Government Cloud Computing: Requirements, Specification and Design of a Cloud-Computing Environment for Police and Law Enforcement

Andreas Tellefsen

Master's Thesis
Master of Science in Information Security
30 ECTS
Department of Computer Science and Media Technology
Gjøvik University College, 2012

Avdeling for
informatikk og medieteknikk
Høgskolen i Gjøvik
Postboks 191
2802 Gjøvik

Department of Computer Science
and Media Technology
Gjøvik University College
Box 191
N-2802 Gjøvik
Norway

Government Cloud Computing: Requirements, Specification and Design of a Cloud-Computing Environment for Police and Law Enforcement

Andreas Tellefsen

2012/07/01

Abstract

As the amount of information found at digital forensic crime scenes increases each year, conventional methods for evidence handling and analysis come under pressure to perform at a sufficient level. New techniques are needed to meet this demand, and cloud computing is a possible solution. However, little has been done to explore this possibility. Cloud computing uses the combined power of multiple servers to create a resource pool, which can then be divided between the cloud's users based on their need for computing resources. Such features could prove to be very useful for digital forensic analysis applications.

To explore the possibilities of a governmental cloud-computing environment, the benefits and drawbacks of such an implementation have been studied. A qualitative approach was used to explore and discuss the individual drawbacks and benefits for possible cloud deployment models. The thesis continues with a study of the requirements for cloud implementations for law enforcement. Here, a guide for creating such requirements was created. Finally, the requirements guide was tested using Microsoft Operations Framework, a project management methodology. This was done to explore the usability of the guide.

The results from the study of benefits and drawbacks show that the private cloud deployment model currently is the most promising alternative for law enforcement implementations. The results also state that conventional environments remain a good choice, and often preferable to cloud environments given their predictability and potential for sound security. However, we believe that this result will change to favour the cloud as the technology gradually matures.

During the second part of the thesis, an elaborate collection of requirements for the planning of a governmental cloud-computing environment was created. This covers, amongst other things, legal, compliance and structural demands. The requirements guide will enable a project to focus on cloud-specific issues, thereby limiting the potential for failure during such projects.

Finally, the requirements guide has been tested using the Microsoft Operations Framework to prove its applicability. The steps in the guide were successfully integrated into the Microsoft methodology, thereby shifting its focus towards cloud issues.

We believe that the results presented in this thesis will both provide a clearer picture of the latent possibilities of cloud computing in our day and time, and aid coming projects to focus on the cloud-specific issues.

Sammendrag

Mengden data som må samles inn på et digitalt åsted øker hvert eneste år, og konvensjonelle metoder for behandling og analyse blir lagt under stadig mer press for å fungere tilfredsstillende. Videre utvikling av metodene trengs for å møte dette behovet, og distribuerte systemer kalt cloud computing er en mulig løsning. Likevel er det gjort lite for å utforske den mulige benyttelsen av cloud computing i politisammenheng.

Cloud computing benytter kombinerte ressurser fra flere enkeltstående servere, et såkalt ressurs magasin, som gjøres tilgjengelig for flere brukere samtidig. Ressurser tildeles brukerne etter behov, og frigjøres for annen bruk når behovet ikke lenger er til stedet. En slik løsning vil kunne være til stor nytte for analyseløsninger for digitale bevis og rammeverk for disse.

For å kunne utforske mulighetene for statlige cloud-miljøer har vi studert fordelene og ulempene ved en slik implementasjon. Kvalitative studier har blitt benyttet for å vurdere og diskutere individuelle fordeler og ulemper ved de forskjellige cloud modellene som finnes. Forskningen fortsatte med en vurdering av hvilke krav som bør settes til cloud-miljøer for politimyndighet. I løpet av dette arbeidet ble det konstruert en guide for kravspesifisering for slik dokumentasjon. For å undersøke dets anvendbarhet ble dokumentet testet ved bruk av prosjektstyringsrammeverket Microsoft Operations Framework.

Utforskingen av fordeler og ulemper viste at private cloud-løsninger er det cloud-miljøet som er best skikket for statlig bruk. Videre viste det seg at konvensjonelle systemer forblir et godt alternativ, som sansynligvis kommer til å bli foretrukket. Grunnen til dette er at slike systemer er forutsigbare og støtter gjennomprøvde sikkerhetsløsninger. Vi tror likevel at cloud-løsningene kommer til å bli enda mer aktuelle etter hvert som teknologien utvikles.

Spesifikasjonsdokument for krav, som be laget i løpet av oppgavens andre del, inneholder blant annet lovmessige, standardbaserte og strukturelle krav. Dokumentet vil kunne hjelpe prosjekter med å holde fokus på cloud-spesifikke problemstillinger, noe som senker sansynligheten for at feil og mangler oppstår underveis.

Til slutt ble guiden testet ved å benytte denne i et fiktivt Microsoft Operations Framework prosjekt. Stegene i guiden ble lagt inn i rammeverkets faser, noe som gjorde at disse ble bedre skikket til å vurdere og dokumentere planleggingen av cloud-implementasjoner.

Vi håper og tror at resultatene som er presentert i denne masteroppgaven vil gjøre det lettere å se hvilke muligheter som per i dag finnes for cloud computing. Videre tror vi at kravspesifiseringsdokumentet vil kunne tilby hjelp i løpet av en planleggingsprosess for implementering av cloud miljøer for statlig bruk.

Contents

List of Figures

List of Tables

List of Abbreviations

CC - Cloud Computing

SLA - Service Level Agreement

SaaS - Software-as-a-Service

PaaS - Platform-as-a-Service

IaaS - Infrastructure-as-a-Service

AAII - Access, Availability, Infrastructure, Integrity

CSP - Cloud Service Provider

ISP - Internet Service Provider

VM - Virtual Machine

VMM - Virtual Machine Monitor

DDoS - Distributed Denial-of-Service

CoC - Chain-of-Custody

PHI - Protected Health Information

PII - Personally Identifiable Information

SMF - Service Management Function

Acknowledgements

First, I would like to thank my supervisor, Professor Katrin Franke, for her help with elaborating the thesis context and scope, providing literature and academic knowledge, and for guidance throughout the thesis period.

Secondly I would like to thank Anders O. Flaglien, who has acted as my unofficial external supervisor and sparring partner, and who provided his experience both from education and the private sector.

I would also like to thank my classmates at GUC, whose input regarding both on- and off-topic issues have aided me in my work and lifted my spirits many a time. I would especially like to thank my opponent, Lars Arne Sand, for his help and support.

Finally I would like to extend my gratitude to my family for their help throughout the years, and to my fiancée, Line Solberg Espedal, for her tireless support and patience.

1 Introduction

This chapter starts with an introduction to the topic of the thesis, then follows the problem description and a motivation section. Next comes a description of the research questions and the methodology, and finally the contributions and the outline of the thesis are presented.

1.1 Topic covered by the Thesis

The amount of data created using information technology (IT) is rapidly growing, and studies propose that this amount might more than double every two years [1]. Naturally, computing power and storage capacity will have to keep up with this pace. For law enforcement this means that the amount of digital evidence collected for a given case might significantly increase each year [2, 3].

Being able to handle this amount of data in the best possible way is a prime concern of any digital forensic analyst, as imaging and analysis of entire partitions is quickly becoming infeasible. Partitions of several terabytes are simply too large to be handled in an effective way, as the resource demands become too high for conventional servers and virtualization platforms. Re-enactments of scenarios for digital investigations are similarly difficult, as the numbers of machines, network traffic and other resource demands make conventional virtual platforms unable to sufficiently run simulations. It is clear that new technology is needed to keep up with the demands of current digital crime scenes.

Cloud computing (CC) is a new frontier in the IT world. By enabling large data centres to provision computing resources and services to the public at their demand, the individual user is able to tap into and utilize resources that are far superior to any single server. Such a reservoir of computing power is in the context of CC known as a *resource pool*, which enables the users of the cloud to boost their computing capabilities temporarily or permanently. The concept can be visualized as a cloud of computers, where the users of the cloud will commission computing resources without knowledge of which computers are used or where these are located.

The strength of a CC system becomes especially visible when there is a change in the resource demands of a part of the system. When a cloud user needs additional resources to do a job, the cloud can be asked to assign these resources in the form of more or faster virtual machines, thereby scaling up the user's infrastructure. In the same way, as the user's needs diminish, the resources can be released backs into the resource pool to be used elsewhere.

Unfortunately, criminal applications of the technology have already been discovered. By utilizing the power of scalability and elasticity, everything from cracking password hashes [4] to distributing botnet services [5] is possible. What makes this a real threat is that it can all be done by one person on one physical machine, where thousands would have been needed before.

Because of such examples, the focus of research in the area of information security has mostly highlighted the problems of conducting digital forensics in the cloud, but little has been done to explore the benefits of using such a system for law enforcement.

It seems likely that expandable resources combined with the versatility of virtualization could be put to good use for digital analysts, and potentially solves many of the emerging problems. The question is: How?

1.2 Keywords

Cloud computing, digital forensics, law enforcement, computational forensics, network security, requirements, project planning framework.

1.3 Problem Description

Finding a solution to the problems of resource deficiency for law enforcement is vital to raising the overall potency of current and future digital forensic techniques. Creating an infrastructure for the analysis of digital evidence or testing of security software that can handle large volumes and scale for large scenarios is a key functionality of the next generation of digital forensic tools. It is in this setting we believe CC will come to its right.

Many implementations of analysis environments are currently created, and academic projects like deLink [6], Jungle [7] and the Information Security Test Laboratory [8] have shown that the applications of virtualization are great in the field of digital forensics, be it malware analysis, software security analysis or scenario building. Many commercial systems have also been built for such purposes, including AD LAB computer forensics software solutions [9] and the FRED stand-alone forensic-analysis workstation [10]. This project aims to assess the applicability of CC for such tasks.

Even though the cloud seems to address the new demands of digital forensics, several problems have to be addressed when planning for CC implementations. As examples, standards, frameworks or best practices have yet to be announced for this technology, as these are still under evaluation by organizations such as the National Institute of Standards and Technology (NIST) and the Cloud Security Alliance (CSA). Furthermore, compliance frameworks have not been specifically created to cater for cloud solutions, and legal and jurisdictional issues constitute severe problem areas. Solutions therefore have to be based on frameworks that are not optimized for the cloud, which could yield problems. These and other arguments state that the cloud may not be the best choice for digital forensics.

1.4 Justification, Motivation and Benefits

CC is here to stay, and most of the corporate and governmental world have already moved some of its services and business infrastructure to the cloud. Following them is the entire criminal society, which seek to reap the benefits of this new technology in its infant stages, as implementations are conducted before the security implications of these are fully understood. In addition, the ever increasing amount of digital evidence in criminal cases implies that new techniques have to be created for the handling of such evidence [2, 3].

This provides a strong incentive for law enforcement to start using cloud systems, as the creation of a cloud environment for forensic analysis will provide the digital investigator with a strong and versatile environment for analysis applications. Such applications could assist in the analysis of digital content at a large scale. As an added bonus, the experience gained while using CC will be helpful in developing new techniques for ana-

lyzing such systems in the future.

One direct example of how digital analysis systems may benefit from introducing cloud-like systems is the Norwegian National Authority for Investigation and Prosecution of Economic and Environmental Crime ØKOKRIM [11]. This Norwegian law enforcement agency decided to base their new digital analysis system on Access Data's AD LAB software, implementing it using resource pooling in a cloud-like fashion. This was said to dramatically improve the abilities of ØKOKRIM to process and analyze data.

In the long run, the advantages of implementing in the cloud is found in the inherent qualities of CC, as a centralized cloud for a country's law enforcement could be used by all offices regardless of size. In this way, abundant computing resources, training and analysis capabilities can be made available for all personnel, and not just the local department.

1.5 Research Questions

In order to specify the problems that should be addressed, some research questions were formed. The main question will be answered based on its sub-questions.

The main question is:

- How can the field of digital forensics and law enforcement benefit from the use of cloud computing?

The sub-questions are:

- What are the potential benefits and drawbacks of implementing forensic applications in the cloud?
- What needs to be included in a specification for a private cloud-computing environment for law enforcement?
- How can one best design a functional private cloud for law enforcement?

To be able to answer these questions the field of CC and digital forensics must first be thoroughly understood. Additionally, a literature study is needed to research the usage of forensic analysis methods, and to identify parts of these that might be interesting to move to a cloud environment. Furthermore, information regarding legal, security, architectural and compliance issues must be researched in order to create a specification for such a system.

Exploration of the benefits and drawbacks of moving governmental or forensic systems to the cloud should then be performed. During this assessment, potential cloud deployment models have to be surveyed and compared to identify possible problems and further elaborate on benefits and drawbacks. Several scenarios are also created and assessed to review a possible selection of governmental cloud applications.

Following this, we must create a list of requirements that needs to be documented and discussed in order to plan the implementation of governmental systems in the cloud. The requirements must then be tested using feasible scenarios.

1.6 Methodology

This thesis will rely on a qualitative approach, as the questions are more general and complex than specific and focused. Compared to Peshkin's categories of qualitative analysis and their outcome [12] the research questions that are proposed fall within two of

these. *Interpretation* is the first one, as the project will elaborate on the existing concepts of CC and digital forensics. It might also provide insights into the identification of problems in the field. *Evaluation* will also be used, as we wish to provide the means to create detailed specifications of CC environments in law enforcement, thereby elaborating on the practices in the field. We also believe that such specifications and their implementation may strengthen the tools of digital forensics by increasing their performance.

The qualitative research design can be categorized as a case study, as explained in [13]. However, the research will be performed somewhat differently compared to the definition. The research will focus on the law enforcement community and its possible utilization of the new CC technology. Observing and documenting a case over time will therefore not be a part of this project.

To be able to answer the presented research questions, a literature study first has to be conducted in the fields of law enforcement, digital forensics and CC. The results of the study are presented in Chapter 2, and create the foundation for further elaboration of the research questions.

The order of the questions is laid out in a way that lets the results of one question be beneficial for us when attempting to answer the next in line. The questions will therefore mostly be answered in sequential order, although some parallel work between the questions will be natural. This will especially be the case with the first and the second question.

After the literature study we will collect the information regarding the advantages and disadvantages of CC for governmental agencies and law enforcement. The goal of this phase is a listing of the benefits and drawbacks as well as a table comparing different CC deployment models against each other. This table provide insight into how the benefits and drawbacks reflect in the choice of deployment models when assessed from our governmental viewpoint, and help us find the ones that provide the best foundation for such CC environments.

Several ways may exist to collect the benefits and drawbacks, and assess the size of the positive and negative effects of these. One possible solution would be to arrange a workshop, collecting participants from several of the departments of Gjøvik University College (GUC), related IT industry and law enforcement. Drawbacks and benefits could thus be collected, discussed and documented. Another approach would be to perform in-depth interviews with selected participants to map benefits and drawbacks. Questionnaires can also be used, and dispatched to a large number of participants in order to collect benefits and drawbacks, and concatenate the results into the final list. Finally, the creation of benefits and drawbacks can be based on an in-depth study of existing literature and cloud solutions. The impact of these would then be discussed at length.

The first approach listed in the preceding paragraph might be the best choice, as it would create a great foundation on which to collect a detailed and justified list of drawbacks and benefits. However, we find that given the limited amount of time and resources available to us, such a solution is not feasible. The same arguments remain for in-depth interviews, as it would require us to either travel to the subject's location, or pay for their travel to GUC. Using a questionnaire for the collection of benefits and drawbacks seemed unfitting, as it would require the respondents to describe their thoughts at great length. Although this was the case, a questionnaire could be more useful if concrete questions could be presented to the respondents.

Because of this we decided to explore possible benefits and drawbacks by doing an in-depth analysis of existing literature, and rate the possible impact levels of such based on discussions in each case. Based on the results we would then create a questionnaire presenting the benefits and drawbacks in the form of questions, and ask respondents to rate the impact levels of these. This would then give us an indication of how our discussed impact ratings compared to the results of the questionnaire.

In order to explore the results presented, the impact levels may be translated into values, which can then be used in calculations. Since we would like to explore the best cloud deployment model in a given situation, the drawback points related to a deployment model can therefore be subtracted from the benefit points of the same deployment model. The results from this simple calculation could then give us a good opportunity for further exploring the combinations of cloud deployment models, and how these compare to each other.

Another approach would be to qualitatively assess and discuss the combinations of benefits and drawbacks that relate to given deployment models. Based on this one can attempt to reach a conclusion for which model is better, for instance by applying relevant scenarios and discuss at length.

The advantage of using calculations for assessing such problems is that the results are easily measured, as a difference in numbers can be more clearly stated than a difference in discussed arguments. A downside of simply calculating is that the details behind the comparison of deployment models might be lost. This is also the benefit of discussion, as all details concerning each benefit and drawback can be exhaustively studied for each deployment model. A downside of this method is that it is far more time consuming.

Given that the study would use five cloud deployment models and four categories of benefits and drawbacks (see 3 for details) we decided that the qualitative approach would be too time consuming. Choosing to calculate the answers would also give us a chance to combine a qualitative and quantitative approach, and we decided to use this method. The details of this approach can be found in Section 3.5.

Work will then continue with defining a list of requirements for a CC environment for law enforcement, which will create the foundation for the potential implementation. The goal will here be to create a guide that can be useful during the planning of similar projects, both within research and the commercial market. In order to do this, the knowledge we gain from the related literature studies regarding CC and digital forensics, and knowledge concerning project frameworks has to be combined.

At last, our requirements document has to be tested using a relevant scenario. Several areas will have to be mapped out during this stage, such as the project planning methodology to be used, what hardware and software the scenario is based on and its goals. The reason for doing this is find out if our requirements document can provide relevant projects with assistance and guidance, regardless of the methodology used. The guide should help to focus the project towards CC by adding steps and tasks specifically tailored for this purpose.

The first two questions will be highly theoretical, and rely on the information found during the literature study. The third question will be approached in a practical way, as it concerns the implementation of a complete environment. After these have been answered, the deliverables will consist of the list of benefits and drawbacks as well as the calculations based on these, the requirements documentation and the documentation of

a fictional test scenario using our requirements guide.

1.7 Contributions

By answering the research questions we seek to investigate the possibility of using the cloud as an environment for governmental CC, and a forensic environment for investigators. As pointed out in Section 1.1, many researchers have discussed the implications of the cloud on forensics, but few, to our knowledge, has addressed the implications of utilizing the cloud for the benefit of law enforcement in the field.

A thorough investigation of benefits and drawbacks will contribute to clarify the premises on which forensic environments in the cloud should rely. Following this, the creation of a guide for exploring cloud-specific requirements will aid in the planning of such implementations. Finally, a test of the guide's applicability will be performed using MOF. The documentation of this can be used for exemplifying the design and planning process.

1.8 Thesis Outline

This section will cover the chapter distribution of the paper and the reason why they are organized as they are. Each chapter is quickly summarized in its own bullet point.

- Chapter two covers the related literature concerning forensics, CC and cloud forensics.
- Chapter three discuss the benefits and drawbacks of implementing forensic applications in the cloud. The results, presented in Section 3.5, culminate in a suggestion for the best suited cloud deployment models, and how these compare to the capabilities of conventional computing environments in several different scenarios
- Chapter four explores the layout of a complete guide for the creation of requirements for a governmental CC environment for law enforcement. The guide, found in Appendix H, is created to be framework agnostic, and to help a project to focus on cloud-specific issues.
- Chapter five covers the testing of the document created in Chapter four by using the MOF project methodology. The requirements guide is applied to the distinct phases of MOF and used to plan and design a fictional law enforcement cloud environment
- Chapter six discusses the results of the research conducted, concludes the master thesis report and provide ideas for further work based on the presented results.
- The appendices of the thesis document support the research and its findings. Amongst other things, the appandices feature tables (Appendix A, B, C and D) used for the calculation of results in chapter three, and an overview of MOF (Appendix I)

2 Related Literature

In this section a related literature study concerning the specific fields covered in this master thesis report is covered. Forensics and several fields linking to the core concepts of CC will be presented. The sections covering CC and digital forensics have in full or parts been selected from my own paper written during my time at GUC [14].

2.1 Forensics

This section covers the field of forensics both traditional and digital. As the focus point of this thesis is on the field of digital forensics, traditional forensics will only be covered briefly.

2.1.1 Definitions

Some expressions will have to be defined, which will be used throughout this report. These cover both conventional law enforcement and digital forensic terms.

Law

The Oxford dictionary define law as ”A system of rules which a particular country or community recognizes as regulating the actions of its members and which it may enforce by the imposition of penalties” [15]. Basically, this means that it is the collection of actions for which one might be punished for doing.

Law Enforcement

The Oxford dictionary definition of enforcement, when limited to law, creates the following: ”The act of compelling observance of or compliance with law” [15]. This means that law enforcement is the observation of actions according to law, and the legal pursuit of those who do not abide by these.

Police

Police can be defined as: ”The civil force of a state, responsible for the prevention and detection of crime and the maintenance of public order” [15]. One can therefore establish that police is the civil force of law enforcement.

Jurisdiction

Jurisdiction is in the Oxford dictionary defined as: ”The official power to make legal decisions and judgments” [15]. In most cases, a jurisdiction is a limited area where a certain set of laws exist. A country is an example of a jurisdiction. Another example is the operational jurisdiction of a police force, for instance the state jurisdictional areas of the United States.

The Forensic Approach

The general forensic approach is outlined in the NIST Special Publication 800-86 regarding forensic techniques in incident response [16]. The model presented in the paper is used by the digital forensic courses at GUC in a revised form. Figure 1 shown underneath is a visual presentation of this refined model, and the one we will be using.

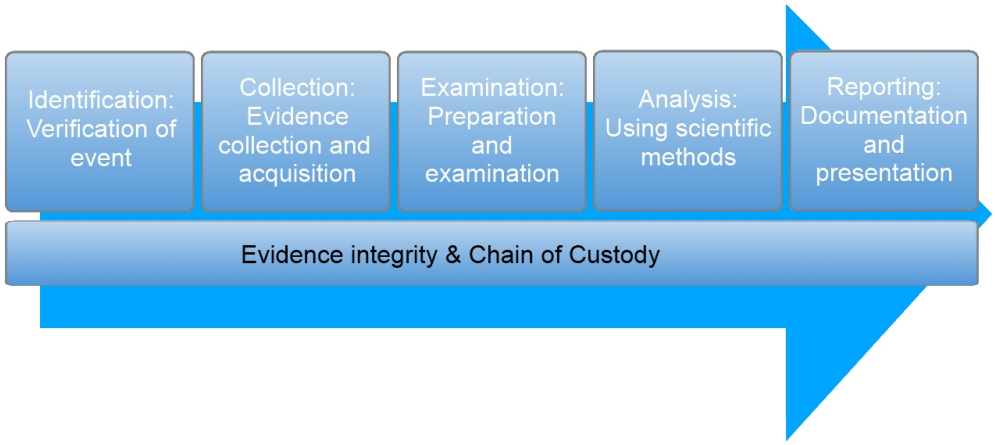

Figure 1: NIST General Forensic Investigation Process. Inspired by [16]

The Digital Forensic Investigation Process

Here we briefly comment on the model from a digital forensic point of view, and use the NIST standard as a reference [16].

Identification

This is the verification of the fact that an incident has occurred, and often the trigger for a forensic investigation. This point is not stated in the NIST standard, but is crucial for digital forensics as one might put processes like intrusion detection/prevention systems (IDS/IPS) in place to detect and mitigate threats, and get a head start on the forensic investigation.

Collection

All identification, collection and labeling of data from all systems that was a part of or monitored the incident. This can for instance include the extraction of all data located on drives, surveillance footage, eye witness statements and other information. In digital forensics one also has to decide whether to seize all hardware or to only copy the needed files from the source.

Examination

The examination process includes processing the data in a way that makes is ready for analysis by both automatic and manual methods. This may include decryption of encrypted materials, restoring damaged files and searching for hidden data.

Analysis

In this section all the data is analyzed for useful information regarding the incident. This might for instance include the creation of time lines based on the information in files and monitoring systems and the identification of participants in the incident like IP addresses, MAC addresses other identification information.

Reporting

The final section gathers all collected information throughout the investigation process and presents it in a report. This will include the entire process from the identification of an event to the detection of all evidence from all sources. It describes tools that were used in the collection, examination and analysis phases, the analysts thoughts surrounding the scenario, time line of the incident, what was taken/changed and why, and can also include recommendations for hardening of the systems software, policies and security measures.

Evidence Integrity

Refers to preservation of evidence in its original form1, and is a crucial part of the entire investigative process. Some of the methods that can be used to obtain this are hashes; which can be used to maintain integrity, and write blockers; which are tools or methods that ensure that the evidence is not altered (willingly or accidentally). The latter can be found in both software and hardware form.

It's worth noting though that in some cases, especially at live forensics the evidence change during acquisition. Volatile data is data that will be lost or changed when a machine shuts down, or as it is used. One can consult with an Order of Volatility (OOV) list to assess what needs to be captured and to capture the most volatile of the interesting data first.

Chain of Custody

Chain of custody (CoC), also referred to as chain of evidence, refers to the documentation of evidence acquisition, control, analysis, disposal (both physical and electronically). The documentation can include paper trails, hashes, event reconstruction images, photography's and so on. To maintain this, we often use lists, notes, reports, time stamps and hash values.

This part of the process must be kept up to date and controlled throughout the investigative process, as it is an important part of the assurance that the evidence is found and kept in a safe way. A slip here might be costly in a court of law, as one could speculate that the evidence is faulty, or question the origin of it. This could in turn render the evidence useless in court [17].

Forensic Readiness

The concept of forensic readiness is not a part of our referenced model, but an important part of the forensic process none the less. The main difference of this task from the previous ones is that the task is performed by the businesses themselves, and not the forensic analysts. By creating or modifying a computer system to assist in the gathering of evidence when this may be needed, one will aid in the identification and collection of events and evidence. A part of establishing forensic readiness also applies to the training of the business workforce, and the creation of appropriate policies.

2.1.2 History

Law Enforcement and Forensics

According to Wikipedia [18] several examples of police-like forces existed in ancient cultures. An example is the prefects of ancient China which existed approximately 1000 years BC. There are also reports that the Chinese started to use fingerprints to identify authors of documents and the creators of sculptures around the year 700 AD, albeit

with no specific classification. The first documented example of medical forensics used is found in the book *Hsi Duan Yu* from 1248, where the distinctions between death by drowning and by strangulation is determined.

In modern times, the first organization into Police offices originate from France in 1667, while George II of England (1737) was the first to pay individuals in London out of tax funds to guard the streets. This shifted the organization of police forces towards governmental control, which set the standard for how police forces are structured today. Although this is the case it is worth to mention that during this time several privately funded police forces also existed in London. During the same time period, in 1784, the first known case where physical matching was used to convict a suspect is documented in England, when a newspaper scrap found on a gun was matched to a paper found in the suspect's pocket.

A lot of new forensic methods were researched at the end of the 19th century, when both methods for poison identification in blood samples and bullet comparison and identification was published. Additionally, the first case of a crime being solved by fingerprint identification was documented in 1880.

By the start of the 20th century the University of Lausanne in Switzerland became one of the first known institutions to offer a forensics curriculum. A couple of years later, in 1908, the Federal Bureau of Investigations (FBI) was founded, who later created their crime laboratory in 1932.

During the 1950's the world saw several new techniques for the capturing of evidence, including the tape lift method (1950) and the invention of the breathalyzer (1954) to test sobriety. During the 1980's the first DNA profiling tests were under way, and by 1986 a crime was solved, as the first one ever, by DNA samples [19]. Even though the admissibility of DNA samples was severely challenged in a 1987 case, it was still widely researched. By 1996 the second National Research Council Committee on Forensic DNA convened, and stated that there should no longer be any reservations against the usage of DNA as admissible evidence in a criminal investigation [20].

Digital Forensics

This section is inspired by the Computer Forensic Timeline [21].

Most of the recorded history of digital forensics originates from the United States, in which the first criminal cases involving computers were registered. In these cases the criminal act mostly related to fraud. By the 1980's it became clear that, in some cases, the available evidence only existed in a digital form.

The eighties spawned a new era for forensics, as tools and techniques had to be created to be able to handle these high-tech cases. Tools like Norton's "Unerase" tool were released, and several law enforcement and academic institutions like SEARC's High-Tech Crime Training Services, the Federal Law Enforcement Training Center and the Association of Certified Fraud Examiners began to train personnel in computer forensics. Programs like FBI's Magnetic Media Program, later to become known as the Computer Analysis and Response Team (CART) were also founded to meet the demands of the digital age. Furthermore, in 1988, the International Association of Computer Investigative Specialists (IACIS) was founded. The IACIS still lives on as a training resource and certification body [22].

By the end of the eighties and the beginning of the nineties there was a surging need

for best practices and specifications concerning digital forensics. Stories like the Cuckoo's Egg [23] can be read for some insight into how unstructured law enforcement in this area was at that time. Many new organizations were founded, and conferences regarding the field of digital forensics were organized. The first international conference on computer evidence was held in 1993, and the International Organization on Computer Evidence (IOCE) was founded in 1995.

In the G8 conference in 1997 it was declared that law enforcement personnel must be trained to handle high-tech crimes, setting a new standard for law enforcement practice. The creation of the principles and best practices used to reach this goal became the job of the IOCE. By the end of the nineties the first INTERPOL Forensic Science Symposium to contain a digital forensics paper was held, further addressing the need for focus on digital forensics.

In the year 2000 the first FBI Regional Computer Forensics Laboratory was created. By 2003 the case load for the CART program had reached 6500 cases containing 782 terabytes of digital evidence. The 2000's also saw advances in forensic software solutions. In 2001, the first version of EnCase was released, and by 2005 the first version of the Forensic ToolKit (FTK) followed suite. As a testimony for the need for advances in digital forensics, it is now estimated that around 80 % of all criminal cases now include some form of digital evidence [24].

2.1.3 Problems of Traditional Law Enforcement in the Digital Age

Several reasons exist for traditional law enforcement methods being less efficient in today's digital society. First of all, evidence is not always available in a physical form. As most traditional methods are based on physically examining evidence, these become unusable, or significantly less efficient. Furthermore, complete and identical copies of digital evidence can be made without leaving any trace. Confiscating a piece of digital evidence will therefore not necessarily mean that the recovered evidence is the only one of its kind. This also means that the evidence will have to be stored differently, as gaining access to the evidence may cause a version of it to be extracted and removed. Evidence can also very easily be deleted when existing in a digital form, without leaving noticeable traces. This will also have implications both for how it is secured and stored.

2.1.4 Digital Forensic Methods

In this chapter we will quickly describe some of the different digital forensic methods which are used today. The sections and the content are inspired by the work done in a GUC course paper [25].

Post-Mortem Analysis

Post-mortem does in this sense mean the analysis of equipment that has been unplugged or shut down. It constitutes the basic method of seizure of evidence performed by most law enforcement to this day. As an example, during a raid several computers are found. To seize these, the police pull the plug on the machines, and thereby freeze some of their current state. This also makes sure that no changes can befall the evidence after its capture, for instance by running processes. The equipment is then brought into a lab for analysis by a forensic analyst.

By capturing evidence in this way, police make sure that the machine is not altered after it is seized, but a lot of information is lost in the process. Volatile information in

accordance to the OOV might be lost, information that could have proved interesting. Open windows, processes, network traffic and data stored in the main memory cannot be retrieved when using this method.

Live Analysis

The opposite of the post-mortem analysis is the live analysis of evidence. In this case, a machine is not shut down upon seizure, but rather kept plugged in and turned on in order to capture or log the current activities of the machine. This can be done by utilizing several tools for live memory capture, network traffic capture and similar techniques. One could also force the machine into hibernation mode, which saves the current state of the machine to a hibernation file. This can be an easy way to assure some live forensic material even if the specific competence in the field is not present.

By performing live forensics on evidence, additional information might be found. There are, however, some downsides to these techniques. For one, the analyst has to perform tasks on the machine that alters its state. In the worst case scenario this could contaminate evidence to the extent that it becomes non-applicable in court, and has to be dismissed.

Data Duplication

Specifically, data duplication is the process of collecting logical evidence from physical evidence. This can for instance be a hard-drive or a digital camera. Data duplication is one of the most used techniques in a digital forensic analyst's arsenal, as it will have to be performed on pretty much all digital data which is going to be analyzed. An analyst will typically refrain from analyzing the original data, as the principle of evidence integrity states that no change should befall the evidence.

Using data duplication the digital forensic analyst is put in a unique position, as multiple identical copies of evidence can be produced. This can be used to test different hypotheses on the evidence without having to worry about harming the evidence, and it can be distributed to other analysts for peer-reviews or cooperation for a more effective investigation. However, in the case of data duplication and distribution, high demands are set for the CoC of the evidence, making sure that no copy is lost or leaked. This could in the worst case scenario damage the investigation and prosecution of a suspect, for instance by deeming the evidence non-admissible.

Data Reduction

As mentioned in the introduction of this report, the increasing amount of digital evidence that is captured in each case makes the analysis hard to handle. By reducing the evidence to be analyzed down from several terabytes to maybe a couple of hundred megabytes the analysis process can be significantly enhanced. There exist several ways to perform data reduction, but blacklisting and whitelisting are the most used methods.

A whitelist is a register of known good files, often by means of their cryptographic hash algorithms, which can then be excluded from the evidence. By doing this, files like known operating system and program files that have not been altered can be excluded, as well as several others. The upside to whitelisting is that only the files known to be common or uninteresting are excluded. The downside is that several other files that are irrelevant for the investigation will remain, making the search set far larger than it needs to be.

Blacklisting is the opposite of whitelisting. Here, known bad files are registered in

the same way as in the whitelist, which can then be used to identify files of interest in digital evidence. An analyst can therefore easily identify files known to be affiliated with criminal behaviour, and single these out from the search set.

The upside to such an analysis is that the interesting files can be singled out very quickly, giving the analyst some clue about what evidence might contain. The downside is that only the registered files are found, and many more files could exist that should be analyzed further.

Because of the inherent strengths and weaknesses of whitelisting and blacklisting, an analyst should use both in succession. By for instance using whitelisting to remove known good files and blacklisting to store known bad files, one is left with what is called gray files. These are files that are not registered, and should be investigated further to be checked out of the case or registered as interesting.

Data Recovery

Often data will be deleted, obfuscated, encrypted or corrupted on the storage medium. As this data can be interesting for an investigator it is important to make it available. The recovery of such evidence can be attempted using several techniques, some of which are presented here.

String searching can be used to search for known words of interest, aliases or other interesting strings in available evidence. An example is the search for passwords in data recovered from main memory.

File carving is a technique for identifying the parts of files and rearranging these to recover complete files. In situations where deleted material has to be recovered, corrupted files have to be fixed, or information in slack space or host-protected areas has to be found, file carving can be applied to attempt the recovery.

Encrypted information may be found during an analysis of evidence, which could prove important to a case. Often the deciphering key to such information is lost or not obtainable, and information has to be decrypted in order to be recovered. Brute forcing and rainbow tables are examples of techniques that can be used in an attempt to find the key for encrypted material, and recover the data.

File Analysis

In file analysis the focus lies in exploring files existing in the evidence. These can for instance be log files like event-, network- or registry-logs and individual files related to applications on the target machine. The analysis of malicious files is often conducted in order to understand the inner workings of malware, which is often of interest for the forensic analyst if the malware is used as a staging point in the criminal activity.

Within the file analysis technique there exist many methods for extracting information. By searching for strings one may specifically find files connected to users, time periods or applications. One can also search for files changed during a given period of time and what was changed. Another focus point can be the analysis of registry files to detect previously opened or executed files, user settings and previously connected hardware.

Network Analysis

The analysis of captured network traffic in the form of network packets is called network analysis. Methods for capturing the traffic are many, and depend on the transfer protocol used and the position of the capturing device. The packets can also be captured for other

purposes, like IDS/IPS. This can in some cases prove to be a viable source for forensic evidence, and can therefore be viewed as a good forensic readiness tool.

After they are captured the packets can be analyzed for interesting information like content, sender and receiver, evidence of network behaviour, emails, communication with other specific nodes on a local or external network and more.

Memory Analysis

Memory analysis is still being researched, and is not, as of yet, an exact science. Its incentive is, however, quite tempting. Passwords and other information which are usually found in an encrypted state is always saved into memory as clear text. Several research projects have been conducted on the matter, including [26, 27]. Tools have also been made available for the capture of main memory [28].

Volatile information like the connections which existed on the computer, processes, passwords and deciphered files can be captured using this type of analysis. Some methods exist to save volatile data, even when applying post-mortem techniques, but these are still in the experimental phases. By for instance freezing down memory modules in liquid nitrogen, the data's persistence on the modules is prolonged, and can be extracted at a later time [29]. However, these techniques are not widespread in use.

2.2 Cloud Computing

In this section we cover the definitions that we use throughout the thesis report, the history of CC, and the general benefits and drawbacks of outsourcing to a CC vendor.

2.2.1 Definitions

One quite broad definition is stated by Lu et. al. is ... *share storage, computation, and services transparently among a mass of users,...* [30]. A more intricate definition is given in a NIST draft from 2011; *Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.* [31]. The definition also includes some essential characteristics and some service and deployment models. This definition model is also used by the Cloud Security Alliance (CSA).

We will use the NIST definition together with its characteristics and models as a basis for this part of our paper. This is because the definition is commonly used in one form or another by almost all papers discussing CC that we have encountered. Furthermore, it is highly detailed and is in our opinion the closest we have to a standard definition of CC. A figure inspired by the NIST CC definition model is given in Figure 2. It shows the essential characteristics, service models and deployment models that forms the foundation of CC.

2.2.2 History

One of the first thoughts regarding a distributed system for services across networks was stated by John McCarthy at MIT Centennial in 1961, where he stated that *"If computers of the kind I have advocated become the computers of the future, then computing may someday be organized as a public utility just as the telephone system is a public utility... The computer utility could become the basis of a new and important industry."* [32], thereby founding the term *Utility computing*.

A couple of years later, in a memo sent from J.C.R Licklider to a group of people

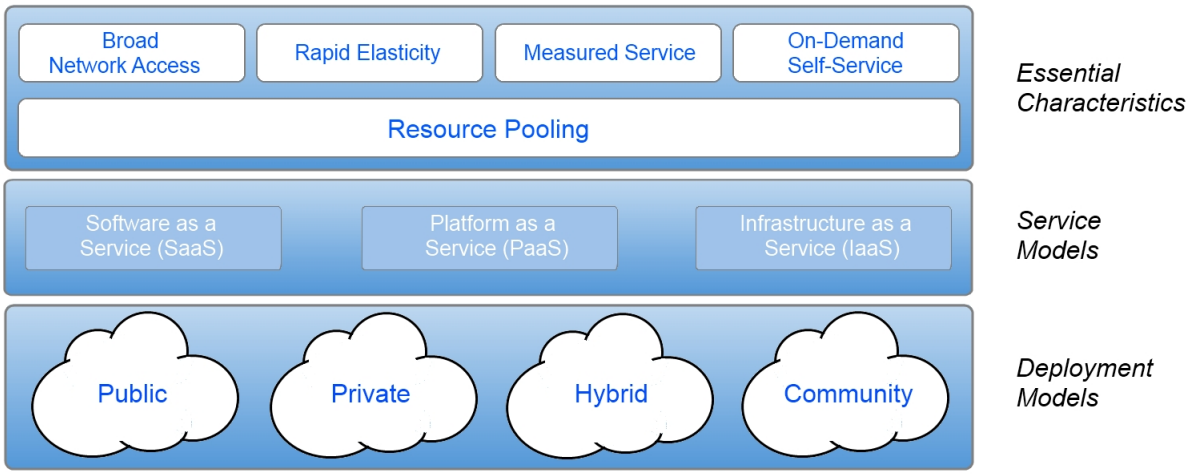

Figure 2: NIST Visual CC definition model. Inspired by [31]

precluding an 1963 ARPA meeting which would ultimately result in the ARPA-net, some more thoughts were brought to light. In the memo Licklider describes a network in which an end node can pull on services from several other systems in order to do its job [33], lining out some of the main thoughts of CC.

The first known use of the term CC on print is found in a Compaq business plan from 1996 called *Internet Solutions Division Strategy for Cloud Computing* [34]. However, after the construction of the Internet as we know it, there have been many synonyms to cloud computing, which explanations more or less differ from CC.

Grid Computing

This is one of these terms which are based on the workings of a power grid. The term was presented by Ian Foster and Carl Kesselman [35], and states that computer power can be made as accessible as the power grid and stated that *The goal is to put full supercomputing capabilities into the hands of anyone who needs it while providing for more efficient use of the supercomputers of tomorrow.* Borrowing some thoughts from *utility computing*, the method of paying based on the amount of computing power used (much like the time-sharing computers used before) is retained.

Distributed Computing

This applies to many of the thoughts stated in grid-computing theory, but differs in the way that it is more like a community than a vendor service. Distributed computing is used mostly as collaboration projects, where a private person, wanting to aid in a project requiring huge amounts of computing power, can configure his or her machine to donate surplus CPU time to working for that project. Projects like Search for Extra-Terrestrial Intelligence(SETI)@home [36], and several other projects including for instance medical research projects have made good use of such services for years.

Virtualization

Virtualization creates the basis for all CC environments, as it allows for the creation of virtual computing nodes existing purely in software. The first examples of resource virtualization was implemented by IBM, when they during the 1960s started to partition their

mainframes in order to run parallel applications [37]. In the early 1990s, the concept of virtualization was introduced for the x86 platform, providing virtualization capabilities at much lower cost than previous examples.

A virtualization platform typically uses a hypervisor, also called a virtual machine monitor, to control the virtual nodes built on top of them. The hypervisors act as an abstraction layer between the hardware and the virtual machines(VMs), translating and sending instructions between these.

There are primarily two virtualization techniques used today: Hardware virtualization and paravirtualization.

Hardware virtualization (also called hardware assisted virtualization or full virtualization) is used to emulate the entirety of a hardware configuration [38]. Unlike its paravirtualized counterparts, hardware virtual machines (HVMs) uses CPU virtualization extensions, and the OSs running on these will not need modified kernels to be virtualization aware. A positive characteristic for hardware virtualization is that any operating system should in theory be able to run on the fully virtualized platform as long as it supports the hardware that is being virtualized. A negative side is overhead, as virtualizing the entire hardware infrastructure is resource-consuming.

