Multilateral Privacy in Clouds: Requirements for Use in Industry

Ina Schiering, Markus Hansen

Multilateral Privacy in Clouds: Requirements for Use in Industry

Ina Schiering[1], Markus Hansen[2]

[1]i.schiering@ostfalia.de, [2]markus.hansen@privacyresearch.eu

Abstract. After the virtualisation of single components of computing systems such as storage, networks or computing devices the next step is the abstraction of the infrastructure as a whole: cloud computing. There are already cloud services on the market, but most of them rely on proprietary technology. Hence standards for cloud computing are needed that realise the requirements we have for present systems. In this context it is important to think of requirements for privacy when personal data are distributed in cloud services and on the other hand on restrictions an owner of computing resources wants to impose. It is important to note that the concepts that enable multilateral privacy are also needed by industry for the flexible realisation of service level agreements and governance to incorporate cloud services in business processes and to be compliant with legal regulations as e.g. SOX, EuroSOX. Therefore the methods that are needed to realise business critical IT services as cloud services are the same as for privacy.

Keywords: multilateral privacy, privacy, data security, data protection, cloud computing, clouds, requirements, identity management, compliance

1 Introduction

Cloud computing refers to methods to dynamically utilise scalable IT services, so called *cloud services,*for a certain purpose over networks, especially the Internet. To achieve this, the abstraction paradigms of virtualisation and scalability are used in combination. While virtualisation allows single physical resources to appear and be used as multiple resources of the same type as the initial single one, scalability allows the cloud users to use IT services as flexible as needed: IT services can be ordered dynamically even for special events as training or testing purposes.

We denote the party (company or private user) that uses a cloud service as a *cloud user*. We concentrate here mainly on companies as cloud users. Cloud computing is offered in the form of a *cloud service*. Cloud services are offered by *cloud providers*. Cloud providers and cloud users are denoted as *interacting partners* in the cloud if we do not need the distinguish between them.

Clouds can be operated by several actors, and the services offered from a cloud can be used in several constellations. In e.g. enterprise environments, spare resources can be offered internally as cloud services to allow for a higher level of utilisation. In this case, where provider and user of the resulting cloud are basically the same instance, the cloud is called an *internal cloud*. On the other hand, cloud services might be offered from an external supplier, e.g. a company that has specialised in operating clouds and sells services or wants to monetise spare resources and operational competencies. In the case of such *external clouds*, all physical resources that are the basis of cloud services are out of physical reach of the cloud users. It is also possible

to extend internal clouds by joining them with external clouds, resulting in *hybrid clouds*.

A cloud service might be a single service as it is the case with storage or compute services as e.g. Amazon S3. That sort of cloud service is named *IaaS* (Infrastructure as a Service). Since for data security and privacy questions we need to describe where the data is located, we denote each cloud provider who owns resources a *resource owner*. Some cloud providers for IaaS cloud services act only as intermediaries, where resource owners rent spare resources to the cloud provider who joins resources from several resource owners to form an IaaS cloud service. But a cloud service can also be the aggregation of multiple physically independent services to appear and be used as a single services. The intention here is to use a combined platform (*PaaS* - Platform as a Service) or even a special software (*SaaS* - Software as a Service) and can lead to the realisation of whole business processes in the form of cloud services.

In more general scenario with cloud providers realising a cloud service based on resource owners and existing cloud services from other cloud providers, a cloud service consists of a dynamically changing network of resource owners, cloud providers and cloud users, the *cloud network* for the cloud service.

Such a cloud network is represented by a finite, directed graph where the vertices denote the cloud users, cloud providers and resource owners. There is an edge from a cloud provider to the cloud user that utilises a cloud service of that provider and there is an edge from a cloud provider resp. resource owner to a second cloud provider, if that second cloud provider incorporates the services or resources of the first one in his own cloud services. The following restrictions concerning graphs representing cloud networks apply: A vertex associated to a cloud user has no successor and a vertex associated to a resource owner has no predecessor.

The subnet of the cloud network servicing one cloud user is named the *cloud subnet* of that cloud user. This subnet is represented by the sub-graph of the cloud network induced by the vertices of the cloud user and all cloud providers, resource owners that are utilised to provide the cloud service for that user.

Cloud users can dynamically decide to begin or end using a cloud service. They can in an automated way request more entities of the cloud service e.g. more resources as storage and system instances. In the case of SaaS the cloud users implicitly scale the cloud service by changing the number of users, transactions or by a different choice of software modules. The cloud provider has to provide the cloud service and needs potentially to involve a dynamically changing number of resource owners and other cloud services as needed. A resource owner or cloud provider might want to sell services or resources only for a certain amount of time, e.g. spare resources that are needed later. Hence we speak of a cloud network or a cloud subnet of a cloud user at a certain point of time.

