Citation:
Chang, V and Ramachandran, M (2015) Towards Achieving Data Security with the Cloud Computing Adoption Framework. IEEE Transactions on Services Computing, 9 (1). 138 - 151. ISSN 1939-1374
DOI: https://doi.org/10.1109/TSC.2015.2491281

Link to Leeds Beckett Repository record:
https://eprints.leedsbeckett.ac.uk/id/eprint/1933/

Document Version:
Article (Accepted Version)

Towards achieving Data Security with the Cloud Computing Adoption Framework

Victor Chang, Muthu Ramachandran, *Member, IEEE*

Abstract— Offering real-time data security for petabytes of data is important for Cloud Computing. A recent survey on cloud security states that the security of users' data has the highest priority as well as concern. We believe this can only be able to achieve with an approach that is systematic, adoptable and well-structured. Therefore, this paper has developed a framework known as Cloud Computing Adoption Framework (CCAF) which has been customized for securing cloud data. This paper explains the overview, rationale and components in the CCAF to protect data security. CCAF is illustrated by the system design based on the requirements and the implementation demonstrated by the CCAF multi-layered security. Since our Data Center has 10 petabytes of data, there is a huge task to provide real-time protection and quarantine. We use Business Process Modeling Notation (BPMN) to simulate how data is in use. The use of BPMN simulation allows us to evaluate the chosen security performances before actual implementation. Results show that the time to take control of security breach can take between 50 and 125 hours. This means that additional security is required to ensure all data is well-protected in the crucial 125 hours. This paper has also demonstrated that CCAF multi-layered security can protect data in real-time and it has three layers of security: 1) firewall and access control; 2) identity management and intrusion prevention and 3) convergent encryption. To validate CCAF, this paper has undertaken two sets of ethical-hacking experiments involved with penetration testing with 10,000 trojans and viruses. The CCAF multi-layered security can block 9,919 viruses and trojans which can be destroyed in seconds and the remaining ones can be quarantined or isolated. The experiments show although the percentage of blocking can decrease for continuous injection of viruses and trojans, 97.43% of them can be quarantined. Our CCAF multi-layered security has an average of 20% better performance than the single-layered approach which could only block 7,438 viruses and trojans. CCAF can be more effective when combined with BPMN simulation to evaluate security process and penetrating testing results.

Index Terms— Cloud Computing Adoption Framework (CCAF),security framework, Business Process Modeling Notation (BPMN), Data security in the Data Center, multi-layered security protection.

1 Introduction

CLOUD Computing and its adoption has been a topic of discussion in the past few years. It has been an agenda for organizational adoption due to benefits in cost-savings, improvement in work efficiencies, business agility and quality of services [1-2]. With the rapid rise in Cloud Computing, software as a service (SaaS) is particularly in demand, since it offers services that suit users' need. For example, Health informatics can help medical researchers diagnose challenging diseases and cancers [3]. Financial analytics can ensure accurate and fast simulations to be available for investors [4]. Education as a Service improves the quality of education and delivery [5]. Mobile applications allow users to play online games and easy-to-use applications to interact with their peers. While more people and organizations use the Cloud services, security and privacy become important to ensure that all the data they use and share are well protected. Some researchers assert that security should be implemented before the use of any Cloud services in place [6-8]. This makes a challenging adoption scenario for organizations since security should be enforced and implemented in parallel with any services. Although organizations that adopt Cloud Computing acknowledge benefits offered by Cloud services, challenges such as security and privacy remain a scrutiny for organizational adoption. While overseeing the importance of security, the software engineering and development process should always design, implement and test security features.

The data centers have encountered challenges of rapid increase in the data [9-11]. For example, in a data center that the lead author used to work with, daily increase of 100 terabytes of data was common. If the organization has encountered a rapid rise of data growth and is unable to respond quickly and efficiently, problems such as data traffic, data security and service level agreement issues can happen [6, 11]. In this paper, we focus on the data security while experiencing a large increase of data, weather they are from the external sources such as attack of viruses or trojans; or they from the internal sources if users or clients accumulate hundreds of terabytes of data per day. This is a research challenge for data security which is essential for the better management of the data center to handle a rapid increase in the data.