Paravirtualization does not need virtualization extensions on host CPUs. Instead, the VMs utilize virtualization aware kernels which are able to generate hypervisor calls. A paravirtualized guest OS must therefore be specifically tailored to the environment in which it is supposed to operate [38]. One strength of using paravirtualization is that the hypervisor and virtualization aware kernel share the complexity of virtualization process, unlike the hardware virtualized systems where the platform must handle all such tasks. Another strength is performance, as paravirtualization lacks the overhead of virtualizing hardware [38]. The drawbacks are the demand for modified kernels, as this constricts possible choices of OSs.

2.2.3 Essential Characteristics

These characteristics are meant to both relate and differentiate CC from other computing methods [31].

Broad Network Access

States that the capabilities should be available over the network, and be accessible through various thick and thin client platforms. This includes, but is not limited to mobile phones, PDAs, laptops, etc. The mobility capabilities of a CC service should be absolute.

Rapid Elasticity

The system should be able to quickly scale up or down based on a customer's needs in storage, computing power, etc. The thought is that additional resources should be available when needed, and in an almost unlimited quantity. Customers will be able to quickly call on additional resources when in need, and pay for them as they are used, and then scale down to a default setting when a normal situation is reinstated.

Measured Service

A CC system should be highly automated in regards to optimization of resources. Monitoring and reporting should be available both from the client and vendor side. This is important for both sides as it eases the controlling of service level agreement (SLA) points, and provides a manual view of the optimization process.

On-Demand Self-Service

A consumer should be able to request (manually or automatically, for instance based on policies) additional resources without having to depend on human interaction. The characteristic is tightly linked to the elasticity principle, and creates the foundation for acquiring resources quickly when in need.

Resource Pooling

Resources should be able to be drawn from all parts of the system, split between the consumers, and pooled where needed. Customers should not be able to pinpoint the origin of the computing resources, but might be able to decide the general location, for instance the country or vendor data centre (DC). This is important if a company has a security policy that blacklists certain countries from managing their data.

2.2.4 Service Models

The models use basic computing terms which are redefined in a CC system. The three terms, Service, Platform and Infrastructure, make up the parts of the *SPI-model* [31]. Figure 3 describes the different service models by how much of the stack is controlled by the CSP, and how much is accessible and configurable for the cloud consumer.

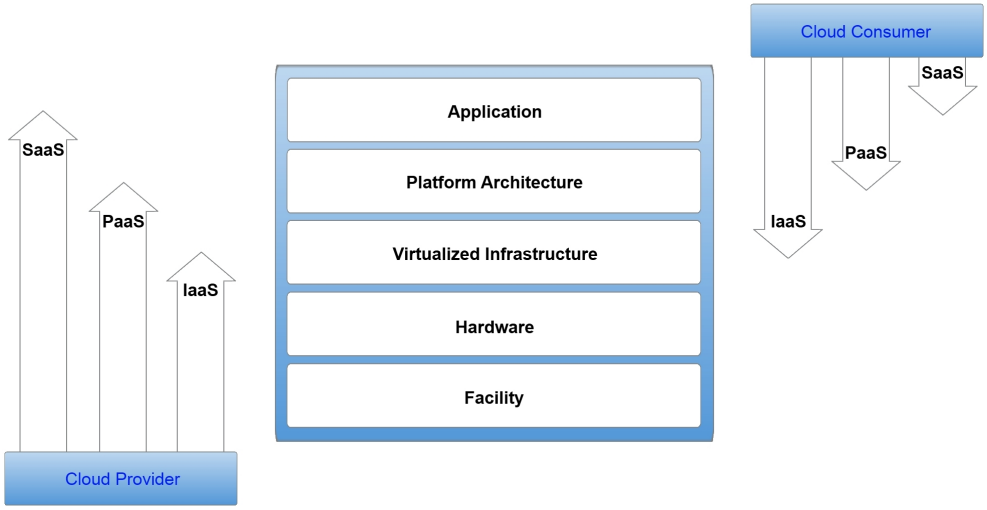

Figure 3: NIST scope and control chart for cloud service models. Inspired by [39]

Software as a Service (SaaS)

SaaS is provided by letting a consumer utilize applications running on the vendors CC system, usually through a web browser. One can say that a SaaS solution enables a user to upload data and use a limited CSP controlled API to handle it in some way. The consumer has no control over where the application is physically run from or the resources it uses (virtual and physical), but might be able to configure the software. One example is a web based e-mail service where the consumer can specify redirection of emails and filters for incoming messages. Another example is web based storage providers like Dropbox.

Platform as a Service (PaaS)

This service model lets a consumer use their own software in a cloud environment which can basically be viewed as a virtualised language environment. The underlying infrastructure and resources are still beyond the consumers' control, and operating systems (OS) are usually prohibited. The model is a good fit for companies that need to deliver a

service of their own making, but do not need the full functionality of a virtual server. A good example of a PaaS solution is Google App Engine (GAE).

Infrastructure as a Service (IaaS)

This gives a consumer control over all system resources but the underlying cloud infrastructure and parts of the networking components. In essence, machines, networks and storage solutions can be virtualised to give users a cloud environment with low-level access rights. The consumer can run their OS of choice (although usually with some restrictions set by the CC vendor) and applications, and can configure some of the networking capabilities. This means that the computer can set up a complete server and/or workstation system with their choice of resources, OS and software, configuring almost all parts of it to their specific needs. A good example of an IaaS solution is Amazon Web Services.

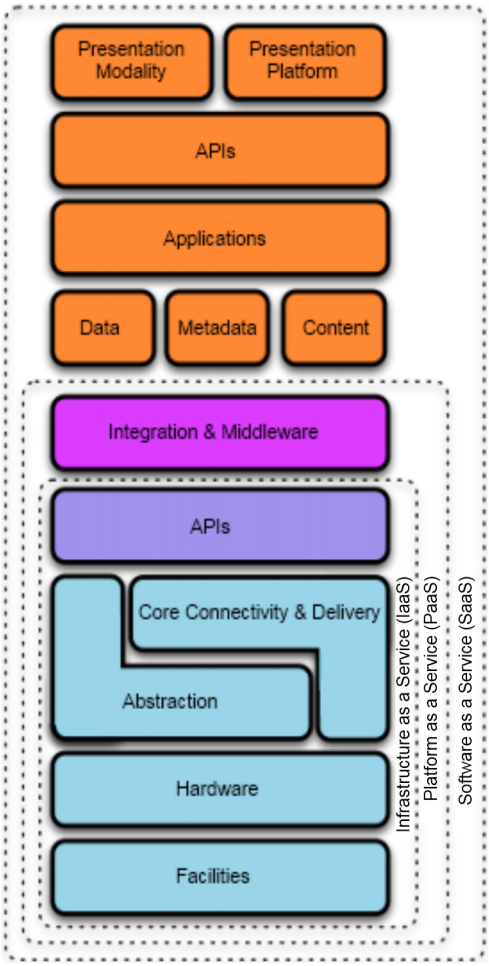

Figure 4: NIST Visual CC stack model [40]

The three parts are building upon each other with the IaaS as a foundation, PaaS on top of that and SaaS at the peak. The model in Figure 4, obtained in the CSA document *Security Guidance for Critical Areas of Focus in Cloud Computing V3.0* [40], shows this relationship and its internal components more clearly. Here, IaaS incorporates the abstraction of resources, and the user's connectivity to these by the use of APIs. PaaS is placed on top of IaaS, and creates an abstraction to the underlying stack by using application development frameworks and middle ware capabilities. SaaS is placed at the summit, and provides the operation environment and user experience.

2.2.5 Deployment Models

These models can be used with all of the service models, and cater to the specific needs of the consumer(s). Figure 5 visualizes the deployment placements.

Public

This infrastructure is used by companies which provide cloud services to the general public or other companies, for instance be CC vendors. According to [39] three broad classes of public clouds exist:

The first class consists of no-cost cloud services supported by advertisement, typically search engines or web mail and for non-commercial use. Non-negotiable terms of use and

service agreements are standard. Security like encrypted connections might be lacking.

The second class consists of fee-based services, free of advertisements. Non-negotiable service agreements are still common, but the security standards are usually better and often configurable by the user. This class is still primarily focused on private consumers.

The third class of public cloud services is also fee-based, but there the SLA is negotiated upon by the organization (client) and the CSP. Cost is often dependent on the amount of or extent of the changes made to a standard SLA for the given agreement plans offered by the CSP. This class focuses specifically on the business market.

Public clouds are multi-tenant systems by design, meaning that the cloud implementation is made to serve several independent clients simultaneously. The same physical resources can therefore serve several users at the same time. The available resources are shared between the users in a flexible way by automatic provisioning based on a client's demands.

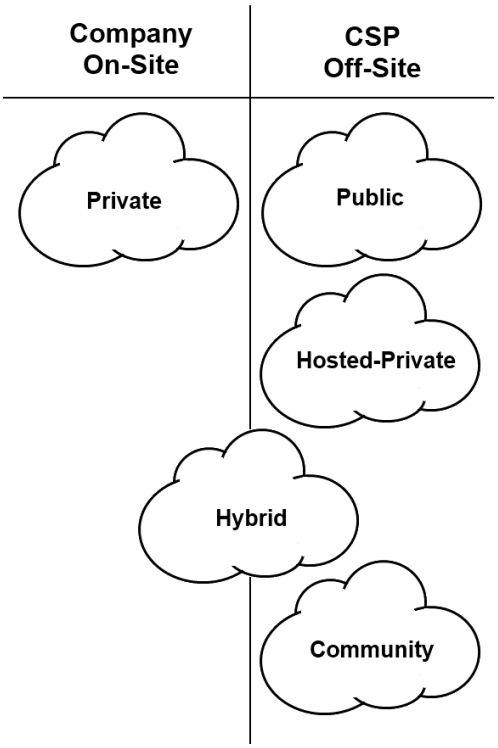

Figure 5: Cloud deployment and placement model

Private

The private model is intended only for serving a single consumer, and is isolated for that purpose. Some may call this type of cloud a single-tenant cloud, although a client can consist of many users. The private cloud will therefore still be able to support multiple tenants of the same client. The infrastructure can be deployed on-site at the customers by themselves, or be outsourced. In the former case we can use open source systems like the *Xen Cloud Platform* [41]. In the latter case we often talk about *hosted*, or *remote private clouds*, as it is maintained outside of the company firewall by a CSP. An example of such a cloud solution is VCODYNE [42].

Hybrid

A hybrid cloud consists of two or more of the other cloud models that are linked together, and share information and/or resources. One can for instance combine a community cloud with a private one to easily share the information obtained in the collaboration with other companies, and use it in the company "isolated" cloud.

Community

This model is meant to be shared between actors which have some kind of collaboration, be it business wise, security wise, etc. The CC system can be run by the consumers, or be outsourced to a vendor in the same way as the other models.

2.2.6 Business Models

This section covers the basic model on which the CSP's sell their cloud services to the consumer market. There are many hybrids, or extended solutions, but these are the foundations for them all.

Pay-as-you-go

The model enables a customer of a cloud service to pay only for the resources that he uses for the amount of time he uses them. CPU clock cycles, memory and storage capacity in addition to network throughput are examples of resources which can be measured. Some resources also have different pricing depending on how they are used. Network traffic could for instance be cheap or free for data transfers within the cloud, while traffic to and from the cloud is more expensive. The upside of this model is the cost for the company, as one will not pay for overhead resources. The downside is that resources cannot be guaranteed, and the upscaling or downscaling of resources may take some time, especially during periods of high resource usage for the entire cloud provider.

Resource reservation

For infrastructure critical to business it may be a good idea to reserve resources so that performance can be guaranteed. This will provide the customer with a stable resource pool which will always be available. The upside is that the resources can be used whenever, and provide a seamless resource consumption in a time need. However, since resources are reserved for the customer, he will be billed for them continuously, even when they are not used.

Per-VM

The per-VM model is quite similar to the resource reservation model, but in this case one or many VM's are reserved. The specifications of the VM's may vary, as will the price, but their reservation will guarantee their availability throughout the reservation period. The upsides and downsides are similar to the resource reservation model, as all reserved resources must be paid for.

2.2.7 Cloud Computing Vendors

Several studies, including papers [43, 44] and technical reports [45] have already covered the comparison of CSPs in detail. This thesis will not seek to repeat such studies, but rather include a short list of some cloud vendors and their main differences, to provide a glance of the current CC landscape.

Xen Cloud Platform

The Xen Cloud Platform(XCP) is an open source server virtualization and CC environment using the Xen hypervisor [46]. The entire spectrum of cloud service models can be built on top of XCP. According to its websites wiki pages [47], Xen is a bare metal hypervisor, also called a virtual machine monitor (VMM). It is the first program that is loaded after the BIOS in a computer system.

Figure 6 presents a graphical representation of XCP. We can here see how the Xen API is controlled by a control interface, which instructs the hypervisor on top of physical hardware. Several physical Xen servers can be arranged into a resource pool to form the cloud. Domain-0 (or dom0) in the figure is a VM that is always started at the boot of the server. It is the only VM that has access to hardware and tools to access the Xen hypervisor. Dom0 therefore acts as a service console which can assign virtual storage space and networks for unprivileged VMs (domUs) present on the server. XCP supports

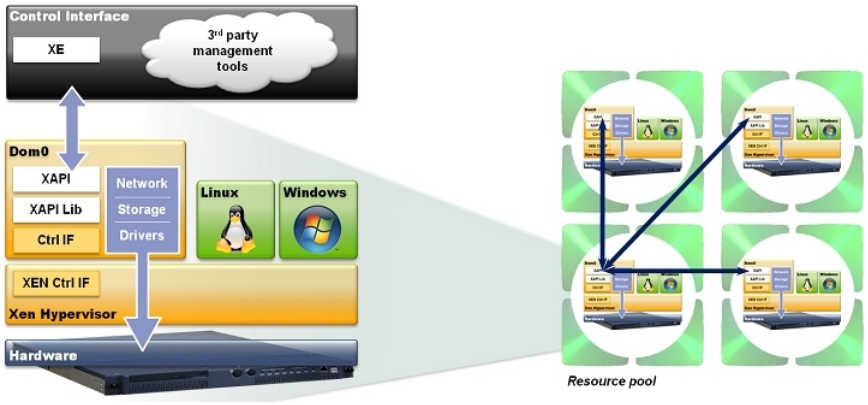

Figure 6: XCP in detail [41]

both HVMs and PVMs, creating a good foundation for wide OS applicability.

CloudStack

CloudStack, which is currently owned by Citrix, is an open source software solution for creating large virtualised networks or cloud implementations [48]. Services spanning all the way from IaaS to SaaS can be built on top of the cloud depending on what the owner requires. One powerful feature of the Cloud Stack software is that it can control several different hypervisors simultaneously. Amongst others, the software supports Oracle VM, KVM and Citrix XenServer.

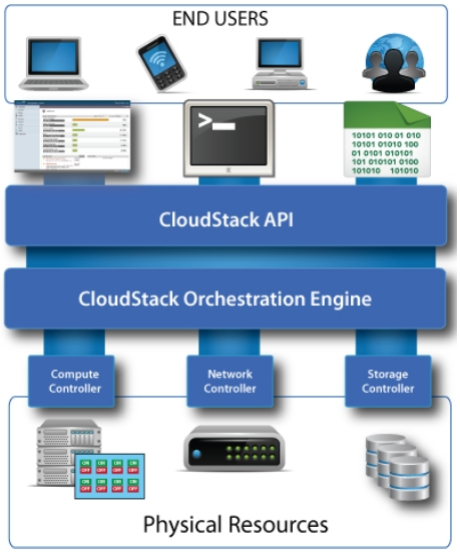

Figure 7: Cloud Stack Model [48]

Figure 7 gives a graphical representation of Cloud Stacks use. An orchestration engine has access to the physical resources of the cloud, which is controlled by the Cloud Stack API. The API is used by both administrative user interfaces and end user self-service portals where these are made available. CloudStack is currently being moved to Apache,

and Citrix Cloud Platform builds on the Apache CloudStack.

Eucalyptus

Eucalyptus (Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems) has both an open-source CC framework and a commercial edition. It provides IaaS private and hybrid solutions implementable on a wide array of computation and storage device combinations [49]. Eucalyptus is created to be as hypervisor independent as possible, but currently uses the Xen hypervisor as a base. Since Eucalyptus has transitioned to a commercial product, some of the features of the open-source distribution have been locked [50], and the solution can therefore not be claimed to be fully open-source.

Figure 8 provides a graphical presentation of an open-source Eucalyptus implementation. The model is inspired by the architecture model found on [51], which consist of several modules: The Cloud Controller (CLC) provides a configuration interface for clusters and instances and translates user input to the Cluster Controller (in the model called CC). The Cluster Controller chooses which compute nodes an instance will run on, and transfer requests to the Node Controllers (NCs). The Walrus controller monitors image buckets and user registration. Internal communication is secured using HTTPS with x509 certificates and SSH.

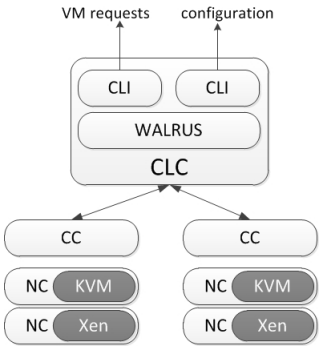

Figure 8: Eucalyptus Deployment Model [51]

Eucalyptus' external interface is based on an Amazon-developed API, which is also used to interact with Amazon EC2 instances. In addition, it supports VMs that run on several of the popular hypervisors, including Xen, KVM and VMware. The high-level system components of Eucalyptus are implemented as stand-alone web services, which use tried APIs, and allow for the use of existing web security features such as WS-Security for communications between the components. Operating systems mainly supported in Eucalyptus are Linux based, but do also support some Windows OSs [52]. Support for Windows instances was added to the enterprise edition of the cloud framework in 2010, but is still lacking in the open-source version [53].

Other Cloud Vendors

Other cloud vendors are:

- Amazon EC2 [54]
- Google Apps [55, 56]

- Microsoft Azure [57]
- Cisco [58]
- Citrix [59]
- IBM [60]
- VMware vSphere [61]
- OpenStack [62]
- STAR [63]

2.3 Cloud Computing Forensics

CC forensics can be split into two categories; Using CC as a forensic tool, thereby aiding the forensic analyst, or performing digital forensics on a crime scene residing in the cloud. This section of the report explains both.

2.3.1 Cloud Computing as a Forensic Tool

As little research has been done concerning how CC can aid a forensic analyst, not much information exists in this field. In [64] Fu et. al. discusses the possibility of using the cloud to deploy network surveillance instances, in order to trace suspicious activities back to their origin. An example of the system's use is to trace users of Tor networks, by dispatching Amazon EC2 based surveillance sentinels at Tor entry- and exit nodes.

Fu's paper suggests that the method has a high probability of catching users of Tor networks using the method, and claims that it is both cost and time effective. However, the study does not cover how the evidence is captured and concatenated from the EC2 sentinels, and to what extent this can be made to comply with evidence integrity and CoC. There are also great ethical implications concerning such methods, as law enforcement cannot be an enabler for illegal activities. Even though the sentinels would here be dispatched to detect criminals, these would also act as gateways and proxies for illegal content. Such enabling of criminal activities cannot be justified, thereby hampering the methods use.

One example of cloud-like systems for law-enforcement that exist was covered in Section 1.4. ØKOKRIM has currently implemented a solution based on the AD Lab [11], which uses resource pooling in a cloud-like fashion. Since ØKOKRIM regularly handles cases with years of digital content, the amount of evidence present in each case is significant. By implementing such a system, ØKOKRIM improved their ability to analyze digital evidence significantly.

2.3.2 Forensic Investigations in the Cloud

Performing investigations in the cloud has seen more attention during the past few years, much due to the increasing interest for businesses to make use of cloud resources. Most research has focused on defining the challenges of performing forensics in CC environments (2.3.2), but some methods and solutions to the problems have been proposed in 2.3.2, 2.3.2, 2.3.2 and 2.3.2.

Challenges

During the work done in [14], a chapter was produced called "Challenges for Forensics in Cloud Computing systems". This section is based on this previous work.

Several papers have currently commented on the challenges that face forensics when it comes to CC, including [65, 66, 64, 67]. In [68] it is stated that the cloud might become a new frontier for digital criminals as the technology is embraced by companies worldwide. It is therefore of pressing concern to create ways of conducting forensics in such environments.

One paper by Wolthusen [69] chose to present six distinct challenges, and further discuss how these must be handled in order to obtain the same forensic capabilities as one would with standalone systems. Due to the well organized presentation of these challenges in comparison to similar papers on the matter, we chose to quickly present these in this section.

Discovery of Computational Structure

Tracking an attacker or node of interest through the use of IP-addresses and ISPs is not an easy task when assessing cloud systems. Such traces will most likely lead to a CSP, and even if a specific VM can be identified by the CSP the placement of the VM and its storage is not easily recovered. Removing complete servers for analysis may also be problematic, as the data of innocent clients can be lost in the process. In addition, the imaging of cloud servers can be close to unfeasible, as several terabytes of data might reside therein.

Wolthusen states that the placement of client data will probably be held by the CSP as a way to improve the performance of the cloud. Such information can therefore be made available for the analyst. However, as the data may reside in many jurisdictions, especially when assessing public clouds, the availability of data will vary greatly. It is further explained that mechanisms for describing what should be captured, and techniques for capturing cloud-server states and volatile content is needed. An understanding of process' dependencies in the cloud is also required.

Attribution of Data

The need to gather provenance in the cloud is also important if a suspect is to be linked to a piece of data. CSPs have different ways of linking their clients to data, and logs containing communication entries will therefore have to be collected to strengthen evidence. Clear methods for creating and logging provenance in the cloud are therefore needed.

Semantic Integrity

When performing forensic analysis of a system, one will often have to create a snapshot of the seized equipment so that this can be loaded into a forensic environment. The snapshot must be safeguarded for integrity, and experiments can be conducted and repeated to verify findings and evidence. As one may not be able to seize cloud infrastructure for this purpose, the original system of such snapshots will change after imaging, damaging the integrity of the system.

Wolthusen identifies that the capture of a systems semantic information, for instance data dictionaries, will be needed in order to preserve evidence integrity in such situations.

Stability of Evidence

Preserving the long term stability of evidence is also of concern. Since one cannot depend on the seizure of physical equipment from the cloud as a viable possibility, Wolthusen claims it to be of vital importance that as much of the extensional information as possible regarding the captured data is collected. Methods for describing such semantic

information must also be created.

Since many cloud services are closed-source, the data formats of information gathered may also be proprietary, and therefore highly problematic. There also exist problems related to the streaming of data, which can be more prominent in distributed systems. Under these circumstances it is crucial to collect both the semantics of the given data and the state of the involved nodes, as thes will be the keys to understanding the nature of the collected evidence.

Presentation and Visualization

Presenting evidence captured from the cloud may be a daunting task. One reason is that the confidence such evidence hold might be limited, based on the methods used for its capture. Until verified methods are created, this problem will remain. Secondly, evidence captured from the cloud can consist of complex data sets, which are not easily visualized and examined.

Techniques for presenting evidence recovered from the cloud must be constructed, in order to give analysts a clear view of the evidence. This will in turn enable analysts to generate hypotheses, prepare experiments based on these, and fill in possible gaps in time lines and data semantics.

Cross-Jurisdictional Aspects

CC adds to the already existing problems of cross-jurisdictional investigations which became prominent when the Internet became a widespread commodity. International cooperative agencies like Interpol have already established methods for conducting investigations across jurisdictions, but these techniques will have to be refined to address the issues of CC.

Forensic Readiness Techniques

Forensic readiness is a concept for the preparatory steps taken to ensure that evidence is collected in the case of crime. These can for instance be the implementation on client and server side systems, set up to collect data continuously or based on different precursors. An example of such a system is an IDS/IPS, which based on a set of parameters can identify an intrusion as it is happening, and start collecting relevant information.

As these techniques require a certain control over the underlying architecture of a system, implementing them on the client side can be hard if the client uses SaaS or PaaS based cloud services. One solution can then be that the services are implemented as default on the server side, or by the choice of the customer when possible. In the cases where IaaS is used, the client will have full control over the OS, and thus have more freedom as to what can be installed and monitored.

Forensic readiness techniques may be used both as such during live forensics to capture data, or even during post-mortem analysis as their logs are uncovered. One possible approach used to plan and implement forensic readiness can be found in the book [17], and this approach is listed underneath. The book can be consulted for a more thorough introduction to the approach.

- Define the business scenarios that require digital evidence
- Identify available sources and different types of potential evidence
- Determine the evidence collection requirements

- Establish a capability for securely gathering and storing evidence in a manner that will make it legally admissible
- Establish a policy for the secure storage and handling of potential evidence and ensure it is properly and regularly tested
- Ensure the monitoring of systems and networks is targeted to both detect and deter major incidents
- Specify the circumstances in which an incident should be escalated to a full formal investigation
- Train all relevant staff in incident awareness
- Document an evidence-based case describing an incident and the impact
- Ensure there is a legal review of the procedures deployed

Some examples of forensic readiness methods will now be quickly introduced. More examples can be found in [14].

Virtual Introspection for Xen is a method proposed in [70] for monitoring the status of Xen VMs. This is done through a Virtual Machine Monitor (VMM) or another virtual node. Virtual Introspection for Xen (VIX) focuses specifically on digital forensics as it can monitor the memory of a VM. The method uses UNIX functions to pause the VM, and extract the needed data as read-only, before the VM is resumed.

The method can be put to great use during a live forensic investigation, as VIX may assist an analyst with the capture of volatile content (ref the OOV). As the VM is frozen and information extracted as read-only, one should also be able to circumvent the greatest downsides of live forensics, namely the changing of the target system as evidence is gathered. A disadvantage with this system is primarily that the usage is limited to Xen systems, but also that it can only be implemented on the server side. In some cases, one would also be able to use the system from a client side system, as a CSP could allow its users to assign one of the VMs as a VMM. In other cases, where a client supports virtualization of their own, the system can be implemented from scratch by a customer as long as it operates in an IaaS environment.

Secure Provenance as proposed by [30] is a technique based on the principle of provenance [71], which means the record of ownership. A monitoring system for provenance needs to be active at the CSP in order for the system to function. Furthermore, it relies on two requirements; Unforgeability and conditional privacy preservation. Unforgeability states that one should not be able to forge new provenance records, or change existing and valid ones without detection. Conditional privacy preservation states that only authorized personnel can recover the true identity of the records owner. Cryptography is utilized to secure the anonymous provenance and logs of modifications done to content. The keys and signatures used in the cryptographic algorithms are generated by using a user identifier, and the records of all user activity will stay anonymized to the user. The provenance will only be made available to authorized entities when needed. If the monitoring system is intact, and can be salvaged during file system analysis, it could prove a valuable post-mortem artifact.

Post-Mortem Forensic Techniques

Post-mortem forensic methods are the ones most frequently applied in forensic investigations, but not all of these are possible to utilize while investigating cloud systems. An introduction to post-mortem forensics was given in 2.1.4. Since the cloud backbone consists of a resource pool of servers, finding the interesting devices for capture may prove impossible. Furthermore, the recovery of any single server in a resource pool could prove to be of no value at all, as the distribution of files and resources means that single devices can only present partial evidence at best. Using Google as an example, which currently secures their Google Apps file system by splitting data into chunks and striping these over several servers [72], no complete files would be recoverable without Google's cooperation. Basic file-system analysis would in this case not have any effect.

More examples of post-mortem techniques being used to analyze cloud environments, both from the client and server side, can be found in [14].

Live-Forensic Techniques

An introduction to live-forensics was given in Section 2.1.4. As such techniques are novelties, we have not found any specific techniques made for live-forensics in the cloud. A classification of existing live-forensic techniques is found in [73], and an overview of how these can be used in a cloud setting is given in [14]. We will quickly cover some of these techniques here.

The collection of a system's date and time information is usually gathered from the BIOS of a computer system. This should therefore also be possible from cloud backbone nodes, and possible time servers that these would be synchronized with. Current network connections would be harder to collect, as it is reasonable to believe that clients of a cloud service are protected by privacy policies. It is therefore unlikely that the network connections of individual clients or VMs are continuously logged. Although this is the case, firewalls related to the cloud data centres are possible targets for detecting network connections that for instance are blocked due to policy.

In the cases of recovering running processes and system memory dumps from cloud systems, the possibilities may be slim. Unless the CSP closely monitors the activities of each VM in their cloud, which might be suspect behaviour from the vendor side, such information will not be available. Detecting open files is similarly difficult, as direct access to the VM should be prohibited for others than the client owning the instance. Detecting the currently logged on users could at the other hand be an easier task, as an essential characteristic of the cloud is measured services. It should therefore always exist detailed logs of when a user utilized the cloud services.

Other Forensic Techniques

Other forensic techniques remain in the arsenal of digital analysts, but their application to a cloud environment are currently limited [14]. Network traffic analysis of live network activity is made difficult by the sheer number of network nodes active in the cloud, making the task of capturing all network packet activity unfeasible. Application Programming Interface (API) analysis, which works by logging the submitted API calls sent to a CC system, could become an interesting method as such are developed. However, there currently exist no distinct methods for this.

During the work performed in [14], a taxonomy of the current state-of-the-art of CC forensics was created. Picture 9 explains the placement of the different forensic tech-

niques within the different forensic disciplines, and if these are applicable to either the client or server side of a cloud system.

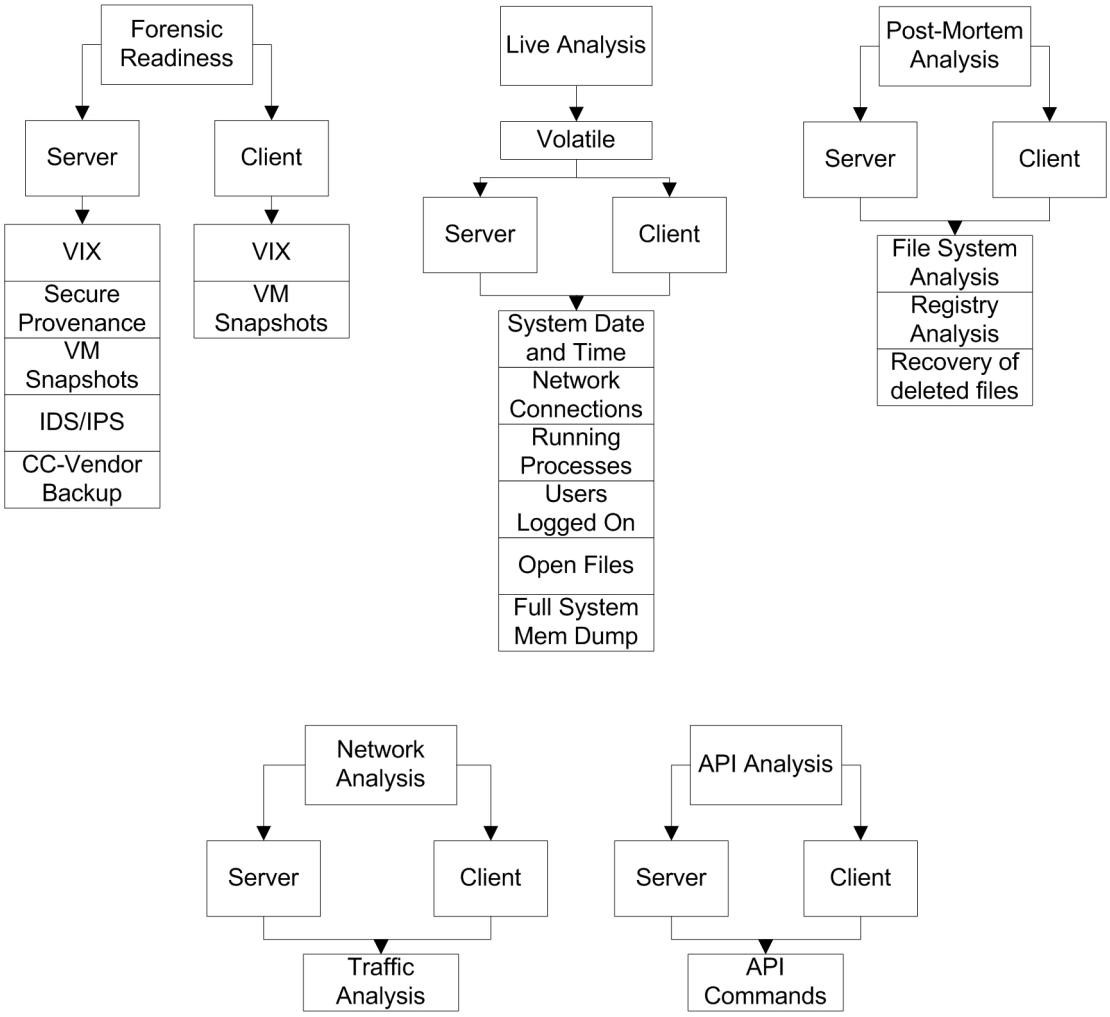

Figure 9: Taxonomy of Forensic Techniques in Cloud Computing [14]

3 Implementing Forensic Applications in the Cloud

In this chapter the benefits and drawbacks of migrating governmental applications to a CC environment will be explored. The CC deployment models will be discussed, and their strengths and weaknesses, both compared to traditional server centric systems and the different deployment models themselves, shall be noted and weighed.

The goal of the chapter is to discuss the pros and cons of cloud migration for government initiatives, especially in the context of law enforcement and digital forensics. Lists of benefits and drawbacks are created based on literature and our subjective opinions. It is therefore important to state that we in no way claim this to be a complete list. The benefits and drawbacks are then rated either ”low”, ”medium” or ”high” based on their impact level in a positive and negative way towards its category. The proposed impact level is also a product of the authors' subjective opinions, backed up by the literature. The product of this chapter provides a foundation on which the different cloud deployment models can be ranked from a governmental and forensic viewpoint.

Community clouds will not be covered in this chapter, as they are basically private clouds shared by cooperating entities. It will therefore have many of the same benefits and drawbacks as a private cloud implementation, with some of the issues of public clouds. We see the potential benefits of using a community cloud for cooperation in law enforcement, but leave this scenario for others to explore. It can also be noted that both private and hosted-private cloud solutions will be covered in this chapter, as they were defined earlier in this thesis. All the deployment models discussed in this chapter are used based on the definitions found in Section 2.2.5.

Furthermore, a deployment model called *conventional* will be used. This represents the server-client on-premises computer systems used in most cases today. By also focusing on the benefits and drawbacks of such implementations we can compare their usage with its CC counterparts.

At the start of the benefits and drawbacks sections a table is presented, which quickly sums up the benefits or drawbacks explored during the chapter. The subsequent sections then discuss these at length.

The benefits (3.2) and drawbacks (3.3) sections are divided into the three deployment models, private, public and hybrid clouds, as well as a sub-section for conventional systems. It is also worth noticing that both ordinary private clouds and hosted-private clouds will be covered. Short definitions of the different models are given in the text.

The deployment model sections are further explored using a set of categories which are inspired by a paper by Paquette [74] covering the security risks associated with governmental use of CC. The categories are access, availability, infrastructure and integrity, hereby referred to as AAII.

With access we mean personnel or systems that can gain insight or rights to modify information. Keeping outsiders out of an on-premise network solution is a well known and containable problem, but in a cloud environment it becomes problematic.

Availability is concerned with the consistency of performance and continued connectivity towards purchased services from the client's viewpoint. Making sure that out-

sourced services will stay available to the company in the same way as on-premise hosted services is an important factor when assessing cloud services.

Infrastructure represents the foundation on which the services are implemented and delivered, for instance how services are delivered to the client nodes. In CC, the infrastructure should reinforce the characteristics of scalability and flexibility, but problems can arise for clients when this infrastructure is changed or of a proprietary design.

Integrity is an umbrella concept for the parts of the computer environment handling data accuracy, validity, security and durability. The efficiency and performance of the services are also included in the concept of integrity [74].

As the audience of this thesis will most likely be prospective or professionals in the field of information security, it seems natural to use the CIA model (confidentiality, integrity and availability) for such a study instead of AAII. However, we believe that the choice of a somewhat non-traditional model will both invigorate the discussion, and inspire us and the reader to assess our work from a different point-of-view.

In this study, we are different by purpose, but all of the security aspects of the CIA model is still present in the AAII model presented by Paquette [74]. It is also possible to map the AAII model to the CIA model. As an example we can use access in the AAII model, which will cover all aspects found in the confidentiality part of the CIA model.

The chapter concludes with the results in Section 3.5 and a discussion in Section 3.5.5.

3.1 Hypotheses

In this section we describe our initial projection for the results of this chapter. These will not be presented as specific expectations, but rather as an outline of our general assumptions regarding how the different deployment models will fare in the study.

Government systems may have to process and store sensitive information concerning both individuals and companies. Especially in the field of law enforcement, systems will often contain information than could do great harm if leaked. Such implications have to be considered when assessing the benefits and drawbacks of CC in this setting.

We predict that the private cloud deployment model will perform well in this study. This is because we believe that cloud solutions which can be implemented using strong security measures are highly beneficial in our setting. We also believe that the hybrid solution should do well for the same reason. The results for the hybrid cloud can also be strengthened by use public resources. Furthermore, we predict that the purely public cloud solutions will not do well in our study. Storing or processing the type of data mentioned above at a third party seems to have several risks associated to it. For this reason we also believe that the hosted-private deployment model will be at a disadvantage in this study. Finally, we believe that the conventional model will produce strong results, as it has proven a predictable environment for the types of systems we assess.

3.2 Benefits

There are many known benefits following the transition to a CC environment. This section presents the ones known to us.

3.2.1 Benefits Overview

Table 1 contains a list of the collected benefits. The columns of the table consist of *category*; specifying for which of the four classes (access, availability, infrastructure and integrity) the benefit applies, *environment*; which specifies the impacted deployment model(s), *benefits*, which quickly outlines the main points of the benefit, *compared with*; stating the deployment models that are not impacted, *possible results*; which quickly notes the potential positive outcome, and *impact*; stating the assessed force of which the benefit will have on the implementation.

Table 1: Benefits overview

Category	Environment	Benefits	Compared with	Possible Results	Impact
Access	Hosted-Private, Private, Conventional	Only one client, smaller attack surface	Public, Hybrid	Less chance of existing vulnerabilities	High
Access	Private, Conventional	Smallest attack surface	Hosted-Private	Less chance of existing vulnerabilities	Low
Access	Private, Conventional	Direct access to hardware	Hosted-Private, Public, Hybrid	Quicker reaction times / mitigation, and controls maintenance	Low
Access	Public	All cloud resources can be reached from anywhere	Conventional, Hosted-Private, Private, Hybrid	Eases cooperation and the work situation for external personnel	Medium
Access	Hybrid	Applications can be distributed to the public or private part of the environment based on criticality	Conventional, Hosted-Private, Private, Public	Flexible accessibility and security	Medium
Availability	Hosted-Private, Private, Conventional	Other clients cannot deplete resources	Public, Hybrid	Not affected by attacks on other clients	Medium
Availability	Private, Conventional	Failure of ISP or CSP is less critical	Hosted-Private, Public, Hybrid	Operational during failure of ISP and CSP	Medium
Availability	Hosted-Private, Private, Public, Hybrid	Applications are not hosted on specific servers	Conventional	The failure of a server will not shut down services	Medium
Availability	Hybrid	Redundancy by implementing both an on- and an off-premise cloud	Conventional, Hosted-Private, Private, Public	Still limited access to resources if one cloud solution goes down	High
Availability	Public	No reserved hardware means that upscaling and downscaling can be performed seamlessly	Conventional, Hosted-Private, Private, Hybrid	Time spent assessing and acquiring hardware is shortened. Applications can be rolled out quicker	High

Continued on next page

Table 1: Benefits overview – continued from previous page

Category	Environment	Benefits	Compared with	Possible Results	Impact
Infrastructure	Conventional, Private	Direct hardware access means that import and export of data can be done more effectively	Hosted-Private, Public, Hybrid	Imports are not charged, and will not be limited by bandwidth	Low
Infrastructure	Public	No reserved hardware means that upscaling and downscaling can be performed seamlessly	Conventional, Hosted-Private, Private, Hybrid	The system will cater to elevations in users and resources effectively. Cost goes down	High
Infrastructure	Hosted-Private, Public	Hardware becomes irrelevant, handled by the CSP	Conventional, Private, Hybrid	Lowers cost (electricity, personnel, real estate), focus on main tasks	Medium
Infrastructure	Hosted-Private, Private, Public, Hybrid	VM-centric, can be frozen, moved and backed up more easily	Conventional	Abilities to capture and move VMs with the current state raises availability	High
Infrastructure	Hybrid	Possible to supplement with public cloud resources when needed	Conventional, Private, Hosted-Private	Temporarily boost performance of private resources	Medium
Integrity	Hosted-Private, Public, Hybrid	Backup on several hot-sites secures the integrity of data	Conventional, Private	Local disasters or attacks will not destroy data	Low

3.2.2 Private Cloud

Given that private clouds are shielded from any other cloud services or clients, some major benefits can be noted.

Access

By design, a private cloud solution is still multi-tenant, but only serves the tenants of one client compared to the servers of a public cloud infrastructure. There is therefore a slimmer chance that unauthorized personnel would gain access to data stored by the client due to information leakage, than in a public cloud. The attack surface of the system becomes smaller, which lowers the chance of vulnerabilities being present. In the case of digital forensics there will always exist data that cannot be made available to the public or unauthorized personnel. The confidentiality of such information is vital to a forensic computer system, so single-tenant systems are highly coveted. The impact of this benefit is therefore assessed to be high.

Private clouds hosted on a client's premises (on-premise) often has the smallest attack surface, as it cannot be caught in the cross-fire of attacks against CSPs or other clients using a CSP-hosted service. Such attacks can be of concern, as they might result in the leakage of information. For a forensic application it would therefore be beneficial that critical parts of the system are hosted on-premise. The benefit is assessed as low, as it gives the impacted systems a slight security benefit over hosted systems. This benefit also affects conventional computer systems.

A feature that can be both a benefit and a drawback for the private cloud hosted on-premise and the conventional computer systems alike is direct access to hardware. If the physical servers fail, one may be able to react quicker than a CSP which could have a complete data centre to bring back on line. A client can also update and upgrade the servers at their own accord. To view the drawbacks see Section 3.3.2.

While direct access to the hardware can be beneficial, for instance to ease compliance issues, it should not be of any significant value for digital forensics. It can even be a bad thing to have full control over hardware, as forensic systems of the past have had a tendency to become extremely outdated before they are finally revolutionized. The revolution often comes as a response after the damage has already been done. This was the case when about a fifth of the Norwegian police forces computers were infected with the Conficker virus in 2009 [75].