Cloud services are a interesting alternative especially for small up to medium size companies. Companies of that size have a limited amount of IT personnel, know-how and a limited IT Budget. Instead of investments in IT it could be an interesting to use cloud services for complex processes e.g. email, customer relationship management (CRM), enterprise content management (ECM), enterprise resource planning (ERP), data archiving, project management or the desktop. Also it could be interesting to use IaaS services for e.g. storage if the cloud provider offers interesting service levels that are difficult to realise as mirroring over different physical sites, off-site backup or high availability of the computing platform.

Beside the advantages of using the know-how and the resources of the cloud services using cloud services incorporates also several risks: The cloud user needs legal warranties concerning data security and privacy from the cloud provider and the whole cloud subnet that realises the cloud service for him at any point in time, since personal and business critical data are operated in the cloud subnet.

In this context we need to consider an adequate generalisation of the concepts of security and privacy: Multilateral security and multilateral privacy. The concept of multilateral security [7] aims at allowing all parties of an interaction to express their security objectives, at recognising conflicting objectives and (automatically) negotiating compromises, and at enforcing objectives within the scope of the compromises negotiated. To enforce the objectives, mechanisms have to be established to allow effective control. Analogously the concept of multilateral privacy refers to clouds that address the privacy (or secrecy in case of legal entities) objectives of all participating parties, with no party taking precedence over another [8].

2 Cloud Requirements

In the case of IaaS the basic functional requirements are concerning type and clock rate of the CPU, the amount of memory or disk space. For SaaS there are functional requirements for the software used, e.g. collaborative work on documents. Beside the functional requirements there are typically operational requirements: The cloud user needs to start, stop and configure the service. For full flexibility of the service automatic provisioning must be possible. Beside these requirements there are non-functional requirements that are normally formulated in the form of an SLA[1] (service level agreement): for example requirements concerning availability, reliability, scalability, data integrity, data security, privacy, access control, legal regulations.

Directive 1995/46/EC of the European Parliament and of the Council (Data Protection Directive) and Directive 2002/58 on Privacy and Electronic Communications (E-Privacy Directive) are EU directives on data protection and privacy. They provide a regulatory framework to the EU member states that must provide legislation accordingly. With regard to the specifics of cloud computing, the most important regulation concerns transfer of personal data[2] to third countries, i.e. countries outside the EU. Personal data may only be transferred outside the EU if those third countries provide an adequate level of privacy protection. For transfer of data to the USA, the Safe Harbour Agreement applies. Companies in the USA can opt-in to Safe Harbour, thereby stating that they follow adequate data protection principles. Then EU companies are - as a general rule - allowed to transfer personal data to them.

In addition to the principle that personal data may only be transferred to countries with adequate protection, further principles that must be complied with according to the Data Protection Directive are that any personal data has to be fairly and lawfully processed, may only be processed for limited purposes, has to be adequate, relevant and not excessive, has to be accurate, must not be kept longer than necessary, may only be processed in accordance with the data subject's rights, and has to be secure.

Examples for other legal regulations cloud users have to comply with are in the USA SOX (Sarbanes-Oxley Act), enacted as a reaction to accounting scandals around companies like Enron, WorldCom, etc. SOX demands e.g. an internal control system for corporations in the USA and all subsidiaries. Similar requirements have evolved in the EU as Directive 2006/43/EC of the European Parliament and the Council of 17 May 2006 on statutory audit of annual accounts and consolidated accounts, and

1 For terminology concerning IT services and service level agreements see [1]

2 The term 'personal data' is defined in the European Data Protection Directive 1995/46/EC, Article 2(a): "'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;"

Directive 2008/30/EC of the European Parliament and of the Council of 11 March 2008 amending Directive 2006/43/EC on statutory audits of annual accounts and consolidated accounts (also named EuroSOX).

To comply with e.g. SOX, EuroSOX organisations need as a prerequisite transparent and documented business processes. Since most processes are supported by IT systems this implies a transparent and documented IT environment. Based on this concrete controls can be defined: For a business process concrete control objectives are formulated, the legal regulation that is the cause for the control objective and the proceeding to monitor the control objective. An example for such a control objective is that an invoice is only paid for if there is a valid quote and the responsible person confirms that the goods resp. services are delivered in correspondence with the quote. Monitoring of control objectives can often be realised in IT systems.

