have to be found for cloud environments. There already exist several approaches, which aim at assuring data integrity and retrievability in the cloud while removing the need to download and check huge chunks of data. Most of the approaches focus on a third-party auditor (TPA) performing audits on cloud infrastructures.

Proof of data possession (PDP) [8] is a concept that uses homomorphic verifiability tags (HVTs) to prove that a remote party is in possession of a file.

This is done by generating metadata about the file prior to upload and verifying possession using a challenge-response protocol later on.

Proofs of retrievability (POR) [9] is a concept to enable an entity to verify the intactness of remotely stored data without requiring to retrieve large files. Thereby the server proves with a high probability to a data owner, that a file can be retrieved even in the case of some parts being corrupted. Most of the presented approaches leverage some kind of POR to enable reliability audits of cloud storage services [10].

In [11] Wang et al. propose a network architecture for secure data storage in the cloud. They acknowledge the previously described availability and privacy problems in current cloud storage systems. In their solution, a trusted TPA conducts audits on behalf of the data owner. They assume, that there is no incentive for the TPA to violate the privacy of the data owner. However, by proposing a solution, where the TPA is denied access to content, data leakage to the TPA shall be prohibited. Therefore, they propose the requirement, that a TPA must not know the contents of data, which is audited. Further goals of their approach are the support of dynamic data updating, batch auditing, and minimization of auditing overhead (e.g., network bandwidth). They propose cryptographic techniques like using homomorphic authenticators, and Merkle hash trees [12] to fulfil those requirements.

Boyang et al. take the concept of public auditability one step further and propose a system, where shared data stored in the cloud can be verified while preserving the identity privacy of each signer of a data block [13]. By using ring signatures to construct homomorphic authenticators the TPA is able to verify data, while not leaking identity information.

Furthermore there exist projects, that include TPAs and public auditability principles to enable secure and trustworthy cloud storage systems [14–17].

Performance auditing and the privacy issues introduced by it have been investigated to a much lesser degree. Large cloud provider such as Amazon, Google or Microsoft usually provide their customers with specialized interfaces (e.g., Amazon CloudWatch) for extracting performance monitoring information. Such information includes, but is not limited to, CPU utilization, network I/O statistics. The degree to which such information is published (e.g., on a per VM basis or details about the actual cloud infrastructure performance) is chosen by the cloud provider. Privacy problems may be linked strongly to the number of details published and the multi-tenant nature of cloud environments. Performance counters published to one customer might be used to deduce information about tenants using the same shared resources.

TPAs auditing cloud performance might need elevated privileges on the examined services (e.g., for measuring loading and saving times in a SaaS scenario). A TPA also has to assess the accounting system of cloud providers. This is usually tied very closely to the collection of usage logs and performance counters. To make a statement about the correctness of a provider's accounting processes, the TPA needs access to such information. However, without proper measures to protect the customer's privacy (e.g., proper level anonymization) a TPA can

easily extract cloud usage profiles (e.g., service interaction and communication) from this information. One thing to consider is the reversal of anonymization by combining multiple sources of anonymized data.

Operational and reliability audits are used to assess a providers procedures, systems, records and activities in order to test the adequacy of controls in place.

3.2 Regulatory Compliance Audit

Another form of audits is for regulatory compliance. Some businesses require cloud service providers to be compliant to or certified against certain regulations. Prime examples of such industries are healthcare, where sensitive medical information is processed, or finance, where sensitive financial data about individual subjects are processed. Typical examples are SAS70 [18] reports or HIPAA [19], which define how such data may be processed. Furthermore, there exist ISO27001 [20] which addresses information security management and CSA CloudAudit [21], which combines several of the previously mentioned audit frameworks to address cloud specific issues. However, these kinds of audits usually contain a lot of non-automatable audit tasks such as questionnaires or interviewing experts at the provider.

Reports generated by these frameworks are often not made available to the public. Amazon for example only releases their SAS70 report, when a customer contacts the Amazon support and requests it specifically. Regarding the privacy of other customers, these reports might reveal security flaws in the provider's processes, which could be exploited.

Regulatory compliance audits are used to assess a providers compliance in order to:

- Test the adequacy of controls in place
- Verify that a provider complies with established policy
- Verify that a provider complies with operational procedures (e.g., COBIT)

3.3 Security and Information Privacy Audit

Cloud security audits are supposed to uncover flaws and vulnerabilities in cloud infrastructures and service delivery chains, for instance: reveal unauthorized access to services and data, destruction of data and denial of service (DoS). The goal of this kind of auditing is to assure an appropriate level of protection especially by following industry best-practices.

