Cloud Computing Security and Privacy

Muhammad Adeel Javaid

Member Vendor Advisory Council, CompTIA

Abstract The cloud computing paradigm is still evolving, but has recently gained tremendous momentum. However, security and privacy issues pose as the key roadblock to its fast adoption. In this paper we present security and privacy challenges that are exacerbated by the unique aspects of clouds and show how they're related to various delivery and deployment models. We discuss various approaches to address these challenges, existing solutions, and future work needed to provide a trustworthy cloud computing environment.

Keywords: Cloud Security, Cloud Threats, Cloud Privacy

1.1 Introduction

Over the last years, something called “cloud computing” has become a major theme in computer science and information security. Essentially, it concerns delivering information technology as a service, by enabling the renting of software, computing power and storage. Some argue that the phenomenon is essentially nothing new; others state that gradual developments have given rise to a qualitatively different computing architecture.

In this contribution, we describe the major security and privacy challenges in cloud computing. In doing so, we take a bird’s eye view.

Firstly, we will discuss the phenomenon of cloud computing in more detail, including its relation to de-perimeterisation, or disappearing boundaries. Secondly,
the ideal of encrypted processing is presented, which could be a “holy grail" for cloud computing, together with its major limitations and practical implications. Thirdly, we focus on the physical security properties that are lost when virtualising information infrastructure, and on the question how to put these back in. Fourthly, we discuss the relation between cloud computing and the changing use of information technology by individuals as a means to establish their identity, notably by means of social networking services. Finally, we discuss ethical implications of disappearing boundaries in information technology, by proposing the concept of informational precaution.

1.2 Cloud computing

1.2.1 Foundations

The foundations of cloud computing lie in the outsourcing of computing tasks to third parties. In the old days of computing, this entailed renting a mainframe for one’s computations. With the rise of the personal computer, these rentals became largely obsolete, except for costly scientific calculations.

When computers were connected to the Internet, a new option for renting computing power became available. It was possible to use the combined power of connected computers for performing one’s calculations. So-called “grids” were established for scientific tasks. A well-known example is the SETI project\(^1\), where cosmic signals are analysed in the processor’s idle time of computers around the world, in order to find signs of extraterrestrial intelligence.

Due to the development of the Internet, and its associated clustering of processing in grids and server rooms, computing has gradually become an “infrastructure”, like the electricity network. Just as companies do not generate their own electricity, they increasingly see computing as something that has to be provided to them. Therefore, they outsource the maintenance of their IT infrastructure, bypassing the need for an in-house IT department. More and more often they will also rent the infrastructure from their provider, relieving them from buying the hardware themselves.

In classical outsourcing, companies can negotiate with their provider to establish the terms of service, including security aspects of the data processing. More and more providers, however, offer IT infrastructure as a commodity, with standard contracts and little room for smaller companies and individuals to negotiate. The combination of a) rental of information processing as-a-service and b) this

\(^1\) SETI Institute Homepage, \url{http://www.seti.org/} (accessed May 10, 2010).
service having the characteristic of a commodity, is what we call cloud computing. According to Gartner, cloud computing is defined as “a style of computing where massively scalable IT-related capabilities are provided ‘as a service’ using Internet technologies to multiple external customers”.

Cloud computing (and outsourcing in general) is one of the main causes of the broader development called de-perimeterisation. This term denotes the fading of the boundaries of organisations and their information infrastructure, thereby invalidating a security approach that focuses on those boundaries. In traditional IT security, many organisations employed such a boundary-based approach, for example based on firewalls. It is then assumed that everything within the boundary is trusted, and everything outside is not. When the organisation’s data is hosted elsewhere, such approaches are not adequate anymore.

Outsourcing and cloud-computing thus lead to de-perimeterisation. Other drives in this direction are the use of mobile devices by employees, and the hiring of consultants from third parties, who have to work within the organisation’s boundaries. All these developments challenge a containment-based approach to information security, and force organisations to implement data-level security instead. Already in 1996, challenges to the existing security paradigm were discussed, and many of those considerations have become only more valid since. Cloud computing thus reinforces existing challenges to current security paradigms.

