2nd IEEE International Conference on Cloud Computing Technology and Science

Privacy, Security and Trust Issues Arising from Cloud Computing

Siani Pearson and Azzedine Benameur
Cloud and Security Research Lab
HP Labs
Bristol, UK
{Siani.Pearson, Azzedine.Benameur}@hp.com

Abstract— **Cloud computing is an emerging paradigm for large scale infrastructures. It has the advantage of reducing cost by sharing computing and storage resources, combined with an on-demand provisioning mechanism relying on a pay-per-use business model. These new features have a direct impact on the budgeting of IT budgeting but also affect traditional security, trust and privacy mechanisms. Many of these mechanisms are no longer adequate, but need to be rethought to fit this new paradigm. In this paper we assess how security, trust and privacy issues occur in the context of cloud computing and discuss ways in which they may be addressed.**

Keywords-cloud computing; privacy; security; risk; trust

I. INTRODUCTION

A cloud may be thought of as a large pool of resources, unified through virtualization or job scheduling techniques; these resources can be managed to dynamically scale up to match the load, using a pay-per-resources business model. These resources are made available through a new cloud computing paradigm that is being increasingly adopted by organizations; the resources include hardware and systems software on remote datacenters, as well as services based upon these that are accessed through the Internet. Key features advertised are elasticity, multi-tenancy, maximal resource utilization and pay-per-use. These new features provide the means to leverage large infrastructures like data centres through virtualization or job management and resource management, but these large pools of resources are not necessarily located in the same country nor even on the same continent. Furthermore, the dynamic expansion or shrinkage of a cloud makes it difficult to keep track of what resources are used and in which country. This makes compliance with regulations related to data handling difficult to fulfill. Auditing is also a challenging task due to the volatility of the resources used. These new features make it hard – and sometimes not possible at all – to reuse traditional security, trust and privacy mechanisms in the cloud. Moreover they raise issues and concerns that need to be fully understood and addressed. Some of these issues will be shared with other paradigms, such as service-oriented architectures (SOA), grid, web-based services or outsourcing, but often they are exacerbated by cloud.

Current cloud services pose an inherent challenge to data privacy, because they typically result in data being present in unencrypted form on a machine owned and operated by a different organization from the data owner. There are threats of unauthorized uses of the data by service providers and of theft of data from machines in the cloud. Fears of leakage of sensitive data [1] or loss of privacy are a significant barrier to the adoption of cloud services [2]. These fears may be justified: in 2007, criminals targeted the prominent cloud service provider (CSP) Salesforce.com, and succeeded in stealing customer emails and addresses using a phishing attack [3]. Moreover, there are laws placing geographical and other restrictions on the processing by third parties of personal and sensitive information. These laws place limits on the use of cloud services as currently designed.

We now consider the privacy, security and trust issues associated with cloud computing in more detail, together with related legal concerns. There is necessarily some overlap and interdependency between such issues, but nevertheless we believe it is still helpful to categorise these issues in this way. In doing this, we take into account the alternative delivery and deployment models for cloud computing, as these influences the risks involved. After this, we look at how the aforementioned issues might be addressed.

II. PRIVACY ISSUES

At the broadest level (and particularly from a European standpoint), privacy is a fundamental human right that encompasses the right to be left alone. In the commercial, consumer context, privacy entails the protection and appropriate use of the personal information of customers, and the meeting of expectations of customers about its use. For organisations, privacy entails the application of laws, policies, standards and processes by which Personally Identifiable Information (PII) of individuals is managed.

When considering privacy risks in the cloud, context is very important as privacy threats differ according to the type of cloud scenario. Some cloud application areas and services might face a very low privacy threat, for example if the service is to process information that is (or is very shortly to be) public. It is only if the service handles personal information, in the sense of collecting, transferring, processing, sharing or storing it, that there could be a privacy risk and privacy needs to be taken into account. However, services that are dynamically personalized – based on people's location, preferences, calendar and social networks, would require privacy to be taken into account a great deal, as the potential risk is high.

978-0-7695-4302-4/10 $26.00 © 2010 IEEE
DOI 10.1109/CloudCom.2010.66

Such services could for example have some sort of embedded tracking and profiling, with inter-device communication and mechanisms to customize the environment and services based on actual individual behaviour.

Public cloud (accessed through the Internet and shared amongst different consumers) is the most dominant architecture when cost reduction is concerned, but relying on a cloud service provider (CSP) to manage and hold your data raises a great many privacy concerns. In the remainder of this section we consider a number of aspects that illustrate best the privacy issues in public cloud: lack of user control, potential unauthorized secondary usage, data proliferation, transborder data flow and dynamic provisioning. Other privacy issues include retention of data and who controls that, ensuring that data has actually been properly destroyed in the cloud (as considered further in subsection IIIB), how to know that privacy breaches have occurred and how to determine who is at fault in such cases. Sensitive data provides a special case: as considered further in section IV, there are special laws concerning treatment of sensitive data, and data leakage and loss of privacy are of particular concern to users when sensitive data is processed in the cloud. Currently this is so much of an issue that the public cloud model would not normally be adopted for this type of information.

A. Lack of User Control

User-centric control seems incompatible with the cloud: as soon as a SaaS environment is used, the service provider becomes responsible for storage of data, in a way in which visibility and control is limited. So how can a consumer retain control over their data when it is stored and processed in the cloud? This can be a legal requirement and also something users/consumers want – it may even be necessary in some cases to provide adequate trust for consumers to switch to cloud services. In cloud computing, consumers' data is processed in 'the cloud' on machines they do not own or control, and there is a threat of theft, misuse (especially for different purposes from those originally notified to and agreed with the consumer) or unauthorized resale. In addition, it is not clear that it will be possible for a CSP to ensure that a data subject can get access to all his/her PII, or to comply with a request for deletion of all his/her data. It can also be difficult to get data back from the cloud, and avoid vendor lock-in.