However, information that is used during a security audit usually contains highly sensitive data, such as access logs and performed actions by customers. Special care has to be taken when providing such information in a public audit system.

Security audits are used to asses a provider's security issues in order to:

- Detect breaches or potential breaches of compliance
- Detect badly configured services

Information privacy audits are used to assess a providers compliance especially with respect to customer data managing in order to:

- Verify that a provider complies with established data policy
- Verify that a provider complies with privacy acts

4 Proposed System Requirements for a Public Cloud Audit System

In this section, we propose a privacy-respecting system for public cloud audits and address the most important requirements.

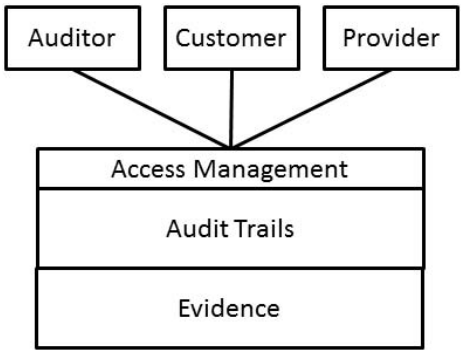

Fig. 2. Privacy-respecting Views on Cloud Audit Reports

Figure 2 depicts a high-level scenario for such a system. The foundation comprises of an evidence base. Evidence is any information needed to provide proof for specific audit tasks. This may include logs, configurations, data provenance information, timestamps, checksums and any other kind of data and metadata, which might be useful in constructing audit trails. Audit trails provide reliable proof about a certain audit task. However, as previously described audit trails might leak PII. Therefore, different views on audit reports must be provided to different stakeholders. This is done using the access management layer, which, depending on the actor requesting the audit report, provides reports with different levels of detail.

Sample Scenario: For example, consider the following scenario:

> A task for auditing the data life-cycle events of a customer's data in the cloud is requested. The sources of the relevant information are the cloud management system (CMS), which tracks high-level events such as provision/deprovision of cloud storage and the storage backend, which provides detailed information about data access, location and usage. A report is generated as a result depicting the audit result.

Mapped to the approach described, three different views on this report are provided by the audit system:

- Customer view:
  The customer view is the most detailed. It provides information about the CMS events, actions performed on the storage, such as who accessed what including timestamps, retrievability checks results.
- Auditor view:
  The auditor's view is less detailed. According to the checks defined in the audit request, the auditor is provided with more high-level results. Evidence accompanying the report is anonymized as needed.
- Provider view:
  The provider's view on the audit report contains mostly information stemming from the CMS.

Cloud Audit System Requirments: From this we derive requirements for the cloud audit system:

- Interfaces: public audit interfaces may differ among providers which complicates interoperability. In complex service provision chains, this hinders efficient auditing (see Section 5).
- Formats: differing data formats for the same information among providers also hinders interoperability. Also, tools used to extract information for specific parts of the audit trail usually use proprietary formats.
- Data collection: auditing requires the collection of data across all architectural levels. However, no more information than actually needed to provide proof for the audit task at hand shall be collected.
- Specific audit tasks: audit tasks shall focus on specific tasks. This shall enable reducing the amount of data needed to provide proof for the task.
- Dynamics: cloud environments are very dynamic. Therefore, a dynamic mechanism for collecting relevant evidence is needed. The audit system must support a mechanism to react address dynamic infrastructures.

5 Audit Challenges in Cloud Service Delivery Chains

Complex cloud service provision scenarios, where multiple service providers are chained for service composition, introduce new audit challenges. Typically, these chains are hidden from the customer. However, the customer's data are transferred along the service provision chain. For example a SaaS service provider may use an IaaS service provider's infrastructure to deliver its service. The IaaS service may be chosen depending on pricing and performance parameters and can also be exchanged without the customer's notice. This is a rather simple example for a cloud service delivery chain. More complex scenarios are thinkable, when multiple SaaS, PaaS and IaaS providers are involved.

Auditing a service and the subsequent services it depends on, introduces the following challenges:

- Audit of each involved service: each service involved in the service provision chain needs to audited.
- Audit of data transfer between services: the communication between services along the provision chain needs to be audited.
- Regulatory compliance: some parts of the provision chain may be located abroad, placing them under different jurisdictions.

