Detlef Hühnlein et al. (Eds.): Open Identity Summit 2015,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2015 153

Towards a Secure Cloud Usage for Financial IT

Marcus Hilbrich[1] Ronald Petrlic[2] Steffen Becker[3]

Abstract: Cloud Computing and Big Data are the current hot topics in research and industry. Based on the enormous amount of preliminary work, ranging from grid and distributed computing to data mining and clustering, to name only a few approaches, cloud computing has become a de-facto standard for computing in general and data-intensive industry tasks in particular. Thus, a lot of questions about how to develop and implement such systems are already answered, but nonetheless, there is reservation to adopt such techniques in some business areas. Most of the reservations are due to security reasons, as in certain areas, like in the banking sector or in the health industry, high levels of security standards have been met for decades and those standards must not be weakened. This is the reason why we investigate—closely together with partners from the industry—how to overcome security concerns in the adoption of cloud computing in the financial industry. An introduction to our strategies is given with this paper.

Keywords: Cloud Computing, Security, Financial IT, Scalability, Elasticity, Distributed Infrastructure, Services, Storage

1 Introduction

In many commercial computing environments it is a common concept to pay for on-demand resource usage. This avoids to over-provision resources and pay for under-utilised hardware. The concept also supports an outsourcing of IT tasks that are not the primary concern of the business. In short, the operation of hard- and software is replaced by services from a provider. Overall, this approach allows for a concentration on the core business and it constitutes a variable and often more cost-efficient usage of exactly the amount of resources that are required in a specific situation. A common solution to this issue is the usage of cloud systems.

In a typical scenario for an IT-assisted enterprise, it is usual to store and process both the customers' and the company's data, which are mainly related to a concrete business. The access to the data has to be restricted. The services can be self-developed and be part of the company's knowledge base or be general or purchased. Depending on the categorization, the services are strongly business-related or be commonly available tools. In most cases, the operation of computing resources, though, does not constitute a core business. Otherwise, it is needed to have enough resources available at all times, even under rare events, which appear just once a year like, e.g. during Christmas shopping period or balancing of accounts. Thus, in case the hardware is provisioned, the mean

[1] Software Quality Lab (s-lab), Universität Paderborn, marcus.hilbrich@uni-paderborn.de
[2] Software Quality Lab (s-lab), Universität Paderborn, ronald.petrlic@uni-paderborn.de
[3] Software Engineering Chair, Technische Universität Chemnitz, steffen.becker@informatik.tu-chemnitz.de

utilisation is often very low.

Considering cost, utilisation, flexibility, and availability, it is often a good decision to use cloud systems with a pay-per-use pricing model. However, cloud technologies also bring new challenges. Many of them are already addressed and have to be adapted to the concrete situation, others need the introduction of innovative ideas. However, we expect that we can handle all the challenges and strongly benefit from using cloud technologies.

The following aspects need to be considered when your data is not under your physical control. You have to avoid a locked-in syndrome [PC09], you have to care about Service Level Agreements (SLAs) [Be11] and you have to establish mutual trust with your service and resource providers.

We are working in the project "Securing the Financial Cloud"[4] (SFC)[5]. The aim of the project is to explore how the advantages of cloud-like systems can be utilised by computing and storage systems of financial IT. This means we have to deal with data from e.g. Automatic Teller Machines (ATMs), bank transfers, balance of accounts, inter-bank transfers, and commercial papers. This data is highly valuable in terms of money, has a very high protection demand defined by the stockholder, and a bulk of legal restrictions.

Besides the data, we have to deal with the analysis performed by services. We have different kinds of aspects of such services. From simple ones that balance an account or calculate interests, which have a low protection demand[6] up to services that estimate the financial standing of a bank-related customer or support strategical investment decisions that hold strong intellectual properties and need an according protection.

To allow stakeholder with different security demands to have separate and shared data in a cooperatively used system, a multi-mandatory support has to be realised. This allows e.g. inter-bank communication (for transferring money from one bank to another one and so on). To avoid a locked-in syndrome, the data have to be sorted by different resource providers and data processing has to be realised independent from a concrete resource provider. In concrete we want to enable a provider independent usage of private and public clouds (hybrid cloud concepts) [Ro11] and cloud brokers [BRC10].

So we have to deal with the already known challenges of scalability, elasticity, and fulfilment of SLAs in a cloud-based environment which has to be matched to a context with very strong demands on safety and security. In concrete, we have to investigate within the project the following concepts:

- Geographical storage locations of all data have to be known and have to be restricted based on SLAs.

[4] Funded by BMBF under grant ID 16KIS0062
[5] http://www.vdivde-it.de/KIS/sichere-ikt/sicheres-cloud-computing/sfc
[6] This holds only for the security level of the algorithm, not for the data the algorithm runs on.