B. Unauthorised Secondary Usage

There is a risk that the data may be put to unauthorized uses. It is part of the standard business model of cloud computing that the service provider may gain revenue from authorized secondary uses of users' data, most commonly the targeting of advertisements. However, some secondary data uses would be very unwelcome to the data owner (such as, for example, the resale of detailed sales data to their competitors). At present there are no technological barriers to such secondary uses.

Furthermore, there is the related issue of financial resilience of the CSPs: for example, risk of vendor demise, and what would happen to the data held by a cloud computing provider that becomes bankrupt or is acquired by another company.

C. Data Proliferation and Transborder Data Flow

Data proliferation is a feature of cloud and this happens in a way that may involve multiple parties and is not controlled by the data owners. CSPs ensure availability by replicating data in multiple datacenters. It is difficult to guarantee that a copy of the data or its backups are not stored or processed in a certain jurisdiction, or that all these copies of data are deleted if such a request is made. This issue is considered further in section III.

Movement of data onto the cloud and potentially across and between legal jurisdictions, including offshoring of data processing, increases risk factors and legal complexity [4,5]. Governance and accountability measures also become more complex as processes are outsourced and data crosses organisational boundaries [6]. The risks that can arise from choosing the wrong business partner can be daunting and very difficult to assess, especially in cloud based environments, where even knowing the jurisdictions involved can be quite difficult [7]. Issues of jurisdiction (i.e. about whose courts would hear a case), which law applies and about whether a legal remedy can be effectively enforced need to be considered [8]. As considered also in the following subsection, a cloud computing service which combines outsourcing and offshoring may raise very complex issues [9]. Hence, it can be difficult to ascertain privacy compliance requirements in the cloud. It can also be possible to violate local laws when transferring data stored in the cloud: cloud computing exacerbates the transborder data flow issue because it can be extremely difficult to ascertain which specific server or storage device will be used, due to the dynamic nature of cloud computing.

D. Dynamic Provisioning

Cloud computing faces many of the same problems as traditional outsourcing, yet the dynamic nature of cloud makes many existing provisions to address this in more static environments obsolete or impractical to set up in such a short timescale. It is not clear which party is responsible (statutorily or contractually) for ensuring legal requirements for personal information are observed, or appropriate data handling standards are set and followed [10], or whether they can effectively audit third-party compliance with such laws and standards. Neither is it yet clear to what extent cloud sub-contractors involved in processing can be properly identified, checked and ascertained as being trustworthy, particularly in a dynamic environment. It is also unclear what rights in the data will be acquired by data processors and their sub-contractors, and whether these are transferable to other third parties upon bankruptcy, takeover, or merger [11].

III. Security Issues

In traditional security models, a security perimeter is set up to create a trust boundary within which there is self-

control over computing resources and where sensitive information is stored and processed. For example, the corporate firewall often marks this boundary. The network provides transit to other trusted end hosts, which operate in a similar manner. This model held for the original Internet, but does not for public and hybrid cloud (a mixture between public and private deployment). The security perimeter becomes blurred in the sense that confidential information may be processed outside known trusted areas as these computing environments often have fuzzy boundaries as to where data is stored and/or processed. On the other hand, in order to obtain the service, consumers need to extend their trust to the cloud service provider, and so this can provide a point of friction, as considered further in section IV.

Public cloud not only raises privacy issues as discussed previously, but also has its share of security concerns. Indeed, security was rated the top challenge of the cloud model in a recent user survey [12]. In this section we present problems that are of high importance for cloud architectures. Private clouds (i.e. restricted to private networks) can to a certain extent guarantee security levels, but the economic costs associated with this approach are relatively high.

At the network, host and application levels, security challenges associated with cloud computing are exacerbated by cloud computing but not specifically caused by it. The main issues relate to defining which parties are responsible for which aspects of security. This division of responsibility is hampered by the fact that cloud APIs are not yet standardized. Customer data security raises a number of concerns, including the risk of loss, unauthorized collection and usage, and generally the CSP not adequately protecting data. There is not the space in this paper to give a comprehensive review of cloud security issues but we do consider here many of the main issues: for further details see for example [13]. More broadly speaking, security risks fit into a broader model of cloud-related risks; for example, according to the Cloud Security Alliance [1], the top threats to cloud computing are: abuse and nefarious use of cloud computing, insecure interfaces and APIs, malicious insiders, shared technology issues, data loss or leakage, account or service hijacking and unknown risk profile. They were unable to reach a consensus on ranking the degree of severity of these risks.

Overall, it may be noted that security need not necessarily suffer in moving to the cloud model, because there is scope for security to be outsourced to experts in security and hence in many cases greater protection than previously can be obtained.

A. Access

Cloud computing may increase the risk of access to confidential information.

First, this may be by foreign governments: there can be increased risks due to government surveillance over data stored in the cloud, as the data may be stored in countries where previously it was not. Governments in the countries where the data is processed or stored may even have legal rights to view the data under some circumstances [14,15], and consumers may not be notified if this happens.

Second, as with other computing models, there is an underlying risk of unauthorised access that may be exacerbated if entities are involved in the provider chain that have inadequate security mechanisms in place (e.g. if they have inadequate vetting of internal IT staff who have highly privileged access). The risk of data theft from machines in the cloud can be by rogue employees of CSPs or by data thieves breaking into service providers' machines, or even by other customers of the same service if there is inadequate separation of different customers' data in a machine that they share in the cloud.

In general, cloud storage can be more at risk from malicious behaviour than processing in the cloud, because data may remain in the cloud for long periods of time and so the exposure time is much greater. On the other hand, there is more potential for usage of encryption in cloud storage, as considered further below.