Using a system that forces the developers to keep up with the rest of the digital world will create a strong motivation to build flexible and effective software. We will therefore rate this benefit as one of low value for digital forensics.

Availability

During DDoS attacks against clients in a public cloud a result could be that those clients either demand excess resources, bogging down the cloud, or crash the servers on which they operate, thereby shutting down the services offered from that node all together. This is a non-issue in a private cloud. As resources are specifically reserved for one client, the client should not have to worry about others depleting their resources. Consequently, this also becomes a benefit that conventional systems have over off-premise clouds.

It should also be noted that if the cloud is hosted on-premise at the client, single points of failure like the ISP and CSP become less of an issue, as only personnel working off-premise will be affected. Additionally, since the cloud is basically built up of identical

servers placed in a resource pool, the failure of individual servers will not have a large impact on the availability of the system.

Securing the continued operation of computers and applications is as important for forensic systems as it is for commercial ones. Since forensic analysis can often be a timecritical mission, for instance to uncover identities and positions in order to eliminate further criminal activities, the availability of the systems is important. We would therefore rank both the preceding benefits as of medium impact.

Services and applications could also benefit from the fact that they are no longer only hosted on single servers, but use a combined pool of resources. If the service is hosted on one of the servers in the pool that fails, it takes little to no effort to move it to another server. Meanwhile, the failed hardware can be replaced with identical equipment and added to the pool to replenish it.

This flexibility means that cloud-based forensic systems have the potential for better redundancy, wich could ensure that the systems remain operational through failures that would cripple their conventional counterparts. This benefit is also ranked with a medium impact level.

Infrastructure

The infrastructural characteristics of a cloud environment constitute its greatest advantages over conventional on-premise server parks. However, some of these benefits might not be applicable for on-premise private cloud implementations in particular.

As a cloud client, hardware should become irrelevant, as the CSP handles the upgrading and upscaling of server parks. Private cloud implementations will not benefit from this. However, a hosted-private cloud will, as its hardware is hosted off-premise. Doing this lowers the cost of electricity and possibly also expenses for personnel, as a part of the workforce might no longer be needed. Storage facilities for equipment could also be made unnecessary, resulting in decreased expences. Removing many of the administrative tasks would also mean that personnel can focus more on specialized areas like security or development.

The migration to CC may release more resources that can in turn be used for forensic analysis. This could directly impact the performance of the forensic analysts, but relies on close cooperation with the CSP. The benefit is therefore ranked as medium impact.

The infrastructure of a CC environment is based on virtualization of servers and clients. As such, the cloud may feature backup-systems that capture both the storage and state of a VM, even freezing the main memory of it. This means that if errors occur on the system, the client's machines can be frozen and resumed at a later stage. This allows for machines to remain quite persistent even during complete power failures. It also allows the VMs to be moved, if necessary, to other data centres, and recover to the same state.

In some cases a VM can be continuously backed up at a hot-site. In this way, if the main data centre hosting the VMs goes down, a handover can be performed to the other data centre, often without any observable implications for the client. An example is Sungard Recover2Cloud, a cloud based recovery system [76].

The power of virtualization is already extensively implemented in some conventional computer systems. In law enforcement, digital forensics has particular use for the method, as it allows them to create environments in which they can recreate scenarios of one or several nodes and networks and analyze data flow. Most other systems are implemented

as hardware though, and clients and servers alike mostly exist as individual machines.

Virtualization can both serve as a good method for redundancy, and bring a new level of flexibility to governmental computing systems. For forensic workstations this could mean being able to create and deploy new instances of a forensic analysis VM when needed, thus not having to wait for time in specialized labs or on certain workstations. We believe this is highly beneficial, and the impact rating is therefore set to high.

By having direct access to hardware, another benefit emerges as this allows one to import or export data more efficiently, not limited by bandwidth. Bandwidth can for instance be affected by the activity level of the ISP and CSP. Moving data into or out from a cloud is ordinarily charged with an extra fee reflecting the amount of data, something that on-premise systems are not affected by. The benefit has a medium impact level compared to other benefits.

Integrity

The integrity of data in a cloud can be substantially safeguarded by backup solutions. Since a cloud provider often has several server centres, data can be backed up on them all. This means that the data is safe from local disasters, and that possible cold- or hot-sites can be easily created.

Making sure that data is not lost during system errors is always important. Creating off-premise backups will be less of an issue for a CSP, as the concept is based on the cooperation of several data centres, unlike conventional implementations. However, contingency and backup plans are always a main concern, and will likely be less of a benefit than other actions. We therefore rank this benefit as one of low impact.

3.2.3 Public Cloud

Operating in a public cloud means that a client can connect from any compatible device connected to the Internet. This has many benefits.

Access

The ability to reach the cloud from anywhere means that a company is able to grant access more readily to external or cooperating personnel. As the security of the connection is handled by the CSP, the only thing the client has to do is to order a new user with an appropriate set of access rights.

This may prove very beneficial to a law enforcement agency, as users can access resources from wherever. What limits the beneficial impact is the type of data that can be accessed through such channels. Even though the connection may be secure, the integrity of the endpoint system will also have to be assessed. Still, access through a public cloud will be beneficial as a channel for non-confidential information, and the impact is therefore assessed to medium.

Availability

Public clouds with distributed and redundant storage have a great advantage in availability, as entire data centres can be taken off-line without disrupting the cloud services. This means that the availability of a cloud service is theoretically 100 %. Less downtime will always benefit the client, as work is interrupted less frequently. This benefit was rated as of medium criticality in 3.2.2.

Rolling out new applications and services should also be quicker, as new capacity and extended resources can be unlocked seamlessly. Not having to wait for hardware

when implementing new features means less time spent on acquiring equipment, and that more time can be spent on development and performance, resulting in faster service availability. It is also highly cost effective, as resources not used can be discarded instantly.

The up- and downscaling characteristics of public CC is one of the main incentives of implementing such a solution, and one that might greatly benefit law enforcement if implemented correctly. For instance, upscaling a task force becomes near effortless, as new users and instances of software can be enabled instantly. The technique has a lot of potential, and is therefore rated with high impact. It also has benefits from an infrastructure viewpoint, as reviewed in Section 3.2.3.

Infrastructure

The benefits noted in 3.2.2 are relevant also for public CC, but compared to an on-premise private cloud, the public cloud is much more flexible. Not being bound by reservations, the client has all the power of the public cloud available at an instant, if need be. The resources made available for a client can be expanded and contracted based on the need for resources, and a client will only pay for the resources that are used at any given time.

This reinforces the foundation of a flexible and distributed system which enables it to stay effective in times of high elevated usage, in regards to the amount of users and the resource demands of the tasks. For these reasons, it is also rated as high from an infrastructure point of view.

A private cloud hosted off-premise at a CSP will be able to benefit somewhat from the up- and downscaling characteristics of cloud infrastructures, but since the hardware always have to be isolated to the client, one have to expect some time delays between ordering the resources and having these made available for the client.

Integrity

The benefits of backups noted in Section 3.2.2 are also relevant for public CC, and will even be more potent in this setting. This is because of the power of distribution, as public CC resources do not have to be isolated for one customer. By avoiding reservation of resources, and letting the CSP organize the placement of data, the price for the service will be lower. In addition, the data will be safer as it can be duplicated several locations without restrictions.

As noted in Section 3.2.2, saving the state of VMs during an outage makes it recoverable in the same state. This means that the integrity of the work done on the machine is kept as well.

3.2.4 Hybrid Cloud

By combining the two cloud deployment models, the benefits of both are within reach.

Access

Using both cloud solutions, the client in a hybrid cloud can host separate services in the two cloud environments as the confidentiality of the data allows it. Access can be granted to both on-premise and external personnel, gaining access to the applications hosted on the given cloud. An example is that non-confidential information and applications can be accessed both from the workplace and outside through web access, while resources demanding a more secure environment are accessible only from within the company

walls.

From a governmental and law enforcement viewpoint, the hybrid cloud would be beneficial for case work, as personnel in the field could access non-confidential information mostly in the same way as it would be done at work. Similar solutions are often implemented for conventional systems using secure remote access like VPN, but the cloud can connect using any client and browser. Whatever solution is to be selected, the important part is to make the connection secure.

The hybrid setup shows a lot of potential for allowing users access to different resources depending on their endpoint security and location, something that warrants a medium impact level.

Availability

Like the other cloud deployment models mentioned, the hybrid cloud will also benefit from the distributed characteristics of a cloud implementation, meaning that applications are distributed to the resource pool, and not to the specific servers. This was earlier rated as a benefit of medium impact.

While private or public CC both have their single points of failure, a hybrid cloud theoretically have redundancy. Operating a cloud both on-premise and at the CSP means that limited availability will still be possible if one of these fail. This will of course depend on the services that are hosted in the cloud solutions. For instance, applications that are usually run in a public cloud might be accessible from the private cloud as well, albeit with limited resources.

This flexibility and redundancy aspect is the main focus of the hybrid cloud deployment model, which lets the client reap the inherent benefits without suffering the drawback of relying on only one cloud model. It has great potential, and is therefore rated high.

Infrastructure

Implementing a hybrid cloud usually means to have some hardware placed at the client's end, lowering the scalability and flexibility potential of the cloud. Having public cloud resources at your disposal could however mitigate this issue, as one may be able to borrow resources from the public cloud when needed. This is often called *cloud bursting* [77]. One can for instance imagine a situation where one encrypt backups of non-confidential material, and store it in the public cloud rather than the private and limited one.

By harnessing the hybrid cloud environment to utilize the abundant resources of a public cloud, many of the inherent issues with private-CC disappear. The main problems with the solution are that not everything can be handled in a public cloud before it is deemed safe, and strict assessments have to be conducted to clear data for public cloud processing in its current state. How this can be done effectively is yet to be uncovered, and many CSPs including [78] try to create secure multi-tenant infrastructures for the corporate or government market. We therefore rate this benefit with a medium impact level.

Integrity

The hybrid cloud shares the potential benefits for integrity as noted in the sections for private 3.2.2 and public 3.2.3, mainly concerning backup and contingency systems. The impact level was assessed as low.

3.2.5 Conventional

As the cloud deployment models are matched against traditional server-client systems, we also have to assess this conventional form of computing infrastructure. The following benefits can be mfound in a conventional system.

Access

The server-client data centre of a company only provides service to the owner, or owners of the hardware. This limitation means that colocation is not an issue, and as an effect the attack surface is smaller. As a result it has high impact on the security of the system. This benefit is evaluated as one with a high impact level.

In addition, a conventional system benefits from being hosted on-premise at the client's location, which grants it a slight security benefit together with a private cloud model. The impact level is therefore assessed to low.

By having direct access to the systems hardware, the client can react quicker to an incident than a CSP will be able to. The client will always be able to prioritize himself, while a CSP might have both their own infrastructure and clients with better uptime clauses in their SLAs to worry about first. This is noted as only a slight benefit, and the impact level is assessed to low.

Availability

None other than the client can use the resources of the computer system. This means that the client will always know the amount of resources he has available, a benefit that is assessed to medium. Furthermore, the conventional systems will remain partially functional if the ISP breaks down, as on-premise access will still be available. This benefit is also evaluated to be of medium impact.

Infrastructure

By having direct access to hardware, the import or export of data is not limited by bandwidth or the activity level of the ISP and CSP, only the activity of the hardware itself. For off-premise cloud implementations, this will mean somewhat longer load times compared with the on-premise implementation, but the benefit does not seem to be very significant. The impact level is therefore assessed to low.

Integrity

No distinct benefits have been found for a conventional system compared to CC.

3.3 Drawbacks

Here the negative implications of implementing a forensic application in the cloud will be covered.

3.3.1 Drawbacks Overview

Table 2, adding up the drawbacks, has been created in the same manner as Table 1. The same columns have been chosen to represent the classes and deployment models they affect, a description of the drawback, the deployment models that are not affected, possible results of the drawback and the assessed impact level.

Table 2: Drawbacks overview

Category	Environment	Drawbacks	Compared with	Possible Results	Impact
Access	Hosted-Private, Public, Hybrid	Shared access gateways to the cloud	Conventional, Private	Attacks of the access gateways might endanger access to the cloud implementation	Low
Access	Hosted-Private, Public, Hybrid	Authorities could demand access to or confiscate hardware at a CSP	Conventional, Private	Loss of or limited access to information	Medium
Access	Conventional, Private, Hybrid	Direct hardware access creates an additional attack vector	Hosted-Private, Public	Malicious insiders might gain access to or destroy hardware	Low
Access	Conventional, Private, Hybrid	Failing hardware can be harder to mitigate for small companies	Hosted-Private, Public	CSP-hosted services always have personnel on-call to handle failure	Low
Access	Public, Hybrid	Leakage due to colocated data	Conventional, Hosted-Private, Private	Data leakage could occur between clients colocated on a server	High
Access	Hosted-Private, Private, Public, Hybrid	Distributed data	Conventional	Data spread over several servers expands the attack surface, and may be a problem for CoC when analyzing case material	Medium
Access	Hybrid	Sharing data between clouds	Conventional, Hosted-Private, Private, Public	Information sharing between clouds could open an additional attack vector	Medium
Availability	Hosted-Private, Public	The CSP becomes a single-point-of-failure. Scheduled and unscheduled downtime can occur without notifying the client or gaining his consent	Conventional, Private	Failure or downtime at the CSP will suspend all cloud services, possibly indefinitely. Loss of data might occur	High
Availability	Hybrid	The CSP becomes a single-point-of-failure for the public part of the hybrid cloud. Downtime can occur without a clients notice or consent	Conventional, Private	Failure or downtime at the CSP will suspend all public cloud services, possibly indefinitely. Loss of data might occur	Medium

Continued on next page

Table 2: Drawbacks overview – continued from previous page

Category	Environment	Drawbacks	Compared with	Possible Results	Impact
Availability	Hosted-Private, Private, Public, Hybrid	Thrashing when a cloud reaches near full capacity	Conventional	Could seriously degrade the level of service from the cloud	Low
Availability	Public, Hybrid	Degenerated service due to colocated data	Conventional, Hosted-Private, Private	Colocated clients demanding high throughput or ones often under attack can degenerate services	Low
Availability	Conventional	High consequences of host errors	Hosted-Private, Public, Hybrid, Private	Full service breakdown can occur because of system errors at the host node	High
Infrastructure	Hosted-Private, Private, Public, Hybrid	Migration issues	Conventional	Applications might not be possible to migrate to the cloud of choice	Medium
Infrastructure	Hosted-Private, Public, Hybrid	Upgrades and Updates	Conventional, Private	CSPs may update or upgrade the cloud, introducing application compatibility issues	Medium
Infrastructure	Hosted-Private, Public, Hybrid	Moving large amounts of data	Conventional, Private	Pricing models can charge significant sums for moving data in and out of the cloud	Medium
Infrastructure	Hosted-Private, Private, Public, Hybrid	Vendor lock-in	Conventional	Cloud applications may only be compatible with one CSP, making it hard to change	Medium
Infrastructure	Hosted-Private, Private, Public, Hybrid	The lack of compliance regulations means that there are no set regulations for cloud services	Conventional	The service level might not be sufficient, and analyzing the situation would be hard	High
Infrastructure	Conventional	Clients nodes have to be powerful to handle possible tasks and programs	Hosted-Private, Private, Public, Hybrid	The cost is higher, and latent computing power is mostly unused at the client side	Medium

Continued on next page

Table 2: Drawbacks overview – continued from previous page

Category	Environment	Drawbacks	Compared with	Possible Results	Impact
Infrastructure	Conventional, Private	Backups have to be handled on-premise	Hosted-Private, Public, Hybrid	Expensive and arduous backup systems have to be implemented and maintained	Low
Integrity	Hosted-Private, Public, Hybrid	Off-premise storage of confidential information	Conventional, Private	CSP must be trusted to handle data	Medium
Integrity	Hosted-Private, Public, Hybrid	No clear custody laws for data in the cloud	Conventional, Private	The CSP could gain the custody of data it is not entitled to	Medium
Integrity	Hosted-Private, Public, Hybrid	Status information	Conventional, Private	Getting correct status information from a CSP is harder than from local sources	Medium

3.3.2 Private Cloud

In a strictly private cloud environment placed within company walls, possible risks basically follow the same rules as for an ordinary on-premise distributed network. However, as a private cloud can also be hosted off-premise by a cloud vendor, the data the company manages in the cloud is not contained and secured by the company itself. This can cause several problems, many of which are noted in this section.

Access

Limiting the access to confidential information in an off-premise hosted-private cloud has some issues associated with it. For one, the segregation of the hosted data from other clients must be upheld. Even though a hosted-private cloud should have its own isolated servers and resources at the CSP, the communication channel to it, the Internet, is common for all users. For instance, sharing a log-on service with other cloud clients can potentially lead to unauthorized personnel gaining access to the cloud solution, as such sites are popular phishing targets. Phishing can be defined as identity theft, often through the use of rogue web pages [79].

Hosted-private cloud solutions like VCODYNE [42] offer VPN access to the client's cluster which is set up in an isolated virtual local area network (VLAN). Such access channels should be offered by hosted-private CSPs, and those not offering such security should not be chosen as hosting services. Because this is a choice of the client the drawback has low impact on the implementation.

A major concern if the private cloud is hosted off-premise, is that the authorities of the countries in which the cloud is hosted can demand access to the files in certain situations. For governmental clients, and especially law enforcement, this may not be a big issue if the solution is hosted domestically. However, if the entire, or parts of the private cloud is hosted in another country, the rules of the host country apply. Storage mediums can therefore be made available to the host countries domestic law enforcement forces.

Where a CSP stores its clients' data can sometimes be hard to uncover, as a domestic CSP could for instance use Amazon EC2 as a foundation for its infrastructure. Such issues should be covered when assessing a CSP, and should therefore be well known prior to a possible choice of CSP. Furthermore, cloud actors like Google with GAE have a file system which is encrypted and distributed over several storage devices [72]. This means that even if the government of the country where the cloud is hosted demands the storage devices delivered, the data should be safe as long as the client has solitary access to the deciphering key. Loss of data may of course still occur.

Hosting data abroad is, according to us, an issue of medium impact. While loss and leakage of data is possible, the solutions consisting of encryption, backups and choosing a domestic hosting service should be sufficient to bring the chances of this happening to a minimum.

If hosted on-premise, having direct access to the hardware could also prove a drawback compared to other cloud models, as it retains the existing attack vector from conventional systems. This could be exploited, for instance by malicious insiders. It should, however, be less of an issue, as internal procedures and continuous assessment of personnel should counter this. The impact is therefore regarded as low.

Another problem of hosting one's own hardware is that if the servers fail outside of office hours, the needed personnel might not be available. A CSP will always have

maintenance staff present to mitigate such situations in a data centre. Even though this is the case, local administrators at the client side could also react faster, especially if the CSP suffers a major problem, and has to assess which systems should be mitigated first. Since both sides of the issue can be argued evenly, the impact is assessed equally to the one set in Section 3.2.2, namely low.

Availability

The largest problem area concerning availability is that the cloud provider becomes a single point of failure. Therefore, if the CSP suffers downtime, so will the company or agency using its services. Some of the problems can be mitigated by creating a system for off line work in the periods when the cloud is down, but this could in turn cause more problems in the form of what should be buffered on the client or corporate side of the cloud, how version handling is done as different versions of documents are created, and synchronization of these as the cloud comes on-line.

There exist many possible causes for downtime at the CSP, ranging from natural disasters, to scheduled and unscheduled maintenance and to focused attacks. In any case, occurrences of temporary down-time have to be accounted for, and a SLA should be written to describe the maximum limit for both scheduled and unscheduled downtime. Another possibility can be that the CSP goes out of business. In that case, the conservation and recovery of the hosted data could be in jeopardy. Such a scenario should always be covered by the SLA.

This problem will persist for hosted-private and public clouds, and will also have some effect on hybrid cloud solutions. There are many examples of CSPs having been forced to discontinue their services, often with little notice to clients. The results can be everything from temporary unavailability to a total loss of all service and data stored in the cloud. As a main concern, the issue has a high impact rating.

Another possible problem concerning availability is *thrashing*, which may occur when a cloud gets close to its maximum capacity. In [74] it is noted that when a cloud reaches about 80 % capacity computers and servers will start moving data rapidly between storage and memory to uphold the characteristics of flexibility and scalability and maintain a resource pool ready for use by clients. This can cause the cloud services to become almost non-responsive. Thrashing may be the most problematic for private clouds, as on-premise hardware is not expanded as easily as in the off-premise clouds. DDoS attacks or extreme changes in resource demands can induce cloud thrashing, but it becomes less and less likely as the technology is improved.

The handling of thrashing has become better in the last few years, and will be limited by making sure that the cloud always has sufficient resource buffers. In addition to this, the automated management of the cloud should be able to foresee such happenings, and to either shut down non-essential consumers of resources, of gracefully degrade the resource usage of applications and services temporarily [74]. In such cases the CSP will most likely reclaim resources from the cloud clients. This could result in unavailable services, but will ensure that the core cloud systems are kept functional. Because of this the chances of thrashing is low. This is reflected in the impact level, which therefore remains low as well.

Infrastructure

Several infrastructural issues may arise when transitioning to a cloud environment, the first of which is migration. Companies considering CC almost always want to move some parts of their already existing computer systems to the cloud. Herein lays a problem, as the cloud infrastructure, like any other computing environment, has its limitations and requirements. Migrating existing systems to the cloud can be hard or near impossible for certain cloud solutions without making changes to the original system. As an example, migrating software to Force.com, the point-and-click programming environment of SalesForce.com, is not possible if not done the SalesForce way. A one-to-one migration, or moving of the application directly to the cloud environment, is simply not possible if written in any other programming language.

This problem should be thoroughly assessed before choosing a CSP. As most clients should know what existing applications or new functionality they want to procure, a dialog with the cloud provider should uncover what can and can't be done. The problem could persist in some way though, which may limit the practicality of the chosen cloud. Because of this and the fact that the issue is relevant for all cloud deployment models, the impact is assessed as medium.

Adding to the issues of compatibility, the CSP may be able to update or upgrade its infrastructure without client consent. This means that the client's own applications running in the cloud could fail after such an event. As the client has little possibility to perform the same types of compatibility tests as would be performed before any major upgrade of the on-premise systems, the company has to rely on the CSP to handle backwards compatibility and continuity in the cloud environment for each upgrade.

Given that this is a private cloud, and thus not shared with the rest of the cloud population, the client should have a say when it comes to the management of the cloud. Clauses can be included in an SLA to indicate that no changes should be done to the cloud that significantly interfere with existing client implemented software and data. Furthermore, it should include specific information concerning the cooperation of CSP and local IT administration, so that meetings and workshops can be held at set intervals. In this way, the client can adapt to the changes of a cloud environment before the actual change occurs, mitigating the problem entirely.

For other cloud deployment models the issue is handled differently, as they might not be able to directly influence a CSP. This is discussed further in 3.3.3 and 3.3.4, and because of the statements noted there, and the ones covered here the impact level is set to medium.

CSPs operate with several different prizing models, some of which will put restraints on the amount of data that can flow into or out from the client's cloud solution. Moving data inside of the cloud is often free, while sending or receiving data from the outside carries a fee. Amazon price their EC2 cloud solution in a similar way [80].

This type of pricing means that every time large transitions of data takes place, it may be both pricey and require large parts of the systems bandwidth for a significant amount of time. Problems like this have sometimes lead to clients sending their data on physical devices to the CSP, rather than attempting a change of the system themselves [81]. Amongst others, Amazon has a service for sending data on physical media [82].

For a government client, such a solution may not be a possible alternative, as most of the data might be highly confidential. Sending the data by a courier, and then ensuring

that only authorized personnel has access to it could be a solution, but it requires that the government has an employee that is stationed at the CSP in such cases, ready to handle the uploading of the data to the cloud.

Non-combined cloud solutions like private or public clouds may not have big issues with this, as most data is stored and handled within the same cloud. However, the problem especially applies for hybrid clouds, as is covered in Section 3.3.4. The issue is believed to have medium impact.

Resuming the problems with interoperability, vendor lock in can be a big issue. As clients start using CC, the competition in the market will mean that one may look to move ones cloud solution to a different CSP. This could be hard, as one might not be able to migrate software and data from cloud to cloud without some major changes. There are no standards that govern CC, which means that any infrastructural proprietary solutions the CSP wishes to implement is allowed. It also means that one could have to rebuild a cloud implementation from scratch if one wants to transition from one cloud to another.

One highly proprietary part of the cloud is the API, which will have to be standardized to counter the problem of vendor lock in [74]. Failing to do so could lead to a situation of monopoly, where a governmental agency would have to refrain from changing CSPs based on the inherent difficulty and high costs of such a transition.

This remains a big problem for CC, and is one of the main reasons why companies refrain from implementing a cloud solution. In a survey performed by IDC in 2011 [83] 80 % of the participants noted the lack of interoperability standards to be a key challenge of cloud services. It also impacts all cloud deployment models including the private on-premise ones. However, cloud initiatives are trying to bridge the gap between CSPs, and cloud environments like CloudSwitch [84] has built a cloud specifically with this in mind.

Because we believe that vendor lock-in is a problem, but one that does not scale up to problems like cloud compliance issues, the impact level is set to medium.

A problem for private cloud implementations is the handling of backups on-premise. This can be an arduous task, including for instance tape backups, or optical media backups. Public cloud providers will have enterprise scale backup systems which will back up the data according to SLA agreements, systems which proficiency is beyond reach for most clients, and even government initiatives. We therefore see this issue as a slight deficiency of the private cloud, with an impact level of low.

Finally, a grave issue is the lack of compliance regulations linked to CC. Most of the regulatory statements, like the Federal Information Processing Standards (FIPS) Publication 199, predate CC. Because of this it has not taken any stand concerning the technology. Transitioning into CC will therefore be done based on existing security standards like ISO27001/2 [85, 86], and the government agencie's own practices and regulations for security and implementation. Without proper compliance organs, implemented systems might have to be significantly modified when they are finally created.

Many CSPs now support some compliance organs, which is a step in the right direction. As an example, Amazon Web Services is compliant with, amongst other standards, PCI DSS, FISMA and ISO 27001. However, these regulations are not made for CC, and will therefore never be able to fully cater to its unique features. As so much of the technological procurement of companies and governments around the world counts on compliance as a sound measurement for security and soundness, basing this on sub-optimal

compliance requirements simply will not do. Until compliance documentations and best practices are created for CC in particular, this remains a high impact issue.

Integrity

Safeguarding the integrity of the system and the data stored within is of critical importance for a government CC environment. In the context of law enforcement, the integrity of case data and evidence is vital for an investigation.

As with any computer storage system, data can become corrupted or lost, and backup systems have to be carefully managed to ensure that information is not lost and that the system keeps copies as close as possible to the current state. For private clouds it is possible that the client can keep these, copying and replacing data as they see fit, but it can also be left to the CSP. In this case the integrity becomes an issue of trust, and no SLA can protect against the possible loss or leakage of confidential information.

For law enforcement it may not be viable to trust a CSP with confidential data, as governmental regulations dictates how such information is to be handled. This will of course have an impact on the usability of clouds hosted off-premises. Encryption may be one solution, as data could be kept at off-premise storage facilities as long as it is only recoverable by authorized personnel. In order to prove this though, a vendor must open the source code to be assessed by the client, something that may be infeasible. This leaves slim possibilities for the usability of off-premise storage and processing by law enforcement. We assess this issue to one of medium impact.

Barring the integrity of stored information, there is also the issue of the integrity of status information. Clients of public CC solutions will need constant information regarding, amongst other things, resource utilization, number of system users, and benchmarking against a given SLA. This information is important both to ensure that the cloud scales correctly compared to the cloud usage patterns, and to be able to assess if permanent upscaling or downscaling is needed. The integrity of this information is vital for upholding high performing cloud systems.

A cloud provider should offer dashboards or management systems that can be used for keeping an overview of such information, enabling the client to promptly react to changes. Most CSPs offer such systems, like the VMware Cloud Management Dashboard and Citrix XenCenter, but the information they provide is different depending on the vendor and implementation plan. Because of this an evaluation criterion should be concerned with the abilities for status information mining, usability and presentation.

No best practices or standards exist on the subject as far as we know, which makes it hard to compare cloud vendors. Seeing that the issue can also create problems for benchmarking of the CSP's services and similar calculations, we estimate that the impact level is at least medium.

Another problem concerning integrity is the law governing CC. The question of who owns the information in the cloud once it is placed there is uncertain, and not yet specified by law [74], which means that carefully crafted SLAs have to be applied to cover all possible implications of this type.

As the SLA is a legally binding contract between the CSP and the client, including such ownership of information in the contract will establish that ownership when all parties agree to and sign the SLA. This mitigates the problem to an extent. However, SLAs should periodically be assessed to reassure all parties of their obligations, and modify the

document to reflect changes in service or recent changes in laws and regulations dictated by the government. The issue covers all cloud deployment models besides the private cloud, and its impact is set to medium.

3.3.3 Public Cloud

Using a public cloud environment means that all the information kept in it is stored off-premise at a CSP. It also means many different clients can share resources like CPU time, memory and storage devices called colocation of clients. Colocation raises some interesting issues.

Access

As with the hosted-private deployment model, a shared access gateway could cause problems for a public implementation. However, securing services and communications across the Internet has long been a focus point for improvements, and tried methods for such do exist. Because of this, the impact level of the drawback remains low.

When utilizing a public or hybrid cloud many issues can be linked to colocated data. For instance, if a client's data is not properly segregated, information might leak from one cloud client to another, or into the host system, VMM or hypervisor managing the cloud clients and resources. This is often called information bleeding, and can potentially result in unauthorized personnel gaining access to confidential data.

Other leakage can be linked to the reuse of hard drives, where data is not purged completely prior to re-instalment to another client. Another problem may be linked to the client's own software, as bugs in the client-induced applications could give an attacker administrative access with which he can transfer information out of the cloud or via his own colocated cloud services. It may prove that it is possible for attackers to successfully identify the cloud servers hosting interesting clients, and to deploy their own client on that same server in order to exploit the situation [87].

The technology used by the CSP has to be sound for the system to be safeguarded from information leakage, and segregation of clients using VLANs or physical locations are currently used together with encryption to uphold this. Still, the technology is young and problems will probably continue to arise as the cloud population of the world expands. For law enforcement this means that public cloud services will remain out of reach until better solutions are found, and we therefore assess this issue as one with high impact for the governmental use of CC.

Using a public cloud, files are most likely distributed over several different servers. While this mostly eliminates the single point of failure, it does mean that the attack surface expands, meaning that more attack vectors become available for possible intrusions. The defence against attacks have been covered in several of the preceding sections.

Additionally, the distribution of files between several servers will be less attractive for the storage of case related data like evidence. This is because of the concept of CoC, which is violated if evidence files are copied and distributed throughout the cloud without being properly accounted for.

The problems of storing digital evidence in the cloud can be emphasized by considering a paper by Ćosić [88]. As evidence can be proved inadmissible in court if the movements of such evidence cannot be proven, logging and time stamping have to be conducted for every change in state. This means that all evidence stored in a cloud environment would have to be time stamped and signed every time the virtual storage server is

moved in the cloud. This would require information about, amongst other things, where the evidence previously resided, when it was moved, why it was moved and where it was moved to. This is information that public cloud vendors are not likely to provide.

Overall, the way clouds process and store information may limit its usability for law enforcement, and therefore has medium impact.

Finally, jurisdictional regulations in some countries could demand that governmental organs, like law enforcement, can gain access to data in certain circumstances. Storing files in servers situated in other countries may therefore make these available to the nation's law enforcement or domestic security forces. This was earlier assessed as an issue of medium impact in Section 3.3.2, and the arguments still apply here.

Availability

As discussed in 3.3.2, the CSP will be a single point of failure. This is particularly problematic for clients using the public cloud, which might lose all access to data during a service outage. The impact remains high. Thrashing was also mentioned in the same section, and will remain an issue for public CC, albeit one of low impact.

As a public cloud is dependent on optimizing the resources available at any given time, data will be stored on several different storage units which may be placed in different data centres and even countries. Having company data colocated with another client that has high resource demands could therefore make that data less accessible. On the other hand, colocation with a client that has low resource demands would mean that resources are more readily available.

Enforcement of such access could be possible through a carefully constructed SLA, which can include sanctions if the CSP cannot deliver the expected level of service, but this does not guarantee the needed access.

Barring the possibility of another company bogging down a client's access to its own data, some companies will be more exposed to attacks than others. Lately the world has seen several DDoS attacks against entities that in some way or another displeased the hacker community. These were both commercial, for instance linked to the WikiLeaks case [89], and governmental, as linked to the Megaupload case [90]. If a client is more prone to an attack given its political or commercial standpoints, the chance for an attack is also higher for the ones sharing storage and resources with the entity.

Periods of less responsive service will occur in any system, even conventional ones. However, compared to conventional systems, the flexibility and scalability of cloud systems can mitigate some of this deficiency. The impact of this drawback is therefore assessed as low.

Infrastructure

The problems with migration, interoperability and vendor lock-in persist in the public cloud, and can even elevate the severity of the problems. This is because, in a public cloud, the client may not be able to work as closely with the CSP as with private cloud implementations. This could lead to the client failing to receive information about infrastructure changes in a timely fashion, which will give him less time to react and adapt. Other issues concerning vendor lock-in and proprietary cloud systems are noted on Section 3.3.2.

A possible solution to this problem would be for the CSP to provide a transitional cloud environment during stages of upgrading or modification. In these cases the clients

should gain access to the modified cloud while still residing in the old version. This would give the client time to adapt and test its applications on the new version of the cloud, while simultaneously delivering services to their users.

The impact level of the problems regarding vendor lock in, migration and upgrades/updates are kept as noted previously.

The issues of importing or exporting large amounts of data will still remain for public cloud environments, and the decision has to be taken on how to transfer such quantities. Section 3.3.2 noted some examples, and the impact level remains at medium.

The compliance issues noted in 3.3.2 remains the same for public clouds.

Integrity

Many of the issues apparent in a private CC environment remain with public CC.

Storage and backup will be managed by the CSP, and the client will most likely not have the opportunity to keep a backup for themselves, at least not in the cases concerning SaaS solutions. Some kind of management concerning the backup might be possible though, where the client can select when and how the data should be backed up. Cloud clients using PaaS and IaaS solutions might be able to back up their own cloud to a location of their choice, but this will in most cases be infeasible, as a CSP often charge large sums for data sent out of and into the cloud. The impact level assessment remains at medium caused by these limitations.

The lack of legal regulations remains, and the ownership of data residing in the cloud have to be stated in an SLA if at all possible. Keeping track of status information could also be problematic for public cloud clients, as one might not be able to tweak information based on default cloud service models. The impact level for both stays the same as in Section 3.3.2.

3.3.4 Hybrid Cloud

Since a hybrid cloud has the advantage of using both a private and public model, one has the opportunity to mitigate some of the problems that are inherent in both individually. However, the problems which are related to CC in general will still be applicable.

Access

Using a hybrid cloud one should be able to construct an environment where important data that need to be accessed consistently is hosted on-premise. Information that needs to be shared more readily, or does not have to be accessible at the same rate, can be placed at the off-premise private cloud. Finally, information that is not critical, does not have high demands for security, and would benefit the most from being easily shared with cooperating entities could be placed in a public cloud solution.

For law enforcement this could for instance mean that case specific information is placed in the on-premise hosted cloud servers, while publicly available information is hosted off-premise in a public cloud. Doing this will limit the threats against case related and other confidential information, but also limit access to flexible and scalable resources that the public cloud provides.

The question of a shared access gateway for resources hosted off-premise remains of little concern for the hybrid model, and the reasons stay the same as stated in 3.3.2 and 3.3.3. Being able to choose what information is made available in a public cloud and what stays protected behind the on-premise firewall supports the chosen impact evaluation.

The problems concerning outsourced storage will stay the same as for public CC. Foreign governments can still demand access to the storage mediums hosting applications, and colocation of data may cause leakage of information. Futhermore, the distribution of data over several servers and data centres can create additional attack vectors. The arguments presented in 3.3.3 remains, and the impact level stays the same.

Barring these problems, one more remains that is specific to a hybrid cloud environment. Sharing information between the private and public cloud means that another attack vector will be present in such implementations, as the public cloud solutions may be open for additional attacks compared to the private cloud. Judging by attacks like the one performed in [87], chances are that compromising the public side could give the attacker a better chance at reaching information kept in the private cloud. However, no such methods for intrusion are known to us.

Since a hybrid model will be a tempting choice for many companies, the implementation and its possibilities for cooperation with public resources and services should be closely evaluated. In this way one may avoid introducing additional vulnerabilities. Since the possibility is present, we would assess the impact level as medium.

Availability

While the CSP is considered a single point of failure for both hosted-private and public clouds, the hybrid cloud will not be totally crippled by CSP downtime. The public part of the hybrid cloud will still be vulnerable though, and the impact can be severe it if occurs.

Colocated clients will exist in a hybrid cloud solution as a way for the CSP to optimize resource usage. Because of this leakage of information, resource deficiencies, service outages and loss of data might occur. Thrashing can also become an issue, both in the private and public cloud environment, and countermeasures like those mentioned in Sections 3.3.2 and 3.3.3 should be evaluated. As previously mentioned, the impact level of both thrashing and colocations remains low.

Infrastructure

Migrating an application made for a conventional system to the cloud will not always be a smooth process, this also applies for migration to a hybrid cloud. In the cases where the application should be compatible with both sides of the cloud, the problems could even escalate. The upgrading and updating of the cloud host will also remain problematic, and close cooperation with the CSP will be important if one wishes to avoid last minute changes and incompatibility issues. Both the problems concerning migration and update/upgrade issues were evaluated as medium impact problems.

Data transfer in and out of a cloud is often charged for, while transferring data within the cloud is free. For hybrid clouds this is an inherent issue, as these interconnect a private and public cloud, and often base their use on the cooperation between the two. Let's for instance say that a company hosts the applications on-premise, but outsources all database storage to a public CSP. In this case the client will be billed for all data transferred between the clouds, which could quickly become the commanding post of the payment plan.

Hybrid clouds using a hosted-private and public cloud combination may be able to circumvent this problem if the public and hosted-private clouds are provided by the same CSP. However, this is usually not the case, and hybrid clouds mostly consist of some kind of on-premise cloud combined with a public solution. This means that the data transfer

issue will remain problematic. Based on the fact that this is not an issue of the same gravity for the other cloud deployment models, it is assessed to medium impact.

The compliance issues noted in 3.3.2 remain the same for hybrid clouds, and so does vendor lock-in issues. One may even be locked to a specific combination of cloud vendors if an application is made to run on both private and public infrastructures.

Integrity

Storage of confidential information is a problem for most off-premise cloud solutions, this also applies to the hybrid cloud. Even though most such information should be kept in the private part of the cloud, some information, like personal identifiable information (PII), could still be present in the public cloud. Making sure that all precautions are made to limit the spreading of such information will thus still be important, and some examples of such are found in the prior sections which discussed integrity. The impact rating is set to medium based the previous discussions on the topic.

The lack of laws and regulations remains for hybrid cloud solutions, which have also been discussed previously. The impact level is evaluated to medium.

In the hybrid cloud two sets of status information have to be collected, the status of the private side, and the public side. Based on how the two sides of the cloud cooperate, the information could even be needed in processes where information should be collected from or handled in the opposite cloud from which it came. The regularity of such information might therefore be crucial. As stated in previous sections the impact of this issue is assessed to medium.

3.3.5 Conventional

As drawbacks exist for CC environments, they can also be found for conventional computer systems. These are the ones we have explored.

Access

While direct access to hardware gives some benefits, one drawback of this access is that it will make the hardware available for malicious insiders. This adds an additional attack vector compared to CC models. The impact level that was assessed as being low in 3.3.2 still applies here for the same reasons. Furthermore, full control over hardware means that the client has to mitigate errors and replace faulty equipment themselves. CSPs, being staffed 24 hours a day for such occasions, may be able to correct errors quicker. The same arguments as discussed earlier apply here, and the drawback is evaluated as of low impact.

Availability

The biggest issue with conventional server-client systems compared to CC deployment models is that single errors in hardware on the server side will potentially shut down all services on the target node. We see this as a huge drawback of conventional systems, which can drastically lower the availability. All CC environments are built on virtualization platforms which allow for the migration of VMs to other host-systems. By doing so, the availability of the system will at best suffer no consequences what so ever, or, at worst, cause limited access or performance. We evaluate this as a serious drawback for conventional systems, with the impact level high.

Scenario Name	Access	Availability	Infrastructure	Integrity
Public Information Services	2	4	1	3
Mobile Information System for Law Enforcement	4	2	3	4
Internal Law Enforcement Cooperation System	3	3	4	4
Digital Forensic Analysis System	3	2	2	4
Public On-line Case Registration	4	2	2	3

Table 3: An overview of the scenarios and category priorities

Infrastructure

As a server-client infrastructure usually places applications on the client nodes, these have to be quite powerful compared to the clients used in a cloud environment. In the cloud, the servers handle the bulk of the computing and data handling, leaving the clients to view the output and issue commands. This makes a conventional system less cost- and resource effective than a cloud implementation, as resources available at the client side is mostly unused. The drawback is prominent, and assessed as having a medium impact level.

As with the private cloud implementation, the handling of backups on-premise could be an issue. We assess this issue to be of the same magnitude as for the private cloud, low.

Integrity

No real drawbacks are found concerning the integrity of data or services hosted on a conventional infrastructure compared to those in cloud environments.

3.4 Scenarios

For better utilization of the results gained from the benefits and drawbacks study, five possible scenarios for cloud implementations for government and law enforcement have been created. These have been composed specifically for the purpose of showing that different systems of different agendas will prioritize the categories we used in the lists above differently. Because of this they will gain more from using certain deployment model than others.

It can, in some cases, be argued that some of the categories are equally important. Putting the categories in ascending order based on importance will therefore not create a complete picture of the scenario. Because of this we chose to rate the importance of each category on a scale from one to four, where higher numbers represent higher levels of importance.

Table 3 presents an overview of the different scenarios and their category priorities. The first column states the name of the each scenario, while columns two till five contains the values assigned to each category in the AAII model for each of the scenarios.

3.4.1 Public Information Services

The governmental agencies of a country have a continuous need to inform the public. Examples of this could be to release recommendations from ministries, white papers and other news meant for the public, or governmental emergency messages during a time of crisis or war. An example of such a service is the Norwegian government's public information website [91].