---

2 As discussed in Paolo Balboni’s presentation at SPCC.


1.2.3 Implementations

The implementation of cloud computing is typically realised by a) invoking the Internet browser to use software, platforms and infrastructure online, and b) virtualising the underlying infrastructure to cope with flexible demand. Virtualisation here refers to the implementation of so-called “virtual machines”, with properties more or less independent of the capacities of the underlying physical machines. This means that a single virtual machine can make use of several physical computers to increase its capabilities, but also that multiple virtual machines can run on a single physical computer. The advantages lie in not having to reserve a single physical machine for a particular task, thereby reducing hardware and operational costs.

Several distinctions have been proposed with respect to cloud services. The first of these has to do with the type of service offered. What is offered can be either software (software-as-a-service, or SaaS, for example Salesforce online bookkeeping or Gmail), a platform for developing and running applications (platform-as-a-service, or PaaS, for example Force.com) or infrastructure that can be rented for processing or storing data (infrastructure-as-a-service, or IaaS, for example Amazon EC2).

Another distinction involves the control over the infrastructure. The infrastructure can be public, private, hybrid, managed, or community-owned. Each of these types involves particular decisions about who owns the infrastructure and who controls it. For example, in a managed cloud, a company owns its own IT infrastructure, but outsources the management to a third party. Obviously, public clouds – both owned and managed by third parties, and possibly accessible to anyone including competitors – are the trickiest ones security-wise. Companies therefore need to evaluate carefully which types of cloud services are suitable for their needs. Privacy legislation can play an important role here.

1.2.4 Security

In general, the transfer of information-related tasks to other parties obviously entails security risks, in terms of confidentiality, integrity and availability of the data and services. There are basically two ways to solve these: trust the provider, or put technical guarantees in place that establish security properties even if the provider

---

6 As discussed in Jean-Pierre Seifert’s presentation at SPCC.
7 As discussed in Filip Schepers’s presentation at SPCC.
is not trustworthy. Usually, only a combination of these is feasible, making cloud security an inherently socio-technical problem.\(^9\)

Tasks that can be performed in the cloud include storage, transfer and processing of information. With respect to storage and transfer, fairly standard security mechanisms can be applied. This does not mean that there cannot be security vulnerabilities, but the question on how to address those is mostly answered within the existing paradigm, called public key infrastructure. Data that is stored or transferred between parties is encrypted with the public key of the intended receiver, who decrypts it using her private key, when required. This means that the service provider that transmits or stores the data will not learn its contents, under the assumptions of the underlying cryptographic system.

For processing, there are no such standard solutions. In order to process data, it needs to be decrypted. Thus, whereas the communication between the data owner and the service provider can be considered secure, the service provider needs to access the plaintext data to do any meaningful processing. Therefore, in order to assume control over the security of the data, the owner needs to trust the service provider not to store the plaintext data, or transmit it to other parties.

In the following sections, we discuss a few noteworthy issues in cloud security and privacy, as apparent from SPCC and other cloud security venues.

### 1.3 The ideal of encrypted processing

Scientists have for a while sought for the ultimate solution to the security of information processing in an untrusted environment. The idea is that if we can build programs that operate on encrypted data, and produce an encrypted version of the correct output, then a service provider can perform calculations without having to possess the data in the clear. Such a system already exists for the simple operation of addition, and it is called homomorphic encryption. This means that if I provide encrypted versions of some numbers, then the system can produce the encrypted sum of the results without having to decrypt the data. Applications of this technique include counting votes in electronic elections.\(^{10}\) Before 2009, the same trick could be done for multiplication, but not with the same system. If we could do

---


both in the same system, we would have so-called *fully homomorphic encryption*, and it can be shown that such a system could perform arbitrary operations on encrypted data.\(^{11}\)

In 2009, the first fully homomorphic encryption system was proposed by Craig Gentry from IBM.\(^{12}\) Unfortunately, the efficiency of the system is so low that it cannot be used for any practical purpose. Whether scientific progress can yield a workable solution in the near future is doubtful. Also, there are security disadvantages to homomorphic encryption, notably the possibility to calculate encrypted versions of certain plaintexts without knowing the associated key. This has consequences for integrity and authenticity of data.

For now, we are stuck with solutions to secure cloud processing for limited situations. An example of this approach is searching in encrypted data.\(^{13}\) This means that we can store an encrypted database with a cloud provider, and we can search in the database without having to download the full database and decrypt it. We need only download and decrypt the results of the search. Similar solutions may be applicable for other specific cases of processing.

