Cloud Computing & Confidentiality

Master of Science graduation thesis Computer Science

Guido Kok | May 24, 2010

UNIVERSITY OF TWENTE.

University of Twente:
dr.ir. W. Pieters
prof.dr. P.H. Hartel

Capgemini

Capgemini:
ing. T. v.d. Meer

Abstract

Cloud computing is an upcoming paradigm that offers tremendous advantages in economical aspects, such as reduced time to market, flexible computing capabilities, and limitless computing power. To use the full potential of cloud computing, data is transferred, processed and stored by external cloud providers. However, data owners are very skeptical to place their data outside their own control sphere.

This thesis discusses to which degree this skepticism is justified, by presenting the Cloud Computing Confidentiality Framework (CCCF). The CCCF is a step-by-step framework that creates mapping from data sensitivity onto the most suitable cloud computing architecture. To achieve this, the CCCF determines first of all the security mechanisms required for each data sensitivity level, secondly which of these security controls may not be supported in certain computing environments, and finally which solutions can be used to cope with the identified security limitations of cloud computing.

The most thorough security controls needed to protect the most sensitive data may not be guaranteed in public cloud computing architectures, while they can be realized in private cloud computing architectures. As the most promising cloud computing approach, this thesis suggests selective cloudbursting, which acts as a hybrid cloud model with selective data transfers between public and private clouds.

Preface

In the process towards my graduation I used various sources of knowledge and experience that pointed me into the direction needed to fulfill the requirements set in the context of this project.

During the main part of my graduation project, I was active within Capgemini NL, where I met several inspiring people. I would like to thank all the professionals who were willing to reserve some time to discuss my graduation topic. In particular, I would like to thank Jan van de Ven, Theo Schalke, Lee Provoost, Martijn Linssen, Martin Kroonsberg, and Hans Peter van Riemsdijk, who gave their intellectual and professional opinion on the development of the framework in this graduation project.

I would also like to thank my fellow graduation students whom I met during my time at Capgemini. With fewer resources due to the economic recession, not many graduation students were active within Capgemini. Therefore, I was happy that I could enjoy the time there with Klaas Tjepkema, Michiel Bax, and Lucas Baarspul. The exchange of ideas and approaches to tackle problems was very helpful. Thanks guys, we will meet again.

Finally, I would like to give special attention to my tutor in the last part of the project. Even though he was not very acquainted with the academic process of master-graduation, he was almost always willing and able to find time to help me, either with a totally independent view on the content of this project, or with the difficulties that every graduation student has in the process of writing his thesis. Together with his wife he was always prepared to support me, and for that I am immensely thankful. Mom and dad, thank you so much.

Guido Kok, May 2010

Table of contents

List of figures

List of tables

1 Introduction

Cloud computing is the collective term for a group of IT technologies which in collaboration are changing the landscape of how IT services are provided, accessed and paid for. Some of the supporting technologies have already been available for quite some time, but it is the combination of several technologies which enables a whole new way of using IT.

There is a lot of discussion of what cloud computing exactly is. The U.S. National Institute of Standards and Technology (NIST) have put an effort in defining cloud computing, and as NIST's publications are generally accepted, their definition of cloud computing will be used in this thesis. The NIST definition of cloud computing is (NIST 2009a):

> Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

To explain the definition in short, "convenient on-demand network access", together with "minimal management effort or service provider interaction," stands for easy and fast network access to resources that are ready to use. With a "shared pool of resources," the available computing resources of a cloud provider are combined as one big collection, to serve all users. The "rapid provisioning and releasing" of computing resources is used to quickly match available resources, with the need for those resources. This rapid provisioning prevents a lack of computing power when the need increases, while rapid release of assigned resources prevents that resources are idle while they may be required elsewhere.

The above definition is by no means exhaustive and it is very hard to find two experts having the same definition of cloud computing. Cloud computing is still an evolving paradigm. The characteristics, deployment and delivery models, as well as the underlying risks, technologies, issues and benefits will be refined by energetic debate by both the public and the private sectors. A more elaborate explanation of these cloud properties will be discussed in chapter 2.

As with most new technologies and paradigms, one tends to look for the functionality first and only later on, one looks after the security of such functionality. However, cloud computing raises such an amount of questions concerning security guarantees that potential users are waiting for clear answers before moving into the cloud.

1.1 Research motivation and objectives

Cloud computing users work with data and applications that are often located off-premise. However, many organizations are uncomfortable with the idea of having their data and applications on systems they do not control. There is a lack of knowledge on how cloud computing impacts the confidentiality of data stored, processed and transmitted in cloud computing environments.

The goal of this thesis is to create a framework that clarifies the impact of cloud computing on confidentiality preservation, by making stepwise recommendations on;

- How data can be classified on confidentiality
- How data classifications relate to the security controls needed to preserve the confidentiality of data

- How the process of security control selection is negatively influenced in cloud computing environments
- How to cope with the negative influences of cloud computing on the protection of data confidentiality.

1.2 Research questions

In order to achieve the research objectives stated above, the necessary knowledge will need to be obtained and combined. The following research questions will guide this research:

- Which data classifications are used today and what are their security requirements with respect to confidentiality?
- Which cloud architectures are available and what security controls do they have in place with respect to confidentiality?
- How can we classify cloud architectures on the area of confidentiality?
- How can we create a mapping from confidential data classes to cloud architectures operating on potentially confidential data?

1.3 Research scope

A broad approach of classifying assets and networks on the topic of security, is investigating the security objectives Confidentiality, Integrity and Availability (CIA). Combining these three objectives in one research project would be too much work for the period of time this research is conducted in. In this thesis we focus on *confidentiality*, as that is where the biggest concerns are at this moment.

Data classification research has already been done extensively (Chen and Liu 2005; Morsi, El-fouly and Badr 2006; Grandison, Bilger, O'Connor et al. 2007), this thesis will use the results of these researches and analyze the security requirements that need to be met in order to protect data confidentiality.

We will elaborate on the research methodology in chapter 3.

1.4 Capgemini

This thesis is conducted as intern at Capgemini NL. Capgemini helps clients deal with changing business and technology issues. Capgemini brings experience, best practices and tools to apply to clients unique requirements.

As the cloud computing paradigm appeared as a new and promising technology, a lack of knowledge on this topic was identified by Capgemini employees. The need for more knowledge on this area was translated to a thesis subject.

Capgemini NL operates in three disciplines (Technology, Outsourcing and Consulting) and is divided in four sectors (Financial Services, Telecom Travel & Utilities, Public and Products), as shown in Figure 1-1.

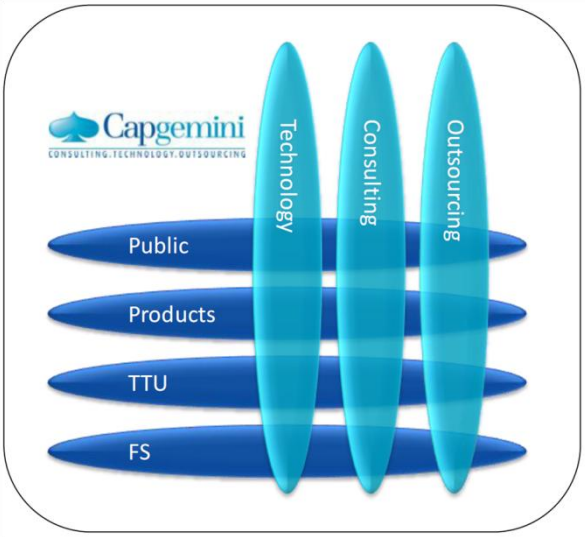


Figure 1-1: Capgemini NL Company structure

This research is executed within the sector Products in which there are 6 practices;

- Products Market Solutions
- Architecture, Governance & Infrastructure (AG&I)
- SAP Process & Industry Solutions
- SAP Netweaver & Intelligence
- TC&OS
- Business Intelligence Management

This thesis is written for the practice Architecture, Governance & Infrastructure, and the section Infrastructure in particular.

1.5 Thesis structure

The thesis is divided into six sections, which will be discussed here one by one.

We will elaborate on the paradigm cloud computing in chapter two, where the key characteristics, service models, deployment models, and security issues related to cloud computing will be discussed. Chapter three discusses the research methodology we will use in this thesis, explaining the tools we will use in the upcoming two chapters.

In the literature review chapter, we conduct a systematic literature search and analysis on topics of cloud computing and confidentiality in order to find answers to the research questions. We need to supplement the knowledge obtained in the literature review, as the literature review does not provide us with all the needed information to construct the framework. This supplementing research involves present-day security practices and our interpretation of them, and will be discussed in chapter five before we present the conceptual framework.

The Cloud Computing Confidentiality Framework (CCCF) is fully presented in chapter six, where we show how the current processes of IT risk management, data & system classification, and security control selection, will identify security problems in cloud environments. With the identified security problems, the CCCF presents a mapping from data classifications to appropriate cloud architectures, and show how the security problems can be anticipated.

The development of the CCCF included the influence of several consultants and security experts in the field. Interviews were conducted to discuss the development and the goal of the framework. These interviews are presented in chapter seven, together with the influence of these interviews on the development of the CCCF.

In the last chapter of the research the discussion takes place. In this chapter the conclusions are presented while practical implications, research limitations and suggestions for further research are discussed.

2 Background

As the paradigm of cloud computing is relatively new, there are various open issues which need to be resolved before cloud computing is fully accepted by the broad community. Before we will dive into the research methodology and the issues this thesis is about, a deeper explanation is needed of what cloud computing encompasses.

The NIST definition of cloud computing mentioned in the introduction will be used as our starting point. To recall the definition:

> Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

The above definition is supported by five key *cloud characteristics*, three *delivery models* and four *deployment models* (NIST 2009a). These supporting properties will be explained below, after which we will discuss various security issues and concerns related to cloud computing.

2.1 Cloud key characteristics

On-demand self-service. Cloud computing resources can be procured and disposed of by the consumer without human interaction with the cloud service provider. This automated process reduces the personnel overhead of the cloud provider, cutting costs and lowering the price at which the services can be offered.

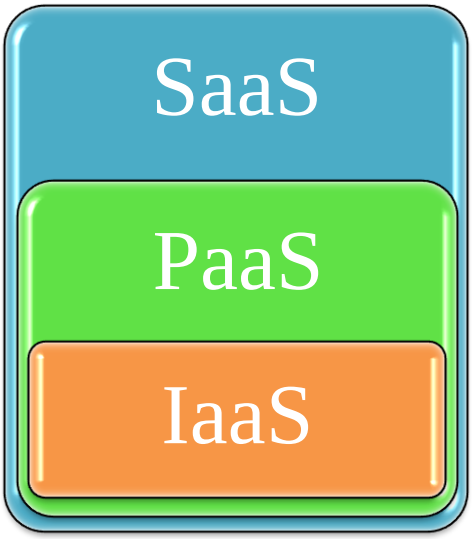
Resource pooling. By using a technique called "virtualization," the cloud provider pools his computing resources. This resource pool enables the sharing of virtual and physical resources by multiple consumers, "dynamically assigning and releasing resources according to consumer demand" (NIST 2009a). The consumer has no explicit knowledge of the physical location of the resources being used, except when the consumer requests to limit the physical location of his data to meet legal requirements.

Broad network access. Cloud services are accessible over the network via standardized interfaces, enabling access to the service not only by complex devices such as personal computers, but also by light weight devices such as smart phones.

Rapid elasticity. The available cloud computing resources are rapidly matched to the actual demand, quickly increasing the cloud capabilities for a service if the demand rises, and quickly releasing the capabilities when the need for drops. This automated process decreases the procurement time for new computing capabilities when the need is there, while preventing an abundance of unused computing power when the need has subsided.

Measured service. Cloud computing enables the measuring of used resources, as is the case in utility computing. The measurements can be used to provide resource efficiency information to the cloud provider, and can be used to provide the consumer a payment model based on "pay-per-use." For example, the consumer may be billed for the data transfer volumes, the number of hours a service is running, or the volume of the data stored per month.

2.2 Cloud service models



Software-as-a-Service (SaaS). The SaaS service model offers the services as applications to the consumer, using standardized interfaces. The services run on top of a cloud infrastructure, which is invisible for the consumer. The cloud provider is responsible for the management the application, operating systems and underlying infrastructure. The consumer can only control some of the user-specific application configuration settings.

Platform-as-a-Service (PaaS). The PaaS service model offers the services as operation and development platforms to the consumer. The consumer can use the platform to develop and run his own applications, supported by a cloud-based infrastructure. "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" (NIST 2009a).

Infrastructure-as-a-Service (IaaS). The IaaS service model is the lowest service model in the technology stack, offering infrastructure resources as a service, such as raw data storage, processing power and network capacity. The consumer can the use IaaS based service offerings to deploy his own operating systems and applications, offering a wider variety of deployment possibilities for a consumer than the PaaS and SaaS models. "The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls)" (NIST 2009a).

2.3 Cloud deployment models

Regardless of which delivery model is utilized, cloud offerings can be deployed in four primary ways, each with their own characteristics. The characteristics to describe the deployment models are; (i) *who owns* the infrastructure; (ii) *who manages* the infrastructure; (iii) *where is* the infrastructure *located*; (iv) and *who accesses* the cloud services.

Public clouds. Public cloud computing is based on massive scale offerings to the general public. The infrastructure is located on the premises of the provider, who also owns and manages the cloud infrastructure. Public cloud users are considered to be untrusted, which means they are not tied to the organization as employees and that the user has no contractual agreements with the provider.

Private clouds. Private clouds run in service of a single organization, where resources are not shared by other entities. "The physical infrastructure may be owned by and/or physically located in the organization's datacenters (on-premise) or that of a designated service provider (off-premise) with an extension of management and security control planes controlled by the organization or designated service provider respectively" (Bardin, Callas, Chaput et al. 2009). Private cloud users are considered as trusted by the organization, in which they are either employees, or have contractual agreements with the organization.

Community clouds. Community clouds run in service of a community of organizations, having the same deployment characteristics as private clouds. Community users are also considered as trusted by the organizations that are part of the community.

Hybrid clouds. Hybrid clouds are a combination of public, private, and community clouds. Hybrid clouds leverage the capabilities of each cloud deployment model. Each part of a hybrid cloud is connected to the other by a gateway, controlling the applications and data that flow from each part to the other. Where private and community clouds are managed, owned, and located on *either* organization *or* third party provider side per characteristic, hybrid clouds have these characteristics on *both* organization *and* third party provider side. The users of hybrid clouds can be considered as trusted and untrusted. Untrusted users are prevented to access the resources of the private and community parts of the hybrid cloud.

Table 2-1 summarizes the four primary cloud deployment models. It should be noted that there are initiatives for deployment models that not necessarily fall inside one of the above categorizations. For example, Amazon offers virtual private clouds, that use public cloud infrastructure in a private manner, connecting the public cloud resources to the organizations internal network (Amazon 2009b).

	Managed by	Infrastructure Owned By	Infrastructure Located	Accessible and Consumed By
Public	3rd party provider	3rd party provider	Off-premise	Untrusted
Private / Community	3rd party provider / Organization	3rd party provider / Organization	Off-premise / On-premise	Trusted
Hybrid	Both Organization & 3rd party provider	Both Organization & 3rd party provider	Both On-premise & Off-premise	Trusted and Untrusted

Table 2-1: Cloud deployment models (Bardin et al. 2009)

The Cloud Security Alliance points out that is difficult to describe an entire cloud service using a single label, because it attempts to describe the following elements (Bardin et al. 2009):

- Who manages it
- Who owns it
- Where is it located
- Who has access to it
- How is it accessed

The answers to the above questions result in multiple flavors of cloud service offerings. The thing to keep in mind is that the above characteristics "that describe *how* Cloud services are deployed, are often used interchangeably with the notion of *where* they are provided; as such, you may often see public and private clouds referred to as 'external' or 'internal' clouds. This can be very confusing" (Bardin et al. 2009).

The way traditional services are offered, is often described in terms of where the security perimeter of the service provider is located. The security perimeter between networks is often implemented as a firewall. When we consider cloud services, using the firewall as a clear demarcation of the security boundary is an outdated concept, as we will explain in the next section.

2.4 Cloud security issues

Although it is important to describe the location of the security perimeter in relation to the assets to be protected, using the terms external clouds and internal clouds would indicate a well-defined perimeter between the outside and the protected inside. This separation is an anachronistic concept due to the de-perimeterization and the loss of trust boundaries resulting from the increasing need of companies to collaborate and provide ubiquitous access to employees, consumers and contractors.

Traditional security controls may be incapable to handle the shift from secure silos of data with strict trust boundaries and well defined access control, to the complex scenarios where access is ubiquitous, information exchange is abundant and data location is often unknown. Cloud computing accelerates this erosion of trust and security boundaries.

With cloud computing, organizations can use services and store data outside their own control. This development raises security questions and should induce a degree of skepticism before using cloud services. In his article, Brodkin discusses a study of Gartner, which points out seven areas of concern around security issues in cloud computing (Brodkin 2008):

Privileged user access

Data stored and processed outside the enterprises direct control, brings with "an inherent level of risk, because outsourced services bypass the physical, logical and personnel controls IT shops exert over in-house programs" (Brodkin 2008). Brodkin advises to get as much information as possible about the people who manage your data and the controls they implement.

Regulatory compliance

Data owners are responsible for the integrity and confidentiality of their data, even when the data is outside their direct control, which is the case with external service providers such as cloud providers. Where traditional service providers are forced to comply to external audits and obtain security certifications, so should cloud computing providers: "Cloud computing providers who refuse to undergo this scrutiny are signaling that customers can only use them for the most trivial functions" (Brodkin 2008).

Most, if not all, of the leading cloud providers do not support on-site external audits on customers request. As a result, some compliances cannot be achieved because on-site auditing is a requirement that cannot be satisfied, for example the Payment Card Industry level 1 compliancy.

Data location

The exact location of data in the cloud is often unknown. Data may be located in systems in other countries, which may be in conflict with regulations prohibiting data to leave a country or union. Gartner advises to investigate if cloud providers will commit to keeping data in specific jurisdictions and whether the providers will make contractual commitments to obey local privacy requirements on behalf of their customers (as cited in Brodkin, 2009).

For example, the EU Data Protection Directive places restrictions on the export of personal data from the EU to countries whose data protection laws are not judges as "adequate" by EU standards (EuropeanCommission 1995a). If not properly attended to, European personal data may be located outside the EU without being compliant to the directive.

Data segregation

The shared, massive scale characteristics of cloud computing makes it likely that one's data is stored alongside data of others consumers. Encryption is often used to segregate data-at-rest, but it is not a cure-all. It is advised to do a thorough evaluation of the encryption systems used by the cloud provider. A proper built, but poorly managed encryption scheme may be just as devastating as no encryption at all, because although the confidentiality of data may be preserved, availability of data may be at risk when data availability is not guaranteed.

Recovery

Cloud providers should have recovery mechanisms in place in case of a disaster. "Any offering that does not replicate the data and application infrastructure across multiple sites is vulnerable to a total failure," Gartner says (as cited in Brodkin, 2009). Cloud providers should provide its guidelines concerning business continuity planning, detailing how long it will take for services to be fully restored.

Investigative support

Gartner warns that "investigating inappropriate or illegal activity may be impossible in cloud computing, because logging and data may be co-located and spread across ever-changing sets of hosts and data centers" (Brodkin 2008). If cloud providers cannot provide customers with a contractual statement specifying support for incorruptible logging and investigation, Gartner says that "the only safe assumption is that investigation and discovery requests will be impossible" (Gartner 2008).

Data Lock-in

Availability of customers data may be at risk if a cloud provider goes broke or is acquired by another organization. Providers should provide procedures how customers can retrieve their data when the needed, and at least as important; in which format the data is presented to the customer. If the data is presented in a format proprietary to the cloud provider, it may be unusable by any other provider. The use of open standards by providers to prevent data lock-in is recommended, but not always supported.

Of the above security issues, the issues related to availability of services are well attended to by researchers and cloud service providers. The largest uncertainties linger around issues related to confidentiality of data, such as data location, access control and regulatory compliance. As such, this thesis focuses on the confidentiality aspects and issues of cloud computing. In the following chapter we will discuss how the research in this thesis is going to be performed.

3 Research methodology

This chapter will describe the approach that has been taken in this research. The steps taken in the subsequent chapters will be explained without diving into the results.

3.1 Orientation

The research starts with the orientation on the area of cloud computing, what is cloud computing about and which security issues are in dire need of investigation. By consulting websites of current cloud service offerings, reading news articles, participating in seminars and discussing cloud computing and security issues with professionals within Capgemini, the research questions of this research are formulated.

To answer the research questions stated in section 1.2, knowledge must be obtained that supplements the information found during the orientation on the topic. As finding information on the web on groundbreaking technologies is a very time-consuming process, this research employs a structured method to obtain high quality information, called a Literature Review.

3.2 Literature review

To explore the available knowledge on the area of cloud computing and confidentiality, a literature review is conducted using a systematic approach. The role of a literature review is depicted in Figure 3-1**.** The objectives of a literature review are:

- To understand the current state of knowledge in a research area
  - What is known/generally accepted
  - What questions remain unanswered
  - Where do conflicting results exist
- To show how the current research project is linked to previous research (cumulative tradition)
- To summarize and synthesize previous research
- To critically analyze previous research: strengths and weaknesses
- To learn from others and stimulate ideas

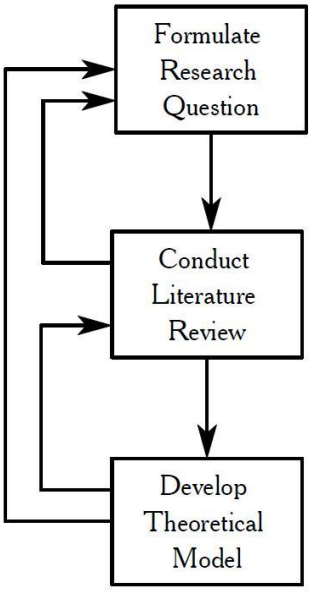


Figure 3-1: Literature Review Role

The first step in a literature review is selecting the top 25 journals to search information in. This ranking is researched and published by several groups, of which the Association of Information Systems is the most recent one (AIS 2009a). The second step is selecting one or more search engines that index these top 25 journals, after which the journals can be examined by searching on a predetermined set of keywords.

Analyzing the results of this top down search will filter out a fair share of results due to irrelevance. Supplementing the shrunken set of results can be achieved by conducting a bottom up search, using both backward and forward citation analysis. The former relates to finding papers referenced by papers found earlier, while the latter is an acronym for finding papers that cite papers we have found earlier, using search engines.

The papers found in the search are analyzed to distill useful concepts with respect to our research. Papers containing topics such as privacy, IT regulation and security in distributed environments, are scrutinized for dimensions to be used in our mapping from confidential data classes to cloud

architectures. The complete process and the results of the Literature Review are presented in chapter 4.

3.3 Design & specification of the framework

The dimensions found in the literature review act as the starting point in the design and specification of the mapping from data classes to cloud architectures. The combination of the dimensions and additional information should result in a model that shows the impact of cloud computing on the protection of sensitive data.

As the dimensions by themselves are not related enough to form a model, additional knowledge must be gained to build a model. With the dimensions in mind, the design and specification phase of this research consists of an ongoing process of discussions with security experts within Capgemini, and more research on existing literature.

The development towards our final model took several revisions, in where the earlier versions were centered around data classification, which was based on standards of the National Institute of Standardization and Technology (NIST). During discussions with security experts about the development of the model, this central position of data classification was found to be too shortsighted. In combination with more literature research on the topics of risk management and security controls, more components were added to the model that would give the model a clear relation with the current approach in IT risk management.

Continuing on this, the literature review dimension should be related to the processes of data classification and security control selection, and peculiarities should be identified that show the unique features of cloud computing that influence the control selection phase, from the confidentiality point of view.

The results of the research described above are integrated into chapters 5 and 6. The research methodology is depicted in Figure 3-2.

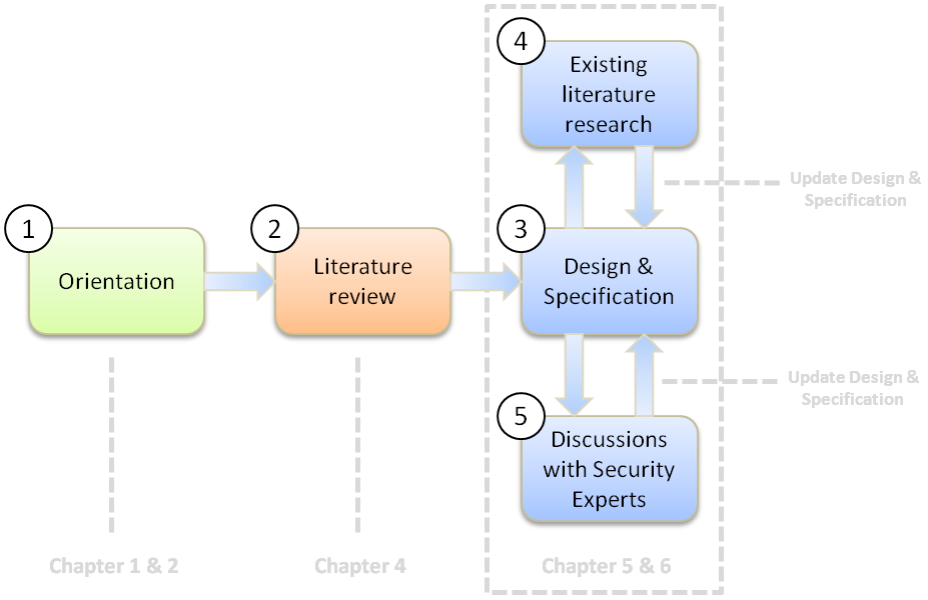


Figure 3-2: Research model used: Start with the orientation on the topic and formulation of the research questions (1). Acquire knowledge from a literature review (2). Produce a detailed design of the framework, based on the literature review (3). Acquire additional knowledge (repeated) about framework components and update the framework design and specification if required (4). Discuss the framework development (repeated) with security experts and update the framework design and specification if required (5).

4 Literature review

In this chapter the process that is used to perform a structured literature search will be presented. The goal of the literature review is to cover all relevant scientific literature of top quality. First, the most important contributions are identified and analyzed. The selected articles are evaluated and the described concepts are synthesized resulting in an overview of the current state of knowledge. Based on the constructs and measures that are mentioned in the articles, a conceptual theoretical model is developed which shows the assumed causal relationships between the constructs.

4.1 Top ranked journal selection

The literature search was started by identifying the top 25 journals on which we can conduct our search. The Association of Information Systems published an overview of 9 journal ranking studies (AIS 2009a). Using this ranking directly has a severe drawback; the total ranking appointed to a journal is the sum of the rankings given to the journal, divided by the number of appearances in the ranking studies. For example, a journal that has been ranked 6th in only one study, ends up as the 6th journal in the overall ranking. In order to filter out this imperfection in the ranking, we decided to only include journals in our ranking that are mentioned in at least three different ranking studies. The results are presented in Table 4-1.

Top 25 MIS Journals	
1.	MIS Quarterly Management Information Systems
2.	Information Systems Research
3.	Communications of the ACM
4.	Management Science
5.	Journal of Management Information Systems
6.	Decision Sciences
7.	Harvard Business Review
8.	IEEE Transactions (various)
9.	European Journal of Information Systems
10.	Decision Support Systems
11.	Information and Management
12.	ACM Transactions on Database Systems
13.	IEEE Transactions on Software Engineering
14.	ACM Transactions
15.	MIT Sloan Management Review
16.	ACM Computing Surveys
17.	Academy of Management Journal
18.	Organization Science
19.	IEEE Transactions on Computers
20.	Information Systems Journal
21.	Administrative Science Quarterly
22.	Data Base for Advances in Information Systems
23.	Communications of the AIS
24.	Journal of the AIS
25.	Journal of Management Systems

Table 4-1: Top 25 MIS Journals (AIS 2009a)

Top 10 CS – Information Systems Journals	
1.	IEEE Transactions on Information Theory
2.	Journal of the ACM
3.	Information Processing Letters
4.	Journal of the American Medical Informatics Association: JAMIA
5.	MIS Quarterly: Management Information Systems
6.	Computer Journal
7.	IEEE Network
8.	Journal of the American Society for Information Science and Technology
9.	Computer Networks, The International Journal of Computer and Telecommunications Networking
10.	IEEE Transactions on Knowledge and Data Engineering

Table 4-2: Top 10 Information Systems Journals (RedJasper 2007)

As cloud computing is such a new paradigm, involving not only the Information Systems research area, but also the Computer Science research area, there was a general belief that the Top 25 MIS Journals would possibly not provide enough sources for literature for our review. We decided to supplement these 25 journals with two top 10 rankings on the research areas of Computer Science – Information Systems and Computer Science – Hardware and Architecture, as published on the

Journal-ranking.com website (RedJasper 2007). The method used for the ranking is the Journal Influence Index, which is the average number of times the published articles have been cited. The top 10 CS – Information Systems and CS – Hardware and Architecture journals are presented in Table 4-2 and Table 4-3, respectively.

Top 10 CS – Hardware and Architecture Journals	
1.	Communications of the ACM
2.	IEEE Transactions on Computers
3.	IEEE/ACM Transactions on Networking
4.	Journal of the ACM
5.	IBM Journal of Research and Development
6.	IEEE Transactions on Neural Networks
7.	IEEE Network
8.	Journal of Computer and System Sciences
9.	Computer
10.	IEEE Micro

Table 4-3: Top 10 CS - Hardware and Architecture Journals (RedJasper 2007)

4.2 Selection criteria

The main criterion for selecting articles is obviously the relevance to the research questions. Besides this main criterion, additional criteria are formulated:

- Articles must be published in the top ranked journals stated in Table 4-1, Table 4-2 and Table 4-3.
- Articles have to be written in English, Dutch or German.
- Articles have to be published in the year 2000 or later.

If selected articles have a very high relevance and/or high value to the research conducted in this paper, forward and backward citation analysis is performed to find more related articles. Forward analysis is the automated search for papers who refer the one found, while backward citation analysis refers to the classic analysis of older work, see Figure 4-1.

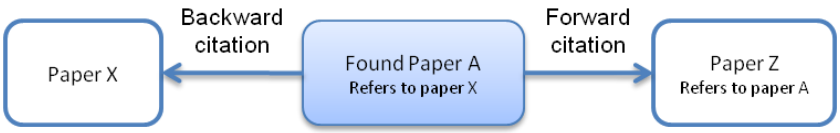


Figure 4-1: Forward and backward citation analysis

The above selection criteria do not apply to these indirect sources of information.

4.3 Search engine selection

Searching in the selected journals in a structured way can be achieved by using a search engine such as Scopus or Web of Science. We chose Scopus.com, based on previous experience with the search machine and the fact that only the following three out of the 36 distinct journals were not indexed by Scopus and had to be searched manually:

- Communications of the AIS (AIS 2009c)
- Journal of the AIS (AIS 2009b)
- Journal of Management Systems (Saee 2006)

Data Secrecy	Network Architecture
Data Classification	Grid Computing
Data Privacy	Virtualization
Confidential Information	

Table 4-4: Keywords with interesting results

4.4 Keyword selection and search query construction

The list of keywords used for the search has been extended quite heavily, as searches showed that there are very few articles which pass our selection criteria on direct keywords such as Cloud

Computing. We performed the search on 40 keywords, but for the sake of clarity we only present the keywords which resulted in interesting articles, as shown in Table 4-4. The complete list of keywords and the search results on them can be found in Appendix A.