For such a system the different categories are prioritized in the following way:

- **Availability** **4**
- **Integrity** **3**
- **Access** **2**
- **Infrastructure** **1**

Availability is prioritized as keeping the system available for the public is vital, especially in times of crisis or war. Next to this the integrity of the information should be secured, as the information presented to the people should be correct. After this, access to the service, meaning the possibility for administrators to log on and influence its content, is prioritized. At the end of the priority list infrastructural demands are placed. Since the information service has quite a monotonous task, there is less need to be flexible concerning new modules or expansions. The only important infrastructural demand is that the system can reach the population of the nation and support their continuous use.

3.4.2 Mobile Information System for Law Enforcement

For a police officer on patrol, or detectives present on a crime scene it is important to be able to access information and log events. Earlier this was only achieved by using radio communication, thereby calling in, or receiving messages live. This created the possibility that messages never reached their targets, as messages were not saved for later use. These days, law enforcement officers get information both through dispatch via radio, and by using a laptop (or equivalent) to access databases and retrieve or input information.

For the computer part of this system, categories are prioritized in the following way:

- **Access** **4**
- **Integrity** **4**
- **Infrastructure** **3**
- **Availability** **2**

The reason for this order is primarily that no unauthorized personnel should be able to access the system, as it will contain large amounts of confidential information. The integrity of the data might here be equally important, as officers should get correct information in order to take the correct actions in any situation. The infrastructure of the system is also quite important, as mobile systems have high demands to durability and connectivity in order to function properly on the move. Finally, availability is placed at

the bottom of the list. This is not because it is unimportant, but mostly because the backup line via radio communication to the dispatch should always be available. It will therefore be possible to access and enter information via dispatch.

3.4.3 Internal Law Enforcement Cooperation System

Cooperation within law enforcement has long been a problem, as the different branches of police forces often have different information systems and regulations for the use of such. This means that they are not easily integrated with other systems. Even though the ideal situation would be to create a combined system for all police services, this might not be a possibility before the issues of regulations are in order.

Until this is done the best solution may be to create a system the different law enforcement branches can use to make information available for each other, in order to more efficiently cooperate in ongoing investigations. The system could enable the police services to create digital task forces in order to construct a combined forum and information exchange point for each case, and limit access to authorized personnel. By including access to the branches' databases it could be an effective system that could for instance include data mining services.

For such a cloud implementation, the categories would be prioritized in the following way:

- **Infrastructure** **4**
- **Integrity** **4**
- **Access** **3**
- **Availability** **3**

The reason that infrastructure is places first here is that the system should be able to interconnect with several different databases and equipment in order to create a good foundation for cooperation. Integrity has also been given the strongest rating, as the information has to be safeguarded to ensure its trustworthiness. If the integrity of the information cannot be trusted, it cannot be used in an ongoing investigation. Access control and configuration is then prioritized, as only authorized personnel should be able to reach this information. Under no circumstances should information concerning an ongoing investigation be leaked, unless such exposure is intended and approved. The availability of the system is also prioritized at level three, as information should be able to reach all actors in the coalition efficiently.

3.4.4 Digital Forensic Analysis System

Systems for the analysis of digital evidence have much to gain from being implemented using CC. Such equipment is usually very resource-consuming during analysis, and dormant the rest of the time. The physical devices used for digital evidence analysis are both expensive and often isolated in use, meaning that few other jobs can be performed on the machine when it is not in use for analysis purposes.

CC is perfect for these types of jobs, as dynamic up-scaling and down-scaling means that the resources can be acquired when needed, and released again when the job is done. Depending on the job and how time-critical the analysis is, more resources can be demanded to finalize the activity quicker.

Several possible implementations of such a system can be mentioned: An in-house private cloud version is one. A version using public- or hybrid-cloud resources during incident response another. For this scenario we will use the in-house example.

We assess that a cloud based analysis system should prioritize in the following way:

- **Integrity** **4**
- **Access** **3**
- **Infrastructure** **2**
- **Availability** **2**

The most important thing concerning the analysis of evidence is to make sure that it remains unchanged throughout the analysis period, and that all analysis performed is logged. Such characteristics directly reflect the principles of CoC and evidence integrity. The integrity of the results is also of the utmost importance, as it have to be usable in a court of law. Next to integrity, access will have to be prioritized. Only authorized personnel should be enabled to reach the evidence or analysis results, and the users that do handle the evidence should be registered together with a thorough documentation of the actions performed.

Next, we prioritize the infrastructure of the system based on the many different types of evidence and analysis systems the implementation should be able to handle. It should be versatile enough to cater for the needs of the analysts, but also resource effective in the way that computing power should not be squandered. Availability is also given an importance level of two, as it is important that such a system remains available for its users, especially during the analysis of evidence that has a deadline (for instance a court date). Even though this is the case, integrity and access takes precedence, given the sensitive nature of the content handled in the system.

3.4.5 Public On-line Case Registration

In addition to handling casework in the cloud, the registration of cases by the public could benefit from using a cloud infrastructure. An example of such a service in use today is the Norwegian police's system for case registration and information [92]. By enabling the public to register their own cases, some of the pressure is lifted from telephone and police station personnel.

A person should be able to file a report regarding criminal activities and send it to the police station of the jurisdiction in which the crime was committed. After filing the report the person should receive a receipt informing that the report has been created and assigned to an officer. It should also contain the date of the report and a deadline for the police response. The system does not need to register its public users, as all other information in the case will be handled through other channels. It will therefore be a one way system. Implementing such systems for emergency services is a good idea, but it would both require two way communication and very high demands for response times. We do not take such use of the system into account for this scenario, and intend it to be used for non-emergency contact only.

For such a system the categories should be prioritized in the following way:

- **Access** **4**

- **Integrity** **3**
- **Availability** **2**
- **Infrastructure** **2**

Access is prioritized, as only the case workers assigned to each case should be able to access the information. Next after this the integrity of the information has to be safeguarded, and should not be modified in the system without proper logging (for instance information regarding the person doing the modification and a time stamp). Availability and infrastructure ends up at the bottom of our list with a level of two for both. The availability will not be as pressing, as police can always be reached both at the station and by telephone. Infrastructure is also of less concern, as the information entered into the system will be quite conventional (text and possibly pictures with a set format and maximum file size).

3.5 Results and Discussion

As the results presented in Sections 3.2.1 and 3.3.1 relied on a qualitative assessment, we found it fitting to rate each benefit and drawback into the classes high, medium or low depending on its impact in a positive or negative direction. Now, in order to quantify the results into something we could measure, we transferred the impact levels into impact values.

We first attempted to rate the benefit and drawback classes by assigning three points to all cases resulting in a high impact value, two points to all cases with medium impact, and one point to all the low impact cases. This, however, produced a result that was very closely packed, giving few distinct features. We therefore decided to square all the original values, resulting in clearer definition. The values therefore ended up in the following way:

- Benefits and Drawbacks
  - High = $3^2 = 9$
  - Medium = $2^2 = 4$
  - Low = $1^2 = 1$

3.5.1 Calculation Description

The results were calculated in Excel, and the Excel file used in this calculation (DrawbBenCompSquared.xlsx) can be found in the archive RQ1calculationsInDepth.zip, which is appended to this thesis. Other calculation methods were also attempted, and the results of these (and the method presented in this section) can be found in the Excel files referenced in Appendix B.

During the calculations we found that the hybrid cloud deployment model was at a significant disadvantage. Possible reasons for this are discussed in Section 3.5.5. We therefore decided to remove the hybrid deployment model from the calculations, thereby also removing any points another deployment model would have gained or lost based on their comparison. The Excel files presented in B include results both with and without the hybrid deployment model.

The file BenDrawTable.xls, found in RQ1calculationsInDepth.zip, includes one sheet for benefits and one for drawbacks, and the impact rating of all of these. The two sheets are used by all other Excel files in the folder to calculate results.

Inside the other files of the .zip-folders the layout is the same, but we will use DrawbBenCompSquared.xlsx as an example since this is the one we are using in this chapter.

The first sheet, ”Explanation”, describes the values used to perform the calculations.

The next five sheets, ”Private”, ”Hosted-Private”, ”Public”, ”Hybrid” and ”Conventional” include the calculations of benefit and drawback points for each of the five deployment models when compared to the others. These are the same type of tables as presented in Appendix A. All these sheets, barring the sheet for the hybrid deployment model, have two tables. The table at the top includes the hybrid model, while the table at the bottom has discarded it. The bottom table was created when the decision to omit the hybrid model was made.

Using the ”Private” sheet as an example, the tables in these Excel sheets consist of five main columns. The first column contains the deployment models we are comparing the private model to. The second column named ”Benefits” contains one sub-column for each of the categories we use. In the cases when the private model has a benefit, and some of the other models have not, the private cloud gains positive benefit points against that deployment model.

In the third column called ”Drawbacks” we also have one sub-column from each of the categories. In the cases where there exists a drawback which impacts the private model negatively, but does not affect some of the other deployment models, the private model gains negative drawback points against these deployment models.

In the forth column called ”SUM” the negative drawback points are subtracted from the positive drawback point for each comparable deployment model in each category. As an example, cell J4 calculates cell B4 minus cell F4. In the fifth column named ”SUM TOTAL” the numbers in the SUM column for each deployment model are summed up into a grand total.

The next four sheets named ”Access Results”, ”Availability Results”, ”Infrastructure Results” and ”Integrity Results” are category comparison tables for each of the four categories. There the numbers found in the ”SUM” columns of the former five Excel sheets are combined, and every permutation of deployment models can be viewed in each category. Also here, two different tables are included, one where the hybrid cloud is included, and one without.

Let’s say we want to compare the public deployment model with the private deployment model with regards to access. In Table 25 found in Appendix C.1 we find the public deployment model in cell 4,1 (column 4, row 1), and the private deployment model we want to compare with in cell 1,2. The cross section between the two is cell 4,2. In this cell the sum of the benefit and drawback points of the private deployment model compared to the public deployment model when focusing on access (cell 10,4 of Table 17), 8, is subtracted from the sum of the benefit and drawback points of the public deployment model compared to the private deployment model when focusing on access (cell 10,3 of Table 19), -10. The calculation then becomes -10-8 = -18, indicating that the public deployment model is at a disadvantage when compared with the private deployment model in the access category.

Comparatively one can view cell 2,4, indicating the opposite view. In this cell the public access points is subtracted from the private access points, 8-(-10) = 18, indicating that the private deployment model has an advantage over the public deployment model when it comes to access.

The last row of the table indicates the sum of all the comparison points of one deployment model against the others. The deployment model with the highest sum gives us an indication of which is the best deployment model for the table category. The collection of the top scores in each category when using the single value calculation method is given in Table 33.

The last sheet of each file called ”Boiled down results” feature the combined ”SUM” rows of all the category comparison tables of the Excel file into one table. The table allows us to view the highest rated deployment model in each category. These tables are similar to the one presented in Appendix D. As before, two separate tables are used, one including the hybrid model, and one without it. Histograms for each table are also included as a visual representation of the results.

3.5.2 Comparing Deployment Models

Different information can be gathered from the data collected using drawbacks and benefits by combining the resulting tables. One of these concerns the deployment model best suited for each of the categories we stated in Chapter 3. One might for instance be interested in what deployment model would suit a CC implementation which focuses on availability and integrity, and can then assess the table for guidance. Table 4 shows the results of this calculation.

Table 4: Deployment Model versus Categories

	Private	H Priv	Public	Conventional	Best deployment model
Access	19	0	-54	**35**	Conventional
Availability	**34**	-18	-16	-18	Private
Infrastructure	-5	-21	**15**	11	Public
Integrity	**22**	-22	-22	**22**	Private/Conventional
SUM	**70**	-61	-77	50	

A graphical representation of this result can be viewed in Figure 10.

From this information we can see that the conventional server/client model rates higher than all the cloud deployment models in the access category. For availability, the private deployment model clearly wins, rating 50 points higher than the next in line. Furthermore, the conventional models win in the infrastructure category, closely followed by the public deployment model. At last, the integrity category is equally dominated by the private and conventional models.

We find that the conventional and private models reach the top of the list in two of the categories each, while the public deployment model is the victor in one of these. Looking at the combined scores for all the categories for each deployment model, we find that the private cloud has the highest overall score with 70, closely followed by the conventional model with a score of 50. The hosted private and public deployment models then follows, both with negative scores.

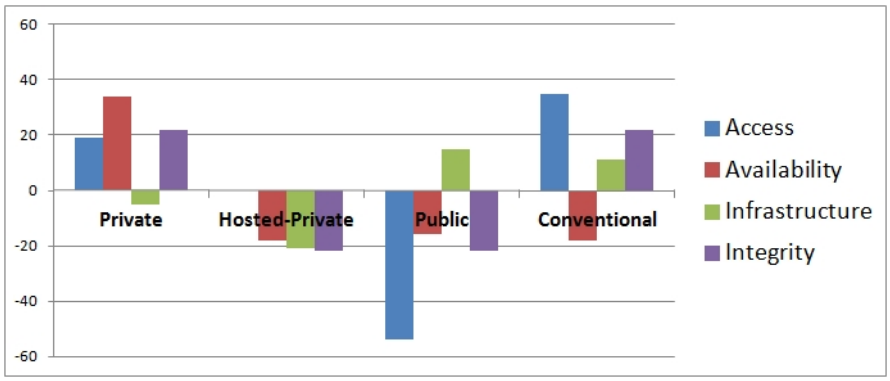

Figure 10: Deployment Model versus Categories

Tables including the hybrid deployment model were also created prior to its omittance, and can be viewed in Appendix D.

3.5.3 Questionnaire

To further explore the choices of impact levels for each of the benefits and drawbacks, a questionnaire was created. In the questionnaire, all the benefits and drawbacks were translated into questions, which resulted in 38 questions. The respondents had to answer each of these to decide on the impact levels. By sending the questionnaire out to a selection of students, we would gain some insight into how our own rating of the benefits and drawbacks matched up with the respondents answers.

The list of questions can be viewed in Appendix F, with a list of four alternatives for each question. Based on the questions content, the alternatives reflect a low, moderate or high alternative. Additionally, an alternative for ”I don’t know” was included. As some of the benefits and drawbacks are complex in nature, the questions tend to be complex as well, resulting in a questionnaire with a fairly low degree of psychological acceptability.

Because of this it was important that the respondents had a strong technical foundation, and preferably some insight into CC. A selection of seven current and former students in the information security track at GUC was therefore made, in addition to a selection of six professors and doctorate stipendiaries. The selection was made in order to decrease the possibility of the questions being misunderstood, which would increase the chance that the respondents could make informed choices.

Questionnaire Results

The question results gathered from Questback using the built-in results generation interface can be found in Appendix G. It presents the answers in percentage for each question, and lets us view what was the most popular answer for each one. Seven out of 15 respondents contributed with their answers.

Looking at the most common answers in each case, we found that 25 % (four out of 16) of benefit answers and 32 % (seven out of 22)of the drawbacks answers coincided with our own assessments.

Another issue that should be noted is that amongst the benefit answers, one ended in a draw with three votes for the medium and high alternatives. Furthermore, amongst the drawbacks answers, four out of the 22 questions ended in a two way draw. Two of

these were between the medium and high alternatives, while the other two resulted in a draw between the low and medium alternatives.

If we further examine the answers data for the benefits, we find that out of the 16 answers which did not correlate with the in-depth study answers, twelve were rated higher (75 %). The same calculation for the drawbacks indicated that twelve out of fifteen answers were rated higher than the original value (80 %).

By generating the same results as in Section 3.5.2 using the questionnaire results, the overview of the best cloud deployment models for each category is changed somewhat. The results are shown in Table 5.

Table 5: Deployment Model versus Categories

	Private	H Priv	Public	Conventional	Best deployment model
Access	37	-46	-44	**53**	Conventional
Availability	**48,75**	-12,25	-10,25	-29	Private
Infrastructure	12,25	-43,75	-7,75	**39,25**	Conventional
Integrity	**26**	-26	-26	**26**	Private/Conventional
SUM	**124**	-128	-88	89,25	

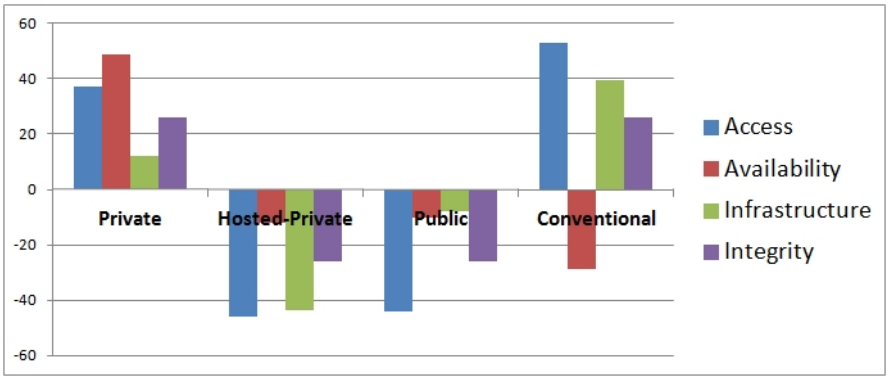

Figure 11: Questionnaire: Deployment Model versus Categories

Here we can see that only the infrastructure category has seen changes in the best suited deployment model. Here the public deployment model has dropped to the third position, and the conventional model has taken its place. The private deployment has risen to the second position in this category. Other than this, the best deployment model results remain the same.

Some changes have occurred to the lower positions. In the access category the hosted-private and public deployment models have swapped places, while the rest of the category remains the same. In the availability category, the hosted-private and conventional models has seen some changes. These were tied for the last position in the in-depth results, but the conventional model seems to have lost much ground, ending up with the least favoured position. In the integrity category all remain the same position-wise.

Scenario Name	Private	Hosted Private	Public	Conventional
Public Information Services	**235**	-159	-223	75
Mobile Information System for Law Enforcement	217	-187	-291	**225**
Internal Law Enforcement Cooperation System	**227**	-226	-238	183
Digital Forensic Analysis System	**203**	-166	-252	179
Public On-line Case Registration	**200**	-144	-284	192

Table 6: In-Depth Study Deployment Model Scores

Looking at the overall scores in the ”SUM” row, we find that the private model still has the best score, even though the conventional model has won one more of the single categories. The conventional maintains the second position in the overall calculation, while the hosted private and public deployment models have switched their positions. This means that the hosted private deploymen model ends up with the lowest overall sum, while the public model takes the third position.

A graphical representation of this result can be viewed in Figure 11.

3.5.4 Calculation of the Best Deployment Models for Scenarios

By using Table 3 which contains the scenario priority values, and the tables containing the category values for each deployment model, we can calculate an overall score for each scenario and deployment model combination. This is done two times: Once for the values collected during the in-depth study (Table 4), and once using the values from the questionnaire (Table 5).

Both calculations have been conducted using matrix multiplication, and an Excel file containing the calculation is described in Appendix E.

Deployment Model Scores based on the In-Depth Study

Table 6 shows that the private deployment model presents strong results for all scenarios. It has the highest score for all scenarios beside the Mobile Information System for Law Enforcement. Here the conventional model has the highest score with 225, eight points more than the private model.

Overall, the conventional and private models have similarly high scores in all scenarios but the Public Information Services scenario, where the private model wins by 160 points. The hosted private deployment model ends up as the third best model in all cases, while the public cloud consistently shows the weakest results.

Deployment Model Scores based on the Questionnaire Study

In Table 7 the private deployment model wins in all categories, closely followed by the conventional model for three of the scenarios. The scores for the Public Information Services scenario stand out again. Here, the private model wins by 252 points over the conventional model which is second in line. The most prominent change in this table when compared to Table 6 is that the public and hosted private models have switched their positions. The results show that the public deployment model show consistently

Scenario Name	Private	Hosted Private	Public	Conventional
Public Information Services	**359,25**	-262,75	-214,75	107,25
Mobile Information System for Law Enforcement	**386,25**	-443,75	-323,75	375,75
Internal Law Enforcement Cooperation System	**410,25**	-453,75	-297,75	333
Digital Forensic Analysis System	**337**	-354	-272	283,5
Public On-line Case Registration	**348**	-374	-290	310,5

Table 7: Questionnaire Study Deployment Model Scores

better results than the hosted private model, thereby taking the third position for all scenarios. This means that the hosted private model has the lowest scores across the board.

3.5.5 Discussion

During this chapter we have assessed a list of drawbacks and benefits from the viewpoint of a governmental agency and performed a calculation of results based on possible implementation scenarios for governmental CC projects.

The chapter started with the collection of the benefits and drawbacks themselves. This resulted in a list of 16 benefits and 22 drawbacks, where two of the drawbacks strictly applied to the conventional, non-cloud model. While we do not propose this list to be complete, it is thoroughly based on the literature that exist in the field. We are confident that the list is a good representation of the benefits and drawbacks that exist.

The noticeably higher amount of drawbacks compared to the benefits could be an indication of the still young technology that CC represents. However, we can assume that as the cloud evolves and improves it will result in the weakening, or even removal of several of the drawbacks. Issues concerning vendor-lock-in, migration of applications and the lack of standards will most likely improve in the near future, thereby removing some significant barriers for the application of CC. This will in turn aid in a strengthening, or possibly growing number of CC benefits, resulting in CC becoming a much better alternative than it is today.

The results of the further study of the rated benefits and drawbacks varied from expected to surprising compared to our initial hypotheses. While, most of the top placing deployment models for each category seem to us like logical choices, the further distribution of deployment models can be a cause of some debate.

For access the on-premise deployment variants topped the scoreboard, most likely because of the focus on drawbacks concerning using the Internet as a communication channel. Furthermore, the conventional model is named the victor because of its non-distributed features. This is because we assessed that the distribution of data on several servers could be seen as a drawback for all CC models if the cloud implementation does not properly segregate clients, which might lead to leakage of data.

One can argument that not all the deployment models are equally susceptible to be severely damaged by leakage events. For instance, the private and hosted-private de-

ployment models would already be segregated to one client, lowering the possibility of the leakage doing significant damage. However, using scenario number four 3.4.4 as an example, the leakage of evidence and analysis results can be severely destructive to a case, even when the leakage is contained within the police force itself, as it breaches the principles of evidence integrity and CoC. In this case it therefore seems that our choice of impact levels is valid, but we understand that this is very specific to each scenario.

Availability was totally dominated by the private model, most likely because it boasts most of the same benefits as the other cloud models in this category, but lacks some of their drawbacks. Problems revolving around CSP errors and colocation of data are non-issues for a private on-premise cloud implementation, giving it the edge over other deployment models in this category. Furthermore, the distributed nature of the private cloud means that it is not as prone to host-related errors as a conventional model, giving it some additional points. We therefore find it to be natural that the private cloud holds this position in the availability category.

In the infrastructure category the results were again dominated by a single actor, namely the public deployment model. Its strong results here are mostly based on the significant amount of resources it can muster at any given time, as a public cloud has fewer restrictions than the comparable models. Highly valued benefits related mostly to its lack of reserved hardware means that it gains a higher score on the benefit side, while sharing many of the drawbacks with the other cloud models against the conventional and private alternatives. Even though it has significantly more drawback points compared to models like private and conventional, these lack the benefit score to take advantage of that fact.

Even though it seems correct that the public cloud has the most to offer on the infrastructure side, this result is probably the one most influenced by our choice to use squared point values. Because of this the drawbacks related to migrational issues, vendor lock-in and environment management loss will not resonate as much as they could. Still, we are content with the impact levels assigned to both the benefits and the drawbacks, leaving the public deployment model at the top of our list.

The integrity category is the only result left in a tie. Here, both the private and conventional models have the same amount of points, leaving these two at the top. However, this is not the only tie noted in this category, as the rest of the deployment models share the second place with minus eight points each.

Even though the gap between the first and second place groups in this category is quite large, we do not feel confident that it necessarily represents the truth. First of all we would have guessed that the conventional server/client model would have ended up at the top of this list, with the private cloud model as a close second. The reason for this is quite simply that conventional computer systems are mature, and their strengths and weaknesses are largely known. CC on the other hand has several areas that still need to be clarified, for instance the segregation of clients and services in software on the same physical device. This should put the conventional model ahead in this category.

We also have some concerns regarding the two cloud deployment models sharing second place. First of all it seems unlikely that the public model should level with the hosted private one, as a fully public infrastructure is widely regarded as the least secure of these due to the exclusive use of third party multi tenant computing power and a communication channel through the Internet. The hosted-private cloud should rank higher,

as its data is always kept on reserved hardware, separated from other clients by tried network segregation methods.

Given the discussion above, we find it likely that not all the sides of the integrity discussion are covered by the benefits and drawbacks used in this study. This statement is backed up by the fact that only one benefit and three drawbacks are noted in this category. More benefits and drawbacks might therefore be found that could differentiate the deployment models better than what our model currently is able to.

The weak results of the hybrid cloud solution are also an issue. As previously stated, results that include the hybrid model can be found in Appendix B. We did, prior to these results, believe that the hybrid cloud model would score much higher on the list, and at least score higher in the integrity category than the public deployment model. This is because of the hybrid models flexibility with placement of sensitive data in a public or private cloud.

The results have most likely been lowered by the fact that the hybrid cloud has been given a lot of the drawbacks of the public and private cloud solution, but missed out on some of the benefits they got based on their non-combined cloud type. Furthermore, many of the benefits and drawbacks that the hybrid cloud model was included in may not have reflected the true impact for the hybrid cloud, as it has received the same impact levels as the other deployment models. Based on this we decided that the model made for assessing the drawbacks and benefits for CC deployment models was unsuited for the hybrid model, and chose to leave it out of the results.

The results of the questionnaire also featured some interensting findings. As stated in 3.5.3, three out of four of the best deployment models correlated with the in-depth study results. This seems to partially verify the results from the in-depth study, however, 25 % correlation of benefit results and 32 % correlation for drawback results cannot be said to be conclusive. Another result to be noted is the fact that the answers that differed from our assessments were generally rated higher than the original values. This amounted to 75 % for benefits and 80 % for drawbacks.

This indicates that the questions asked in the questionnaire were generally perceived as highly beneficial or highly non-beneficial for the benefits and drawbacks respectively. In this case, the language used to illustrate the questions might have unintentionally influenced the respondents to perceive the benefits and drawbacks as stronger than the reality should suggest. The results could also indicate that the in-depth study has not covered all aspects during the assessment of benefits and drawbacks.

Reaching a definitive conclusion concerning which of these suggestions, if any, are correct is not easy with the information we currently possess. In order to extract more information from the questionnaire, one could perform in-depth interviews with some of the respondents. This would give us insight into why they answered the way they did, and what assumptions they made prior to their answers.

Finally, due to the limited number of respondents, it is challenging to draw conclusive results. However, the data collected provide indications of what may be the case, and supports our other results to an extent. The results can therefore act as a foundation for further studies, or as a base-line for a more comprehensive study of the same type.

At the end of this discussion section we have the results gathered from combining the best deployment model tables (4 and 5) with the scenario category priority table (3) using matrix multiplication. Based on the results found in Section 3.5.4 the only possible

cloud choice would be the private cloud deployment model. The model boasts the best scores of all the cloud models in all scenarios, leaving the remaining cloud deployment models as unattractive alternatives. This is the case when using the in-depth study values as well as the questionnaire values. The result coincides with our hypothesis, as it stated that the private cloud would do better than the hosted private and public alternatives.

One can argue that some of the scenarios would be be better off using other environments. As an example, in scenario number one (Section 3.4.1), the public information system needs to uphold contact with the population, thus the need for system redundancy in the case of errors is very important. It also needs to distribute all its information via the Internet. A public or hybrid version could therefore work better for filling its requirements. However, such an implementation would include a potential external single point of failure, the CSP. This would come in addition to the ISP, which will remain a potential single point of failure regardless of the chosen cloud model. Such issues coupled with the fact that it can be hard to prove the security of a public cloud suggests that the private cloud or a conventional model is the better choice.

The conventional model presents very strong results in most of the scenarios presented, suggesting that it still serves as a better choice than most cloud models. In the cases where the private cloud model provide significantly better results, it is mostly caused by the conventional models weak potential for availability combined with a scenario that prioritizes this category. This is the case with the public information services scenario.

Given the overall strong results, it is reasonable to say that the conventional model is be a a good choice. When we add these results to the reputation of such a platform, which is proven over time to be predictable and reliable, possible clients may still choose the conventional model over novel cloud technologies. Even though this is the case we find it highly probable that the cloud results will strengthen over time, as the technology matures. This will also make the additional cloud deployement models better choices.

4 Requirements for a Governmental Law Enforcement Cloud Environment

In this chapter, a series of requirements for a cloud implementation seen from a governmental and law enforcement viewpoint will be presented. The requirements will in turn be arranged in a document that constitutes a complete requirements document and guide for the planning of CC environments. Planning documentation created by consorting the guidance document creates a good foundation for the implementation of a CC solution. The resulting document is created to be framework independent, but examples from existing frameworks and methodologies are given where appropriate.

Each section of the chapter will explain different requirements and why they should be included in the resulting document. The sections will also state specific steps that should be implemented in the document based on its content. These propositions will be marked with **"Additions to the Requirements Document"**. At the end of the chapter, a disposition of the requirements document is given. In addition to this, the complete requirements document is included in Appendix H. This constitutes a guide with descriptions and examples for each step.

4.1 Setting the Stage: Digital Forensic Analysis System in the Cloud

It is very hard to assess anything from a strictly generic standpoint, something that also applies to cloud implementations. To elaborate on the importance of setting the stage for such a study, a quote from NIST's special publication 800-144 can be used:

> *"The outlook on cloud computing services can vary significantly among organizations, because of inherent differences in such things as the intended purpose, assets held, legal obligations, exposure to the public, threats faced, and tolerance to risk. For example, a government organization that predominantly handles data about individual citizens of the country has different privacy and security objectives from a government organization that does not. Similarly, the security objectives of a government organization that prepares and disseminates information for public consumption are different from one that deals mainly with classified information for its own internal use. From a risk perspective, determining the suitability of cloud services for an organization is not possible without understanding the context in which the organization operates and the consequences from the plausible threats it faces."* [39].

To obtain a sense of purpose we can apply a setting or scenario to create reference points. Based on the scenarios created in the previous chapter, in section 3.4, we can assess what a cloud needs to deliver to fill its requirements. For this chapter we have chosen to use scenario number four, Digital Forensic Analysis System 3.4.4. Such systems are some of the most sophisticated in a police forces digital suite, and they specialize in analysis and storage of highly confidential information. The information has to be handled securely and in coherence with the principles of CoC and evidence integrity.

Implementing such features in a cloud environment will be daunting, as information like the results of an analysis, the evidence itself or metadata concerning the evidence should under no circumstance be exposed to environments where unauthorized personnel could access it.

Even though we will primarily be using scenario number four as an example throughout this chapter, we will in certain cases provide other examples where we find that this is appropriate.

Additions to the Requirements Document

To properly explain the project to be implemented we propose that sections for problem description and scope should be added to the requirements document. The problem description should explain what problem the project is attempting to fix or mitigate, and the scope should limit the project to a specific area. The scope could contain the entire problem or part(s) of it. A section stating the goals of the project should also be created, as it will further specify what the project is attempting to accomplish in order to solve the problem area defined in the scope.

In addition to these steps, a section outlining the possible benefits and drawbacks of the proposed solution in relation to existing or alternate implementations should be created, as it forces the project to assess alternatives and possible outcomes early.

Both of these proposals have been added into the requirements guide(H.1) under step 1.(a).

4.2 High-Level Requirements

When specifying the project further, it is important to start at the high-level requirements of the project and then further explore to a high-level of granularity. This type of top down approach lets us first explore the large building blocks of a planned system, before breaking these down into their smaller parts. The two subsections of this section present some examples on how to map high-level requirements.

4.2.1 Users and User Stories

By exploring the types of actions that the users of the proposed system would most likely perform, the high-level requirements of the system can be mapped. User groups, or actors in the system, should be explored, and the different user groups should be mapped to the main activities they would like to perform on the system.

One way of mapping such activities is to make text based explanations of the requirements called user stories [93], often used in agile development frameworks. A way of creating a user story can for instance be: ”I, as an ACTOR, should be able to TASK because REASON”. An example of this is: ”I, as an Administrator, should be able to log into the system because I need to be authorized before I perform other tasks”.

The user stories should be created in dialogue with the potential users of the implementation, as these are the ones that know what they need. The stories should not include how the goal of the user story is reached.

Additions to the Requirements Document

Based on this we propose that a section for mapping user groups and their user stories should be added to the requirements document. This has been added to the guide(H.1)

in step 2.(a)i. Examples of tables for user stories and groups can be found in Section 5.3.1 Tables 9 and 10.

4.2.2 Functional and Non-Functional Requirements

Based on the users and user stories found in Section 4.2.1, a list of high-level functional and non-functional requirements can be created.

Functional requirements are the functions or services in a system that supports the user's goals, tasks or activities [94]. An example could be the ability to log into the system, or, in the case of scenario number four, the ability to load evidence for analysis.

Non-functional requirements can be defined as restraints or qualities [94]. Qualities are characteristics of the system that are coveted by the projects stakeholders, for instance performance. Constraints are set limitations of the system that need to be complied with, and thus cannot be omitted. One example is that the system cannot use a mouse for I/O.

Some other, and more elaborate examples on non-functional requirements can be given that are more focused on law enforcement.

Cross-Country Coverage

Governmental computer systems can be constructed to provide services for local agencies or a nationwide initiative. This will impose different demands on the system to secure transactions between internal and across unsecured networks.

An online case registration system would for instance be exposed to the Internet to enable the public to access it and register cases. It would then be natural that this service would be available to the entire country.

Cross-Jurisdictional Abilities

Many initiatives like Europol and Interpol necessitate the use of cross-jurisdictional computer systems that comply with the laws and regulations of all included jurisdictions. This will greatly enhance the reaction times of investigators by cutting through the red tape using pre-defined agreements rather than relying on a swift handover of case material, custody and responsibility.

Databases maintained by Interpol can serve as an example [95]. Here, the 190 member countries have the ability to access global databases with information concerning most areas of law enforcement without having to start a paper trail and complete an entire authorization process every time. Access is made available through the secured I-24/7 network, which links together the National Central Bureaus (NCBs) of each member country.

In scenario four, information will not flow across borders. However, one must accept the possibility that the information might end up in a global database like the ones used by Interpol. It is therefore of the utmost importance that the information derived from the analysis of evidence is validated and handled according to the principles of CoC and evidence integrity.

Throughput/Performance

Performance is never an absolute concept, and different systems will have different ways of defining good or bad performance. The term is most often used to describe either the

amount of tasks the system can handle in a certain time period or the speed at which each single task is completed. The threshold for good or bad performance should then be based on the feasible results obtainable by using the given hardware, or by the usability and appearance for the users of the system.

Given an example like scenario number five, an on line case registration system, the system should be able to register a large amount of cases from different parts of the country of origin without failing or generating long waiting periods. In the example of Norwegian law enforcement, 394 000 individual cases were registered in 2011 [96]. This amounts to about 1080 cases a day, or a little under 45 cases per hour in Norway. Even though the pressure on the system will be higher at certain times of the day, or days of the year, the numbers are not high compared to the numbers from larger countries. A goal for the performance of such a system in Norway could therefore possibly be 24 hour access to the system through the Internet, and zero fails during case registration throughout the year due to system overload.

Additions to the Requirements Document

We propose that a section for both functional and non-functional requirements should be included in the requirements document. These steps were added in steps 2.(a)ii. and iii. of the guide.

4.3 Detailed Requirements

After the high-level requirements have been mapped and agreed upon, detailed requirements can be created based on the documentation. Here the findings from the high-level requirements study are both linked together to explore interconnectivity between functionality, and broken down into detail.

4.3.1 Use-Case Diagrams

By using use-case diagrams we can present the different use-cases and how these are connected. This enables us to further explore the requirements.

Figure 12 shows an example of a use-case diagram for the fictive forensic analysis system in scenario four. The word «extends» on some of the edges signifies that when a parent use-case is performed, any of the child use-cases may be performed. In addition, Table 13 in Chapter 5 presents an advanced use-case description and flow of events.

We identify two different users, or actors of the system, administrators and case workers, both members of a given jurisdictional authority.

The administrator will be able to log into the system, and, in this case, only have one main task, managing the users of the system. He currently has four ways of doing this: List, which returns a list of all the users entered into the system; Add users, which allows the administrator to add new user accounts; Remove users, which gives the administrator the possibility of removing already authorized user accounts from the system; and Manage privileges, which allows the administrator to configure specific attributes of the account, like access to different analysis methods and the disabling and enabling of accounts.

Authorized personnel will also be able to log into the system, and conduct one main task, analyzing case data. This use-case has three extensions: uploading case data, or evidence; choosing an analysis method to conduct; and extracting the results of the ana-

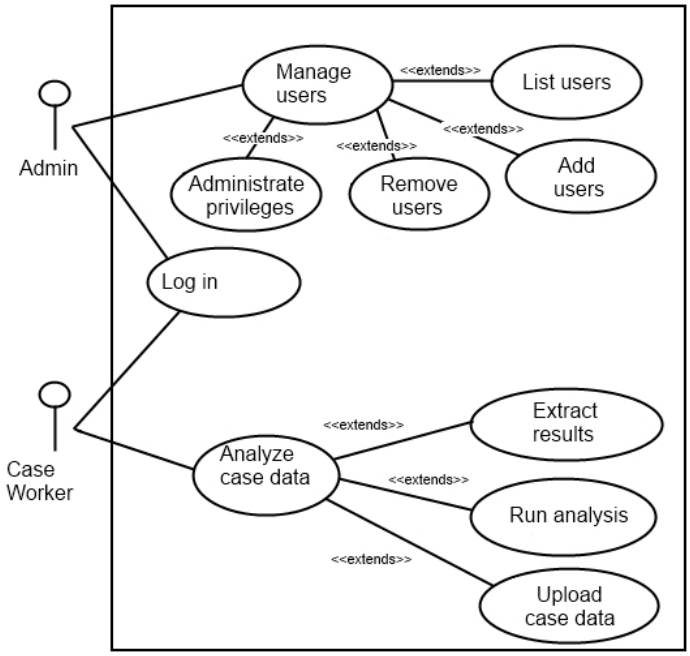

Figure 12: UML Use-Case diagram for the forensic analysis system

lysis.

Additions to the Requirements Document
A section for use-case models, capable of mapping the functional features of the proposed implementation in a project, should be included in the requirements document. This can be found in Step 2.(b)i. in Appendix H.1.

4.3.2 Data Flow

Mapping data flow is another good way of visualizing a systems components and their interoperability, mapping where data flows during operation. It is an important step towards identifying areas that need to be additionally secured and avoid possible leakage of information into other parts of a system or out of the systems boundaries.

For CC this can include the flow of data from an internal private system to public infrastructure given a hybrid deployment model. Similarly a private cloud could use both on-premise resources and a hosted private cloud situated at a CSP, making the same problems relevant for private implementations.

Data flow can be based on use-cases, mapping these with more detail, and visualized using data flow diagrams (DFDs). At the lowest level, a DFD shows how output values are derived from input values [97], but it does not need to be presented at this granularity. Flow charts can also be used to represent a use-case at a detailed level, making it easier to visualize functionality. Figure 13 presents an example of such, by mapping a login function.

Additions to the Requirements Document
A section for data flow mapping should be included in the requirements document to

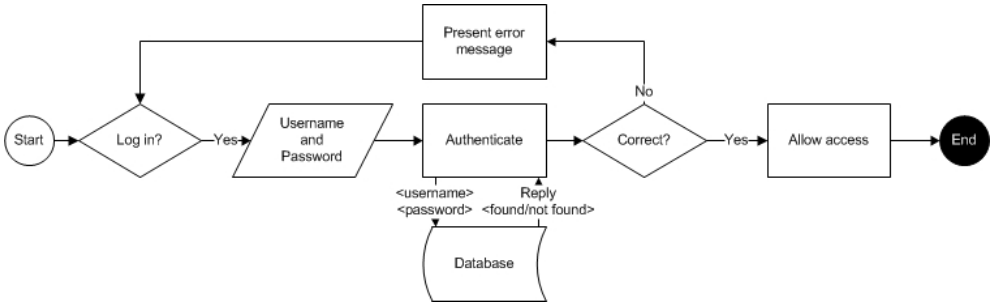

Figure 13: Flow chart visualizing log in functionality

Asset Name	Asset Type	Description
Case data	Data	Evidence files to be analyzed by the system
Upload case data	Process	Uploading the case data to be analyzed
Run analysis	Process	Conducting the analysis of the case data
Extract results	Process	Retrieving the log from a performed analysis
Result log	Data	File containing the results from the analysis
Log in	Process	Authorization process conducted to gain access to the system
List users	Process	Listing the user accounts
Add users	Process	Adding one or several new user accounts to the system
Remove users	Process	Removing one or several existing user accounts from the system

Table 8: Asset list for a digital forensic analysis system

assist in the detailing of the high-level requirements. This is added into Step 2.(b)ii. in the Appendix H.1.

4.3.3 Assets

Assets also need to be mapped prior to a cloud implementation, and can basically be split into two categories: data or applications/functions/processes [40]. The construction of use-cases and data flow models create a good foundation for exploring the assets of a system.

As an example we have noted nine assets concerning our scenario, which are deduced from the use-case diagram and presented in Table 8. We do not propose that the list is complete, or that this is the only correct way to map assets, but rather samples of possible assets and a presentation method.

Additions to the Requirements Document

By mapping the assets we now have a sufficiently detailed list that represents the projects proposed implementation, giving us a foundation on which to assess possible cloud environments. We therefore propose that a listing and description of assets should be included in the requirements document. This has been implemented in Steps 2.(b)iii., iv. and v.

4.4 Assessing Cloud Applicability

A detailed study of the requirements of the system should now exist, and based on these it is possible to assess the applicability of cloud environments. This study should be split into three different studies: choosing a cloud service model, choosing a cloud deployment model, and choosing a CSP. Keep in mind that an initial legal 4.5 and compliance 4.6 study should also have been conducted prior to this step 4.9. The results of these studies will also have to be taken into account when conducting the studies concerning cloud applicability.

