Designing a Secure Cloud Architecture: The SeCA Model

Thijs Baars, Department of Information and Computing Sciences, Utrecht University, The Netherlands
Marco Spruit, Department of Information and Computing Sciences, Utrecht University, The Netherlands

ABSTRACT
Security issues are paramount when considering adoption of any cloud technology. This research proposes the Secure Cloud Architecture (SeCA) model on the basis of data classifications which defines a properly secure cloud architecture by testing the cloud environment on eight attributes. The SeCA model is developed using a literature review and a delphi study with seventeen experts, consisting of three rounds. We integrate the CI3A—an extension on the CIA-triad—to create a basic framework for testing the classification inputted. The data classification is then tested on regional, geo-spatial, delivery, deployment, governance & compliance, network, premise and encryption attributes. After this testing has been executed, a specification for a secure cloud architecture is outputted.

Keywords: cloud computing, cloud security, information security, delphi study, ci3a, information systems, secure cloud architecture

INTRODUCTION: IN SEARCH FOR A SAFE DELIVERY MODEL
The Cloud is called by some a paradigm-shift in computing (Voas & J. Zhang, 2009), by others it doesn't even exist (Reuters, 2008). It is in this light that the presented research tries to formulate the complexities of cloud security. This new phenomena called the cloud does exist, however it is not a brand new technology. The cloud has always been here, under the name of "the internet", and the idea of utilizing the internet as a storage and computing power provider isn't new either. In 1993, Eric Schmidt, then CTO of Sun Microsystems, said in an email "When the network becomes as fast as the processor, the computer hollows out and spreads across the network" (Gilder, 2006). This "the network is the computer" concept is basically what the cloud is all about. Utilizing all the power that makes up the all-encompassing internet for better productivity and scalability. That being said, some still call it the new paradigm of computing. That is because the cloud is a new delivery model, or as Mulholland, Pyke, & Fingar state: "The big deal is that cloud computing is a disruptive *delivery* model. It's an economic, not technological shift!" (Mulholland, Pyke, & Fingar, 2010 p.24).

The National Institute for Standards and Technology (NIST) defines cloud computing as: "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models" (Mell & Grance, 2010 p.1). The European Network and Information Security Agency (ENISA) defines the cloud similarly

(Hogben & Catteddu, 2009), and these definitions will be used continuously in this paper when referring to the cloud.

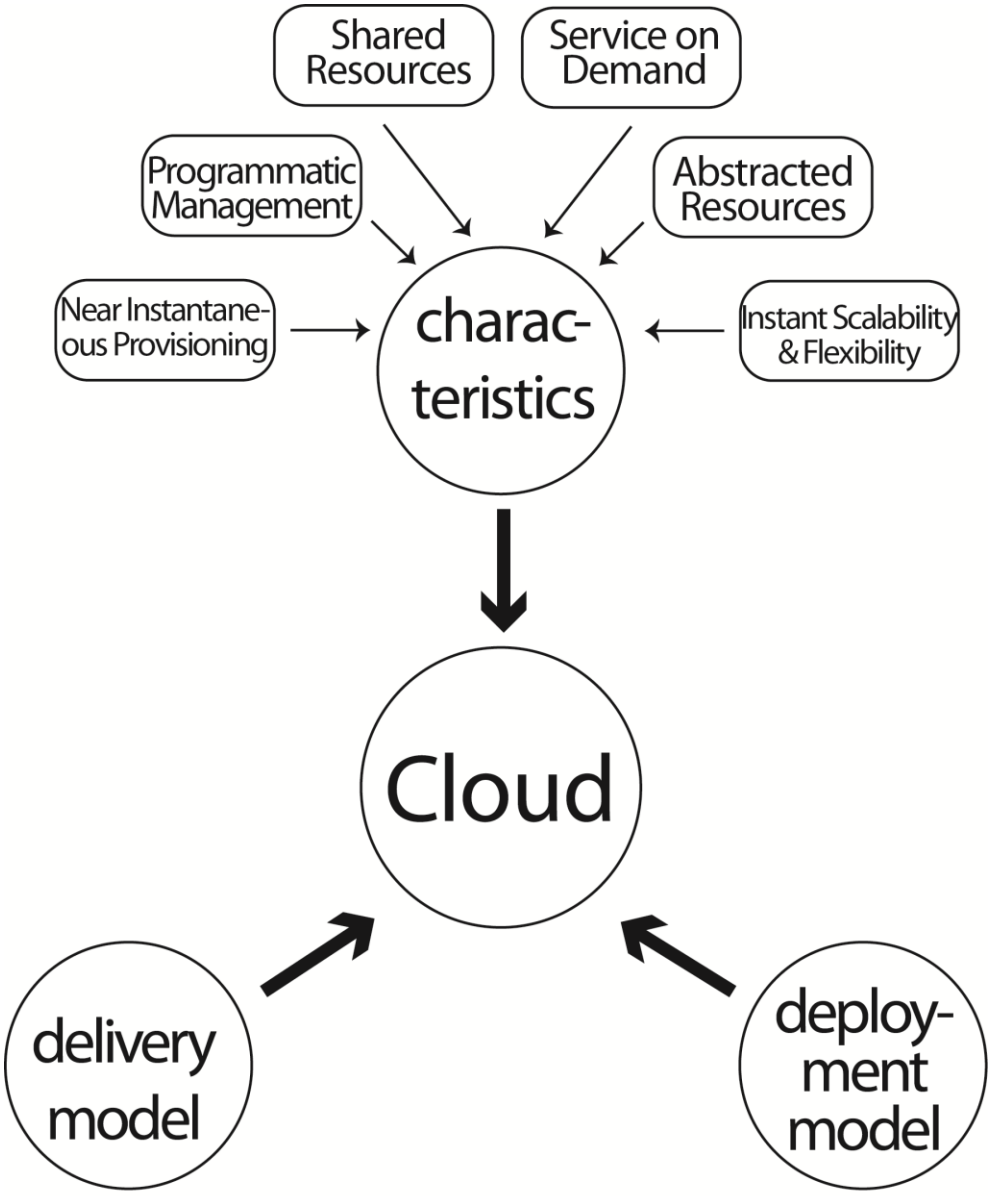

Figure 1: Key cloud computing characteristics.

The definitions show that ASPs are more or less a part of the Cloud and that Software as a Service (SaaS) is actually a model within a cloud environment. Table 1 below explains these characteristics of the cloud.

Cloud characteristic	**Description**
Abstracted resources	Using virtualisation, resources can be 'created' and scaled on the spot over one or more physical resources.
Instant scalability & flexibility	The ability to add or remove virtual resources with the click

	on a button.
Near instantaneous provisioning	The ability to supply resources, services and such nearly instantaneous.
Shared resources	Multiple tenants can share resources.
Service on demand	Get the services needed on demand, and pay only for what you use (pay per hour basis, pay per use etc.).
Programmatic management	APIs provide interfaces to manage the cloud environment. E.g. via web interfaces.

Table 1: Cloud characteristics explained (Mell & Grance, 2010).