Apart from the data center security management for

- *Victor Chang works for School of Computing, Creative Technologies and Engineering, Leeds Beckett University, Headinely campus, Leeds LS6 3QR, UK. He is affiliated with Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK. E-mail: V.I.Chang@leedsbeckett.ac.uk (corresponding author).*
- *Muthu Ramachandran is with School of Computing, Creative Technologies and Engineering, Leeds Beckett University, Headinely campus, Leeds LS6 3QR, UK. E-mail: M.Ramachandran @leedsbeckett.ac.uk*

*Please note that all acknowledgments should be placed at the end of the paper, before the bibliography (**note that corresponding authorship is not noted in affiliation box, but in acknowledgment section**).*

rapid growth in data, the software engineering process should be robust enough to withstand attacks and unauthorized access. The entire process can be further consolidated with the development of a framework to tighten up the technical design and implementations, governance and policies associated with good practices. This motivates us to develop a framework, Cloud Computing Adoption Framework (CCAF), to help organizations successfully adopt and deliver any Cloud services and projects. In this paper, we demonstrate our security design, implementation and solution for CCAF. We use penetration testing and related experiments to validate its robustness and measure precision, recall and F-measure to justify advantages over other approaches. The breakdown of this paper is as follows. Section 1.1, 1.2 and 1.3 present literature related to Cloud application security. Section 2 presents security overview under CCAF. Section 3 describes CCAF security in details, including the code, multi-layered approach and component for each layer. Section 4 explains how to protect data security and predicts likely consequences by using Business Process Modeling Notation (BPMN) simulations. Section 5 uses penetration testing against the CCAF multi-layered security and compares with other similar approaches. Section 6 presents Conclusion.

1.1 Cloud applications security literature and overview

We review a few selected literatures that are relevant for Cloud application security described as follows. Existing literature [7-9, 11-12] define cloud application service security as threats, vulnerabilities and protection of cloud operational services and software as a service applications Liu et al [7] has proposed an agent-oriented modeling framework for analyzing security requirements. However, it is perceived as yet another modeling language than security requirements capturing framework. Mather et al [8] provides a detailed definition and description on various cloud security and privacy issues. However, there is no clear framework to follow from security requirements. Cebula and Young [12] further classify cloud applications security engineering and its implementation into two major groups: ***software acquisition security*** (which includes the security specifications in all processes to buy, rent, or interchange software to use in an enterprise) and ***systems & software development security*** (which include the security specifications in all processes to develop information systems). However, there is no clear framework to be adopted to classify security requirements and then to feed towards implementation. A framework with a holistic approach of offering an integrated solution and multi-layered security is required.

1.2 Data security for the private clouds hosted in the Data Center

As discussed in the introduction, the rapid data growth poses challenges for data security for the private clouds hosted in the data center. Literatures for different security solutions are as follows. Zhang et al [11] provide review of the Cloud Computing and explain the research challenges associated with security. However, they only provide an overview of important security challenges but do not provide a full detailed solution on Cloud security. Liu et al [7] explain their software security analysis with their rationale and an example. However, there is a lack of details about the software design and implementation process involved, and empirical results to evaluate its performance and effectiveness of their proposed solution, which looks like the combination of UML and workflows. Yu et al [13] and Wang et al [14] propose their fine-grained security model for Cloud storage. Both are similar, except that proposal from Yu et al [14] are more in details and they explain theories and users associated with their proof-of-concept. However, both proposals [13, 14] do not have any experiments, simulation and empirical data to prove the effectiveness and robustness of their fine-grained security model. Thus, both proposals do not address in-depth data security issues, when the rapid growth of data is a challenge for the Data Center.

There are common observations in the security proposed methods: Each paper [7-8, 10, 12, 14] only proposes a single solution. In the event of fraud, cyber criminal activities and unauthorized hack, the security solution is insufficient to protect the data security and the data center if only a single solution is adopted. Hence, a better alternative is required. We proposed the multi-layered security to integrate security techniques to illustrate the essence and effectiveness of the framework with advantages of doing so. First, the strength of each technique is enhanced. Second, since each technique cannot always fully prevent hacking or provide a full solution without fallacy, the multi-layered security can improve the extent of security since it is more difficult for viruses and trojans to break different types of security in one go. The aim is to maximize security protection and reduce the threats.

To demonstrate the data security of the private clouds hosted in the data center, we propose the use of ethical hacking to demonstrate whether our CCAF multi-layered security can withstand a large amount of viruses and trojans attacks, if the rapid data increase is from the external malicious hacking. We will provide detailed process and results in Section 5.