Cloud providers and resource owners on the other hand have requirements concerning monitoring, measuring, reporting and billing for services. They are interested in an easy way to integrate services to create new cloud services on the basis of existing services. Cloud providers have to comply with legal regulations for their services, e.g. export control regulations. So there are restrictions concerning the countries where a cloud provider is allowed to sell services.

3 Methods

To realise the requirements of legal regulations in a cloud environment, e.g. internal control systems, similar mechanisms are needed as for ensuring data security and privacy: F*ederated identity management can* realise access control and monitoring and reporting on access. Since it does not correspond to the flexibility and dynamic of cloud services if the cloud user has to negotiate an SLA with each cloud provider in the form of a contract, there must be an automatic process for the communication of these requirements in the *cloud interface*, oftena *cloud API (application programming interface)*. Finally the cloud user needs *control and certification mechanisms* to check that the requirements are fulfilled. In the following we describe cloud interfaces and control and certifications mechanisms in more detail. For an overview about federated identity management see [2].

3.1 Cloud Interface

Concerning cloud interfaces resp. cloud APIs there are currently two different approaches: For SaaS, a web browser is mainly used as interface. In the case of IaaS several APIs exist that are specific for the respective cloud provider, e.g. Sun Cloud API, Amazon EC2 API, etc. They are mainly based on XML or JSON (JavaScript Object Notation). They are generally used to represent functional requirements. Therefore it represents a risk to use these cloud services for business critical environments where at least requirements concerning compliance, availability, privacy and data security have to be assured. In addition, as each provider uses his own API, changing the cloud provider will lead to a change of the software of the cloud user as a different API has to be used. Hence the goal is the development of standardised cloud APIs that allow the formulation of non-functional requirements.

There are initiatives that try to develop cloud APIs for at least IaaS environments where it is possible to formulate non-functional requirements as e.g. the Open Cloud Computing Interface Working Group (OCCI-WG). The OCCI-WG works on an API for IaaS cloud services based on cloud APIs in industry. Some draft documents do

already exist that line out use cases [4]. They rely on the RESERVOIR architecture where the architecture consists of resource owners, cloud providers that work as intermediaries, and cloud users [3]. For further examples of initiatives that work on cloud APIs in the IaaS field see [5] (SNIA), [6] (DMTF).

Each interacting partner in the cloud network has requirements that need to be fulfilled. Because of the dynamic change of the cloud network the requirements have to be interchanged and checked automatically. Hence they can be formulated as in the example of the OCCI-WG in XML.

As a first step to the formulation of requirements in an API they must be categorized: categories as e.g. high, medium or low availability are created where each category is documented by the service provider. A cloud user begins using a cloud service. Hence he requests the cloud service from the cloud provider where requirements are expressed in XML. The cloud provider checks if all requirements are fulfilled. If that is the case, he acknowledges the request. Otherwise he starts requests to all direct successors in the cloud network that are needed to provide the service with the defined quality. These requests should be derived automatically. The requests are tagged with the initial cloud provider and a number for the request. Each cloud provider and resource owner answers only once to each request and stores all requests and answers. This assurance process is executed recursively. It terminates since the graph representing the cloud network is finite. At least all resource owners, whose corresponding nodes in the graph do not have predecessors, can acknowledge or non-acknowledge the requirements. When the cloud provider has received all acknowledge or non-acknowledge messages from his direct successors in the graph, he derives from the messages if he can deliver the service with the requested quality or not. Hence he can acknowledge resp. non-acknowledge the request. The request is acknowledged if the cloud user receives an acknowledge message. Then the cloud subnet delivering the cloud service for that cloud user is represented by the sub-graph induced by the following nodes: The cloud providers and resource owners that acknowledged the requirements and where there is a path in the graph from the node corresponding to that interacting partner to the cloud user such that all nodes on the path have also acknowledged the requirements.

A cloud user can e.g. express the requirement that any data may only reside and be processed on systems located within the European Union, that only systems and services from companies outside (or, respectively, inside) a certain jurisdiction may be used, that systems and services from a business competitor may not be part of the specific cloud subnet, or that all parties have to have signed the Safe Harbour Agreement. The cloud user would submit these requirements as an XML document through the cloud API, the cloud provider would then select the resources that match the requirements in appropriate quantity and join them into the specific cloud subnet. Analogously, the resource owners can themselves also define their specific requirements to be matched against through the API, e.g. that any resources must not be used for military purposes, or that no medical data may be stored. Also, the cloud providers may have certain requirements that can be expressed and matched alike. Thus, in IaaS scenarios, security and privacy requirements can be expressed and interpreted in an automated process when initiating a cloud subnet. For SaaS scenarios, a similar approach can be followed by adding meta-data to the data to be processed to express e.g. purpose limitations that the SaaS environment has to enforce.