According to both commercial reports and academic research, security issues are paramount when adopting the cloud (Foster, Zhao, Raicu, & Lu, 2008; Ghinste, 2010; Mowbray & Pearson, 2009), while no clear model exists to determine security issues and solutions. Therefore this paper will provide an overview of the security issues and describe the Secure Cloud Architecture (SeCA) model to determine the security issues one might expect in a certain cloud environment and what solutions might be used to secure those issues. The framework is developed as an answer to the following question:

Can the Cloud be a safe alternative for the storage and execution of organizational confidential data?

RELATED RESEARCH: TOWARDS SECURE USAGE OF CLOUD TECHNOLOGIES

Although the Cloud is still in development (Mulholland et al., 2010), it has already caught the attention of the research community. Its definition (over 20 are known) has been researched in (Vaquero, Rodero-Merino, Caceres, & Lindner, 2008) and is also discussed in (Chen, Paxson, & Katz, 2010). Vaquero et al. (2008) manage to give an overview of the features a cloud should have and discuss the differences with a computing grid. The NIST definition (Mell & Grance, 2010), which was defined in May of 2009, five months after the publication of Vaquero et al. (2008), is a more open definition, while still preserving the key characteristics of the cloud.

Mulholland et al. (2010) give an overview on Cloud computing and its facets for enterprises in their book, but fail to mention any security related topics. Jericho Forum (2009) has modelled the cloud in order to help users to understand the different facets of the cloud and support a secure use of cloud technologies. It will be discussed in the second-to-last section "SeCA: The Secure Cloud Architecture Model".

The security of the cloud is an issue that is well in the centre of cloud research. Chen et al., (2010) give an introduction to security issues in the cloud and discuss which issues are specifically new in the cloud and which are issues that that are related to traditional forms of computing. Stating that "arguably many of the incidents described as "cloud security" in fact just reflect traditional web application and data-hosting problems [..] such as phishing, downtime, data loss, password weaknesses, and compromised hosts running botnets." (p. 4, internal references removed). They hold that most cloud security issues are not new, but do need new implementations to provide the security wanted. Benson, Sahu, Akella, & Shaikh (2010) discussed security issues from a cloud provider's support division perspective. Kaliski Jr & Pauley (2010) discuss risk assessment of the cloud, stating that "[t]he very characteristics that make cloud computing attractive also tend to make it hard to assess" (p.2). Richter et al. (2011) discuss VM retrospection, a method for inspecting previous VM states in order to perform forensics, debugging, troubleshooting and so forth. Wang, Wang, Li, Ren, & Lou (2009) discuss

the necessity of a Third Party Actor (TPA) to assure security standards and to provide transparency in the security controls. Christodorescu, Sailer, Schales, Sgandurra, & Zamboni (2009) discuss methods of securing Clouds at the Virtual Machine (VM) level. They provide a short overview of known VM issues and solutions, and then propose their system which protects VMs in a cloud against malware and rootkits using a white/blacklist method. Jensen, Schwenk, Gruschka, & Iacono (2009) describe technical security issues related to cloud, using the Amazon EC^2 cloud as a case. Although all issues discussed are related to the cloud, all of them were already apparent before the coming of the cloud, some just have found new grounds to be relevant again. Ko, Jeon, & Morales (2011) developed the HybrEx Model which shows how more resources from public clouds can be added to private clouds without concerns for privacy and security. Vigfusson & Chockler (2010) discuss in "Clouds at Crossroads: Research Perspectives" research topics in the cloud, including privacy related issues. Discussing the trust problems that arise with the complex models of some cloud environments, it provides a few solutions to suggestions which might be very feasible.

This research overlaps with current research in that it provides a global overview of the cloud, introducing and explaining it, but it also expands the current research with proposed solutions to analyse the cloud and model a secure cloud architecture. With the current research explaining either very technical details on protecting the cloud, where the mere describes arbitrary issues that are not specific to the cloud, or giving overviews of the cloud where security issues are touched lightly, this research will focus on the security issues in depth that come with the cloud in a more practical sense. We specify cloud specific issues and general security issues that have found new terrain in the cloud environment.

To conclude, this research provides users with a model that navigates them through the security checkpoints in a cloud environment, outputting an architecture specific for their data classification.

RESEARCH METHOD: A THREE-STAGE DELPHI STUDY

To conduct the research this paper describes, a literature review has been conducted first. By reading the overall themes in security, followed by cloud specific topics, an overview has been created that is used as the starting point in the development of the SeCA model. This model has been verified by an expert panel. The experts were selected on the basis of their function, publications and knowledge of security, the cloud or a combination thereof. This group of experts, 26 in total, coming from organizations within the business to consumer, business to business, business to government industries and governmental organizations. They were interviewed using the Delphi tool developed at the Wharton Business School (Wharton Business School, 2005).

#	Position/function	Organization type	Cloud	Non-cloud	Security
1	Consultant	Enterprise integrator	X	X	
2	Director	IT consultancy		X	X
3	Security consultant/architect	IT security firm		X	X
4	Researcher	American University,	X	X	X
5	Enterprise Architect	Enterprise transportation		X	
6	Sr. manager	Large accounting firm	X	X	X
7	Security advisor	Transportation firm		X	X

8	IT Architect	IT consultancy	X	X	
9	Manager	Security solutions	X		X
10	Security manager	Utilities		X	X
11	Consultant	IT consultancy	X	X	
12	Security manager	Government		X	X
13	Security manager	Healthcare products		X	X
14	Manager	IT consultancy		X	X
15	Consultant	Enterprise integrator		X	X
16	Security manager	Utilities		X	X
17	IT auditor	Accounting	X	X	X

Table 2: The experts (filtered on those that did all three rounds) in the delphi study

A delphi study has been considered to be the best method for research in this paper, as it provides the researchers with a qualitative data set which would allow to create and verify the model, and would allow for the experts to see anonymized answers and be able to respond to these answers in upcoming rounds (Dalkey & Helmer, 1963). The first expert answer in question one was not the same first expert in question two per se. This way, a consensus is reached on the acceptance of the model. The delphi method was executed consisting of three rounds of surveys with qualitative questions. Three rounds were chosen instead of two (which is more common (Skulmoski, Hartman, & Krahn, 2007), so that a first round could be used to obtain general information on the topic, not specifically regarding to the model to be developed, while still having enough rounds to reach a consensus. The first round consisting of open questions where the experts were questioned on their experience with security and the cloud, issues and concerns regarding security in the cloud. These questions gave a wide result set that strengthened the results of the literature research earlier performed. Seventeen respondents answered the questions in the survey in all three rounds, a rate of 65%. See Table 2 for an overview.