B. Control over Data Lifecycle

Another major issue for cloud is to ensure that the customer has control over the lifecycle of their data, and in particular deletion, in the sense of how to be sure that data that should be deleted really are deleted and are not recoverable by a cloud service provider. There are currently no ways to prove this as it relies on trust, and the problems is exacerbated in cloud because there can be many copies of the data, potentially held by different entities.

More specifically, this risk depends very much on the cloud service model being used. Using Infrastructure as a Service (IaaS) or Platform as a Service (PaaS), one or more virtual machines are created in order for a program to be run within those – when the task is finished, the virtual machines and the disk space are released. However, the media may not be wiped – if it is not, it may be possible for the next user to recover the previous user's data. As space is virtually allocated users do not know what happens to the physical volume supporting their virtual storage. Using a Software as a Service (Saas) approach, the customer is one of the users of a multi-tenant application developed by the CSP, and the customers' data is stored in the cloud, to be accessible the next time the customer logs in. The data would only be deleted at the end of the lifecycle of the data, if the customer wishes to change service provider, etc.

C. Availability and Backup

It is not easy to guarantee adequate availability and backup in the cloud. When data are hosted remotely in the cloud, backup is critical for businesses to recover in case of failure. But cloud providers enforcing resilience of their infrastructure might rely on seamless backups. This is a high security issue as these backups might be done without the customer's active informed consent and could lead to serious threats from an insider or external attacker. One of the top threats identified by CSA [1] is 'data loss or leakage', where records may be deleted or altered without

a backup of the original content. A record might be unlinked from a larger context, making it unrecoverable, data could be stored on unreliable media and if there is a key management failure then data could be effectively destroyed.

The autonomic aspect of cloud also poses new risks, namely self-optimization and self-healing. Self-optimization grants a degree of autonomy in decision making, e.g. automatically adapting services to meet the changing needs of customers and service providers; this challenges enterprises' abilities to maintain consistent security standards. Self-healing allows CSPs to provide appropriate business continuity, recovery and back-up, but it may not be possible to determine with any specificity where data processing takes place within the cloud [16]. Autonomic aspects of cloud computing are one of its assets but need to be tailored to be compliant with privacy and legal issues.

D. Lack of Standardization

All paradigms struggle when emerging, mainly because of a lack of standardization. Grid computing failed to gain rapid adoption for Virtual Organisations because of this. After some time middleware for grid interoperability made it this possible. SOA on the other hand tried to correct past mistakes by establishing standards as early as the concepts. However, it struggled because of too much standardisation, ending up with more than 30 WS-* standards – making it hard to understand and even having some overlapping standards. Cloud computing, as of today, lacks interoperability standards. There is no standardised communication between and within cloud providers and no standardized data export format which makes it difficult to leave a cloud provider. The lack of standards also makes it difficult to establish security frameworks for such heterogeneous environments and forces people for the moment to rely on common security best practice.

E. Multi-Tenancy

Multi-tenancy is an architectural feature whereby a single instance of software runs on a SaaS vendor's servers, serving multiple client organizations. The software is designed to virtually partition its data and configuration so that each client organization works with a customized virtual application instance.

As considered above in section B, the cloud service model affects security risks: in particular, in the SaaS model, customers are users of multi-tenant applications developed by CSPs, it is likely that personal data and even financial data are stored by CSP in the cloud, and it is the responsibility of the CSP to secure the data.

Some providers use job scheduling and resources management [17], but most cloud providers use virtualization to maximize hardware utilisation. Virtual machines (VMs) are sandboxed environments and therefore completely isolated from each other. This assumption makes it safe for users to share the same hardware. However, this security can sometime break down, allowing attackers to escape the boundaries of this sandboxed environment and have full access to the host [18]. The use of virtualisation can introduce new security vulnerabilities, such as cross-VM side-channel attacks to extract information from a target VM on the same machine [19]. Therefore strong emphasis on virtualization software is needed.

F. Audit

CSPs need to implement internal monitoring controls, in addition to an external audit process. The cloud computing environment presents new challenges from an audit and compliance perspective, but existing solutions for outsourcing and audit can be leveraged. Transactions involving data that resides in the cloud need to be properly made and recorded, in order to ensure integrity of data and the data owner needs to be able to trust the environment that no untraceable action has taken place. However, provision of a full audit trail within the cloud, particularly in public cloud models, is still an unsolved issue. In addition, transactional data is a byproduct with unclear ownership, and it can be hard to anticipate which data to protect, as even innocuous-seeming data can turn out to be commercially sensitive [6].

IV. TRUST ISSUES

The speed and flexibility of adjustment to vendor offerings, which benefits business and motivates cloud computing uptake, brings a higher risk to data privacy and security. This is a key user concern, particularly for financial and health data, and the associated lack of trust can be a key business inhibitor for cloud computing in domains where confidential or sensitive information is involved.

Since customers lack control of cloud resources, they are not in a good position to utilize technical mechanisms in order to protect their data against unauthorized access or secondary usage or other forms of misuse. Instead, they must rely on contracts or other trust mechanisms to try to encourage appropriate usage, in combination with mechanisms that provide compensation in the event of a breach, such as insurance, court action, or penalties for breach of service level agreements (SLAs).

Trust is a complex concept for which there is no universally accepted scholarly definition. Evidence from a contemporary, cross-disciplinary collection of scholarly writing suggests that a widely held definition of trust is as follows [20]: "Trust is a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviour of another". Yet this definition does not fully capture the dynamic and varied subtleties involved: trust is a complex notion and a multi-level analysis is important in order to try to understand it. There are many different ways in which on-line trust can be established: security may be one of these (although security, on its own, does not necessarily imply trust [21]). Some would argue that security is not even a component of trust: Nissenbaum argues that the level of security does not affect trust [22]. On the other hand, an

example of increasing security to increase trust comes from people being more willing to engage in e-commerce if they are assured that their credit card numbers and personal data are cryptographically protected [23].