4.4.1 Cloud Service Model Study

The first cloud factor that has to be discussed is the service model. These are usually divided into SaaS, PaaS and IaaS as discussed in Section 2.2.4. The assessment is important because its results might create several restrictions for the remainder of the project. Choosing a SaaS model would of course limit the selection of CSPs (as would any other service model choice), but it would also limit the amount of control the stakeholders of the project would have over the security and manageability of the completed system. It would also limit the choices for the developers of the project, as they will be limited to the development functionality present in a SaaS CSP. On the other hand, choosing an IaaS model would leave most options other than CSPs wide open. The choice of a service model should be based on the lowest level of control the client needs in the cloud stack order to use or create a service(ref. Figure 3). For instance, if a client only needs to use a prespecified application, a SaaS solution may be reasonable. On the other hand, if the client wishes to deliver SaaS or PaaS services, control over the Virtual Infrastructure is required. This would demand an IaaS solution.

Additions to the Requirements Document

A study of the applicable cloud service models should be conducted, where each alternative is assessed against the requirements and assets found in prior studies. This has been implemented in Steps 3.(a), (b) and (c).

4.4.2 Cloud Deployment Model Study

The cloud deployment models, explained in Section 2.2.5, are in most cases applicable to all the cloud service models. These will therefore have to be assessed based on the same criteria as in 4.4.1, and some of the same limitations can be created based on the choice. For instance, by choosing a public cloud infrastructure, one will limit the possibilities of securing the implementation at hardware levels. Likewise, choosing a private model could restrict the flexibility and scalability characteristics of the implementation.

Additions to the Requirements Document

Applicable cloud deployment models should be defined and discussed against the requirements and assets, and has been added to the guide as Steps 4.(a), (b) and (c).

4.4.3 Cloud Service Provider Study

Based on the results of the studies in 4.4.1 and 4.4.2, the choice of a CSP will be limited to some degree. The project group will have to define possible alternatives that suit all the criteria in the prior studies, discuss those alternatives and make a choice. The study may consist of as little as reading the technical documents of the CSP, and making a

choice based on this, or arranging meetings with CSP officials.

If one of the requirements of the project is that an open source alternative should be used, support from the vendor will not always be available. In these cases it is up to the project group to collect useful information from the vendor's website and possible developer and community forums.

For a government project of the magnitude specified in our scenarios 3.4, it would be advisable to arrange meetings with possible CSPs (if at all possible), as this would create great opportunities to thoroughly assess the CSPs abilities to comply with the requirements of the project. Guided inspections of a CSPs data centres is also advisable if public, hybrid or hosted-private deployment models are chosen as alternatives.

In addition to assessing CSPs based on the requirements, one can assess cloud alternatives against scoring frameworks. Popular advisory firms like Forrester and Gartner can provide their customers with such criteria, which can be presented to the CSP for assessment

Additions to the Requirements Document

A thorough study of all CSP alternatives should be conducted by the project group, making it possible to make an informed decision. In Appendix H this is done in Steps 5.(a), (b) and (c).

4.5 Legal Issues

Before establishing or migrating to a cloud environment, the legality of the transition should be explored to ensure legal compliance. This should be done as part of the due diligence a governmental, law enforcement, or any other entity has to perform. The privacy of personally identifiable information (PII), health care information etc. are critical points of interest, also for the areas of governmental and law enforcement computing. Many regulations and laws cover the electronic handling of such content.

A big issue for CC is that no laws, regulations or even best practices have so far been created (although some compliance organizations, new as well as old, are attempting to become such 4.6), leaving us to apply older regulatory demands to CC. This is not always entirely possible, leaving CC in somewhat of a jurisdictional limbo.

In this thesis we will base the discussion on European Union and US laws and regulations. We will also take into account the cross-jurisdictional implications is often encountered by CC. For instance, a cloud provider residing in one jurisdiction will often be required to follow several sets of jurisdictional regulations in order to provide services to clients on both sides.

4.5.1 European Union Regulations

In Europe the Data Protection Directive (95/46/EC) [98] and ePrivacy Directive (2002/58/EC) [99] are the main laws covering information kept in digital information systems.

The Data Protection Directive is created to protect the individual when processing or moving PII. Important articles that can be noted is; article six, which elaborates on how and when information can be gathered lawfully for a legitimate purpose and the quality of the data, article 12 concerning a subjects right to insight in his/her PII content and processing, and article 14 covering the users right to deny the collection and processing of data in certain cases.

The directive states, in article four, that the provisions adopted through it should be applied together with the national laws in the countries where the information is processed and the country of origin, unless the origin is outside of the EU. The directive is not applied where information is present on equipment used simply for the transport of information. Additionally, article three states that the directive is not to be applied in cases of national and public security and criminal law. This means that some government tasks, including our digital analysis system is exempt from the directive. As a result it does not cover cloud providers situated in countries outside of the EU with the result that the US Patriot Act will allow the US government insight into the data handled by the US based cloud providers.

The ePrivacy Directive 2002/58/EC aims to harmonize the level of protection for all member countries in the context of privacy. It complements Directive 95/46/EC, thereby following, amongst other things, the jurisdictional limitations specified. Directive 2002/58/EC elaborates on some sections that were not included, or covered sufficiently in the directive from 1995, including confidentiality, billing and traffic data [100].

Because of the dated regulations that are currently active, and the increasing number of supporting directives (like 2002/58/EC) that over the years has been made to make the law applicable to the new technology and trends, the EU Commission in 2012 proposed a complete reform of the EU data protection, collecting several laws and regulations under one law [101]. Special attention is also given to technology advancements like CC [102], which is used as a prime example for the potential of elevated cooperation within the member countries and against non-members. The main pointers of this part of the reform are to ensure that EU rules will apply also for non-members when using information originating from the EU. It is also stated that the directive will include regulation of police cooperation and criminal justice, but no further information is given.

An EU-directive that does cover some of the implications of law enforcement agencies handling PII and other sensitive information is the Framework Decision 2008/977/JHA. The directive is created to ensure the protection of the rights and freedom of private persons, especially privacy, when processing personal data for police and judicial cooperation in criminal matters [103].

2008/977/JHA limits its demands to data protection under and after transferring data between EU member states, for instance the processing of such data at the receiving end. This is for instance constantly done during cases involving Europol. Note that the directive does not state any demands for the handling of evidence or data in criminal investigations limited to a single member state. Important sections of Framework Decision 2008/977/JHA concerns who has the right to handle the data (Article 3), logging and documentation of transfers and processing (Article 10), for what purpose the data may be processed (Article 11) and security demands for the processing of such data (Article 22).

Framework Decision 2008/977/JHA have been active for some time, but it seems that proposals are now being created to replace the directive. In 2011 a draft for a new directive, ”The Police and Criminal Justice Data Protection Directive”, was leaked according to out-law.com [104]. The draft proposes to restrict the situations in which personal data can be collected and processed for law enforcement purposes, and include demands for strictly domestic collection and processing as well as cross-border evidence handling. This specific directive has not been confirmed by any EU-commission officials, but an EU-

commissions press release proposing to reform several of the data protection directives in use today stated that:

> *"A new Directive will apply general data protection principles and rules for police and judicial cooperation in criminal matters. The rules will apply to both domestic and cross-border transfers of data."* [101].

Finally, Council Directive 89/391/ECC [105] may be relevant for cloud implementations, as it concerns an employers duty to ensure the safety of his/her workforce. Given in article one of the directive, the member countries have to take measures to avoid occupational risks and encourage improvements to their workplace. Risk assessment and policy management are important processes that can assist in ensuring workplace safety, and as IT systems might leak PII to unauthorized actors, the safety of employees will be relevant here as well.

4.5.2 US Regulations

In USA, several laws and regulations applies, where the most prominent in the later years has been the Gramm-Leach-Bliley Act(GLBA)[106] and the Health Insurance Portability and Accountability Act(HIPAA)[107].

The GLBA is primarily aimed at financial institutions and services, and covers the demands for the security of sensitive data and their customer's rights to information regarding the use of such. It contains many guidelines stating how the affected companies should generate privacy notification documents informing customers of what information is gathered and how it is used and (possibly) shared. It also elaborates on the customers rights to deny the use of their information in some or all of the intended ways. It also includes important regulations for how people receiving sensitive information by a mistake can or cannot disclose that information, keeping them responsible for any illegal disclosure.

In addition to the GLBA, the Sarbanes-Oxley Act regulates financial institutions in the USA [108]. The act will most likely impact CC implementations used by financial institutions, and can also impact European based cloud solutions with American corporate clients, as compliance has to cover all parts of a business. CSPs will therefore be forced to comply if they want to deliver services to the American market.

The HIPAA concerns the collection and use of protected health information(PHI), especially concerning the disclosure and sharing. The main goals are to prevent health care fraud, and simplify the administration of PHI. The direct effects of HIPAA was that patients now had the right to be informed and consent to a hospitals privacy policy, which had to be created for all institutions affected by the directive. Several new patient rights were also created, including the right to demand copies of your medical records, and place restrictions on who will be able to access these. NIST SP-800-66 has been specifically created to provide guidance to companies or government agencies implementing HIPAA [109].

Several other state laws apply depending on the placement of a CC environment's physical equipment. Laws like the Massachusetts Security Regulations are examples of state laws that apply in this case. The law applies to all businesses that store or handle PII concerning a Massachusetts resident, regardless of the place of the business[110]. This basically means that the law applies internationally. The law demands that the affected

businesses implement an information security program, and that all PII covered by it is secured by encryption.

In a speech made by Peter Hustinx, the European Data Protection Supervisor, during the third European Cyber Security Awareness Day the following was stated: ”A cloud provider established in the EU - or acting as processor for a controller established in the EU - will in principle be ’caught’ by EU law.” ”A cloud provider which uses equipment (such as servers) in an EU Member State - or acting as processor for a controller using such equipment - will also be caught.” Finally, ”A cloud provider in other cases - even if it mainly and mostly targets European citizens - would not be caught by EU law.” [111] This is in contrast to the Massachusetts law stating that all must comply, regardless of placement.

We understand that several additional laws and legislations might apply, including for instance the following which are noted in [40]:

- Clinger-Cohen Act of 1996
- OMB Circular No. A-130(Appendix 3)
- The Privacy Act of 1974
- The E-Government Act of 2002
- FISMA 2002
- NARA regulations(Federal Records Act (44 U.S.C Chapters 22,29,31,33)(Title 36 Code of Federal Regulations, Chapter 12, Sub chapter B)

Additions to the Requirements Document
Regardless of the laws that apply, we propose that an assessment of such is conducted during the planning process. Assessments should be conducted both in the preliminary stages, and later on when the project is defined in detail. The initial legal study is included in the guide as Steps 1.(b)i. and ii., while a thorough legal study can be found in Step 6.

4.6 Compliance Demands

Next after laws and regulations, compliance standards are the greatest regulatory force. Such standards are often created as an answer to laws, which constructs a common ground on which the affected companies or governmental agencies can rely for guidance. This will help organizations to abide by the laws, directives and regulations of their respective fields. Together with best practices, compliance standards eases the transformation of existing systems, assists in building new compliant systems and organizations from the ground up and makes it significantly easier to audit these through standardized methods and techniques.

According to the Security Guidance for Critical Areas of Focus in Cloud Computing V3.0 made by the Cloud Security Alliance, few regulations have been created specifically to cater for virtualised or CC based systems [40]. Before these are created and made available, makeshift modifications and interpretations of existing laws, regulations and best practices has to provide the foundation for compliance. This makes documents like the NIST Guidelines on Security and Privacy in Public Cloud Computing (SP 800-144)

[39] important to be able to navigate through the forest of legal and compliance demands.

Some compliance assessment organizations have modified or created documentation and standards to assist in the implementation of CC. Others should still be complied with even though they were created ahead of CC. This section will cover some of these.

4.6.1 Daubert Criteria

In light of our current scenario, the Daubert criteria is important to consider. Daubert's four factors for admissibility of expert testimony [112], presented during a 1993 US court case [113], currently sits as one of the primary standards for ensuring the credibility of evidence.

Daubert states that the scientific method used for the analysis must support hypothesis testing. One should then be able to refute or prove the hypothesis based on testing. The next criteria states that the error rate of the scientific technique should be known and verified. This will provide the court with a clear statement of the likelihood that the method will produce erroneous evidence.

Third, the techniques used should be subject to peer reviews and be published. This will allow the technique to be scrutinized and thoroughly contested to prove its applicability. Finally, the forth criteria states that methods should be generally accepted in the scientific community. If the collected scientific community in the field of interest finds the method to be trustworthy, it strengthen the possibility that this is the case.

The users of the forensic analysis system should be able to provide such information, should evidence analyzed by it have to be presented in a court of law.

4.6.2 FedRAMP

The Federal Cloud Computing initiative developed the Federal Risk and Authorization Management Program (FedRAMP) as a risk management program for the US government focusing on securing CC systems [114]. It provides a standard complete with a set of controls for security assessment that is used to perform conformity assessment of for instance CSPs. The list of controls consists of 17 categories, each with a number of individual controls which includes parameters for the requirements and guidance. This list can be freely downloaded and reviewed [115]. Specific control categories that can be noted are access control, audit and accountability, identification and authentication and system and communications protection.

By demanding CSPs to use qualified (accredited by FedRAMP) third party assessment organizations (3PAO) to perform initial and periodical assessments of the cloud environment, compliance is demonstrated and maintained. The 3PAO is required to comply with the requirements of FedRAMP as well as the requirements for assessment organizations stated in ISO/IEC 17020:1998.

4.6.3 ISO/IEC 27001/2:2005

ISO/IEC 27001:2005 (hereby referred to as ISO27001) and ISO/IEC 27002:2005 (hereby referred to as ISO27002) concerns the creation and management of information security management systems (ISMS), where ISO27001 provides a specification for an ISMS while ISO27002 is the related code of practice and framework for best practice in information security. These were derived from the preceding British Standards BS7799 (ISO27002) and BS7799-2 (ISO27001), in order to redefine the standards for interna-

tional use[116].

ISO27001 has the goal of providing information security personnel with the tools necessary to identify and mitigate the possible risks that poses a threat to the information assets of a given system. The standard can be used to, amongst other things, formulate security requirements and objectives, perform risk management and capturing and assessing status information from the target system [85]. Compliance with ISO27001 will also aid the organization in complying with the demands set forth by regulatory directives such as the Combined Code, the Turnbull Guidance and SOX.

ISO27002 contains the guidelines and best practices that can aid information security personnel in reaching the goals presented in ISO27001. Spanning from the creation of security policies and access control to asset management and incident response [86], the standard covers a broad spectrum of control objectives and security management areas. ISO27002 also introduce a plan, do, check, act (PDCA) cyclic process, with a list of sub-points, for structuring the implementation of the standard.

Even though it is not specifically created with CC in mind, the security principles of ISO27001 and guidelines of ISO27002 can be of great help for securing any cloud implementation. They are also fully implementable, proven by the certification of several major cloud providers like SalesForce.com and Amazon Web Services.

4.6.4 Architecture: Common Criteria

The Common Criteria is developed as a collaboration project between many of the most influential countries in the world, including the UK and USA [117]. It is created under the auspices of ISO, to make sure that it has the quality to be instated as the standard security evaluation criteria framework. Because of this, the Common Criteria is now also known as ISO15408. The standard is capable of evaluating a wide area of components in the IT infrastructure, and defines its potential targets as "... a set of software, firmware and/or hardware..." [118]. Examples given in the documentation range from evaluation of an operating system to smart card integrated circuits.

Common Criteria consist of three parts [119]:

Part 1, Introduction and general model, introduces Common Criteria, and defines templates for security objectives, requirements and high-level specifications for IT systems.

Part 2, Security functional requirements, presents tools for documenting the security functional requirements of systems to be evaluated.

Part 3, Security assurance requirements, defines assurance requirements and presents the evaluation assurance levels used in Common Criteria.

For developers, Common Criteria can assist in the mapping and evaluation of targets of evaluation, or TOEs, document these in a security target (ST) and use it to prove compliance with security requirements and/or user/client security demands. It also emphasizes on the comparability of evaluation results within an IT system and between similar systems, proposing standard ways to evaluate and rate TOEs.

4.6.5 Architecture: Neumann

In 1975, Saltzer and Schroeder wrote the paper ”The Protection of Information in Computer Systems” [120], containing eight design principles for strengthening protection mechanisms in computer systems. The principles apply both to software and hardware implementations, focusing on simplicity and restrictions. The philosophy is that simple systems are less prone to error or vulnerabilities, while focus on restricting errors lowers the chance of unauthorized entities gaining access.

The Saltzer and Schroeder principles created a foundation for secure design and development, and were later explored and expanded by Neumann in the study ”Principled Architectures for Survivable Systems and Networks” [121]. The work by Neumann focused on the survivability and security of systems and the mitigation of threats, and presented 16 security principles, where some of the original principles where left out, redefined or split into additional principles. In addition, some new principles where introduced. For a detailed description of the principles, see [121].

The principles are neither meant as a standard nor warrant any compliance demands, and are created strictly to grant guidance throughout the development and use of a computer system. It elaborates on the main focus points of the implementation of a computer system, emphasizing their importance, thereby leaving the reader to find the best suited methods or tools needed to secure the system.

4.6.6 NIST SP800-27

SP800-27, or the ”Engineering Principles for Information Technology Security”, also called EP-ITS for short, presents system-level security principles created to strengthen the security of existing systems or throughout the development and life-cycle of new ones [122]. The principles are created for use by federal agencies in the US, but can be voluntarily implemented by non-governmental initiatives. Its content is primarily based on NISTs SP800-14 and the Common Criteria.

EP-ITS consists of 33 security principles ranging from the creation of a security policy to the utilization of open standards for migrational and interoperability purposes. All the principles are discussed in relation to a list of life-cycle stages, which states when and to what extent the principle should be applied in the life-cycle. The life-cycle phases, defined in SP800-14 are the initiation phase, development/acquisition phase, implementation phase, operation/maintenance phase and disposal phase. The phases are applicable both to complete system implementations and single assets within a system.

4.6.7 PCI DSS

According to the documentation of the standards latest release, the Payment Card Industry Data Security Standard (PCI DSS) offers technical and operational requirements for the specific purpose of securing and protecting card holder information [123]. The standard was created as a cooperation project between many of the major actors in the payment card industry, including VISA and MasterCard. PCI DSS applies for all entities that handle, store or transmit such information in some way, and must be followed for all parts of the complete system handling the data.

The standard consists of twelve requirements, each with an (often substantial) amount of sub-requirements that must be followed to be in compliance with the standard. Businesses handling card holder information at a grand scale have to be audited by third party assessors, while actors operating at a smaller scale can assess themselves for com-

pliance through a self-assessment questionnaire. VISA and MasterCard both demand that their partners are compliant with PCI DSS.

Additions to the Requirements Document

Regardless of which apply in different cloud projects we propose that compliance studies should be one of the requirements in the document. These should be conducted both as preliminary studies and as a complete study later in the project lifespan, in the same way as with legal studies. In the guide (Appendix H) the initial compliance study can be found in Step 1.(c), while the steps of the thorough study is found in Step 7.

4.7 Security Demands

The security demands of a computer system should be based on several criteria, for instance: who will be using the system, internal personnel only or external personnel as well?; what is the nature of the data to be stored or handled in it, publically available and open data, or sensitive PII?; where is the system placed physically, in a server behind its owners firewall or in a public cloud? The answer to these, and many other questions has to be found in order to produce a complete profile of the security needs and strategy for any given computer system.

With CC, the planning of security may become more complex. Characteristics like multi-tenancy, cross-organizational clouds linked by unsecure lines, and sensitive data stored by public cloud providers make the environment to be assessed more complex. Existing security assessment techniques has to evolve to support cloud implementations. This section will cover some of the available techniques for such security work.

This security demands section is divided in the following way.

4.7.1 Securing Data

The most prominent part of a system that needs to be secured is the data it contains, both information stored as part of a systems main functionality (the information in a digital case file would be an example from scenario number four) and data concerning the users of the given system. Regardless of the type of data, it all follows the same basic life cycle during its presence in the computer system.

One representation of this is the data security life cycle [124], which follows data from its creation to its destruction. It enables one to assess when data is exposed to security threats, and take the necessary precautions [40]. The stages of the life cycle are as follows:

1. Create: The generation of new, or modification of existing digital content
2. Store: Retaining the data in storage for later use
3. Use: The viewing or processing of the data in ways other than modification
4. Share: Making the data available in some way to personnel other than the existing owner of the content
5. Archive: Long term storage of data after it's removed from active use
6. Destroy: The act of utterly destroying data in a secure way, making sure that it is not recoverable by any means

As well as determining where in the data security life cycle data is currently placed, its physical location and sensitivity level is of importance. Data kept on a device in an internal network has different needs for security than data residing in a public cloud. For instance, a sensitive piece of data can be generated at an external computer (for instance at the home of personnel), then stored at work. After a while it is moved to the companies cloud solution to be used actively in functionality, and then archived in a storage system linked to the cloud for possible reinstation at a later time. Finally, when it is decided that the data is no longer needed, it is destroyed, removing all traces completely.

In total the asset in our example has been transferred three times, and traversed four different systems in its lifetime. At all these systems the data may still be present, starting a new life cycle, and during its transportation (using mobile storage devices or networks) the data can be picked up and stored by both malicious and benign monitoring equipment. Making sure that data resides in the correct places, and only in those places, then becomes a genuine problem.

The CSA Guidance documentation covers many of the important cases for CC in which data should be secured, the first of which is hindering unnecessary movement of data to the cloud. Detection of such movements could be possible using database and/or file activity monitoring (DAM/FAM), and hindered using data loss prevention (DLP) techniques or URL-filtering [40]. For instance, a business can apply a DLP system with a clause stating that no data marked as sensitive can be moved to other clouds than the ones used and cleared by the company or government agency. One can therefore block the migration of data marked as sensitive to for instance Dropbox or similar file storage and sharing services.

For legitimate movement of data to or within a cloud, the CSA Guidance recommends three methods for securing the data transfer:

- Client/Application encryption, meaning endpoint encryption using built-in, or agent-based methods
- Link/Network Encryption, with examples like SSH, SSL and other VPNs
- Proxy-Based encryption, which means transferring data to a proxy used to encrypt and transfer data to a given location

The different methods should be reviewed in the context of the assets (data and functionality) to be moved or used in the cloud.

Next, data must be secured in the cloud, making sure that it does not leak to other parts of it, and that it is not accessible to unauthorized personnel. For IaaS solutions the security responsibility is mostly in the hands of the client, which can implement volume and object storage encryption to secure data from events like the loss of physical drives, and unauthorized access. For PaaS the clients possibilities of securing data (in addition to the CSP based security services) are limited due to the loss of control over the infrastructure. Nevertheless, client/application encryption, and any other encryption method capable of running on the target operating system will still apply, together with proxy encryption methods. For SaaS solutions one is basically in the mercy of the CSPs security implementation, although proxy encryption might still apply.

At the end of the life cycle the data must be destroyed, thereby permanently removing access to it on the targeted storage media. Such results can be reached by directly

overwriting the areas of the hard-drive holding the data. Although the method varies (overwriting one or several times using zeros or random bits), overwriting once with zeros (quickest and less resource demanding) is usually enough as this removes all hope of recovering files using software based recovery techniques. Some advanced techniques like Magnetic Force Microscopy [125] may still be able to reconstruct data based on the analysis of magnetic residue on hard-drive platters, but the method is not completely validated. In addition, NIST has stated that it assesses overwriting once with zeros as a sufficient method for the destruction of data [126].

The actions discussed here usually demands access to the storage medium, which may in some cases (private or hosted-private cloud solutions) be available. In other cases (for instance SaaS and PaaS) the destruction of data on the physical storage has to be left to the CSP, and it's the client's job to ensure that the methods used by the CSP are sufficient.

It should be noted that even if hard drives are obtainable, chances are that the data will be hard to recover by any means. CSPs like Google has encryption built into the file system, while at the same time distributing the files over several independent disks [72]. Similar methods can be expected from other cloud vendors.

4.7.2 Securing Equipment

Most cloud clients will have little possibility of influencing the security of physical equipment because of its placement at the CSP. In these cases it is up to the clients to assure themselves that the CSPs security regulations are sufficient, for instance by using auditing. In the cases where the client does have direct access to the equipment (for instance the private, on-premise part of a hybrid cloud) the same degree of security should be enforced here as is expected for all other equipment based on the sensitivity of the data it handles.

For bring-your-own-device (BYOD) solutions, for example a public cloud allowing connections from any supported device, security mechanisms, endpoint security and network access control systems can be applied.

4.7.3 Securing Users

The security of users means more than the blocking of unauthorized personnel from access to sensitive information. It also includes that the clients of the system can rely on their own session to be secured from eavesdropping or hijacking, and the ability to act as an anonymous user. The process of handling and managing the users of computer systems are often popularly referred to as Identity and Access Management (IAM). The application of such processes to CC are complicated by the fact that users log on from everywhere, meaning that the identification and authorization process may not be centralized (de-perimeterized identification) using several different directory services (DSs).

Some cloud providers have chosen to create services based on these issues, and companies like Lighthouse Gateway offers complete IAM-as-a-Service solutions. Research has also been conducted where IAM like processes are built into the user authentication processes of clouds [127]. We will, however, approach this area from a more general perspective.

There will typically be several points of focus when handling the identification and authentication of entities in any system. We start with an identity, defining the unique user. This must in turn be based on several different assertions, the attributes of the identity. Furthermore the process of entitlement, stating the privileges or security clearance

of the entity, must be performed, and fit into the set of existing access management rules. For more information on the organization of these processes, we refer to [40]

For SaaS implementations, identity federation is often used to provide identities [40]. Identity federation is basically an interconnection of several different DSs, which may or may not include the clients own DS. Federations typically use the Security Assertion Markup Language (SAML), which is defined as a framework for exchanging security information between online business partners [128]. Such systems can often be used for single sign-on (SSO), where a user that signs in at one of the members of the federation is automatically be authenticated at the remaining members if needed.

Similar standards for federated identity does not exist for IaaS and PaaS solutions [40], which will most likely be the ones interesting for governmental and law enforcement use. Federation of identity might also be problematic for federal agencies, as the highly distributed model could be hard to assess with security in mind. It is therefore unlikely that a government would chose similar identification services for highly sensitive systems.

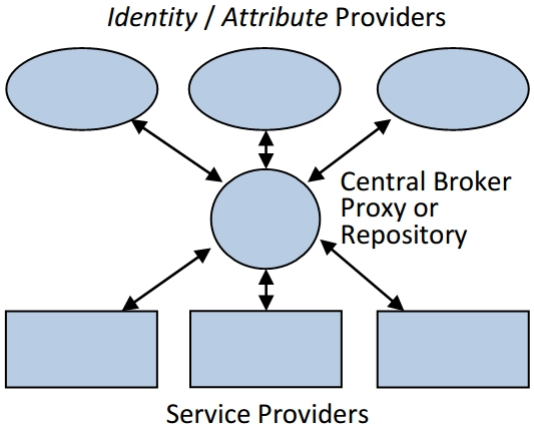

Figure 14: Hub and Spoke model [40]

One possible solution to the identification issues for governments and law enforcements is the ”hub and spoke” model shown in Figure 14. Here the cloud is able to communicate with the internal systems of the client that will handle all user authentications, for instance a repository via SAML. This allows the client to fully control the authentication service as well as all logging of the system, easing monitoring initiatives. Even though security is prioritized, it is important to note that the solution could impact the performance of the system to some extent, creating a bottle neck in the authentication process.

Additions to the Requirements Document

A section assessing the security solutions for all identified assets and physical devices is an important part of the planning process, and is therefore included in the requirements document in Step 8.(a), (b) and (c).

4.8 Risk Assessment

Barring the risk assessment all projects should perform prior to initiation, a risk assessment of the product to be implemented and its constituents should also be performed. This is done to explore what parts of the solution are most prone to risk, and to what degree an occurrence can negatively impact the implementation or its owner. Using assets as a base for assessing risk is a good starting point, as noted in Section 4.7.

This study will not further explore the details of risk assessments, and how these are performed. We therefore limit ourselves to pointing out that such assessments should be based on proven standards like [129]. Figure 15 presents the ISO/IEC 27005:2011 information security risk management process, which is a possible alternative for risk assessment. Here, the risk evaluation, treatment and acceptance is laid out as an iterative model.

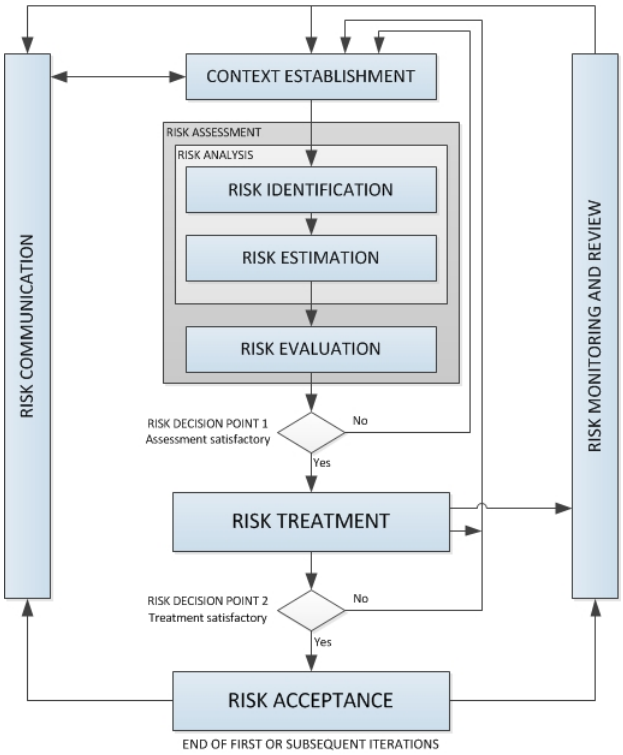

Figure 15: ISO 27005 Risk management flowchart. Inspired by [129]

Additions to the Requirements Document

Risk assessments should be conducted during the planning stages of a project, and will be included in our requirements in Step 9 of Appendix H.

4.9 Requirements Document Disposition

Based on the work in this chapter, a disposition of a requirement specification document has been created. Each sub-point has an individual list of tasks that needs to be fulfilled. The complete list with explanations and examples can be viewed in Appendix H.

- Prerequisites
  - Initial description of project
  - Initial legal study
  - Initial compliance study
- Requirements, functional and non-functional
  - High-level requirements
  - Detailed requirements
- Requirements for possible cloud environments
  - Cloud service model study
  - Cloud deployment model study
  - Cloud service provider study
- Legal requirements
  - Identifying relevant law
  - Assess law against requirements
- Compliance requirements
  - Identifying relevant compliance standards
  - Assess compliance demands against requirements
  - Create a plan for auditing compliance
- Security requirements
  - Securing data
  - Securing equipment
  - Securing users
- Risk assessment
  1. Description

4.10 Discussion

The goal of this chapter was to create a guide that list the steps needed to find and document the requirements of a CC implementation for governmental agencies. The results are presented in Section 4.9 as a quick disposition of the main chapters in the document, and Appendix H consisting of a complete requirements guide. The guide contains several tasks, descriptions of goals and examples of products and methods that can be used. One of the goals of this chapter was that the resulting document should remain framework independent, making it applicable to a wide audience.

In its current state the document represents the main steps of a planning phase, in which most steps closely resemble conventional methodologies. Where they differ is in

the content, as the functional and non-functional requirements of a CC implementation will for instance include cloud-specific demands. By explaining each step of the requirements guide H, the cloud-specific content for each of these are emphasized, making sure that important factors are not disregarded.

One might argue that the detail level of the document is not sufficient, and therefore not a good basis on which to conduct a planning phase. However, we believe that adding more details, thus restricting the freedom of choice in each section, would limit the applicability for many of the project guidance methodologies and frameworks which are used today.

By keeping the requirements document general, focusing on the sequences and reliabilities of the tasks rather than how these are specifically performed, it can easily fit into, or be customized to apply to most methodologies. It will then be up to the user to consult the document and perform its tasks throughout the applicable phases or work cycles of the chosen methodology.

We believe that this document could be a valuable tool for the planning of all cloud projects, not only the ones regarding governmental implementations.

5 Prototypical Specification

In this chapter we use the requirements document created in Chapter 4 to plan for the implementation of a governmental cloud environment. The planning is done using the Microsoft Operations Framework (MOF), which gives us a chance to test the requirements applicability to a project management methodology. The chapter starts with an introduction to the project scenario in Section 5.1, and continues with a description of what has been done in the MOF phases in Sections 5.2 and 5.3. In Section 5.5 the chapter's content is discussed.

For this chapter we use scenario number five (explained in Section 3.4.4), which is the same one used as the main example in Chapter 4. This serves as a starting point for our fictional government project. An elaboration on the scenario now follows in Section 5.1.

5.1 Project Scenario

A law enforcement agency situated in a country with an EU-membership has decided to implement a forensic analysis system using a CC environment. A description of the analysis system is found in Section 3.4.4. A project group has been gathered to plan and implement the solution using Microsoft Operations Framework (MOF) as a project methodology, which by policy is the required methodology for the law enforcement agency.

The agency will also be using a guide document (Appendix H) to assist in the process. It contains the steps required to map the requirements for a governmental CC application, and the project group has to ensure that all the sections in the document are handled during the project period.

As the project group is limited to the MOF methodology (quickly explained in Appendix I), the phases and Service Management Functions (SMFs) of MOF has to be followed. The following sections cover the phases of MOF, and the work performed by the fictional project team during each of these. As project planning and implementation is handled in the Deliver Phase 5.3 of MOF, the Plan Phase 5.2 will only be covered briefly.

5.2 MOF Plan Phase

The Plan Phase mostly covers mapping the existing IT infrastructures and needs of the agency, in addition to general policies that serve as rules and guidelines for the agency. This section will therefore only be briefly commented on.

During a meeting reviewing the Business/IT Alignment SMF, it was decided that because of the rising amount of digital evidence, and the need for the analysis of such, the existing analysis systems would soon struggle to meet the demands bestowed upon the law enforcement agency by the government. The need for a new forensic analysis system was identified.

Based on the findings which were reported in the management reviews (MRs) from the Plan Phase, it was decided that a new project group should be established to start planning a replacement implementation by mapping the needs it had to fill, and present

the findings. Given the governments focus on CC as of late, a prerequisite of the project is that the new system should be cloud based.

5.3 MOF Deliver Phase

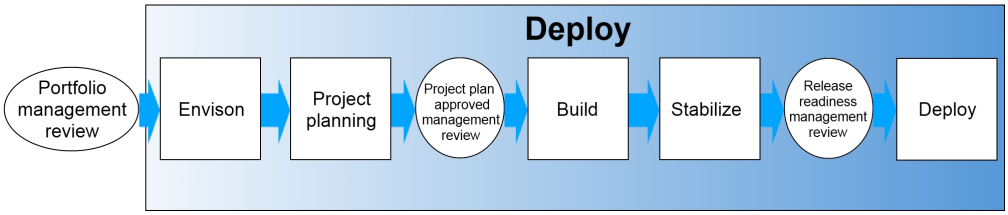

Figure 16: Quick overview of the deliver phase. Inspired by [130]

Figure 16 shows what the Deliver Phase of the MOF methodology is composed of, and how the checkpoint reviews (ellipses) and SMFs (rectangles) depends on each other. The portfolio management review, which documented the need for enhancements concerning the digital analysis system, starts the chain of activities that makes up the Deliver Phase.

This is the MOF phase that is most prominent for our scenario, and will therefore be the focus of this chapter.

5.3.1 Envision

Here follows the work done during the processes of the envision SMF.

Process 1: Organize the Core Team

A core team was assembled using resources in the agency. The specifics, the members and their roles are not of any significance for our study.

The project structure document is completed in this process, which includes:

- Project goals, objectives, assumptions, and constraints
- Project scope
- Roles and responsibilities
- Project protocols
- Risk and issue assessment
- Project glossary

The document will contain a rough draft of what will eventually become the vision and scope document created in Section 5.3.1. Because of this most of the content of project goals, objectives, assumptions and constraints, scope and risk and issue assessment will be elaborated on there. We therefore chose to skip this draft section, and jump straight to Process 2.

Process 2: Vision and Scope

The vision and scope document consists of several sub-points:

- Business opportunity
- Solutions concept

- Scope
- Solution design strategies

We will add two more sections to the document to coincide with the requirements document in Appendix H:

- Initial legal study
- Initial compliance study

This part of the envision phase incorporates all of the steps in Section number 1.(a) of Appendix H. Here, Steps 1.(a)i., iii. and iv. becomes the business opportunity and solutions concept sections, and Step 1.(a)ii. remains the scope.

Business Opportunity and Solutions Concept
This project has been initiated to map the requirements for a cloud based forensic analysis system, in order to assess replacements for existing systems. The following bullet points further explain the project.

- Problem Description
  During the last couple of years, police forces have seen a significant increase in crimes being performed completely or partially on computers and computing networks. Digital crime scenes now exist in 80 % of all criminal cases being handled in the US [24], and the amount and size of digital evidence is increasing at a high rate. Because of this, the traditional ways of analyzing evidence using dedicated servers suffers, especially during times of high pressure, as more hardware must be installed to fill demands. As a traditional server structure is basically used for one purpose at the time, the servers used for analysis will stay dormant and unused when the work load returns to normal, making it hard to justify the investment.

  The project group would like to explore the possibilities of a forensic analysis system implemented in a cloud environment in order to find out if the flexibility and scalability of such a system will better suit the task.

- Goals
  The main goal is to plan for an implementation of a forensic analysis system in the cloud to such an extent that a prototypical implementation could be based on it.

  The software should be able to:

  - Securely and reliably authorize users
  - Securely and reliably load evidence
  - Present a selection of analysis methods that can be conducted
  - Securely and reliably delivers results to the analyst
  - Log all activity (besides results) in coherence with CoC

  The cloud environment should be able to:

  - Support a wide variety of operating systems

  - Support management consoles for ease of upscaling, downscaling and deployment of virtual servers and clients
  - Make its resources available for other use during times of need while upholding a forensic analysis capability

- Analysis of expected benefits/drawbacks/gains
  Some specific benefits are expected for this project to be a success.

  The system should be:

  - Able to operate at the same performance level as conventional systems
  - Easy to maintain, update and upgrade with software and hardware
  - Capable of running server instances in parallel which should not interfere with the primary task
  - Should lower the overall costs to its owner compared to the system it replaces

Scope

The project group would like to explore the requirements of a forensic analysis system and map these to possible cloud environments in order to assess its applicability. The requirements will be based on both the features of the software-, possible hardware-, legal- and compliance demands of such a system. The police force in question is a part of the European Union, and will therefore have to comply with EU-regulations.

The forensic analysis system should be able to authenticate and administrate users to minimize the risk of unauthorized use. It should accommodate safe loading of evidence into the system, with access limited to the given analyst of the evidence, and have a set of analysis methods that can be applied to it. The system should log all user and analysis activity besides test results, which should be presented to the analyst in a secure manner and then be purged from the system. All actions performed on the system should be in coherence with the principles of evidence integrity and chain-of-custody, as well as the law and compliance demands that apply.

Solution Design Strategies

In order to provide a preview of the concept that is the cloud based digital analysis system, high-level requirements can be created and documented. The high-level requirements section from the requirements document will therefore be used during this activity of the envision SMF.

- High-Level Requirements
  This section will constitute the foundation on which a technical solution can be designed. It should focus on what should be done rather than how to do it.

  - *Identify users and usage patterns*
    During dialog with the projects stakeholders, a representative group of potential users of the solution and the project group, a list of user stories were created. Table 9 represents these findings.

    Based on the list we have identified two user groups, administrators and users. Table 10 presents these.

I as an Admin should be able to log into the system because I need to be properly authenticated to perform my other tasks
I as a User should be able to log into the system because I need to be properly authenticated to perform my other tasks
I as an Admin should be able to add a user because the users of the system needs to be strictly regulated
I as an Admin should be able to remove a user because users that are not needed on the system should not persist in it
I as an Admin should be able to add an admin because the admins of the system needs to be strictly regulated
I as an Admin should be able to remove an admin because admins that are not needed on the system should not persist in it
I as an Admin should be able to list all users because it will ease the task of administrating these
I as an Admin should be able to list all admins because it will ease the task of administrating these
I as an Admin should be able to search in a list of users because it will ease the task of administrating these
I as a User should be able to load digital evidence into the system because it has to be present on the system in order to analyze it
I as an Admin should be able to add an analysis method because new methods will be created, and should be made available to the users
I as an Admin should be able to modify an analysis method because it might need to be updated or fixed
I as a User should be able to list available analysis methods because it will enable me to assess how to perform the analysis
I as a User should be able to choose an analysis method because it will enable me to specify my analysis
I as a User should be able to recover the results of an analysis because it will enable me to review and save these
I as a User should be able to recover the list of my recent activities (except analysis results) because it will enable me to prove the occurrence of a test
I as an Admin should be able to retrieve the logs of all user activities (except analysis results) because of auditing purposes

Table 9: Identified user stories

Administrators	Administrating users, other administrators and the analysis methods present in the system
Users	Loading evidence into the system, analyzing it and recovering the analysis results

Table 10: Identified user groups

- *Identify major functional requirements based on usage patterns*
  Based on the study of users and usage patterns we have established a list of high-level functional requirements. Table 11 presents these.

  FROM HERE
- *Identify major non-functional requirements based on usage patterns*
  Based on the study of users and usage patterns we have established the high-level non-functional requirements presented in Table 12.

A secure method for logging into the system must be present for all user groups
Administrators should be able to administrate the users of the system
Administrators should be able to administrate the analysis methods of the system
Users should be able to analyze digital evidence using a chosen analysis method
Users should be able to recover the results of an analysis
The system should not allow a user or administrator of the system to view the results of other user's analysis
The system should log user's activities, but not the results, and make this available only to the user in question and administrators

Table 11: High-level functional requirements

NOTE: Since a reference digital forensic analysis system is not present, we have no way of measuring the performance of it. We therefore have no way of ensuring the compliance of our identified non-functional requirements. These will therefore serve only as examples of possible non-functional requirements.

The system should be able to handle several analysis jobs simultaneously
The system should be able to process the analysis jobs quicker than, or as quick as the existing system
The system should be highly transparent throughout the entire cloud stack, to provide the best possible foundation for creating and implementing security and logging mechanisms

Table 12: High-level non-functional requirements

Initial Legal Study

The proposed system will handle digital evidence in the interest of law enforcement, including information that is defined as PII. Because of this several EU-directives can be omitted, but may still be complied with to an extent.