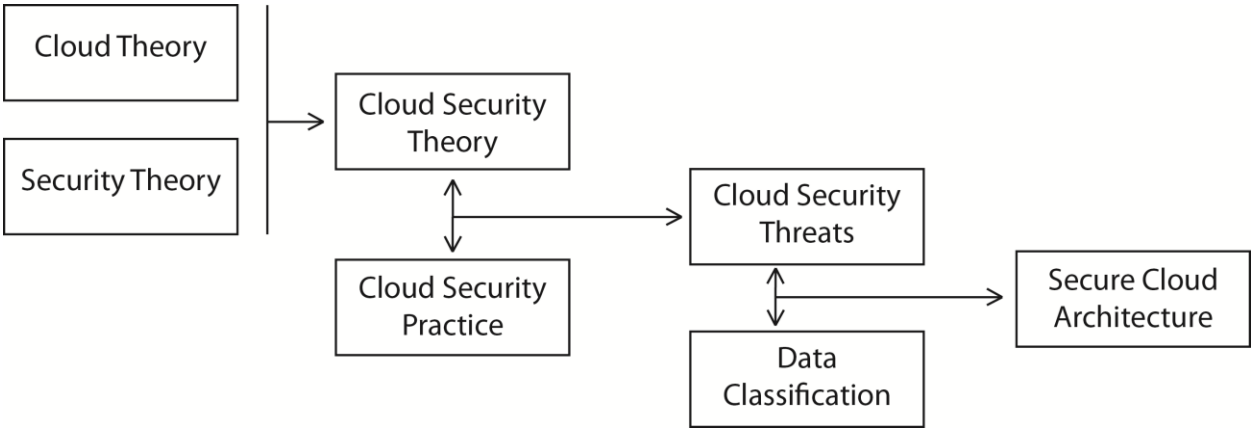

Figure 2: Conceptual model of the research presented

RESULTS: FROM EXPLORATION TO MODEL VALIDATION

The result set that the first round created delivered the starting point for the second round, in which some questions were asked again in order to give the experts the option to rephrase their answers after having seen the answers of round one. Some questions were designed after noticing a consensus or discrepancy in the answers from round one, where others were

completely new and had no specific relation to the questions asked in round one. The questions for round one were fed by literature review and informal meetings during conferences and congresses.

In the second round, an initial version of the SeCA model was presented. The goal of this model is to provide implementers, decision makers and experts in the field with a framework that they can use in order to assess cloud environments to their security needs. The feedback on this initial model was then used to improve it. The author of this paper originally expected that a framework where raw data is inputted would be best. This as the data will be hosted in the cloud (or at least has that intention.) However, it was found that the SeCA model looked at the data from an unusual perspective for its target audience and that a more architectural point of view was needed in order to be usable in the field. In round three, an improved SeCA model was introduced; the SeCA model was remodelled to accept data classifications as input, instead of raw data. This as raw data is classified within an organization and that for each classification, a different system architecture might be needed to host and execute data safely, as prescribed by the pertaining classification. Therefore, encryption is inserted as an attribute as this changes among the other attributes per classification and thus architecture.

The SeCA model allows for any user to assess the cloud environment from two perspectives. Either the user looks at its current data and the inherent classification and decides how the cloud environment should be configured to meet its requirements. Or one reverses that action and sees what data can be used by taking a cloud environment and on that basis determine what can go in. This paper will describe only the forward movement, thus taking data classifications as an input and determine on that basis how the cloud environment should be configured.

The burn chart below shows the amount of consensus reached in the study, per topic. White cells represent no questions on that topic were asked in that round; grey that consensus was reached; checkered pattern a consensus in part was reached; black no consensus reached.

Topic	**Round1**	**Round2**	**Round3**	**Consensus**	**Comments**
Security issues					All issues are accounted for in the model
Locationlessness					Location is a new issue and thoroughly discussed.
Trust issues					Outsourcing/insourcing/cloud differences are in discord
Encryption					Different knowledge levels; study done through literature.
Feasibility					Not a technical/security issue. Topic abandoned.
Model					Model validated and approved by the experts.
Auditing					Issues reached consensus; added to the CI3A.

Table 3: burn chart of the consensus reached in the delphi study.

As one can see, not all topics reached consensus. This was due to the fact that in the expert selection, business knowledge or technical knowledge on some topics were not taken into

account. For example, the field of encryption is a very technical field that can be hard to fully understand and apply. Although some answers were very useful, other answers were dismissed in the same round as unfeasible, simpleminded or simply not true. This meant that the experience or knowledge between the experts varied too greatly to reach consensus. Subsequent research was done through literature review in the applicable topics.

RISKS & THREATS IN CLOUD COMPUTING

As previous research has shown, many risks in the cloud are not specifically cloud related, but browser, user or framework related (Jensen et al., 2009). This paper will not discuss those issues at large, but will instead focus on the design and implementation of the cloud environment from the security perspective at the server side

CI3A

Because of the complexities the cloud presents as dictated by a majority of the participants of the Delphi study, the *de facto* CIA triad, which is used for testing the confidentiality, integrity and availability in systems, data flows and so forth, was found to be too constrained. For that reason, the CI3A was developed. The CI3A is an extension on the CIA triad, comprising of confidentiality, integrity, availability, accountability and auditability. The proposed model utilizes CI3A to assure the right level of security is maintained within the environment. This section will describe the CI3A, following separate sections on locationlessness and trust chains.

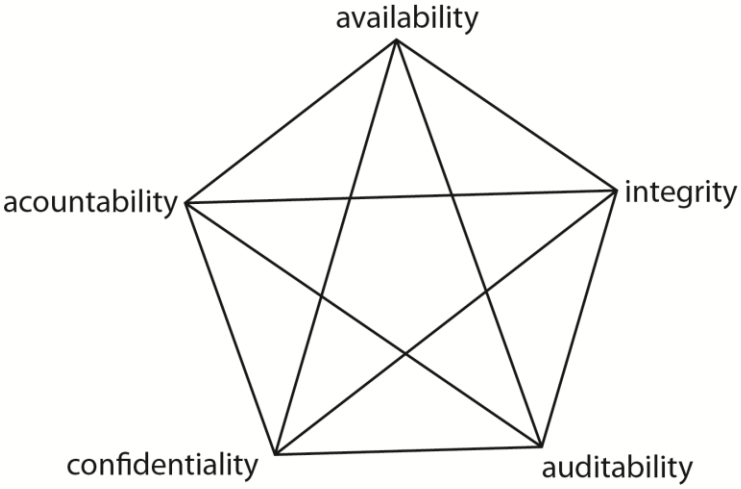

Figure 3: The CI3A visualized

Parts of the CI3A

Confidentiality is reached by proper authentication/authorization controls and encryption methods such as secured computing and two-factor encryption. Preventing data leakage is a central part within the confidentiality strategy. The choice of distribution and delivery model influences the level of confidentiality and the methods needed to assure confidentiality.

Integrity assures only authorized actors have access to certain data and that data gets distributed to only authorized persons. Within that distribution, any editing or changes within the data should only be made by the right persons. Governance and Compliance influence the

integrity of the data; a fully compliant environment is more likely to assure integrity. As with Confidentiality, the chosen delivery and distribution model influences the level of integrity.

Availability comprises of measures to prevent unauthorized actors from deleting and moving data, or accessing those files, minimizing downtime of the environment. These measures could be a HA infrastructure, strong authentication servers, disaster recovery or external hosted service such as CloudPolice (Popa, Yu, Ko, Ratnasamy, & Stoica, 2010). Availability plays a big role within the cloud environment, as servers can be hosted anywhere in the world, at multiple locations. Although an advantage in the eyes of HA and disaster recovery; latency, desynchronization and vulnerabilities in the transceiver links can pose threats. Also, ownership of data is a part of availability. Availability is linked to regional, geo-spatial, network, premise and to the delivery and distribution models.