2 Security overview under cloud computing adoption framework (CCAF)

The current challenges facing cloud community on cloud security is enormous. Therefore, we need a clear framework, which provides an integrated approach to study cloud service performances before the implementation, the one that supports clear implementation of cloud security attributes at the implementation level, and the one that can be adopted by both cloud users and cloud providers. The use of the framework is a suitable approach illustrated by Zhang et al. [15], who propose a user-based security framework for collaborative computing systems. They explain their rationale, background, core technologies, usage scenarios, experiments, results and their interpretations. Their approach is heavily focused on the use of XML to transfer and interpret data through their security mechanism. The use of the framework is an suitable approach provided with careful and clear explanations. We have proposed our own framework, Cloud Computing Adoption Framework (CCAF), to address the security challenge.

The CCAF is a comprehensive model for adopting and applying cloud security principles systematically. The outcome of each activity is shown inside the parenthesis. These best practice techniques will keep grow as the framework has been in various applications. It is a conceptual framework like ITIL version 3 to guide organizations for the best practices. Additionally, such a framework can integrate with Cloud Computing services to provide added values for adopting organizations [16]. It is also an architecture framework focused on the delivery of a security service, in the form of developing a multi-layered security for data centers. Zhang et al. (2008) explain their rationale, background, core technologies, usage scenarios, experiments, results and their interpretations. Their approach is heavily focused on the use of XML to transfer and interpret data through their security mechanism. Framework is an appropriate method provided with careful and clear explanations. This section presents the background work and overview for our proposed Cloud Computing Adoption Framework (CCAF).

2.1 Overview

We generalize the areas for security overview. The following are categories of CCAF security aims to cover:

- Application software security which deals with how we can build systems that can automatically protect themselves.
- Network (LAN, MAN, GAN), wireless network security and platform security include Operating Systems, Virtualization and systems software.
- Convergence network security where converging, multi-network media infrastructures, social networks and technologies, which is one of the emerging areas of research.
- Service-oriented security where issues related to system services such as denial of service attacks, distributed denial of services, and web services.
- Cloud security deals with services security, data security and privacy so that services delivered and assets are protected.
- Open-source software security includes issues such as trust, certification and qualification models.
- Software components and architecture, security which deals with building components and architectures with security can be used as plug-ins.
- Web services security is essential to ensure secure services are delivered with integrity.
- Systems & Software security engineering deals with building security in CCAF right from requirements. This is also considered developing software applications with CCAF.

Recommendations from McGraw [17] provide a comprehensive framework for systems engineering methods and concepts. However, it does not offer a complete solution for Cloud Computing. This motivates us to have a comprehensive design, implementation and service for Cloud security under the CCAF recommendation. CCAF is a framework for organizations that we have previously demonstrated how CCAF can be offered in healthcare [18], finance [19] and other types of businesses. It is our goal to provide guidelines and recommendation for security and privacy. Computer security has been classified into a number of general concepts and processes such as identification, which identifies objects, functions, and actions, authentication, authorization, privacy, integrity and durability. We have so far well established basic security features with identification, authentication, authorization, digital security encryption and decryption techniques. Key features with their explanations are as follows.

Identification is a basic and first process of establishing and distinguishing amongst person/user & admin ids, a program/process/another computer ids, and data connections and communications. Often we use alphanumerical string as user identification key and some may use your email as the user identification key and this can be checked against when a user login into the system. Authentication and authorization are two distinct forms of access controls to access any information in the system.

Privacy is the key to maintaining the success of cloud computing and its impact on sharing information for social networking and team work on a specific project. This can be maintained by allowing users to choose when and what they wish to share in addition to allowing encryption and decryption facilities when they need to protect specific information/data/media content.

Integrity is the basic feature of human being as a process of maintaining consistency of actions, communications, values, methods, measures, principles, expectations, and outcomes. Ethical values are important for cloud service providers to protect integrity of cloud user's data with honesty, truthfulness and accuracy at all time. In cloud computing terms, we can achieve integrity by maintaining regular redundancy checks and digital certification in addition to other basic security features of maintaining identification, authentication, and authorization.

Durability is also known as, persistency of user actions and services in use should include sessions and multiple sessions.