First and foremost one can state that the Data Protection Directive (95/46/EC) will not apply for this system as of article 3.2, stating that "This Directive shall not apply to the processing of personal data: - in the course of an activity which falls outside the scope of Community law, such as those provided for by Titles V and VI of the Treaty on European Union..."[98].

Title V refers to provisions on a common foreign and security policy, while title VI refers to provisions on police and judicial cooperation in criminal matters [131]. This means that supplemental directives like 2002/58/EC [99], which has the same jurisdictional limitations, can also be disregarded.

Since this system is primarily created to handle domestic crimes, it should not have to comply with the Framework Decision 2008/977/JHA, which covers data protection demands when transferring and handling transferred data from other member countries. It would however be very wise to comply with the directive, as evidence received from other member countries or through Europol would not be eligible for processing in the system.

Even though some of the current directives state that data handling for criminal case work is exempt, it would be wise to comply with all the directives to the possible extent. The directives are currently being revised, and will be reorganized in one directive for the security of data handling in the EU, with a section specified for law enforcement.

As this might impose many of the demands of the currently exempt directives on law enforcement activities, these should be consulted and assessed for compliance.

Other than this, the national law of the country in which the system shall be situated will apply in the cases where EU law does not cover the said area, and when the national law is stricter than the demands imposed by such regulations. Since our implementation is not country specific, national laws will not be covered here.

The next three lists consist of a run through of the applicable directives and their articles. In the case of directives that does not demand compliance, the project group has assessed which of the articles may be complied with and how, and which articles should not be complied with based on the nature of the project. In the case where a directive is mandatory, all articles has been assessed for compliance.

- 95/46/EC compliance
  - As of article six the system must be built to ensure that information containing personal data (as defined in the directive) holds a high quality and is collected and processed solely on the basis of why it was originally procured. The data must be kept up to date if necessary, safeguarded from alterations and unauthorized access, and kept only for as long as necessary. One should only be able to connect a piece of data to a person for the time this is deemed strictly necessary, and any storage beyond this point should only be performed after the data is sanitized, removing personally identifiable features. Sanitized data should only be kept as long as deemed necessary.
    The project group believes that the system may fully comply with this article through thorough security controls for system access, logging and procedures for reassessing data stored in the system or system affiliated databases at set intervals. The specifics of the reassessment procedures should depend on whether the data can be linked to a person or not, and if it is part of a current of prior criminal case. When no reason for keeping said data is found during assessment, the data should be securely disposed of. No such data should be present in the system for an extended amount of time, but the procedures of the law must be followed during this limited time period.
  - It is the belief of the project group that since the system is built for law enforcement, it will not need to be in compliance with articles seven, eight and nine.
  - It is the belief of the project group that the demands of articles 10, 11 and 12, related to a subjects rights to be informed of the processing of data related to his person, will not need to be complied with in its entirety for the forensic analysis system. This caused by the fact that subjects under surveillance should not be informed of such things during the surveillance period, etc. In addition, the system will not store evidence data or its logs over long periods of time, rather moving such information to other systems. A person's right to insight into how his/her data has been processed will therefore most likely rather apply to other systems instead.
  - Article 13 states a member nations rights to use their own legislations to limit the rights of subjects in several articles, for instance when the data is used for

law enforcement. As national laws will not be covered here, we will not assess this article further, but note that it is an important part of legal compliance.

- The system should be in compliance with articles 16 and 17 given the implementation of proper security controls. The systems purpose and functionality should, in the belief of the project group, not impose any problems for this compliance.
- The system will be exempt from articles 18, 19, as the notification of subjects that has had their information processed by the system, and what information this notification should contain, is handled by other sections of the law enforcement system. However, logs from the system should be created that correctly states what processing has taken place, how the data has been handled in the system and by whom, so that such information can be retrieved in cases of notification. Such information can also be used for the cases noted in article 21, for which the system should also be expected to comply with to an extent. Publishing the complete process of analyzing evidence will in most cases not be favourable, as such information could be used to obfuscate evidence and fool the analysis process.
- The system will be exempt from article 20 of the directive, as the data will be cleared for processing prior to reaching the analysis system.
- The rest of the articles concerns how the member states will uphold the obligations stated in the directive, and does not concern the system in particular. Articles concerning a data subjects rights have also been omitted, as they are out of the systems scope.

- 2002/58/EC compliance
  - As noted, the directive is not mandatory as it follows the same rules for applicability as 95/46/EC according to article one.
  - Article three further states that the directive applies for ".. the processing of personal data in connection with the provision of publicly available electronic communications services in public communications networks" [99]. As the system will not include any publically available communication channels as part of the implementation, the project group rule out this directive in its entirety.

- 2008/977/JHA compliance
  - Article one states that the directive is created to ensure the proper security of data transferred between EU member states for the purpose of processing by law enforcement. It also states that the subject's rights to freedom and privacy shall be maintained throughout this process. This means that the system should comply with the directive when data transferred from another member state is to be processed by it. The project group finds that this directive is to be deemed as mandatory for the system, as the probability of such actions taking

place is highly likely, and that failure or avoidance of compliance would place severe limitations on the systems scope.

- Article three describes that the collected data can only be processed for the purpose of which it was collected. It furthermore states that any other processing can be done as long as it is not incompatible with the original purpose. A clear statement of the analysis methods present in the system should therefore be created, so as to ensure that the methods are compatible with each other, thereby providing the option of running additional analyses based on the results of others.
- Article four, paragraph two will be applicable to the implementation as any personal data processed by the system should be properly erased when it is no longer needed. Controls should be implemented to ensure that the analysis system thoroughly and securely erases the data from the system, ensuring that it cannot be retrieved from it at a later stage. It is the project group's belief that such controls and their technical solutions can be successfully implemented to comply with the article.
- Article five and nine should not become issues, as the system is not created to store any evidence over time.
- Article six, seven, 21 and 22 should be complied with by appropriate security controls throughout the analysis process, meaning the procurement of evidence, analysis, result extraction, logging and secure removal of the analyzed data.
- Article eight should not be applicable, as all data should be quality checked before it is cleared for analysis.
- Article 10 concerning the documentation of data handling should be complied with by ensuring thorough logging of the handling of data, and the secure handling of the logs themselves.
- Articles 11, 12, 13 and 14 regarding the lawful purposes of sending data between member countries will be omitted, as it concerns the processing of data before it reaches the analysis system.
- Article 15 states that the senders of data can ask for documentation on how the data has been handled. The documentation of the analysis system, and logs from the systems processes should therefore be created to be able to meet the demands.
- The rest of the articles does not concern our implementation.

Initial Compliance Study

A study of compliance regulations should now be conducted based on the legal study and the requirements created for the project so far. We will not conduct a thorough study off applicable compliance standards, but some thoughts can be noted. Out of the compliance standards and regulations covered in Section 4.6 we can presume some possible choices. ISO27001 and 27002 are obvious choices, as the ability to identify and mitigate risk together with the creation of security requirements are important factors to consider.

The Common Criteria could also be important for the system, as it assists in the creation of security requirements as well. NIST SP800-27 would also be a good candidate, as it would be applicable to the US counterpart to our implementation. Finally, the Daubert criteria must be assessed and its demands resolved. This will directly influence the admissibility of evidence handled by the system.

Risk Assessment

A project specific assessment of risks is performed. Its specifics will not be documented here, but more information concerning how it can be conducted can be found in Appendix I.4.

Process 3: Approve the Vision and Scope Document

At the end of the Envision Phase, a meeting is scheduled between the project group and the stakeholders of the project within the law enforcement agency. The agenda is to assess the vision and scope document, and ensure that the project group is progressing on schedule and according to the original plan. In this case the document is approved, together with the additional sections regarding legal and compliance studies, and the risk assessment. This allows the project group to advance to the next SMF of the phase.

5.3.2 Project Planning

During this SMF a detailed planning process is conducted, which is based on the work performed in the Envision SMF.

Process 1: Evaluate Products and Technologies

This process is concerned with evaluating possible technologies which can be used in the solution. It consists of two sub-activities.

Customer baseline

This process starts with creating the customer base line, which basically is the assessment of the available environment in which a solution can be placed. This equipment is the same equipment that was used in [132].

- Two HP Proliant DL320 G5
  - Dual core Intel Xeon 64-bit CPU(2.4GHz)
  - 2GB of RAM
  - One 250GB hard drive
  - Dual ethernet interfaces
- Two Dell Optiplex GX260
  - Intel Pentium 4 CPU(2.66GHz)
  - 512MB of RAM
  - One 40GB hard drive
  - Single ethernet interface
- One Dell Optiplex 960

  - Intel Core 2 Duo CPU(3.0GHz)
  - 4GB of RAM
  - One 300GB hard drive
  - Single ethernet interface

- Network infrastructure
  - Switch for interconnectivity

- Other equipment
  - KVM-switch for I/O-equipment

The environment has free access to the Internet if needed.

The system seems capable of running a private cloud environment using two servers in a resource pool, and the remaining servers for possible additional features (storage, remote administration, etc).

Evaluation of products and technologies

The MOF methodology suggests assessing the technology based on the vision and scope documentation and the customer baseline. Some sections from the requirements document can therefore be incorporated in this SMF.

- Requirements for possible cloud service models
  A digital analysis system requires that the owner of the solution has full control over both the software and hardware environments on which the analysis is run. This is because of the principles of CoC and evidence integrity. None other than authorized personnel should be able to access the evidence, analysis results or the system on which the analysis solution is built.
  - Defining the cloud service model alternatives
    The only possible alternative for this solution will be an IaaS implementation, as the owner of the system will need control over as much of the cloud infrastructure as possible. This will make it easier to implement and assess customized security and logging mechanisms, something the implementation and legal regulations demand. A PaaS solution would only give the developers control at the OS layer and up, and a SaaS solution would be even more restricted. For a description of the service model we refer to Section 2.2.4.
  - Discussing each cloud service alternative against the high-level requirements
    In Section number 5.3.1 a list of goals was noted for both software and hardware. In the list of software goals one explained that all activity should be logged. By ensuring that the owner of the system has full control over the OS, and the choice of which OSs to use, the opportunity for logging can be enhanced. In the list concerning cloud environment goals, it is stated that the solution should support a wide variety of OSs. This is typically only the case on an IaaS-environment supporting both hardware supported and paravirtualised instances. Furthermore, applying to the requirements featured in Tables

9,10 and 11 are trivial, as the supported VMs should not limit the ability of the proposed software in any particular way. The requirements stated in Table 12 will have to be covered during the discussions regarding deployment models and CSPs.

  - Choosing a cloud service model
    Given that IaaS is the only service model that sufficiently meets the requirements of the proposed implementation, the project group declares that the IaaS service model should be used.

- Requirements for possible cloud deployment models
  The deployment model specifies where the physical devices of the cloud will be situated, something that effects how data will be transported between the cloud and its users.

  - Defining the cloud deployment model alternatives
    A description of cloud deployment models are found in Section 2.2.5.
  - Discussing each cloud deployment alternative against the high-level requirements
    Based on the large number of confidential assets, and the general sensitive nature of a forensic analysis system, a private deployment model is the logical choice. This must trump the superior flexibility and overwhelming resource potential of public and hybrid clouds, which such a system would benefit massively from. However, given the state of the art of CC, such a system cannot operate in a public cloud environment, as one cannot argue that such environments are secure enough yet.

    Furthermore, since most public cloud providers use closed source systems, there is no way of knowing if the information is kept secret or if it can be made available for the CSP. The issues of colocation of data and spread data centres over several countries might also conflict with 2008/977/JHA, which was assessed in the legal study. We therefore propose that all use of public CC environments should be avoided. This eliminates the use of public and hybrid cloud deployment models.

    The two types of private clouds found are on-premise private clouds and hosted private clouds. While hosted-private clouds show a lot of promise concerning the complete segregation of clients on the hardware level, this would still mean that a lot of sensitive information is processed at a third party. This would in turn put high demands on the third party CSP.
  - Choosing a cloud deployment model
    Based on the discussion in the previous step, we state that the project group decide on an IaaS cloud environment. In addition to this the on-premise private deployment model is chosen, as the group do not see the technology as mature enough to risk off-premise processing of digital evidence.

- Requirements for specific vendors
  As the previous step specified that the implementation will consist of an IaaS cloud

using an on-premise private configuration, the choice of vendors is limited. Additionally, the project demands a great deal of system transparency to ensure that information leakage or other errors concerning applications running in the cloud can be observed and corrected. This emphasizes that open source solutions should be assessed. Some open source vendors have been selected as alternatives.

- Discuss alternatives
  Several alternatives can be mentioned, amongst others CloudStack, Xen Cloud Platform and Eucalyptus. A presentation of the cloud vendors can be found in Section 2.2.7. All meet the requirements for deployment models, and, because of the choice of IaaS as a service model, the high-level requirements are covered as well. As IaaS gives the owners access to the virtual hardware, limitations to what can be created is only limited by the supported OSs. The hardware goals set forth in Section 5.3.1 can be used to further assess the different vendors. We will not be performing an elaborate assessment here, but some examples of discussion points can be mentioned. The open-source version of Eucalyptus lacks support for Windows VMs, meaning that it does not fit the requirement for wide OS support. In addition, parts of the cloud are closed-source due to the enterprise edition which has been created.
  We are then left with CloudStack and XCP, which are harder to distinguish, which should warrant a thorough investigation. A point in the favour of CloudStack is that it supports several hypervisors, including the Xen hypervisor supported by XCP. A point for XCP is its devotion to one hypervisor, as Xen is an integrated part of the cloud framework. One can therefore claim that XCP is more likely to fully support its hypervisor.
- Verify choice
  Both XCP and CloudStack fill the requirements for a cloud vendor presented in the preceding steps of the deliver phase. However, due to prior testing, it is also established that XCP is supported by the hardware we have available. Based on this we state that XCP was chosen as the cloud vendor.

Process 2: Write the Functional Specification

This part of the project planning SMF consists of further documentation of the project requirements into the functional specification document, and the creation of design documents Section 2.(b), Detailed requirements, in the requirements document (Appendix H) is applied to this MOF process. The process consists of several sub-activities which are presented in this section.

Documenting the Project Requirements

MOF states four different categories of requirements that have to be covered in the requirements document: Business, user, operational and system requirements. Most of the business requirements would here be elaborated from the organizational policies and expectations to IT services stated in Section 5.2. Since these are out of scope for this study, the business requirements will be omitted. Instead we will include security requirements as a part of this documentation, which will be added to the end.

The rest of the requirements will be based on the vision and scope document written in process two of the envision SMF 5.3.1. The user requirements will be based on the

user stories that were presented in Table 9, the user groups defined in Table 10 and the functional requirements stated in Table 11.

User Requirements

A large part of the user requirements consists of creating use-cases. Table 13 gives an example of an advanced use-case table based on the two first user stories of Table 9.

Use-case 4	**Log in**
Actor	Administrator and User
Goal	Gaining access to the system with user group specific rights
Summary	A user or administrator would like to be authenticated based on their group affiliation, by logging in using user name and password
Type	Essential
Preconditions	Wants to log into the system
Postconditions	Successfully logged into the system with the correct permissions
Eventfloat	
Actor	System
1. Open system interface 3. Enter user name 4. Enter password 5. Submit information	2. Presents login screen 6. Acquire credentials 7. Check access permissions 8. Grant access to system
Alternative eventfloat	
Steps 1 - 7 is performed	8. Access denied 9. Present error message 10. Present login screen

Table 13: Advanced use-case table

Use-cases like these, and use-case diagrams should be created for all the user requirements, to properly map them. Flow diagrams can also be added to further map the use-cases. An example of this was given in Figure 13.

Operational Requirements

These requirements will, as the name suggests, map the expected capabilities of the implementation. Where the user requirements covered the functional demands, the operational requirements concerns the non-functional ones. Some non-functional requirements were listed in Table 12:

- The system should be able to handle several analysis jobs simultaneously
- The system should be able to process the analysis jobs quicker than, or as quick as the existing system

- The system should be highly transparent throughout the entire cloud stack, to provide the best possible foundation for creating and implementing security and logging mechanisms

These would have to be further assessed to come up with technical solutions for how to meet the demands. The handling of several simultaneous analysis jobs would for instance indicate that the system should accept a number of simultaneous log ins. Such a solution could in turn indicate that a log in should be presented to users through remote login services or web services. The second non-functional requirement warrants the need for benchmarking results from the existing solution, and studies of the system requirements. The expected performance capabilities of the new system would then be able to be assessed against the existing figures. From there, the need for additional resources could be emphasized in order to meet the demands. The third requirement is upheld by the fact that an open source solution is selected. It will then be up to the developers to explore the source code for possibilities.

System Requirements

The system requirements will consist of demands for hardware and software, which are based on the user and operational requirements. Specifically, the system requirements will ensure that the system is able to provide for all the usability and resource needs of its users, and that it has the ability to expand, be upgraded and updated as were designed.

In our case the system requirements are restricted, as the available hardware has already been defined in Section 5.3.2. This will often be the case when limited financial resources exist, or if a predefined test environment is the only available equipment. User and operational requirements will then have to be assessed with this in mind, as upgrading the available systems might not be a viable option.

Security Requirements

After all the use-cases have been mapped, all the assets of the implementation can be found. By documenting the assets it can be easier to resolve what has to be protected in the system, as stated in AppendixH Section 2.(b)iii. Table 8 in Section 4.3.3 present a list of possible assets.

It will be particularly important to assess the sensitivity of the assets. By answering the questions stated in Section 2.(b)iv. in Appendix H, and rating the answers from low sensitivity to high sensitivity, we can assess each asset by themselves. An example is given in Table 14.

By answering the questions in this example, we deduced that the sensitivity of the asset is ”medium”. This is based on the overall combination of results for each of the questions. The sensitivity level of the asset could just as well be based on the highest rated question answer, if this was warranted by the project group. In that case it would give the asset a sensitivity rating of ”high”. Table 15 provides an example of how the asset can be presented. Optionally it could be expanded with a description like in Table 8.

When every asset is assessed it creates a foundation on which the security requirements of the implementation can be based. As covered in Section 4.7 one can divide this documentation into the security of data, equipment and users (Note that, if compliance

Sensitivity assessment of Asset 1: Log in		
Question	Answer	Sensitivity rating
How would we be harmed if the asset became widely published and widely distributed?	If the source code for the login program was distributed, it could create an opportunity for potential attackers to study the program and look for bugs or information concerning system components. This would be unfortunate, but probably not severe, as no login information should be present in the code. Furthermore, programmed in a secure way, the authentication process should not be vulnerable, even if the code is known.	Medium
How would we be harmed if an employee of our CSP accessed the asset?	As the cloud will be deployed privately, the operators of the cloud will consist of agency personnel. The consequences should therefore be small.	Low
How would we be harmed if the process or function were manipulated by an outsider?	The consequences of this happening could be devastating, as criminal cases which evidence presentation is based on the analysis system might be dismissed.	High
How would we be harmed if the process or function failed to provide expected results?	This depends on how the function would fail. A log in program that would reset to a ”default allow” setting when failing in some way could be devastating, as all attempts to log in would result in access to the system. A ”default deny” state would however be no more than a nuisance, as users of the system would be unable to log in and do their job.	Medium
How would we be harmed if the information/data were unexpectedly changed?	If the program was changed to potentially include a back door, or allow access in some other way, it would create a possibility for great harm. Dismissal of evidence and leakage of PII could occur.	High
How would we be harmed if the asset were unavailable for a period of time?	Users of the system would not be able to use it, thus limiting their ability to do their job. It would, however, not cause any direct harm unless the system remained down for a considerable amount of time.	Low

Table 14: Sensitivity assessment of asset 1

with the standards like the Common Criteria is a demanded, these will have their own way of assessing the security requirements).

Document the Functional Specification

The functional specification document will collect the user-, operational-, system- and security requirements into one document. The document should sound foundation on which the development of the implementation can be based.

Document the Design

Asset 1	**Log in**
Type	Functionality
Associated use-cases	4
Importance	Critical
Sensitivity	Medium

Table 15: Asset 1 description table

The design documentation is gradually created and detailed through three steps. The first step consists of a conceptual design, where all the ideas are created and possible solutions are sketched out. For a quick introduction to this process, this video from Nokia capture the core ideas [133]. The conceptual design could already be completed during the envision phase, specifically during Process two in the Envision SMF (Section 5.3.1), but should nevertheless be included in this documentation for further reference and traceability of concepts.

Next on the list is the logical design. Here the conceptual design is arranged into a more structured form. Examples of this documentation can be domain models or class diagrams. The components of the domain and class diagrams will in turn be used as alternatives for software components. Other than this, sequence or state diagrams can be created to further describe use-cases.

The last part of the design document consists of the physical design. Here the layout of the hardware, its software suites and the network topology is mapped out and defined. An overview of our available hardware was given in Section 5.3.2 during the mapping of the customer baseline. Based on this, a model can be created to visualize the hardware setup. Figure 17 explains how parts of the equipment was set up during the work in [132], which we have modified for our case. The figure shows a XCP setup using two cloud hosts in a resource pool running the Linux based Xen hypervisor, a network storage server and a server providing a third party management UI. Furthermore the model visualizes log in functionality for users of the system, a firewall and a Web UI that can provide users with the services that the cloud offers. We note that the model could be expanded with more features, such as a log server and a user repository. Another model should also be constructed based on the virtualised hardware and network setup.

5.4 Final Notes

As all the sections of the requirements document were successfully implemented into the MOF framework, the mapping is complete. The tables given in Appendix J present the sections of the requirements document, and where in the MOF framework they were implemented during this process.

5.5 Discussion

During this chapter we decided to apply our requirements document to the MOF project management standard, in order to test our demand of a framework independent product. In order to do this we had to map all the sections of the requirements document to the phases, SMFs and activities of the MOF framework.

Since the requirements document was generated to guide in the specification of requirements during the planning of a CC implementation, most of the document applied

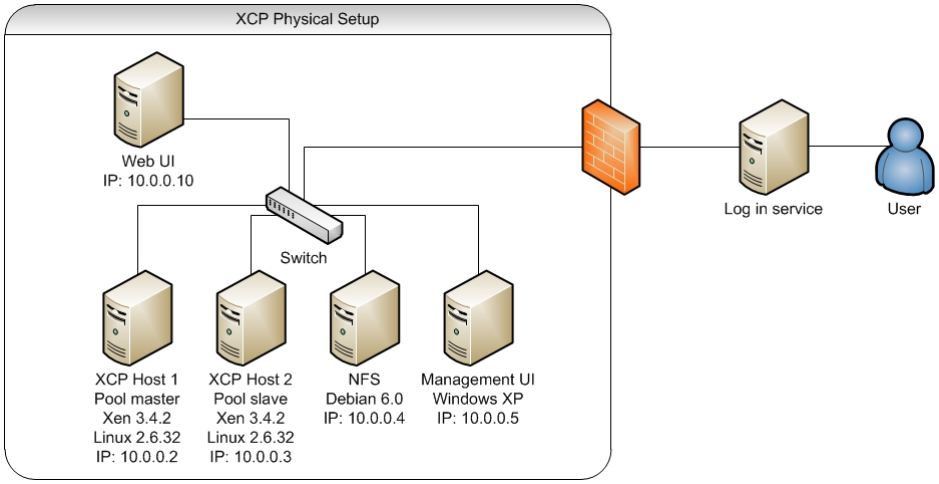

Figure 17: Quick overview of the hardware implementation [132]

to the Project Planning SMF of the Deliver phase of MOF.

Only one step was mapped to the Plan phase, as most of the phase is not limited to singular projects. The authority step of the requirements document was placed in the Business/IT Alignment and Reliability SMFs as this is where new projects are initially discussed and authorized. This worked out well, but as authorization is performed at several stages of the MOF process, this is not the only place it resides.

Most of the initial description steps of the Prerequisites/Constraints section were placed into the Envision SMF of MOF, as the three first steps of this section closely resembled the content of the vision and scope document in the Envision SMF. No problems were encountered here. The high-level requirements were added into the Solution design strategies part of the vision and scope document, and proved to fill its needs nicely by applying user stories, actor lists and high-level functional and non-functional requirements.

Furthermore, initial legal and compliance studies were added to the vision and scope document. This was not a part of the MOF model, but it seemed appropriate that the issues concerning these demands were handled at an early stage. No issues have been noted regarding this. Risk assessment also found its place in this SMF, as it is already mentioned as a part of MOF. Furthermore, the Authority step returned during Process number three of the Envision phase, as an approval of the vision and scope document is needed here.

As the initial legal and compliance studies are handled in the early stages of the project, one could argue that the premises of the project is not sufficiently stated, and that the foundation of the studies are weaker than if such were to be performed at a later stage. Even though this is true, it is important for the organization to approach such requirements early, thus uncovering clear breaches that might lead to problems later on in the project. We see no reason why such studies cannot be performed during the later stages of the project, but these should be performed in addition to the initial studies to be certain that the premises of the implementation is valid from the start.

We then move on to the Project Planning phase which starts with the choice of technology. This meant that the Resources section of the requirements document was placed

here in order to map the available hardware resources. This aligned nicely with the MOF section called Customer Baseline, which amongst other things mapped the customer's infrastructure for assessment with possible products. The requirements for possible cloud deployment models, cloud service providers and cloud vendors then followed, which could be assessed against the customer baseline.

After this, the Detailed requirements of the requirements document was used in its entirety, as these related well to the user, operational and system requirements of the MOF model. The security requirements were also incorporated into this section to assess the security demands of the noted requirements. This concluded the mapping of the requirements guide to the MOF phases, as the Operate phase was out of scope for this study.

One problem encountered was that MOF advice that a choice of technology (deploymentmodel/servicemodel/CSP) is chosen before detailed requirements are created. We decided to try the MOF approach, and chose the technology before exploring detailed requirements. This shifted around some of our requirements sections, and we had to rely on the high-level requirements to guide the choice of technology.

A benefit of this approach is that one will be able to tailor the detailed requirements to the chosen technology. A drawback is that the choice of technology is based on a weaker foundation. No additional issues emerged because of this, and we find that choosing the technology early is a good idea especially if all, or parts of the technology are already stated in the preliminary planning stages.

We find that the sections of the requirements document were incorporated into the existing MOF methodology without any severe problems. This strengthens the belief that the guide is framework independent, and shows the flexibility and adaptability of the requirements document. We cannot, however, state that this is a proof of framework independence, and several other methodologies would need to be tested in order to solidify the claim. In addition, the tests should be performed during a real project, and not be limited to fictional cases.

6 Discussion, Further Work and Conclusion

In this chapter we first present a quick summary of the thesis results. After this the theoretical and practical implication of our research is presented, discussing how the results may improve the process of planning and developing governmental cloud implementations. At the end of this chapter the conclusion of the thesis is presented, and possible areas for further studies are proposed.

6.1 Summary

At the initial stages of this study we found many potential benefits and drawbacks of governmental cloud implementations. The results we got by attempting to process these seem to imply that the private cloud deployment model is the best suited cloud alternative for law enforcement, while conventional infrastructures still seem to be well suited. The discussion emphasizes that some refinements should be conducted to strengthen the quality of the method, but that it shows promise as an assessment method for cloud applicability studies. The listed benefits and drawbacks represent the primary arguments both for and against CC in our day and time.

In coherence with research question number two, we set out to map the requirements for CC implementations, using a private cloud for forensic analysis as an example. This early on proved to be a difficult task, as requirements may vary greatly based on the scenario that is applied in each case. By focusing instead on how this problem could be solved for all scenarios, we decided that constructing a guide for the creation of such requirements would be the best approach. The result was a modifiable list of steps, created to be methodology agnostic, that would assist in the building of requirements during the planning stages of an implementation: A requirements guide. By focusing on cloud issues, and providing the user with examples of solutions and related literature, the guide should be well suited as a tool for planning cloud implementations.

To test the document and answer the last of our research questions, we used MOF as a project management methodology, and applied our newly created guide. The steps of the guide fit into the MOF planning phases with some adjustments, and thus provided an example of how one can plan for a cloud environment for law enforcement. It also showed promise for the flexibility and methodology independence of the guide and its steps.

6.2 Theoretical Implications

The research performed during this master thesis started by reviewing the upsides and downsides of CC as viewed from the standpoint of a governmental law enforcement agency. Such work requires a lot of insight into the fields of CC and law enforcement. This work resulted in an extensive collection of benefits and drawbacks, but it is likely that more may be found. Because of this, a benefits and drawbacks list created by a law enforcement agency project group could look different, but we are confident that this list represents most of the current benefits and drawbacks that exist.

The benefits and drawbacks were categorized using the AAII categories by Paquette [74] in relation to the well proven CIA model. By doing this, Chapter 3 of the thesis feature a new way of assessing the applicability of cloud environments. The AAII categories proved to be well suited for our task, but might be an unfamiliar system for most information security professionals. Organizing the same benefits and drawbacks using the CIA model should be possible, however, we believe that using a different model has helped us in creating a more nuanced view of CCs application for law enforcement.

The resulting list of CC drawbacks outnumbers the benefits, which is most likely a result of the still young technology that CC represents. We also believe that as the technology improves, these numbers will shift in favour of the benefits. The list of benefits and drawbacks can thus be used as a baseline for CC, and by updating it regularly it could help to clarify the improvements of the cloud as it evolves.

As we used an in-depth research approach while rating the benefits and drawbacks, the results should reflect the current applicability of CC for law enforcement. Judging by the results of the calculations conducted in 3.5.2, the private cloud seems most suited, while conventional systems still has the edge over cloud systems in all but the areas supporting availability.

While the results of the questionnaire were deemed to be less reliable based on the low number of respondents, the results can still show indications of the applicability of the cloud in our setting. Seeing that the end results of best suited deployment models closely resembled the results of the in-depth study, this is likely to be the case.

The next step of this master thesis was to explore the requirements for governmental CC implementations. In order to make this applicable to a large amount of possible projects, we decided that it would be useful to create a guide for exploring the creation of such requirements in each case. As the creation of requirements is highly related to the individual projects, the requirements guide that was created represents a general approach to the problem. This ensures a high degree of applicability to any project.

Since the task was not to create any form of methodology or framework, the guide, called a requirements document, had to be created so as to be methodology-independent. This is in tune with our goal of high applicability, and is intended to ensure that projects limited to a given framework are not discouraged from using the guide.

The requirements document found in Appendix H is the product of this section, while Chapter 4 itself documents why we believe that the steps in the requirements document should be implemented. The requirements document borrows several examples like use case creation and data flow modeling from methodologies and methods like RUP and UML. It also incorporates cloud-specific steps for the choice of cloud environments and vendors, as well as tips concerning the applicable legal regulations and compliance standards.

The implications of applying demands for standards to a technology which is not yet standardized cannot be underestimated. This is one of the vices of using new technology, and the only way to mitigate the issue is to use thorough methods to map the needs of the technology. By consulting the requirements guide made here, such issues can be explored and limited.

In order to test the application of the requirements document to project management methodologies, we decided to create a fictive MOF project. The project would use the MOF phases and SMFs to drive the project documentation, while the steps in the

requirements document were to be implemented where suitable. Some problem areas were found, as noted in Section 5.5, but overall the steps in the requirements guide slotted nicely into the MOF framework. In the cases where the steps did not match, this was mostly based on a difference in sequence between the steps in the guide and in MOF. By altering the sequence of the guide to fit the MOF methodology, these issues were mitigated. As a result, we managed to apply the entire requirements document to MOF, which proves some of its flexibility. The result does not prove the applicability to all frameworks, but it shows promise for such a characteristic.

6.3 Practical Considerations

We believe that the method used may greatly benefit the assessment or planning processes of CC implementation projects. There are several reasons for this statement. First, the method requires that a wide area of aspects concerning CC is assessed. If the project is required to rate such a list of benefits and drawbacks using a well established scenario or project description, it will force the project to acknowledge possible problem areas of the cloud implementation as well as possible upsides. Secondly, how a project would choose to rate the different benefits and drawbacks could assist in focusing the priorities of the project, as the highly rated benefits and drawbacks should reflect the prioritized requirements for the given project. The method for calculating the best deployment model could also prove a useful feature for such work.

If implemented together with an existing project management framework, the requirements guide ensures that cloud-specific issues are under focus throughout the process. While not extensive in describing how a task should be conducted, the guide provides relevant examples of solutions, encouraging the user to apply the methods best suited in each task. By also adding references to useful documentation for legal and compliance demands, the user of the guide is encouraged to assess these to ensure conformity. By elaborating on all the tasks stated in the requirements guide, its users will be aided in making informed decisions.

6.4 Conclusion and Further Work

The goal of this master thesis study was to research the applicability of CC for governmental agencies. We decided to study this by providing new knowledge regarding the benefits and drawbacks of implementing governmental computer systems in the cloud, and what should be covered during the planning of such projects. By creating a list of benefits and drawbacks and rating the impact each of these have on a CC implementation, a map of the current status of CC usability for governmental agencies has been created. Furthermore, calculations have been conducted to study how CC compares to conventional computing models. Even though the results favoured the conventional computing model for most of the categories that were used, the private cloud deployment model ended up with the best results in most of the scenarios we assessed. This is most likely caused by the private clouds high overall scores in all the categories used, compared to the conventional systems lacking results in the availability category. We believe that the applicability of cloud systems will increase the future, as the technology matures and best practices are created.

The second part of our study set out to create a guided document for the creation of requirements for governmental CC implementations. The goal was to identify the areas of

a planning phase which demanded additional attention in the case of CC. Special attention was given to the choice of cloud-specific technologies and vendors, as well as probable law and compliance regulations that would apply. Additionally, common methods for the planning process, such as the creation of high-level and detailed requirements as well as security requirements, have been commented on from a cloud and governmental viewpoint. This resulted in a document that can be consulted during the planning and designing of a CC environment, not only for governmental needs, which focuses the work on cloud-specific issues and compliance with standards and legal demands.

The document was also created to be applicable to a wide range of project management frameworks, by keeping the overall layout of the document general. The applicability was tested using a fictional CC planning project, with MOF as the project framework. The study showed that the created guide can be implemented into the framework's phases without much modification, and adds to the ability of the framework to manage cloud-based projects. The requirements guide seems well fit to aid existing frameworks with cloud planning and documentation.

Through these three steps we have added to the awareness of the issues and profits of moving services to the cloud, and provided guidance on how the planning of such projects can be performed.

As CC is still a young technology, several ways exist in which the study conducted here could be built upon. Additionally, some improvements to the study can also be mentioned.

The list of benefits and drawbacks created during this study is thorough, but more may still be found. The further exploration of benefits and drawbacks should continue in order to ensure that the list evolves in line with the technology's advancements. Such advancements will also change the premises for the existing benefits and drawbacks, and the impact levels of these should be re-assessed as the cloud matures. This will ensure that an up-to-date picture of cloud applicability is retained.

Ways of including the hybrid cloud environment better into such a list would also be beneficial, as the existing list and rating could not assess the hybrid cloud's strengths and weaknesses in an adequate way. Doing this combined with rating more of the benefits and drawbacks specifically for the hybrid cloud model based on its composite features would both strengthen the method and the results.

Another point of further study for this section of the thesis would be to explore other ways of using the results from the rating. These could for instance be used to assess cloud service models, as well as CSPs.

Furthermore, the questionnaire should be refined to better reflect the benefits and drawbacks that exist, for instance by simplifying and clarifying the questions asked. A larger number of respondents should participate to better reflect a population's collective perception. This could be achieved by motivating the respondents to answer the questionnaire. Offering a gift in return for an answer, or entry into a raffle with a more substantial prize could improve the numbers.

One could also benefit from conducting in-depth interviews with the respondents to understand the choices made in the questionnaire. This would enable us to document the assumptions each respondent made prior to answering questions, and gain better insight into each individual's impression of CC.

Regarding the creation of the requirements document and its testing, some possible

improvements can also be mentioned here. First of all, assessing and reviewing the steps of the guide based on existing project documentation from successful implementations could result in improvements of these. Additional steps could also be added as a result. New examples for the requirements guide might also be found as a result, improving the guide's ability to assist during CC project work.

The guide also needs to be tested using additional frameworks, in order to explore and improve its ability to apply to a broad spectrum of project frameworks. Additionally, documents explaining how such documentation can be best applied to specific frameworks would also be of value. Such documentation could potentially be recreated as modified frameworks specifically for CC project management.

Bibliography

[1] Gantz, J. & Reinsel, D. June 2011. Extracting value from chaos. *IDC research report IDC research report, Framingham, MA, June. Retrieved September* http://tinyurl.com/3f56u9t, 19, 1–12. Accessed: 2011.12.7.

[2] Cameron, S. August 2011. Digital evidence. http://tinyurl.com/cchzuf8. Accessed: 2011.11.29.

[3] Brunker, M. November 2011. Digital evidence becoming central in criminal cases. http://tinyurl.com/bldrfqv. Accessed: 2011.11.15.

[4] Leyden, J. November 2010. German 'hacker' uses rented computing to crack hashing algorithm. http://tinyurl.com/385bj5g. Accessed: 2011.12.07.

[5] Samson, T. September 2011. Botnet rentals reveal the darker side of the cloud. http://tinyurl.com/3wt5qrh. Accessed: 2011.12.07.

[6] Flaglien, A. O. Cross-computer malware detection in digital forensics. Master's thesis, Gjøvik University College, 2010.

[7] Borgenholt, G., Alseth, T. I., & Øksnekvad, O. T. 2011. Jungle. Gjøvik University College Bachelor Thesis.

[8] Mallasvik, A. F., Rosenlund, R. D. H., Flaglien, A. O., & Evensen, K. E. May 2008. Information security test laboratory. Gjøvik University College Bachelor Thesis.

[9] AccessData Group. 2012. Ad lab. http://accessdata.com/products/computer-forensics/lab. Accessed: 2012.06.13.

[10] Digital Intelligence. 2012. Fred dx. http://www.digitalintelligence.com/products/freddx2r/. Accessed: 2012.06.13.

[11] AccessData Group. 2011. Case study: økokrim. http://accessdata.com/downloads/media/CASE_STUDY_OKOKRIM.pdf. Accessed: 2012.06.22.

[12] Peshkin, A. 1993. The goodness of qualitative research. *Educational Researcher*, 22(2), 23–29.

[13] Leedy, P. D. & Ormrod, J. E. March 2004. *Practical Research: Planning and Design (8th Edition)*. Prentice Hall, 8 edition.

[14] Tellefsen, A. June 2011. Cloud computing forensics: State of the art. Course paper at Gjøvik University College. IMT 4022 Digital Forensics 2.

[15] Oxford Dictionaries. 2011. http://oxforddictionaries.com/. Accessed: 2012.01.20.

[16] Kent, K., Chevalier, S., Grance, T., & Dang, H. Guide to integrating forensic techniques into incident response. Technical report, National Institute of Standards and Technology, Gaithersburg, MD 20899-8930 http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf, August 2006. Accessed: 2011.05.05.

[17] Jones, A. & Valli, C. 2008. *Building A Digital Forensic Laboratory: Establishing and Managing a Successful Facility*. Butterworth Heinemann. Elsevier Science.

[18] Wikipedia. 2012. History of criminal justice — wikipedia, the free encyclopedia. http://tinyurl.com/cjte7zr. Accessed: 2012.06.01.

[19] Elvidge, S. June 2011. Forensic Cases: Colin Pitchfork, First Exoneration Through DNA. http://tinyurl.com/cmp4vyv. Accessed: 2012.06.01.

[20] Thompson, W. 1996. Accepting lower standards: The National Research Council's second report on forensic DNA evidence. *Jurimetrics*, 37, 405.

[21] Kedziora, M. August 2010. Computer forensics history. http://tinyurl.com/d72pkqy. Accessed: 2012.01.18.

[22] The International Association of Computer Investigative Specialists. 2012. Home: About us. https://www.iacis.com/. Accessed: 2012.06.01.

[23] Stoll, C. 1989. *The cuckoo's egg*. Doubleday New York.

[24] Rogers, M. K. 2011. DCSA: A Practical Approach to Digital Crime Scene Analysis. http://www2.tech.purdue.edu/cit/Courses/cit556/readings/DCSA.pdf. Accessed: 2012.06.13.

[25] Sand, L. A. & Tellefsen, A. December 2011. Forensic Report of the events leading up to the terrorist attack 22 july - Norway. Course paper at Gjøvik University College. IMT 4881 Specialization course.

[26] Brendan & Dolan-Gavitt. 2008. Forensic analysis of the windows registry in memory. In *The Proceedings of the Eighth Annual DFRWS Conference*, volume 5, Supplement, S26 – S32.

[27] Sutherland, I., Evans, J., Tryfonas, T., & Blyth, A. April 2008. Acquiring volatile operating system data tools and techniques. *SIGOPS Oper. Syst. Rev.*, 42, 65–73.

[28] Volatile Systems, LLC. 2006. The volatility framework: Volatile memory artifact extraction utility framework. www.volatilesystems.com/default/volatility. Accessed: 2011.09.29.

[29] Halderman, J., Schoen, S., Heninger, N., Clarkson, W., Paul, W., Calandrino, J., Feldman, A., Appelbaum, J., & Felten, E. 2009. Lest we remember: cold-boot attacks on encryption keys. *Communications of the ACM*, 52(5), 91–98.

[30] Lu, R., Lin, X., Liang, X., & Shen, X. S. 2010. Secure provenance: the essential of bread and butter of data forensics in cloud computing. In *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '10, 282–292, New York, NY, USA. ACM.

[31] Mell, P. & Grance, T. January 2011. NIST Special Publication 800-145: The NIST Definition of Cloud Computing (Draft). `http://tinyurl.com/4dmorxe`. Accessed: 2011.04.28.

[32] CloudRecruiting.net. 2011. The future is in the cloud. `http://cloudrecruiting.net/the-future-is-in-the-clouds/`. Accessed: 2011.05.04.

[33] Licklider, J. April 1963. Memorandum for: Members and affiliates of the intergalactic computer network subject: Topics for discussion at the forthcoming meeting. advanced research projects agency(arpa). `http://tinyurl.com/85k6xcb`. Accessed: 2011.04.28.

[34] Compaq. 1996. Internet solutions division strategy for cloud computing. `http://www.technologyreview.com/business/38987/`. Accessed: 2012.01.19.

[35] Foster, I. & Kesselman, C. 1999. *The grid: blueprint for a new computing infrastructure*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[36] Search for Extraterrestrial Intelligence (SETI). 2011. What is SETI@home? `http://setiathome.berkeley.edu/`. Accessed: 2011.06.18.

[37] VMware. 2012. History of virtualization. `http://www.vmware.com/virtualization/history.html`. Accessed: 2012.05.27.