Accountability defines the measures taken to assure that no actor can make actions without a record. This is needed for forensics and governance. The measures needed to assure accountability greatly depend on the delivery model, but also on the distribution model and compliance in general. Chen Wang & Zhou (2010) have found accountability of paramount importance in the cloud, proposing a method for transferring accountability onto an external host in order to perform accountability in a multitenant platform.

Auditability, the ability of the environment to be audited, is directly related to governance and compliancy. Without a decent grade of auditability, compliance cannot be achieved. Auditability is influenced by the delivery and distribution model, as well by the geo-spatial and geographic boundaries.

Locationlessness

Because of the virtualized environment in which the cloud runs, geographic location does not tend to be an issue in the eyes of the beholder. The end user might not, or in some opinions doesn't need to, know where his data physically resides. This locationlessness behavior of the cloud can be a serious risk that outweighs the benefits, according to a consensus of all the participating experts. We define a locationless cloud as: a cloud environment in which the end user has no awareness of where his data physically resides. Issues of a locationless environment are plenty.

First of all, your data has to reside somewhere physically in order for any system to get it, even though it seems to the end user his data is located ubiquitous. This lack of control makes compliancy nearly impossible; physical security becomes hard to control.

Further auditability issues arise, as different countries have different legal systems that will require different solutions. Then, there is the issue of availability. With no knowledge of where data resides, and no need for providers to provide the user with that information, it might happen that your data will reside at the other end of the world from one moment to another resulting in a high latency or even time-outs.

Trust chains

Trust is a major issue in any relation, be it personal or professional. Although this is trivial, cloud computing can create trust chains, in which the end user is not always aware which other links are present in his chain of trust. This pertains especially towards delivery models. With IaaS, the tenant is in direct contact with the owner of the infrastructure (in some cases there might be a reseller in between) who can have outsourced duties associated with the maintenance of the physical systems. In a SaaS model, one is not aware if the SaaS provider also owns the platform,

or the infrastructure. This means that there might be a variety of different actors working on the cloud, whom all might be able to access the data that is being used in the SaaS in some way or another. Actors whom the tenant initially didn't trust have now become a part of his organizational network. This might result in actions that are a threat to the data. Although doing business is about making relations and trust, making this issue not insurmountable, it is a risk factor.

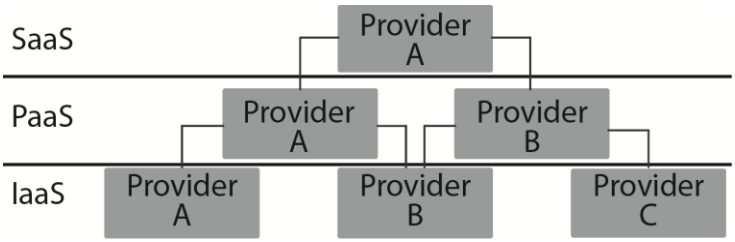

Figure 4: Trust chains in cloud architectures.

Regional

Regional describe the boundaries that signify separate legal systems. These boundaries include cities, states, countries and territories. Some changes in legal systems might be significant, such as the difference in respect to privacy between the European Union and China; some might be incremental such as the difference between county and state laws in the United States. These differences however do impose a risk if your data gets placed on a physical server crossing such a boundary. Next to that, different legal systems have different perspectives on privacy, the use of subpoenas on data extraction from datacenters. As one expert commented: "bringing privacy information out of the European Union can be [a] violation of local or European law". This would mean that keeping in compliance with laws, be it local, national or international, will become more difficult without knowledge of the physical location of the data store and computing unit. Peterson & Gondree (2011) provide an elaborate view on the importance of data location awareness from an American perspective.

Geo-spatial

With regional or geo-spatial risks, the distance of objects "relating to the relative position [..] on the earth's surface" (Collins English Dictionary, 2009) is meant, in this case the distance between servers, but also the location of each server. This can be of importance in the case of disaster recovery, but also with regards to physical security as presented in security norms such as the ISO 2700x series. In the light of location, one could also consider other features such as the building type, the accessibility of the server etc. Geographic location should also be taken into account in the light of latency and propagation speeds, as emphasized by Tiwana, Balakrishnan, Aguilera, Ballani, & Mao (2010).

Delivery model

The cloud has three distinct platforms on which a cloud environment can be offered. They are stackable, meaning that if you have a Software as a Service (SaaS) solution, chances are that your provider manages a Platform as a Service (PaaS), but takes services from an Infrastructure as a Service (IaaS) provider. (See the section on "Trust Chains" for an elaboration.) This, however, does not mean that every SaaS solution is running as the top of a stack of cloud

platforms. A SaaS solution can run on a traditional hardware stack with no further cloud environment attached. Figure 3 shows the hierarchy within the cloud. IaaS can be used to deploy PaaS solutions; PaaS can be used to deploy SaaS solutions.

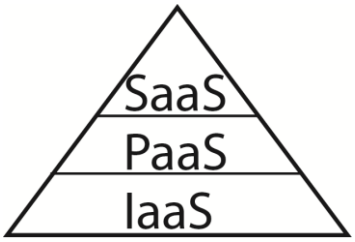

Figure 5: The delivery models visualized in a triangle

IaaS

The lowest platform in the above displayed delivery method pyramid, IaaS, or Infrastructure as a Service, provides the infrastructure of a server park. "[V]irtual machines and other abstracted hardware and operating systems which may be controlled through a service API." (Hogben & Catteddu, 2009 p.15). Virtual Private Servers hosted as a cloud environment are often IaaS services. Rapid Elasticity comes into place as more resources are required, the IaaS provider can then add more Virtual Machines to the subscription, and are wound down when no longer needed (Mulholland et al., 2010). (Williams et al., 2011) propose a modular type of IaaS which allows for extending current IaaS architectures.

PaaS

The Middle layer, Platform as a Service, "allows customers to develop new applications using APIs deployed and configurable remotely. The platforms offered include development tools, configuration management, and deployment platforms. Examples are Microsoft Azure, Force and Google App engine." (Hogben & Catteddu, 2009 p.15), "The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations" according to the NIST definition (Mell & Grance, 2010 p.2).

SaaS

SaaS is software available on subscription, or as ENISA defines it: "software offered by a third party provider, available on demand, usually via the Internet configurable remotely." (Hogben & Catteddu, 2009 p.15). NIST explains it as: "The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser. [..] The consumer does not manage or control the underlying cloud infrastructure [..] with the possible exception of limited user-specific application configuration settings" (Mell & Grance, 2010 p.2).

OaaS

Apart from the above three mentioned, there are multiple parties that tend to acknowledge more levels In the pyramid. These "Others as a Service" include BPMaaS (Business Process Management as a Service (Mulholland et al., 2010), Security as a Service (McAfee, 2009), Disaster Recovery as a Service (Wood et al., 2010) and Storage as a Service. These are just a

small selection of popular services in the cloud. Some of them float between the aforementioned three layers. Security as a Service for example might exist on all levels, as its nature encompasses the infrastructure to end-user authentication. This paper limits itself to IaaS, PaaS and SaaS, as results might be very usable on other services than discussed in this paper.

Deployment model

The cloud comes in four different deployment models, these are private, public, hybrid and community/partner clouds. The difference between these four models is the openness of the cloud to its tenants. Below is a description of each these models.