To search with the search engine Scopus in the top ranked journals above, Scopus' Advanced Search offers the user a highly customized search query. The query we use is presented in Figure 4-2.

```
TITLE-ABS-KEY(keyword)
AND ( LIMIT-TO(EXACTSRCTITLE,"MIS Quarterly Management Information Systems" )
OR LIMIT-TO(EXACTSRCTITLE,"Information Systems Research" )
OR LIMIT-TO(EXACTSRCTITLE,"Communications of the ACM" )
OR LIMIT-TO(EXACTSRCTITLE,"Management Science" )
OR LIMIT-TO(EXACTSRCTITLE,"Journal of Management Information Systems" )
OR LIMIT-TO(EXACTSRCTITLE,"Decision Sciences" )
OR LIMIT-TO(EXACTSRCTITLE,"Harvard Business Review" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Transactions on Computers" )
OR LIMIT-TO(EXACTSRCTITLE,"European Journal of Information Systems" )
OR LIMIT-TO(EXACTSRCTITLE,"Decision Support Systems" )
OR LIMIT-TO(EXACTSRCTITLE,"Information and Management" )
OR LIMIT-TO(EXACTSRCTITLE,"ACM Transactions on Database Systems" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Transactions on Software Engineering" )
OR LIMIT-TO(EXACTSRCTITLE,"ACM Transactions" )
OR LIMIT-TO(EXACTSRCTITLE,"ACM Computing Surveys" )
OR LIMIT-TO(EXACTSRCTITLE,"Academy of Management Journal" )
OR LIMIT-TO(EXACTSRCTITLE,"Organization Science" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Transactions on Computers" )
OR LIMIT-TO(EXACTSRCTITLE,"Information Systems Journal" )
OR LIMIT-TO(EXACTSRCTITLE,"Administrative Science Quarterly" )
OR LIMIT-TO(EXACTSRCTITLE,"Data Base for Advances in Information Systems" )
OR LIMIT-TO(EXACTSRCTITLE,"Sloan Management Review" )
OR LIMIT-TO(EXACTSRCTITLE,"MIT Sloan Management Review" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Transactions on Information Theory" )
OR LIMIT-TO(EXACTSRCTITLE,"Journal of the ACM" )
OR LIMIT-TO(EXACTSRCTITLE,"Information Processing Letters" )
OR LIMIT-TO(EXACTSRCTITLE,"Journal of the American Medical Informatics Association: JAMIA" )
OR LIMIT-TO(EXACTSRCTITLE,"Computer Journal" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Network" )
OR LIMIT-TO(EXACTSRCTITLE,"Journal of the American Society for Information Science and Technology" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Transactions on Knowledge and Data Engineering" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE/ACM Transactions on Networking" )
OR LIMIT-TO(EXACTSRCTITLE,"IBM Journal of Research and Development" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Transactions on Neural Networks" )
OR LIMIT-TO(EXACTSRCTITLE,"Journal of Computer and System Sciences" )
OR LIMIT-TO(EXACTSRCTITLE,"Computer" )
OR LIMIT-TO(EXACTSRCTITLE,"IEEE Micro" ) )
```

Figure 4-2: The Scopus search query

This query searches for a match of the given keyword in the titles, abstracts and keywords of all the articles in the database, which are published in the given journals. The *Sloan Management Review* appears twice in the query, as this journal was renamed to *MIT Sloan Management Review* in 1996.

4.5 Search results

We performed our search on the 40 distinct keywords, using the search engine Scopus and manually consulting the three journals stated in Section 4.3. The results were meager, only 15 out of the 40 keywords returned any articles that were published in the selected journals. The articles found via these 15 keywords were screened on the other selection criteria of section 4.2, after which the titles and abstracts of these articles were analyzed to identify the relevance. Only 6 keywords out of the total set of 40 produced relevant articles. The results of the 6 useful keywords and two promising keywords are presented in Figure 4-3.

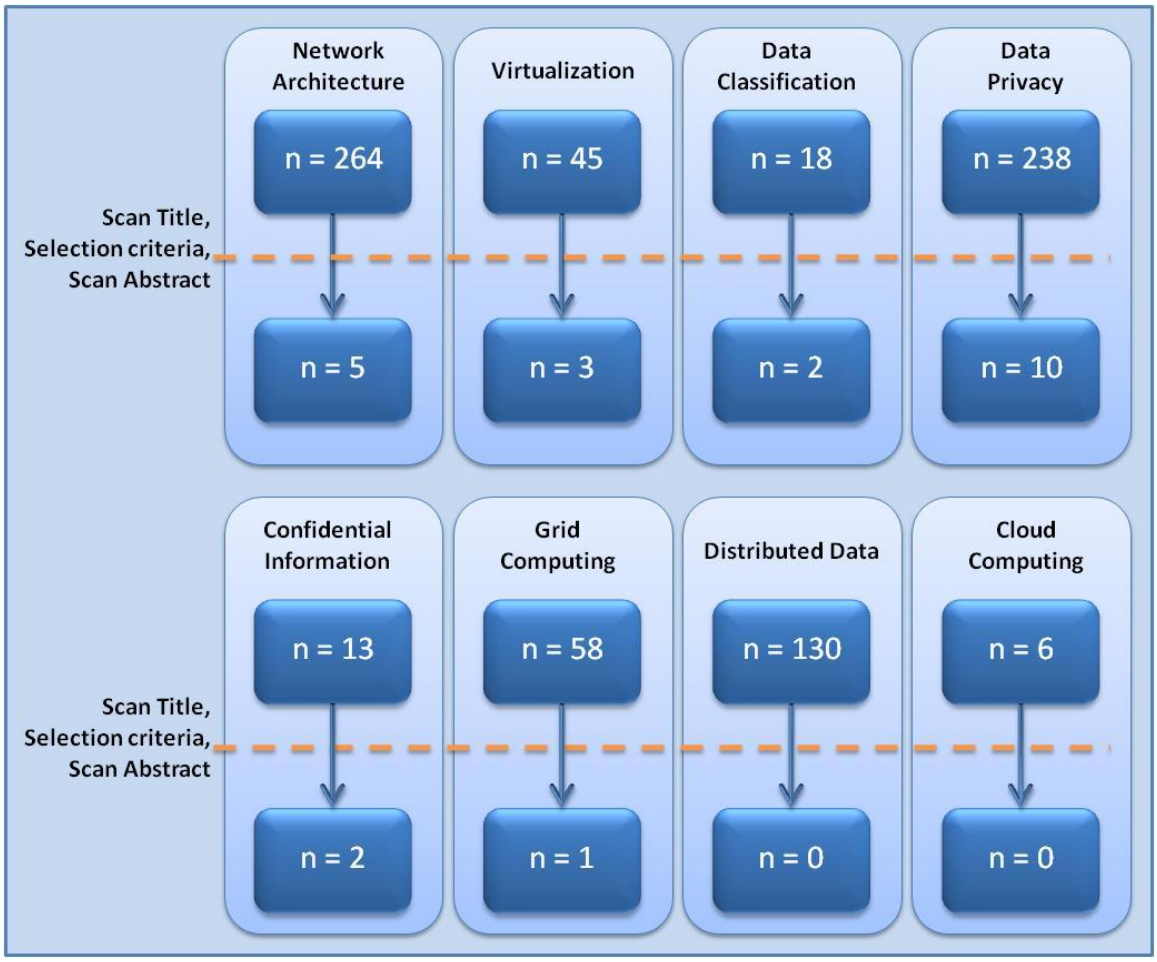


Figure 4-3: Literature search results

For example, the keyword "Network Architecture" had 264 hits in the selected set of journals. After analyzing the title of each article, a relevant subset was screened on whether the articles were published in the year 2000 or later. If so, the remaining set of articles had to be written in English, German or Dutch. If all these requirements were met, the abstract of each article was analyzed to finally decide if the article was to be included in our literature review. After applying all these steps for the 264 articles found on the keyword "network architecture," only 5 articles were selected for further scrutiny. Promising keywords such as "cloud computing" and "distributed data" did not relate to any interesting articles at all. The list of results for each of the 40 keywords can be found in Appendix A.

The twenty-three articles that did pass the above tests were thoroughly analyzed for interesting concepts and ideas to be included in our further research. The summaries and the relevance of these 23 articles are described in Appendix B. The most interesting articles and concepts are presented in the following section.

4.6 Literature analysis

In this section we will identify concepts in each article and evaluate the relevance of the concepts to our research. This will help us to identify the dimensions we can use in our model, while discarding irrelevant concepts.

When one talks about computer security, one automatically thinks of *what* needs to be secured and in *which* way. An asset has an implicit or explicit value, and the higher the value, the more protection for

the asset is warranted. What's new is that the environment in where the data and its protection mechanisms are located, has changed. In cloud environments it is possible that the data and the data protection mechanisms are no longer under the direct control of the data owner. The concepts described below will be used as points of departure for further research.

4.6.1 Data protection concept

Spiekermann and Cranor (Spiekermann and Cranor 2009) discuss personal privacy, which can be protected in a policy-based approach, to a more restrictive approach by architectural mechanisms:

- *Privacy-by-policy;* based on implementation of notice and choice principles of Fair Information Practices (FIP), on which European privacy legislation is based.
- *Privacy-by-architecture:* Using mechanisms to anonymize any information, resulting in little or no personal data being collected at all.
- *Hybrid approach:* The combination of the above two, where privacy-by-policy is enhanced through technical mechanisms that audit or enforce policy compliance.

These approaches are used to make architectural choices on two axes:

- *Network Centricity:* The degree of control a network operator has over client's operations
- *User Identifiability:* The degree to which data can be directly related to an individual

Figure 4-4 shows the relation between network centricity, user identifiability, and the protection mechanisms from a personal privacy perspective. When it is harder to identify a person based on a set of data, privacy friendliness increases. When a second party, such as a network operator, has less influence on the network a person is active on, privacy friendliness is also increased.

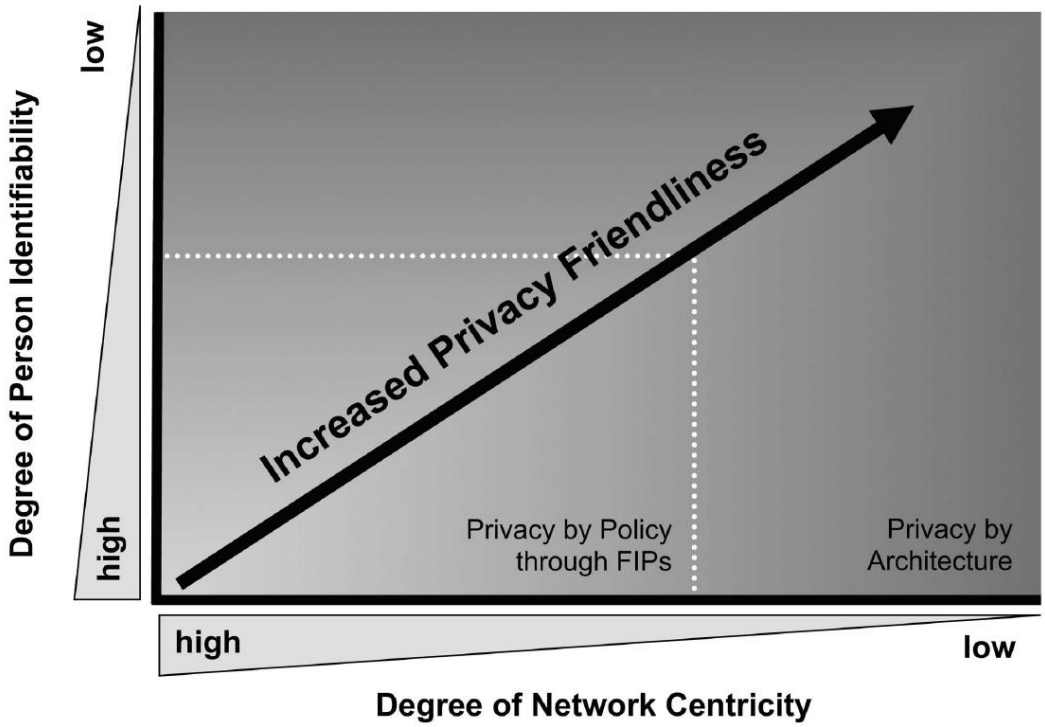


Figure 4-4 Personal Privacy protection (Spiekermann et al. 2009)

The mechanisms named for privacy-by-architecture are focused on client-centric architecture and anonymous transactions, which are mechanisms pointed in the opposite direction of the network-

centric architecture of Cloud Computing. The increasing protection from privacy-by-policy to privacy-to-architecture is a notion that we can use as severity of protection in our research.

Cody et al (Cody, Sharman, Rao et al. 2008) approach data protection in grid computing environments, in which data control is decentralized. The authors place their framework in relation to three types of grid computing systems, each with their own vulnerabilities:

- *Computational Grid*: Focused on computing power, solving complex problems
- *Data Grid*: Used to store and access large volumes of data, often distributed across multiple domains
- *Service Grid*: A grid which provides services that are not available on a single machine

The classification framework consists of four main categories, each having unique properties how to accomplish grid security and to what situations they best apply to:

- *System Solutions* deal with manipulations of software and hardware directly in order to achieve security. There are two subcategories:
    - *System Security for Grid Resources* focuses on protecting grid resources, such as hardware, applications, data and communication channels. Solutions in this category address Data grids and Service grids.
    - *Intrusion Detection Systems* (IDS) function in the computational and service grids.
- *Behavioral Solutions* use policy and management controls in order to maintain security in the grid. Behavioral Solutions are intangible and intuitive and are based on policies and/or trust:
    - *Comprehensive Policy Controls* govern a wide range of grid computing actions, instead of focusing on one area of activity. Policies function best in computational grids.
    - *Trust-based security solutions* function in computational and data grids. Trust solutions can be used to lower security overhead. If trust-levels are too low then additional security mechanisms are enacted.
- *Hybrid Solutions* is a category that combines System solutions and Behavioral solutions. Authentication and Authorization based solutions fall in this category.
- *Related Technologies* are taken from areas other than grid computing, in which the security solutions bear similarity to those required by grid computing. The described related technologies could function within data and service grids.

Figure 4-5 shows the cohesion between the protection approaches.

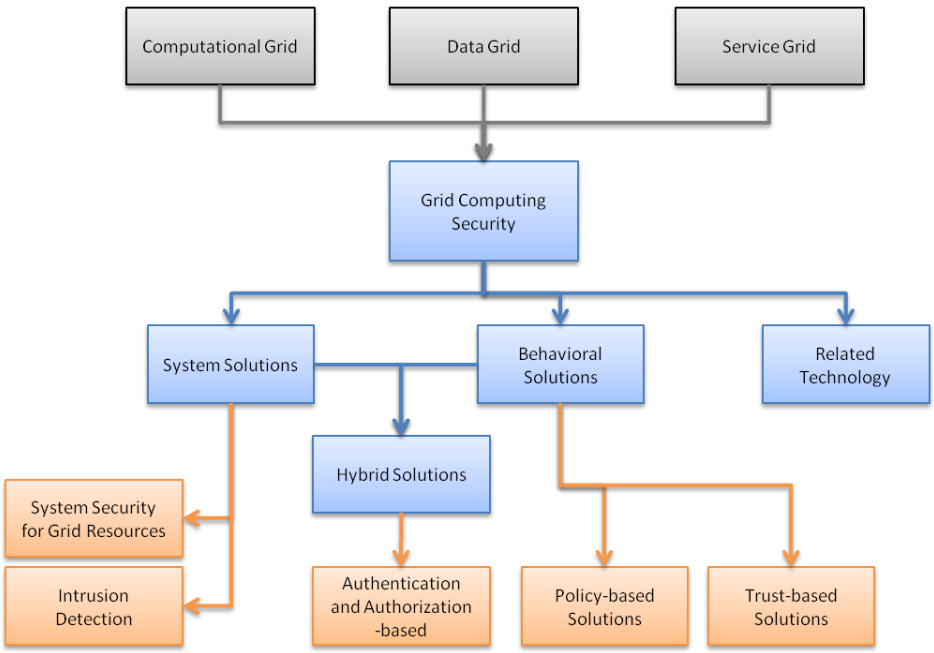


Figure 4-5: Grid computing security classifications (Cody et al. 2008)

4.6.2 Data location concept

Next to the protection concept in section 4.6.1, the authors of the article *Engineering Privacy* discuss the notion of personal privacy in relation to where the personal data is located (Spiekermann et al. 2009). They categorize the location of personal data in relation to the data owner, into three *spheres*:

- *User sphere*; location of data is fully controllable by a user, the user is responsible
- *Recipient sphere*; company-centric sphere of control, control lies with the company
- *Joint sphere*; companies hosting people's data and providing services. Users and providers have a joint control about access to data

The authors demand that system engineers should bear the responsibility of designing privacy friendly systems. They summarize the privacy spheres and resulting engineering responsibilities in a three-layer privacy responsibility framework, see Table 4-5.

This location-dependant variable of privacy can be used in relation to the topic of cloud computing**,** where one can make a clear demarcation on how much control a data owner has over his data.

Privacy Spheres	Where Data is Stored	Engineer's Responsibility	Engineering Issues
User Sphere	Users' desktop PCs, laptops, mobile phones, RFID chips	• Give users control over access to themselves (in terms of access to data and attention)	• What data is transferred from the client to a data recipient? • Is the user explicitly involved in the transfer? • Is the user aware of remote and/or local applications storing data on his system? • Is data storage transient or persistent?
Joint Sphere	Web service provider's servers and databases	• Give users *some* control over access to themselves (in terms of access to data and attention) • Minimize users' future privacy risks	• Is the user fully aware of how his data is used and can he control this?
Recipient Sphere	Any data recipients: servers and databases of network providers, service providers or other parties with whom data recipient shares data	• Minimize users' future privacy risks	• What data is being shared by the data recipient with other parties? • Can the user expect or anticipate a transfer of his data by the recipient? • Is personal data adequately secured? • Is data storage transient or persistent? • Can the processing of personal data be foreseen by the user? • Are there secondary uses of data that may be foreseen by the user? • Is there a way to minimize processing? (e.. by delegating some pre-processing to User Sphere)

Table 4-5: Three-Layer Privacy Responsibility Framework and Engineering Issues (Spiekermann et al. 2009)

4.6.3 System task concept

In *Engineering Privacy*, the authors use the notion of *System Activities*; what kind of action does a system perform on data (Spiekermann et al. 2009). They distinguish three types of system tasks, all relating to personal data, using personal privacy as the point of view;

- *Data Transfer:* Is there and *explicit* involvement of the user when his personal data is being transferred, or is the user not aware of any transfer of his data and thus *implicitly* involved in the transfer.
- *Data Storage:* If the personal data is stored outside the direct control of the user, the *persistency* of the data storage is an important privacy factor. If the data is stored in a *persistent* way, data is available for a longer period. If data is stored in a *transient* way, the data is stored for the purpose of an immediate transaction and then deleted. Transient data storage has minimal privacy implications, while persistent data storage can raise significant privacy concerns (Karat and Blom 2004).
- *Data Processing:* Procession of personal data often occurs outside the users' sphere of influence. Privacy concerns arise when personal data is processed without the user's consent, which happens often in secondary uses of the data. Some privacy laws regulate such secondary uses in the European Union (EuropeanCommission 1995a), and in the United States (Rotenberg 1998).

4.7 Literature review conclusion

Although we performed a literature search on 36 different, high ranked journals, we found little to zero literature linked to keywords closely related to our research, such as *Network Classification, Cloud Computing* and *De-perimeterization*. This proves that our research area is in its infancy and there are a lot of open issues to be answered.

During the literature review, some papers were found on the topic of data classification, but these were written before the year 2000 and as such did not pass our selection criteria. The most promising paper was the "Trusted Computer System Evaluation Criteria" of the American Department of Defense, more commonly known as the Orange Book (NCSC 1985). In the Orange Book, technical criteria and evaluation methodologies are given to evaluate the security of military systems. Although the Orange Book is a very interesting source of information, the book was only written for the American Department of Defense, and not for corporations or even other governmental agencies. The other disadvantage was that the book was published in 1985, and in 24 years the world of Information Technology has changed dramatically.

The literature review did not produce the information needed to answer the research questions regarding classification of data and what their security requirements are. However, we did find three interesting concepts on *how* data is *used*, *where* data is *used* and *how* data can be *protected* in a distributed environment. In the following chapter we explore how these concepts can be integrated in our research, by mapping them to dimensions in the cloud computing context.

To answer the research questions more precisely and devise a way how we can make recommendations on how confidential data can be used in cloud services without losing confidentiality, we performed additional research in the next chapter.

5 Towards an extended risk management framework

This chapter will explain how and why the concepts obtained in the literature review, are mapped to dimensions related to cloud computing. With the dimensions we want to show in which way this research approaches problems and solutions in cloud computing.

In chapter 4 we concluded that the concepts will not provide us with enough information to answer the research questions. Therefore, we need to supplement the information from the literature review, by performing practical research on which standards and best practices are used in present-day systems. The selection, combination, and verification of these information sources will be explained and motivated in this chapter.

5.1 Literature dimensions

In this section we will present three paradigms from the literature review in Chapter 4, in the form of dimensions to the model in the context of cloud computing. The goal of these three dimensions is to identify the uniqueness of the cloud computing paradigm. The presented dimensions are taken from the area of privacy and grid computing, which came up as results when we searched for high quality sources of information on the topic of cloud computing and confidentiality, using the keywords in Appendix A.

The three dimensions describe *how* data is used in subsection 5.1.1, *where* data is located in subsection 5.1.2, and *how data is protected* in distributed environments in subsection 5.1.3.

5.1.1 System tasks dimension

Systems perform one or more of the following tasks on data, each with its own concerns regarding privacy (Spiekermann et al. 2009).

Transfer

Disclosure of sensitive data during transfer from one party to the other is a concern that has been addressed quite extensively with the use of encryption. Encryption of data during transport is a well known concept and is sufficient, on the presumption both sender and receiver are trusted parties. In the article of Spiekermann et al, the authors are more concerned about the difference between transfers with and without explicit user involvement. Sending sensitive information with the users' involvement, such as filling a form with private information, in order to gain access to a service, has lower privacy concerns than information that is transferred without users' involvement, such as cookies and other information requested by the receiver.

When we translate these privacy concerns to the cloud computing paradigm, one can make a difference between information-*push* to the cloud and information-*pull* from local resources to the cloud, where the latter has more concern. Information-pull is initiated by the cloud service provider, and depending on the service, with or without user involvement.

Storage

Storage of data can occur inside or outside the user's or corporation's direct control. When the data is stored outside the direct control, the data owner can exercise *separation of duties*, by encrypting the data before storing it externally, while keeping the means of decryption in the owners control. This separation of duties does not work when stored data needs to be processed externally.

It may be useful to distinguish between persistent and transient storage. *Persistent* storage stores data on a long-term basis, like normal hard disks. Persistent storage brings more data retention concerns than *transient* storage, where data is deleted when the initial purpose of the data has been completed. The notion of transient storage can be implemented by preventing software to store the data on hard disks and only keep the data in memory, which is done in one of the products of cloud service provider Gigaspaces.com.

Processing

Processing refers to any use or transformation of data. In the context of personal privacy, privacy concerns are raised when data is used for purposes not foreseen by users. Under European privacy laws, users must be informed up front of all secondary uses of data and given an opportunity to provide or withhold their consent (EuropeanCommission 1995a). In the US, sector-specific legal requirements regulate secondary use of data (Rotenberg 1998).

When processing needs to take place within the cloud, data cannot be protected by the same means as data at rest and data in transit (e.g. encryption). Data needs to be in readable form in order to be processed. As such, proper data access controls need to be in place to preserve the confidentiality of data being processed externally.

There is ongoing research on the possibility of processing data in encrypted form, which is called *homomorphic* encryption (Gentry 2009). Homomorphic encryption enables data owners to have their encrypted data processed by another entity, while preventing the processing party to find out what the data is in unencrypted form. This theory is very interesting for the cloud computing paradigm, but the researcher Craig Gentry admits that it may take up to 40 years before the theory becomes practical (Gentry 2009).

5.1.2 Data location dimension

One may make a distinction on where data is located from the data owners perspective. Data location can be placed in one of three control domains: the data owner sphere, the joint sphere and the recipient sphere (Spiekermann et al. 2009).

The *data owner sphere* encompasses the company's or users' devices on which the data is located. The data is fully controllable by the data owner and data should not be able to flow in and out of these devices without the owners being able to intervene.

The joint sphere is the situation where a provider hosts the data and may provide additional services, but where the provider and the owner have a joint say as to the degree of access allowed to the data. This includes access to the data by the data host itself, for purposes other than to what the data owners agreed to. For example, Google received strong criticism for mining its users' e-mail accounts for advertisement purposes (Zetter 2004). A more recent example of data owners expecting and demanding control of who accesses their remotely stored data, are privacy issues concerning social networking site Facebook. Millions of users protested when it became publicly known that Facebook shares users' personal information with 3rd party developers without the users' consent (Schmidt 2009).

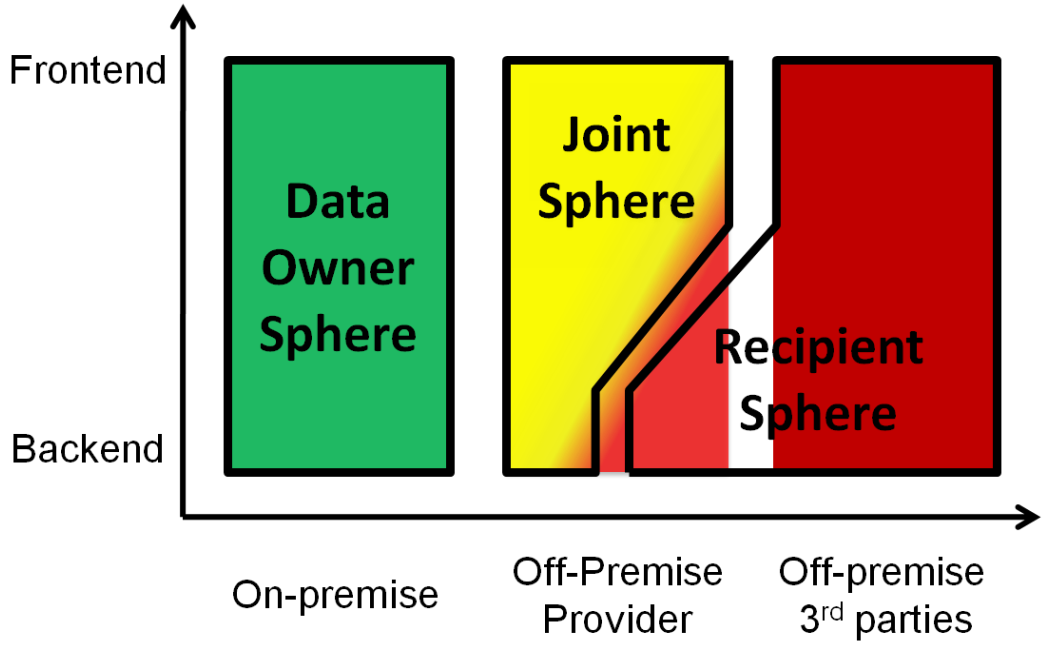


Figure 5-1: Data owner control depends on data location

The *recipient sphere* encompasses an external party-centric sphere of data control in which data owners have no direct control over their data. This sphere involves 3rd party companies and sometimes the backend infrastructure of service providers and data sharing networks. As data owners have no control over their data in this sphere, either the security measures put in place by the data custodian must be trusted by the data owner, or measures must be taken to prevent data flowing into this sphere.

Figure 5-1 shows how much control data owners can execute in the various spheres. The data owner sphere speaks for itself, but the joint and recipient sphere depend on how much control data owners have over their data located on the provider's systems. If data owners do not have enough control over their remote data, or cannot place enough trust in the correct execution of the controls placed by the service provider, data location should be considered to be in the "recipient sphere."

5.1.3 Data protection dimension

Cody et al. discusses security options in Grid Computing environments. In grid computing, resources are owned and managed by multiple entities, creating a challenge to offer a secure environment. The authors distinguished three classifications of *security solutions* that may be useful in relation to cloud computing (Cody et al. 2008). We will discuss these security solutions below and relate them to cloud computing.

System Solutions are based on the physical layer of an information system, directly manipulating the software and hardware in order to achieve security. As system based solutions are responsible for the security at the lower levels of the technology stack, these security mechanisms enable the use of other security solutions, like the behavioral and hybrid solutions discussed below. System based solutions such as cryptography act as building blocks for behavioral solutions. An example of a system solution is an Intrusion Detection System (IDS), which detects security breaches by monitoring data transfers and executions of functionality.

Behavioral Solutions act on a higher plane of abstraction than the system solutions described above. As the name says, the behavioral solutions are focused on the behavior of the users of an information

system. The behavior is controlled in the form of policies-based solutions which limit the users' access to an information system, and trust-based solutions in where other security mechanisms are only needed if the user is not trusted enough.

Hybrid Solutions are a category of solutions that combine system and behavioral solutions. Examples of hybrid solutions are authentication and authorization mechanisms.

Figure 5-2 presents these categories of security solutions.

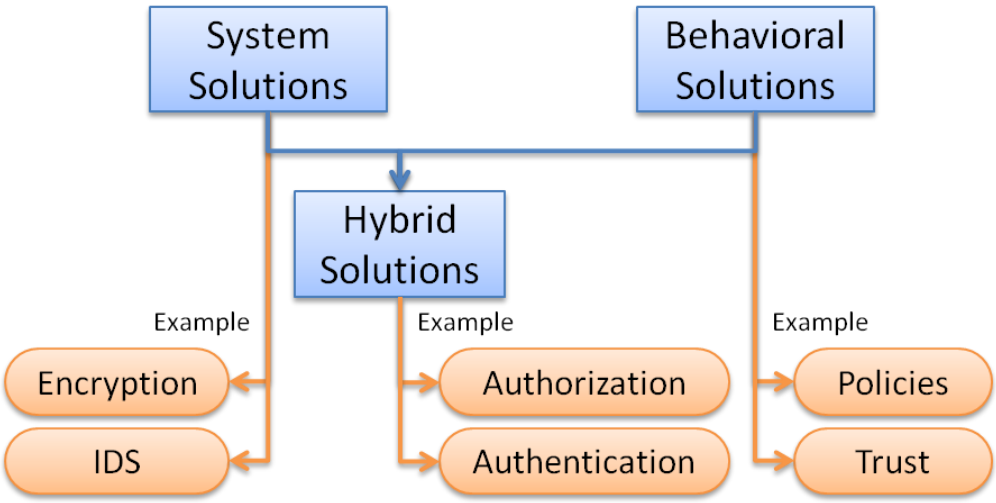


Figure 5-2: Security Solution categories in the protection dimension

5.2 Present-day information security practices

The literature review did not provide enough insights to answer our research questions, so we need to perform research on other sources of information. This further research is a repeating process of selecting and combining the information, while performing verification of the resulting idea's with security experts. First, we explain how and why we chose the information resources used next to the three dimensions from the literature, after which we explain how these resources shape our framework presented in chapter 6.

In the conclusion of chapter 4, we discussed the literature research results concerning which data classification standards are used today. We identified the Orange Book as a starting point for our research, but as this work was written in 1985, we found this too old to cite from in the fast evolving world of IT (NCSC 1985).

As such, the need was identified for more recent IT security guidelines, written for a bigger audience than just the U.S. Department of Defense. Guidelines targeted for the public or private sector as a whole was preferred. We identified two interesting organizations that develop present-day guidelines and/or standards on the topic of IT security:

- **ISO**, the *International Organization for Standardization*, the largest developer and publisher of international standards. The ISO is a non-governmental organization, with members from 162 countries and a central secretariat located in Geneva, Switzerland (ISO 2009). Related to our research area, the ISO works on their 27000 series of standards, centered around describing the Information Security Management System.

- **NIST**, the *National Institute of Standards and Technology*¸ is an U.S. federal agency with the goal to promote U.S. innovation and industrial competitiveness (NIST 2009). Related to our research area, the NIST has a Computer Security division, which develops a family of recommendations in the Special Publication 800 series.

Although ISO standards are better known than the NIST recommendations, we continued the research by using NIST recommendations as information source for present-day information security practices. There are various reasons for the choice of NIST as our prime source of present-day IT security information: The NIST publications are authoritative and implemented in all U.S. federal information systems, and as such have proved their usefulness. NIST publications are also often used in the field of IT security, and every security expert we had discussions with was acquainted with the NIST standards and publications. The most important reason of all is that to our opinion, the accessibility and readability of NIST resources are better than ISO resources.

The NIST Security Division issues documents in two variations; The *Federal Information Processing Standards* (FIPS), which are *mandatory* standards for all federal agencies in the United States, while the *Special Publications* (SP) are issued by the NIST as *implementation guides* and *recommendations* and have a less strict nature.