Another component of online trust is reputation. Reputation is perhaps a company's most valuable asset [22] (although of course a CSP's reputation may not be justified). Brand image is associated with trust and suffers if there is a breach of trust or privacy.

There can be differing phases in a relationship such as building trust, a stable trust relationship and declining trust. Trust can be lost quickly: as Nielsen states [24]: "It [trust] is hard to build and easy to lose: a single violation of trust can destroy years of slowly accumulated credibility".

When assessing trust in relation to cloud computing, it may be useful to distinguish between social and technological means of providing persistent and dynamic trust, as all of these aspects of trust can be necessary [25]. *Persistent trust* is trust in long-term underlying properties or infrastructure; this arises through relatively static social and technological mechanisms. *Dynamic trust* is trust specific to certain states, contexts, or short-term or variable information; this can arise through context-based social and technological mechanisms.

Persistent social-based trust in a hardware or software component or system is an expression of confidence in technological-based trust, because it is assurance about implementation and operation of that component or system. In particular, there are links between social-based trust and technological-based trust through the vouching mechanism, because it is important to know who is vouching for something as well as what they are vouching; hence social-based trust should always be considered.

In assessing cloud computing provision, mechanisms to provide dynamic technological-based trust need to be used in combination with social and technological mechanisms for providing persistent trust: if software processes provide information about the way in which information is stored, accessed and shared within a cloud, that information can only be trusted if entities that are trusted vouch for the method of providing the information and assessing the information. Depending upon the context, these entities could be consumer groups, auditors, security experts, regulators, companies with proven reputation, established CSPs, etc.

A. Weak Trust Relationships

Trust relationships at any point in the cloud service delivery chain may be weak, but exist in order that a service can be provided quickly. Significant business risk may be introduced in a way that is not transparent when a cloud transaction is initiated, due to loss of control in passing sensitive data to other organisations and the globalised nature of cloud infrastructure. Organisations that contract out key business processes may not even know that contractors are sub-contracting, or even if they do, contract requirements regarding data protection measures may not be propagated down the contracting chain.

Trust along the chain from the customer to cloud providers at all levels may be non-transitive, and in particular the customer may not trust some of the subcontractors (XaaS providers). Indeed, due to a lack of transparency they may not even be aware of the identity of the cloud providers in this chain. In particular, 'on-demand' and 'pay-as-you-go' models may be based on weak trust relationships, involve third parties with lax data security practices, expose data widely, and make deletion hard to verify. In order to provide extra capacity at short notice or in real time, new providers could be added to the chain for which there is not sufficient chance to make adequate checks about their identity, practices, reputation and trustworthiness.

B. Lack of Customer Trust

When it is not clear to individuals why their personal information is requested, or how and by whom it will be processed, this lack of control and lack of visibility of the provider supply chain will lead to suspicion and ultimately distrust [26]. There are also security-related concerns about whether data in the cloud will be adequately protected [10]. As a result, customers may hold back from using cloud services where personally identifiable information is involved, without an understanding of the obligations involved and the compliance risks faced, and assurance that potential suppliers will address such risks. This is particularly the case where sensitive information is involved, for example financial and healthcare information.

Ultimately, usage of the cloud is a question of tradeoffs between security, privacy, compliance, costs and benefits. Trust is key to adoption of SaaS, and transparency is an important mechanism. Furthermore, trust mechanisms need to be propagated right along the chain of service provision.

V. LEGAL ASPECTS

Legal frameworks have been instrumental and key to the protection of users' personal and sensitive information. For example, in Europe there is national legislation based upon EU Directive, in US there is a patchwork of legislation according to sector, information and/or geographical area, and in many other countries worldwide analogous frameworks apply. The fundamental concepts of such frameworks are in the main technology neutral, and their validity would still apply to cloud computing. Nevertheless, such frameworks – along with the associated tools, advice and national legislation – need to be constantly updated and adjusted with current and future technologies in mind. There is currently a dialogue between organizations, regulators and stakeholders to ensure that the regulatory framework does adapt to new frameworks and business models without eroding consumers' trust in the systems that are deployed. In particular, the dynamically changing nature of cloud computing, potentially combined with cross-jurisdictional

interactions, introduce legal aspects that need to be carefully considered when processing data. In this section we discuss these aspects.

First of all, not knowing which routes transnational traffic will take makes it very difficult to understand the particular laws which will apply; furthermore, it is difficult to know the exposure of the data transferred, because information passing through some countries (including US, as permitted by the US Patriot Act) can be accessed by law enforcement agencies. Even if this is known, because of the global nature of cloud computing and the many legislations in place around the world, it can be complex and difficult to ensure compliance with all the legislation that may apply in a given case.

Second, it is a common requirement under the law that if a company outsources the handling of personal information or confidential data to another company, it has some responsibility to make sure the outsourcer uses "reasonable security" to protect those data. In the case of cloud computing, the CSP needs to implement "reasonable security". Different companies may be involved in the cloud supply chain, and this can make it difficult to ensure that such security is provided all the way along the chain. At present, clients often only know the initial CSP and the standard terms and conditions of cloud computing service providers do not include any clauses ensuring the level of security provided: they provide no guarantee as to the security of data, and even deny liability for deletion, alteration or loss related to data that is stored. As current terms of service are very much set in favour of the CSP [9], if anything goes wrong it is often the customer that will be made liable.

Another aspect is litigation: a CSP may be forced to hand over data stored in the cloud, as illustrated by the US vs. Weaver case [27], where Microsoft was requested via a trial subpoena rather than a warrant to provide e-mails handled by their Hotmail service. A government only needs to show the requested material is relevant to the case for a subpoena, whereas for a warrant, probable cause must be demonstrated. In order to avoid a similar situation occurring with non-governmental entities, subscribers to cloud services could include contractual provisions in the service agreement that govern the CSP's response to any subpoena requests from such entities.