2.2 CCAF Security Design

This section describes the system design required by CCAF. Capturing and identifying requirements for security explicitly is one of challenges in Cloud security for SaaS, which has an impact on the functionality of the system. Therefore, we need to be able specify security requirements explicitly throughout the security-specific lifecycle phases as part of achieving CCAF (security requirements, design for security, security testing and securability testing). Tondel et al. [20] has provided an extensive survey on security requirements methods which help to identify security requirements systematically and structure them. For example, Mead [21] for the SEI's (software Engineering Institute) has identified a method known as SQUARE (Secure Quality Requirements Engineering) which has been extended SysSQUARE (Systems Engineering SQUARE) towards systems security engineering method. Our extended method consists of ten steps as follow:

- **Agree on definition** to define a set of acronyms, definitions, and domain-specific knowledge needs to be agreed by stakeholders. This will help identify and validate security-specific requirements clearly by stakeholders.
- **Identify security goals** to clearly define what is expected of the system with respect to security of business drivers, policies and procedures.
- **Develop artefacts** to develop scenarios, for examples, misuse cases and templates for specifications and forms.
- **Perform risk assessments** to conduct risk analysis for all security goals identified, conduct threat analysis.
- **Select an elicitation technique** to include systematic identification and analysis of security requirements from stakeholders in the forms of *interviews, business process modeling and simulations, prototypes, discussion and focus groups*. As part of this phase, one should identify level of security, cost-benefits analysis, organizational culture, structure and style.
- **Elicit security requirements** to include activities such as producing security requirements document based security specific principle structure as part of our goal of developing CCAF earlier, risk assessment results, and techniques identifies for analysis such as *business process modeling and simulations, threat modeling, and misuse cases*, etc.
- **Categorize security requirements** to include activities that (1) classify and categorize security requirements based on company-specific requirements specification templates and (2) use our recommended security principles as this will help Systems Engineers to apply CCAF and (3) track security-specific requirements to validate & verify at all stages of the systems engineering life-cycle.
- **Identify systems data security requirements** to include activities on extracting and carefully identifying data security and relevant sub-systems such as data centers, servers, cloud VMs, and software security, SQL security, and other types of security that are relevant to the data. This separation of concerns allows systems engineers to integrate, track, design, and develop data security as part of enterprise wide systems development.
- **Prioritize security requirements** to include activities of selecting and prioritizing security requirements based on business goals as well as cost-benefit analysis.
- **Inspect security requirements** to conduct requirements validation process using requirements inspection and review meetings.

To achieve an integrated security for the iterated requirements, one can select keywords as objects and components. System and software components should contain a CCAF multi-layered security and each layer has its own security focus. Details will be presented in Section 3 and 5.

Most of the security attributes and principles identified earlier are clearly applicable to developing cloud services with a systems engineering focus. However, there are some cloud-specific security related issues such as security in virtualization and server environments. Cloud security attributes can be found in many-fold as shown in Figure 1. Although there are many attributes available, they can be further categorized as follows:

Confidentiality, Privacy and Trust – These are well known basic attributes of digital security such as authentication and authorization of information as well protecting privacy and trust.

Cloud services security – This includes security on all its services such as SaaS, PaaS, and IaaS. This is the key area of attention needed for achieving cloud security.

Data security – This category is again paramount to sustaining cloud technology. This includes protecting and recovering planning for cloud data and service centers. It is also important to secure data in transactions.

Physical protection of cloud assets – This category belongs to protecting cloud centers and its assets.

The above cloud security attributes/characteristics are essential and useful to understand non-functional aspects of services development and service provision. These attributes are also useful for building CCAF and maintaining security.

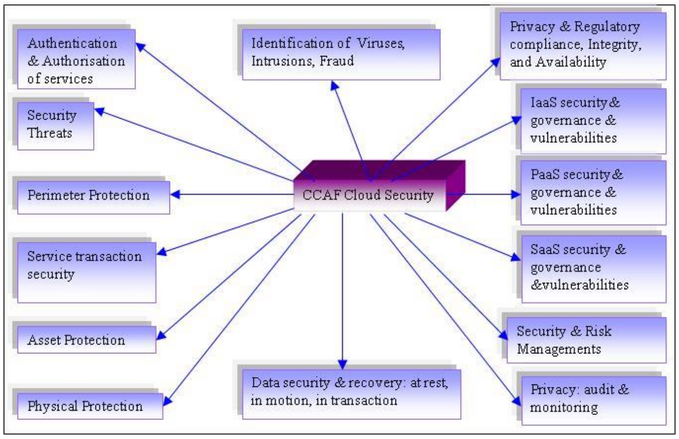
Authentication & Authorisation of services
Identification of Viruses, Intrusions, Fraud
Privacy & Regulatory compliance, Integrity, and Availability
Security Threats
IaaS security & governance & vulnerabilities
Perimeter Protection
CCAF Cloud Security
PaaS security & governance & vulnerabilities
Service transaction security
SaaS security & governance &vulnerabilities
Asset Protection
Security & Risk Managements
Physical Protection
Data security & recovery: at rest, in motion, in transaction
Privacy: audit & monitoring