In our search to identify which data classifications are used in today's information systems, we found out that FIPS 199 and SP 800-60 describe the data classifications and their mandatory usage by U.S. federal agencies. The details of these data classifications are divulged in chapter 6.

We started to develop a model, incorporating the literature dimensions and the data classifications and examine the usefulness of such an approach of the model. In discussions with security experts and cloud computing experts from within Capgemini, we realized that the model lacked a clear identification of problem areas when related to cloud computing. The model based on data classifications and the literature dimensions alone was insufficient to answer more research questions and as such, needed a broader perspective.

FIPS standard	Standard full name	Related Special Publication	Special Publication Full Name
FIPS 199 (NIST 2004a)	Standards for Security Categorization of Federal Information and Information Systems	SP 800-60 (NIST 2008a)	Guide for Mapping Types of Information and Information Systems to Security Categories
FIPS 200 (NIST 2006)	Minimum Security Requirements for Federal Information and Information Systems	SP 800-53 (NIST 2009b)	Recommended Security Controls for Federal Information Systems and Organizations

Table 5-1: Relevant NIST Information security Standards and guidelines

The so called *categorization of information and information systems*, described in FIPS 199 and SP 800-60, is recommended to be followed up by the selection of security controls that are mandated in FIPS 200 and guided in SP 800-53. See Table 5-1 for the relationships between these publications.

Feedback from security experts advised us to put these publications in a broader perspective, by showing their place in risk management strategies. Further research showed that these standards and guidelines are steps in an IT risk management strategy that the NIST calls the *Risk Management Framework*.

5.2.1 Risk management

Data classification and the selection of security controls are part of an organization-wide information security program for the management of risks. This risk is related to the impact an information system has on the organizational operations and assets, as well as on individuals and other organizations. This information security program to manage risks is translated by the NIST into a *Risk Management Framework*, as depicted in Figure 5-3. We will explain each step in short, after which we explain the rationale for our focus on first two steps.

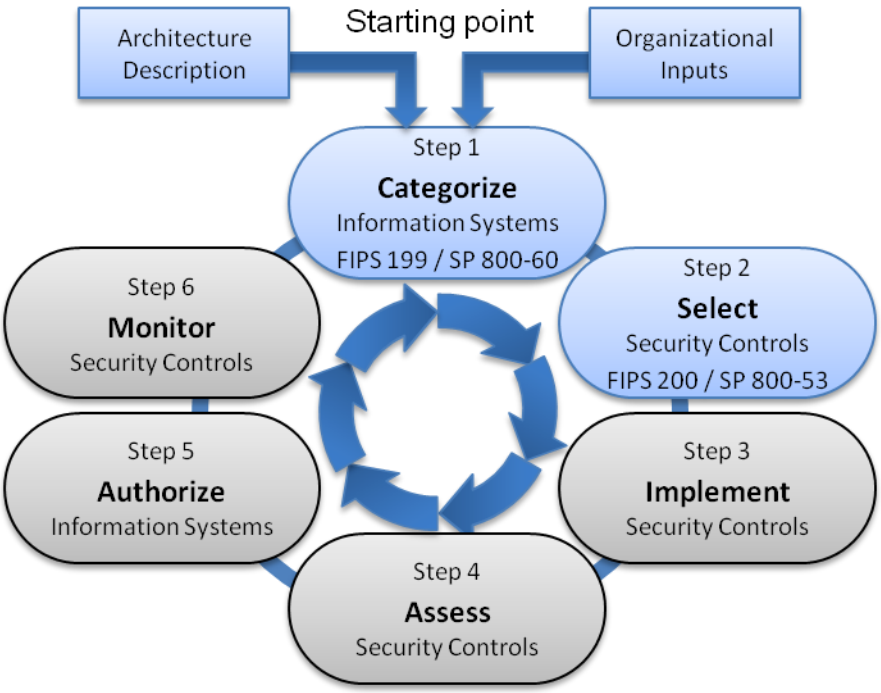


Figure 5-3: The Risk Management Framework (NIST 2009b)

The NIST framework consists of the following steps (NIST 2009b):

1. **Categorize** the information systems.
   With the use of architectural descriptions of the information systems and organizational inputs such as business goals and objectives, an organization should categorize his data and information systems. The categorization is based on the impact level of data processed, stored and transmitted by the information systems, and is mandated in FIPS 199 (NIST 2004a) and guided by the recommendations in SP 800-60 (NIST 2008a).
2. **Select** the security controls.
   With the categories of information systems, appropriate security controls must be selected to protect the information systems and data. The security control selection starts with a basic set of controls that match the impact level of the information system, after which this baseline is tailored and supplemented to meet the business specific requirements. The security control selection is mandated in FIPS 200 (NIST 2006), and guided by the recommendations in SP 800-53 (NIST 2009b).
3. **Implement** the security controls.
   The set of selected security controls is implemented in the information system, together with the creation of specifications of how and where the controls are implemented. These specifications are needed in the other phases of the risk management framework.
4. **Assess** the security controls.

The implementation of the security controls must be assessed in order to get a clear view of the extent in which the controls are implemented as specified in the security requirements. The assessment can be performed by implementers, testers, and internal- and external auditors. Guidance for assessing the security controls can be found in SP 800-53*A* (NIST 2008c).

5. **Authorize** the information systems.
   Before an information system can be used responsibly, the information system and the attached security controls must be authorized as ready for deployment. The entity responsible for authorizing an information system, decides if the risk to the organizational assets and operations is at an acceptable level. Guidance for authorizing information systems can be found in (NIST 2004b)
6. **Monitor** the security controls.
   After deployment of the information system, the system is continually monitored for security control effectiveness and changes in the computing environment or the information system itself, which may lead to needed alterations in the security plan of the system. The results of the monitoring phase are used as input in the categorization phase of the risk management framework, effectively closing the continuous loop of risk management. Guidance for monitoring an information system, can be found in (NIST 2008c).

Due to the scoping of our research, we focused on categorization and control selection steps in the above framework, and keep the other steps of the risk management framework outside our research.

5.3 Extending the risk management framework

Within the risk management framework presented above, we want to zoom in on the problem areas that arise when this risk management framework is applied in a cloud computing environment.

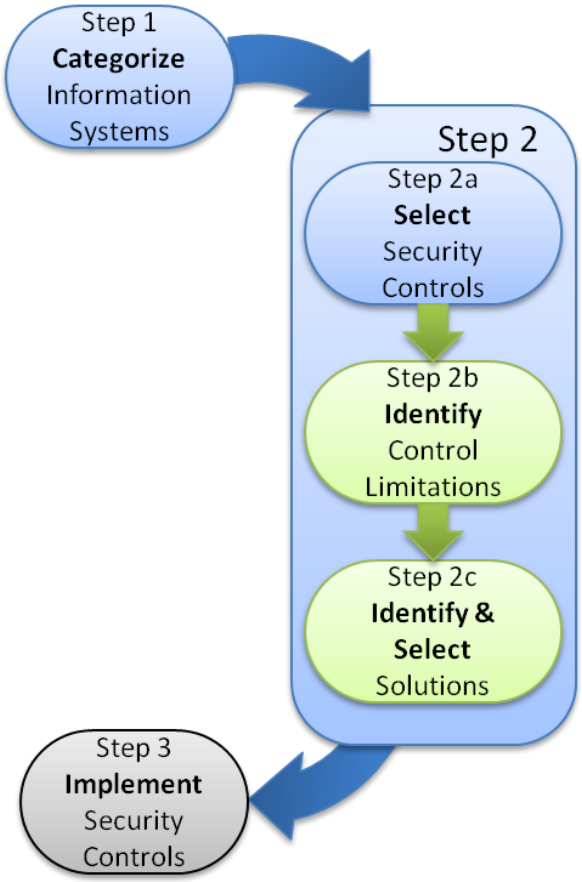


Figure 5-4: The cloud control limitation and solution extension within the Risk Management Framework

During our research it became apparent that it is not the data sensitivity itself that poses problems and limitations in cloud computing environments, but rather the mechanisms that should protect the data in these environments. During discussions with security experts it became clear that examining the controls used to protect data and information systems, would be most promising in our search to find the differences between traditional security, and security in cloud computing.

The identification of limitations that occur when these controls would be applied in a cloud computing environment would be of great value to the scientific community and to the industry. Once the

limitations that occur in the cloud have been identified, further research on the solutions for these limitations is required.

It is our opinion that this is the area where the security differences between cloud computing and traditional environments are best described. We depict this extension to the well known risk management framework, in the green blocks in Figure 5-4. In chapter 6 we will discuss the first two steps of the risk management framework with the extension in detail, where the dimensions from section 5.1 have their influence on how controls are selected, where the limitations occur and which possible solutions there exist for these limitations.

6 The Cloud Computing Confidentiality Framework

In this chapter we will present the Cloud Computing Confidentiality Framework (CCCF), which will enable companies to review the possibilities to engage in cloud based services, based on the confidentiality of the data used within the company.

The goal of the framework is to explain the differences between security in cloud computing environments, and the security in present-day information security practices. This explanation is done by describing the first steps of the IT risk management strategy, which we introduced in chapter 5, in detail and identify which differences will appear when these steps are performed in a cloud computing environment and propose possible solutions to compensate the differences. As it is a good practice for every enterprise to follow such a risk management strategy to secure their data and information systems, the framework we present here will be relevant to every entity interested to work with cloud-based information systems.

Based on the work of Shaw et al. on the topic of integrated network analysis and design, we approach our framework from a top-down perspective to ensure that security development is consistent with organizational goals and objectives and overall information system goals and objectives (Shaw and Yadav 2001). In this top-down approach, we start by explaining the need of IT security in the context of strategic goals of the business. From this abstract high level we go down to more concrete parts of the framework. Via a Business Impact Analysis (BIA) we obtain the business processes and information systems that are deemed important to the business, both in terms of criticality and confidentiality.

With the identified information systems supporting these processes and the information types involved in these information systems, we classify each information type on the topic of confidentiality. When all information types involved in a system have been classified, we can label the confidentiality of an information system by low, moderate or high confidentiality impact level. This classification is based on recommendations of the NIST on the categorization of US Federal information and information systems (NIST 2008a).

With the confidentiality labels associated with the information systems, we can ascertain the risk involved, and define which controls are needed for each confidentiality level. For the definition of this basic set of controls we use the recommendations that the NIST published in their Special Publication 800-53, while focusing on controls that are relevant to confidentiality protection (NIST 2009b).

We adjust these basic recommendations for cloud computing environments, by involving knowledge from our literature review in the form of three dimensions. These dimensions are:

- Protection mechanisms, which refers to the controls that protect information systems and data.
- Data location, which refers to the amount of control the data owner can exert over the data itself, depending on where the data is located.
- System tasks, which refers to whether the data is processed, transferred, stored, or a combination of the three.

Each dimension has its peculiarities in relation to cloud computing, which will be explained later on in this chapter. Data protection concerns the layers of protection, from higher abstract level controls to the low technical and physical controls.

The framework is presented in Figure 6-1. The gray boxes are described in short in sections 6.1 and 6.2, as although they are vital ingredients in our framework, they are considered to be outside the scope of this research because these processes have no direct relation to cloud computing and are identical in both cloud environments and traditional environments.

The blue boxes represent the present-day information security practices, in the form of recommendations concerning data classification and control selection. We will discuss the system and data classification in section 6.3, while the security control selection will be discussed in section 6.4.

The green rectangles represent important variables in our framework, in the form of the dimensions from the literature review, and trust related issues. These variables either have their effect on the control selection in section 6.4, or on identification of cloud control limitations, which will be discussed in section 6.5.

The possible solutions for the cloud control limitations will be presented in section 6.6.

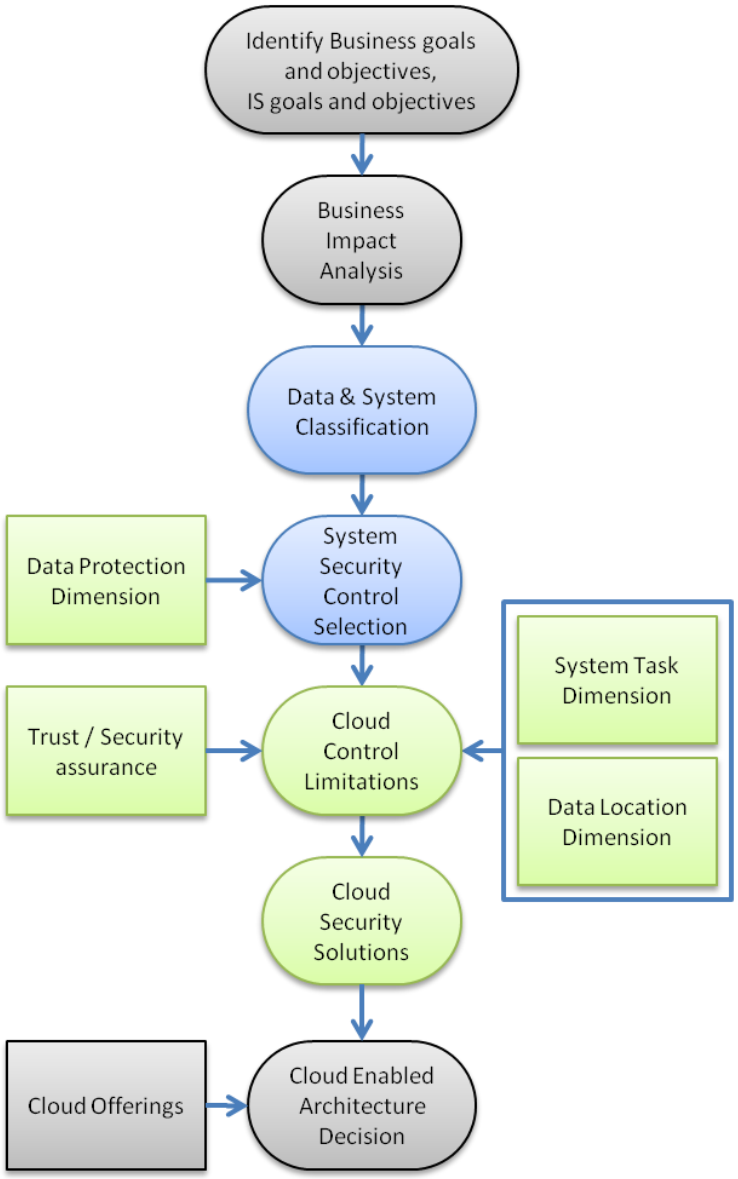


Figure 6-1: The Cloud Computing Confidentiality Framework

6.1 Identify business and information system goals and objectives

The first step in our framework is to identify goals and objectives at two levels. This step is needed to properly establish the context and scope of the confidentiality framework. If an enterprise is interested in cloud computing, it should start (and probably already has done so) with establishing business goals and objectives, which may be in the form of a mission statement. After the business goals are established, an organization should establish information system goals and objectives in a manner that will help the organization to attain the business level goals (Shaw et al. 2001).

Most of today's enterprises are driven by their information assets, which are the most critical (and often the most valuable) possessions a business has. According to Grandison et al, the reasons for this are three-fold (Grandison et al. 2007):

- Information represents the know-how of an enterprise
- Business processes operate on information
- Trusted relations are maintained by exchanging (possibly sensitive) information

Security-conscious enterprises understand that IT Security is the critical factor of their business resilience and continuity strategy. Lack of proper IT security controls place the entire enterprise at risk. In most of today's landscapes, IT security practices are not correlated to business objectives. This absence makes it difficult to determine the right level of IT security to be employed by an organization en near impossible to justify investment levels in IT security controls (Grandison et al. 2007).

6.2 Business impact analysis

The second step in our framework is a Business Impact Analysis (BIA). The reason we incorporate the BIA in our framework, is that part of the BIA process is the identification of systems and processes within an organization.

A BIA is a step commonly used in Business Continuity Planning (BCP), which is the creation of a logistical plan that specifies how an organization can and will recover in case of a disaster with disrupted business functions. A BIA will result in a prioritization between critical and non-critical organization functions and processes (NIST 2001). A BIA assigns each critical function an acceptable amount of time to restore a function once it is disrupted, and an acceptable amount of data that cannot be restored after a failure of the system.

The assigned values above are related to the availability of a system, while this research is focused on the confidentiality of data within these systems. The BIA provides the identification of information systems, which are used as input in our next step in the framework; the data and system classification.

6.3 Data & system classification

The third step in our framework is the classification of data and the information systems handling the data. The goal of the classification process is to identify *what* needs to be secured and *how valuable* the data and information systems are. When the data and systems are classified, the appropriate security controls can be selected to protect these assets, in section 6.4. Without the proper classification of the IT assets within an enterprise, two unwanted results can occur:

- An enterprise undervalues its assets, leading to inadequate security mechanisms, effectively leaving the assets at a too high risk to be compromised; or
- An enterprise overvalues its assets, leading to investing in costly security mechanisms to protect assets that do not require very extensive security.

The National Institute of Standards and Technology suggests a classification scheme based on four impact levels on the CIA properties of information and systems (NIST 2008a). We base the classification step of our framework on this NIST classification scheme. Their classification process consists of four steps to classify information in federal information systems, see Figure 6-2.

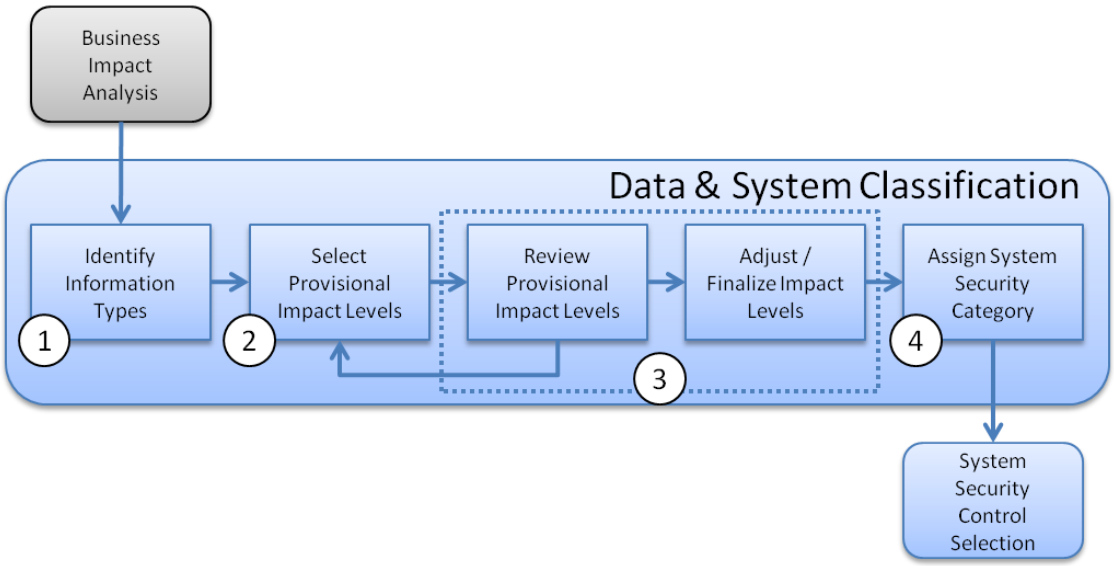


Figure 6-2: The NIST Security Categorization Process (NIST 2008a)

The NIST Special Publication 800-60 contains the basic guidelines for mapping information types and information systems to security categorizations.

As we are focusing on confidentiality in this thesis, the Integrity and Availability properties presented in this guide are outside the scope of the thesis and are not presented here.

Before we will involve the cloud computing paradigm, we explain the data and information system classification process of NIST SP 800-60 below. It is recommended to document each of the four steps of the classification process, in a format described in subsection 6.3.5.

6.3.1 Classification step 1: Identify information types

The first step in the security categorization process, is to "Identify all the information types that are representative of input, stored, processed, and/or output from each information system" (NIST 2008a).

6.3.2 Classification step 2: Select Provisional Impact Levels

The second step in the security categorization process is to select the security impact levels for the identified information types in step 1. "The provisional impact levels are the original impact levels assigned to the security objectives of an information type before any adjustments are made" (NIST 2008a). The impact levels for information types range from *Not Applicable* to *High*. See Table 6-1 for an explanation of the possible impact levels for confidentiality according to FIPS 199.

	POTENTIAL IMPACT		
SECURITY OBJECTIVE	LOW	MODERATE	HIGH
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information.	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

Table 6-1: FIPS 199 Categorization of Federal Information and Information Systems on confidentiality(NIST 2004a)

6.3.3 Classification step 3: Review provisional impact levels, adjust and finalize

The third step in the classification process, is to "(i) review the appropriateness of the provisional impact levels based on the organization, environment, mission, use, and data sharing; (ii) adjust the impact levels based on special factors" (NIST 2008a).

According to the NIST, *factors* that can influence the adjustment of the confidentiality impact levels, could be:

- Information life cycle. For example, contract information may be classified with a moderate confidentiality impact level during the time a contract is active, while it can be downgraded to a low impact level when the contract is terminated (NIST 2008a).
- Associating mission-based information. For example, common information types that are used with very sensitive mission-based information types may have higher impact levels than the same common information used with less sensitive mission-based information types (NIST 2008a).
- Configuration and security policy related information may have a higher impact level if these types of information control other very sensitive information types (NIST 2008a).
- Special handling due to legal or statutory reasons. Some information types are prone to special treatment due to the fact these information types are regulated (NIST 2008a). These regulations differ per country and will not be discussed further in this research.

6.3.4 Classification step 4: Assign system security category

The last step in the classification process, is to "(i) review identified security categorizations for the aggregate of information types; (ii) determine the system security categorization by identifying the security impact level high water mark for confidentiality; (iii) adjust the security impact level high water mark for each system security objective, as necessary, by applying the *system security adjustment factors* discussed below" (NIST 2008a).

System security adjustment factors

There are a wide range of other factors that may influence the overall system security confidentiality impact level. These factors are:

- **Aggregation:** Aggregated information can be more sensitive than every piece of information in isolation.
- **Critical system functionality:** Although a system compromise may be low impact in isolation, dependencies of other systems on the compromised system may have a exacerbating effect on overall system impact.
- **Privacy information**: When a system handles information that is protected by privacy regulations, such as Personal Identifiable Information (PII), system security categorization must be adjusted accordingly. The confidentiality impact level should generally fall in the *moderate* level.
- **Trade Secrets**: There are several laws in the United States that prohibit the unauthorized disclosure of trade secrets. Systems that store, process or communicate trade secrets should generally be assigned at least a *moderate* level of confidentiality impact level.

6.3.5 Documenting the security categorization process

Essential to the security categorization process is documenting the research, key decisions and approvals, and supporting rationale driving the information system security categorization. This information is key to supporting the security life cycle and will need to be included in the information system's security plan.

An example of the documentation of the security categorization in a power plant's information system, is given in Figure 6-3. As the source of the categorization process is (NIST 2008a), Integrity and Availability factors are also given.

<table>
<tr><th colspan="4">Information System Name: SCADA System [and Agency specific identifier]</th></tr>
<tr><td colspan="4">Business and Mission Supported: The SCADA (supervisory control and data acquisition) system provides real-time control and information supporting the main power plant. The power plant provides critical distribution of electric power to the military installation.</td></tr>
<tr><th colspan="4">Information Types</th></tr>
<tr><td>[D.7.1] Energy Supply</td><td colspan="3">Sensor data monitoring the availability of energy for the Military installation and its soldiers and command authority. This function includes control of distribution and transfer of power. The SCADA remote control capabilities can take action such as initiating necessary switching actions to alleviate an overloading power condition. The impacts to this information and the SCADA system may affect the installation's critical infrastructures.</td></tr>
<tr><td>[C.2.8.12]General Information</td><td colspan="3">The SCADA information system processes routine administrative information.</td></tr>
<tr><td rowspan="3">Step 1
Identify Information Types</td><td colspan="3">Step 2 [Provisional] / Step 3a [Adjustments]</td></tr>
<tr><td>Confidentiality Impact</td><td>Integrity Impact</td><td>Availability Impact</td></tr>
<tr><td colspan="3">Step 3b- Impact Adjustment Justification</td></tr>
<tr><td rowspan="2">Energy Supply</td><td>L / M</td><td>L / H</td><td>L / H</td></tr>
<tr><td>Disclosure of sensor information may seriously impact the missions if indications & warnings of overall capability are provided to an adversary.</td><td>Severe impacts or consequences may occur if adversarial modification of information results in incorrect power system regulation or control actions.</td><td>Due to loss of availability, severe impact to the mission capability may result and may in-turn have overall catastrophic consequences for the facility's critical infrastructures and possible loss of human life.</td></tr>
<tr><td rowspan="2">General Information</td><td>L</td><td>L</td><td>L</td></tr>
<tr><td>No adjustments</td><td>No adjustments</td><td>No adjustments</td></tr>
<tr><td rowspan="2">Step 4 System Categorization:</td><td>Moderate</td><td>High</td><td>High</td></tr>
<tr><td colspan="3">Overall Information System Impact: High</td></tr>
</table>

Figure 6-3: Example of documented Security categorization of all CIA properties (NIST 2008a)

6.4 System security control selection

In this section, we will describe the control selection process. We start by describing security controls classes and which security control families there are. Then we will describe the control selection process, presenting a recommended baseline of controls for each impact level of an information system. We will also show how this baseline can be refined to match the specific requirements of an organization. The result will be a list of required technical controls to match the security requirements of an information system given the confidentiality impact level of the system.

The control selection procedure described in this section is based on the NIST SP 800-53, which recommends security controls for Federal Organizations in the USA (NIST 2009b).

Security controls, when used correctly, can prevent, limit or deter threat-source damage to organization. Security controls can be placed into three classes:

Technical security controls

Technical controls can be used to protect against specific types of threats. These controls can range from simple to complex measures and consist of a mix of software, hardware and firmware. Next to standalone controls, technical controls also support the management and operational controls described below.

Management security controls

Management security controls are implemented to manage and reduce risks for the organization and to protect an organization's mission. Management security controls can be considered of the highest level of controls, focusing on the stipulation of policies, standards and guidelines, which are carried out by operational procedures to fulfill the organization's goals and missions.

Operational security controls

Operational security controls are used to correct operational deficiencies that might be exploited by potential attackers. These controls are implemented following good industry practices and a base set of requirements in the form of technical controls. Physical protection procedures and mechanisms are examples of operational security controls.

Within these three control classes, seventeen control families have been identified by (NIST 2009b). They are presented in Table 6-2. Due to the scope of this research, we want to focus on the technical layer of controls. Operational controls, which govern the physical protection and personnel management, will not differ much in a cloud environment with respect to the traditional computing environments. We do not include management controls in this section as they either do not differ much from the traditional environments, or they have already been described in sections 6.1, 6.2 and 6.3. Therefore we omit the operational and management classes of controls in the rest of this section.

IDENTIFIER	FAMILY	CLASS
AC	Access Control	Technical
AT	Awareness and Training	Operational
AU	Audit and Accountability	Technical
CA	Certification, Accreditation, and Security Assessments	Management
CM	Configuration Management	Operational
CP	Contingency Planning	Operational
IA	Identification and Authentication	Technical
IR	Incident Response	Operational
MA	Maintenance	Operational
MP	Media Protection	Operational
PE	Physical and Environmental Protection	Operational
PL	Planning	Management

PS	Personnel Security	Operational
RA	Risk Assessment	Management
SA	System and Services Acquisition	Management
SC	System and Communications Protection	Technical
SI	System and Information Integrity	Operational

Table 6-2: The Security Control Families

The technical control families we are focusing on are Access Control, Audit & Accountability, Identification & Authentication, and System and Communication Protection. These four families have a strong relation to the three categories of protection solutions from the literature review (section 5.1.3). In the literature review we identified the data protection dimension as an ingredient of our framework. The four technical control families have an 1-on-1 relation with a data protection solution, with the exception of Access Control and Audit & Accountability, which both are similar to the Behavioral solutions, see Table 6-3.

FAMILY	Data Protection Dimension
Access Control	Behavioral Solutions
Audit and Accountability	Behavioral Solutions
Identification and Authentication	Hybrid solutions
System and Communications Protection	System based solution

Table 6-3: Mapping of technical control families to data protection solutions

When organizations start the selection process, there are three steps to be executed sequentially:

1. Selecting the initial security control baseline (section 6.4.1)
2. Tailoring the security control baseline (section 6.4.2)
3. Supplementing the tailored security controls (section 6.4.3)

This process is depicted in Figure 6-4: The security control selection process and the following sections discuss these steps in greater detail.

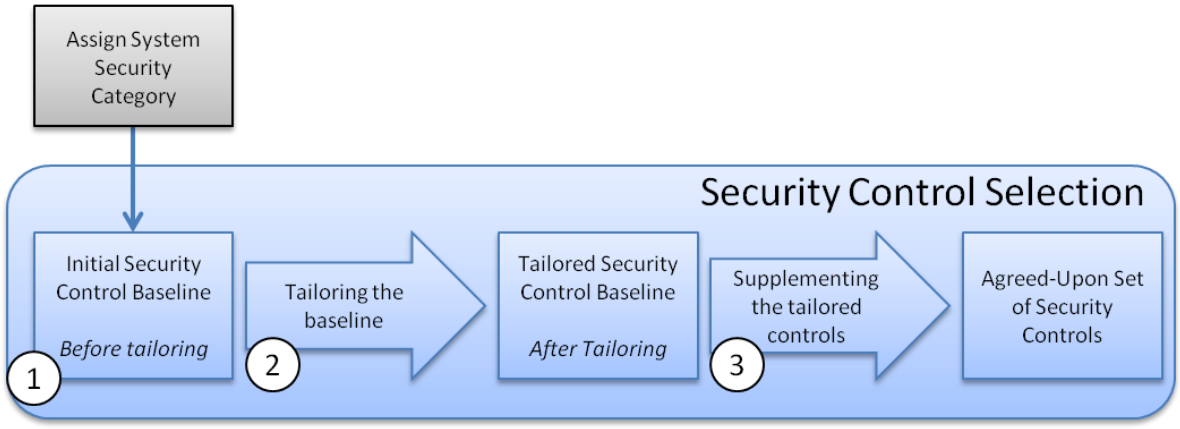


Figure 6-4: The security control selection process (NIST 2009b)

6.4.1 Selecting the initial security control baseline

The selection process begins with a baseline of controls, which are later on tailored and supplemented when the need arises. NIST provides a baseline of technical controls per impact level of an

information system. The controls presented in the table below are a subset of the complete baseline of technical controls. The complete list of controls can be found in Appendix C.

The table below contains the baseline controls that are recommended for each impact level. Controls that aren't deemed to be part of the baseline for an impact level are designated *"Not Selected."* Control enhancements, which are used to supplement security controls, are indicated by the numbers between parentheses. As can be seen below, control enhancements are often recommended when a system is classified with a higher impact level. Note that controls and control enhancements not mentioned in the table below (such as AC-9 or control enhancement AC-17 (6)), are not considered to be part of the baseline for any impact level, but are available for use by organizations if needed. This need can arise in the tailoring and supplementing steps of the control selection process, which are described below.

A full explanation of each technical control, including control enhancement descriptions and supplementing and implementing guidelines per control, can be found in the technical control catalog in Appendix Appendix D. Note that the first control of each control family is based on Policies and Procedures for that family, combining all the policies and procedures for that family into one control.