If encrypting is not feasible, we may at least wish to anonymise data before processing, in order to comply with privacy requirements. Techniques in this direction are being developed.\(^{14}\)

### 1.4 Putting physical limitations back in place

Another principal question in cloud computing relates to the technology of virtualisation. When physical machines are replaced by virtual machines, what does this mean for security?\(^{15}\) Similar questions have appeared in electronic voting (a physical ballot box versus a digital ballot box), and led to a lot of controversy about

---


whether digital voting systems would be able to meet the same criteria as paper-based ones. Especially, the scale of fraud has been a source of concern, as hacking a digital ballot box would potentially allow one to steal all the votes in an election, whereas one would have to physically manipulate the ballot box in each district in case of paper voting. Moreover, in case of Internet voting, a digital attack could be launched from any location.

The technology of virtualisation has raised similar concerns in relation to cloud computing. If a vulnerability would exist in the software that creates and manages virtual machines, the so-called hypervisor, this would enable an attacker to compromise virtual machines around the world. In particular, the question has often been raised whether it would be possible to access the data of another virtual machine from one’s own virtual machine, given that they run on the same physical computer.

Another type of concerns are those about location and time. On a privately owned device, one can be sure where the data is stored, from where it can be accessed, and when it will be made available, or made unavailable by deletion. These constraints are no longer present when data is stored in the cloud, and policies concerning location and time may need to be explicitly enforced. Such physical properties turn from inherent into imposed properties.\footnote{Blakley, R., The emperor’s old armor. In Proc. New Security Paradigms '96. ACM Press, 1997.}

For location, research is done into location-based or location-aware access control.\footnote{See e.g. Ardagna, C.A., Cremonini, M., Damiani, E., De Capitani di Vimercati, S. and Samarati, P., Supporting location-based conditions in access control policies. In Proc. of the ACM Symposium on Information, Computer and Communications Security (ASIACCS’06), pages 212-222, ACM, 2006.} In such an approach, requirements can be put in place for the location of the user as well as the location of the data. This means that someone may not be able to access the data if she is in the Netherlands, but it may also mean that someone may not be able to access the data if the data resides in the Netherlands. Such a policy can thus enforce accessibility only within certain jurisdictions. How to implement the appropriate mechanisms is another question, and for most systems reliance is necessary on secure sensing of location, as one can always try to fake the signal representing location information. Other approaches can measure the time of certain communications and thereby determine distance.\footnote{Pavlovic, D. and Meadows, C., Quantifying pervasive authentication: the case of the Hancke-Kuhn protocol. Technical Report No. RR-09-09. OUCL, 2009. http://www.comlab.ox.ac.uk/files/2437/RR-09-09.pdf (accessed May 10, 2010).}

For the property of time, the most notable development is that of secure deletion. Secure wiping of privately owned storage devices has existed for a while, but such mechanisms can of course not be enforced if one does not own the storage, as in cloud computing. In that case, one would have to trust that the provider would
do a secure wipe upon request, which will typically not be the case if storage is offered as a commodity.

The idea that has been developed is to encrypt the data, and make sure that the decryption key is destroyed after a certain time period. This destruction then has to be delegated to a trusted party,19 or based on some inherent property, like changing nodes in a peer-to-peer network.20 The most obvious limitation of the approach is that, in order to use the data, it needs to be decrypted. It can then also be stored in decrypted form, which would make the destruction of the key useless. Therefore, storage of the decrypted data should be discouraged to make the approach work, just like in the prevention of copying music with digital rights management.

The transition from inherent to imposed security for these physical properties introduces new risks, as the mechanisms now need to be designed and implemented rather than using e.g. physical distance as a safeguard.21 Vulnerabilities in the systems proposed are likely to be found. It remains therefore to be seen whether such approaches will really be used in practice.

1.5 Outsourced identity

One of the apparent challenges in offering IT as a service is the identification of users and administrators. In the past, it could be enforced that administrator access would only be granted with a local login. In the cloud, there is no such thing as “local”, and the administrator of a virtual machine is forced to log on remotely. Meanwhile, administrators of the provider will manage the physical machines, giving rise to multiple levels of administrator roles.22 Security risks have always been for a significant part due to employees of the organisation itself, with administrators as a prominent example. The entanglement

---


of their responsibilities in the cloud also leads to new types of insiders, who have
or can obtain the necessary credentials to endanger information security.
Insiders are no longer only found within a single organisation, but spread among
business partners and service providers, such as those of cloud services.