- All data is stored encrypted, so even in case of an SLA violation the data are protected.
- Data is never overwritten or deleted, updates are performed by providing an additional version of a file.
- Analysis processes are realised as services which allow that different companies can use the same software [PC09].
- Multiple execution zones have to be supported, e.g. local computing centre, private clouds, public clouds and hybrid clouds.
- The integrity of the data can be validated based on cryptographic methods.

2 Preconditions and Actual Situation

During the arrangement and the beginning of the SFC project we worked out the conditions for a later realisation phase. One of the terms we have to fulfil is a legal one. For a large part of the data, the geographical location of storage and processing is restricted [Re00, Hi06], e.g. to the country the data is accrued. This can be considered by cooperating with local cloud operators that ensure a concrete data location.

Another aspect is the required security level from the stakeholder of the data. Often the protection level can be ensured by SLAs [BA11]. In some cases, especially for financial data it is also common to demand that an external resource provider is not able to read the data. In this case a cryptographically secure solution has to be established.

For the project, so-called Hardware Security Modules (HSMs) can be used. These are trust anchors that can store cryptographic keys and perform encryption and decryption. Based on the hard- and software based security arrangement, these operations can be considered as secure because in case of an attempt to breach the modules, their data are destroyed. The modules provide security even when not physically controlled by the stakeholders. So they can be offered as a special service by a cloud provider which needs a secured initialization process. In a preceding work, members of the project group have already covered the security aspects of the client-side, i.e. security of ATMs, making use of Trusted Platform Modules (TPMs) that served as trust anchors [PS14].

Moreover, we make use of a relatively new cryptographic approach called attribute-based encryption (ABE). Such ABE schemes allow for a fine-grained access control. Data are encrypted under certain access structures and only users/processes that possess private keys with the corresponding attributes that fulfil the access structure of the ciphertext allow for a decryption. The private keys can only be issued by a central key server.

3 Envisioned Target Architecture

As a starting point, a general architecture which mainly focuses on the functional view was developed. This architecture can receive input data that are financial transactions, e.g. from ATMs or inter-bank communication. The output data are e.g. account balances or the result of automatic analysis processes, which is often determined knowledge about the bank customer's behaviour or knowledge-based decisions. To realise the in- and output mechanisms, a communication layer is intended (see Fig.1).

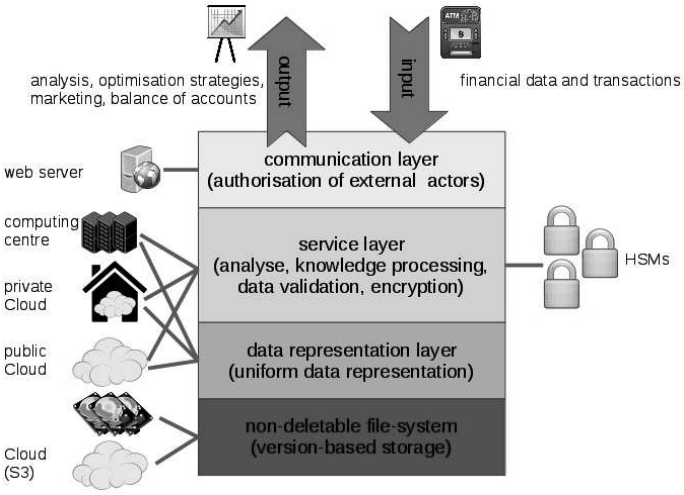

Fig. 1: Layer-based view of the SFC Concept

An additional layer is the service layer. This layer realises information and knowledge generation and processing, which are the demanded services from the stakeholders. Another important task is the validation of the input data, e.g. checking source signatures or transaction order, which is also realised as a service. Based on the fact that this layer has to do the data processing, this layer needs to be able to decrypt the data. This is realised by using the HSMs. The main reason for making use of HSMs in our scenario is due to a separation between the cloud provider, i.e. the provider performing the computations on the data, and the data owner, i.e. the bank that outsourced the processing of its data to the cloud provider. The HSM is in full control of the bank, i.e. the cloud provider is not able to retrieve the bank's private keys that are stored on the HSM. Based on the fact that the other layers do not realise data processing, it is not needed to decrypt data in any other layer. Thus, this is the only layer which needs access to the HSMs. We will establish a multi-agent-based system to realise flexible and efficient data processing combined with a blackboard design pattern for communication and data access. From a research perspective, there are two challenging tasks that we currently deal with at the service layer. First, we need to find a way to virtualize HSMs.

The hardware HSMs constitute the trust anchors of those virtualized HSMs. Therefore, we need to find a way to provide as strong security guarantees for virtualized HSMs as they hold for hardware HSMs. The second challenge is to find out to which extent we can implement our developed attribute-based encryption scheme on state-of-the-art HSMs.