IMPACT LEVEL	LOW	MEDIUM	HIGH
ACCESS CONTROL			
Access Control Policy and Procedures	AC-1	AC-1	AC-1
Account Management	AC-2	AC-2 (1) (2) (3) (4)	AC-2 (1) (2) (3) (4)
Access Enforcement	AC-3	AC-3	AC-3
Information Flow Enforcement	Not Selected	AC-4	AC-4
Separation of Duties	Not Selected	AC-5	AC-5
Least Privilege	Not Selected	AC-6 (1) (2)	AC-6 (1) (2)
Unsuccessful Login Attempts	AC-7	AC-7	AC-7
System Use Notification	AC-8	AC-8	AC-8
Concurrent Session Control	Not Selected	Not Selected	AC-10
Session Lock	Not Selected	AC-11	AC-11
Permitted Actions without Identification or Authentication	AC-14	AC-14 (1)	AC-14 (1)
Remote Access	AC-17	AC-17 (1) (2) (3) (4) (5) (7) (8)	AC-17 (1) (2) (3) (4) (5) (7) (8)
Wireless Access	AC-18	AC-18 (1)	AC-18 (1) (2) (4) (5)
Access Control for Mobile Devices	AC-19	AC-19 (1) (2) (3)	AC-19 (1) (2) (3)
Use of External Information Systems	AC-20	AC-20 (1) (2)	AC-20 (1) (2)
Publicly Accessible Content	AC-22	AC-22	AC-22
AUDIT & ACCOUNTABILITY			
Audit and Accountability Policy and Procedures	AU-1	AU-1	AU-1
Auditable Events	AU-2	AU-2 (3) (4)	AU-2 (3) (4)
Content of Audit Records	AU-3	AU-3 (1)	AU-3 (1) (2)
Audit Storage Capacity	AU-4	AU-4	AU-4
Response to Audit Processing Failures	AU-5	AU-5	AU-5 (1) (2)
Audit Review, Analysis, and Reporting	AU-6	AU-6	AU-6 (1)
Audit Reduction and Report Generation	Not Selected	AU-7 (1)	AU-7 (1)
Time Stamps	AU-8	AU-8 (1)	AU-8 (1)
Protection of Audit Information	AU-9	AU-9	AU-9
Non-repudiation	Not Selected	Not Selected	AU-10
Audit Record Retention	AU-11	AU-11	AU-11
Audit Generation	AU-12	AU-12	AU-12 (1)

IDENTIFICATION & AUTHENTICATION			
Identification and Authentication Policy and Procedures	IA-1	IA-1	IA-1
Identification and Authentication (Organizational Users)	IA-2 (1)	IA-2 (1) (2) (3) (8)	IA-2 (1) (2) (3) (4) (8) (9)
Device Identification and Authentication	Not Selected	IA-3	IA-3
Identifier Management	IA-4	IA-4	IA-4
Authenticator Management	IA-5 (1)	IA-5 (1) (2) (3)	IA-5 (1) (2) (3)
Authenticator Feedback	IA-6	IA-6	IA-6
Cryptographic Module Authentication	IA-7	IA-7	IA-7
Identification and Authentication (Non-Organizational Users)	IA-8	IA-8	IA-8
SYSTEM & COMMUNICATION PROTECTION			
System and Communications Protection Policy and Procedures	SC-1	SC-1	SC-1
Application Partitioning	Not Selected	SC-2	SC-2
Security Function Isolation	Not Selected	Not Selected	SC-3
Information in Shared Resources	Not Selected	SC-4	SC-4
Denial of Service Protection	SC-5	SC-5	SC-5
Boundary Protection	SC-7	SC-7 (1) (2) (3) (4) (5) (7)	SC-7 (1) (2) (3) (4) (5) (6) (7) (8)
Transmission Integrity	Not Selected	SC-8 (1)	SC-8 (1)
Transmission Confidentiality	Not Selected	SC-9 (1)	SC-9 (1)
Network Disconnect	Not Selected	SC-10	SC-10
Cryptographic Key Establishment and Management	SC-12	SC-12	SC-12 (1)
Use of Cryptography	SC-13	SC-13	SC-13
Public Access Protections	SC-14	SC-14	SC-14
Collaborative Computing Devices	SC-15	SC-15	SC-15
Public Key Infrastructure Certificates	Not Selected	SC-17	SC-17
Mobile Code	Not Selected	SC-18	SC-18
Voice Over Internet Protocol	Not Selected	SC-19	SC-19
Secure Name /Address Resolution Service (Authoritative Source)	SC-20 (1)	SC-20 (1)	SC-20 (1)
Secure Name /Address Resolution Service (Recursive or Caching Resolver)	Not Selected	Not Selected	SC-21
Architecture and Provisioning for Name/Address Resolution Service	Not Selected	SC-22	SC-22
Session Authenticity	Not Selected	SC-23	SC-23
Fail in Known State	Not Selected	Not Selected	SC-24
Protection of Information at Rest	Not Selected	SC-28	SC-28
Information System Partitioning	Not Selected	SC-32	SC-32

Table 6-4: The recommended technical control baseline per information system impact level (NIST 2009b). Controls not selected in the baseline for any impact level, are omitted in this table.

6.4.2 Tailoring the security control baseline

After selecting the initial security control set from Table 6-4 and/or Appendix C, the organization continues the selection process by tailoring this baseline to their specific business conditions. Tailoring a baseline consists of three steps.

Scoping guidance

Scoping guidance is applied to the initial baseline controls, which will help organizations to determine to implement *only* those controls that are essential for delivering the protection for an information system. There are several scoping considerations to be made that may affect how the baseline controls are applied and implemented by organizations. The most relevant considerations are discussed below, while the full list of considerations can be found in (NIST 2009b).

- *Policy & regulatory related considerations*
  Security controls related to information and information systems that are governed by laws, directives, policies and regulations (such as AC-22 Publicly Accessible Content), are required to be implemented only if the implementation of the control is consistent with the information and information systems covered by the laws, directives, policies and regulations.
- *Security objective related considerations*
  Some security controls only support one or two of the confidentiality, integrity or availability security objectives (CIA). These controls may be downgraded to the control in a lower baseline level (or eliminated from selection if it is not defined for the lower baseline level), if and only if the downgrading of the control: (1) is consistent with the classification of that security objective before high watermarking was applied to the overall information system (subsection 6.3.4); (2) is supported by the overall risk assessment; (3) does not have a negative influence on the level of protection provided to the information system. NIST SP 800-53 recommends only two technical controls related to confidentiality, as candidates for downgrading, being SC-4 Information in Shared Resources, and SC-9 Transmission Confidentiality (NIST 2009b).
- *Public access related considerations*
  When public access is allowed to an information system, security controls related to personal identification and authentication are only applicable in a limited manner. For example, while these controls offer identification and authentication of personnel that maintain a publicly available website, those controls are not needed for access to public available information.

Compensating controls

It is possible that a baseline control cannot be implemented or the costs of implementing the control outweigh the benefits of the protection the control provides. For example, separation of duties prevents that an employee can plan a payment and at the same time authorizes the payment. If an organization is so small it does not have enough personnel to separate these duties, the organization may strengthen the audit, accounting and personnel security controls within the same information system. The strengthening of these controls act as compensating controls for the separation of duties control.

A compensating control is a management, operational or technical control that acts as a replacement for the baseline control, providing a comparable or equivalent level of protection for the information system.

Organization-defined parameters

After applying the scoping guidance and compensating controls, an organization selects or assigns the variables that are part of the control descriptions in the security catalog of NIST SP 800-53. Many control descriptions in this security catalog offer flexibility in the form of organization-defined parameters, where organizations can tailor a control to support specific business, mission or operational needs. For example, the second control enhancement of the Account Management control

states “The information system automatically terminates temporary and emergency accounts after [Assignment: organization-defined time period for each type of account].” This assignment statement offers an organization the option to tailor a control to support specific business, mission or operational needs.

6.4.3 Supplementing the tailored security controls

The tailored security control baseline acts as the starting point for determining whether or not this selection of controls provides enough security for the information system. This is done by comparing the organizations assessment of risk and what is required to sufficiently mitigate the risks to the organization. In many cases, additional controls and control enhancements must be selected to supplement the tailored security control baseline. Two approaches can be taken to identify which additional controls and control enhancements must be included in the final agreed-upon set of controls; the r*equirements definition* approach and the *gap analysis* approach, which will be explained next.

Following the *requirements definition* approach, the organization investigates possible threats and acquires credible and specific information about what adversaries may be capable of, as well as what damage human errors may inflict. With this assessment of possible threats, additional security can be obtained by adding controls and control enhancements from Appendix D.

In contrast to the above requirement definition approach, the *gap analysis* approach begins with an assessment of the current security capabilities, followed by a determination of what threats can be expected. This approach identifies the *gap* between the current security capabilities and selects additional controls and control enhancements from Appendix D.

The result of the whole control selection process will be the list of required technical security controls to match the requirements of an information system given the confidentiality impact level of the system. Due to the scoping of our research, we only describe the technical controls and omit the management and operational security controls. We continue our research by analyzing the technical security controls, with respect to cloud computing. In the next section, this analysis will identify limitations of required security controls when the controls are applied in cloud computing environments.

6.5 Cloud control limitations

In the previous subsections we discussed existing practices in the risk management framework. In this section, we show that the application of these security controls in cloud environments can have limitations.

The controls selection process we discussed in section 6.4 is a list of selected controls that protect an information system. Five properties influence the applicability of a security control, depending on the deployment of the information system and inherently, the deployment of the control itself. The five properties depend on the following questions:

- Who owns the information system?
- Who manages the information system?
- Where is the information system located?
- Who has access to the information system?
- How is the information system accessed?

One might notice that these 5 questions are also used to describe the cloud computing deployment models in chapter 2. The first three questions should be answered from the perspective of the data owner, where the information system in question processes, transfers or stores the data. The last two questions should be answered from the perspective of the information system user.

The *ownership* of the information system and the underlying infrastructure of the information system, lies with either the owner of the data inside the system, or lies with a 3rd party.

The *management* of the information system and the underlying infrastructure is either done by the data owner, by a party managing the information on behalf of the data owner, or by another party who has no official relation with the data owner.

The *location* of the information system and the underlying infrastructure, is either within the organizational boundaries of the entity owning the data, or is located external to the data owner's location.

Who accesses the information system and the data within the information system can be divided into three groups; 1) Public users, who do not have to be identified or authenticated as they access only public information, 2) Non-organizational users, who do not belong to the organization, but access information not deemed as public and as such, require identification and authentication. For example, a website owner needs to log in before he can change the content of his website, 3) Organizational users, who are either employees of the organization, or users deemed to have equal status as employees (e.g. contractors, guest researchers). These three groups are displayed in Table 6-5. The reason we make this difference between user groups, is that each group require different controls to access an information system. As we will present in section 6.5.1, some cloud control limitations only occur in controls that protect functionality for organizational users only. These privileged functions are not available for the other, less trusted user groups.

Access to	Accessed by
Public information and functionality	Public users
Private information and functionality	Non-Organizational Users
	Organizational Users

Table 6-5: Grouping of types of users accessing information systems

How the information system is *accessed*, is described by the type of connection a user has to the information system and data. The type of connection has a strong relation to the traditional organizational boundary of an organization. The connection to an information system can be divided in local and network based access. Network based access can be further divided in access via internal networks (e.g. LAN, WAN and VPN connections), and access via external networks (e.g. Internet, Dial-in, Wireless, Broadband). How an information system is accessed, is an important factor in the matter how much the access can be trusted to be secure. Figure 6-5 summarizes the categorization of how systems are accessed.

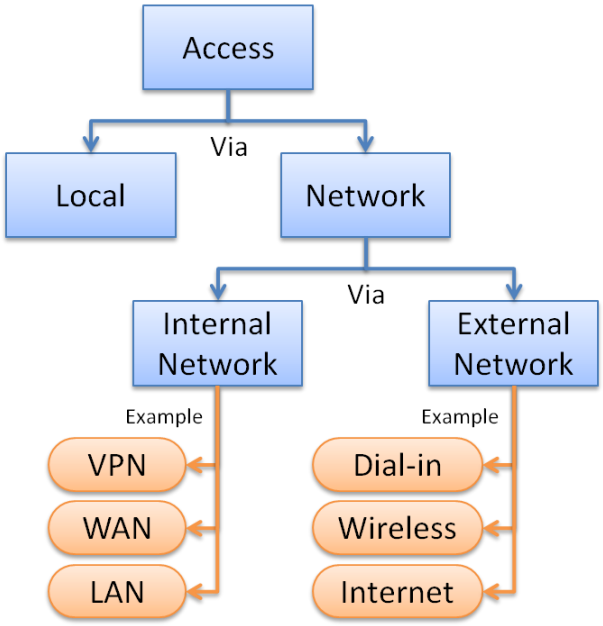


Figure 6-5: Categorization of access connections

The above 5 system properties are ingredients to determine *how much control the data owner has* over his data. In section 5.1.2, we introduced the data location dimension as an indicator how much control the data owner has over his data. To recall, the dimension described three spheres of control over the data:

- *Data owner Sphere*, where the data owner has full control over who accesses his data. In this sphere, the data owner has full control over the information system in which the data is located, and as such can influence the information system infrastructure to support recommended security controls.
- *Joint Sphere*, where a second party hosts the information system and the data, but where the provider and data owner have a joint say as to the degree of access allowed to the data.
- *Recipient Sphere*, which is an external party-centric sphere of data control in which data owners have no direct control over their data.

The description of the limitations of these controls not only depends on the impact of the information system they are protecting, but also on the environment the controls are operating in. Some controls demand a high degree of control over the information system the controls are implemented in, and that can be a limitation if there is less influence over the information system. Therefore, we do not only approach cloud control limitations via the 5 questions above, but also display the limitations of controls depending on the data control spheres.

In the following subsections, we analyze each of the 78 technical security controls published in the NIST security control recommendations (NIST 2009b). A separation is made between limitations that occur within controls that are designated as baseline controls, and controls and control enhancements that are optional and as such not part of the minimum set of controls. This separation is made to make clear distinctions between limitations that are encountered in any case, and limitations that are encountered when a choice has been made to implement additional controls and/or control enhancements.

In subsection 6.5.1 we present the limitations that occur in the recommended baseline of controls, if these controls are to be implemented in cloud computing environments.

In subsection 6.5.2 we present the limitations that exist in the controls and control enhancements that are not part of the recommended baseline of controls, if these optional controls and control enhancements are to be implemented in cloud computing environments.

As a result of the limitation identification within each control, we identify three areas of inter-control cloud limitations, which are based on limitations that occur in multiple controls. As some of the control limitations can be related to the same topic, we aggregate these limitations into three problem areas, which will be discussed in subsection 6.5.3.

6.5.1 Baseline security control limitations

Table 6-6 presents the limitations that exist in *baseline* control or control enhancements, if these are placed in a cloud computing environment. If there are multiple limitations within the same control, this is donated by the roman numerals in the third column. The full NIST recommendation and guidelines for these baseline controls with limitations, are given in Appendix Appendix D.

Baseline control	Baseline control full name	Limita-tion nr.	Limitation description
AC-17	Remote Access	-	This control states that communications with the information system via external networks, need to be encrypted for systems with Moderate and High impact levels. If encryption is not supported and assured, this requirement poses a limitation on the applicability of Moderate and High impact information systems.
AC-20	Use of External Information Systems	**I**	Control enhancement 1 states that the organization needs to verify control implementation on Moderate and High impact systems, or needs to have approved connection or processing agreements with the hosting entity
		II	This control states that the organization needs to establish terms and conditions with external hosting entities. This is a problem in the recipient sphere, in where there are no mutual agreements between the data owner and the hosting entity
IA-2	Identification and authentication for Organizational Users	-	This control states that multifactor authentication is needed, depending on the way of access (network / local), type of account (privileged / non-privileged) and the impact of the information system: • Low impact systems only need multifactor authentication for network access to privileged systems • Moderate impact systems require multifactor authentication for all access to privileged accounts, and requires multifactor authentication to privileged accounts for local access. • High impact systems require multifactor authentication for all types of access and all types of accounts. If multifactor authentication is not supported by the hosting party, this severely limits the possibilities for types of accounts, types of access, information system levels, or a combination thereof

Baseline control	Baseline control full name	Limita-tion nr.	Limitation description
SC-7	Boundary protection	-	Control Enhancement 1 states that publicly accessible information system *components* of Moderate and High impact levels should be allocated to separate subnetworks with separate physical network interfaces. An example of such a component is a public web server. Physical separation of components or systems is normally not supported by cloud providers, only logical separation of the virtual systems.
SC-32	Information System Partitioning	-	This control for Moderate and High impact systems requires the physical partitioning of information system components as part of a defense-in-depth strategy. This partitioning can be guided by the security categorization process. This control cannot be executed when physical partitioning is not supported by the hosting entity

Table 6-6: Baseline security control limitations

Some controls or enhancements are only baseline for moderate and high systems and as such, limitations that occur in these impact-dependant systems can be seen as a limitation for the usage of moderate or high systems if these controls cannot be implemented correctly. This is a limitation based on the impact level of the system.

Another dimension on which limitations can be described, is the notion of the control spheres. As we have described in section 2.3, it is shortsighted to approach cloud computing as just type of computing environment, with no control over your data once you have put it "in the cloud." With the data location spheres, we want to describe a finer-grained notion of control over data than full control or no control at all.

With these two dimensions to categorize the limitations, we can present the limitations on two axes; one being the impact of the information system, the other the environment of the information system. It is our belief that this representation makes a clear indication which limitations occur in which cloud computing setting. The baseline limitations in relation to the control spheres and impact level, are presented in Table 6-7.

Impact of system	Low	Moderate	High
Data Owner Sphere			
Joint Sphere	IA-2	AC-20(I) IA-2 SC-7 SC-32	AC-20(I) IA-2 SC-7 SC-32
Recipient Sphere	Same as Joint Sphere + **AC-20(II)**	Same as Joint Sphere + **AC-17** **AC-20(II)**	Same as Joint Sphere + **AC-17** **AC-20(II)**

Table 6-7: Baseline control limitations categorized by sphere and impact level

6.5.2 Optional security control limitations

When the 78 technical controls were scrutinized for limitations and interesting notations in relation to cloud computing, limitations in baseline controls were identified, as well as limitations that exist in controls and control enhancements that are optional with respect to the baseline controls. As described in section 6.4, these optional controls and control enhancements are selected in the tailoring or

supplementing phase of the control selection process. As it is unrealistic to assume that present-day systems have a sufficient security plan consisting of only baseline controls, it is important to also pay attention to these optional controls and control enhancements.

The *optional* control and control enhancements that have limitations when applied to cloud environments are presented in Table 6-8. The full NIST recommendation and guidelines for these optional controls and control enhancements with limitations, are given in Appendix Appendix D.

Optional control	Optional control full name	Limit-ation nr.	Optional control limitation description
AC-3	Access Enforcement	-	Optional control enhancement 6 states that information should be encrypted or stored off-line. If this enhancement is selected to be integrated in the security system, encryption is often the only option, as off-line storage is not supported in environments external to the data owner's environment.
AC-4	Information Flow Enforcement	-	Optional control enhancement 4 states that encrypted data are prevented from bypassing content-checking mechanisms. If this enhancement is selected to be implemented, one should keep in mind that this will conflict with the idea that all transferred data should be encrypted to prevent disclosure of information. To enable both encryption of transferred data and content-checking mechanisms, one would have to encrypt data after content-checking is performed, before transmitting the data.
AC-16	Security Attributes	-	This optional control states that the information system supports security attributes bound to information in storage, in process, and in transmission. The fact that this control is not stated in the control baseline is disputable, when the operational environment of cloud computing is considered.
AU-9	Protection of Audit Information	**I**	Enhancement 1 states that the information system should produce its audit records on hardware-enforced, write-once media. This typically cannot be implemented in a cloud provider's information system, as such functionality is rarely supported, if supported at all.
		II	Enhancement 4 of this control, protects audit information from compromise by privileged users, by requiring that privileged access further defined between audit-related privileges and other privileges, effectively limiting the number of users with audit-related privileges. Further reduction of this risk can be achieved by performing audit activity on a separate information system than the one being audited. If access control within cloud-based information systems is not detailed enough to separate access levels between cloud-based information systems, a good practice is to perform auditing on or store auditing information in an internal information system.
SC-4	Information in Shared Resources	-	Enhancement 1 states that sharing resources with systems of different security levels is disabled. Cloud computing is based on shared resources via virtualization.
SC-12	Cryptographic Key	-	Enhancement 3 concerns the use of NSA-approved key management technologies. US laws may forbid the export

Optional control	Optional control full name	Limit-ation nr.	Optional control limitation description
	Establishment and Management		of NSA-proprietary key management techniques
SC-13	Use of Cryptography	-	Enhancement 2 states the use of NSA-approved cryptography, which may be forbidden by US laws to be exported outside the US.

Table 6-8: Optional control limitations

6.5.3 Three general security limitations

In this subsection we will present the limitations, which spanned multiple baseline and optional controls, and as such deserve more attention. Some control limitations share the same problem area. It is important to abstract from limitations on the technical security control level, and present the common problems on a higher level. This presentation will discuss these problems on a more general level. This generalized level is interesting for readers who are not that interested in problems on the control level, but who want to understand what the more general security issues are on cloud computing, when cloud computing is looked upon from a confidentiality point of view. This generalization of limitations is depicted in Figure 6-6.

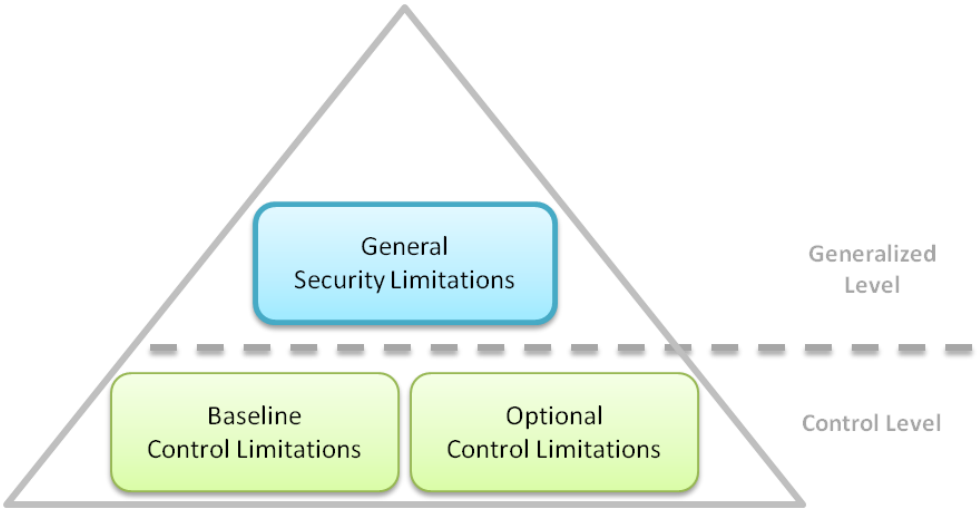


Figure 6-6: Control limitation generalization

We have identified the three problem areas that have their roots in multiple technical controls, and as such deserve further attention. These three problem areas are:

- *Access related limitations*
- *Security assurance limitations*
- *System separation limitation*

Access related limitations

The first problem area we want to discuss, are the access limitations that occur when the information systems are placed in a cloud computing environment.

A very important distinction is the difference between access by external or internal networks. If the infrastructure used to access an information system, is not under the control of the information system owner, the security of the transmissions over such infrastructure cannot be guaranteed and as such, it is marked as an external network.

In order to work with *external network* access to an information system, there are three options available:

a. Prohibit access to information systems with Moderate or High impact levels and only allow access to Low impact information systems.
b. Facilitate encryption to protect the confidentiality and integrity of the information transmitted. The encryption strength of the mechanism is based on the security classification of the information.
c. Facilitate the implementation of Virtual Private Network (VPN) technology, that not only protects the confidentiality and integrity of the transmissions, but also requires organization controlled end-points of the connection. If a VPN is implemented in this way, the network access is classified as *internal network* access by the NIST (NIST 2009b).

These options have a very restrictive manner; either moderate and high impact systems are prohibited, or this requirement produces no limitation on the final answer whether or not to allow moderate and high impact systems. It is not the case that if encryption and/or VPN are supported and adequately implemented, that all classes of information systems are allowed to be accessed via external networks: it just meets one of the criteria that need to be met before moderate and high systems may be accessed in such a way.

The central issue in cloud computing security is the amount of support for encryption and whether both end-points are organization controlled or not. In relation to cloud computing, this leads us to the following access related limitation:

> **If the cloud service is accessed via external networks, and no transmission encryption is supported, then cloud computing is limited to low impact and public access systems.**

Security assurance limitations

Although confidence and trust are used interchangeable in most literature, they are not the same in a security setting. In his paper, Pieters discusses the difference between confidence in a system, and trust in a system (Pieters 2006). Pieters defines *confidence* as the kind of assurance that a person or organization *needs* to have in a system, simply because there is no alternative. For example, people need to have confidence in the democratic voting system using paper ballots, if there is no alternative.

However, if electronic voting machines become a possibility, there is a choice between paper and electronic systems and people do not have to have confidence in one of the two systems. The result is that each of the voting systems must prove its *trustworthiness*, in order to gain *trust* of the people who have a choice and *want* to use one of the two alternatives.

The same applies to information systems and cloud computing; at the time there was no alternative for running an information system on an organization's mainframe, the organization and its employees needed to be confident that the security is handled appropriately within this environment. When

external hosting and cloud computing became possible computing environments, these systems need to prove their trustworthiness, as they provide the organization a choice on where to host their information systems.

If organizations want to use information systems hosted by a cloud provider, they want "the assurance that the risk from using the external systems is at an acceptable level, which depends on the level of *trust* the organization places in the external service provider" (NIST 2009b). The level of trust can depend on two factors:

- The degree of direct control an organization has on the external provider with regard to the employment of security controls and the effectiveness of these controls.
- The security assurance that selected security controls are implemented and are effective in use.

The degree of direct control is traditionally established in the service level agreement with the service provider. This degree can range from extensive direct control, in where the SLA specifies detailed control requirements for the service provider, to very limited direct control, where only commodity services are provided with no specific control requirements.

In cloud computing environments that are outside the Data Owner Sphere, the degree of control an organization has on cloud providers is usually very limited. According to Lee Provoost, senior consultant at Capgemini during this research, the big cloud providers such as Amazon and force.com offer most of their services in a 'one-size-fits-all' approach, where there is no space for customized security requests. These big cloud providers rely on a mass product to the public with a standard Service Level Agreement, without space for negotiation over additional security controls the provider should implement.

The other factor that creates trust of an organization in the security of a system is security assurance, which is the confidence that security controls implemented in an information system are effective in their operation. Organizations interested in cloud services should place security assurance requirements on cloud service providers in order to gain trust in the cloud provider. Assurances can be obtained through information supplied by developers, implementers and operators during the control selection phase, or by security control assessors during the assessment and monitoring phase of the risk management framework. We will discuss the information supplied in the control selection phase below. The security assurances required in the assessment and monitoring phases are outside the scope of this thesis, for more information concerning the assessment of security controls (e.g. by testers, auditors, information system owners), we refer to NIST Special Publication 800-53A (NIST 2008c).

NIST SP 800-53 elaborates on *minimum assurance requirements* for security controls in low-impact, moderate-impact and high-impact (NIST 2009b). The higher the impact of a system, the more extensive the security assurance requirements are.

"For security controls in a *low-impact* information system, the emphasis is on the control being in place and that it can be assumed that no obvious errors exist. If security flaws exist, they are attended to in a timely manner" (NIST 2009b).

"For controls in a *moderate-impact* system, in addition to the above requirements, the emphasis is on the increased confidence that the control is working correctly." (NIST 2009b) This increased

confidence is achieved by developers/implementers providing a description of the functional properties of the control, with enough detail to enable analysis and testing of the control.

High-impact systems have control assurance requirements that, in addition to requirements stated above, not only demand documentation of the functional properties of a control, but also documentation of the design and implementation of the control. "The control implementer/developer should implement capabilities that can prove the control continually and consistently meets its required function or purpose, and capabilities that support improvement in the effectiveness of the control" (NIST 2009b).

The increasing assurance requirements for controls protecting higher impact systems come with a serious drawback. It is common that the above requirements are met when the controls are implemented in a traditional system, where the organization has a high degree of control. But with cloud computing, it is usually the case that these assurances should be given by external cloud providers. However, cloud providers are very reluctant in divulging information that would assure customers of a secure environment. Jian Zhen, Director of Cloud Solutions at VMware, says about the big cloud providers that "neither Amazon Web Services nor Google App Engine are providing any type of transparency through reports or logs" (Zhen 2009).

This lack of transparency by big cloud providers has two consequences. The first consequence is that the cloud providers are not trusted enough to host information systems with moderate or high impact levels, and maybe not even trusted enough to host low impact systems. The second consequence is that security minded parties will seize every opportunity to force security transparency by cloud providers. The most prominent example of this transparency demand, is the use of the so called *right to audit* clauses that are part of most contracts.

The right-to-audit clauses allow customers to demand transparency on the security plan of service providers by various information requests, while the service provider must comply to these requests when the contract is active. These audit requests range from questionnaires on how security objectives are met by the provider, to possible on-site auditing teams that visit the physical location of the datacenters of a service provider.

These audits cause a high workload for the cloud providers if they are acted upon by the customer. Chris Hoff, Director of Cloud and Virtualization Solutions at Cisco Systems, says: "Most customers have traditionally not acted on these clauses as they used them more as contingency/insurance options. With the uncertainty relating to confidentiality, integrity and availability of Cloud services, this is no more. Almost all of the Cloud providers I have spoken to are being absolutely hammered by customers acting on their 'right to audit' clauses in contracts" (Hoff 2009).

Because the big cloud providers rely on mass service with massive infrastructure with as low overhead as possible, the duty to satisfy the audit requests requires considerable amounts of time, money and resources of the cloud providers. "Cloud providers continue to lament that they really, really want a standardized way of responding to these requests," Chris Hoff says (Hoff 2009). The lack of a standardized way for audit-responses by service providers result in two trends; providers charge customers heavily for responding to audit requests, or worse, request to audit clauses are removed from contracts altogether.

Contracts without right to audit clauses might be a solution for service providers to lower their costs, but it does not contribute to the security assurance that providers should give to customers.

System separation limitations

Some controls in both the baseline controls (Boundary Protection control and the Information System Partitioning control) as well as in the optional controls (Access Enforcement control, Protection of Audit Information control) demand either physical separation of systems, or storage of information on physical off-line media.

In the traditional computing models, network topology consists of network zones and tiers, which perform both logical and physical separation of information systems. For example, it is normal for a development system and a production system to be logically separated from each other on the network level. This logical separation is generally supported by a physical separation on the host level, where each system runs on a different physical machine.

With cloud computing however, this separation is not so clear anymore. The usage of virtualization in cloud computing makes it possible to run development and production systems on the same physical system, while logical separation is performed on the host level in domains. For further information on virtualization and the security aspects involved, we refer to (Price and Partners 2008; Vaughan-Nichols 2008).

The problems in the baseline controls arise on the requirements that systems should be physically separated, while limitations in optional controls either demand the storage of sensitive information in an encrypted form, or on off-line, write-once media.

The interesting result of the introduction of virtualization in cloud computing, is that demanding physical separation is not as standard as it is in traditional systems. When the NIST recommendations on physical separation of systems and/or components are used as guideline in a cloud environment, there are two options available:

- Demand physical separation of systems by the cloud provider. Although physical separation of systems is not part of the standard offerings of cloud providers, the customer demands that the cloud provider supports and implements physical separation of the systems of the costumer. This requirement does involve the security assurance problems described earlier in this subsection, with the possibility that the provider is unwilling or unable to support the physical separation of systems.
- Denote the physical separation requirement as obsolete. The required physical separation of systems was designed as a security mechanism in a time it was not perceived that virtualization would become so popular and influential. As a result, recommendations such as physical separation can be seen as standards and regulations that are out of date and not realistic with respect to the fast-paced developments in science and technology.

6.6 Cloud security solutions

In the previous section we discussed the limitations that arise within the security controls and on a more general level, when an information system is hosted in a cloud environment. The goal of this section is to describe the solutions and choices available to either counter these limitations, or accept the limitations.

The common and simple perception of cloud computing at the beginning of this thesis is that cloud computing has two variants; either the organization uses the cloud computing paradigm in-house (private cloud), or it is hosted by other entities with a lot of uncertainty of how the security is handled

by the external entities (public cloud). Hybrid clouds can be seen as a combination of private and public cloud. When we put this common perspective in relation to the Data Location dimension of section 5.1.2, private clouds are considered to be in the data owner sphere, in where there is full control by the organization on how the data belonging to the organization is handled. Public cloud can be placed in both the joint and recipient sphere, where there may be some, or no control at all. This perception is depicted in Figure 6-7.