Private Cloud

This cloud infrastructure is operated for just one organization. This does not mean that it has to be managed by that organization. The management of the private cloud can be done by a third party, and the cloud itself can be physically located on the premises of that organization, or can be hosted somewhere else (sometimes called a virtual private cloud.) The cloud can exist behind a firewall of the organization, and thus only accessible within its private network, but can also be hosted off-premise on dedicated hardware (thus no multitenancy with other organizations). The main difference between a mainframe or internal traditional datacentre and a private cloud is that there is a virtualization layer that can be used to host SaaS applications, rapid deployment and other benefits of cloud computing. Even though private clouds conventionally lack the flexibility of their public equivalents, a model has been proposed to allocate public cloud space for private clouds, giving it the full flexibility as public clouds with the added security of private clouds (Ko et al., 2011).

Community or Partner Cloud

In a community cloud, a community or group of organizations share the same cloud infrastructure. These communities have shared concerns, such as a mission, goal and/or policy. The cloud can be managed by one of the organization within the community, or by a third party, and may exist on or off premise. (Mulholland et al., 2010) An example of a community cloud is the Eucalyptus Community Cloud. It is a platform for software engineers to test drive Eucalyptus (cloud architecture software). It features all possibilities that a self-hosted eucalyptus cloud would have, thus making it possible to use eucalyptus before investing in the full implementation of it (Eucalyptus, 2011).

Public Cloud

The cloud is open for use to a large group of tenants, which do not need to know each other. The Cloud is a ran by a cloud service provider. An example can be a majority of offerings from Force.com and Google's Gmail to VPS.net, Rackspace cloud hosting and other public services, free or on a subscription basis.

Hybrid Cloud

This cloud is a composition, or hybrid if you will, of two or more clouds of the types mentioned above. They are unique entities, but tied together with APIs to enable the exchange of data and applications. Due to the nature of the different clouds the hybrid consists of, it can be deployed both on and of premise, and be partly behind the firewall of an organization. An example is the

announced app store HP is building in order to communicate and process sales with public and business relations. It features a public cloud in which consumers can buy products and services, and a private cloud for product development and administration. (Hewlett-Packard Development Company, L.P., 2011)

Governance & Compliance

Executing governance and compliance is according to our experts is a much debated issue. Because governance and compliance greatly depend on the infrastructure of the system and the above-mentioned boundary issues, this topic is much under the discretion of the chosen cloud environment.

Depending on the chosen delivery model, compliance can be completely out of hand. A SaaS application depends on their vendors for governance and compliance. For PaaS it is partly the same, any compliance and governance within the software and how it handles data is on the part of the developer. The governance of the infrastructure and platform on which the application relies is in the hands of the provider. As with SaaS, negotiations need to take place with the provider in order to secure compliance. For IaaS, most of the governance and compliance lays in the hands of the tenant. The IaaS provider has to take care of the compliance to standards such as SAS-70, but many issues like privacy, data encryption and authentication are the responsibility of the tenant.

Concerning deployment models, compliance and governance in a public cloud can be difficult, as you are limited to your VM instance, whereas in a private cloud your negotiation position will be stronger as there are no other tenants to take into account. In a partner cloud, one can imagine that governance and compliance is a shared goal.

Concerning boundaries, the major aspect is the geographic location of the servers. The easiest option is of course in the same region as the organization resides, most knowledge of laws and executing governance/assuring compliance will be readily available. Auditing will not be an issue, as you can identify an auditing partner with whom you can easily communicate. That being said, the hardest option is obviously a cloud environment dispersed over the globe. Although disaster recovery wise there will be no issue complying with the toughest guidelines, getting audited and governance worldwide will be tougher. Although experts in the survey were wary of the fact that it could be done, in a personal interview with a Chief Information Security Officer of a large utilities company, it was made clear that a global audit is unprecedented.

Encryption

Encryption plays a vital role within the cloud environment. It is affected by all but the geo-spatial attributes in the SeCA model and affects the regional, delivery and deployment model. Although encryption is a broad topic that has been covered in many papers, theses and books, there are some aspects that are specifically related to the cloud. VPN tunnels, together with SSH can provide secure access to the cloud environment. Two-factor authentication can be very helpful for the cloud environment. Many institutions are using hardware key-tokens or SMS gateways in order two provide the second form of authentication apart from keying in a password. Authentication servers using protocols as RADIUS in combination with LDAP, Kerberos or Active Directory can handle all access requests in a proven manner as they are no different from any LAN/WAN setup at a traditional environment. The author therefore believes

that in terms of access control, authentication and authorization, no cloud specific issues are at hand.

Apart from the aforementioned, an encryption method specifically pertaining to the cloud is secure computing. Secure computing offers a solution to issues that arise when multiple systems have to use secure information in transactions and computations, in essence described by Yao's (1982) Millionaires' problem.

This research has been extended by (Goldreich, 2000), who researched the problem with multiple actors (called SMC, Secure Multi-party Computations). Recent research involves SMC geometry, researching transactions of polygons on convex hulls. See (Wang, Luo, & Huang, 2008) for an overview.

It is known that any multi-party computational problem can be solved using the generic technique of Yao (Yao, 1982). To overcome the overhead with Yoa's Millionaires' problem, and thus SMC, it seems that algorithms designed to compute a special task need to be written (Feigenbaum, Pinkas, Ryger, & Saint Jean, 2004; Goldreich, 2000). Using encryption methods such as homomorphic encryption and public key encryption, several algorithms have shown to be applicable to the cloud (Das & Srinathan, 2007; Hu & Xu, 2009; Troncoso-Pastoriza & Pérez-González, 2010) and have proven to provide the security needed for the cloud within test situations approaching real life cloud environments.

These methods of secure computing would allow the creation of a chain of trust that is secure, even though not all parties within the chain know each other nor trust each other. This could overcome any trust issues that might be in the field of cloud environments. Together with the enhanced and proven techniques of authentication and authorization already available, encryption can make the cloud a very secure architecture.

Apart from the above mentioned, (Mowbray & Pearson, 2009) have developed a privacy manager that can obfuscate data in effort to protect it from malicious providers.

The following table shows how encryption affects and is affected by the choice of certain in the model.

SeCA attribute	Confidentiality	Integrity	Availability	Accountability	Auditability
Regional	X			X	X
Geo-spatial			X		
Compliance	X	X			X
Delivery model	X	X	X		X
Deployment model	X	X	X	X	X
Encryption	X	X		X	X
Network	X		X	X	
Premises	X	X	X		

Table 4: Encryption and how it is affected or affects the other attributes in the SeCA model

Network

Network indicates the boundary of an organizational computer network. This is an important factor, as some information is not wanted outside the corporate network, such as trade secrets. Keeping a cloud environment within the boundaries of the network can be reached by keeping it on premise and thus physically in the network, or it can be reached by creating a VPN connection (as elaborated by (Wood, Gerber, Ramakrishnan, Shenoy, & Van der Merwe, 2009))

or a VLAN (in case of internal networks, or in public as described by (Hao, Lakshman, Mukherjee, & Song, 2010)) in order to keep the information within the network.