The data representation layer offers a uniform and up-to-date view to the data. Based on the current state of the project, it is planned to represent the data as a POSIX-based file system. This allows a very simple and general interface to realise and adapt services to the infrastructure of SFC. Another aspect is the sufficient performance and scalability of modern distributed file systems known from cloud and High Performance Computing (HPC) context. This even allows a communication of services on different locations via the uniform file-system view. The data representation layer also guarantees that data is not deleted or overwritten. To realise this property, an additional layer is used. This additional layer is a non-deletable data storage which holds all versions of a value. The data representation layer provides a view to this data storage which only holds the last version of the data. As already described, the data storage and representation layer only work on encrypted data and do not need to access HSMs. Data validation is not part of this layer. A high level validation based on cryptographic integrity tests is provided as a service which probably needs access to HSMs and low level data safety is provided by the file system layer described next.

To realise the non-deletable file-system, open source tools like Ceph[7] will be evaluated. These tools also have to offer an additional property which is demanded by the SFC system. For realising a safe storage of the data, it is needed to realise a replication-based physical storage strategy. Therefore, it is needed to distinct between different geographical locations to avoid that copies of data are written to the same physical location. This concept of replication and distinction of geo-locations is e.g. supported by Ceph. This avoids to reimplement a distributed storage strategy as part of the SFC-project.

In the layers of the SFC Infrastructure, different execution environments will be supported. An example is the service execution. Therefore, it is needed do distinct between the service description, which holds all the information to run the service, and the execution system. The service description can be a virtual machine image or a container image. Such an image can be started on a server or a container execution system (e.g. Docker[8]) to operate an instance of the service. So the same service can be executed in different environments, e.g. local computing centres or the public cloud. In this way it is even possible to have the same service with different security contexts, depending on the execution environment. Another example is the data storage where different storage systems like e.g. cloud storage, servers with disks, and nodes with network attached storage can be used to form a uniform file system.

[7] http://ceph.com/
[8] http://www.docker.com/

4 Conclusion and Future Work

To provide a relevant contribution to the field of secure cloud architectures, in future work, we will show that even financial data can be processed. Thus, we develop a cloud-based infrastructure taking into account security constraints in particular. As part of future work, we will investigate how attribute-based encryption can be combined with high security modules in an efficient way, i.e. we will need to analyse which tasks need to be performed on the HSMs and which tasks can be executed on the "ordinary" machines. Moreover, we also need to come up with a holistic security approach that includes organizational security aspects additionally to technical measures. Based on the security concept we also develop a prototype which will be able to process and store data in a scalable, elastic and efficient manner. So we can prove to benefit by using cloud environments even under hard security and safety constraints.

References

[Ba11] Badger, Lee; Bohn, Robert; Chu, Shilong; Hogan, Mike; Liu, Fang; Kaufmann, Viktor; Mao, Jian; Messina, John; Mills, Kevin; Sokol, Annie; Tong, Jin; Whiteside, Fred; Leaf, Dawn: NIST Special Publication 500-293, US Government Cloud Computing Technology Roadmap, Release 1.0 (Draft), Volume II Useful Information for Cloud Adopters, November 2011.

[Be11] Bernsmed, K.; Jaatun, M.G.; Meland, P.H.; Undheim, A.: Security SLAs for Federated Cloud Services. In: Availability, Reliability and Security (ARES), 2011 Sixth International Conference on. pp. 202 –209, aug. 2011.

[BRC10] Buyya, Rajkumar; Ranjan, Rajiv; Calheiros, Rodrigo: InterCloud: Utility-Oriented Federation of Cloud Computing Environments for Scaling of Application Services. In (Hsu, Ching-Hsien; Yang, Laurence; Park, Jong; Yeo, Sang-Soo, eds): Algorithms and Architectures for Parallel Processing, volume 6081 of Lecture Notes in Computer Science, pp. 13–31. Springer Berlin / Heidelberg, 2010.

[Hi06] Hildebrandt, Mireille: Profiling: From data to knowledge. Datenschutz und Datensicherheit - DuD, 30:548–552, 2006. 10.1007/s11623-006-0140-3.

[PC09] Parameswaran, A. V.; Chaddha, A.: Cloud Interoperability and Standardization. In: SETLabs Briefings, Vol. 7, 2009.

[PS14] Petrlic, Ronald; Sorge, Christoph: Establishing user trust in Automated Teller Machine Integrity. IET Information Security, 8(2):132–139, 2014.

[Re00] Rehm, Gebhard Marc: Just Judicial Activism? Privacy and Informational Self-Determination in U.S. and German Constitutional Law, January2000. Available at SSRN: http://ssrn.com/abstract=216348 or http://dx.doi.org/10.2139/ssrn.216348.

[Ro11] Rochwerger, B.; Breitgand, D.; Epstein, A.; Hadas, D.; Loy, I.; Nagin, K.; Tordsson, J.; Ragusa,C.;Villari, M.;Clayman,S.;Levy,E.;Maraschini,A.; Massonet,P.;Muõandoz, H.; Tofetti, G.: Reservoir - When One Cloud Is Not Enough. Computer, 44(3):44 –51, march 2011.