Figure 1: CCAF Cloud Security Attributes serving for the community

2.3 CCAF Data security

Data security address most of the cloud computing security challenges either you consider architectural and technological concerns nor process and regulatory security challenges; all of them comes down to data in many forms such as information (deals with identity management), data in transition and transaction, data in modification, privacy of user data, and data at rest on servers and storages. However, the selections of a number of recommendations [7-9; 20-24] have identified about eight key data security issues that are:

- Data tampering deals with issues of unauthorized modification to a transaction. For example, if you add 100 times to a simple transaction of £/$1000.00 this equals to £/$100K. Oracle [22] presents that 80% of security breaches are caused by insider attacks than any other forms of security attacks.
- Eavesdropping and Data Theft deal with stealing critical personal data (personal and financial information such as credit card) during data transmission. Network and packet sniffers can be used to steal such information.
- Falsifying User Identities deals with identity theft by gaining access to data and can also threaten digital signatures with non-repudiation attacks
- Password-related threats deals with stealing and cracking passwords.
- Authorized access to tables, columns, and rows deals with security at the database level.
- Lack of accountability deals with system administrators for monitoring and protecting data access and user account management.
- Complex User Management Requirements deal with user account management strategies.
- Multi-tier Systems deal with providing access to other services and application layers.
- Scaling the security administration of multiple Systems poses extra complexity of managing cloud security as it deals with providing multiple accesses to multiple applications.

3 CCAF data security in details

This section describes different types of system development and process development for CCAF. The content includes the code syntax to proceed with the CCAF security, the architecture and the proposal of the multi-layered security.

3.1 CCAF Security Schema by XACML

This section describes the software scheme required by CCAF. Extensible Access Control Markup Language (XACML) is the language that can define the rule, permission, function and interactions in the use of SaaS and Cloud security. A proposed XACML section type, Rescue, is described here as an example. "Rescue" is used to block virus, trojans and attacks such as denials of services and unauthorized access. In the event of hacking, all the files are backed up and retrieved from secure ports such as 22 for secure FTP and 443 for secure HTTPS. Instead of displaying IP addresses in the traditional method, the IP addresses in all virtual machines are assigned at runtime.

There is an OVF ID that handles processing of the DR request. The syntax is ovf:id ="rescue" presented in Table 1. All the OVF IDs can be mapped to the required IP addresses when a VM is deployed. This allows "Rescue" to describe not just a single VM behavior, but expected communications and actions between VMs required for rescued actions. Another feature in Table 1 shows ovf:required="true", which means Rescue action is on. What triggers Rescue is when the security software detects activities from unknown IPs in the list of unknown hosts to ensure Rescue can protect all the users in real-time.

Table 1: The CCAF Security Software Schema

```
<ns: Rescue ovf:required="true" xsi:type="ovf:Rescue_Type">

<Info> Rescued actions for SaaS security </Info>
<Rule>
   <Info> Retrieve data and put them in safety </Info>
  <Protocol> tcp </Protocol>
  <DstAddr ovf:id="rescue" />
  <DstPort> 22 </DstPort>
  <DstPort> 443 </DstPort>
  <SrcAddr> any </SrcAddr>
  <SrcPort> any </SrcPort>
 </Rule>

 <Rule>
  <Info> Destinations for quarantined files if infected </Info>
  <Protocol> tcp </Protocol>
  <DstAddr ovf:id="rescue" />
  <DstPort> 3306 </DstPort>
  <SrcAddr ovf:id=" rescue" />
  <SrcPort> any </SrcPort>
 </Rule>

 <Origin>
  <Info> Firewall protection for all VMs </Info>
  <DateAdded> 2013-01-18 </DateAdded>
  <AddedBy name="Administrator" role="creator"</>
 </Origin>

</Rescue>
```

3.2 CCAF multi-layered security

CCAF security software implementation is demonstrated by its multi-layers of security mechanism to maximize protection. It also ensures reduction in the infections by trojans, virus, worms and unauthorized access and denial of service attacks. Each layer has its own protection and is in charge of one or multiple duties in the protection, preventive measurement and quarantine action presented in Figure 2.