Because some configurations stretch the extension of the enterprise network, additional risks are incurred due to this stretch, as some of our experts mentioned in the survey. This stretch in the network is also noticeable in the added amount of actors which have to be trusted. The cloud provider will probably have access to your network, or the possibility to illegally gain so.

An added risk is the uncertainty of the WAN infrastructure at the provider's side. Connecting with the cloud provider might create vulnerabilities that could threaten the corporate network. Next to that, multitenancy might also be considered within the range of network boundaries. Although multitenancy should never be a threat to the virtual machine, in that it shouldn't have the possibility of other tenants to enter your VM, it has been proven that a vulnerability on the OS level could provide access to other VMs. Ristenpart, Tromer, Shacham, & Savage (2009) describe ways to discover where nodes are hosted on Amazon's EC2 cloud, following with a discussion how to place a co-resident on that physical server in order to able to reach the hardware a selected node is on. By then compromising the system, the selected node might be entered. This is a risk that has to be considered, how small it seems to be (see (Asadoorian, 2007; Mehta & Smith, 2007; Ormandy, 2007) for an overview).

Premise

Organizational premises play a role in the physical location of the cloud environment. One can either wise choose to have the hardware reside on or off organizational premises. For high security purposes keeping the hardware on premise, and thus fully in own control, might provide a benefit; personnel can be screening, extended access control to the datacenter and forensics are some of those. This extends the discussion on the geographic location of the server, presenting a trade off in security between on-premise servers versus geo-spatial choices.

SECA: THE SECURE CLOUD ARCHITECTURE MODEL

Resulting from the research conducted, we can summarize that the cloud can be secure, as long as its policies and SLAs are correctly in place and enforced. The different factors and risks involving cloud computing make it difficult to pinpoint to one secure cloud. In fact that is impossible, due to the diversity of cloud architectures and the data that is being stored on it. To circumvent this problem, a model has been designed. This model has been validated in the final round of the Delphi study.

The model described below gives an abstract overview of all the characteristics of the cloud. It defines a secure cloud architecture for a specific data classification.

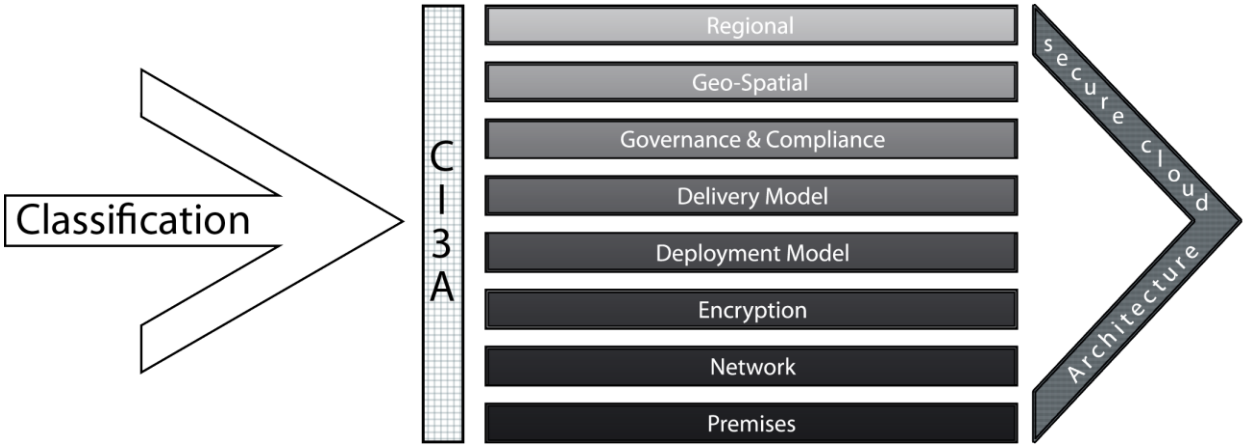

Figure 6: The SeCA Model

The model outputs guidelines for the Cloud environment and to which specification a cloud solution should adhere. In the appendix is a template that could be used for assessing the cloud architecture following the SeCA model. The flow chart below shows where this assessment ordinarily should take place. It is assumed that organizations already have classifications tailor made for their data in an earlier stage. For each classification a cloud architecture is assessed using the SeCA model. Once this is done, a list of cloud providers who can adhere to the results from the assessment is created. This can be done manually or using proposed services such as CloudCMP (Li, Yang, Kandula, & Zhang, 2010). Ultimately a cloud provider is selected, arrangements are made and the data can be placed in the cloud.

Figure 7: The position of the SeCA model in the process of moving data into the cloud (simplified)

It can occur that each classification has a different output from the assessment. (It is actually most likely to do so.) In that case several options are open. For each classification a different list of cloud providers is made in order to find and select the right cloud provider who can provide the cloud architecture needed. These can be combined in Hybrid Clouds. One can also decide that for certain classifications it is simply not feasible to transfer that data into the cloud and thus stay with the solutions already in place.

The model does not provide the intelligence which classifications could be hosted at the same cloud architectures. It is for the assessor to decide which cloud architectures that are the result from the assessments can be merged.

For example, when assessing a fictional classification named ‘private’, which would fit to any private information in an organization which is not mission critical and for internal purposes only, the following results can be outputted (simplified for exemplary purposes).

Classification: Private	**Secure Cloud Architecture specification**
Attribute	**Value**
Regional	Cloud environment physically within the same region as the organisation.
Geo-spatial	A High Availability architecture is preferred, at least one backup location in

	a different building on separate power net.
Governance/compliance	No need to adhere to specific standards. Annual audit is required to assure proper protection.
Delivery model	Any
Deployment model	Any
Encryption	a proper authentication and authorisation system should be in place for any actions. Alignment with the LDAP server in place is preferred.
Network	Any
Premises	Any

Table 5: results from the cloud analysis

CONCLUSIONS & FURTHER RESEARCH

Defining something as secure depends on many factors. Depending on the sort of data, the classification of that data and taking that wholly into perspective of the cloud environment, it can be said that the cloud is secure in certain situations. Depending on the outcomes of investigations, there should always be a cloud architecture that fits one's security needs. Better yet, the cloud can provide additional layers of security by utilizing virtualization, elasticity and HA architectures. Even though the additional layer of virtualization on the system might provide additional hazards, looking at the scarcity of exploitations in this layer one can rationally say that the virtualization layer adds more protection than threats.

By using the SeCA model described above, each and every classification can be checked to see how a cloud architecture should be designed in order to meet the security standards needed. It will, however, depend on the cloud provider whether it can deliver the architecture that is needed.

For the upmost secure classifications, a private cloud, hosted on premise, within the network, with mirroring on a different physical location (branch office) utilizing the needed encryption methods will provide a very secure architecture whilst maintaining the flexibility the cloud has to offer.

For every architecture counts that data location awareness is essential. Without the full knowledge of where the data resides and is processed, issues will arise in all actors of the CI3A. Data location awareness will also provide the means for compliance, legally and to security standards. These standards are being adopted by all major vendors, including Amazon, Google and Microsoft, with smaller ones following. This facilitates full compliance to the de facto security and auditing standards such as SAS 70, ISO 27000 series, PCI and COBIT. It depends, once again, on the configuration of the cloud architecture and, where applicable, the willingness of the cloud provider to allow for audits. If the selected cloud architecture features datacentres in widely spread different parts of the world, auditing might be more complicated. This of course also applies to the compliance to legal systems (privacy, intellectual property and auditing regulations) which can vary between jurisdictions. It is because of these implications that so-called locationless clouds are not preferable. They have an opaque layer that hides the user from vital knowledge in order to gain assurance with respect to the CI3A.