Location matters from a legal point of view as different laws may apply depending on where information exists, but in cloud computing the information might sometimes be in multiple places simultaneously, it may be difficult to know exactly where it is or it may be in transit. As already discussed, a complicating factor is that there are multiple copies of data located in the cloud. Furthermore, these copies can be managed by different entities: a back-up SP, a provider used to respond to peak capacity needs, specialised services, etc.

Putting data in the cloud may impact privacy rights, obligations and status: for example it may make it impossible to comply with some laws such as the Canadian Privacy Act or health laws. Legal protection can be reduced, and trade secrets may be impacted.

Since cloud technology has moved ahead of the law, there is much legal uncertainty about privacy rights in the cloud and it is hard to predict what will happen when existing laws are applied in cloud environments.

Nevertheless, there are existing legal constraints on the treatment of users' private data by cloud computing providers. Privacy laws vary according to jurisdiction, but EU countries generally only allow personally-identifiable information to be processed if the data subject is aware of the processing and its purpose, and place special restrictions on the processing of sensitive data (for example, health or financial data), the explicit consent of the data owner being part of a sufficient justification for such processing [28]. They generally adhere to the concept of *data minimization*, that is, they require that personally identifiable information is not collected or processed unless that information is necessary to meet the stated purposes. In Europe, data subjects can refuse to allow their personally identifiable data to be used for marketing purposes [29]. Moreover, there may be requirements on the security and geographical location of the machines on which personally identifiable data is stored [30]. European law limiting cross-border data transfers also might prohibit the use of cloud computing services to process this data if data would be stored in countries with weak privacy protection laws [31].

Furthermore, as discussed in subsection IIB, it is difficult to enforce transborder data flow regulations within the cloud. Cloud computing can exacerbate the problem of knowledge of geographic location of where cloud computing activities are occurring, as due to its dynamic nature this can be extremely difficult to find out.

Finally, not least because many CSPs rely upon secondary use of data as part of their revenue, it will be necessary for consumers and CSPs to make legally binding agreements as to how data provided to CSPs may be used. As we consider further in the following section, it is likely that in future such agreements might be enforceable in a technological sense. This will help enhance trust and mitigate the effects of the blurring of security boundaries considered above.

In general, the legal situation is subject to change: legislation has not yet been updated to address the challenges above and courts have not yet ruled many cases specifically related to cloud computing.

VI. Addressing Privacy, Security and Trust Issues

We previously presented several issues about cloud computing. In this section we present some initial steps that help address these concerns.

A. Data Handling Mechanisms

In order to help protect sensitive corporate and customer data, an organization considering using cloud services should take the following procedural measures: classifying its information assets to clarify which ones are confidential; before selecting a CSP, determining their data protection and business continuity capabilities; once a

CSP is determined, making sure that certain aspects are explicitly stated within negotiated agreements; where appropriate, trying to restrict transfer of data to CSP to the non-confidential data. In particular, these contractual aspects would include:

- stating the CSP's obligations to protect the organisation's data (based on the organisation's privacy policies)
- holding the CSP liable for failure to satisfy those obligations
- clarifying in what geographical regions the data (including replication, backup and others) is located, and requiring notification before any changes in that situation.
- requiring the CSP to comply with applicable data protection and privacy laws
- clarifying ownership of the data and what will happen when the agreement ends
- clarifying what will happen if data is lost
- defining policies for data retention and destruction

B. *Data Security Mitigation*

To a large extent, privacy and security issues go away if data is not revealed in the clear within the cloud. Encrypting personal data is feasible, and strongly advisable, if using an IaaS cloud service for simple storage. However, data-at-rest used by a cloud-based application is generally not encrypted, because encryption would prevent indexing or searching that data. Moreover, the data cannot be encrypted if processed in the cloud, as it is not yet possible to process encrypted data in an efficient way. It is possible to do this however in a non-efficient way, notably via Yao's protocol for secure two-party computation [32] and Gentry's fully homomorphic encryption scheme [33]. In a few years' time it is likely that the latter would provide the basis for a practical commercial means of data security mitigation. In the meantime, and as an alternative approach, for some applications a trade-off can be made of efficiency against security, so that it is possible to obfuscate some of the data before transferring it to the cloud, with the degree of this obfuscation dependant upon the context [34]. Unlike Yao and Gentry's approaches, with this method it is not necessary to assume that client devices are able to carry out a large amount of storage or computation, nor that cloud computing providers are willing to rewrite their applications or to translate them into binary circuits. However, this approach does not fit all types of cloud application, so in general customers will still need to pay close attention to the security of their data in the cloud.

Another approach is to run cloud applications without putting sensitive data in the cloud. An example is Google Secure Data Connector [35], which allows programs developed by an organization (based upon analysis of the data structures involved) to access information behind its firewall.

C. *Design for Privacy*

Current privacy concepts such as the Fair Information Principles [36] are applicable to cloud computing scenarios and can mitigate the risks considered in section II, but it is necessary to implement mechanisms within the cloud to underpin these concepts. An initial framework was provided in [37], building upon the generic approach of design for privacy (see for example [38]), and subsequently this approach has been refined for cloud computing [39].

D. *Standardisation*

Standardised solutions and options for cloud that are currently missing include provision of trust, assurance and audit frameworks including associated APIs to expose what CSPs are actually doing, i.e. what's going on 'behind the scenes': adequate assurance needs to be given about the way in which cloud providers process and protect data, and data tracking mechanisms need to be provided along the chain of provision. Standardisation is needed for the assessment of CSPs both by customers and by other CSPs. Currently there are no standards for how data is stored, access controls, performance metrics, etc. Vendors, analysts and security professionals are currently trying to work on developing such standards, including for SLAs.