[38] Ferre', M. R., Pomeroy, D., Wahlstrom, M., & Watts, D. *Virtualization on the IBM System x3950 Server*. IBM, June 2006. `http://www.redbooks.ibm.com/redbooks/pdfs/sg247190.pdf` Accessed: 2011.08.06.

[39] Jansen, W. & Grance, T. Guidelines on security and privacy in public cloud computing. NIST Special Publication 800-144, National Institute of Standards and Technology, December 2011. `http://www.nist.gov/manuscript-publication-search.cfm?pub_id=909494` Accessed: 2012.01.30.

[40] CSA. Security guidance for critical areas of focus in cloud computing v3.0. Technical report, Cloud Security Alliance, December 2009. `http://tinyurl.com/8ykzd2x` Accessed: 2012.03.21.

[41] Xen. 2011. Xen cloud platform. `http://www.xen.org/products/cloudxen.html`. Accessed: 2011.05.24.

[42] VCODYNE. Remote private cloud. `http://www.vcodyne.com/private-cloud/remote-private-cloud.html`. Accessed: 2011.08.04.

[43] Peng, J., Zhang, X., Lei, Z., Zhang, B., Zhang, W., & Li, Q. dec. 2009. Comparison of several cloud computing platforms. In *Information Science and Engineering (ISISE), 2009 Second International Symposium on*, 23 –27.

[44] Iosup, A., Yigitbasi, N., & Epema, D. may 2011. On the performance variability of production cloud services. In *Cluster, Cloud and Grid Computing (CCGrid), 2011 11th IEEE/ACM International Symposium*, 104 –113.

[45] Maguire, J., Vance, J., & Harvey, C. 85 cloud computing vendors shaping the emerging cloud. Technical report, iTManagement Tech.Rep, August 2009. `http://tinyurl.com/79uk9or` Accessed: 2012.04.24.

[46] CitrixSystems. 2011. Xen cloud platform. `http://www.xen.org/products/cloudxen.html`. Accessed: 2011.06.18.

[47] Xen. 2011. Xen overview. `http://wiki.xensource.com/xenwiki/XenOverview`. Accessed: 2011.05.24.

[48] Citrix Systems Inc. 2012. Cloudstack - Open Source Cloud Computing Software. `http://cloudstack.org/cloudstack-overview.html`. Accessed: 2012.05.21.

[49] Nurmi, D. et. al. 2009. The eucalyptus open-source cloud-computing system. In *CCGRID'09. 9th IEEE/ACM International Symposium on Cluster Computing and the Grid*, 124–131. IEEE.

[50] Foran, J. 2012. Comparing open source cloud platforms: Openstack versus eucalyptus. `http://tinyurl.com/6w9yand`. Accessed: 2012.05.22.

[51] Cerbelaud, D., Garg, S., & Huylebroeck, J. 2009. Opening the clouds: qualitative overview of the state-of-the-art open source vm-based cloud management platforms. In *Proceedings of the 10th ACM/IFIP/USENIX International Conference on Middleware*, Middleware '09, 22:1–22:8, New York, NY, USA. Springer-Verlag New York, Inc.

[52] Eucalyptus Systems Inc. 2012. Eucalyptus 3 compatibility matrix. `www.eucalyptus.com/eucalyptus-cloud/iaas/compatibility`. Accessed: 2012.05.22.

[53] Shekhawat, D. April 2012. Windows server 2003 instance problem. `https://engage.eucalyptus.com/customer/portal/questions/281877`. Accessed: 2012.05.22.

[54] Amazon.com. 2011. Amazon Elastic Compute Cloud (Amazon EC2). `http://aws.amazon.com/ec2/`. Accessed: 2011.05.24.

[55] Google. 2011. Google apps for business. `http://www.google.com/apps/intl/en/business/`. Accessed: 2011.05.24.

[56] Google. 2011. Google cloud services: Homepage for gae. `http://www.google.com/enterprise/cloud/appengine/`. Accessed: 2011.11.14.

[57] Microsoft. 2011. Windows Azure. http://www.microsoft.com/windowsazure/. Accessed: 2011.05.24.

[58] Cisco. 2011. Private cloud. Website for Cisco Cloud Solutions http://www.cisco.com/en/US/netsol/ns983/index.html. Accessed: 2011.11.14.

[59] Citrix. 2011. Citrix cloud platform. http://tinyurl.com/nnos7p. Accessed: 2011.11.14.

[60] IBM. 2011. Introduction to the IBM SmartCloud solution. http://www.ibm.com/cloud-computing/us/en/index.html. Accessed: 2011.11.14.

[61] VMware. 2012. VMware vSphere™ for Enterprise. http://tinyurl.com/lo6kes. Accessed: 2012.06.14.

[62] OpenStack. 2012. Open source software for building private and public clouds. http://openstack.org/. Accessed: 2012.06.14.

[63] Star. 2011. Cloud computing services for uk business. http://www.star.co.uk/. Accessed: 2012.01.19.

[64] Fu, Xinwn et al. June 2010. Cyber crime scene investigations (c'2) through cloud computing. In *Distributed Computing Systems Workshops (ICDCSW), 2010 IEEE 30th International Conference on*, 26 –31.

[65] Reilly, D., Wren, C., & Berry, T. nov. 2010. Cloud computing: Forensic challenges for law enforcement. In *Internet Technology and Secured Transactions (ICITST), 2010 International Conference for*, 1 –7.

[66] Ahmed, S. & Raja, M. December 2010. Tackling cloud security issues and forensics model. In *High-Capacity Optical Networks and Enabling Technologies (HONET), 2010*, 190 –195.

[67] Hobson, E. W. 2010. Digital investigations in the cloud. http://tinyurl.com/cbos9xk. Accessed: 2011.05.07.

[68] Biggs, S. & Vidalis, S. nov. 2009. Cloud computing: The impact on digital forensic investigations. In *Internet Technology and Secured Transactions, 2009. ICITST 2009. International Conference for*, 1 –6.

[69] Wolthusen, S. D. 2009. Overcast: Forensic discovery in cloud environments. In *Proceedings of the 2009 Fifth International Conference on IT Security Incident Management and IT Forensics*, 3–9, Washington, DC, USA. IEEE Computer Society.

[70] Hay, B. & Nance, K. April 2008. Forensics examination of volatile system data using virtual introspection. *SIGOPS Oper. Syst. Rev.*, 42, 74–82.

[71] Lynch, C. A. January 2001. When documents deceive: Trust and provenance as new factors for information retrieval in a tangled web. *J. Am. Soc. Inf. Sci. Technol.*, 52, 12–17.

[72] Ghemawat, S., Gobioff, H., & Leung, S.-T. October 2003. The Google File System. http://labs.google.com/papers/gfs.html. Accessed: 2011.05.13.

[73] Jones, K. J., Bejtlich, R., & Rose, C. W. 2006. *Real Digital Forensics: Computer Security and Incident Response*. Addison-Wesley Professional.

[74] Paquette, S., Jaeger, P. T., & Wilson, S. C. 2010. Identifying the security risks associated with governmental use of cloud computing. *Government Information Quarterly*, 27(3), 245–253.

[75] Zachariassen, E. March 2009. Derfor stoppet politiet. http://www.tu.no/it/article204794.ece. Accessed: 2012.02.22.

[76] Sungard. 2012. Recover2cloud. http://tinyurl.com/7579gsm. Accessed: 2012.02.22.

[77] Chandrasekhar, B. March 2011. What is cloudbursting? http://cloudsecurity.trendmicro.com/what-is-cloudbursting/. Accessed: 2012.02.24.

[78] Cisco. 2010. Service provider builds secure public cloud for mission-critical applications. http://tinyurl.com/6q3h3zj. Accessed: 2012.02.24.

[79] Garera, S., Provos, N., Chew, M., & Rubin, A. D. 2007. A framework for detection and measurement of phishing attacks. In *Proceedings of the 2007 ACM workshop on Recurring malcode*, WORM '07, 1–8, New York, NY, USA. ACM.

[80] Amazon. 2012. Amazon ec2 pricing. http://aws.amazon.com/ec2/pricing/. Accessed: 2012.02.24.

[81] Armbrust, Michael et al. Above the Clouds: A Berkeley View of Cloud Computing. http://radlab.cs.berkeley.edu/publication/285, February 2009. Accessed: 2012.02.09.

[82] Amazon. 2012. Aws import/export. http://aws.amazon.com/importexport/. Accessed: 2012.02.24.

[83] Tapper, D. March 2011. Prepare and assess readiness for cloud services. http://tinyurl.com/7h8qrzn. Accessed: 2012.03.01.

[84] Cloudswitch. 2012. The enterprise gateway to the cloud. http://www.cloudswitch.com/. Accessed: 2012.03.01.

[85] ISO. 2005. ISO/IEC-27001:2005. http://tinyurl.com/3lp3ott. Accessed: 2012.01.19.

[86] ISO. 2005. ISO/IEC-27002:2005. http://www.iso.org/iso/catalogue_detail?csnumber=50297. Accessed: 2010.12.03.

[87] Ristenpart, T., Tromer, E., Shacham, H., & Savage, S. 2009. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In *Proceedings of the 16th ACM conference on Computer and communications security*, CCS '09, 199–212, New York, NY, USA. ACM.

[88] Cosic, J. & Baca, M. may 2010. (im)proving chain of custody and digital evidence integrity with time stamp. In *MIPRO, 2010 Proceedings of the 33rd International Convention*, 1226 –1230.

[89] McMillan, R. December 2010. Group used 30,000-node botnet in mastercard, paypal attacks. http://tinyurl.com/7njw9qa.
Accessed: 2012.02.08.

[90] Tristin, H. January 2012. Hackers attack fbi, justice department websites after megaupload shutdown. http://tinyurl.com/6v4e2ut.
Accessed: 2012.02.08.

[91] Government Administration Services. 2012. Government.no information from the government and tthe ministries. http://www.regjeringen.no/en.html?id=4.
Accessed: 2012.02.28.

[92] Norwegian Police forces. 2012. Anmeld et forhold.
https://www.politi.no/tjenester/anmeld_et_forhold/.
Accessed: 2012.02.28.

[93] Ambler, S. W. 2009. Introduction to user stories.
http://www.agilemodeling.com/artifacts/userStory.htm.
Accessed: 2012.05.03.

[94] Malan, R. & Bredemeyer, D. 2001. Defining non-functional requirements. http://www.bredemeyer.com/pdf_files/NonFunctReq.PDF.
Accessed: 2012.05.01.

[95] Interpol. 2012. Databases.
http://www.interpol.int/INTERPOL-expertise/Databases.
Accessed: 2012.03.26.

[96] Statistics Norway. March 2011. Offences reported to the police, 2010.
http://www.ssb.no/english/subjects/03/05/lovbrudda_en/.
Accessed: 2012.03.27.

[97] Yourdon, E. 2000. *Modern structured analysis*. Prentice Hall PTR.

[98] EUR-Lex. 1995. Data protection directive.
http://tinyurl.com/6gpkrav. Accessed: 2011.09.13.

[99] EUR-Lex. 2002. eprivacy directive.
http://tinyurl.com/66wghdd. Accessed: 2012.03.21.

[100] European Data Protection Supervisor. 2012. Legislation.
http://tinyurl.com/7kmo6rk. Accessed: 2012.03.21.

[101] European Commission. January 2012. Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses. http://tinyurl.com/7mxg9n7. Accessed: 2012.04.17.

[102] European Commission. January 2012. How will the eu's data protection reform make international cooperation easier? http://tinyurl.com/766pj5q. Accessed: 2012.04.17.

[103] EUR-Lex. November 2008. Council framework decision 2008/977/jha. http://tinyurl.com/6rg5exq. Accessed: 2012.04.29.

[104] Pinsent Masons LLP. December 2012. Further data protection proposals for law enforcement will clarify existing law, expert says. http://tinyurl.com/8643k8d. Accessed: 2012.04.29.

[105] European Commission. June 1989. Council directive 89/391/eec. http://tinyurl.com/6nzckb4. Accessed: 2012.05.05.

[106] Bureau of Consumer Protection. 2012. Gramm-leach-bliley act. business.ftc.gov/privacy-and-security/gramm-leach-bliley-act. Accessed: 2012.03.08.

[107] U.S. Department of Health & Human Services. 2012. Summary of the HIPAA Privacy Rule. www.hhs.gov/ocr/privacy/hipaa/understanding/summary/index.html. Accessed: 2012.03.08.

[108] The Senate and House of Representatives of the United States of America. July 2002. Public law 107–204: Sarbanes-oxley act of 2002. www.sec.gov/about/laws/soa2002.pdf. Accessed: 2012.03.22.

[109] Scholl, Matthew et al. October 2008. An Introductory Resource Guide for Implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule. By NIST. http://tinyurl.com/cd29c2. Accessed: 2011.06.17.

[110] Smedinghoff, T. J. & Hamady, L. E. 2009. New data security regulations create compliance challenges for business. http://tinyurl.com/84s7l33. Accessed: 2012.03.22.

[111] Hustinx, P. April 2010. Data protection and cloud computing under eu law. http://tinyurl.com/6wdxffp. Accessed: 2012.03.21.

[112] Mahle, S. 1999. An introduction to daubert v. merrell dow. http://www.daubertexpert.com/basics_daubert-v-merrell-dow.html. Accessed: 2012.06.15.

[113] Legal Information Institute: Cornell University. March 1993. Daubert v. merrell dow pharmaceuticals (92-102), 509 u.s. 579 (1993). http://www.law.cornell.edu/supct/html/92-102.ZS.html. Accessed: 2012.06.15.

[114] U.S. General Services Administration. December 2011. FedRAMP Third Party Assessment Organization (3PAO). www.gsa.gov/graphics/staffoffices/3PAO_Program_Description.docx. Accessed: 2011.12.18.

[115] U.S. General Services Administration. January 2012. Fedramp security controls. http://tinyurl.com/7l63f55. Accessed: 2011.12.18.

[116] Calder, A. & Watkins, S. 2008. *IT Governance: A Manager's Guide to Data Security and ISO27001/ISO27002*. Kogan Page Limited.

[117] CCRA. 2010. Common Criteria. http://www.commoncriteriaportal.org/cc/. Accessed: 2011.09.13.

[118] Common Criteria. July 2009. Part 1: Introduction and general model. www.commoncriteriaportal.org/files/ccfiles/CCPART1V3.1R3.pdf. Accessed: 2012.03.26.

[119] Common Criteria. October 1999. User guide. www.niap-ccevs.org/cc_docs/cc_users_guide.pdf. Accessed: 2012.03.23.

[120] Saltzer, J. & Schroeder, M. September 1975. The protection of information in computer systems. *Proceedings of the IEEE*, 63(9), 1278 – 1308.

[121] Neumann, P. G. June 2000. Practical architectures for survivable systems and networks (phase-two final report). www.csl.sri.com/users/neumann/survivability.pdf. Accessed: 2010.11.21.

[122] Stoneburner, G., Hayden, C., & Feringa, A. Engineering principles for information technology security. NIST Special Publication 800-27 Rev A, National Institute of Standards and Technology, June 2004. csrc.nist.gov/publications/nistpubs/800-27A/SP800-27-RevA.pdf Accessed: 2011.09.13.

[123] PCI Security Standards Council. October 2010. PCI DSS Requirements and Security Assessment Procedures Version 2.0. https://www.pcisecuritystandards.org/documents/pci_dss_v2.pdf. Accessed: 2012.03.26.

[124] Mogull, R. & Adrian, L. September 2011. Data security lifecycle 2.0. https://securosis.com/blog/data-security-lifecycle-2.0. Accessed: 2012.03.27.

[125] Gutmann, P. 1996. Secure deletion of data from magnetic and solid-state memory. In *In Proceedings of the 6th USENIX Security Symposium*, 77–89.

[126] Kissel, Richard et al. September 2006. NIST Special Publication 800-88: Guidelines for Media Sanitization. csrc.nist.gov/publications/nistpubs/800-88/NISTSP800-88_rev1.pdf. Accessed: 2011.06.11.

[127] Choudhury, A., Kumar, P., Sain, M., Lim, H., & Jae-Lee, H. dec. 2011. A strong user authentication framework for cloud computing. In *Services Computing Conference (APSCC), 2011 IEEE Asia-Pacific*, 110 –115.

[128] OASIS Security Services TC. SAML Technical Overview. Technical report, OASIS, March 2008. http://tinyurl.com/psu8tg Accessed: 2012.03.28.

[129] ISO. 2011. ISO/IEC-27005:2011.
http://www.iso.org/iso/catalogue_detail?csnumber=56742.
Accessed: 2012.05.25.

[130] Microsoft. 2012. Microsoft operations framework (mof) 4.0.
http://www.microsoft.com/en-us/download/details.aspx?id=17647.
Accessed: 2012.05.03.

[131] European Commission. July 1992. Treaty on European Union.
http://eur-lex.europa.eu/en/treaties/dat/11992M/htm/11992M.html.
Accessed: 2012.04.30.

[132] Tellefsen, A. August 2011. A forensic environment based on the xen cloud platform. findings and challenges. Course paper at Gjøvik University College. IMT 4641 Computational Forensics.

[133] NokiaDevForum. December 2009. Conceptual design in 2 minutes.
http://www.youtube.com/watch?v=Jo2SG4JhohQ. Accessed: 2011.04.28.

[134] Rosen, K. H. 2002. *Discrete Mathematics and Its Applications*. McGraw-Hill Higher Education, 5th edition.

A Benefits and Drawbacks Result Tables

This appendix consists of two sections, A.1 where the impact values from the in-depth study (3.2,3.3) is used, and A.2 where the impact values from the questionnaire results (G) are used. Five main tables (disregarding Table 16) are presented in each case, one for each of the cloud deployment models we use and one for conventional server/client implementations.

The tables consist of five major columns, the first from the left indicates (from the top) the deployment model we focus on, which is followed by the models we compare it with. The second column, ”Benefits”, shows us the number of positive benefit points the deployment model in question has gathered against all the other deployment models in each of our four categories. The third column, ”Drawbacks”, will in turn show us the number of negative drawback points the deployment model has gathered. The forth column from the left, called ”SUM”, shows us the calculation of benefit points minus drawback points in each category compared with the given deployment model, while the fifth column gives us an overall score, combining all the scores from each of the categories.

In this appendix both the benefit and drawback values has been raised to the power of two. The model states that the benefits or drawbacks rated as ”High” should have a significant impact on the results, while the subsequent rating levels (medium and low) should have lower influence. Table 16 explains the values that are used.

Table 16: Overview of benefit and drawback values for the Squared Values model

Impact Level	Benefits	Drawbacks
High	9	9
Medium	4	4
Low	1	1

A.1 Benefits and Drawbacks Result Tables: In-depth Study

Table 17: Private Cloud Results Overview with Squared Values

Private Cloud	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Hosted-Private	2	4	1	0	2	0	1	0	0	4	0	0	4
Public	10	8	1	0	2	0	1	0	8	8	0	0	16
Conventional	0	4	9	0	4	1	17	0	-4	3	-8	0	-9

Table 18: Hosted-Private Cloud Results Overview with Squared Values

Hosted-Private Cloud	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Private	0	0	4	1	5	9	8	12	-5	-9	-4	-11	-29
Public	9	4	0	0	0	0	0	0	9	4	0	0	13
Conventional	0	4	13	1	9	10	25	12	-9	-6	-12	-11	-38

Table 19: Public Cloud Results Overview with Squared Values

Public Cloud	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Private	4	9	13	1	14	10	8	12	-10	-1	5	-11	-17
Hosted-Private	4	9	9	0	9	1	0	0	-5	8	9	0	12
Conventional	4	13	22	1	18	11	25	12	-14	2	-3	-11	-26

Table 20: Conventional Results Overview with Squared Values

Conventional	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Private	0	0	0	0	0	9	4	0	0	-9	-4	0	-13
Hosted-Private	2	4	1	0	2	9	5	0	0	-5	-4	0	-9
Public	10	4	1	0	2	9	5	0	8	-5	-4	0	-1

A.2 Benefits and Drawbacks Result Tables: Questionnaire

This appendix consists of the same table layout as in Appendix A.1; five main tables, one for each of the cloud deployment models we use and one for conventional server/client implementations. The same numbers are used for the calculations as in Table 16. For more information, read the introduction to A.1.

Table 21: Private Cloud Results Overview with Squared Values

Private Cloud	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Hosted-Private	13	9	9	0	8	0	4	0	5	9	5	0	19
Public	8	15,25	9	0	8	0	4	0	0	15,25	5	0	20,25
Conventional	0	9	9	0	4	1	22	0	-4	8	-13	0	-9

Table 22: Hosted-Private Cloud Results Overview with Squared Values

Hosted-Private Cloud	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Private	0	0	9	9	18	6,25	18	22	-18	-6,25	-9	-13	-46,25
Public	4	6,25	0	0	0	0	0	0	4	6,25	0	0	10,25
Conventional	0	9	18	9	22	7,25	40	22	-22	1,75	-22	-13	-55,25

Table 23: Public Cloud Results Overview with Squared Values

Public Cloud	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Private	9	9	18	9	27	10,25	18	22	-18	-1,25	0	-13	-32,25
Hosted-Private	9	9	9	0	9	4	0	0	0	5	9	0	14
Conventional	9	18	27	9	31	11,25	40	22	-22	6,75	-13	-13	-41,25

Table 24: Conventional Results Overview with Squared Values

Conventional	Benefits				Drawbacks				SUM				
	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	Access	Avail.	Infra.	Integrity	SUM TOTAL
Private	0	0	0	0	0	9	6,25	0	0	-9	-6,25	0	-15,25
Hosted-Private	13	9	9	0	8	9	10,25	0	5	0	-1,25	0	3,75
Public	8	5,5	9	0	8	9	10,25	0	0	-3,5	-1,25	0	-4,75

B Result Tables: Other calculation methods

Since the number of tables that were generated during this thesis is quite high, we have decided to only include the most important ones text form. All other tables have been included in two .zip-folders containing two sets of Excel sheets.

RQ1calculationsInDepth.zip includes all tables from the in-depth study.

RQ1calculationsQuest7Resp.zip includes all tables from the questionnaire study.

Both the Excel folders are laid out in the following way:

The BenDrawTable.xls includes one sheet for benefits and one for drawbacks, and the impact rating of all of these. The two sheets are used by all other Excel files in the folder to calculate results.

The five remaining Excel files all calculate the same results using different impact rating values:

- DrawbBenCompDoubleBen.xlsx uses double the original values for benefits
- DrawbBenCompDoubleDraw.xlsx uses double the original values for drawbacks
- DrawbBenCompSingleBoth.xlsx uses the original values for both benefits and drawbacks
- DrawbBenCompSquared.xlsx uses squared values for both benefits and drawbacks
- DrawbBenCompSquaredBenDoubleDraw.xlsx uses squared values for benefits, and double the original values for drawbacks

The calculations made in these Excel files were explained in Section 3.5.1.

C Category Comparison Tables

The tables presented here switches the focus towards each of the categories in the AAII model, explained in Chapter 3, by using the ”SUM” column from the tables in Appendix A.1 and A.2. A detailed description of the calculation method is found in Section 3.5.1.

Here follows the category comparison tables for the squared calculation method. For an overview of the category comparison tables of all the calculation methods used, these are found in the Excel files referenced to in Appendix B.

C.1 Comparison Table for Deployment Models using Squared Values

Table 25: Access Comparison Table using Squared Values

Access	Private	Hosted-Private	Public	Conventional
Private	X	-5	-18	4
Hosted-Private	5	X	-14	9
Public	18	14	X	22
Conventional	-4	-9	-22	X
SUM	19	0	-54	35

Table 26: Availability Comparison Table using Squared Values

Avail.	Private	Hosted-Private	Public	Conventional
Private	X	-13	-9	-12
Hosted-Private	13	X	-14	1
Public	9	-4	X	-7
Conventional	12	-1	7	X
SUM	34	-18	-16	-18

Table 27: Infrastructure Comparison Table using Squared Values

Infras.	Private	Hosted-Private	Public	Conventional
Private	X	-4	5	4
Hosted-Private	4	X	9	8
Public	-5	-9	X	-1
Conventional	-4	-8	1	X
SUM	-5	-21	15	11

Table 28: Integrity Comparison Table using Squared Values

Integrity	Private	Hosted-Private	Public	Conventional
Private	X	-11	-11	0
Hosted-Private	11	X	0	11
Public	11	0	X	11
Conventional	0	-11	-11	X
SUM	22	-22	-22	22

C.2 Comparison Table for Deployment Models using Squared Values: Questionnaire Values

This represents the same calculation performed in C.1, using the values collected from the questionnaire A.2.

Table 29: Access Comparison Table using Squared Values from the Questionnaire

Access	Private	Hosted-Private	Public	Conventional
Private	X	-23	-18	4
Hosted-Private	23	X	-4	27
Public	18	4	X	22
Conventional	-4	-27	-22	X
SUM	37	-46	-44	53

Table 30: Availability Comparison Table using Squared Values from the Questionnaire

Avail.	Private	Hosted-Private	Public	Conventional
Private	X	-15,25	-16,5	-17
Hosted-Private	15,25	X	-4	-1,75
Public	16,5	1,25	X	-10,25
Conventional	17	1,75	10,25	X
SUM	48,75	-12,25	-10,25	-29

Table 31: Infrastructure Comparison Table using Squared Values from the Questionnaire

Infras.	Private	Hosted-Private	Public	Conventional
Private	X	-14	-5	6,75
Hosted-Private	14	X	9	20,75
Public	5	-9	X	11,75
Conventional	-6,75	-20,75	-11,75	X
SUM	12,25	-43,75	-7,75	39,25

Table 32: Integrity Comparison Table using Squared Values from the Questionnaire

Integrity	Private	Hosted-Private	Public	Conventional
Private	X	-13	-13	0
Hosted-Private	13	X	0	13
Public	13	0	X	13
Conventional	0	-13	-13	X
SUM	26	-26	-26	26

D Best Deployment Model Tables

These collection of figures presents the SUM rows of the tables in Appendix C.1 and C.2 in Table 33 and Table 34 respectively. The tables only represent the results from using the squared calculation method, but the rest of these can be found in the Excel files referenced to in Appendix B.

Table 33: Best Deployment Model when using Squared Values

	Private	H-Priv.	Public	Conv.	Best Deployment Model
Access	19	0	-54	**35**	Conventional
Availability	**34**	-18	-16	-18	Private
Infrastructure	-5	-21	**15**	11	Public
Integrity	**22**	-22	-22	**22**	Private/Conventional

Table 34: Questionnaire: Best Deployment Model when using Squared Values

	Private	H-Priv.	Public	Conv.	Best Deployment Model
Access	37	-46	-44	**53**	Conventional
Availability	**48,75**	-12,25	-10,25	-29	Private
Infrastructure	12,25	-47,75	**-7,75**	39,25	Public
Integrity	**26**	-26	-26	**26**	Private/Conventional

E Scenario Score Calculation

The Excel file ”ScenarioCalculations.xslx” is appended to this document. It contains two Excel sheets: One using the results from the in-depth study, and one using the results from the questionnaire. The layout of the sheets are similar, where the tables marked with ”1” are both Table 3, the category priority values for each scenario. The tables marked with ”2” are the result Tables 4 (for the in-depth study) and 5 (for the questionnaire study). These tables are explained in their respective sections in the document.

The tables marked with ”3” in each sheet are the results of a matrix multiplication calculation between Table ”1” and ”2” in the given sheet. This is done by multiplying the rows of Table ”1” with the columns of Table ”2”. A thorough introduction to matrices and matrix multiplication can be found in the book by Rosen [134].

F Questionnaire for Research Question One

1. Single-tenant systems, which is a principle stating that a hardware device only delivers its services to a single client at the time, is in your opinion how important for limiting vulnerabilities in cloud computing?
    - Low importance
    - Medium importance
    - High importance
    - I don't know
2. Systems hosted on-premise behind a clients firewall does not have the main communications channel exposed to the Internet. How effective do you perceive this type of implementation to be for lowering possible vulnerabilities?
    - Low effect
    - Medium effect
    - High effect
    - I don't know
3. Clients hosting systems on-premise has direct access to hardware, and full control over its maintenance. To what extent do you believe that this impacts the possibility to correct errors and failures in a positive direction compared to off-premise hosted systems?
    - Low
    - Medium
    - High
    - I don't know
4. Using a public cloud infrastructure means that one might access a hosted resource from anywhere, which might ease the process of cooperation initiatives and an external workforce. To what extent do you think public cloud implementations elevate user's ability to access their resources?
    - Low
    - Medium
    - High
    - I don't know
5. By having access to both a private and a public cloud (hybrid cloud computing), one might distribute the workload and storage based on the confidentiality of the data handled. How would you describe the impact of which implementing such a solution elevates the possibility of making resources available for authorized personnel?

- Low
- Medium
- High
- I don't know

6. How would you assess the importance of implementing single-tenant systems in order to assure that other clients cannot deplete or otherwise disrupt the resource flow?

   - Low importance
   - Medium importance
   - High importance
   - I don't know

7. If a major failure occurs at the Internet service provider or the cloud service provider, on-premise systems will still be operational and reachable by on-premise users. How important would you claim this argument to be for keeping the cloud resources available for its users?

   - Low importance
   - Medium importance
   - High importance
   - I don't know

8. In cloud computing, applications and services are not hosted on specific servers, rather utilizing a resource pool of several physical servers. Hosted content will therefore potentially not be discontinued if a hardware device fails. Which level of importance would you assess this statement to have when discussing cloud computing versus conventional server/client implementations?

   - Low importance
   - Medium importance
   - High importance
   - I don't know

9. When discussing hybrid clouds, to what would you value the ability of still using one part of the cloud when the other part is incapacitated?

   - Low value
   - Medium value
   - High value
   - I don't know

10. Knowing that clients using public cloud infrastructures can order and deploy virtual machine instances on the fly without considering hardware, how substantial would do you think this characteristic is for deploying and maintaining services in a computer network?

- Low
- Medium
- High
- I don't know

11. How important would you claim the ability to locally import and export data into a computer system to be?

    - Low importance
    - Medium importance
    - High importance
    - I don't know

12. Choosing clouds that have no reserved hardware might result in lower costs from avoiding the purchase of hardware, and the personnel to operate it. It also means that the client can focus the additional resources on the main tasks of the company. How important would you claim this to be for a company?

    - Low importance
    - Medium importance
    - Highly important
    - I don't know

13. How important would you assess a public cloud infrastructures ability to upscale and downscale all consumed resources of a computer system to be, when comparing them to the limited abilities of other cloud infrastructures and conventional server/client systems?

    - Low importance
    - Medium importance
    - High importance
    - I don't know

14. Virtualization allows for entire networks of computers to exist strictly in software, a technology that cloud computing thoroughly exploits. What degree of importance would you claim the possibility of creating, saving, moving or deleting computer instances quickly and easily to be when planning a transition to the cloud?

    - Low importance
    - Medium importance
    - High importance
    - I don't know

15. Hybrid cloud implementations might borrow resources from the public cloud in order to temporarily enhance the performance of the on-premise cloud solution (cloud boosting). How important do you reckon this statement is in the case for hybrid cloud solutions?

- Low importance
- Medium importance
- High importance
- I don't know

16. Most cloud service providers have intricate backup systems built into their infrastructure, and might make that technology available to the clients as built-in services. To what extent do you believe this impact the user's ability to secure their data?

    - Low
    - Medium
    - High
    - I don't know

17. Several clients or parts of a clients company might utilize the same public log-in gateways, regardless of authorization level, making the log-in gateway a prime target for attacks. To what extent do you believe this might impact the security of the cloud solution?

    - Low impact
    - Medium impact
    - High impact
    - I don't know

18. By using cloud services in a given country, the authorities of that nation might demand access to the data, for instance during a forensic investigation. What severity level would you allot to this fact, given that this might result in loss of, or limited access to information stored or services hosted in the cloud?

    - Low
    - Medium
    - High
    - I don't know

19. On-premise hosted private clouds will have direct access to hardware in the same way as conventional server/client implementations, something that might open additional attack vectors, for instance for malicious insiders. How severely do you believe this impact the security of the implementation?

    - Low
    - Medium
    - High
    - I don't know

20. The failure of hardware for systems hosted on-premise might be harder to mitigate for smaller companies in comparison to dedicated cloud providers, which will always have personnel on-call to handle failures. To what extent do you believe this might lever the decision towards avoiding on-premise hosted implementations?

    - Low
    - Medium
    - High
    - I don't know

21. Colocation of several clients on one physical device (so called multi-tenancy) is common in cloud computing. This might lead to leakage of data from one client to another through targeted attacks or faulty cloud infrastructure. To which criticality level would you assess multi-tenancy from a security standpoint?

    - Low
    - Medium
    - High
    - I don't know

22. By distributing data storage and processing over several servers or data centres expands the attack surface of the clients cloud implementation, possibly adding to the number of vulnerabilities. How strongly do you believe the distribution of data processing counts against implementing a cloud solution?

    - Low
    - Medium
    - High
    - I don't know

23. Hybrid clouds share data between private and public cloud solutions, something that might expose the complete implementation to additional attack vectors. How would you describe the level of reservations you might have against implementing a hybrid cloud based on this information?

    - Low
    - Medium
    - High
    - I don't know

24. The cloud service provider becomes a single point of failure for all cloud services a client uses. How severe would you assess this issue to be?

    - Low severity
    - Moderate severity
    - High severity

- I don't know

25. Using a hybrid cloud solution, the public parts of the cloud would still be susceptible to errors on the cloud service provider side. What severity level would you assign to this problem area for of hybrid deployment models?

    - Low
    - Medium
    - High
    - I don't know

26. Cloud implementations might experience thrashing when a cloud reaches near full capacity (about 80 % or more). This can lead to data being moved rapidly from server to server in an attempt to uphold flexibility, occupying additional resources, which might degrade the level of service. How would you describe the concerns of thrashing when considering a migration to the cloud?

    - Low
    - Medium
    - High
    - I don't know

27. When evaluating availability, how would you assess the risks related to colocation of data? As an example, colocated clients that have high resource demands or ones under attack might degenerate the services for other colocated clients if the cloud resource allocation is not sufficiently segregated.

    - Low
    - Medium
    - High
    - I don't know

28. Cloud computing is redundant by default, distributing services between several physical servers and sites. Conventional server/client infrastructures on the other hand, might be more sensitive to singular errors. To what extent would you claim this to be a negative trait of conventional computing systems?

    - Low
    - Medium
    - High
    - I don't know

29. After a cloud platform is chosen, it might not be possible to migrate all applications present in a conventional system to its cloud counterpart, as it might need to be reprogrammed or significantly changed. To what severity level would you claim this issue to be?

- Low
- Medium
- High
- I don't know

30. A cloud provider might update and upgrade the infrastructure on their own accord, possibly introducing compatibility issues for already existing client applications. How critical would you claim this issue to be for cloud clients?

    - Low
    - Medium
    - High
    - I don't know

31. Cloud providers often have intricate pricing models that charge the activity of the client. A model used frequently charges companies for all data sent out from or into the cloud. To what extent do you believe such pricing models have a negative effect on the choice of cloud providers?

    - Low effect
    - Medium effect
    - High effect
    - I don't know

32. Vendor lock-in means that a client might have a very hard time moving data and applications from one cloud platform to another based on the proprietary state of some of these. How severe would you claim the issues of vendor lock-in to be for cloud computing?

    - Lo
    - Medium
    - High
    - I don't know

33. Cloud computing lacks standards and compliance regulations which could ensure the security and interoperability of such computing environments. How strongly does this speak against implementing cloud solutions?

    - Low
    - Medium
    - High
    - I don't know

34. In conventional server/client based computer networks, the client nodes have to be sufficiently powerful in order to handle all possible tasks it might be used for, often resulting in a lot of latent resources going unused. To what extent do you think this impact negatively on conventional systems compared to cloud computing?

- Low
- Medium
- High
- I don't know

35. Using cloud computing, backup systems are often an inherent part of the system. With on-premise implemented systems on the other hand, backup systems will have to be implemented and administered completely by the client. In your opinion, to what extent does this speak against implementing on-premise solutions?

    - Low
    - Medium
    - High
    - I don't know

36. If confidential information is stored in the cloud, the cloud provider has to be trusted to keep that information safe. How would you define your reservations against trusting a third party with your confidential data?

    - None to Low reservations
    - Moderate reservations
    - Strong reservations
    - I don't know

37. The lack of clear laws and regulations for cloud environments means that there is no clear custody or ownership of data residing in the cloud. What severity level would you allocate to this issue, given that it might mean that a cloud provider can gain custody of client data residing in their cloud?

    - Low
    - Medium
    - High
    - I don't know

38. Keeping track of the status information in the cloud, for instance available resources, might be harder than from local sources depending on the quality of the self-service and administrative tools the cloud provides. To what extent do you believe this could negatively impact your decision to transition to a cloud implementation?

    - No to low impact
    - Moderate impact
    - Strong impact
    - I don't know

G Questionnaire Results

Here follows the results from the questionnaire as generated by Questback. Notice that the results say 7 responses (1 unique). This is because the questionnaire was distributed by a link, and therefore, all responses are seen as the same respondent.

Benetifs and drawbacks of cloud computing

Published from 02.06.2012 to 30.06.2012
7 responses (1 unique)

1. Single-tenant systems, which is a principle stating that a hardware device only delivers its services to a single client at the time, is in your opinion how important for limiting vulnerabilities in cloud computing?

Alternatives	Percent	Value
1 Low importance	14,3 %	1
2 Medium importance	57,1 %	4
3 High importance	14,3 %	1
-1 I don't know	14,3 %	1
Total		7

2. Systems hosted on-premise behind a clients firewall does not have the main communications channel exposed to the Internet. How effective do you perceive this type of implementation to be for lowering possible vulnerabilities?

Alternatives	Percent	Value
1 Low effect	0,0 %	0
2 Medium effect	28,6 %	2
3 High effect	57,1 %	4
-1 I don't know	14,3 %	1
Total		7

3. Clients hosting systems on-premise has direct access to hardware, and full control over its maintenance. To what extent do you believe that this impacts the possibility to correct errors and failures in a positive direction compared to off-premise hosted systems?

Alternatives	Percent	Value
1 Low	28,6 %	2
2 Medium	42,9 %	3
3 High	0,0 %	0
-1 I don't know	28,6 %	2
Total		7

4. Using a public cloud infrastructure means that one might access a hosted resource from anywhere, which might ease the process of cooperation initiatives and an external workforce. To what extent do you think public cloud implementations elevate user's ability to access their resources?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	28,6 %	2
3 High	71,4 %	5
-1 I don't know	0,0 %	0
Total		7

5. By having access to both a private and a public cloud (hybrid cloud computing), one might distribute the workload and storage based on the confidentiality of the data handled. How would you describe the impact of which implementing such a solution elevates the possibility of making resources available for authorized personnel?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	57,1 %	4
3 High	14,3 %	1
-1 I don't know	28,6 %	2
Total		7

6. How would you assess the importance of implementing single-tenant systems in order to assure that other clients cannot deplete or otherwise disrupt the resource flow?

Alternatives	Percent	Value
1 Low importance	14,3 %	1
2 Medium importance	42,9 %	3
3 High importance	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

7. If a major failure occurs at the Internet service provider or the cloud service provider, on-premise systems will still be operational and reachable by on-premise users. How important would you claim this argument to be for keeping the cloud resources available for its users?

Alternatives	Percent	Value
1 Low importance	28,6 %	2
2 Medium importance	14,3 %	1
3 High importance	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

8. In cloud computing, applications and services are not hosted on specific servers, rather utilizing a resource pool of several physical servers. Hosted content will therefore potentially not be discontinued if a hardware device fails. Which level of importance would you assess this statement to have when discussing cloud computing versus conventional server/client implementations?

Alternatives	Percent	Value
1 Low importance	0,0 %	0
2 Medium importance	14,3 %	1
3 High importance	71,4 %	5
-1 I don't know	14,3 %	1
Total		7

9. When discussing hybrid clouds, to what would you value the ability of still using one part of the cloud when the other part is incapacitated?

Alternatives	Percent	Value
1 Low value	0,0 %	0
2 Medium value	71,4 %	5
3 High value	28,6 %	2
-1 I don't know	0,0 %	0
Total		7

10. Knowing that clients using public cloud infrastructures can order and deploy virtual machine instances on the fly without considering hardware, how substantial would do you think this characteristic is for deploying and maintaining services in a computer network?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	28,6 %	2
3 High	71,4 %	5
-1 I don't know	0,0 %	0
Total		7

11. How important would you claim the ability to locally import and export data into a computer system to be?

Alternatives	Percent	Value
1 Low importance	28,6 %	2
2 Medium importance	28,6 %	2
3 High importance	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

12. How important would you assess a public cloud infrastructures ability to upscale and downscale all consumed resources of a computer system to be, when comparing them to the limited abilities of other cloud infrastructures and conventional server/client systems?

Alternatives	Percent	Value
1 Low importance	0,0 %	0
2 Medium importance	28,6 %	2
3 High importance	57,1 %	4
-1 I don't know	14,3 %	1
Total		7

13. Choosing clouds that have no reserved hardware might result in lower costs from avoiding the purchase of hardware, and the personnel to operate it. It also means that the client can focus the additional resources on the main tasks of the company. How important would you claim this to be for a company?

Alternatives	Percent	Value
1 Low importance	0,0 %	0
2 Medium importance	42,9 %	3
3 Highly important	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

14. Virtualization allows for entire networks of computers to exist strictly in software, a technology that cloud computing thoroughly exploits. What degree of importance would you claim the possibility of creating, saving, moving or deleting computer instances quickly and easily to be when planning a transition to the cloud?

Alternatives	Percent	Value
1 Low importance	0,0 %	0
2 Medium importance	14,3 %	1
3 High importance	85,7 %	6
-1 I don't know	0,0 %	0
Total		7

15. Hybrid cloud implementations might borrow resources from the public cloud in order to temporarily enhance the performance of the on-premise cloud solution (cloud boosting). How important do you recon this statement is in the case for hybrid cloud solutions?

Alternatives	Percent	Value
1 Low importance	0,0 %	0
2 Medium importance	42,9 %	3
3 High importance	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

16. Most cloud service providers have intricate backup systems built into their infrastructure, and might make that technology available to the clients as built in services. To what extent do you believe this impact the user's ability to secure their data?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	28,6 %	2
3 High	71,4 %	5
-1 I don't know	0,0 %	0
Total		7

17. Several clients or parts of a clients company might utilize the same public log-in gateways, regardless of authorization level, making the log-in gateway a prime target for attacks. To what extent do you believe this might impact the security of the cloud solution?