Further research can be conducted in the legal field. This was out of scope of this research, but the legal issues surrounding auditing, SLAs and NDAs are of paramount importance for the security in the cloud. SLAs especially, are of profound importance as they describe what measures a cloud provider should undertake for the security of the cloud. This paper

unfortunately has not had the possibility to explore the provider side of the cloud environment much.

Related to this is auditing in international/worldwide clouds. Auditing certifications, governance and compliance to legal systems in these environments mean that auditing firms, datacentre owner, providers and application owners all need to work together in order to perform a successful audit. In international and worldwide clouds these relations might become very complex, not to mention that multiple audit firms and offices have to work together. The issues raised with datacentres situated in different legal regions, such as China and the United States, are worth more research. Auditing plays also here a major role.

A pressing issue not discussed, but deserving further research regards the third party appliances that are currently installed in traditional datacentres. These appliances cannot be directly converted to the cloud, as the cloud does not offer any place for such appliances. It seems that at the moment of writing many of these appliances are converted to the cloud by their developers. It is nonetheless interesting to see what impact these appliances have on the adoption of the cloud computing concept. On a similar note, (Krautheim, 2009) has developed an infrastructure called PVI for the cloud that automates provisions depending on security settings. It would be interesting to see how the SeCA model can be connected to the presented PVI.

Although some cloud providers are certified, the impact of that certification on the real security of the services the provider offers is not always known. SAS70 for example does not offer any concrete security, it only offers a framework for auditing internal controls. The cloud provider will need to list its internal controls for any user to see what has been audited. It might be interesting to see how cloud providers use that information, what they do with it and whether the certifications really add up to extra level of security that is said it adds.

REFERENCES

Asadoorian, P. (2007). Escaping From The Virtualization Cave. *PaulDotCom*. Retrieved July 12, 2011, from http://www.pauldotcom.com/2007/07/31/escaping_from_the_virtualizati.html

Benson, T., Sahu, S., Akella, A., & Shaikh, A. (2010). A First Look At Problems In The Cloud. *Proceedings of the 2nd USENIX conference on Hot topics in cloud computing* (pp. 1-7). Boston, Massachusetts: USENIX Association. doi:10.1.1.150.2883

Chen, Y., Paxson, V., & Katz, R. H. (2010). *What's New About Cloud Computing Security*. *University of California, Berkeley Report No. UCB/EECS-2010-5 January* (Vol. 20). Berkeley, CA. Retrieved from http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-5.pdf

Christodorescu, M., Sailer, R., Schales, D. L., Sgandurra, D., & Zamboni, D. (2009). Cloud Security Is Not (Just) Virtualization Security. *Proceedings of the 2009 ACM workshop on Cloud computing security - CCSW '09* (p. 97). New York, New York, USA: ACM Press. doi:10.1145/1655008.1655022

Dalkey, N., & Helmer, O. (1963). An Experimental Application of the DELPHI Method to the Use of Experts. *Management Science, 9*(3), 458-467. JSTOR. doi:10.1287/mnsc.9.3.458

Das, A. S., & Srinathan, K. (2007). Privacy Preserving Cooperative Clustering Service. *15th International Conference on Advanced Computing and Communications (ADCOM 2007)* (pp. 435-440). Guwahati, India: Ieee. doi:10.1109/ADCOM.2007.52

Feigenbaum, J., Pinkas, B., Ryger, R. S., & Saint Jean, F. (2004). Secure computation of surveys. *EU Workshop on Secure* (pp. 1-6). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.3258&rep=rep1&type=pdf

Forum, J. (2009). *Cloud Cube Model: Selecting Cloud Formations for Secure Collaboration*. Retrieved from http://www.opengroup.org/jericho/cloud_cube_model_v1.0.pdf

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud Computing and Grid Computing 360-degree Compared. *2008 Grid Computing Environments Workshop* (pp. 1-10). Austin, TX: Ieee. doi:10.1109/GCE.2008.4738445

Ghinste, B. V. (2010). Gartner: Private Cloud Computing Plans From Conference Polls. *MSDN Blogs*. Retrieved June 27, 2011, from http://blogs.msdn.com/b/architectsrule/archive/2010/05/07/gartner-private-cloud-computing-plans-from-conference-polls.aspx

Gilder, G. (2006). The Information Factories. *Wired.com*. Retrieved June 27, 2011, from http://www.wired.com/wired/archive/14.10/cloudware_pr.html

Goldreich, O. (2000). Secure Multi-Party Computation. *Working Draft*. New York, New York, USA: Citeseer. doi:10.1145/508172.508174

Hao, F., Lakshman, T., Mukherjee, S., & Song, H. (2010). Secure Cloud Computing With A Virtualized Network Infrastructure. *Proceedings of the 2nd USENIX conference on Hot topics in cloud computing* (pp. 16–16). Boston, Massachusetts: USENIX Association. doi:10.1234/12345678

Hogben, G., & Catteddu, D. (2009). *Cloud Computing: Benefits, Risks and Recommendations for Information Security*. (C. Serrão, V. Aguilera Díaz, & F. Cerullo, Eds.)*Web Application Security* (Vol. 72). Crete, Greece: Springer Berlin Heidelberg. Retrieved from http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing-risk-assessment/at_download/fullReport

Hu, H., & Xu, J. (2009). Non-Exposure Location Anonymity. *2009 IEEE 25th International Conference on Data Engineering* (pp. 1120-1131). Shanghai, China: IEEE. doi:10.1109/ICDE.2009.106

Jensen, M., Schwenk, J., Gruschka, N., & Iacono, L. L. (2009). On Technical Security Issues in Cloud Computing. *2009 IEEE International Conference on Cloud Computing* (pp. 109-116). Bangalore, India: IEEE. doi:10.1109/CLOUD.2009.60

Kaliski Jr, B. S., & Pauley, W. (2010). Toward Risk Assessment As A Service In Cloud Environments. *Proceedings of the 2nd USENIX conference on Hot topics in cloud computing* (pp. 13–13). Boston, Massachusetts: USENIX Association. Retrieved from http://portal.acm.org/citation.cfm?id=1863116

Ko, S. Y., Jeon, K., & Morales, R. (2011). The HybrEx Model for Confidentiality and Privacy in Cloud Computing. *Proceedings of the 2011 conference on Hot topics in cloud computing*. Portland, OR. Retrieved from http://www.usenix.org/event/hotcloud11/tech/final_files/Ko.pdf

Krautheim, F. J. (2009). Private Virtual Infrastructure For Cloud Computing. *Proceedings of the 2009 conference on Hot topics in cloud computing* (pp. 5–5). San Diego, CA: USENIX Association. Retrieved from http://portal.acm.org/citation.cfm?id=1855538

Li, A., Yang, X., Kandula, S., & Zhang, M. (2010). CloudCmp: Shopping For A Cloud Made Easy. *Proceedings of the 2nd USENIX conference on Hot topics in cloud computing* (pp. 5–5). Boston, Massachusetts: USENIX Association. Retrieved from http://portal.acm.org/citation.cfm?id=1863108