It is therefore important to develop access control models and accountability
mechanisms that can establish strong links between identity and activity, when
necessary. We see the development of such architectures for example in Electronic
Health Record (EHR) systems. Here, the right kind of authorisation is in principle
necessary to obtain sensitive data. However, in case of an emergency, such a
mechanism needs to be overridden, and accountability is established after the fact.

Apart from issues of authentication, cloud computing also raises questions con-
cerning the nature of identity itself. Cloud use is not limited to companies. In-
creasingly, individuals store their information online instead of on their own de-
vices. Identities are managed in Facebook, phone companies store users’ SMS
messages online and different forms of online cooperation are offered by Google
and others. Considering the high availability rates of the Internet nowadays and
the professional backup system usually connected to online storage, these services
may indeed be quite attractive.

All this comes at a price, though, which is called loss of control. Unlike in busi-
nesses, control may not be such an explicit consideration for individuals. Indeed,
not being in control may relieve the individual of unnecessary burdens. However,
it is generally not the case that individuals would allow service providers to do
with their data what they like. There are limitations based on privacy laws, but
even then, users may wish to explicitly weigh the advantages of online storage and
processing against the exposure of their data.

For example, many Dutch citizens have objected to the online transmission of
their health data via the electronic patient file system. Facebook users have forced
the provider to adapt default privacy settings and even shut down a new service
that allowed friends to track one’s shopping behaviour. And Google has explained
that the target advertisements popping up with users’ e-mails do not actually in-
volve human reading of the e-mail contents – for what it’s worth, for who has ac-
cess to the machines actually processing these e-mails to match the ads?

One the one hand, individual use of cloud services seems to be merely a matter
of convenience. If I can more reliably and more easily work with my information
online, then why wouldn’t I do so? On the other hand, the tendency to put per-

\begin{itemize}
\end{itemize}
sonal information online for other purposes than processing it oneself is certainly new, and has to be evaluated more carefully. The use of Facebook, Twitter, LinkedIn is not comparable to other forms of cloud computing by individuals, as it concerns deliberate publication of personal information for social reasons.

Two trends have come together to produce the latter phenomenon. The first is the demise of traditional group memberships in society (called “ontzuiling” in Dutch, meaning “de-pillerisation”). Where individuals were born into or easily led into certain groups, based on family, gender, class, et cetera, this is now much less straightforward. The second is the rise of the Internet and associated cloud services. Combining the two, online services like Facebook have emerged to provide online management of one’s identity. One now needs to explicitly group oneself, and social networking services are a tool to do precisely this. We call this phenomenon outsourced identity.

Three combined characteristics of outsourced identity distinguish it from earlier forms of external identity (diaries, village gossip, etc.).\(^{25}\) Firstly, the outsourcing is intentional; secondly, the information is public or semi-public; and thirdly, the form of the information makes it easy to de-contextualise it.\(^{26}\)

The consequences of this phenomenon are that the problems of profiling associated with online information about individuals cannot only be discussed in terms of social sorting. In outsourced identity, individuals want to get sorted. It is therefore necessary to distinguish good from bad sorting. The question is how we could make this distinction.

Informed consent may be the key concept here: if individuals are allowed to deny group membership, they maintain their autonomy in the face of undesirable assignments to groups. This of course implies that 1) they know about the group membership and 2) the membership can (at least theoretically) be denied. The former condition does not hold if decisions are being made in secret or based on secret profiling information; the latter does not hold if it concerns conditions such as age, race or gender, but also probabilistic groups like “those who are likely not to pay their bills”.

There is thus a difference in protecting privacy in the cloud from a company perspective and from an individual’s perspective. From a company perspective, privacy-sensitive data needs to be processed in accordance with privacy laws, and the confidentiality of this data should be guaranteed as far as possible. This should then make sure that people do not suffer from undesirable consequences of loss of privacy, in particular in terms of discrimination. From an individual perspective, people use the cloud and its applications precisely to discriminate themselves; to


make themselves stand out from the crowd and be assigned to groups simultaneously. This outsourced identity cannot be addressed from the same perspective as privacy in traditional applications, and care is needed when judging the phenomenon from an ethical point of view.