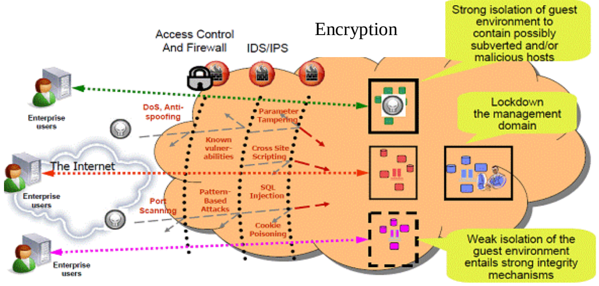

Figure 2: The CCAF multi-layered security in a nut shell

All the features in CCAF multi-layered include access control, intrusion detection system (IDS) and intrusion prevention system (IPS), this fine-grained security framework introduced fine-grained perimeter defense. The layer description is as follows.

- The first layer of defense is **Access Control and firewall** to allow restricted members to access.
- The second layer consists of the **IDS and IPS**. The aim is to detect attack, intrusion and penetration, and also provide up-to-date technologies to prevent attacks such as DoS, anti-spoofing, port scanning, known vulnerabilities, pattern-based attacks, parameter tampering, cross site scripting, SQL injection and cookie poisoning. The identity management is enforced to ensure that right level of access is only granted to the right person.
- The third layer, being an innovative approach, **Encryption**, enforces top down policy based security management; integrity management. This feature monitors and provides early warning as soon as the behavior of the multi-layered entity starts to behave abnormally; and end-to-end continuous assurance which includes the investigation and remediation after an abnormality is detected.

Although Yu et al. [13] have illustrated a similar example, their proposal is focused on theoretical concepts rather than services on offer and implementation. They focus on access control and do not have a comprehensive approach in providing multi-layered security. The details in each layer of security are presented as follows.

3.3 Layer 1: Firewall

This section describes the intrusion protection used in CCAF to ensure that all data is safeguarded all the times. The Intrusion Prevention System (IPS) is used with the core syntax includes:

```
crypto key pubkey-chain rsa
named-key realm-cisco.pub signature
key-string
```

While typing these three lines, an encrypted key-string is generated to protect the data from potential malicious hack. The key-string may look like this:

```
B199ABCB D34ED0F9 085FADC1 359C189E F30AF10A C0EFB624
7E0764BF 3E53053E
```

Once the key generation is done, the IPS configuration can be saved. Similar to "Rescue" XML tag in Section 3.1, the next step is to create a rule for IPS, followed by configuring IPS signature storage location. The final step includes IPS event notification. Their respective steps are presented as follows.

```
ip ips name <rule name> < optional ACL>
router#configure terminal
router(config)# ip ips name iosips

ip ips config location flash:<directory name>
router(config)#ip ips config location flash:ips

ip ips notify sdee
router(config)#ip ips notify sdee
```

3.4 Layer 2: Identity Management

The identity management is divided into three roles: users, CCAF server and the security manager as follows.

Users

Users can encrypt each key from his block and his own key. They can split files into blocks, encrypt them with the key, followed by signing the resulting encrypted blocks and creating the storage request. For each file, this key will be used to decrypt and rebuild the original file during the retrieval phase. The user also uses single sign-on to access each block with a compact signature scheme.

CCAF Server

Three roles are offered by the server. First, it can authenticate users during the storage/retrieval phase. Second, it can access control. Third, it can encrypt/decrypt data between users and their cloud. The data can be further encrypted to prevent dictionary attacks before being forwarded to the metadata manager (MM). Blocks are decrypted and the server verifies the signature of each block with the user's public key during the retrieval phase.

Security Manager (SM)

Security Manager (SM) stores metadata which include block signatures, encrypted keys and process identity management check. While SM checks and verifies the right identity, the CCAF security proceeds to convergent encryption, which serves as the third layer of security. SM has a link list and a small database, where the link list is as follows.

– Each node in the linked list represents a data block. The identifier of each node is obtained by hashing the encrypted data block received from the server.