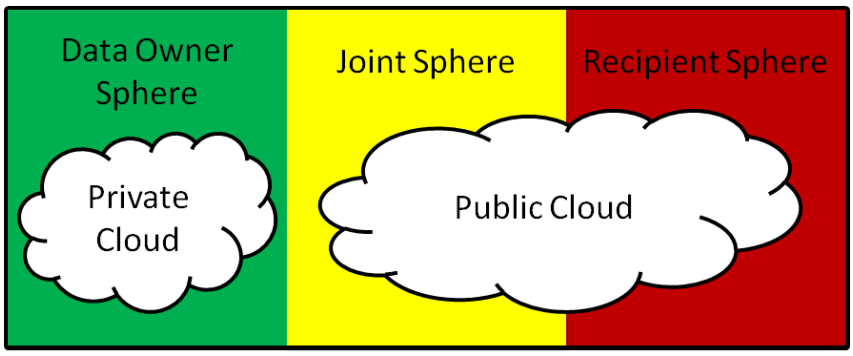


Figure 6-7: The common perception of cloud computing

This perception of private versus public cloud computing suggests that it would be unwise to let any but public information to enter public cloud environments. With the process described in the cloud computing confidentiality framework, an organization can assess which security controls there need to be in place to protect the information system in question.

When an organization considers a cloud service offering as operational environment for the information system in question, both parties can perform a gap analysis to determine which security controls are required for the information system, and which security controls the cloud service provider supports. The difference between the required controls and the supported controls is called the security gap. To reduce the organizational risk that the security gap imposes, the NIST recommends the following three options to close the gap between what security is needed and what security is offered by external service providers:

- "Use the existing contractual vehicle to require the external provider to meet the additional security control requirements established by the organization" (NIST 2009b)
- "Negotiate with the provider for additional security controls (including compensating controls) if the existing contractual vehicle does not provide for such added requirements" (NIST 2009b)
- "Employ alternative risk mitigation measures within the organizational information system when a contract either does not exist or the contract does not provide the necessary leverage for the organization to obtain needed security controls" (NIST 2009b)

If the additional controls demanded by the organization can be implemented by the cloud provider, the public cloud environment of the provider meets the security requirements set by the organization. In this case, the public cloud environment of the provider can be perceived as being in the joint sphere, see Figure 6-8. In this case, "Public cloud" may be a confusing term, because public is often associated with public access while that is strictly controlled now by both the cloud provider and the organization owning the data.

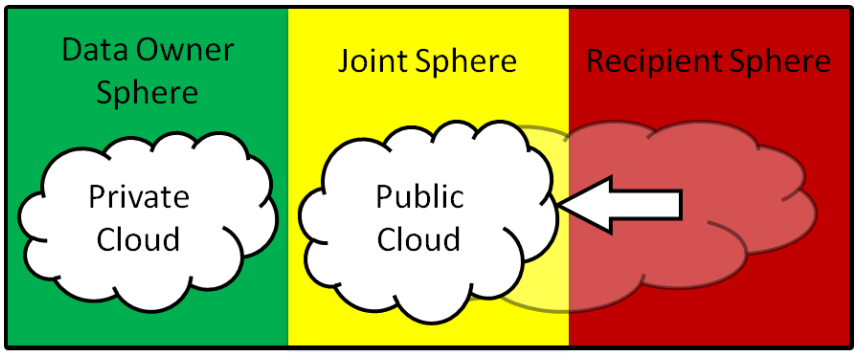


Figure 6-8: Perception of public cloud when meeting the security requirements of the data owner

However, if the security gap between what controls the organization requires and what the cloud provider supports cannot be closed by additional controls or supplementing controls, the risk involved must be mitigated by other ways than via contractual agreements. The risk mitigation can take various forms, which we will discuss next.

Option 0: Don't enter the cloud

The most obvious and easiest option, but at the same time the most short-sighted and least satisfying one, is to exclude cloud computing as a possible computing environment.

Option 1: Private cloud only

The second option is the deployment model where the organization has full control of his information system, from application to the infrastructure level. This private cloud deployment model has the best security promises of any cloud deployment, as there is no external provider involved in either owning the system, managing the system or hosting the system. The private cloud security resembles the security level of traditional computing environments.

The three general security limitations mentioned in 6.5.3 do not occur in private clouds; encryption and VPN implementations are available so there are no access related limitations, security assurance can be proved, and the required physical separation of systems can be included in the architecture of the organization's cloud.

Although private clouds is the best option available within the cloud computing paradigm with regard to security, the usability is low compared to the other cloud deployment options. The total computing capacity is limited by the physical capacity of the organization's own datacenter, and the procurement phase of new capacity is comparable traditional datacenters when the workload of the cloud demands extra physical infrastructure.

Option 2: Adopt hybrid clouds

The best option to cope with the security limitations from section 6.5, while still being able to use the full potential of the cloud computing paradigm, is the use of the hybrid cloud deployment model. In this model, both a private and one or more public clouds are used in conjunction with each other. As described in option 1, the private cloud part has the properties of the Data Owner sphere, while the public cloud part is situated in the Joint and Recipient spheres. For clarification see Figure 6-9.

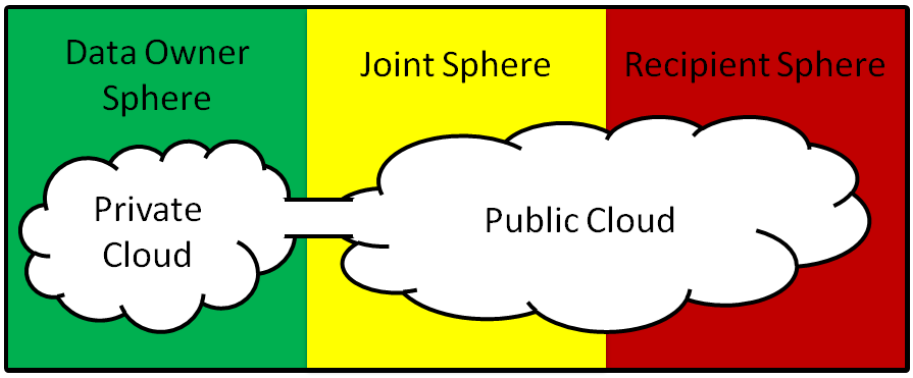


Figure 6-9: Hybrid cloud computing; the combination of clouds in multiple control spheres

This setup is very useful as a workaround for the control problems that occur in both the joint and recipient sphere. Also the three general problems that are described in 6.5.3, can be handled by this setup.

For example, when the public cloud providers do not support physical separation of systems that are classified as Moderate and High systems, these systems should be hosted in the private part of the hybrid cloud, where these limitations do not occur.

Likewise, if the encryption requirements are not properly supported by the public cloud providers, the Moderate and High systems should operate in the data owner sphere, while Low systems and data could operate in both parts of the Hybrid cloud.

It is important to magnify on the connection between the private and public cloud. This gateway between the two control spheres is of critical importance to the usability of the hybrid deployment model. This gateway is responsible for the following functions and security objectives;

- Allow information systems and data flow between the public and private cloud parts, in order to support the independent resource pooling and rapid elasticity characteristics of cloud computing.
- Prevent information systems and data to flow from the private part to the public part, if the security for those systems and data cannot be guaranteed by the public cloud provider.

This gateway has the responsibility for the trade-off between usability of the public cloud, and security of the private cloud.

7 Framework validation

The development of the Cloud Computing Confidentiality Framework (CCCF), which is presented in chapter 6, included the influence of several consultants and security experts in the field. Interviews were conducted to discuss the development and the goal of the framework. These interviews include professional opinions on the framework which criticize and validate the framework.

This chapter will describe each of the major development versions of the framework and will present in which way each of the interviews have had their influence on each of the versions of the CCCF. Section 7.1 will present the approach taken in the interviews, while sections 7.2 through 7.4 contain the interviews per development version of the CCCF.

7.1 Validation approach

There were three major versions of the framework that were discussed in interviewees. Per version of the framework, we will describe the framework in short, after which we describe the interviews that were conducted with that version of the framework in mind. These rounds of validation will contain the influences of these interviews on the succeeding version of the framework.

Six professionals of Capgemini NL were interviewed, supplemented with an interview with one external business information security officer. At least one security expert was interviewed per version of the model.

Semi structured interviews were conducted, which are interviews that start with some specific questions and the rest of the interview is a more natural conversation. We used this setup because the framework was still under construction and each version of the framework would have different discussable subjects.

There was one single question that was asked in each of the interviews:

- *What should be in the framework that is not there yet?*

The interviews were conducted on an individual basis and in a face-to-face setting.

7.2 First round of validation

In the first round of interviews, three interviews were held around the progress of developing the CCCF. The input for these three interviews is the model presented in Figure 7-1. The model will be explained first, after which the results of the interviews will be presented.

As can be seen in Figure 7-1, the model is centered around the NIST SP 800-60 guidelines on mapping information systems and information types to security categories. This process is also known as data classification. The blue boxes in the figure depict these NIST guidelines, which are preceded by the gray box depicting the relation between data classification and business goals.

The green and red boxes represent the added value of this research project in relation to what was already available in the form of the NIST processes. The step “Review Provisional Impact Levels” within the SP 800-60 elaborates on the adjustment of the provisional impact levels, in order to match the specific situation the organization itself is in. In this version of the model we want to highlight the regulations and statuary factors, as these factors regularly appeared in articles concerning cloud computing. For example, the concerns on the possible spread of data across legal boundaries, are mentioned that often, that we want to include these concerns in our model.

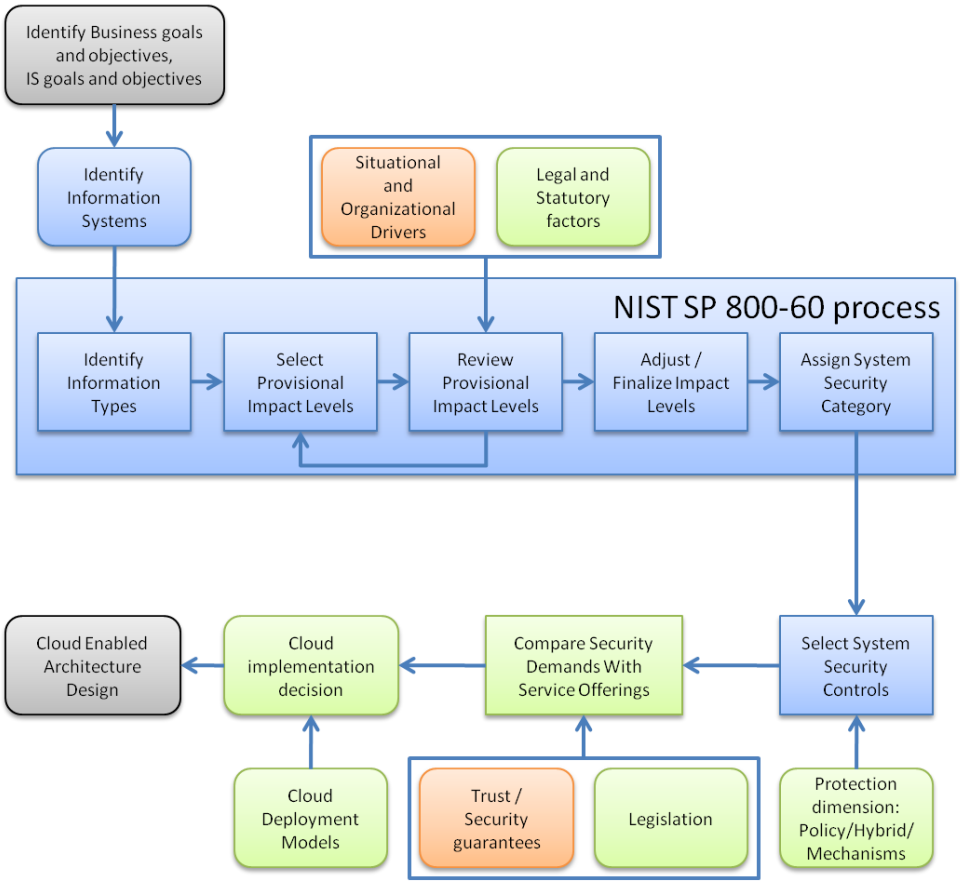


Figure 7-1: The CCCF for the first round of validation interviews

In the NIST SP 800-60 guidelines the next step "select system security controls" is mentioned. At the time that this version of the framework was the most recent one, thorough research was not yet conducted on the topic of security control selection. However, we did want to relate our protection dimension to the security control selection step, even when it was uncertain how the process of control selection was going to look like. The result of the security control selection, which would be the security requirements set by the organization, would need to be matched to the cloud service offerings available. Trust and security guarantees would be important ingredients in this decision making, as well as the legislation constraints that might limit where data could be stored.

With the result of the comparison between security demands and available cloud service offerings, a decision could be made which cloud deployment would be best to meet the organization's demand.

When one looks back at the development of the CCCF, it is easy to indicate what is incomplete or what is plain wrong with the above version of the model. However, for the validation of the final version of the framework it is important to know how each version is influenced by the opinions of professionals and security experts.

The model presented in Figure 7-1 was validated in three interviews. These three interviews were conducted with colleagues from Capgemini, who are either directly involved in information security, or are direct colleagues who were very interested in the research as a whole, and were willing to give their point of view on the framework. The three interlocutors are stated in Table 7-1.

Interview	1	2	3
Interviewee	Theo Schalke	Jan van de Ven	Martijn Linssen
Company	Capgemini NL	Capgemini NL	Capgemini NL
Interviewee's position	Security Architect (CISSP certified)	Enterprise Infrastructure Architect	Enterprise Integration Architect
Interview date	September 7, 2009	September 7, 2009	September 7, 2009
Interview location	Capgemini, Utrecht	Capgemini, Utrecht	Capgemini, Utrecht

Table 7-1: The interlocutors for the first round of validation

First, we will present the answers of the interviewees on the prepared question for the interview, after which we will state other discussion observations. The adjustments to the framework resulting from these interviews will be presented in the description of the next version of the framework in section 7.3.

What should be in the framework that is not there yet?

One of the things that Mr Schalke is missing in this framework, is the cooperation of Cloud Computing, Information Security, and Risk Management. "Risk Management is never mentioned in this model, but the result of risk management decides whether or not you want to operate in the cloud." On top of that Theo has the opinion that the gray block at the start of the framework (Identify Business Goals and Objectives, IS Goals and Objectives) has been paid too little attention to. This particular block has to clarify that IT security is not just a IT issue, but that should be considered a business issue. The security of IT should be of prime concern for the whole organization.

Mr Van De Ven is concerned that the value and impact of data is shed insufficiently light on. He argued that the value and impact of data should be more approached, because "the impact and the value of the data, is the operational management of an organization. Information is the most important asset of an organization and if the management of information is not executed properly, the continuity of the whole organization is at risk."

Mr Linssen indicated that the connection of the classification labels onto protection mechanisms does not receive enough attention. He either expects that this connection would result in interesting findings, or that the protection mechanisms themselves would be very interesting to research in the context of cloud computing.

Other points of discussion

Next to the answers on the general question, other interesting opinions and advises were expressed.

Mr Van De Ven and Mr Schalke expressed their opinion that the concepts presented in the framework are too loosely coupled. The texts accompanying the framework lacked cohesion and according to Jan "the texts are too universal."

Mr Van De Ven stresses that it should be determined what the business impacts and CIA ratings are, of both data and processes. This provides a solid basis for determining the appropriate protection mechanisms.

During the conversation with Mr Linssen it became clear that the framework should remain high level. In this sense the block 'Compare Security Demands With Service Offerings' is too specific and would involve a quantitative analysis of the cloud services available. Comparing all the cloud offerings is a too big quantitative research: "I'm afraid you cannot make the connection. Creating an automatic coupling of supply and demand is a bridge too far. It would be best if you can roughly tell, at the end of your research, which [cloud] model would be best to use."

7.3 Second round of validation

Before this second round of validation was conducted, the framework presented in section 7.2 was adjusted according to the interviews of the first round of validation. Not all advises from the first three interviews were implemented in the new version of the framework, as the time between the first and the second round of interviews was less than a week. The original interview plans included that the first 5 interviews would be conducted on the model presented in section 7.2, but due to rescheduling needs of two interviewees, it was possible to adjust the model based on some of the feedback of the first three interviews. This round of validation consists of two interviews, discussing the evolved framework presented in Figure 7-2.

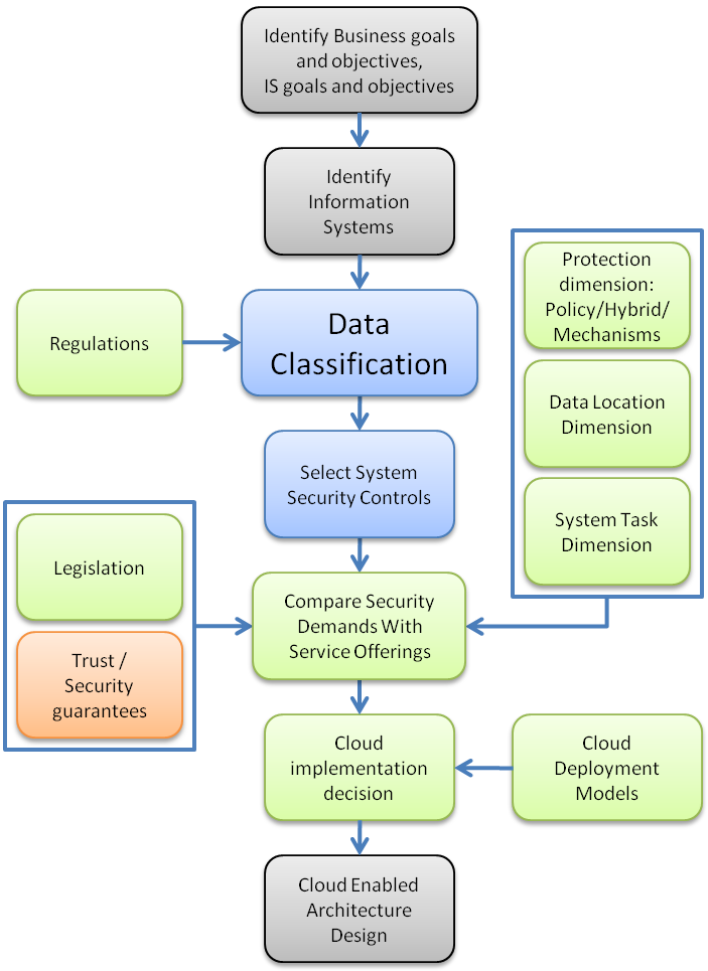


Figure 7-2: The CCCF for the second round of validation

Although the revision time for this version of the framework is less than a week, some important changes has been made to the version of Figure 7-1.

The vertical representation of a flow diagram has been chosen to present the framework more as a step-by-step plan for organizations to get to a well-founded decision on how to enter cloud computing.

The schematic importance of the NIST SP 800-60 process has been scaled down. The focus of first version of the framework seemed to be on the data classification process, as it was the perception that the analysis of this process in cloud computing environments might result in interesting conclusions. Mr. Linssen gave the advise that the selection of security controls, depending on the data

classifications, may be very interesting to research in the context of cloud computing. This is why 'Data Classification' was given almost the same magnitude as security control selection. The downscaling of the data classification process included the removal of explicitly naming the situational and organizational drivers as factor in the NIST SP 800-60 guidelines.

All the literature concepts have been added in this version of the framework, as dimensions in the comparison of security demands and service offerings. The exact way how these dimensions would influence the comparison was still uncertain, because the block 'Compare Security Demands With Service Offerings' was already a point of discussion in the interview with Mr. Linssen.

The model presented in Figure 7-2 was validated in two interviews. These two interviews were conducted with colleagues from Capgemini, who are either directly involved in information security, or are active in the fields of Web 2.0 and Cloud computing in general. The two interlocutors are stated in Table 7-2.

Interview number	4	5
Interviewee	Martin Kroonsberg	Lee Provoost
Company	Capgemini NL	Capgemini NL
Interviewee's position	IT Governance & Information Risk Management Consultant (CISSP certified)	IT Strategist - Web 2.0, social media, cloud computing
Interview date	September 11, 2009	September 14, 2009
Interview location	ING, Amsterdam	Capgemini, Utrecht

Table 7-2: The interlocutors for the second round of validation

We will first present the answers of Mr. Kroonsberg and Mr. Provoost on the question what they think should be added to the framework, after which we will present other discussion observations. The adjustments to the framework, resulting from these interviews, will be presented in the description of the next version of the framework in section 7.4.

What should be in the framework that is not there yet?

Mr. Kroonsberg is the security expert in this round of evaluation, and as such he misses the notion of Business Impact Analysis (BIA) in the framework. With a BIA systems and data are identified and the impacts if those systems and data are unavailable, is determined. The identified data can be used in the data classification process to derive the CIA labels. Together with a risk analysis, one can devise a list of required security controls. Mr. Kroonsberg advises to include the BIA as a block in de framework.

Mr. Provoost advised that adding layers of users may give an interesting perception in the framework, where each layer of users could be using different systems. The layers of users create a distinction between basic employees, middle management and executive management. The interviewee gave an example of a company where the e-mail environments of the executive management are hosted on on-premise exchange servers, while the e-mail of basic employees was hosted on external cloud-based servers. The reasons for this distinction were, next to the cost perspective, backup efficiency and compliancy.

Other points of discussion

In the interview Mr. Kroonsberg pointed out that the data classification and system security control selection, together with the gray blocks preceding them, are part of the risk management strategy.

Therefore, he advised to include the NIST SP 800-30 "Risk Management" guidelines into the research.

Although Mr. Kroonsberg has no knowledge or experience with cloud computing, he pointed out that it is important to match two groups of knowledge. One group would be what services are possible with cloud computing and which security measures has been taken for those services, while the other group of knowledge would be the outcome of the data classification and security control selection. These two need to be matched, which is also called a gap analysis. Mr. Kroonsberg: "You need to ask 'which elements do I need of [cloud] providers to match the [security] demands to the [security] guarantees?'"

Because the cloud computing knowledge of Mr. Kroonsberg is limited, Mr. Provoost was asked which security guarantees there are within the cloud. He answered that there is no standardized way that cloud providers use to publish information on security guarantees. If these kinds of information would need to be obtained, "you would have to examine the fine prints of each cloud provider." This would mean an extensive quantitative research of each of the service level agreements of each provider. Lee warned that even if this quantitative research is conducted, it would be a lot of work, and the results could be out-dated at the time they are published.

Mr. Provoost also noticed that the suggestion of Mr. Kroonsberg to do a gap analysis to identify the extra measures needed, is a good suggestion but with severe limitations: "It is important to know which security demands you have, regardless of the cloud provider. However, you cannot approach a cloud provider with these security demands, because this does not work. The big cloud providers have an idea of how their service should look like, and they do not implement specific demands." According to Lee, all major cloud providers offer a standard, basic model, with the idea of 'One-size-fits-all.' "You may represent a very big company with specific needs, but if you do not fit the model used by the provider, you are plain unlucky," Mr. Provoost says. If an organization wants to have specific features, those features have to be implemented by a thirds party, or by the customer himself.

7.4 Final round of validation

Before the last round of validation was conducted, the framework in section 7.3 was adjusted according to the interviews in the second round of validation. This round of validation consists of two interviews, discussing the evolved framework presented in Figure 7-3.

Several important changes have been made to the 2nd version of the CCCF. We will describe the changes in a top-down approach, following the steps of the framework.

We acceptedthe advice of Mr Kroonsberg, by replacing the 'Idenfity Information Systems' block with the Business Impact Analysis block. Business Impact Analysis is a basic part of a business wide risk analysis. Risk analysis is likely to have already been conducted within an organization.

For this version of the model, it was decided that the discussion on regulations and legislation would be combined in the 'Legislation' block, to prevent confusion when legal issues are mentioned in multiple places.

Since the development of the second version of the CCCF, the 'System Security Control Selection' block became an important part of the framework. Using the advice of Mr. Kroonsberg on the topic of Risk Management, it became clear that the NIST SP 800-53 guidelines would be a very solid basis on which a list of required security controls could be established. Combined with the Data Protection Dimension, it became possible to relate required security to the classification of data and systems.

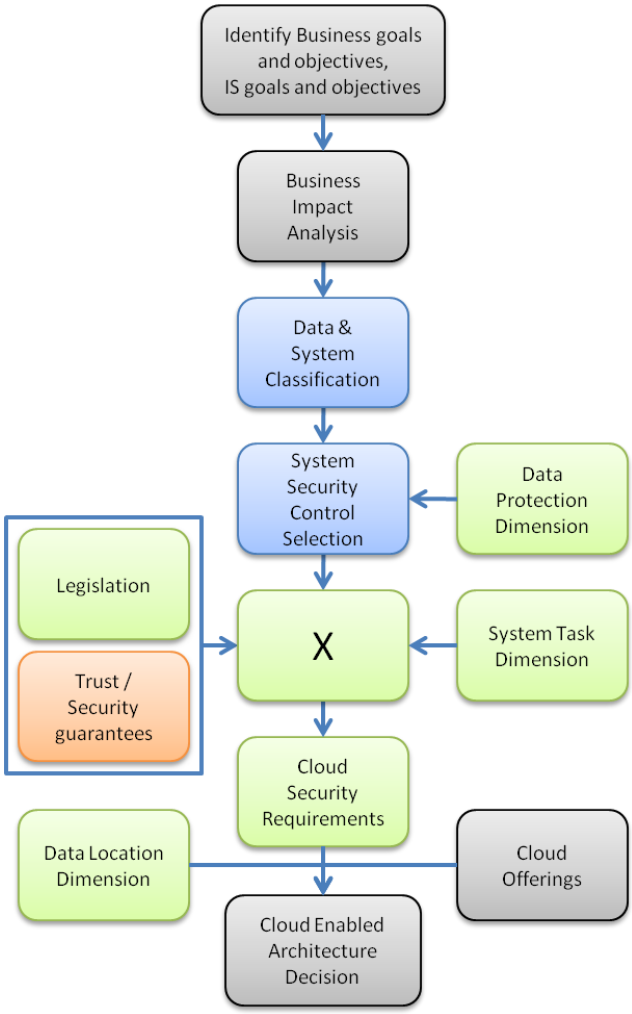


Figure 7-3: The CCCF for the final round of validation

During the interviews with both Mr. Linssen and Mr. Provoost it became clear that a quantitative research on cloud service offerings would be too extensive, and would result in quickly out-dated information. Therefore we removed the 'Compare Security Demands With Service Offerings' block. Although it was known that this part of the model would be the hub of interesting research results, it was uncertain how to call it at this point. Research was already being conducted on the problems that might occur when security controls are related to the system tasks of Processing, Transfer, and Storage in the cloud. The underlying perception was that applying security controls in a cloud computing environment could result in limitations.

Some problems were identified, but there was no fruitful cohesion between the problems on which we might make strong conclusions or recommendations. Therefore, the interviewees in the final round of validation would be asked for their opinion on the used approach.

Whereas the literature dimensions were combined as a single input in the model of Figure 7-2, in this model they are more related to their relevant areas. The data protection dimension has a direct relation with security controls, while the system tasks dimension is related to the research mentioned in the previous paragraph. The data location dimension was envisioned to be incorporated in the decision which cloud architecture would be best to fit the cloud security requirements.

The model presented in Figure 7-3 was validated in two interviews. One interview was conducted with a well-respected colleague from Capgemini , while the other interviewee is a security expert who was active for 5 years as external IT auditor and is now active as an information security officer . The two interlocutors are stated in Table 7-3.

Interview number	6	7
Interviewee	Herman Hartman	Hans-Peter van Riemsdijk
Company	Capgemini NL	RAET
Interviewee's position	Certified Enterprise Architect – Principal Consultant	Business Information Security Officer (CISSP certified)
Interview date	October 22, 2009	October 25, 2009
Interview location	Capgemini, Utrecht	RAET, Amersfoort

Table 7-3: The interlocutors for the final round of validation

We will first present the answers of Mr. Hartman and Mr. Van Riemsdijk on the question what they think should be added to the framework, after which we will present other discussion observations.

What should be in the framework that is not there yet?

Mr. Hartman acknowledges that the identification of applicability problems, which may arise when the security controls are applied in a cloud environment, is a good approach. However, Mr. Hartman says it is a good idea to have an aggregated level of problems, next to the very technical problems the CCCF is identifying now. According to Mr. Hartman, "it is doubtful that all organizations assess cloud providers using a list of technical control problems. An interesting paradigm that is often used, is the difference between information producing organizations, and organizations producing physical goods". The former are totally reliant on information, while the latter are less dependent on information. The result is that information producing organizations could assess cloud providers on a security control basis, while organizations that produce physical goods are hardly interested in the technical details of security and prefer a higher level overview of security problems.

Mr. Van Riemsdijk indicates that the identification of security control limitations is a good approach and that this approach can result in interesting conclusions. From his experience, he knows that problems with security controls are almost always related to the following five properties; Who owns the data, who manages the data, where is the data located, who accesses the data and how is the data accessed. The first three properties can be directly related to the properties of the cloud deployment models. Mr. Van Riemsdijk advises that these five properties should be considered in the identification and description of control limitations.

Other points of discussion

When Mr. Hartman was asked for his opinion on the unfruitful approach with the System Task dimension in the identification of security problems, he advised that it would be best to continue the work on a higher aggregated level.

In the interview with Mr. Van Riemsdijk, the possibility was discussed to alter the approach taken in the identification of the security limitations. Instead of relating security limitations to the system task dimension, it might be more fruitful to relate the security limitations to the data location dimension. This data location dimension could be in turn related to the cloud deployment models. Combined with the suggested five properties in the previous section, this approach was deemed to be very promising.

Mr. Van Riemsdijk mentioned that the 'Data Classification' and the 'Security Control Selection' blocks are two steps in the bigger picture of Risk Management. As an information security officer, he

advised to place the CCCF in perspective with the well-known Risk Management paradigm, which would result in that the CCCF has a very solid, well-known basis. The research project could then be explained as a research to identify the anomalies that occur when the normal risk management approach is applied to cloud computing.

8 Conclusions and further work

The usage of cloud computing as a computing environment for information systems and data can place data outside the data owner's control. The amount of protection needed to secure data is directly proportional to the value of the data. When the value of data increases, the number and extensiveness of needed security controls also increase. It could be a problem if these security controls are not supported by the cloud provider. The uncertainty of how security can be guaranteed in external computing environments raises several security questions concerning the availability, integrity, and confidentiality of data in these cloud computing environments. This thesis focused on the confidentiality issues in cloud computing environments.

The goal of this research project was to create a framework that clarifies the impact of cloud computing on the confidentiality of data placed in such environments. The framework should make recommendations on

- How data can be classified on confidentiality
- How data classifications relate to the security controls needed to preserve the confidentiality of data
- How the process of security control selection is negatively influenced in cloud computing environments
- How to cope with the negative influences of cloud computing on the protection of data confidentiality.

We managed to make the above recommendations using the following research questions:

- Which data classifications are used today and what are their security requirements with respect to confidentiality?
- Which cloud architectures are available and what security controls do they have in place with respect to confidentiality?
- How can we classify cloud architectures on the area of confidentiality?
- How can we create a mapping from confidential data classes to cloud architectures operating on potentially confidential data?

We will present the answers to the research questions in section 8.1, after we will discuss the recommendations in section 8.2. The contributions of this research are presented in section 8.3. Options for further research are presented in section 8.4.

8.1 Conclusions

The investigation on confidentiality preservation and data classifications, started with a literature review (see chapter 4). The literature review has been conducted in order to search all relevant scientific literature of top quality. The relevant academic and peer-reviewed information on the above topics is very limited at the time of writing.

During the literature review, three concepts were distilled that were related to the cloud computing paradigm in the form of dimensions (see chapter 5). These dimensions relate to *how* data *is used*, *where* data *is located* in relation to the data owner, and *how* data *is protected*. Each of these dimensions is used to answer the research questions below.

Research question 1

Which data classifications are used today and what are their security requirements with respect to confidentiality?

In the conclusion of chapter 4, we discussed the literature research results concerning which data classification standards are used today. We identified the Orange Book as a starting point for our research, but as this work was written in 1985, we found this too old to be useful in the fast evolving world of IT security. In section 5.2, we identified the ISO security standards and the NIST security standards and recommendations as present-day, reliable, acknowledged and very popular guidelines for data classification and protections.