McAfee. (2009). Technical FAQ. *Security*. McAfee: Santa Clara, CA. Retrieved from http://www.mcafee.com/us/resources/white-papers/wp-saas-faq.pdf

Mehta, N., & Smith, R. (2007). VMWare DHCP Server Remote Code Execution Vulnerabilities. *IBM Internal Security Systems*. Retrieved July 12, 2011, from http://www.iss.net/threats/275.html

Mell, P., & Grance, T. (2010). *NIST Definition of Cloud Computing v15*. Washington, DC. Retrieved from http://www.csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

Mowbray, M., & Pearson, S. (2009). A Client-Based Privacy Manager For Cloud Computing. *Proceedings of the Fourth International ICST Conference on COMmunication System softWAre and middlewaRE - COMSWARE '09* (p. 1). Dublin, Ireland: ACM Press. doi:10.1145/1621890.1621897

Mulholland, A., Pyke, J., & Fingar, P. (2010). *Enterprise Cloud Computing*. Tampa, FL: Meghan-Kiffer Press.

Ormandy, T. (2007). An Empirical Study Into The Security Exposure To Host Of Hostile Virtualized Environments. *Proceedings of CanSecWest Applied Security Conference*. Vancouver., Canada: Citeseer. doi:10.1.1.105.6943

Peterson, Z., & Gondree, M. (2011). A Position Paper on Data Sovereignty: The Importance of Geolocating Data in the Cloud. *Proceedings of the 2011 conference on Hot topics in cloud computing*. Portland, OR. Retrieved from http://www.usenix.org/event/hotcloud11/tech/final_files/Peterson.pdf

Popa, L., Yu, M., Ko, S. Y., Ratnasamy, S., & Stoica, I. (2010). CloudPolice: Taking Access Control Out Of The Network. *Proceedings of the Ninth ACM SIGCOMM Workshop on Hot Topics in Networks* (p. 7). Monterey, CA: ACM. Retrieved from http://portal.acm.org/citation.cfm?id=1868454

Reuters. (2008). What On Earth Is "Cloud Computing"? *Reuters*. Retrieved June 27, 2011, from http://blogs.reuters.com/mediafile/2008/09/25/what-on-earth-is-cloud-computing/

Richter, W., Ammons, G., Harkes, J., Goode, A., Bila, N., de Lara, E., Bala, V., et al. (2011). Privacy-Sensitive VM Retrospection. *Proceedings of the 2011 conference on Hot topics in cloud computing* (pp. 1-6). Portland, OR. Retrieved from http://www.usenix.org/events/hotcloud11/tech/final_files/Richter.pdf

Ristenpart, T., Tromer, E., Shacham, H., & Savage, S. (2009). Hey, You, Get Off My Cloud: Exploring Information Leakage In Third-Party Compute Clouds. *Proceedings of the 16th ACM conference on Computer and communications security* (pp. 199–212). Chicago, IL: ACM. doi:10.1.1.150.681

School, W. B. (2005). Delphi Decission Aid. Retrieved October 5, 2010, from http://armstrong.wharton.upenn.edu/delphi2/

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi Method for Graduate Research. *Journal of Information Technology Education*, *6*, 1-21. doi:10.1.1.151.8144

Tiwana, B., Balakrishnan, M., Aguilera, M. K., Ballani, H., & Mao, Z. M. (2010). Location, Location, Location!: Modeling Data Proximity in the Cloud. *Proceedings of the Ninth ACM SIGCOMM Workshop on Hot Topics in Networks* (p. 15). Monterey, CA: ACM. doi:10.1145/1868447.1868462

Troncoso-Pastoriza, J. R., & Pérez-González, F. (2010). CryptoDSPs for Cloud Privacy. *Workshop on Cloud Information System Engineering (CISE'10)* (pp. 1-12). Hong Kong, China. doi:10.1.1.185.429

Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. (2008). A Break In The Clouds: Towards A Cloud Definition. *ACM SIGCOMM Computer Communication Review*, *39*(1), 50–55. ACM. Retrieved from http://portal.acm.org/citation.cfm?id=1496100

Vigfusson, Y., & Chockler, G. (2010). Clouds At The Crossroads. *Crossroads*, *16*(3), 10-13. doi:10.1145/1734160.1734165

Voas, J., & Zhang, J. (2009). Cloud Computing: New Wine or Just a New Bottle? *IT professional*, *11*(2), 15–17. IEEE. doi:10.1109/MITP.2009.23

Wang, Chen, & Zhou, Y. (2010). A Collaborative Monitoring Mechanism for Making a Multitenant Platform Accountable. *Proceedings of the 2nd USENIX conference on Hot*

Topics in Cloud Computing (pp. 18-25). Boston, Massachusetts: ACM. Retrieved from http://www.usenix.org/event/hotcloud10/tech/full_papers/WangC.pdf

Wang, Qian, Wang, Cong, Li, J., Ren, K., & Lou, W. (2009). Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing. In M. Backes & P. Ning (Eds.), *LNCS, ESORICS 2009* (5789th ed., Vol. 5789, pp. 355-370). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-04444-1

Wang, Qi, Luo, Y., & Huang, L. (2008). Privacy-preserving Protocols for Finding the Convex Hulls. *2008 Third International Conference on Availability, Reliability and Security*, (070412043), 727-732. Ieee. doi:10.1109/ARES.2008.11

Williams, D., Elnikety, E., Eldehiry, M., Jamjoom, H., Huang, H., & Weatherspoon, H. (2011). Unshackle the Cloud! *Proceedings of the 2011 conference on Hot topics in cloud computing*. Portland, OR. Retrieved from http://www.cs.cornell.edu/~djwill/pubs/unshackle.pdf

Wood, T., Cecchet, E., Ramakrishnan, K., Shenoy, Prashant, Van Der Merwe, J., & Venkataramani, A. (2010). Disaster Recovery As A Cloud Service: Economic Benefits & Deployment Challenges. *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing* (pp. 8–8). USENIX Association. Retrieved from http://portal.acm.org/citation.cfm?id=1863111

Wood, T., Gerber, A., Ramakrishnan, K., Shenoy, P., & Van der Merwe, J. (2009). The Case For Enterprise-Ready Virtual Private Clouds. *Proceedings of the 2009 conference on Hot topics in cloud computing* (pp. 4–9). Monterey, CA: USENIX Association. Retrieved from http://portal.acm.org/citation.cfm?id=1855537

Yao, A. C. (1982). Protocols for Secure Computations. *23rd Annual Symposium on Foundations of Computer Science* (pp. 160-164). Chicago, IL: IEEE. doi:10.1109/SFCS.1982.38

APPENDIX

SeCA Data classification Template	Date: _____
Classification Name/ identification: _____	Expert's Name:_____
Regional:	
Geo-spatial:	
Governance & Compliance:	
Delivery Model:	○ IaaS ○ PaaS ○ SaaS
Deployment model:	○ Private ○ Partner/Community ○ Public ○ Hybrid
Encryption:	
Network:	○ Within ○ Outside ○ Any
Premises:	○ On premise ○ Off premise ○ Any