Standardisation also links in with the discussion on trust in section IV. Due to a lack of information and time, together with the huge complexity of IT security, it is impossible for cloud consumers to themselves identify the level of security offered by individual CSPs. They need to rely upon the reliability of that provider being assessed by experts via evaluation and certification procedures, and they can also take into account reputation gathered and conveyed via electronic systems. This type of approach is currently being improved for cloud computing: for example, ENISA is developing standards for security assessment of cloud providers [40], and other parties are working on development of reputation systems to assess cloud providers.

E. *Accountability*

For businesses having data lost, leakage or privacy violation is catastrophic but what could be worse is to have no clear entity to blame for it. Accountability in the cloud is a very important concept that needs to be supported from both a legal and technical viewpoint. The way to achieve accountability in such a dynamic and worldwide infrastructure is to have a strong emphasis on auditing. Audit should be able to keep track of where the data has been outsourced, who processed it and for what purpose. These steps are essential in ensuring accountability and gaining user trust.

Solutions to privacy risks in the cloud involve reintroducing an element of control. For the corporate user, privacy risk in cloud computing can be reduced if organisations involved in cloud provision use a combination of privacy policies and contractual terms to create accountability in the form of transparent, enforceable commitments to responsible data handling.

Specifically, accountable organisations will ensure that obligations to protect data (corresponding to user, legal and company policy requirements) are observed by all processors of the data, irrespective of where that processing occurs.

Through contractual agreements, all organizations involved in the cloud provision would be accountable. While the corporate user, as the first corporate entity in the cloud provision, would be held legally accountable, the corporate user would then hold the initial service provider (SP1) accountable through contractual agreements, requiring in turn that SP1 hold its SPs accountable contractually as well. This is analogous to some existing cases in outsourcing environments, where the transferor is held accountable by regulators even when it is the transferee that does not act in accordance with individuals' wishes [27]. For further details of this approach, see [41]. An alternative approach could be based upon the mechanisms described in [42].

F. Cross-Layer Interaction and Enforcement

One aspect that is needed is for consumer and service provider requirements to be enforced along the chain of cloud providers. One way of doing this is via accountability by contract (as explained in the previous subsection), together with machine-readable policies propagated with data through the cloud and integrated risk assessment. Natural language policies in the contract can be associated with lower-level machine readable policies that define usage constraints of the associated PII, are transmitted through the cloud associated with PII, and are acted on automatically within the cloud without the need for human intervention. Privacy protecting controls can be built into different aspects of the business process: there is an ongoing process of review throughout the contractual chain, combined with risk assessment and decision support to assess harm. One aspect is bridging the policy level gap from legal and regulatory policies and risk assessment, to technical policies and policy languages like XACML, P3P, EPAL and RBAC. Within the EnCoRe Project [43] we are building a conceptual model for this mapping from high level to low-level policies, using templates for different policy requirements [44].

The high-level policies that apply to a given situation may be deduced with the help of a decision support system [45]. If lower-level policies are then generated that correspond to these high-level policies as described above, then these could be enforced within the cloud by means of sticky policies, data tagging and privacy-enhanced access control.

Decision support tools may also be used to assess risk and privacy harms based on context before transferring PII in the cloud, to advise on obligations to be passed to a CSP (within the contract, and associated with data) or to assess the suitability of a CSP before using it [46].

G. Mechanisms for Increasing Trust

Both social and technological mechanisms to increase trust are necessary for cloud computing, quite apart from additional social guarantees of privacy and security.

Social mechanisms, behaviour and values contributing to persistent trust include sanctions, assurance and vouching (including seals of approval): such examples are of infrastructural mechanisms that may vary over time, but in general are relatively stable. Technological mechanisms include underlying security infrastructure, well-known practices and the technological features corresponding to static social mechanisms; these can involve the following, for example: certified hardware (for example, tamper-resistant hardware); protocols; certified cryptographic techniques; assurance; other security features; audit and enforcement.

The content of social mechanisms giving rise to dynamic trust would be liable to substantial change at short notice, such as brand image, look and feel, reputation and history of interactions. The technological mechanisms supporting dynamic trust give confidence that a particular environment or system state is trusted (at a given time, for a particular purpose). A system's behaviour can change according to a given context, and in particular if it has been hacked, and in some cases system behaviour can be driven by policies (dictated by people, business needs or even malicious people) that change over time. For example, dynamic trust could be affected by the following information being divulged: a particular cloud component has been compromised; the jurisdiction of stored data has changed; software is in a certain state; policy enforcement has not been carried out.

Clouds can outsource processing, storage or maintenance of data seamlessly. In this situation it is important that technical solutions are put in place to keep this chain of trust relevant. Binding the legal contractual agreements with the data using sticky policy mechanisms [47] is a way to achieve it. However, such techniques require common elements for key management and the lack of standardization makes this difficult to conceive at present.

Ann Cavoukian, the Privacy Commissioner in Ontario, suggests four fundamental technological approaches towards assuring confidence and trust in the privacy of PII in the cloud [48], namely:

- attaching individual privacy rights, conditions and preferences directly to identity data
- getting personal devices like cell phones, PDAs, smart cards and other tokens under our physical control to interface with the cloud and act on our behalf
- using intelligent software agents within these devices or within the cloud to automatically and continuously scan, negotiate, do our bidding, reveal identity information and act on our behalf
- having trusted identity providers acting as privacy infomediaries and carrying out audit and enforcement

We are still some way from having developed these technological approaches, but we do have the foundations for providing these, i.e. 'sticky policies' [47], digital rights-type technologies, agent technologies, federated identity systems, etc. One way to build customer trust for a CSP is to have appropriate governance frameworks in place to assure customers that it will fulfill the promises written in the terms of service. This goal can be achieved by having the CSP certified against ITIL, COBIT, etc. This is just an initial step in building the trust relationship: more mechanisms need to be in place to enable data rights management. One approach is cloud rights management, based upon mechanisms for uniquely identifying users (such as OpenID [49]), etc. and then using a classic identity management system to control rights. In general, providing interfaces that allow users to control their privacy settings in a transparent and non-complex way is important, as well as the back-end enforcement of those choices. For example, users might want to set up privacy rules based on time, location and groups of friends, and this could be linked into location obfuscation and provision of an 'invisible mode' within the cloud service. Other technologies also have a role to play: for example, provision of marketplaces for CSPs [50] and reputation management systems for cloud providers.