While multilateral security includes mechanisms for automated negotiation and, therefore, compromises (e.g. about what cryptographic algorithms and what key lengths are to be applied), privacy objectives usually are not open to compromise. The process of deciding whether a certain resource can be a node within the cloud of a certain cloud user therefore is a simple binary function, a the resource can only meet the requirements from the privacy objective of the cloud user or not.

3.2 Certification and Control

A means to allow control can be to make use of certification. Systems and services forming the cloud can be certified to meet certain security and privacy standards. Certification according to e.g. IT-Grundschutz [9] or ISO 27001 could replace actual hands-on control for security while the ICPP Privacy Seal [10] can certify privacy compliance. These certificates could be handed through from each resource to the cloud providers and the cloud users using the API. Therefore, cloud users would not have to check the resources from the resource owners for compliance themselves but would rather rely on trusted third parties, i.e. the certification authorities. Protocols using e.g. Trusted Computing components could then be used to allow remote attestation of the state of any system joining the cloud and to allow detection in case the state of a system is not according to certification or contracts.

Still, even when certified, a closed source resource can not actually be controlled and therefore has always to be regarded as a security risk, although probably a low one as for the certification. But as closed source resources also ease vendor lock-in situations, it might be wiser for cloud users to avoid them.

4 Conclusion

In case cloud providers and resource owners take care that only resources certified to meet security standards are integrated into cloud subnets, they can offer transparent and well documented IT to the cloud users that e.g. also allows to establish the location of data. Cloud users can then rely on the certification to use applications on that IT that process personal data. If such applications have received certification as for that they comply with privacy legislation, they can furthermore be offered in an SaaS scenario. But if today's certification frameworks are already capable of representing the specific requirements of dynamically interacting system is currently an open question.

Comprehensive use of combined security and privacy certification could allow SaaS to be a valid business model for processing personal data. Using the API and the certificates, cloud providers can automatically generate clouds for which certain requirements have been proven to be met. Another option would be to only offer certified clouds. Providers of certified software in SaaS making use of hardware offered by other parties have to make sure, that they will use certified systems to still be able to prove that requirements are met. Still, cloud users will have to make sure within their scope that privacy requirements for the processing of personal data are fulfilled.

References

[1] ITIL IT Service Management - Glossary of Terms and Definitions, OGC, 2007, http://www.itsmfi.org/files/ITILV3_Glossary_English_v1_2007_0.pdf.

[2] E. Maler, D. Reed, *The Venn of Identity: Options and Issues in Federated Identity Management*, IEEE Security and Privacy, Vol. 6, No. 2, March/April 2008, pp. 16-23.

[3] B. Rochwerger, D. Breitgand, E. Levy, A. Galis, K. Nagin, I. Llorente, R. Montero, Y. Wolfsthal, E. Elmroth, J. Caceres, M. Ben-Yehuda, W. Emmerich, and F. Galán. *The RESERVOIR Model and Architecture for Open Federated Cloud Computing*, IBM Journal of Research & Development, Volume 53, Number 4, 2009.

[4] Open Cloud Computing Interface WG (OCCI-WG), http://forge.ggf.org/sf/projects/occi-wg.

[5] SNIA Cloud Data Management Interface, http://www.snia.org/cloud.
[6] DMTF Cloud Incubator, http://www.dmtf.org/about/cloud-incubator.
[7] H. Federrath, A. Pfitzmann, Bausteine zur Realisierung mehrseitiger Sicherheit, in: G. Müller, A. Pfitzmann (Hrsg.): Mehrseitige Sicherheit in der Kommunikationstechnik, Addison-Wesley-Longman, pp. 83-104, http://www.semper.org/sirene/publ/FePf_97MehrsSicher.inBuch.ps.gz
[8] R. Cissée, An agent-based approach for privacy-preserving information filtering, dissertation, 2009, http://deposit.ddb.de/cgi-bin/dokserv?idn=994920466&dok_var=d1&dok_ext=pdf&filename=994920466.pdf
[9] IT-Grundschutz, https://www.bsi.bund.de/cln_155/EN/topics/ITGrundschutz/ITGrundschutzHome/itgrundschutzhome_node.html
[10] ICPP Privacy Seal, https://www.datenschutzzentrum.de/guetesiegel/index.htm