6 Conclusion

In this paper we presented public cloud audit as a possible solution to increase trust in cloud computing by providing a proof and higher cloud transparency. We thereby focused on privacy concerns which arise in public cloud audit scenarios and addressed them in our proposed solution. By providing cloud stakeholders with different access views on cloud audit reports and therefore privacy of cloud customers can be protected.

In our future work, we will refine the high-level requirements of this system. Additionally, more complex service delivery chains, which involve multiple cloud service providers will be analyzed in detail to make public cloud audits practicable and service delivery chains auditable.

Acknowledgment. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no: 317550 (A4Cloud).

References

1. Mell, P., Grance, T.: The NIST Definition of Cloud Computing. Technical report, National Institute of Standards and Technology, Information Technology Laboratory (2011)
2. CNN: Google fires engineer for privacy breach (September 2010), `http://edition.cnn.com/2010/TECH/web/09/15/google.privacy.firing/`
3. Cloud Security Alliance (CSA): Security, Trust & Assurance Registry (STAR), `https://cloudsecurityalliance.org/star/`
4. Amazon: Amazon's service health dashboard, `http://status.aws.amazon.com/`
5. Privacy Rights Clearinghouse: Amsterdam Hospitality Group, `https://www.privacyrights.org/data-breach-asc?title=amsterdam`
6. Pearson, S., Benameur, A.: Privacy, security and trust issues arising from cloud computing. In: 2010 IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), pp. 693–702 (2010)
7. Badger, L., Bohn, R., Chu, S., Hogan, M., Liu, F., Kaufmann, V., Mao, J., Messina, J., Mills, K., Sokol, A., Tong, J., Whiteside, F., Leaf, D.: US Government Cloud Computing Technology Roadmap Volume II Release 1.0 (Draft) - Useful Information for Cloud Adopters. Technical report, National Institute of Standards and Technology, Information Technology Laboratory (2011)

8. Ateniese, G., Burns, R., Curtmola, R., Herring, J., Kissner, L., Peterson, Z., Song, D.: Provable data possession at untrusted stores. Cryptology ePrint Archive, Report 2007/202 (2007), http://eprint.iacr.org/
9. Bowers, K.D., Juels, A., Oprea, A.: Proofs of retrievability: theory and implementation. In: Proceedings of the 2009 ACM Workshop on Cloud Computing Security, CCSW 2009, pp. 43–54. ACM, New York (2009)
10. Juels, A., Kaliski Jr., B.S.: Pors: proofs of retrievability for large files. In: Proceedings of the 14th ACM Conference on Computer and Communications Security, CCS 2007, pp. 584–597. ACM, New York (2007)
11. Wang, C., Ren, K., Lou, W., Li, J.: Toward publicly auditable secure cloud data storage services. IEEE Network 24(4), 19–24 (2010)
12. Merkle, R.C.: Protocols for public key cryptosystems. In: IEEE Symposium on Security and Privacy, pp. 122–134 (1980)
13. Wang, B., Li, B., Li, H.: Oruta: Privacy-preserving public auditing for shared data in the cloud. In: 2012 IEEE 5th International Conference on Cloud Computing (CLOUD), pp. 295–302 (2012)
14. Li, L., Xu, L., Li, J., Zhang, C.: Study on the third-party audit in cloud storage service. In: 2011 International Conference on Cloud and Service Computing (CSC), pp. 220–227 (2011)
15. Patel, H., Patel, D.: A review of approaches to achieve data storage correctness in cloud computing using trusted third party auditor. In: 2012 International Symposium on Cloud and Services Computing (ISCOS), pp. 84–87 (2012)
16. Zhu, Y., Hu, H., Ahn, G.J., Yau, S.S.: Efficient audit service outsourcing for data integrity in clouds. J. Syst. Softw. 85(5), 1083–1095 (2012)
17. Shah, M.A., Swaminathan, R., Baker, M.: Privacy-preserving audit and extraction of digital contents. Cryptology eprint archive, report 2008/186 (2008)
18. SAS70: SAS70, http://sas70.com/sas70_overview.html
19. U.S. Government Printing Office: HIPAA, http://www.gpo.gov/fdsys/pkg/PLAW-104publ191/html/PLAW-104publ191.htm
20. ISO: ISO27001:2005, http://www.iso.org/iso/catalogue_detail?csnumber=42103
21. Cloud Security Alliance (CSA): CloudAudit A6 Cloud Security Alliance, http://cloudaudit.org/