We chose the NIST standards and recommendations as main source of classification and security requirements information. The data classification used in this research project consists of three security objectives and three or four impact levels. The three security objectives are Confidentiality, Integrity, and Availability (CIA), while the impact levels of the data to organizations are either High, Moderate, or Low. The confidentiality security objective of data can have the fourth impact level Not Applicable, which relates to no impact to the organization and as such, there is no need for protection of this class of data.

The aggregate classifications of all the data types used in an information system dictate how the information system is classified on confidentiality, integrity and availability. Although we tried to focus on the confidentiality aspects of data and system security, further separation of the security objectives became a problem. The NIST sources we used to establish the relation between classifications and protection mechanisms are based on the total impact level of an information system. The total impact level of a system is the high water mark of the confidentiality, integrity, and availability security objectives together (see section 6.3). The impact level of the whole information system dictates the baseline set of security controls that protect the system (see section 6.4). The biggest drawback of this approach is that information systems with a low confidentiality classification may be protected by the most severe security mechanisms, simply because the availability may be classified as high and as such, the whole system is classified as high impact.

As we focused on confidentiality issues within cloud computing environments, we decided to take the aggregate confidentiality impact levels of all information type within an information system as the overall impact level. This would result in the best identification of the security requirements on the topic of confidentiality, which are needed in the rest of our research.

Research question 2

Which cloud architectures are available and what security controls do they have in place with respect to confidentiality?

We decided to focus on the categorization of cloud architectures by looking at the deployment and service models used to describe cloud computing. Researching which architectures are available is hard to do because cloud computing offerings are emerging at a rapid pace. It is not only hard to keep track of all the cloud computing offerings, the cloud computing paradigm itself is also rapidly evolving. Finding out which services are offered and deriving which architectures are used for each of these services, is an impossible goal to achieve in a market this big and developing this fast. As a result, creating a mapping of which security controls are implemented by each cloud provider is a goal aimed too high as well.

However, it is possible to categorize cloud architectures. Using the definition of cloud computing presented in chapter 2, cloud architectures can be categorized by either the service model or the deployment model.

The service models defined in this thesis are Software-as-a-Service, Platform-of-a-Service, and Infrastructure-as-a-Service, depending on which level of the technology stack the cloud service is offered (see section 2.2). Although the service models describe to which extent the consumer has control over some but not all functionalities, the service models are not used in this research. We decided to focus on the cloud deployment models described below, because they offer better insights on where confidentiality problems can occur in cloud computing environments.

Cloud environments can also be categorized with the deployment model, which describes first of all which party owns the infrastructure, secondly which party manages the infrastructure, and finally at whose location the infrastructure is located (see section 2.3).

The security controls implemented in each deployment model differ per cloud provider. However, there are limitations on which security controls can be implemented by external cloud providers. These reasons for these limitations and the limitations themselves are discussed in the following research question and in section 8.2.

Research question 3

How can we classify cloud architectures on the area of confidentiality?

As discussed in the previous section, cloud architectures can be categorized by the cloud deployment model in use. The most common cloud deployment models - private and public clouds - are often described with respect to which side of the organization's protective boundary they are. Private clouds are inside the organization's boundary (on-premise), while public clouds are seen as outside the organization's boundary (off-premise).

We used the cloud deployment models to classify cloud architectures on confidentiality, by determining the degree to which the cloud deployment model can comply with the security requirements set by the data owner. The security requirements depend on the data classification, and are expressed by the number and extensiveness of security controls needed to protect the information system operating on the data.

We classified the cloud deployment models by placing them in Data Location spheres, which describe the amount of control the data owner has over his data (see section 5.1.2). The distinction between full control by the data owner (data owner sphere), shared control by both data owner and hosting party (joint sphere), and no distinguishable control by the data owner (recipient sphere), was made based on which controls the hosting party is unwilling or unable to support and implement. The unwillingness or disability to support the required security controls, results in a lack of trustworthiness of the cloud provider, and directly relates to which cloud deployment model is used. (see section 6.6 and section 8.2).

Research question 4

How can we create a mapping from confidential data classes to cloud architectures operating on potentially confidential data?

We conducted a thorough literature review in chapter 4, by which we were unable to find mappings from data classifications to cloud computing architectures. As a result, we conclude that such

mappings do either not exist or have not been peer-reviewed and published yet. We used part of the NIST risk management framework to construct the Cloud Computing Confidentiality Framework (CCCF, see chapter 6), which is a step-by-step framework that creates the mapping from data to the most suitable cloud architecture as computing environment. This framework determines which security is required by the data, which security cannot be guaranteed in which computing environments and which solutions are available for these shortcomings, via the following steps:

1) Identify the information systems used within the organization
2) Identify the data types used in each information system
3) Classify the data types and use the data classifications to classify the information system
4) Select and tailor the security controls, based on the classification of the information system
5) Identify the problems that occur when these security controls are required in cloud computing environments
6) Identify the cloud environment that supports the required security controls and/or copes with the limitations identified in step 5).

8.2 Results

The process of finding answers to the research questions produced the following results.

The relevant academic and peer-reviewed information on the topics of cloud computing, distributed security mechanisms, and data classification is very limited at the time of writing (see chapter 4).

The security controls needed to protect data and information systems, have limitations when applied in external cloud environments. The limitations per technical security control are summarized in sections 6.5.1 and 6.5.2. Some of these security control limitations are the foundation of the following three major problem areas:

- Information systems classified as moderate or high impact level, require extensive security controls that may not be part of the standard set of controls supported by the cloud provider. This lack of supported controls limits the possible information systems that can be run on external computing environments such as public clouds.
- Cloud providers have problems gaining the trust of potential customers, because providers are very reluctant in offering customized security plans, and do not offer detailed information on how the security plans are exactly implemented. On the other hand, customers demand provider transparency on security before they denote a service offering as trustworthy and usable.
- Security controls exist that require a physical separation of information systems and component, while the focus of cloud computing is on virtual usage of infrastructure, systems and data. It is unclear to which degree public cloud providers support the physical separation requirements. However, this requirement should start the discussion on whether some security requirements, set by standards, may be considered so old-fashioned that they should be rendered obsolete and removed from standards and guidelines.

If the public cloud provider can comply with the security requirements of the data owner, the public cloud is considered to be in joint control of both the cloud provider and the data owner. This is the most ideal situation, in which none of the above security limitations exist.

However, when the implemented security controls of a public cloud provider cannot meet the security requirements of the data owner, the data and information systems of the data owner cannot run completely in public cloud environments.

To visualize this restricting situation, the cloud deployment models of public, private and hybrid clouds are related to the amount of influence a data owner has on the security of his data (see section 6.6). This mapping leads to the following options that handle the security limitations that may occur:

- By only employing a *private* cloud environment, the data owner is in full control of how his systems are configured and protected. A private cloud is by definition designed to serve one organization, offering the possibility to customize the infrastructure in such a way that the above limitations can be coped with. Considering a security limitation perspective, private clouds closely resemble traditional computing environments. However, it must be kept in mind that private clouds, just like traditional in-house solutions, cannot use the full functional potential that public clouds can. The flexibility and apparent limitless computing capabilities are limited in private clouds. Computing capabilities in private cloud environments are limited by the size of the dedicated datacenters, which are of comparable size to those used in traditional in-house datacenters.
- In the case that some, but not all security requirements can be met by the public cloud provider, the *hybrid* cloud model is a very promising cloud deployment model. In this model, part of the deployment is considered as a private cloud, while the other part is considered public cloud. The main distinction is made on the amount of control the data owner has, that is the data owner has full control on the security mechanisms needed to security his highly confidential data in the private part of the hybrid cloud, while the data owner has less control on the security mechanisms in the public part of the hybrid cloud.
  With this approach, it is possible to work around the security control limitations that can occur when moderate and high impact information systems are applied in cloud environments. If public cloud providers cannot prove the trustworthiness of moderate and high impact systems, a data owner can cope with this problem by hosting the higher impact systems in the private part of the hybrid cloud, while he can 'outsource' his less sensitive information systems and data to the public part of the cloud.
  Cloudbursting is a "dynamic deployment of a software application that runs on internal organizational computing resources to a public cloud to address a spike in demand" (Andrzejak, Kondo and Anderson 2010). In order to preserve confidentiality, not all information systems and data can be burst to the public part of the hybrid cloud. We recommend starting bursting less sensitive information systems and data to the public part of a hybrid cloud, if the private infrastructure does not provide enough computing capacity to handle the workload. We call this *selective cloudbursting.*

8.3 Contributions

This section will describe the contributions of this research project to science and practice.

A literature review has been conducted on the approached of security in cloud computing environments. The results of this literature research show that cloud computing is in its infancy, because the structured search revealed minimal peer-reviewed information on these topics. Regardless, this research has distilled three concepts from the literature and transformed them into dimensions that are usable in the context of cloud computing. The most contributing dimension is based on concept from the research area of personal privacy. The concept of the location of personal information in relation to the location of the person itself is used to create the Data Location

Dimension. This dimension is used to categorize how much control a data owner can exert over his own data, which depends on the location of the data. This is important because the location of data in cloud computing could be very dynamic, and to our opinion there was no proper way to describe the problems that relate to the amount of control a data owner has over his own data.

Another contribution of this research is related to the popular and well-respected Risk Management Framework (RMF). This research extends the RMF by determining anomalies when two steps of this framework are applied in a cloud computing environment. The steps Data Classification (NIST FIPS 199 and SP 800-60) and Security Control Selection (NIST FIPS 200 and SP 800-53) are used in this research, in which we extend the Security Control Selection with our own work. The extensions that made to this control selection step are the steps "Identify Control Limitations" and "Identify and Select Solutions."

With the identification of control limitations, this research shows which limitations can occur when technical security controls are applied in a cloud computing environment. This type of research has not been conducted before and it is important for entities that wish to know what the security limitations of cloud computing are on this technical level. Some limitations are also aggregated onto a higher level, which is important for entities who want to know the security limitations of cloud computing on a more abstract level.

This research also contributes by making recommendations on the cloud deployment models that should be used, which act as solutions to the security limitations. The recommendation of *selective cloudbursting* serves as an important tradeoff between the safety of a private cloud hosted in a fully controlled environment, and the flexibility and apparent limitless computing power of public clouds.

8.4 Further research

Only technical security controls were analyzed in this thesis. In future research on the topic of confidentiality preservation in cloud computing, the Cloud Computing Confidentiality Framework presented in this thesis can be extended by adding the analysis of operational and management security controls. Such an investigation could lead to supplemental controls for limitations that might occur in cloud computing environments.

As discussed in the previous section, *hybrid* cloud computing is a very promising cloud deployment model that can cope with the security limitations occurring in a public cloud environment, while still being able to support many of the economical advantages of public cloud computing. Hybrid clouds depend heavily on the gateway between the private part of the hybrid cloud and the public part of the hybrid cloud. The gateway between the private and public parts of an hybrid cloud is a interesting point for further research. To make this deployment model successful, the following research areas presented as further research:

- The gateway must prevent information systems and data to flow from the private part to the public part, if the security for those systems and data cannot be guaranteed by the public cloud part provider. It is possible that automated information flow enforcement, in combination with security attributes on data and information systems, is a vital security control combination for a secured gateway. For more information on information flow enforcement and security attributes, we refer to security controls AC-4 and AC-16 in (NIST 2009b), respectively.
- Internet bandwidth may become the major bottleneck for the hybrid cloud deployment model, as identified in (Armbrust, Fox, Griffith et al. 2009). Computing power is almost doubling

every 18 months, also known as Moore's Law. Storage capacity is increasing at the same speed, sometimes referred to as the Kryder's Law (Walter 2005). Internet bandwidth does not grow exponentially like computing power and storage space, which is known as Nielsen's Law (Nielsen 1998). Further research is needed on the impact of this bottleneck and the optimization of data transfers via the hybrid cloud gateway.

As a last recommendation for further research, we suggest that the security assurance limitations identified in 6.5.3 might be solved by developing a structured way for cloud providers to answer auditing questions. Such a structure can greatly help potential cloud users to verify the state of security of cloud providers.

9 References

AIS. (2009a). Association for Information Systems - MIS Journal Rankings. Retrieved July 9, 2009, from http://ais.affiniscape.com/displaycommon.cfm?an=1&subarticlenbr=432.

AIS. (2009b). Journal of the Association for Information Systems. Retrieved July 27, 2009, from http://aisel.aisnet.org/jais/.

AIS. (2009c). Communications of the Association for Information Systems. Retrieved July 27, 2009, from http://aisel.aisnet.org/cais/.

Amazon. (2009b). Amazon Virtual Private Cloud (Amazon VPC). Retrieved December 28, 2009, from http://aws.amazon.com/vpc/.

Andrzejak, A., Kondo, D. and Anderson, D. (2010). Exploiting Non-Dedicated Resources for Cloud Computing. In Proceedings of 12th IEEE/IFIP Network Operations & Management Symposium (NOMS 2010), Osaka Japan.

Antón, A., Bertino, E., Li, N. and Yu, T. (2007). A roadmap for comprehensive online privacy policy management. *Communications of the ACM*, *50*(7): 116.

Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R. et al. (2009). Above the clouds: A Berkeley view of cloud computing. *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28*.

Baralis, E. and Chiusano, S. (2004). Essential classification rule sets. *ACM Transactions on Database Systems*, *29*(4): 635-674.

Bardin, J., Callas, J., Chaput, S., Fusco, P., Gilbert, F. et al. (2009). Security Guidance for Critical Areas of Focus in Cloud Computing v2.1, Retrieved January 28, 2010, from Cloud Security Alliance, from http://www.cloudsecurityalliance.org/guidance/

Breaux, T. and Ant, A. (2008). Analyzing Regulatory Rules for Privacy and Security Requirements. *IEEE Trans. Softw. Eng.*, *34*(1): 5-20.

Brodkin, J. (2008). Gartner: Seven cloud-computing security risks, Retrieved September 23, 2009, from Network World, from http://www.networkworld.com/news/2008/070208-cloud.html

Chen, K. and Liu, L. (2005). Privacy preserving data classification with rotation perturbation. In Proceedings of Fifth International Conference of Data Mining, 589–592. IEEE.

Chockler, G. V., Keidar, I. and Vitenberg, R. (2001). Group communication specifications: A comprehensive study. *ACM Computing Surveys*, *33*(4): 427-469.

Cody, E., Sharman, R., Rao, R. H. and Upadhyaya, S. (2008). Security in grid computing: A review and synthesis. *Decision Support Systems*, *44*(4): 749-764.

Dawson, S., De Vimercati, S. C. D., Lincoln, P. and Samarati, P. (2002). Maximizing sharing of protected information. *Journal of Computer and System Sciences*, *64*(3): 496-541.

EuropeanCommission (1995a). 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal of the EC. 23*.

Garfinkel, R., Gopal, R. and Goes, P. (2002). Privacy Protection of Binary Confidential Data Against Deterministic, Stochastic, and Insider Threat. *Management Science*, INFORMS: Institute for Operations Research. *48***:** 749-764.

Gartner (2008). Assessing the Security Risks of Cloud Computing, Retrieved December 5, 2009, http://www.gartner.com/DisplayDocument?id=685308

Gentry, C. (2009). Fully Homomorphic Encryption Using Ideal Lattices, Retrieved October 12, 2009, from Stanford University and IBM Watson, from http://boxen.math.washington.edu/home/wstein/www/home/watkins/CG.pdf

Goo, J., Kishore, R., Rao, H. R. and Nam, K. (2009). The role of service level agreements in relational management of information technology outsourcing: An empirical study. *MIS Quarterly: Management Information Systems*, *33*(1): 119-146.

Grandison, T., Bilger, M., O'Connor, L., Graf, M., Swimmer, M. et al. (2007). Elevating the Discussion on Security Management: The Data Centric Paradigm. In Proceedings of *2nd IEEE/IFIP International Workshop*, 84-93.

Hoff, C. (2009). Incomplete Thought: The Crushing Costs of Complying With Cloud Customer "Right To Audit" Clauses. *Rational Survivability*, Retrieved September 20, 2009, from http://www.rationalsurvivability.com/blog/?p=877.

ISO. (2009). International Organization for Standardization, main website. from http://www.iso.org/iso/home.htm.

Karat, C. and Blom, J. (2004). *Designing personalized user experiences in e-Commerce*, Kluwer Academic Publishers.

Kesh, S. and Ratnasingam, P. (2007). A knowledge architecture for IT security. *Communications of the ACM*, *50*(7): 103-108.

Lysne, O., Reinemo, S., Skeie, T., Solheim, A., Sodring, T. et al. (2008). Interconnection Networks: Architectural Challenges for Utility Computing Data Centers. *Computer*, *41*(9): 62-69.

Morsi, W., El-fouly, T. and Badr, A. (2006). Using IPSec to Secure Multi-Level Data Classification in MLS Networks. In Proceedings of 6th International Conference on ITS Telecommunications, 817-821.

NCSC (1985). Trusted Computer System Evaluation Criteria. *Orange Book*. D. o. Defense, from http://csrc.nist.gov/publications/history/dod85.pdf

Nicholson, S. and Smith, C. (2007). Using lessons from health care to protect the privacy of library users: Guidelines for the de-identification of library data based on HIPAA. *Journal of the American Society for Information Science and Technology*, *58*(8): 1198–206.

Nielsen, J. (1998). Nielsen's Law of Internet Bandwidth. Retrieved February 11, 2010, from http://www.useit.com/alertbox/980405.html.

NIST. (2001). Risk Management Guide for Information Technology Systems, SP 800-30 Retrieved November 23, 2009, from http://csrc.nist.gov/publications/PubsSPs.html.

NIST. (2004a). FIPS 199: Standards for Security Categorization of Federal Information and Information Systems. Retrieved August 28, 2009, from http://csrc.nist.gov/publications/fips/fips199/FIPS-PUB-199-final.pdf.

NIST. (2004b). Guide for the Security Certification and Accreditation of Federal Information Systems, SP 800-37. Retrieved January 30, 2010, from http://csrc.nist.gov/publications/nistpubs/800-37/SP800-37-final.pdf.

NIST. (2006). FIPS 200: Minimum Security Requirements for Federal Information and Information Systems. Retrieved December 1, 2009, from http://csrc.nist.gov/publications/fips/fips200/FIPS-200-final-march.pdf.

NIST. (2008a). Guide for Mapping Types of Information and Information Systems to Security Categories, Volume I, SP 800-60 Rev. 1. Retrieved August 27, 2009, from http://csrc.nist.gov/publications/nistpubs/800-60-rev1/SP800-60_Vol1-Rev1.pdf.

NIST. (2008c). Guide for Assessing the Security Controls in Federal Information Systems, SP 800-53A. Retrieved November 11, 2009, from http://csrc.nist.gov/publications/PubsSPs.html.

NIST. (2009). National Institute of Standards and Technology, main website. from http://www.nist.gov.

NIST. (2009a). NIST Working Definition of Cloud Computing v15. Retrieved October 7, 2009, from http://csrc.nist.gov/groups/SNS/cloud-computing/index.html.

NIST. (2009b, 12 August 2009). *Recommended Security Controls for Federal Information Systems and Organizations,* SP 800-53 Rev 3 from http://csrc.nist.gov/publications/nistpubs/800-53-Rev3/sp800-53-rev3-final-errata.pdf.

Pieters, W. (2006). Acceptance of Voting Technology: Between Confidence and Trust. In K. Stølen (Eds.), iTrust. Lecture Notes on Computer Science 3986, pp. 283-297.

Polat, H. and Du, W. (2005). Privacy-preserving top-N recommendation on horizontally partitioned data. In Proceedings of. Citeseer.

Pollach, I. (2007). What's wrong with online privacy policies? *Commun. ACM, 50*(9): 103-108.

Price, M. and Partners, A. (2008). The Paradox of Security in Virtual Environments. *Computer, 41*(11): 22-28.

RedJasper. (2007). Red Jasper Ltd. Journal Ranking. from www.journal-ranking.com.

Robbin, A. and Koball, H. (2001). Seeking explanation in theory: Reflections on the social practices of organizations that distribute public use microdata files for research purposes. *Journal of the American Society for Information Science and Technology, 52*(13): 1169-1189.

Rotenberg, M. (1998). The Privacy Law Sourcebook: United States Law, International Law, and Recent Developments. *Washington, DC: Electronic Privacy Information Center*.

Saee, J. (2006). Journal of Management Systems. Retrieved July 27, 2009, from http://www.aom-iaom.org/jms_new.html.

Sarathy, R. and Muralidhar, K. (2006). Secure and useful data sharing. *Decision Support Systems, 42*(1): 204-220.

Schmidt, S. (2009). Commissioner rules Facebook has 'serious' privacy gaps. *Edmonton Journal*. Edmonton, Canwest News Service.

Schneier, B. (2005). Risks of third-party data.

Shaw, N. and Yadav, S. (2001). DEACON: an integrated approach to the analysis and design of enterprise architecture-based computer networks. *Communications of the Association for Information Systems*, *7*(1): 11.

Spiekermann, S. and Cranor, L. F. (2009). Engineering Privacy. *IEEE Transactions on Software Engineering 35*(1): 67-82.

Susan, L. (2008). Privacy and Security A Multidimensional Problem. *Communications of the ACM*, *51*(11): 25-26.

Thuraisingham, B. (2005). Directions for security and privacy for semantic e-business applications. *Communications of the ACM*, *48*(12): 73.

Tsarouchis, C., Denazis, S., Kitahara, C., Vivero, J., Salamanca, E. et al. (2003). A policy-based management architecture for active and programmable networks. *IEEE Network*, *17*(3): 22-28.

Vaughan-Nichols, S. (2008). Virtualization sparks security concerns. *Computer*, *41*(8): 13-15.

Walter, C. (2005). Kryder's Law. *August 2005 Scientific American Magazine*.

Zetter, K. (2004). Free E-Mail With a Steep Price? , Retrieved August 22, 2009, from http://www.wired.com/techbiz/media/news/2004/04/62917.

Zhen, J. (2009). Security and Compliance in the Age of Clouds. *Zhen 2.0*, Retrieved January 13, 2010, from http://www.zhen.org/zen20/category/security-compliance/.

Appendix A Literature review search results

In this appendix, the results of the literature review are presented. The keywords noted in the leftmost column were used to search the journals on, see chapter 4.

The keywords produced results ("found column"), after which 3 steps were taken before an article was denoted as "interesting":

1. The title of the article was scanned for relevance
2. The selection criteria were applied to the articles, which are:
    a. Articles must be published in the top ranked journals stated in Table 4-1, Table 4-2 and Table 4-3
    b. Articles have to be written in English, Dutch or German.
    c. Articles have to be published in the year 2000 or later.
3. The abstract of each article that passed the above steps, was read to decide of the article should be read completely

Table 9-1: Articles found per keyword

Keyword Search Results, number of articles found and articles found interesting						
Search Engine	Scopus		Communications of the AIS		Journal of the AIS	Journal of Management Systems
Keyword	Found	Interesting	Found	Interesting	Found	Found
Network Architecture	262	**4**	2	**1**	0	0
Data classification	18	**2**	0	0	0	0
Data privacy	238	**10**	0	0	0	0
Confidential information	13	**2**	0	0	0	0
Grid computing	58	**1**	1	0	0	0
Virtualization	44	**3**	1	0	0	0
Data secrecy	1	0	0	0	0	0
Cloud computing	6	0	0	0	0	0
Confidentiality requirements	1	0	0	0	0	0
Secrecy requirements	3	0	0	0	0	0
Classify architecture	1	0	0	0	0	0
Distributed security	2	0	0	0	0	0
Security framework	1	0	0	0	0	0
Network framework	5	0	0	0	0	0
Distributed data	129	0	1	0	0	0
Secret data	7	0	0	0	0	0
Cloud classification	5	0	0	0	0	0
Network classification	14	0	0	0	0	0

Keyword Search Results, number of articles found and articles found interesting						
Search Engine	Scopus		Communications of the AIS		Journal of the AIS	Journal of Management Systems
Keyword	Found	Interesting	Found	Interesting	Found	Found
Confidentiality in the cloud	0	0	0	0	0	0
De-perimeterization	0	0	0	0	0	0
Computer architecture classification	0	0	0	0	0	0
Architecture classification	0	0	0	0	0	0
Cloud Confidentiality	0	0	0	0	0	0
Cloud Computer Confidentiality	0	0	0	0	0	0
Confidentiality in Grid computing	0	0	0	0	0	0
Privacy in Cloud Computing	0	0	0	0	0	0
Handling confidential distributed data	0	0	0	0	0	0
Handling secret data	0	0	0	0	0	0
Security classifications	0	0	0	0	0	0
Securing remote data	0	0	0	0	0	0
Confidentiality Integrity Availability	0	0	0	0	0	0
Network security class	0	0	0	0	0	0
Distributed data protection	0	0	0	0	0	0
Secure distributed data	0	0	0	0	0	0
Modeling a cloud	0	0	0	0	0	0
Cloud modeling	0	0	0	0	0	0
Cloud certification	0	0	0	0	0	0
Security model Cloud	0	0	0	0	0	0
Cloud architecture	0	0	0	0	0	0
CIA aspects cloud	0	0	0	0	0	0

Appendix B Literature Analysis

This appendix contains the 23 articles identified in the Literature review as relevant enough to completely read and analyze them. Per article a summary is given, after which a short description of the exact relevance of these papers is given.

Title: Engineering Privacy (Spiekermann et al. 2009)

Search Keyword: Data Privacy

This paper consists of two parts.

The first part discusses privacy requirements by looking at it from two perspectives:

- System Activities; what information system task is being performed
  - Data Transfer
  - Data Storage
  - Data Processing
- Four impact factors on privacy variables
  - *How* are the tasks performed
  - *What type of data* is involved
  - *Who* uses the data
  - *In which of the spheres* (privacy responsibility domains) does the activity occur:
    - *User sphere*; location of data is fully controllable by a user, the user is responsible
    - *Recipient sphere*; company-centric sphere of control, control lies with the company
    - *Joint sphere*; companies hosting people's data and providing services. Users and providers have a joint control about access to data

The second part shows guidelines for building privacy-friendly systems based on three approaches:

- *Privacy-by-policy;* based on implementation of notice and choice principles of Fair Information Practices (FIP), on which European privacy legislation is based.
- *Privacy-by-architecture:* Using mechanisms to anonymize any information, resulting in little or no personal data being collected at all.
- *Hybrid approach:* The combination of the above two, where privacy-by-policy is enhanced through technical mechanisms that audit or enforce policy compliance.

These approached are used to make architectural choices on two dimensions:

- *Network Centricity:* The degree of control a network operator has over client's operations
- *User Identifiability:* The degree to which data can be directly related to an individual

Relevance to our research: High, a lot of the above concepts can be considered for porting to the cloud computing paradigm, if we can substitute "personal privacy" with "corporate privacy" and "user" to "data owner". The mechanisms named for privacy-by-architecture are focused on client-centric architecture and anonymous transactions, which are mechanisms pointed in the opposite direction of the network-centric architecture of Cloud Computing. The increasing protection from privacy-by-policy to privacy-to-architecture is a notion that we can use as severity of protection in our research.

Title: What's wrong with online privacy policies (Pollach 2007)

Search keyword: Data Privacy

Privacy policies published on websites seem to be drafted with the threat of privacy litigations in mind rather than commitment to fair data handling policies. The legalistic nature of the published policies discourages internet users to read them, missing an opportunity to build trust.
Some narrative policies are written with linguistic patterns that are either the result of poor writing skills, or are aimed at deceiving and confusing readers to what happens to their data.
Relevance to our research: It is important to write policies for all interested parties. The model built must be clear and transparent for all interested parties.

Title: Analyzing regulatory rules for privacy and security requirements (Breaux and Ant 2008)

Search keyword: Data Privacy
The methodology analyses an entire regulation (such as HIPAA), to extract security requirements in a systematical way. With the methodology data access requirements are extracted and priorities between data access requirements can be acquired. The method makes it easy to prove that a system is accountably compliant with the law. As a case study, the authors use their methods on the HIPAA regulation.
Relevance to our research: The method may be used as reference once cloud providers and/or users have to be compliant with a regulation. If that is the case, the method presented can be used to see what the system requirements are and if the cloud provider is able to meet these requirements. The case study on HIPAA can be used to identify protection mechanisms on every layer of the technology stack.

Title: A Knowledge architecture for IT Security (Kesh and Ratnasingam 2007)

Search keyword: Network Architecture
The paper presents the Information Security Knowledge Architecture (ISKA) which can be used by organizations to assess their IT security knowledge needs by determining the current and preferred future knowledge architecture. The ISKA assists organizations to determine the quality, completeness and effectiveness of their IT security Knowledge. The model consists of four components and six interfaces between the components:

- *Stakeholders:* Users that should have some IT security knowledge. Stakeholders can be internal and external to an organization, with different knowledge needs.
- *Knowledge dimensions:* What information must be known to deploy IT security.
- *Knowledge characteristics:* Types of knowledge. Knowledge can be declarative, procedural, individual (tacit knowledge), social, conditional, relational and pragmatic.
- *Knowledge resources:* How the knowledge can be obtained, and how can tacit knowledge be made explicit and transferable.

The interfaces are either between the stakeholders and each KM component (primary interfaces), or between the KM components, excluding the stakeholders (secondary interfaces).

- *Stakeholder and Knowledge dimension* interface represents which security responsibilities lie with the stakeholder.
- *Stakeholder and Knowledge resources* interface represents the access of the stakeholder to the correct IT security information.
- *Stakeholder and Knowledge characteristics* interface lets an organization determine the characteristics of what type of knowledge each stakeholder has, so that organizations can provide the correct resources for converting tacit knowledge to explicit knowledge.

- *Characteristics and dimensions* interface tries to relate characteristics such as tacit or explicit knowledge, to knowledge dimensions like legal and ethical issues related to security.
- *Characteristics and resources* interface relates knowledge characteristics to the resources where the knowledge is available.
- *Resources and dimensions* interface explores whether knowledge dimensions are linked to knowledge resources.

ISKA can help organizations to involve all the stakeholders in the security management process. This framework lets organizations identify *who* needs or has knowledge, *what* information is known, *what type* of knowledge is available and *how* the IT security knowledge can be obtained.
Relevance to our research: ISKA can help organizations to identify the components needed once an organization has decided to integrate their IT with cloud computing. What cloud computing knowledge needs to be obtained by whom in the organization and by which way. This concept can be used in the explanation of our model to the related parties.

Title: DEACON: An integrated approach to the analysis and design of enterprise architecture-based computer networks (Shaw et al. 2001)

Search keyword: Network Architecture
The methodology presented can be used by organizations to develop computer networks that are integrated with business requirements and business goals. Design of Enterprise Architecture-based Computer Networks (DEACON) uses the following steps in the design:

1. *Problem Definition:* Define organizational goals and objectives, IS goals and objectives, and network goals and objectives
2. *Requirement Specification:* Model business processes and organizational data
3. *Location Model:* Construct location connectivity diagrams, and use data-location and process-location matrices to refine location connectivity diagrams, process models and data models (what processes and data are active in which locations)
4. *Network Architecture:* Design architecture diagram from the location connectivity diagram, assign processes and data to nodes and match available technology to the architecture diagram
5. *Network Performance Evaluation:* Simulate network operations to identify bottlenecks, optimize and refine the network architecture
6. *Implementation:* Implement the architecture and convert to the new network.

Relevance to our research: Although confidentiality is not a part of DEACON, the approach presented may come to use when a company wished to upgrade or build a new network, taking cloud computing into mind.

Title: Security in Grid Computing: a review and synthesis (Cody et al. 2008)

Search keyword: Grid Computing
The paper provides a detailed review and analysis of current literature and presents a classification framework for the analysis of current grid computing security solutions. The authors place their framework in relation to three types of grid computing systems, each with their own vulnerabilities:

- *Computational Grid*: Focused on computing power, solving complex problems
- *Data Grid*: Used to store and access large volumes of data, often distributed across multiple domains
- *Service Grid*: A grid which provides services that are not available on a single machine

Combinations of multiple types are possible, inheriting the vulnerabilities of each type.