1.6 Informational precaution

In the face of the development towards cloud computing, the need to separate pieces of information becomes more and more profound. If information is not properly secured, some parties may become too powerful, and trust relations may be broken. Unfortunately, the role of information security in providing this separation is poorly understood. On the one hand, technical solutions are being developed to secure information; on the other hand, policies are developed without much regard to what is technically possible, and sometimes policies seem to run ahead of the possibility of function creep, by allowing all kinds of extended usage scenarios on forehand. Substantial future research is needed to clarify the relation between technical forms of information security and societal goals, both from a social science and from a philosophical perspective.

In the meantime, we need a means to convince policymakers, both in government and in industry, of the real ethical dimensions of the issues. Implementing another distributed IT application is not just an increase in convenience; we are really changing the world here. Where environmental concerns are high on the agenda, the data protection community has not yet succeeded in providing a clear vision on the need of protecting something like information, with a view to social effects of information processing. One approach to bridge this gap would be the translation of norms from environmental ethics to information technology. Earlier, we have proposed to do this for the precautionary principle.\(^{27}\) This principle states that parties should refrain from actions in the face of scientific uncertainties about serious or irreversible harm to public health or the environment. It further holds that the burden of proof for assuring the safety of an action falls on those who propose it.\(^{28}\) The precautionary principle has been applied successfully in the European Union, but is less


popular in the United States, as it obviously implies government interference with what is desirable and what is not.

A similar baseline for computer ethics as something that contributes to sustainability does not exist yet. Even if the precautionary principle is not universally accepted in environmental ethics, it does provide a basis from where to start discussions on new technologies. Generalised to information technology, it can serve as a trigger for governments to at least consider the social implications of IT developments.

Whereas the traditional precautionary principle targets environmental sustainability, informational precaution would target social sustainability.\(^{29}\) Clear definitions of this concept are however lacking, and more precise definitions are necessary. One could say that social sustainability relies on maintaining stable trust and power relations in society, and information is a key asset there. With that in mind, information security can indeed contribute to social sustainability, and is therefore an indispensable feature in ethical cloud scenarios – including government initiatives.

These considerations are part of larger-scale developments, and disappearing boundaries do not only occur in information security. The precautionary principle may play a role for society in general in dealing with these dependencies. As such, the cloud is an instance of our self-created dependence upon large-scale infrastructures, and thereby on the ethical behaviour of those who manage them. Precaution can serve to enforce this ethical behaviour, by designing technology such that it stimulates the right kinds of actions.\(^{30}\)

However, also the precautionary principle itself needs to be treated with precaution. It easily becomes another tool in the hand of a few to control the many, and some authors even criticize what they call a “precaution state”.\(^{31}\) Therefore, informational precaution deserves our further attention and discussion in the cloud era.


1.7 Conclusions

In cloud computing, information storage, transmission and processing are purchased as a commodity from a service provider. Although security issues in storage and transmission can be addressed to a reasonable extent using standard tools, protecting data being processed is another story. Although fully homomorphic encryption would enable processing of encrypted data, the theoretical breakthrough of 2009 is far from practical. Therefore, efforts need to be put into securing data processing in limited cases, such as searching in encrypted data.

Another fundamental issue in cloud computing is the abolishment of physical constraints that helped securing the data in the past. Like in a transition from paper voting in polling stations to Internet voting, this brings additional security challenges, and it is not even clear beforehand that all of these can be solved. In order to simulate physical constraints in a virtualised infrastructure, proposals like location-based access control and secure deletion have been put forward.

Meanwhile, the increasing outsourcing and specialisation not only affects companies, but also individuals. In this case, cloud computing not only means a different way for individuals to store their data, but also a different way to manage the external representations of their identity. In this “outsourced identity”, people intend to get themselves socially sorted, and social sorting cannot be seen as problematic in and of itself. Instead, rules should be put in place to distinguish good from bad sorting, notably by requiring informed consent.

In order to provide ethical foundations for the complex interaction of governments, industries and the social environment in the age of cloud computing, research is needed in the relation between information technology and social sustainability. The precautionary principle may serve as a tool to translate successful approaches from environmental ethics to information ethics in the cloud.

1.8 References