Alternatives	Percent	Value
1 Low impact	14,3 %	1
2 Medium impact	28,6 %	2
3 High impact	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

18. By using cloud services in a given country, the authorities of that nation might demand access to the data, for instance during a forensic investigation. What severity level would you allot to this fact, given that this might result in loss of, or limited access to information stored or services hosted in the cloud?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	28,6 %	2
3 High	71,4 %	5
-1 I don't know	0,0 %	0
Total		7

19. On-premise hosted private clouds will have direct access to hardware in the same way as conventional server/client implementations, something that might open additional attack vectors, for instance for malicious insiders. How severely do you believe this impact the security of the implementation?

Alternatives	Percent	Value
1 Low	28,6 %	2
2 Medium	42,9 %	3
3 High	28,6 %	2
-1 I don't know	0,0 %	0
Total		7

20. The failure of hardware for systems hosted on-premise might be harder to mitigate for smaller companies in comparison to dedicated cloud providers, which will always have personnel on-call to handle failures. To what extent do you believe this might lever the decision towards avoiding on-premise hosted implementations?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	71,4 %	5
3 High	28,6 %	2
-1 I don't know	0,0 %	0
Total		7

21. Colocation of several clients on one physical device (so called multi-tenancy) is common in cloud computing. This might lead to leakage of data from one client to another through targeted attacks or faulty cloud infrastructure. To which criticality level would you assess multi-tenancy from a security standpoint?

Alternatives	Percent	Value
1 Low	14,3 %	1
2 Medium	28,6 %	2
3 High	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

22. By distributing data storage and processing over several servers or data centres expands the attack surface of the clients cloud implementation, possibly adding to the number of vulnerabilities. How strongly do you believe the distribution of data processing counts against implementing a cloud solution?

Alternatives	Percent	Value
1 Low	14,3 %	1
2 Medium	85,7 %	6
3 High	0,0 %	0
-1 I don't know	0,0 %	0
Total		7

23. Hybrid clouds share data between private and public cloud solutions, something that might expose the complete implementation to additional attack vectors. How would you describe the level of reservations you might have against implementing a hybrid cloud based on this information?

Alternatives	Percent	Value
1 Low	42,9 %	3
2 Medium	42,9 %	3
3 High	0,0 %	0
-1 I don't know	14,3 %	1
Total		7

24. The cloud service provider becomes a single point of failure for all cloud services a client uses. How severe would you assess this issue to be?

Alternatives	Percent	Value
1 Low severity	14,3 %	1
2 Moderate severity	42,9 %	3
3 High severity	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

25. Using a hybrid cloud solution, the public parts of the cloud would still be susceptible to errors on the cloud service provider side. What severity level would you assign to this problem area for of hybrid deployment models?

Alternatives	Percent	Value
1 Low	42,9 %	3
2 Medium	42,9 %	3
3 High	14,3 %	1
-1 I don't know	0,0 %	0
Total		7

26. Cloud implementations might experience thrashing when a cloud reaches near full capacity (about 80% or more). This can lead to data being moved rapidly from server to server in an attempt to uphold flexibility, occupying additional resources, which might degrade the level of service. How would you describe the concerns of thrashing when considering a migration to the cloud?

Alternatives	Percent	Value
1 Low	71,4 %	5
2 Medium	14,3 %	1
3 High	14,3 %	1
-1 I don't know	0,0 %	0
Total		7

27. When evaluating availability, how would you assess the risks related to colocation of data? As an example, colocated clients that has high resource demands or ones under attack might degenerate the services for other colocated clients if the cloud resource allocation is not sufficiently segregated.

Alternatives	Percent	Value
1 Low	14,3 %	1
2 Medium	57,1 %	4
3 High	0,0 %	0
-1 I don't know	28,6 %	2
Total		7

28. Cloud computing is redundant by default, distributing services between several physical servers and sites. Conventional server/client infrastructures on the other hand, might be more sensitive to singular errors. To what extent would you claim this to be a negative trait of conventional computing systems?

Alternatives	Percent	Value
1 Low	28,6 %	2
2 Medium	28,6 %	2
3 High	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

29. After a cloud platform is chosen, it might not be possible to migrate all applications present in a conventional system to its cloud counterpart, as it might need to be reprogrammed or significantly changed. To what severity level would you claim this issue to be?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	57,1 %	4
3 High	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

30. A cloud provider might update and upgrade the infrastructure on their own accord, possibly introducing compatibility issues for already existing client applications. How critical would you claim this issue to be for cloud clients?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	42,9 %	3
3 High	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

31. Cloud providers often have intricate pricing models that charge the activity of the client. A model used frequently charges companies for all data sent out from or into the cloud. To what extent do you believe such pricing models have a negative effect on the choice of cloud providers?

Alternatives	Percent	Value
1 Low effect	28,6 %	2
2 Moderate effect	28,6 %	2
3 High effect	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

32. Vendor lock-in means that a client might have a very hard time moving data and applications from one cloud platform to another based on the proprietary state of some of these. How severe would you claim the issues of vendor lock-in to be for cloud computing?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	14,3 %	1
3 High	85,7 %	6
-1 I don't know	0,0 %	0
Total		7

33. Cloud computing lacks standards and compliance regulations which could ensure the security and interoperability of such computing environments. How strongly does this speak against implementing cloud solutions?

Alternatives	Percent	Value
1 Low	14,3 %	1
2 Medium	28,6 %	2
3 High	57,1 %	4
-1 I don't know	0,0 %	0
Total		7

34. In conventional server/client based computer networks, the client nodes have to be sufficiently powerful in order to handle all possible tasks it might be used for, often resulting in a lot of latent resources going unused. To what extent do you think this impact negatively on conventional systems compared to cloud computing?

Alternatives	Percent	Value
1 Low	14,3 %	1
2 Medium	42,9 %	3
3 High	42,9 %	3
-1 I don't know	0,0 %	0
Total		7

35. Using cloud computing, backup systems are often an inherent part of the system. With on-premise implemented systems on the other hand, backup systems will have to be implemented and administered completely by the client. In your opinion, to what extent does this speak against implementing on-premise solutions?

Alternatives	Percent	Value
1 Low	14,3 %	1
2 Medium	42,9 %	3
3 High	28,6 %	2
-1 I don't know	14,3 %	1
Total		7

36. If confidential information is stored in the cloud, the cloud provider has to be trusted to keep that information safe. How would you define your reservations against trusting a third party with your confidential data?

Alternatives	Percent	Value
1 None to Low reservations	0,0 %	0
2 Moderate reservations	14,3 %	1
3 Strong reservations	85,7 %	6
-1 I don't know	0,0 %	0
Total		7

37. The lack of clear laws and regulations for cloud environments means that there are no clear custody or ownership of data residing in the cloud. What severity level would you allocate to this issue, given that it might mean that a cloud provider can gain custody of client data residing in their cloud?

Alternatives	Percent	Value
1 Low	0,0 %	0
2 Medium	14,3 %	1
3 High	85,7 %	6
-1 I don't know	0,0 %	0
Total		7

38. Keeping track of the status information in the cloud, for instance available resources, might be harder than from local sources depending on the quality of the self-service and administrative tools the cloud provides. To what extent do you believe this could negatively impact your decision to transition to a cloud implementation?

Alternatives	Percent	Value
1 No to low impact	14,3 %	1
2 Moderate impact	57,1 %	4
3 Strong impact	28,6 %	2
-1 I don't know	0,0 %	0
Total		7

H Requirements Document

This documentation of requirements is created in order to equip the reader with a list of tasks that can be followed in order to plan and document a cloud implementation for governmental and law enforcement use. By completing the steps below you will assess all the major areas that should be documented and fully understood before undertaking such a task, and create a foundation from which you can assess possible cloud environments in a clear and structured way.

Keep in mind that the document is general in nature, and should be used to complement existing regulations, rather than taking their place. Existing planning and implementation frameworks applied by the client should be consulted in parallel with this document, making sure that the tasks are integrated into the framework at the appropriate stages of the project.

The document does not include the bureaucratic process of getting a project approved by a government, linking of the project to political agendas, establishing the funding for the project, etc.

H.1 Requirements guide

1. Prerequisites/Constraints
   This section present the system that is to be implemented or migrated to the cloud. It is defined as preliminary studies and documentation that will have to be conducted before a project can be assessed, verified and approved. The following subpoints specify this section.

   (a) Initial description of project
   Several sections are included in this part of the project, basically summing up the background, motivation, high-level goals and backing of the project. The task should culminate in a preliminary project description that can be presented to possible stakeholders etc.(to sell in the project).

      i. Problem description
      General description of the problem at hand, for instance describing the main reasons for implementing a given system in the cloud or migrating an existing system. Should emphasize the motivation for the project.

      ii. Scope
      A general, high-level description of the system and its general functionality. If the project revolves around the migration of existing solutions to the cloud, this section should link to already existing documentation. If only sections of the system is to be moved, the scope document should emphasize on those sections, what functionality and data is moved to the cloud and, if relevant, how the cloud implementation will communicate with the internal system on the client's premises.

iii. Goals
Should consist of a bullet-point list of high-level goals, including, for instance, the main functionality of the system and the expected benefits of placing such functionality into the cloud. References to preceding projects of the same kind (if such exist), and the status of these can help to provide some idea of the possible benefits. Either way, this section will serve as a yardstick for achievements and benchmarking for system tests and assessment at project completion.

iv. Analysis of expected benefits/drawbacks/gains
A study should be conducted to support the goals of the project, thereby ensuring both the stakeholders and the architects of the potential and projected gains of moving to the cloud compared to other possible solutions.

v. Resources
Expected use of economical, technical, and human resources throughout the project planning, implementation, initialization and life-cycle (maintenance, training).

vi. Authority
This section should describe the governmental entity which owns the project, and by which the project is sanctioned.

(b) Initial legal study
A preliminary study of the legality of the project should be conducted and presented as part of the prerequisite tasks for the project.

i. Execution of the study
As work has not yet been conducted on creating a detailed list of functional and non-functional requirements, the data to be used in the system and how that data shall be processed, the study must be based on the initial project description. Section five in this document can be helpful for finding law texts that are relevant for this study.
This task might be hard, given the vagueness of the document at hand. However, one should be able to verify if a project is legally feasible at an early stage, to avoid any legal implications as the project progresses.
The result of the task should be a document predicting the legality of the project against relevant laws and legislations in the area of operation (EU, US, national specific legislations), and it should specifically emphasize possible problem areas and scenarios.

ii. Assessment and verification of legal study
It will now be up to the project group to assess the stated problems in the documentation from 1.(b)i., and to decide if the project description should be changed to avoid these problem areas. It must also be stated if the problems are manageable, and how. Documentation stating the choices of the project group should be presented, and, if necessary, the documentation should also state who is responsible for mitigating the problems that are deemed to be manageable. The task is completed when the project group decides that the project coincides with the law, or accepts potential problem areas.

(c) Initial compliance study
Government agencies are often bound by compliance standards that assist in abiding the laws and ensure the quality of the work conducted. A preliminary study should be conducted to ensure that the project will comply with these compliance standards, and to assess possible standards if such are not already stated.

i. Execution of the study
As in step 2.(b)i., the compliance study has to be based on the project description from Step 1.(a). Section six of this specification document draws a rough outline of what should be included in a compliance study, and can be helpful during this step. Special attention should be given to possible problem areas that could render the solution non-compliant with the relevant standards. The result of the task should be a document assessing the existing or possible compliance standards that are, or will be relevant during the planning and implementation of the solution.

ii. Assessment and verification of compliance study
Based on the study in step 1.(c)i., the project group has to assess the risks of the problem areas stated (if any), and decide if the project description should be altered to better fit the compliance demands, or if the risks can be managed. The task should produce documentation stating the choice of the project group, and, if relevant, who accepts the risks that should be managed. The task is completed when the project group decides that the project is in compliance with relevant standards, or accepts potential risks.

2. Requirements, functional and non-functional
In this section the project must explore and lock down the requirements of the solution to be implemented. Starting with sub-point 2.(a), the high-level requirements are mapped and agreed upon. Following this, sub-point 2.(b) describes the requirements at a lower level of granularity, breaking the requirements into use-cases and assessing these. Finally, in sub-point (c), the choice of a cloud environment can be discussed on a sound foundation of defined assets.

(a) High-level requirements
This section will constitute the foundation on which a technical solution can be designed. It should focus on what should be done rather than how to do it.

i. Identify users and usage patterns
Mapping the user groups needed in the system and their allowed activities is a good way to map initial requirements for the system in the form of functionality and security needs. Can be text based, for instance: ”I as an ACTOR should be able to TASK because REASON”. This example is called a user story [93], and is frequently used in agile project frameworks. This initial survey should remain focused on the major tasks for each user group, and rather explain what should be included in the implementation rather than how it should be implemented

ii. Identify major functional requirements based on usage patterns
By following each usage pattern and assessing the functionality that has to be present for it to work as intended, the requirements of the system are further specified. The result should be a list of functional requirements that roughly covers all aspects of the system.

iii. Identify major non-functional requirements based on usage patterns
Non-functional requirements must also be kept in mind throughout the planning and implementation process. Non-functional requirements can be defined as restraints or qualities. Qualities are characteristics of the system that are coveted by the projects stakeholders, for instance performance. Constraints are set limitations of the system that need to be complied with, and thus cannot be omitted. An example could be that the system cannot use a mouse for I/O.

iv. Verification of high-level requirements
This task should be performed during one, or a series of meetings, in which the high-level requirements are found, discussed and agreed upon by the stakeholders, its users and the architects. If an agreement is reached, work can progress with detailed requirements.

(b) Detailed requirements
In this section the requirements are further explored to create a detailed list. All the assets of the system should be explored and mapped into either data or functionality. This is important because one will always have to assess each asset for its sensitivity and security demands. This work must be based on the high-level requirements created in 1.(b). In order to identify all assets, the following steps can be followed.

i. Create use-case diagrams to map functionality
By creating a use-case diagram, the high-level functionality of the system is mapped. This makes it easier to assess which functionality assets the system will be composed of. Both an overall graphical representation of the use-cases, and a table for each specific use-case should be created to thoroughly visualize and describe the functionality. An example of a graphical representation is given in Appendix one of this guide, while an example of a use-case table is given in the second Appendix.

ii. Map data flow
By creating a data flow diagram, it is easier to see how functionality assets cooperate, in turn making it easier to map what data needs to be created and used. This is also important for assessing security features to safeguard the data and functionality. Flow charts and sequence diagrams (modelling work flow charts) can be used to achieve this.

iii. Create a list of assets based on the use-case and data flow diagrams
By listing all the assets of the system, one will create a great foundation for further assessment of the needs (technology and security wise) of the system to be implemented. The list should include the asset's name, its type (data/functionality), with which use-cases the asset is associated, and its expected importance for the system (critical role, supporting role,

usability/cosmetic)

iv. Discuss the sensitivity of the assets
Each asset should now be thoroughly assessed concerning the sensitivity of its actions/content. This will both clarify if the asset is liable for cloud use in its current state, and be useful when discussing possible cloud deployment models and cloud service providers (CSPs). This section can be conducted as a full-scale risk assessment study, in which case a solid risk assessment framework should be used.
At minimum, the project should provide answers to the following list of questions, which is defined in CSA v3.0 [40]:

- How would we be harmed if the asset became widely published and distributed?
- How would we be harmed if an employee of our CSP accessed the asset?
- How would we be harmed if the process or function was manipulated by an outsider?
- How would we be harmed if the process or function failed to provide expected results?
- How would we be harmed if the information/data was unexpectedly changed?
- How would we be harmed if the asset was unavailable for a period of time?

v. Verification of asset list
One or a series of meetings should be held to verify the documented assets, both data and functions, to lock down the requirements. By doing so, an agreement is reached between the stakeholders and architects of exactly what the solution should be able to do, and what data it should manage and contain. The group should also have a clear indication of the sensitivity of each asset

3. Requirements for possible cloud service models
Under this sub-point a discussion regarding the choice of a service model should be conducted. A comparison of possible choices should be done, and the content of sub-points 2.(a) and 2.(b) should be used to clarify the choice.

   If a cloud service model has not been chosen in advance, discuss all relevant cloud models (SaaS, PaaS, IaaS) and how they meet the requirements.

   If a cloud service model is already selected, and this section cannot support the choice, the project group must re-assess the assets to be placed in the cloud, or the cloud service model to be implemented.

   (a) Defining the cloud service model alternatives
   A definition of the service models that are seen as alternatives for the solution should be created and agreed upon by the project group and the stakeholders. This ensures that the understanding regarding the models is distributed

throughout the group. Existing definitions should be consulted when performing this task. The result of this task should be a document defining the choice of alternatives made by the project group.

(b) Discussing each cloud service alternative against the verified assets
By evaluating the alternatives against each asset, one will be able to assess which cloud service model best fits the requirements of the implementation. This might be significantly time consuming if the alternatives provided do not cater to all the requirements. The project group must then either return to step 2.(c)i. to re-assess the alternatives, or to step 2.(b)iv. to re-evaluate the assets and how these can be altered to fit the alternatives. If the assets still cannot be mapped to possible alternatives, the project group has to decide whether step 2 should be re-evaluated in its entirety.

The result of this task will be a document stating where the model alternatives meet, and do not meet the stated requirements.

(c) Choosing a cloud service model
After evaluating the different alternatives in Section 2.(c)ii., the project group should have a sound foundation on which a choice of service model alternative can be conducted.

The task should result in a document explaining the reason for the choice, and stating why the other alternatives have been eliminated. It should also consist of an agreement between the stakeholders and the project group, stating that the choice is accepted and understood by all parties.

4. Requirements for possible cloud deployment models
Under this sub-point a discussion regarding the choice of a deployment model should be conducted. A comparison of possible choices should be done, and the content of sub-points 2.(a) and 2.(b) should be used to clarify the choice.

If a cloud deployment model has not been chosen in advance, discuss all relevant cloud models (public, private, hybrid, community, hosted-private) and how they meet the requirements.

If a cloud deployment model is already selected, and this section cannot support the choice, the project group must re-assess the assets to be placed in the cloud, or the cloud deployment model to be implemented.

(a) Defining the cloud deployment model alternatives
By starting with defining each possible cloud deployment model, one ensures that both the stakeholders and the architects have the same view of the different models. Taking the example of a private cloud, it is important to define if the cloud should be hosted on-premises at the client, be a hosted-private implementation at a CSP, or a combination of the two. The combined model would then be, or not be defined as a hybrid model, etc.

The result of the task should be a document containing definitions of the cloud model alternatives agreed upon by the entire project group.

(b) Discussing each cloud deployment alternative against the verified assets
By evaluating the alternatives against each asset, one will be able to assess

which cloud model best fits the requirements of the implementation. This can be significantly time consuming if the alternatives provided do not cater to all the requirements. The project group must then either return to step 2.(c)i. to re-assess the alternatives, or to step 2.(b)iv. to re-evaluate the assets and how these can be altered to fit the deployment alternatives. If the assets still cannot be mapped to possible alternatives, the project group has to decide whether step 2 should be re-evaluated in its entirety.

The result of this task will be a document stating where the model alternatives meet, and do not meet the stated requirements.

(c) Choosing a cloud deployment model
After evaluating the different alternatives in Section 2.(c)ii., the project group should have a sound foundation on which a choice of deployment model alternative can be conducted.

The task should result in a document explaining the reason for the choice, and stating why the other alternatives have been eliminated. It should also consist of an agreement between the stakeholders and the project group, stating that the choice is accepted and understood by all parties.

5. Requirements for possible cloud service providers
The choice of a CSP has to be based on the choice of deployment and service models, as a CSP often offers only one combination of the two. It is therefore important that steps three and four are completed before the choice of a CSP is attempted. It the CSP is already chosen in advance, one must still make sure that it can support the service and deployment models required by the planned solution.

(a) Defining the cloud service provider alternatives
There will most likely exist several possible CSPs for each of the deployment models selected as alternatives. These should be studied, and a short introduction should be created for each of them.

(b) Discuss each alternative against the verified assets
The alternatives should all be explored to verify if they meet the demands of the requirement specification. This can be done by studying the verified assets and assess if the CSP can comply with their features. Also, a study of the security and usability features the CSP offers should be performed. One can consult Section seven of this requirement document to get an idea of what is important to cover in a security evaluation, for instance verifying that the CSP in question supports the controls necessary to secure data and users according to your demands, and a thorough SLA (the SLA could for instance be customizable to attend to your specific needs).

(c) Choosing a cloud service provider
Findings should be presented, and a decision should be reached.

6. Legal requirements
This section should cover all legal obligations of the owner of the system and the CSP used. Special attention should be given to the data the system will be handling, in addition to how and where that data is processed.

The task requires that a detailed specification of the project is created and agreed upon. However, it can still be helpful during preliminary studies like the one conducted in step 1,b.

Note that separate laws can apply for where the data originates, where it is processed, and where the CSP is situated. Given the cloud's distributed nature, the position of the client, data centres, storage facilities, transport equipment and CSP might be different, and under multiple jurisdictions. It is important to map these in order to assess the legality of any cloud project.

(a) Identifying relevant law
As applicable law depends on the jurisdictions that the cloud implementations are fully or partially placed in, it is not feasible to mention all relevant legislations. Instead, the major legislations of the European Union and the United States of America have been noted to give some examples. These, as well as other relevant national or international legislations, should be considered when discussing the legality of the planned cloud implementation:

- European Union directives
  - 95/46/EC
  - 2002/58/EC
  - 2006/24/EC
  - 2008/977/JHA
- US legislations
  - Gramm-Leach-Bliley Act
  - Health Insurance Portability and Accountability Act
  - Sarbanes-Oxley Act
  - Clinger-Cohen Act of 1996
  - OMB Circular No. A-130(Appendix 3)
  - The Privacy Act of 1974
  - The E-Government Act of 2002
  - FISMA 2002
  - NARA regulations(Federal Records Act (44 U.S.C Chapters 22,29,31,33)(Title 36 Code of Federal Regulations, Chapter 12, Subchapter B)

(b) Assess law to uncover barriers early
As the cloud scope and components are mapped, review the following to find possible legal predicaments. NOTE: Keep in mind that Section 3.(a) will be a natural section to assess in parallel with this study.

i. For on-premises equipment and services
This part of the system will follow the same regulations as any other on-premise hosted service, with special attention to what, if anything, is exposed to the outside of the client's perimeter. If a CSP is assigned to host the cloud inside of the client's perimeter, special attention should be given to ensure that the same legal obligations apply for this actor. Service level

agreements (SLAs) should be created to legally bind the parties with obligations used to handle disputes. Legal obligations should also be passed on to users, both internal and external, that utilize the system.
NOTE: If a strictly public cloud deployment model is chosen, part 2.(b)i. can be omitted.

ii. For off-premises (CSP) equipment and services
This part should cover both the client and provider's legal obligations to content outsourced to the CSP. This should include both the laws and regulations of relevant jurisdictions, as well as SLAs. As in 2.(a)i., legal obligations should also be passed on to users, both internal and external, that utilize the system.
NOTE: If a strictly private cloud deployment model is chosen, part 2.(a)ii. can be omitted.

iii. For data assets and data flow
Working with for instance personally identifiable information (PII) or protected health information (PHI) will limit the way the information can be used and transferred.

7. Compliance requirements
Laws and legislations are often followed by new or improved compliance standards which purpose often is to ease the adherence of a given, or set of laws. The standards can also serve as an additional security measure, often going beyond the minimum requirements of laws.

(a) Identifying relevant compliance standards
Compliance standards that may be relevant (amongst others) are:

- Daubert Criteria
- Federal Risk and Authorization Management Program (FedRAMP)
- ISO27001/27002
- Common Criteria
- Payment Card Industry Data Security Standard (PCI DSS)

(b) Assess compliance demands to uncover barriers early
Compliance demands might be obligatory for parts of the system based on the type of data handled, where it is processed and who owns it. The client could also have prior obligations to follow certain compliance demands, and the new cloud implementation could be expected to comply with these. Make sure that all such scenarios are thoroughly explored.
NOTE: Keep in mind that Section 2.(a) may be a natural section to assess in parallel with this study.

i. For on-premises equipment and services
This section will be influenced by the client's compliance demands. It will also be influenced by the CSPs compliance demands if the CSP is hosting parts of the cloud on-premises at the client side.

ii. For off-premises (CSP) equipment and services
The equipment referred to in this section will have to comply with the

CSP's compliance demands, but can also have to comply with the client's compliance demands if such demands also adhere to outsourced equipment.

iii. For data assets and data flow
These compliance demands will apply more to the applications deployed in the environment than the cloud environment itself. Companies that frequently develop software will probably have introduced compliance demands for safe programming, and procedures for logging and collaboration in this process. These will have to be followed. For some well established principles on how to strengthen software (and hardware) solutions, one can consult the Saltzer and Schroeder principles [120], or other established best practice documentation.

(c) Create a plan for auditing compliance
Planning the audit compliance of the implementation depends on the implementation itself, and the compliance regulations applied. Please consult the documentation for each specific compliance regulation or standard to assess how this must be done in order to rech certification and uphold a continuous audit procedure. Remember that separate compliance regulations may apply for the two sections of your cloud implementation:

i. For systems the client reserves the responsibility for

ii. For systems the CSP reserves the responsibility for

8. Security requirements

In order to uphold a high degree of security in the implemented system, the security requirements should be created and implemented at early stages rather than being an afterthought or add-on to a finished product. Several sections of the system should be thoroughly assessed to map such requirements:

(a) Securing data

i. Discuss each data asset against the data security life cycle
All parts of the life cycle of data have to be secured in order to ensure its security from inception, through use and storage, and at last deletion and disposal.

ii. Map each data asset to security controls and methods

(b) Securing equipment

i. Create a plan for securing hardware under the responsibility of the client

ii. Create a plan for assessing the security of off-premise equipment (SLAs and audits)

(c) Securing users

i. Define the solution for authentication in the given cloud environment

ii. Define the entitlement rules for different user groups (link to use-cases and data flow)

9. Risk assessment In order to identify possible risks during the project period, and after an implementation has been put into production, risk assessments must be performed.

   (a) When to perform risk assessment
   Risk assessments should be conducted both in the preliminary stages of the project, and during the design stages for the components of the proposed system. Risk assessments concerning the team members and general time and resource problems will not be covered in this guide. The focus will rather be on the assessment of system components. One natural step in which a risk assessment could be particularly useful is during Section 2.(b)iv. of this guide. Here the assets of the implementation have been mapped, and is in the process of being assessed for sensitivity. The degree of sensitivity can be useful when assessing the possible risks.

   (b) How to perform the risk assessment
   How the risk assessment is conducted is less important, as long as it is thorough and applies to the best practices of the procedure. To be safe, standardized and proven methods are advisable. Most project management frameworks have risk assessments built into it, or references to possible alternatives. One alternative is ISO/IEC 27005:2011 information security risk management process, as presented in Figure 20, which can be found in Section H.2.4 of the appendices to this guide.

H.2 Requirement Appendices

H.2.1 Example use-case and use-case diagram

Figure 18 shows two actors and nine use-cases. Some of the use-cases are grouped under other use-cases to visualize the similar nature of these. This signifies that the use-cases might also be grouped during development and implementation.

H.2.2 Example use-case table

Use-case 4	Log in
Actor	Administrator and User
Goal	Gaining access to the system with user-group specific rights
Normal event float	1. Open system interface 2. Enter user name 3. Enter password 4. Submit information 5. Gain access
Variations	Access is denied if credentials are not correct

Table 35: Example use-case table

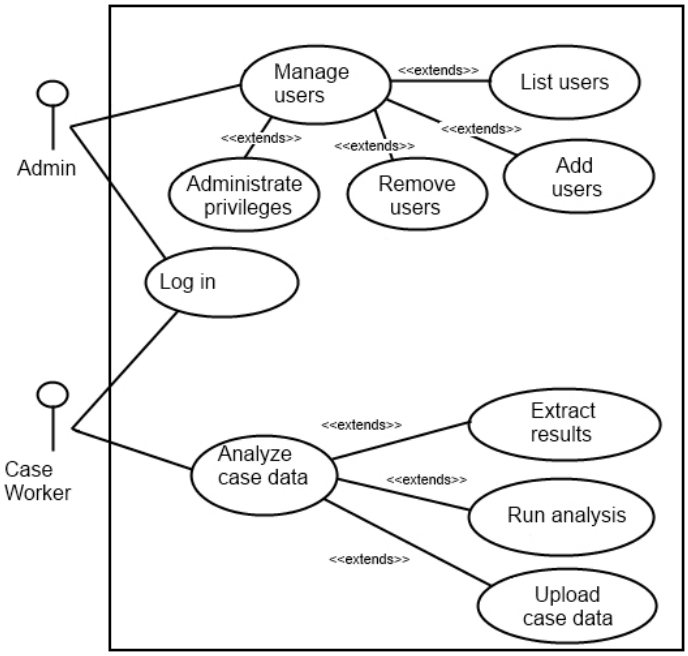

Figure 18: Guide: UML Use-Case diagram for the forensic analysis system

In Table 35 we have a numbered use-case with a name, a list of actors or user groups who should be able to use, a specific goal that one should be able to link to high-level requirements, a description of the normal stages during runtime and, if relevant, possible variations to the event float. Creating a table for a use-case can be helpful to create a more detailed picture of functionality, and the data needed in the process.

H.2.3 Example Data Flow Model

Figure 19 shows an example of a flow chart explaining a login process.

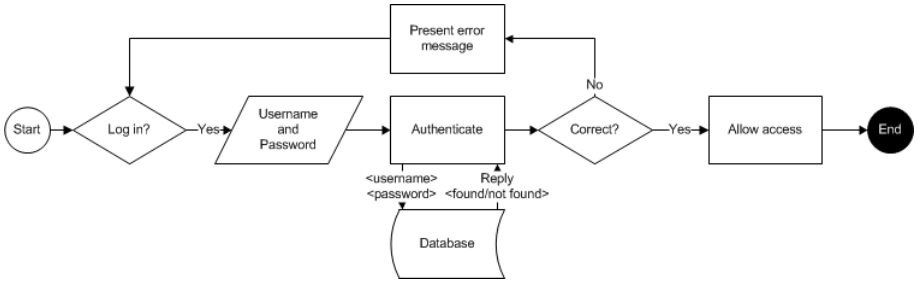

Figure 19: Guide: Flow chart visualizing log in functionality

H.2.4 Example Risk Assessment Method

The figure illustrates how risk assessment can be modelled as an iterative process. The model is initiated by establishing the context of the risk assessment, and then covers the identification, assessment and treatment of risks. Included in the process is also a knowledge base (Risk communication), for storing knowledge about explored risks for later use.

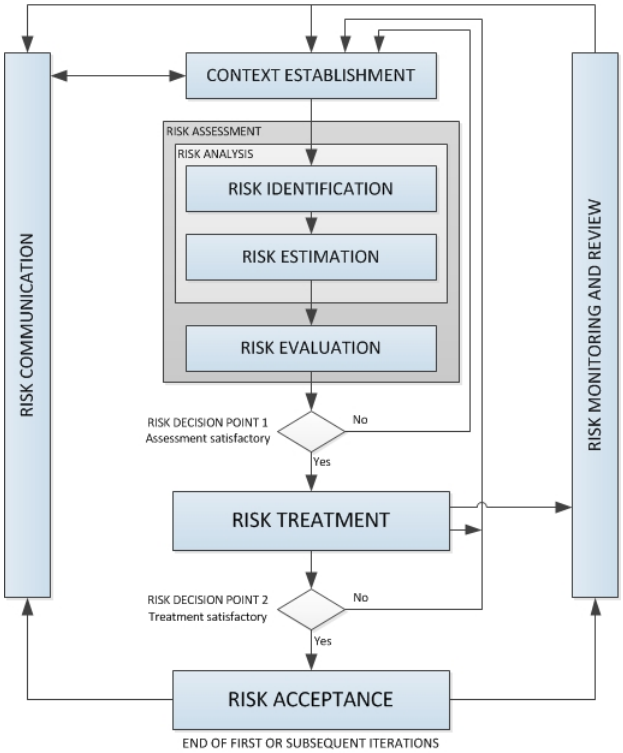

Figure 20: Guide: ISO 27005 Risk management flowchart. Inspired by [129]

I Microsoft Operations Framework

As the requirements document H is intended to be framework independent, one should be able to use its guidelines in parallel with any framework. To showcase the use of the requirements document, a fictional project for a CC implementation is conducted using the well known project management methodology: Microsoft Operations Framework(MOF).

MOF is and iterative methodology consisting of best practices, principles and activities that serve as guidelines for IT-related projects [130]. The methodology is divided into three main phases: ”Plan”, ”Deliver” and ”Operate”, with a ”Manage” phase continuously overseeing the other three. The phases themselves contain several Service Management Functions (SMFs), which in turn contain the processes and activities that support the phase in which the SMF is placed. The framework is intended to support projects ranging from new conceptions and operation of existing implementations, to the retirement of systems currently in operation.

Figure 21 shows the IT service life cycle used by MOF, the three phases, and the main SMFs within each phase.

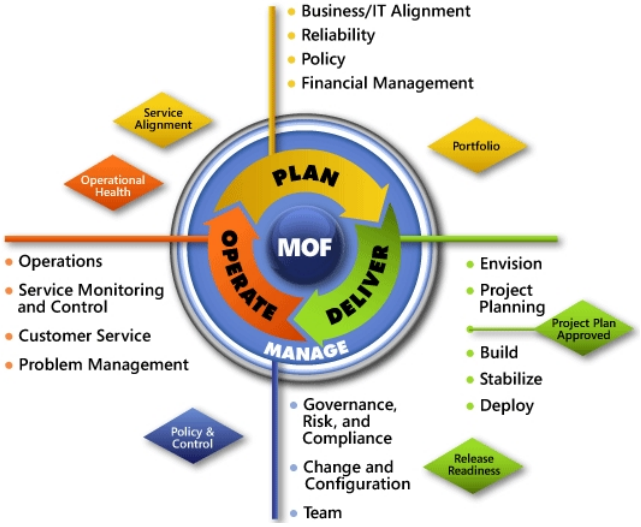

Figure 21: Overview of the MOF phases and SMFs. Inspired by [130]

I.1 Plan Phase

The primary goals of the plan phase are to guide the planning of IT service strategy optimization. By using the phase’s SMFs, the processes, resources and activities required to align an IT service to stakeholder’s requirements are conducted. This creates a foundation on which new projects or improvements to existing IT systems can be planned and conducted. The SMFs of this phase are:

- **Business/IT Alignment**

The SMF is used to map a business's IT strategy and IT services and demands currently in place. It is then used to relate the IT services to the requirements in order to uncover if any of the requirements are not met. The SMF culminates in an IT service portfolio, which includes the present and planned IT services for the business.

- **Reliability**
  This SMF is created to ensure that the IT services in a business are reliable and available for its users. By reviewing the business's requirements, building plans to meet their expectations and continuously monitoring the status of IT services, the reliability SMF aims to keep services running, and to keep recovery plans up to date in case of errors.

- **Policy**
  The policy SMF is used to evaluate the needs for policies and the creation, maintenance and enforcement of these. Its primary goals are to ensure up-to-date guidelines that comply with the desired actions and behaviours of the organization, and it also helps mitigating risks.

- **Financial Management**
  This SMF concerns the correct evaluation of the business value of IT services. Planning budgets, managing finances and accounting are some of the tasks performed in this SMF. The work done here is instrumental for the acceptance of projects, as the expected value for each project has to be acceptable before these are approved and throughout a project's lifetime.

I.2 Deliver Phase

In the deliver phase, IT services are planned, designed and implemented. Both new projects and changes or modifications to services already in production use the deliver phase. Some of the goals of the phase are to ensure that every project creates functional specifications and solution design, and to make sure that new services hold a high standard and are stable when deployed.

The phase builds on the foundation created in the planning phase, and has the following SMFs:

- **Envision**
  In the Envision SMF, business needs and requirements are analyzed to uncover new needs or changes to the IT infrastructure. The SMF concerns the creation of high-level goals and restraints, as well as risk assessment and the first draft of a project timeline. Main goals of the Envision SMF are to ensure that all project actors and stakeholders have the same understanding regarding the project to be conducted, and that risks are documented and handled.

- **Project Planning**
  During Project Planning the high-level requirements created in the envision SMF are further explored to create the complete functional specification and design. Together with work plans and a list of deliverables, this constitutes the master plan for a project. The primary goals of this SMF are to create a realistic project schedule, and to get the detailed project plan approved by customers and stakeholders.

Defining the roles and responsibilities of team members is also included in this SMF.

- **Build**
  The Build SMF covers the actual development processes of the project, including coding, developer documentation and documentation of infrastructures that support the development. The main goals of the Build SMF are to develop the project components to the specifications created in the prior SMFs, ensuring that they meet the expectations and that the system is free from bugs. Another goal is to ensure that the project is progressing on schedule.

- **Stabilize**
  The Stabilize SMF has three main processes; stabilize a release candidate, create a project pilot and review the release management review. This means that most of the work in this SMF is spent on testing. The goals of the Stabilize SMF include the testing of a complete solution, the deployment of a release candidate to a pilot group and handling of pilot feedback.

- **Deploy**
  In this SMF a version of the stabilized implementation is moved from a test environment to a production environment, where stabilization will continue. The Deploy goals are to implement a stable deployment of the project components in production and get approval from the customer and stakeholders stating that the components are implemented according to the requirements. The final part of this SMF and the Deliver phase is to hand over the implementation to operations and support teams.

I.3 Operate Phase

The operate phase focus on how to maintain IT services and solutions that are already in place in the environment. The phase specifically aims to ensure that services are managed properly, that the status of the services is available and monitored, and that faulty services are restored to a healthy state. The teams created and operating in this phase are usually dedicated to a specific task, usually without any projected end of project time in mind. The SMFs of this phase are:

- **Operations**
  This SMF concerns the daily operations of IT services running in a production environment, and the tasks that are required for their maintenance. The SMF does not propose specific tasks, but gives examples on how to best create and use these based on the services in question. The goals of the operate SMF are to identify the tasks needed to efficiently operate a given service, enable proactive tasks to reduce reactive work, minimize errors and streamline recurring tasks.

- **Service, Monitoring and Control**
  The service, monitoring and control SMF concerns live monitoring of IT services, and how to use the information that such activities produce. Much of this information will be useful to other SMFs in the MOF methodology. The major goals of this SMF are to observe the health of services, take action to remediate errors, understand the infrastructure foundation of the services and assist in the optimization

of these. Doing this could lower the probability of further errors and SLA or OLA breaches, as well as improve the availability of services.

- **Customer Service**
  The customer service SMF covers tasks that are relevant for IT support and service desk personnel. Its goals are to create and provide a positive experience and effective management of customer assistance and complaints. Integration of self-service portals for initial automated assistance, and effectively collecting, classifying and handling user requests are main activities of customer service.

- **Problem Management**
  The SMF provides guidance for handling advanced problems concerning the IT solutions. This is done by identifying the problem areas, studying the nature of the issues and creating solutions or workarounds. The goals of problem management are both to reduce the probability and occurrences of problems in specific systems, but also to build knowledge that will assist during the creation and configuration of new services.

I.4 Management Layer

The Management layer is not defined as a phase, as it rather provides a foundation for the three main phases of MOF. The layer is not time-specific, and applies continuously to all the activities of all the phases. The goals of the Management layer is to assist in the coordination of life cycle activities by applying decision-making processes, performing risk management and defining the accountabilities of the roles used during the implementation of tasks. The layer uses the following SMFs:

- **Governance, Risk, and Compliance**
  This SMF concerns the assessment of the trade-off value between reaching IT goals and creating business value. Governance specifies the chain of command, accountabilities and performance evaluations that should result in consistent policies and clear objectives and goals. Risk concerns the possibility that actions, or the lack of such, may negatively influence set objectives and goals. Risks are identified for all tasks, and how these are handled depends on the specifics of the risks and if they pose a threat to highly prioritized or less important objectives. Compliance aims to make all parts of an organization aware of the regulations, policies and procedures that are obligatory, based on jurisdictional law and senior management decisions. It can also be used to evaluate the organization based on the goals stated by management.

- **Change and Configuration**
  The Change and Configuration SMF is used to keep track of the status configurations, manage the changes done to these and reduce the risk of such changes. The SMF does not only apply to changes done to IT services, but can include the management of change to an entire IT organization. The assessment of a change has to balance the risk of implementing change versus the risk of leaving the system or organization as it is. It also has to determine how appropriate controls may mitigate risks. The goals of the SMF include improved reliability or customer satisfaction, reductions in cancelled or delayed projects, clear procedures for reverting to older configurations if needed and decreased problem-solving time frames.

- **Team**
  This SMF is used throughout the life-cycle phases to ensure that each member of a project is assigned a clear role, has the skills necessary to inhabit the role and understand the tasks and responsibility they represent. Teams assembled with this in mind will be better suited to complete their tasks and reach their goals effectively, as team member tasks and responsibilities are known throughout. The goals of the Team SMF are to create role types and roles, establishing principles and best practices, and identifying and ensuring that the correct person inhabits a role.

J MOF Mapping Tables

This chapter contains three tables, one for each of the phases that were used when mapping the requirements document steps to the MOF framework. The Operate phase was not used, and is therefore omitted. Table 36 presents the requirements sections mapped to the Plan phase SMFs, Table 37 shows the requirements sections mapped to the Deliver phase SMFs, and Table 38 presents the requirements sections mapped to the Management layer SMFs.

MOF Plan Phase	
SMF	**Requirements**
Business/IT Alignment	Authority
Reliability	Authority
Policy	None
Financial Management	None

Table 36: MOF Plan Phase Mapping

MOF Deliver Phase	
SMF	**Requirements**
Envision	Problem description Scope Goals Authority High-level requirements Legal requirements Compliance requirements Risk assessment
Project Planning	Resources Initial legal study Initial compliance study Create use-case diagrams to map functionality Map data flow Create a list of assets based on the use-case and data flow Discuss the sensitivity of the assets Verification of the asset list Requirements for possible cloud deployments Requirements for possible cloud service providers Requirements for specific vendors Security requirements
Build	None
Stabilize	None
Deploy	None

Table 37: MOF Deliver Phase Mapping

MOF Management Layer	
SMF	**Requirements**
Governance, Risk, and Compliance	Compliance requirements
Change and Configuration	None
Team	None

Table 38: MOF Management Layer Mapping