The classification framework consists of four main categories, each having unique properties how to accomplish grid security and to what situations they best apply to:

- *System Solutions* deal with manipulations of software and hardware directly in order to achieve security. There are two subcategories:
  - *System Security for Grid Resources* focuses on protecting grid resources, such as hardware, applications, data and communication channels. Solutions in this category address Data grids and Service grids.
  - *Intrusion Detection Systems* (IDS) function in the computational and service grids.
- *Behavioral Solutions* use policy and management controls in order to maintain security in the grid. Behavioral Solutions are intangible and intuitive and are based on policies and/or trust:
  - *Comprehensive Policy Controls* govern a wide range of grid computing actions, instead of focusing on one area of activity. Policies function best in computational grids.
  - *Trust-based security solutions* function in computational and data grids. Trust solutions can be used to lower security overhead. If trust-levels are too low then additional security mechanisms are enacted.
- *Hybrid Solutions* is a category that combines System solutions and Behavioral solutions. Authentication and Authorization based solutions fall in this category.
- *Related Technologies* are taken from areas other than grid computing, in which the security solutions bear similarity to those required by grid computing. The describes related technologies could function within data and service grids.

The authors point out that further research in the area of high security vs. high performance is of high importance. The literature review only found one study which identifies high-performance, high-security computing as its primary goal.

Relevance to our research: A lot of presented concepts can be imported to the area of cloud computing, but with caution; nodes in grid computing are independent and outside the centralized control. System solutions have a lot in common with privacy-by-architecture, while behavioral solutions have a lot in coming with privacy-by-policy, while Hybrid solutions seem to be a combination of privacy-by-policy and privacy-by-architecture.

Title: Using lessons from health care to protect the privacy of library users: Guidelines for the De-Identification of library data based on HIPAA (Nicholson and Smith 2007)

Search keyword: Data Privacy

Methods to create de-identified library data, based on Health Insurance Portability and Accountability Act (HIPAA). User-information is de-identified once the usage of library sources has been completed by the user. Direct identifiers are removed, while indirect identifiers are generalized to be remove the possibility of aggregate data to identify individuals.

Library privacy policies should contain three clear fields to educate the user:

- What data fields are and are not collected
- What the data will and will not be used for
- How users can remove their own data from the system

Relevance to our research: The paper is focused on personal privacy, not mentioning corporate privacy concerns. Both these privacy issues are of relevance in cloud computing environments. The Urge to create clear policies was also covered in “What’s wrong with online privacy policies.” Clear communication to users is required.

Title: A roadmap for comprehensive online privacy policy management (Antón, Bertino, Li et al. 2007)

Search keyword: Data Privacy

A framework is presented that supports the complete privacy policy lifecycle. The framework consists of two sides, the Enterprise side and the User side. The Enterprise side part of the framework has three top-down tiers:

- *Top Tier*: High level privacy policies that describe what goals to achieve, not how. This tier communicates with the user.
- *Middle Tier*: Traditional privacy policies that enforce the high-level privacy policies on application level. Middle tier policies govern the privacy protecting policies on the bottom tier.
- *Bottom Tier*: Fine-grained policies for enforcement in the physical layer.

The User side of the framework consists of agents that help to specify the users privacy preferences in a formal language, so that the matching between enterprises' policies and users' preferences can be done automatically, greatly enhancing the usability of the framework.

How a lot of the paradigms and communications are going to be specified, are presented as challenges and proposed for further research.

Relevance to our research: Although the paper presents a lot of open issues, the suggested top-down policy management and automated communication with the user, might be used in our model how to clarify operations of cloud providers and how Enterprises policies might automatically be matched to a cloud user. Quite interesting.

Title: Secure and useful data sharing (Sarathy and Muralidhar 2006)

Search keyword: Data Privacy

The paper identifies potential Operations Research/Management Science (OR/MS) research opportunities, which are motivated by confidentiality and privacy concerns in organizations and government agencies. At present, most information is either shared without concern for (or knowledge of) security issues, or not shared at all due to security concerns. OR/MS tries to find compromises between the two extremes, ensuring that data sharing is both useful to organizations and is secure. A framework is presented for secure and useful data sharing, from top to bottom:

- *Context*: In which context is data shared. Others contexts are similar to one of these four:
    - Government Data Dissemination (Government to public data sharing)
    - Government Record Linkage (Combining multiple sources of data)
    - Data Provider (Collecting and reselling of data as business model)
    - Data Exchange
- *Data Abstraction*: What type of data is shared and what risks does each type have, independent of context. Not all abstractions are used in each context
    - Multivariate Dataset (de-identified set of categorical and numerical variables)
        - Risk of disclosing confidential information
        - Risk of re-identification
    - Individual Record
        - Risk of disclosing confidential information
    - Subset (specific de-identified subset of data, meeting requirements of the client)
        - Risk of disclosing confidential information
        - Risk of re-identification
- *Tools and Techniques:* The paper classifies existing research in two categories:

- *Disclosure Prevention Mechanisms*, containing the techniques:
  - *Concealment*
  - *Suppression*
  - *Camouflage*
- *Record-Linkage*, containing the techniques:
  - *Re-Identification*
  - *Consolidation*

The rest of the paper identifies the research opportunities on the tools and techniques, related to the problems originating in the data abstraction layer.

Relevance to our research: The separation between the context (in what situation is the data used) and the data abstraction (what kind of data) may be interesting to incorporate in our model. But it is only useful when cloud users and providers are willing to share in the first place.

Title: Maximizing sharing of protected information (Dawson, De Vimercati, Lincoln et al. 2002)

Search keyword: Data Classification

The paper presents a technique that classifies database objects and the relation between the data objects, in order to prevent inference and data association attacks that are serious threats to database systems. The classification algorithm is based on three constraints:

- *Basic constraints* reflect the sensitivity of the information in the data object
- *Inference constraints* are used to prevent bypassing of basic constraints by data inference. Data inference refers to the possibility to determine a high-classified attribute value from the values of one or more low-classified attributes.
- *Association constraints* constrict the combined visibility of multiple attributes, when the combination of multiple values is considered to be of higher classification.

This construction permits it to specify the relationships between security levels of a set of one or more attributes and the level of another attribute or explicit level. The technique described prevents overclassification of data, in favor of usability. The algorithm can compute a minimal classification of an unclassified set of data, or it can optimize a classified set of data.

Relevance to our research: Paper presents an algorithm to efficiently produce a minimal classification of data. The basic classification of each data object is presumed to be already there. The paper can be relevant if we need a data classification algorithm as input for the model.

Title: Essential Classification Rule Sets (Baralis and Chiusano 2004)

Search keyword: Data Classification

Given a class model constructed from a set of labeled data, classifiers assign new unclassified data to the appropriate class. The classifier is a set of associative rules upon which classification decisions are taken. The problem with today's expanding data sets is that the set of association rules is expanding at a much higher rate, with an increasing number of redundant rules. By way of *rule compression*, the paper suggests a way to distill an essential rule set from a complete rule set, while both contain the same classification information. The essential rule set is a general-purpose compact rule set which can be used to generate various associative classifiers. Existing classifiers can exploit essential rule sets to improve their efficiency by reducing the number of association rules.

Relevance to our research: Highly mathematical paper. The ideas might become relevant if a highly complex model is built, but it is probably only destined for further research, if relevant at all.

Title: The role of Service Level Agreements in Relational Management of Information Technology Outsourcing: An Empirical Study (Goo, Kishore, Rao et al. 2009)

Search keyword: Network Architecture

This paper analyses SLA's and Relational Governance and identifies that these two paradigms act as complements, rather than as substitutes as some other studies indicate. The paper analyses 11 formal contract elements and categorizes them into the three categories Foundation Characteristics, Change Characteristics and Governance Characteristics. In the presented research model, the characteristics interact with Relation Governance, which is based on trust and social identification. The paper concludes that well-structures SLA's not only provide a way for measuring a service providers performance, but also provide a way to effectively manage IT outsourcing engagements through the development of relational governance. The combination of well-structured SLA's and Relational Governance can deliver a much higher exchange performance than each governance choice in isolation.

Relevance to our researcher : The study shows that, in the context of cloud computing, it is likely that SLA's nor relational governance on its own is successful enough to govern the relation between cloud service user and cloud service provider. The combination of both has the highest chance of success to remove the paranoia of distrust in placing business critical data outside one's own control.

Title: A policy-based management architecture for active and programmable networks (Tsarouchis, Denazis, Kitahara et al. 2003)

Search keyword: Network Architecture

The paper describes a policy-based network management (PBNM) architecture as part of the Future Active IP Networks (FAIN) project. The FAIN project has as main objective the development of an Active Network architecture oriented toward dynamic service deployment in heterogeneous networks. The PBNM has three actors:

- *Active Network Service Provider (ANSP),* which is the primary owner of network resources and provides facilities for the deployment and operation of services offered to SP's. The whole offering takes the form of a virtual network.
- *Service Providers (SP)* buy network resources from the ANSP, deploys services and offers the services to Consumers.
- *Consumers (C)* are the end users of a service offered by a SP and can take the form of traditional end users, an internet application or even another SP.

The PBNM is based on a two-tiered architecture: The network management level containing the network management system (NMS), and the element management level, containing the element management system (EMS). Service Level Agreements (SLAs) enter the architecture at the NMS, which then create the correct decision policies and enforcement policies. The decision and enforcement policies are mapped to the lower tier policies in the EMS level.

The management architecture described above is deployed in virtual management instances, which manage virtual environments. This construction enables the deployment of multiple virtual networks and their virtual management architectures, on the same physical infrastructure.

The PBNM system of the ANSP can be used to instantiate another management system for a SP, so SPs do not have to build their own management architectures from scratch. This paradigm is recursive, so SPs can delegate their own management instances to Consumers.

Relevance to our research: The PBNM architecture and its actors have a lot in common with today's Cloud Computing paradigm. Delegation of management architectures from network providers to service providers and subsequently to consumers has very appealing properties.

Title: The paradox of security in virtual environments (Price et al. 2008)

Search keyword: Virtualization

The paper describes virtualization with its advantages, but also its disadvantages and threats new in IT security. The advantages are lower administrative overhead, easier management and new ways to combat OS-level security vulnerabilities because there is a control layer between the OS and the physical layer. The disadvantages are VM proliferation with more and more insecure, unpatched or improperly configured VM's that are stored on the shelf and are not deleted when other improved VM's are available.

Virtualization brings new vulnerabilities. The Virtual Machine Monitor (VMM) is a single point of failure for multiple systems. If the VMM is compromised, all virtual systems managed by the VMM are compromised, while the virtual systems can't detect this compromise due to the mechanics of virtualization.

Relevance to our research: Security minded people that are interested in Cloud Computing should be aware of the technologies used in the lower layers of the technology stack, including the vulnerabilities that are there and should inform themselves on how these vulnerabilities are protected. Proper knowledge may stimulate cloud adoption. The paper presents advantages and disadvantages of virtualization in an objective way.

Title: Risks of Third-Party Data (Schneier 2005)

Search keyword: Data Privacy

Third-party data has a privacy risk and an identity theft risk. Data that was once under direct control is now controlled by others. Users have no option but to trust companies with their security and privacy, even when they have no incentive to protect them. The author Bruce Schneier argues that users should be able to control their own data, regardless of where it is stored.

Relevance to our research: Plea for data-centralized control. No new information.

Title: Privacy-Preserving Top-N Recommendation on Distributed Data (Polat and Du 2005)

Search keyword: Data Privacy

A method is suggested to share corporate privacy protected data with other entities in order to increase knowledge while protecting data owners' privacy. Participants are able to set and find an equilibrium among accuracy, privacy and efficiency. Very mathematical paper.

Relevance to our research: Method is meant for the data-mining market, specifically targeting binary data sets. Does not consider personal privacy but does consider corporate privacy. Not relevant for our research.

Title: Privacy and Security; A Multidimensional Problem (Susan 2008)

Search keyword: Data Privacy

Column pointing out that the rush toward releasing a product results in little economic incentive to spend time on properly designing privacy and security in systems. Legal and policy systems do not keep up with technology advances, resulting in a bad state of privacy and security mechanisms, of which the technologists who built the systems, bear part of the responsibility.

Relevance to our research: Low relevance, it is the same point of view we have.

Title: Directions for security and privacy for semantic E-business applications (Thuraisingham 2005)

Search keyword: Data Privacy

The Semantic Web has capabilities such as inference capabilities, which exacerbate privacy and security problems. The author pleas for a focused research program that addresses security for the Semantic Web and its layers. Privacy implications due to Semantic Web mining also needs attention.

Relevance to our research: Low, plea made in 2005. No real content.

Title: Privacy protection of binary confidential data against deterministic, stochastic, and insider threat (Garfinkel, Gopal and Goes 2002)

Search keyword: Confidential Information

The authors enhance the Confidentiality Via Camouflage (CVC) technique to protect confidentiality on the database query level. The technique offers three protections

- *Deterministic protection:* A confidential value cannot be determined *exactly* from any set of queries
- *Stochastic protection:* A confidential value cannot be guessed with a high probability.
- *Insider threat protection:* Confidential values are protected from someone with some knowledge of the confidential values.

The model is proposed as a software layer between users&administrators and the database itself.

Relevance to our research: The model is specifically tailored for prevention of information inference. The model assumes that direct access to confidential information is handled appropriately and is therefore not part of the paper. The model is too specific for our research; the model is made for the lower parts of the technology stack (database access via SQL).

Title: Seeking explanation in theory: Reflections on the social practices of organizations that distribute public use microdata files for research purposes. (Robbin and Koball 2001)

Search keyword: Confidential Information

Big survey about what Statistical Disclosure Limitation (SDL) methods organizations use before releasing public use microdata files, containing longitudinal, administrative or contextual data. There are two approaches, being *restricted data* and *restricted access*. Restricted data is achieved by statistical techniques that alter the data itself, while restricted access are administrative procedures to provide access control. The survey was meant as a small exploratory survey, but the results were so unexpected the authors searched for explanations.

Relevance to our research: Eight year old survey on confidentiality preserving practices within organizations that produce public data. The results are so various, no clear overview is given of the results.

Title: Group Communication Specifications: A comprehensive Study (Chockler, Keidar and Vitenberg 2001)

Search keyword: Network Architecture

The authors analyzed 30 Group Communication System (GCS) specifications and present a framework for classifying, analyzing and comparing Group Communication Systems. Group Communication is a means for providing multipoint to multipoint communication, by organizing processes in groups. The framework can be used by builders of group communication systems to

understand and specify their service semantics, while the survey of the 30 GCS's enables the builders to compare their service to others.
Relevance to our research: While the abstract seemed relevant, GCS have little in common with the cloud computing paradigm and is primarily about group membership and synchronized multicast messaging.

Title: Interconnection networks: Architectural challenges for utility computing data centers (Lysne, Reinemo, Skeie et al. 2008)

Search keyword: Virtualization
The paper discusses the advances made on the area of virtual interconnection networks within utility computing data centers. The goal is to support the same seamless virtualization found in other parts of hardware, such as CPUs. The challenges on virtual networks that must be handled by further research are:

- *Flexible Partitioning:* Assigning partitions of the network to each job, while preserving the job requirements and preventing misbehaving jobs from using resources needed by other jobs.
- *Fault Tolerance*: Handling faulty components in the network should be handles as invisible as possible, affecting the least number of jobs en leaving other jobs uninterrupted
- *Predictable Service*: Sharing switches and links between virtual servers requires Quality of Service guarantees. Current QoS mechanisms such as link-level flow control introduces congestions, which reduce the network's overall performance.

Relevance to our research: The paper scrutinizes previous research on the topic of virtual networks, identifying rough areas for further research. The research is on the lowest level of the technology stack (virtualization and physical layer). The paper is not relevant to our specific research.

Title: Virtualization sparks security concerns (Vaughan-Nichols 2008)

Search keyword: Virtualization
The paper points out that virtual systems can not as easily be protected as physical systems. Only OS vulnerabilities can be handled the same way as normal physical systems.
Hypervisor vulnerabilities pose the biggest threat as a compromised hypervisor could compromise all the virtual servers running on the hypervisor.
Configuration management can become a problem as the number of Virtual Machines rises, making it difficult to keep each VM up-to-date with security patches, virus and spyware signatures.
Traditional security mechanisms that inspect network packets, such as Intrusion Prevention / Detection Systems cannot inspect packets between Virtual Machines on the same physical host.
In response to these security issues, vendors are developing virtualization-security tools that automate patch-management, provide intrusion detection within one hypervisor and other services. The development of these virtualization security-tools may take several years to mature.
Relevance to our research: Security minded people that are interested in Cloud Computing should be aware of the technologies used in the lower layers of the technology stack, including the vulnerabilities that are there and should inform themselves on how these vulnerabilities are protected. Proper knowledge may stimulate cloud adoption. The paper presents disadvantages of virtualization and approaches to counter these disadvantages in a subjective way, quoting people and highlighting new products.

Appendix C Technical control baseline - summary

This is a subset of the complete security control baseline symmary, as published in NIST SP 800-53, Appendix D (NIST 2009b). This subset only contains the technical controls, and does not mention the operational and management security controls.

The table shows for each technical control, if it is recommended to be in the baseline for the security plan of an information system classified as Low, Moderate, or High. A control mentioned as 'Not Selected' for a certain impact level, means that the control is not needed to be in the baseline for such an information system. A number between parentheses, such as AC-2 (1), means that that control is recommended, together with the first control enhancement.

The full description of each control and control enhancement, is given in Appendix D.

The priority column stands for the order in which the controls should be implemented, where P1 has the highest priority, P3 the lowest and P0 has no priority as those controls are not part of any baseline. A full explanation of the priority codes is given in Appendix D.

CNTL NO.	CONTROL NAME	PRIORITY	CONTROL BASELINE		
			LO W	MOD	HIGH
Access Control					
AC-1	Access Control Policy and Procedures	P1	AC-1	AC-1	AC-1
AC-2	Account Management	P1	AC-2	AC-2 (1) (2) (3) (4)	AC-2 (1) (2) (3) (4)
AC-3	Access Enforcement	P1	AC-3	AC-3	AC-3
AC-4	Information Flow Enforcement	P1	Not Selected	AC-4	AC-4
AC-5	Separation of Duties	P1	Not Selected	AC-5	AC-5
AC-6	Least Privilege	P1	Not Selected	AC-6 (1) (2)	AC-6 (1) (2)
AC-7	Unsuccessful Login Attempts	P2	AC-7	AC-7	AC-7
AC-8	System Use Notification	P1	AC-8	AC-8	AC-8
AC-9	Previous Logon (Access) Notification	P0	Not Selected	Not Selected	Not Selected
AC-10	Concurrent Session Control	P2	Not Selected	Not Selected	AC-10
AC-11	Session Lock	P3	Not Selected	AC-11	AC-11
AC-12	Session Termination (Withdrawn)	---	---	---	---
AC-13	Supervision and Review—Access Control (Withdrawn)	---	---	---	---
AC-14	Permitted Actions without Identification or Authentication	P1	AC-14	AC-14 (1)	AC-14 (1)
AC-15	Automated Marking (Withdrawn)	---	---	---	---
AC-16	Security Attributes	P0	Not Selected	Not Selected	Not Selected
AC-17	Remote Access	P1	AC-17	AC-17 (1) (2) (3) (4) (5) (7) (8)	AC-17 (1) (2) (3) (4) (5) (7) (8)
AC-18	Wireless Access	P1	AC-18	AC-18 (1)	AC-18 (1) (2) (4) (5)
AC-19	Access Control for Mobile Devices	P1	AC-19	AC-19 (1) (2) (3)	AC-19 (1) (2) (3)
AC-20	Use of External Information Systems	P1	AC-20	AC-20 (1) (2)	AC-20 (1) (2)
AC-21	User-Based Collaboration and Information Sharing	P0	Not Selected	Not Selected	Not Selected
AC-22	Publicly Accessible Content	P2	AC-22	AC-22	AC-22
Audit and Accountability					

CNTL NO.	CONTROL NAME	PRIORITY	CONTROL BASELINE		
			LO W	MOD	HIGH
AU-1	Audit and Accountability Policy and Procedures	P1	AU-1	AU-1	AU-1
AU-2	Auditable Events	P1	AU-2	AU-2 (3) (4)	AU-2 (3) (4)
AU-3	Content of Audit Records	P1	AU-3	AU-3 (1)	AU-3 (1) (2)
AU-4	Audit Storage Capacity	P1	AU-4	AU-4	AU-4
AU-5	Response to Audit Processing Failures	P1	AU-5	AU-5	AU-5 (1) (2)
AU-6	Audit Review, Analysis, and Reporting	P1	AU-6	AU-6	AU-6 (1)
AU-7	Audit Reduction and Report Generation	P2	Not Selected	AU-7 (1)	AU-7 (1)
AU-8	Time Stamps	P1	AU-8	AU-8 (1)	AU-8 (1)
AU-9	Protection of Audit Information	P1	AU-9	AU-9	AU-9
AU-10	Non-repudiation	P1	Not Selected	Not Selected	AU-10
AU-11	Audit Record Retention	P3	AU-11	AU-11	AU-11
AU-12	Audit Generation	P1	AU-12	AU-12	AU-12 (1)
AU-13	Monitoring for Information Disclosure	P0	Not Selected	Not Selected	Not Selected
AU-14	Session Audit	P0	Not Selected	Not Selected	Not Selected
Identification and Authentication					
IA-1	Identification and Authentication Policy and Procedures	P1	IA-1	IA-1	IA-1
IA-2	Identification and Authentication (Organizational Users)	P1	IA-2 (1)	IA-2 (1) (2) (3) (8)	IA-2 (1) (2) (3) (4) (8) (9)
IA-3	Device Identification and Authentication	P1	Not Selected	IA-3	IA-3
IA-4	Identifier Management	P1	IA-4	IA-4	IA-4
IA-5	Authenticator Management	P1	IA-5 (1)	IA-5 (1) (2) (3)	IA-5 (1) (2) (3)
IA-6	Authenticator Feedback	P1	IA-6	IA-6	IA-6
IA-7	Cryptographic Module Authentication	P1	IA-7	IA-7	IA-7
IA-8	Identification and Authentication (Non-Organizational Users)	P1	IA-8	IA-8	IA-8
System and Communications Protection					
SC-1	System and Communications Protection Policy and Procedures	P1	SC-1	SC-1	SC-1
SC-2	Application Partitioning	P1	Not Selected	SC-2	SC-2
SC-3	Security Function Isolation	P1	Not Selected	Not Selected	SC-3
SC-4	Information in Shared Resources	P1	Not Selected	SC-4	SC-4
SC-5	Denial of Service Protection	P1	SC-5	SC-5	SC-5
SC-6	Resource Priority	P0	Not Selected	Not Selected	Not Selected
SC-7	Boundary Protection	P1	SC-7	SC-7 (1) (2) (3) (4) (5) (7)	SC-7 (1) (2) (3) (4) (5) (6) (7) (8)
SC-8	Transmission Integrity	P1	Not Selected	SC-8 (1)	SC-8 (1)
SC-9	Transmission Confidentiality	P1	Not Selected	SC-9 (1)	SC-9 (1)
SC-10	Network Disconnect	P2	Not Selected	SC-10	SC-10
SC-11	Trusted Path	P0	Not Selected	Not Selected	Not Selected
SC-12	Cryptographic Key Establishment and Management	P1	SC-12	SC-12	SC-12 (1)
SC-13	Use of Cryptography	P1	SC-13	SC-13	SC-13
SC-14	Public Access Protections	P1	SC-14	SC-14	SC-14
SC-15	Collaborative Computing Devices	P1	SC-15	SC-15	SC-15

CNTL NO.	CONTROL NAME	PRIORITY	CONTROL BASELINE		
			LO W	MOD	HIGH
SC-16	Transmission of Security Attributes	P0	Not Selected	Not Selected	Not Selected
SC-17	Public Key Infrastructure Certificates	P1	Not Selected	SC-17	SC-17
SC-18	Mobile Code	P1	Not Selected	SC-18	SC-18
SC-19	Voice Over Internet Protocol	P1	Not Selected	SC-19	SC-19
SC-20	Secure Name /Address Resolution Service (Authoritative Source)	P1	SC-20 (1)	SC-20 (1)	SC-20 (1)
SC-21	Secure Name /Address Resolution Service (Recursive or Caching Resolver)	P1	Not Selected	Not Selected	SC-21
SC-22	Architecture and Provisioning for Name/Address Resolution Service	P1	Not Selected	SC-22	SC-22
SC-23	Session Authenticity	P1	Not Selected	SC-23	SC-23
SC-24	Fail in Known State	P1	Not Selected	Not Selected	SC-24
SC-25	Thin Nodes	P0	Not Selected	Not Selected	Not Selected
SC-26	Honeypots	P0	Not Selected	Not Selected	Not Selected
SC-27	Operating System-Independent Applications	P0	Not Selected	Not Selected	Not Selected
SC-28	Protection of Information at Rest	P1	Not Selected	SC-28	SC-28
SC-29	Heterogeneity	P0	Not Selected	Not Selected	Not Selected
SC-30	Virtualization Techniques	P0	Not Selected	Not Selected	Not Selected
SC-31	Covert Channel Analysis	P0	Not Selected	Not Selected	Not Selected
SC-32	Information System Partitioning	P0	Not Selected	SC-32	SC-32
SC-33	Transmission Preparation Integrity	P0	Not Selected	Not Selected	Not Selected
SC-34	Non-Modifiable Executable Programs	P0	Not Selected	Not Selected	Not Selected

Appendix D Technical control catalog with limitations

The catalog of technical security controls in this appendix provides a range of safeguards and countermeasures for organizations and information systems. The catalog presented here is a subset of the complete security control catalog of NIST Special Publication 800-53 Revision 3 (NIST 2009b), as only the technical controls *with cloud limitations* are presented here.

Section Appendix D.1 will present the baseline controls which have cloud limitations, while section Appendix D.2 contain the full descriptions of the optional controls and/or control enhancements that have cloud limitations.

The security control structure of the controls mentioned below, consist of the following components:

1) *Control* section, which provides a brief statement of the security capabilities needed to protect a particular aspect of an information system
2) *Supplemental guidance* section, which provides additional information about the security control, but contains no requirements. This section provides considerations for implementing security controls in the context of mission requirements and the organization's operational environment. Supplemental Guidance sections may contain references to related controls. Control enchantments may also contain supplemental guidance, named under *Enhancement Supplemental Guidance* subsections. These subsections contain additional information for specific control enhancements that are not applicable to the control in general.
3) *Control Enhancements*, which provide security capabilities to:
    a. Build additional functionality to a control; and/or
    b. Increase the strength of a control

    Control Enhancements are numbered sequentially. For example, if the first three enhancements are selected for control Remote Access (AC-17), the will be named AC-17 (1)(2)(3).
4) *References* section. This section contains references to applicable and relevant federal laws, directives, policies, standards, and guidelines (e.g. FIPS and NIST Special Publications).
5) *Priority and Baseline allocation* section provides a priority listing for the implementation of baseline controls, where controls marked as *P1* should be implemented first. See the table below for an explanation of the priorities.
    The *baseline allocation* provides the initial allocation of controls and control enhancements for Low, Moderate, and High impact systems.

Priority Code	Sequencing	Action
Priority Code 1 (P1)	FIRST	Implement P1 security controls first.
Priority Code 2 (P2)	NEXT	Implement P2 security controls after implementation of P1 controls.
Priority Code 3 (P3)	LAST	Implement P3 security controls after implementation of P1 and P2 controls.
Unspec. Priority Code (P0)	NONE	Security control not selected for baseline.

Appendix D.1 Baseline controls with cloud limitations

AC-17 REMOTE ACCESS (NIST 2009b)

Control: The organization:

a. Documents allowed methods of remote access to the information system;

b. Establishes usage restrictions and implementation guidance for each allowed remote access method;

c. Monitors for unauthorized remote access to the information system;

d. Authorizes remote access to the information system prior to connection; and

e. Enforces requirements for remote connections to the information system.

Supplemental Guidance: This control requires explicit authorization prior to allowing remote access to an information system without specifying a specific format for that authorization. For example, while the organization may deem it appropriate to use a system interconnection agreement to authorize a given remote access, such agreements are not required by this control. Remote access is any access to an organizational information system by a user (or process acting on behalf of a user) communicating through an external network (e.g., the Internet). Examples of remote access methods include dial-up, broadband, and wireless (see AC-18 for wireless access). A virtual private network when adequately provisioned with appropriate security controls, is considered an internal network (i.e., the organization establishes a network connection between organization-controlled endpoints in a manner that does not require the organization to depend on external networks to protect the confidentiality or integrity of information transmitted across the network). Remote access controls are applicable to information systems other than public web servers or systems specifically designed for public access. Enforcing access restrictions associated with remote connections is accomplished by control AC-3. Related controls: AC-3, AC-18, AC-20, IA-2, IA-3, IA-8, MA-4.

Control Enhancements:

(1) The organization employs automated mechanisms to facilitate the monitoring and control of remote access methods.

Enhancement Supplemental Guidance: Automated monitoring of remote access sessions allows organizations to audit user activities on a variety of information system components (e.g., servers, workstations, notebook/laptop computers) and to ensure compliance with remote access policy.

(2) The organization uses cryptography to protect the confidentiality and integrity of remote access sessions.

Enhancement Supplemental Guidance: The encryption strength of mechanism is selected based on the security categorization of the information. Related controls: SC-8, SC-9, SC-13.

(3) The information system routes all remote accesses through a limited number of managed access control points.

Enhancement Supplemental Guidance: Related control: SC-7.

(4) The organization authorizes the execution of privileged commands and access to security-relevant information via remote access only for compelling operational needs and documents the rationale for such access in the security plan for the information system.

Enhancement Supplemental Guidance: Related control: AC-6.

(5) The organization monitors for unauthorized remote connections to the information system [*Assignment: organization-defined frequency*], and takes appropriate action if an unauthorized connection is discovered.

(6) The organization ensures that users protect information about remote access mechanisms from unauthorized use and disclosure.

(7) The organization ensures that remote sessions for accessing [*Assignment: organization-defined list of security functions and security-relevant information*] employ [*Assignment: organization-defined additional security measures*] and are audited.

Enhancement Supplemental Guidance: Additional security measures are typically above and beyond standard bulk or session layer encryption (e.g., Secure Shell [SSH], Virtual Private Networking [VPN] with blocking mode enabled). Related controls: SC-8, SC-9.

(8) The organization disables [*Assignment: organization-defined networking protocols within the information system deemed to be nonsecure*] except for explicitly identified components in support of specific operational requirements.

Enhancement Supplemental Guidance: The organization can either make a determination of the relative security of the networking protocol or base the security decision on the assessment of other entities. Bluetooth and peer-to-peer networking are examples of less than secure networking protocols.

References: NIST Special Publications 800-46, 800-77, 800-113, 800-114, 800-121.

Priority and Baseline Allocation:

P1	**LOW** AC-17	**MOD** AC-17 (1) (2) (3) (4) (5) (7) (8)	**HIGH** AC-17 (1) (2) (3) (4) (5) (7) (8)

AC-20 USE OF EXTERNAL INFORMATION SYSTEMS (NIST 2009b)

Control: The organization establishes terms and conditions, consistent with any trust relationships established with other organizations owning, operating, and/or maintaining external information systems, allowing authorized individuals to:

a. Access the information system from the external information systems; and

b. Process, store, and/or transmit organization-controlled information using the external information systems.

Supplemental Guidance: External information systems are information systems or components of information systems that are outside of the authorization boundary established by the organization and for which the organization typically has no direct supervision and authority over the application of required security controls or the assessment of security control effectiveness. External information systems include, but are not limited to: (i) personally owned information systems (e.g., computers, cellular telephones, or personal digital assistants); (ii) privately owned computing and communications devices resident in commercial or public facilities (e.g., hotels, convention centers, or airports); (iii) information systems owned or controlled by nonfederal governmental organizations; and (iv) federal information systems that are not owned by, operated by, or under the direct supervision and authority of the organization. For some external systems, in particular those systems operated by other federal agencies, including organizations subordinate to those agencies, the trust relationships that have been established between those organizations and the originating organization may be such, that no explicit terms and conditions are required. In effect, the information systems of these organizations would not be considered external. These situations typically occur when, for example, there is some pre-existing sharing or trust agreement (either implicit or explicit) established between federal agencies and/or organizations subordinate to those agencies, or such trust agreements are specified by applicable laws, Executive Orders, directives, or policies. Authorized individuals include organizational personnel, contractors, or any other individuals with authorized access to the organizational information system and over which the organization has the authority to impose rules of behavior with regard to system access. The restrictions that an organization imposes on authorized individuals need not be uniform, as those restrictions are likely to vary depending upon the trust relationships between organizations. Thus, an organization might impose more stringent security restrictions on a contractor than on a state, local, or tribal government.