H. Combination of Solutions

As cloud computing exhibits many different aspects, privacy and security solutions need to address a combination of issues, and this may require new and even unique mechanisms rather than just a combination of known techniques for addressing selected aspects. For example, privacy problems when transferring PII across borders within a group of companies can be addressed via Binding Corporate Rules, and yet this approach would not be available to a corporation seeking to adopt a cloud computing solution where PII will be handled by third party CSPs. Similarly, the approach of Model Contracts that is used at present to satisfy EU adequacy requirements can take months to set up, and so is not suitable for cloud environments where new providers need to be brought in very quickly.

I. Contextual Choice

In section II we discussed the importance of context to privacy requirements. As a result solutions often need to be tailored to a specific context. In general, customers considering cloud services should consider their organisation's operational, security, privacy and compliance requirements to see what approach would best suit them. Technology may help this decision making process. For example, privacy and security requirements could be gathered and matched to service provisioning in an automated or semi-automated way, and on an ongoing basis. Some work has been carried out already related to this (e.g. [45,49]); in future, more research is needed to help cloud service provisioning be carried out at the different layers in a more optimal way.

VII. CONCLUSIONS

Cloud providers need to safeguard the privacy and security of personal data that they hold on behalf of organisations and users. In particular, it is essential for the adoption of public cloud systems that consumers and citizens are reassured that privacy and security is not compromised.

Responsible management of personal data is a central part of creating the trust that underpins adoption of cloud-based services – without trust, customers will be reluctant to use cloud-based services. Privacy protection builds trust between service providers and users: accountability and privacy by design provide mechanisms to achieve the desired end effects and create this trust. This management can span a number of layers: policy, process, legal and technological. It is universally accepted as best practice that such mechanisms should be built in as early as possible into a system's lifecycle. We are currently carrying out research on ways to improve the protection of private data and thereby enable further deployment of cloud technologies; these mechanisms include identity management, sticky policies and data obfuscation. By these means users and citizens can be provided with reassurance that their personal data will be protected, and cloud deployments can be made compliant with regulations, even within countries where such regulation is relatively strict.

Conforming to legal privacy requirements and meeting client privacy and security expectations with regard to personal information require corporations to demonstrate a context-appropriate level of control over such data at all stages of its processing, from collection to destruction. The advantages of cloud computing – its ability to scale rapidly (through subcontractors), store data remotely (in unknown places), and share services in a dynamic environment – can become disadvantages in maintaining a level of assurance sufficient to sustain confidence in potential customers. In this paper we have assessed some of the key issues involved, and set out the basis of some approaches that we believe will be a step forward in addressing this situation.

DISCLAIMER

The opinions expressed here do not necessarily represent the view of Hewlett-Packard Company or its affiliates.

REFERENCES

[1] Cloud Security Alliance, "Top Threats to Cloud Computing", v1.0, March 2010.

[2] Horrigan, J.B. "Use of cloud computing applications and services". Pew Internet & American Life project memo, Sept 2008.

[3] Greenberg, A. "Cloud Computing's Stormy Side". Forbes Magazine, 19 Feb 2008.

[4] Abrams, M. "A Perspective: Data Flow Governance in Asia Pacific & APEC Framework" 2008.

[5] Fratto, M. *Internet Evolution,* The Big Report, Cloud Control, 2009.

[6] Hall, J.A. & Liedtka, S.L. "The Sarbanes-Oxley Act: implications for large-scale IT outsourcing", *Communications of the ACM*, 50(3), 2007, pp. 95-100.

[7] Reidenberg, J. "Technology and Internet Jurisdiction", *University of Pennsylvania Law Review 1 SSRN eLibrary*, 2005.

[8] Kohl, U. *Jurisdiction and the Internet*, Cambridge University Press, 2007.

[9] Mowbray, M. "The Fog over the Grimpen Mire: Cloud Computing and the Law". *Script-ed Journal of Law, Technology and Society*, vol. 6, no.1, April 2009.

[10] Cloud Security Alliance, "Security Guidance for Critical Areas of Focus in Cloud Computing". 2009.

[11] Gellman, R. *Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing*. World Privacy Forum. http://www.worldprivacyforum.org/pdf/WPF_Cloud_Privacy_Report.pdf, 2009.

[12] IDC, Enterprise Panel, September 2009. http://www.slideshare.net/JorFigOr/cloud-computing-2010-an-idc-update

[13] Mather, T., S. Kumaraswamy, and S. Latif, *Cloud Security and Privacy*, O'Reilly, Sebastopol, 2009.

[14] Regulation of Investigatory Powers Act, Part II, s 28, UK, 2000.

[15] Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT ACT) Act, Title V, s 505, 2001.

[16] McKinley, P.K., Samimi, F.A., Shapiro, J.K., Chiping T.: Service Clouds: A Distributed Infrastructure for Constructing Autonomic Communication Services. Dependable, Autonomic and Secure Computing, IEEE, pp.341-348, 2006.

[17] Google App Engine. http://code.google.com/appengine

[18] Kortchinsky, K. "CLOUDBURST: A VMWare Guest to Host Escape Story", BlackHat, USA, 2009.

[19] Ristenpart, T., E. Tromer, H. Shacham, and S. Savage, "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds", *CCS'09*, ACM, Chicago, Illinois, November 2009.

[20] Rousseau, D., S. Sitkin, R. Burt and C. Camerer, "Not so Different after All: a Cross-discipline View of Trust", *Academy of Management Review*, 23, no. 3, 1998, pp. 393-404.