This control does not apply to the use of external information systems to access public interfaces to organizational information systems and information (e.g., individuals accessing federal information through www.usa.gov). The organization establishes terms and conditions for the use of external information systems in accordance with organizational security policies and procedures. The terms and conditions address as a minimum; (i) the types of applications that can be accessed on the organizational information system from the external information system; and (ii) the maximum security categorization of information that can be processed, stored, and transmitted on the external information system. This control defines access authorizations enforced by AC-3, rules of behavior requirements enforced by PL-4, and session establishment rules enforced by AC-17. Related controls: AC-3, AC-17, PL-4.

Control Enhancements:

(1) The organization permits authorized individuals to use an external information system to access the information system or to process, store, or transmit organization-controlled information only when the organization:

(a) Can verify the implementation of required security controls on the external system as specified in the organization's information security policy and security plan; or

(b) Has approved information system connection or processing agreements with the organizational entity hosting the external information system.

(2) The organization limits the use of organization-controlled portable storage media by authorized individuals on external information systems.

Enhancement Supplemental Guidance: Limits on the use of organization-controlled portable storage media in external information systems can include, for example, complete prohibition of the use of such devices or restrictions on how the devices may be used and under what conditions the devices may be used.

References: FIPS Publication 199.

Priority and Baseline Allocation:

P1	**LOW** AC-20	**MOD** AC-20 (1) (2)	**HIGH** AC-20 (1) (2)

IA-2 IDENTIFICATION AND AUTHENTICATION (ORGANIZATIONAL USERS) (NIST 2009b)

Control: The information system uniquely identifies and authenticates organizational users (or processes acting on behalf of organizational users).

Supplemental Guidance: Organizational users include organizational employees or individuals the organization deems to have equivalent status of employees (e.g., contractors, guest researchers, individuals from allied nations). Users are uniquely identified and authenticated for all accesses other than those accesses explicitly identified and documented by the organization in AC-14. Unique identification of individuals in group accounts (e.g., shared privilege accounts) may need to be considered for detailed accountability of activity. Authentication of user identities is accomplished through the use of passwords, tokens, biometrics, or in the case of multifactor authentication, some combination thereof. Access to organizational information systems is defined as either local or network. Local access is any access to an organizational information system by a user (or process acting on behalf of a user) where such access is obtained by direct connection without the use of a network. Network access is any access to an organizational information system by a user (or process acting on behalf of a user) where such access is obtained through a network connection. Remote access is a type of network access which involves communication through an external network (e.g., the Internet). Internal networks include local area networks, wide area networks, and virtual private networks that are under the control of the organization. For a virtual private network (VPN), the VPN is considered an internal network if the organization establishes the VPN connection between organization-controlled endpoints in a manner that does not require the organization to depend on any external networks across which the VPN transits to protect the confidentiality and integrity of information transmitted. Identification and authentication requirements for information system access by other than organizational users are described in IA-8.

The identification and authentication requirements in this control are satisfied by complying with Homeland Security Presidential Directive 12 consistent with organization-specific implementation plans provided to OMB. In addition to identifying and authenticating users at the information-system level (i.e., at logon), identification and authentication mechanisms are employed at the application level, when necessary, to provide increased information security for the organization. Related controls: AC-14, AC-17, AC-18, IA-4, IA-5.

Control Enhancements:

(1) **The information system uses multifactor authentication for network access to privileged accounts.**

(2) **The information system uses multifactor authentication for network access to non-privileged accounts.**

(3) **The information system uses multifactor authentication for local access to privileged accounts.**

(4) **The information system uses multifactor authentication for local access to non-privileged accounts.**

(5) **The organization:**

(a) **Allows the use of group authenticators only when used in conjunction with an individual/unique authenticator; and**

(b) **Requires individuals to be authenticated with an individual authenticator prior to using a group authenticator.**

(6) **The information system uses multifactor authentication for network access to privileged accounts where one of the factors is provided by a device separate from the information system being accessed.**

(7) **The information system uses multifactor authentication for network access to non-privileged accounts where one of the factors is provided by a device separate from the information system being accessed.**

(8) **The information system uses [*Assignment: organization-defined replay-resistant authentication mechanisms*] for network access to privileged accounts.**

Enhancement Supplemental Guidance: An authentication process resists replay attacks if it is impractical to achieve a successful authentication by recording and replaying a previous authentication message. Techniques used to address this include protocols that use nonces or challenges (e.g., TLS), and time synchronous or challenge-response one-time authenticators.

(9) **The information system uses [*Assignment: organization-defined replay-resistant authentication mechanisms*] for network access to non-privileged accounts.**

Enhancement Supplemental Guidance: An authentication process resists replay attacks if it is impractical to achieve a successful authentication by recording and replaying a previous authentication message. Techniques used to address this include protocols that use nonces or challenges (e.g., TLS), and time synchronous or challenge-response one-time authenticators.

References: HSPD 12; OMB Memorandum 04-04; FIPS Publication 201; NIST Special Publications 800-63, 800-73, 800-76, 800-78.

Priority and Baseline Allocation:

P1	**LOW** IA-2 (1)	**MOD** IA-2 (1) (2) (3) (8)	**HIGH** IA-2 (1) (2) (3) (4) (8) (9)

SC-7 BOUNDARY PROTECTION (NIST 2009b)

Control: The information system:

a. Monitors and controls communications at the external boundary of the system and at key internal boundaries within the system; and

b. Connects to external networks or information systems only through managed interfaces consisting of boundary protection devices arranged in accordance with an organizational security architecture.

Supplemental Guidance: Restricting external web traffic only to organizational web servers within managed interfaces and prohibiting external traffic that appears to be spoofing an internal address as the source are examples of restricting and prohibiting communications. Managed interfaces employing boundary protection devices include, for example, proxies, gateways, routers, firewalls, guards, or encrypted tunnels arranged in an effective security architecture (e.g., routers protecting firewalls and application gateways residing on a protected subnetwork commonly referred to as a demilitarized zone or DMZ).

The organization considers the intrinsically shared nature of commercial telecommunications services in the implementation of security controls associated with the use of such services. Commercial telecommunications services are commonly based on network components and consolidated management systems shared by all attached commercial customers, and may include third-party provided access lines and other service elements. Consequently, such interconnecting transmission services may represent sources of increased risk despite contract security provisions. Therefore, when this situation occurs, the organization either implements appropriate compensating security controls or explicitly accepts the additional risk. Related controls: AC-4, IR-4, SC-5.

Control Enhancements:

(1) The organization physically allocates publicly accessible information system components to separate subnetworks with separate physical network interfaces.

Enhancement Supplemental Guidance: Publicly accessible information system components include, for example, public web servers.

(2) The information system prevents public access into the organization's internal networks except as appropriately mediated by managed interfaces employing boundary protection devices.

(3) The organization limits the number of access points to the information system to allow for more comprehensive monitoring of inbound and outbound communications and network traffic.

Enhancement Supplemental Guidance: The Trusted Internet Connection (TIC) initiative is an example of limiting the number of managed network access points.

(4) The organization:

(a) Implements a managed interface for each external telecommunication service;

(b) Establishes a traffic flow policy for each managed interface;

(c) Employs security controls as needed to protect the confidentiality and integrity of the information being transmitted;

(d) Documents each exception to the traffic flow policy with a supporting mission/business need and duration of that need;

(e) Reviews exceptions to the traffic flow policy [*Assignment: organization-defined frequency*]; and

(f) Removes traffic flow policy exceptions that are no longer supported by an explicit mission/business need.

(5) The information system at managed interfaces, denies network traffic by default and allows network traffic by exception (i.e., deny all, permit by exception).

(6) **The organization prevents the unauthorized release of information outside of the information system boundary or any unauthorized communication through the information system boundary when there is an operational failure of the boundary protection mechanisms.**

(7) **The information system prevents remote devices that have established a non-remote connection with the system from communicating outside of that communications path with resources in external networks.**

Enhancement Supplemental Guidance: This control enhancement is implemented within the remote device (e.g., notebook/laptop computer) via configuration settings that are not configurable by the user of that device. An example of a non-remote communications path from a remote device is a virtual private network. When a non-remote connection is established using a virtual private network, the configuration settings prevent *split-tunneling*. Split tunneling might otherwise be used by remote users to communicate with the information system as an extension of that system and to communicate with local resources such as a printer or file server. Since the remote device, when connected by a non-remote connection, becomes an extension of the information system, allowing dual communications paths such as split-tunneling would be, in effect, allowing unauthorized external connections into the system.

(8) **The information system routes [*Assignment: organization-defined internal communications traffic*] to [*Assignment: organization-defined external networks*] through authenticated proxy servers within the managed interfaces of boundary protection devices.**

Enhancement Supplemental Guidance: External networks are networks outside the control of the organization. Proxy servers support logging individual Transmission Control Protocol (TCP) sessions and blocking specific Uniform Resource Locators (URLs), domain names, and Internet Protocol (IP) addresses. Proxy servers are also configurable with organization-defined lists of authorized and unauthorized websites.

(9) **The information system, at managed interfaces, denies network traffic and audits internal users (or malicious code) posing a threat to external information systems.**

Enhancement Supplemental Guidance: Detecting internal actions that may pose a security threat to external information systems is sometimes termed extrusion detection. Extrusion detection at the information system boundary includes the analysis of network traffic (incoming as well as outgoing) looking for indications of an internal threat to the security of external systems.

(10) **The organization prevents the unauthorized exfiltration of information across managed interfaces.**

Enhancement Supplemental Guidance: Measures to prevent unauthorized exfiltration of information from the information system include, for example: (i) strict adherence to protocol formats; (ii) monitoring for indications of beaconing from the information system; (iii) monitoring for use of steganography; (iv) disconnecting external network interfaces except when explicitly needed; (v) disassembling and reassembling packet headers; and (vi) employing traffic profile analysis to detect deviations from the volume or types of traffic expected within the organization. Examples of devices enforcing strict adherence to protocol formats include, for example, deep packet inspection firewalls and XML gateways. These devices verify adherence to the protocol specification at the application layer and serve to identify vulnerabilities that cannot be detected by devices operating at the network or transport layer.

(11) **The information system checks incoming communications to ensure that the communications are coming from an authorized source and routed to an authorized destination.**

(12) **The information system implements host-based boundary protection mechanisms for servers, workstations, and mobile devices.**

Enhancement Supplemental Guidance: A host-based boundary protection mechanism is, for example, a host-based firewall. Host-based boundary protection mechanisms are employed on mobile devices, such as notebook/laptop computers, and other types of mobile devices where such boundary protection mechanisms are available.

(13) **The organization isolates [*Assignment: organization defined key information security tools, mechanisms, and support components*] from other internal information system components via physically separate subnets with managed interfaces to other portions of the system.**

(14) **The organization protects against unauthorized physical connections across the boundary protections implemented at [*Assignment: organization-de*fined list of managed interfaces].**

Enhancement Supplemental Guidance: Information systems operating at different security categories may routinely share common physical and environmental controls, since the systems may share space within organizational facilities. In practice, it is possible that these separate information systems may share common equipment rooms, wiring closets, and cable distribution paths. Protection against unauthorized

physical connections can be achieved, for example, by employing clearly identified and physically separated cable trays, connection frames, and patch panels for each side of managed interfaces with physical access controls enforcing limited authorized access to these items. Related control: PE-4.

(15) The **information system routes all networked, privileged accesses through a dedicated, managed interface for purposes of access control and auditing.**

Enhancement Supplemental Guidance: Related controls: AC-2, AC-3, AC-4, AU-2.

(16) **The information system prevents discovery of specific system components (or devices) composing a managed interface.**

Enhancement Supplemental Guidance: This control enhancement is intended to protect the network addresses of information system components that are part of the managed interface from discovery through common tools and techniques used to identify devices on a network. The network addresses are not available for discovery (e.g., not published or entered in the domain name system), requiring prior knowledge for access. Another obfuscation technique is to periodically change network addresses.

(17) **The organization employs automated mechanisms to enforce strict adherence to protocol format.**

Enhancement Supplemental Guidance: Automated mechanisms used to enforce protocol formats include, for example, deep packet inspection firewalls and XML gateways. These devices verify adherence to the protocol specification (e.g., IEEE) at the application layer and serve to identify significant vulnerabilities that cannot be detected by devices operating at the network or transport layer.

(18) **The information system fails securely in the event of an operational failure of a boundary protection device.**

Enhancement Supplemental Guidance: Fail secure is a condition achieved by the application of a set of information system mechanisms to ensure that in the event of an operational failure of a boundary protection device at a managed interface (e.g., router, firewall, guard, application gateway residing on a protected subnetwork commonly referred to as a demilitarized zone), the system does not enter into an unsecure state where intended security properties no longer hold. A failure of a boundary protection device cannot lead to, or cause information external to the boundary protection device to enter the device, nor can a failure permit unauthorized information release.

References: FIPS Publication 199; NIST Special Publications 800-41, 800-77.

Priority and Baseline Allocation:

P1	**LOW** SC-7	**MOD** SC-7 (1) (2) (3) (4) (5) (7)	**HIGH** SC-7 (1) (2) (3) (4) (5) (6) (7) (8)

SC-32 INFORMATION SYSTEM PARTITIONING (NIST 2009b)

Control: The organization partitions the information system into components residing in separate physical domains (or environments) as deemed necessary.

Supplemental Guidance: Information system partitioning is a part of a defense-in-depth protection strategy. An organizational assessment of risk guides the partitioning of information system components into separate physical domains (or environments). The security categorization also guides the selection of appropriate candidates for domain partitioning when system components can be associated with different system impact levels derived from the categorization. Managed interfaces restrict or prohibit network access and information flow among partitioned information system components. Related controls: AC-4, SC-7.

Control Enhancements: None.

References: FIPS Publication 199.

Priority and Baseline Allocation:

P0	**LOW** Not Selected	**MOD** SC-32	**HIGH** SC-32

Appendix D.2 Optional controls with cloud limitations

AC-3 ACCESS ENFORCEMENT (NIST 2009b)

Control: The information system enforces approved authorizations for logical access to the system in accordance with applicable policy.

Supplemental Guidance: Access control policies (e.g., identity-based policies, role-based policies, attribute-based policies) and access enforcement mechanisms (e.g., access control lists, access control matrices, cryptography) are employed by organizations to control access between users (or processes acting on behalf of users) and objects (e.g., devices, files, records, processes, programs, domains) in the information system. In addition to enforcing authorized access at the information-system level, access enforcement mechanisms are employed at the application level, when necessary, to provide increased information security for the organization. Consideration is given to the implementation of an audited, explicit override of automated mechanisms in the event of emergencies or other serious events. If encryption of stored information is employed as an access enforcement mechanism, the cryptography used is FIPS 140-2 (as amended) compliant. For classified information, the cryptography used is largely dependent on the classification level of the information and the clearances of the individuals having access to the information. Mechanisms implemented by AC-3 are configured to enforce authorizations determined by other security controls. Related controls: AC-2, AC-4, AC-5, AC-6, AC-16, AC-17, AC-18, AC-19, AC-20, AC-21, AC-22, AU-9, CM-5, CM-6, MA-3, MA-4, MA-5, SA-7, SC-13, SI-9.

Control Enhancements:

(1) [Withdrawn: Incorporated into AC-6].

(2) **The information system enforces dual authorization, based on organizational policies and procedures for [*Assignment: organization-defined privileged commands*].**

Enhancement Supplemental Guidance: Dual authorization mechanisms require two forms of approval to execute. The organization does not employ dual authorization mechanisms when an immediate response is necessary to ensure public and environmental safety.

(3) **The information system enforces [*Assignment: organization-defined nondiscretionary access control policies*] over [*Assignment: organization-defined set of users and resources*] where the policy rule set for each policy specifies:**

(a) **Access control information (i.e., attributes) employed by the policy rule set (e.g., position, nationality, age, project, time of day); and**

(b) **Required relationships among the access control information to permit access.**

Enhancement Supplemental Guidance: Nondiscretionary access control policies that may be implemented by organizations include, for example, Attribute-Based Access Control, Mandatory Access Control, and Originator Controlled Access Control. Nondiscretionary access control policies may be employed by organizations in addition to the employment of discretionary access control policies.

For Mandatory Access Control (MAC): Policy establishes coverage over all subjects and objects under its control to ensure that each user receives only that information to which the user is authorized access based on classification of the information, and on user clearance and formal access authorization. The information system assigns appropriate security attributes (e.g., labels/security domains/types) to subjects and objects, and uses these attributes as the basis for MAC decisions. The Bell-LaPadula security model defines allowed access with regard to an organization-defined set of strictly hierarchical security levels as follows: A subject can read an object only if the security level of the subject dominates the security level of the object and a subject can write to an object only if two conditions are met: the security level of the object dominates the security level of the subject, and the security level of the user's clearance dominates the security level of the object (no read up, no write down).

For Role-Based Access Control (RBAC): Policy establishes coverage over all users and resources to ensure that access rights are grouped by role name, and access to resources is restricted to users who have been authorized to assume the associated role.

(4) **The information system enforces a Discretionary Access Control (DAC) policy that:**

(a) **Allows users to specify and control sharing by named individuals or groups of individuals, or by both;**

(b) **Limits propagation of access rights; and**

(c) **Includes or excludes access to the granularity of a single user.**

(5) **The information system prevents access to [*Assignment: organization-defined security-relevant information*] except during secure, nonoperable system states.**

Enhancement Supplemental Guidance: Security-relevant information is any information within the information system that can potentially impact the operation of security functions in a manner that could result in failure to enforce the system security policy or maintain isolation of code and data. Filtering rules for routers and firewalls, cryptographic key management information, key configuration parameters for security services, and access control lists are examples of security-relevant information. Secure, nonoperable system states are states in which the information system is not performing mission/business-related processing (e.g., the system is off-line for maintenance, troubleshooting, boot-up, shutdown).

(6) **The organization encrypts or stores off-line in a secure location [*Assignment: organization-defined user and/or system information*].**

Enhancement Supplemental Guidance: The use of encryption by the organization reduces the probability of unauthorized disclosure of information and can also detect unauthorized changes to information. Removing information from online storage to offline storage eliminates the possibility of individuals gaining unauthorized access via a network. Related control: MP-4.

References: None.

Priority and Baseline Allocation:

P1	**LOW** AC-3	**MOD** AC-3	**HIGH** AC-3

AC-4 INFORMATION FLOW ENFORCEMENT (NIST 2009b)

Control: The information system enforces approved authorizations for controlling the flow of information within the system and between interconnected systems in accordance with applicable policy.

Supplemental Guidance: Information flow control regulates where information is allowed to travel within an information system and between information systems (as opposed to who is allowed to access the information) and without explicit regard to subsequent accesses to that information. A few examples of flow control restrictions include: keeping export controlled information from being transmitted in the clear to the Internet, blocking outside traffic that claims to be from within the organization, and not passing any web requests to the Internet that are not from the internal web proxy. Information flow control policies and enforcement mechanisms are commonly employed by organizations to control the flow of information between designated sources and destinations (e.g., networks, individuals, devices) within information systems and between interconnected systems. Flow control is based on the characteristics of the information and/or the information path. Specific examples of flow control enforcement can be found in boundary protection devices (e.g., proxies, gateways, guards, encrypted tunnels, firewalls, and routers) that employ rule sets or establish configuration settings that restrict information system services, provide a packet-filtering capability based on header information, or message-filtering capability based on content (e.g., using key word searches or document characteristics). Mechanisms implemented by AC-4 are configured to enforce authorizations determined by other security controls. Related controls: AC-17, AC-19, AC-21, CM-7, SA-8, SC-2, SC-5, SC-7, SC-18.

Control Enhancements:

(1) **The information system enforces information flow control using explicit security attributes on information, source, and destination objects as a basis for flow control decisions.**

Enhancement Supplemental Guidance: Information flow enforcement mechanisms compare security attributes on all information (data content and data structure), source and destination objects, and respond appropriately (e.g., block, quarantine, alert administrator) when the mechanisms encounter information flows not explicitly allowed by the information flow policy. Information flow enforcement using explicit security attributes can be used, for example, to control the release of certain types of information.

(2) **The information system enforces information flow control using protected processing domains (e.g., domain type-enforcement) as a basis for flow control decisions.**

(3) **The information system enforces dynamic information flow control based on policy that allows or disallows information flows based on changing conditions or operational considerations.**

(4) **The information system prevents encrypted data from bypassing content-checking mechanisms.**

(5) **The information system enforces [*Assignment: organization-defined limitations on the embedding of data types within other data types*].**

(6) **The information system enforces information flow control on metadata.**

(7) **The information system enforces [*Assignment: organization-defined one-way flows*] using hardware mechanisms.**

(8) **The information system enforces information flow control using [*Assignment: organization-defined security policy filters*] as a basis for flow control decisions.**

Enhancement Supplemental Guidance: Organization-defined security policy filters include, for example, dirty word filters, file type checking filters, structured data filters, unstructured data filters, metadata content filters, and hidden content filters. Structured data permits the interpretation of its content by virtue of atomic elements that are understandable by an application and indivisible. Unstructured data refers to masses of (usually) digital information that does not have a data structure or has a data structure that is not easily readable by a machine. Unstructured data consists of two basic categories: (i) bitmap objects that are inherently non language-based (i.e., image, video, or audio files); and (ii) textual objects that are based on a written or printed language (i.e., commercial off-the-shelf word processing documents, spreadsheets, or emails).

(9) **The information system enforces the use of human review for [*Assignment: organization-defined security policy filters*] when the system is not capable of making an information flow control decision.**

(10) **The information system provides the capability for a privileged administrator to enable/disable [*Assignment: organization-defined security policy filters*].**

(11) **The information system provides the capability for a privileged administrator to configure [*Assignment: organization-defined security policy filters*] to support different security policies.**

Enhancement Supplemental Guidance: For example, to reflect changes in the security policy, an administrator can change the list of "dirty words" that the security policy mechanism checks in accordance with the definitions provided by the organization.

(12) **The information system, when transferring information between different security domains, identifies information flows by data type specification and usage.**

Enhancement Supplemental Guidance: Data type specification and usage include, for example, using file naming to reflect type of data and limiting data transfer based on file type.

(13) **The information system, when transferring information between different security domains, decomposes information into policy-relevant subcomponents for submission to policy enforcement mechanisms.**

Enhancement Supplemental Guidance: Policy enforcement mechanisms include the filtering and/or sanitization rules that are applied to information prior to transfer to a different security domain. Parsing transfer files facilitates policy decisions on source, destination, certificates, classification, subject, attachments, and other information security-related component differentiators. Policy rules for cross domain transfers include, for example, limitations on embedding components/information types within other components/information types, prohibiting more than two-levels of embedding, and prohibiting the transfer of archived information types.

(14) **The information system, when transferring information between different security domains, implements policy filters that constrain data structure and content to [*Assignment: organization-defined information security policy requirements*].**

Enhancement Supplemental Guidance: Constraining file lengths, allowed enumerations, character sets, schemas, and other data object attributes reduces the range of potential malicious and/or unsanctioned content. Examples of constraints include ensuring that: (i) character data fields only contain printable ASCII; (ii) character data fields only contain alpha-numeric characters; (iii) character data fields do not contain special characters; or (iv) maximum field sizes and file lengths are enforced based upon organization-defined security policy.

(15) **The information system, when transferring information between different security domains, detects unsanctioned information and prohibits the transfer of such information in accordance with the security policy.**

Enhancement Supplemental Guidance: Actions to support this enhancement include: checking all transferred information for malware, implementing dirty word list searches on transferred information, and applying the same protection measures to metadata (e.g., security attributes) that is applied to the information payload.

(16) **The information system enforces security policies regarding information on interconnected systems.**

Enhancement Supplemental Guidance: Transferring information between interconnected information systems of differing security policies introduces risk that such transfers violate one or more policies. While security policy violations may not be absolutely prohibited, policy guidance from information

owners/stewards is implemented at the policy enforcement point between the interconnected systems. Specific architectural solutions are mandated, when required, to reduce the potential for undiscovered vulnerabilities. Architectural solutions include, for example: (i) prohibiting information transfers between interconnected systems (i.e. implementing access only, one way transfer mechanisms); (ii) employing hardware mechanisms to enforce unitary information flow directions; and (iii) implementing fully tested, re-grading mechanisms to reassign security attributes and associated security labels.

(17) The information system:

(a) Uniquely identifies and authenticates source and destination domains for information transfer;

(b) Binds security attributes to information to facilitate information flow policy enforcement; and

(c) Tracks problems associated with the security attribute binding and information transfer.

Enhancement Supplemental Guidance: Attribution is a critical component of a security concept of operations. The ability to identify source and destination points for information flowing in an information system, allows forensic reconstruction of events when required, and increases policy compliance by attributing policy violations to specific organizations/individuals. Means to enforce this enhancement include ensuring that the information system resolution labels distinguish between information systems and organizations, and between specific system components or individuals involved in preparing, sending, receiving, or disseminating information.

References: None.

Priority and Baseline Allocation:

P1	**LOW** Not Selected	**MOD** AC-4	**HIGH** AC-4

AC-16 SECURITY ATTRIBUTES (NIST 2009b)

Control: The information system supports and maintains the binding of [*Assignment: organization-defined security attributes*] to information in storage, in process, and in transmission.

Supplemental Guidance: Security attributes are abstractions representing the basic properties or characteristics of an entity (e.g., subjects and objects) with respect to safeguarding information. These attributes are typically associated with internal data structures (e.g., records, buffers, files) within the information system and are used to enable the implementation of access control and flow control policies, reflect special dissemination, handling or distribution instructions, or support other aspects of the information security policy. The term security label is often used to associate a set of security attributes with a specific information object as part of the data structure for that object (e.g., user access privileges, nationality, affiliation as contractor). Related controls: AC-3, AC-4, SC-16, MP-3.

Control Enhancements:

(1) The information system dynamically reconfigures security attributes in accordance with an identified security policy as information is created and combined.

(2) The information system allows authorized entities to change security attributes.

(3) The information system maintains the binding of security attributes to information with sufficient assurance that the information--attribute association can be used as the basis for automated policy actions.

Enhanced Supplemental Guidance: Examples of automated policy actions include automated access control decisions (e.g., Mandatory Access Control decisions), or decisions to release (or not release) information (e.g., information flows via cross domain systems).

(4) The information system allows authorized users to associate security attributes with information.

Enhanced Supplemental Guidance: The support provided by the information system can vary from prompting users to select security attributes to be associated with specific information objects, to ensuring that the combination of attributes selected is valid.

(5) The information system displays security attributes in human-readable form on each object output from the system to system output devices to identify [*Assignment: organization-identified set of special dissemination, handling, or distribution instructions*] using [*Assignment: organization-identified human readable, standard naming conventions*].

Enhancement Supplemental Guidance: Objects output from the information system include, for example, pages, screens, or equivalent. Output devices include, for example, printers and video displays on computer terminals, monitors, screens on notebook/laptop computers and personal digital assistants.

References: None.

Priority and Baseline Allocation:

P0	**LOW** Not Selected	**MOD** Not Selected	**HIGH** Not Selected

AU-9 PROTECTION OF AUDIT INFORMATION (NIST 2009b)

Control: The information system protects audit information and audit tools from unauthorized access, modification, and deletion.

Supplemental Guidance: Audit information includes all information (e.g., audit records, audit settings, and audit reports) needed to successfully audit information system activity. Related controls: AC-3, AC-6.

Control Enhancements:

(1) The information system produces audit records on hardware-enforced, write-once media.

(2) The information system backs up audit records [*Assignment: organization-defined frequency*] onto a different system or media than the system being audited.

(3) The information system uses cryptographic mechanisms to protect the integrity of audit information and audit tools.

Enhancement Supplemental Guidance: An example of a cryptographic mechanism for the protection of integrity is the computation and application of a cryptographic-signed hash using asymmetric cryptography, protecting the confidentiality of the key used to generate the hash, and using the public key to verify the hash information.

(4) The organization:

(a) Authorizes access to management of audit functionality to only a limited subset of privileged users; and

(b) Protects the audit records of non-local accesses to privileged accounts and the execution of privileged functions.

Enhancement Supplemental Guidance: Auditing may not be reliable when performed by the information system to which the user being audited has privileged access. The privileged user may inhibit auditing or modify audit records. This control enhancement helps mitigate this risk by requiring that privileged access be further defined between audit-related privileges and other privileges, thus, limiting the users with audit-related privileges. Reducing the risk of audit compromises by privileged users can also be achieved, for example, by performing audit activity on a separate information system or by using storage media that cannot be modified (e.g., write-once recording devices).

References: None.

Priority and Baseline Allocation:

P1	**LOW** AU-9	**MOD** AU-9	**HIGH** AU-9

SC-4 INFORMATION IN SHARED RESOURCES (NIST 2009b)

Control: The information system prevents unauthorized and unintended information transfer via shared system resources.

Supplemental Guidance: The purpose of this control is to prevent information, including encrypted representations of information, produced by the actions of a prior user/role (or the actions of a process acting on behalf of a prior user/role) from being available to any current user/role (or current process) that obtains access to a shared system resource (e.g., registers, main memory, secondary storage) after that resource has been released back to the information system. Control of information in shared resources is also referred to as object reuse. This control does not address: (i) information remanence which refers to residual representation of data that has been in some way nominally erased or removed; (ii) covert channels where shared resources are manipulated to

achieve a violation of information flow restrictions; or (iii) components in the information system for which there is only a single user/role.

Control Enhancements:

(1) **The information system does not share resources that are used to interface with systems operating at different security levels.**

Enhancement Supplemental Guidance: Shared resources include, for example, memory, input/output queues, and network interface cards.

References: None.

Priority and Baseline Allocation:

P1	**LOW** Not Selected	**MOD** SC-4	**HIGH** SC-4

SC-12 CRYPTOGRAPHIC KEY ESTABLISHMENT AND MANAGEMENT (NIST 2009b)

Control: The organization establishes and manages cryptographic keys for required cryptography employed within the information system.

Supplemental Guidance: Cryptographic key management and establishment can be performed using manual procedures or automated mechanisms with supporting manual procedures. In addition to being required for the effective operation of a cryptographic mechanism, effective cryptographic key management provides protections to maintain the availability of the information in the event of the loss of cryptographic keys by users.

Control Enhancements:

(1) **The organization maintains availability of information in the event of the loss of cryptographic keys by users.**

(2) **The organization produces, controls, and distributes symmetric cryptographic keys using [*Selection: NIST-approved, NSA-approved*] key management technology and processes.**

(3) **The organization produces, controls, and distributes symmetric and asymmetric cryptographic keys using NSA-approved key management technology and processes.**

(4) **The organization produces, controls, and distributes asymmetric cryptographic keys using approved PKI Class 3 certificates or prepositioned keying material.**

(5) **The organization produces, controls, and distributes asymmetric cryptographic keys using approved PKI Class 3 or Class 4 certificates and hardware security tokens that protect the user's private key.**

References: NIST Special Publications 800-56, 800-57.

Priority and Baseline Allocation:

P1	**LOW** SC-12	**MOD** SC-12	**HIGH** SC-12 (1)

SC-13 USE OF CRYPTOGRAPHY (NIST 2009b)

Control: The information system implements required cryptographic protections using cryptographic modules that comply with applicable federal laws, Executive Orders, directives, policies, regulations, standards, and guidance.

Supplemental Guidance: None.

Control Enhancements:

(1) **The organization employs, at a minimum, FIPS-validated cryptography to protect unclassified information.**

(2) **The organization employs NSA-approved cryptography to protect classified information.**

(3) **The organization employs, at a minimum, FIPS-validated cryptography to protect information when such information must be separated from individuals who have the necessary clearances yet lack the necessary access approvals.**

(4) The organization employs [*Selection: FIPS-validated; NSA-approved*] cryptography to implement digital signatures.

References: FIPS Publication 140-2; Web: CSRC.NIST.GOV/CRYPTVAL, WWW.CNSS.GOV.

Priority and Baseline Allocation:

P1	**LOW** SC-13	**MOD** SC-13	**HIGH** SC-13