[21] Osterwalder, D., "Trust Through Evaluation and Certification?" *Social Science Computer Review*, 19, no. 1, Sage Publications, Inc., Spring 2001, pp. 32-46.

[22] Nissenbaum, H., "Can Trust be Secured Online? A theoretical perspective", *Etica e Politica,* no. 2, Dec 1999.

[23] Giff, S., *The Influence of Metaphor, Smart Cards and Interface Dialogue on Trust in eCommerce*, MSc project, University College London, 2000.

[24] Nielsen, J., "Trust or Bust: Communicating Trustworthiness in Web Design", *Jacob Nielsen's Alertbox*, 1999. Available via http://www.useit.com/alertbox/990307.html.

[25] Pearson, S., M. Casassa Mont and S. Crane, "Persistent and Dynamic Trust: Analysis and the Related Impact of Trusted Platforms", *Trust Management*, Proc. iTrust 2005, LNCS 3477, ed: Peter Herrmann, Valérie Issarny, Simon Shiu, pp. 355-363, 2005.

[26] Tweney, A. & Crane, S. "Trustguide2: An exploration of privacy preferences in an online world", *Expanding the Knowledge Economy*, IOS Press, 2007.

[27] Goldberg, N.M. and Wildon-Byrne, M., "Securing Communications on the Cloud", Bloomberg Law Reports – Technology Law, vol. 1, no. 10, 2009. http://www.infolawgroup.com/uploads/file/Goldberg%20Article.pdf

[28] Organization for Economic Co-operation and Development (OECD): Guidelines Governing the Protection of Privacy and Transborder Flow of Personal Data. OECD, Geneva, 1980.

[29] EU Data Protection Directive (95/46/EC), 1995.

[30] Salmon, J. "Clouded in uncertainty – the legal pitfalls of cloud computing". *Computing magazine*, 24 Sept 2008.

[31] Crompton, M., Cowper, C., Jefferis, C. "The Australian Dodo Case: an insight for data protection regulation", *World Data Protection Report*, vol. 9, no. 1, BNA, 2009.

[32] Yao, A.C., "How to Generate and Exchange Secrets", *27th Symposium on Foundations of Computer Science (FoCS)*, IEEE Press, New York, pp.162-167, 1986.

[33] Gentry, C., "Fully Homomorphic Encryption Using Ideal Lattices". *41st ACM Symposium on Theory of Computing (STOC)*, ACM, New York, pp. 169-178, 2009.

[34] Pearson, S., Y. Shen and M. Mowbray, "A Privacy Manager for Cloud Computing", *Proc. 1st CloudCom 2009*, ed. M.G. Jaatun, G. Zhao, C. Rong, Beijing, Springer LNCS 5931, pp. 90-106, December 2009.

[35] Google, Secure Data Connector, http://code.google.com/securedataconnector/docs/1.0/overview

[36] Federal Trade Commission, Privacy Online: Fair Information Practices in the Electronic Marketplace: A Federal Trade Commission Report to Congress, Washington DC, May 2000.

[37] Pearson, S. "Taking Account of Privacy when Designing Cloud Computing Services". *ICSE-Cloud'09*, Vancouver, IEEE. Also available as HP Labs Technical Report, HPL-2009-54, http://www.hpl.hp.com/techreports/2009/HPL-2009-54.html, 2009.

[38] Information Commissioners Office, *Privacy by Design*, Report. www.ico.gov.uk, 2008.

[39] NEC Company Ltd and Information and Privacy Commissioner, Ontorio, Canada, "Modelling cloud computing architecture without compromising privacy: A privacy by design approach", June 2010.

[40] ENISA, *Cloud Computing: Benefits, risks and recommendations for information security*, Ed. Daniele Catteddu and Giles Hogben, November 2009.

[41] Pearson, S. and A. Charlesworth, "Accountability as a Way Forward for Privacy Protection in the Cloud", *Proc. 1st CloudCom 2009*, ed. M.G. Jaatun, G. Zhao, C. Rong, Beijing, Springer LNCS 5931, pp. 131-144, December 2009.

[42] Weitzner, D., Abelson, H., Berners-Lee, T., Hanson, C., Hendler, J.A., Kagal, L., McGuinness, D.L., Sussman, G.J., Waterman, K.K., "Transparent Accountable Data Mining: New Strategies for Privacy Protection", *Proceedings of AAAI Spring Symposium on The Semantic Web meets eGovernment*, AAAI Press, 2006.

[43] EnCoRe: Ensuring Consent and Revocation project, http://www.encore-project.info, 2010.

[44] Casassa Mont, M., S. Pearson, S. Creese, M. Goldsmith and N. Papanikolaou, "EnCoRe: Towards a conceptual model for privacy policies", Primelife 2010.

[45] Pearson, S., P. Rao, T. Sander, A. Parry, A. Paull, S. Patruni, V. Dandamudi-Ratnakar and P. Sharma, "Scalable, Accountable Privacy Management for Large Organizations", *INSPEC 2009*, IEEE, pp. 168-175, September 2009.

[46] Pearson, S. and T. Sander, "A Mechanism for Policy-Driven Selection of Service Providers in SOA and Cloud Computing", *Proc. PROMASC 2010*, IEEE, 2010.

[47] Casassa-Mont, M., S. Pearson and P. Bramhall, "Towards Accountable Management of Identity and Privacy: Sticky Policies and Enforceable Tracing Services", *Proc. DEXA 2003*, IEEE Computer Society, pp. 377-382, 2003.

[48] Cavoukian, A. "Privacy in the Clouds", *Identity Journal Ltd*, Springer, 2008.

[49] OpenID Official Site, 2010. http://openid.net/

[50] Li, J., B. Stephenson, H.R. Motahari-Nezhad and S. Singhal, "A Data Assurance Policy Specification and Enforcement Framework for Outsourced Services", HP Labs Research Report, HPL-2009-357, 2009.