–A link between two nodes, for example, nodes A and B, corresponds to the file identifier and the encryption of the key to decrypt the data block B.

SM can check whether a user is authorized to retrieve a file that he/she has requested. This offers an additional access control. Additionally, SM can communicate with the cloud service provider (SP) to store and retrieve data blocks.

3.5 Layer 3: Convergent Encryption

After the identity management phase, all data has to undergo the security test offered by Convergent encryption (CoE), which uses the hash of plaintext to work out the encryption key (K). Here is a sample example to illustrate how it works. Adam obtains the encryption key from his message M such that K = H(M), where H is a cryptographic hash function; he uses this key to encrypt his message, hence: CoE = E(K, M)= E(H(M);M), where E is a block cipher. By applying this technique, two different users with two identical plaintexts will obtain two identical ciphertexts since the encryption key is the same. This allows the cloud storage provider to perform efficient storage (such as deduplication, which means the same file is only stored and archived at one place without duplication) on such ciphertexts without having any knowledge on the original plain-texts. We then illustrate to encrypt the ciphertexts with other encryption algorithm using the same keying material for all input to prevent attacks against. The benefit is that the deduplication requirement can be compatible with CoE.

3.6 The core code to deploy security

This section explains the core code to proceed with multi-layered security to check the status of the CCAF security and introduces the state of 0 and 1. The status 0 means all activities and all files are manageable and cannot be fully controlled. The analogy is like human bodies: while there are also bad/cancerous cells, the percentage is so tiny that they are controlled. But until to a certain status trigger the body immunity, bad/cancerous cells cannot be controlled. To offset his, our human body triggers the alarm for body defense. Similar to our security design, status=1 means that an alarm is triggered and the remedy action begins. The system manager can also manually trigger it if the data center is under the threat before the system detection turns positive.

Table 2: The code syntax for CCAF security

```
While trigger(status(job)) do
check(status(job)); //to check the status is 0 or 1
 if (security == 1)
   firewall(status(job));
   identity(status(job));
   encryption(status(job));
else
   action((status(job));
   quarantine(status(job));
   report(status(job)); //report the system; do not stop CCAF
end;
end;
```

If ‘security’ is equal to 1, which means the CCAF security process is kicked off as shown in Table 2. If ‘security’ is equal 0, it means the CCAF recognizes there is a low risk and threat. The term “status(job)” means that the CCAF security is offering real-time protection and actions for quarantine. All these CCAF commands enable the functioning of multi-layered security. Explanations of other parts of the security process are as follows.

- “trigger(status(job))” is to enable the triggering of the contingency action. It is the first step to trigger a list of actions for maintaining system and data security.
- “check(status(job)” is to check the status of security is 0 or 1. The status 0 is the controlled status and status of 1 is the triggered status due to security breach or threats.
- “firewall(status(job))” is to enable firewall on.
- “identity(status(job))” is to enable identity management to be active.
- “encryption(status(job))” is to enable encryption on. By default, the first three are on.
- “quarantine((status(job))” is when the CCAF system finds the Trojans or viruses, it begins the isolate trojans and viruses and attempt to kill them or retain them to be completely isolated.
- “action((status(job))” is to manually make the above commands.
- *“report(status(job))” is to report to the system at once after “action((status(job))” or “quarantine((status(job))” are done.*

3.7 Isolation and quarantine if trojans and viruses are detected

This section describes the actions taken if trojans and viruses are found. All malicious files and signatures are first isolated. The strong isolation and integrity management is used to protect user safety while using the CCAF security service. Strong isolation is required while detecting vulnerabilities in any of the cloud services, including the block of unauthorized IPs and attack points/ports. While these malicious files and unauthorized access attempts happen, quarantine is the next step to ensure the safety and security. It first backups the data safely and then attempts to quarantine infected data. If a quarantine action is unsuccessful, the files can be kept under “quarantine area”, or chosen to be deleted. In the quarantine area, the infected files are locked up until further notice.

3.8 The integrated solution – checking all the files and data on one go

Descriptions in Section 3 present how to deal with malicious files and unauthorized access in each layer of CCAF security. Our CCAF proposal can also illustrate the integrated approach which checks all layers in one go. This is an important step due to the following reasons. First, the insider threat is an issue if the leaving employees or someone with a good knowledge of the security system can find ways to sneak through the security check [25]. Second, each layer has its own “gatekeeper” for security. There is a possibility that well-written malicious code, either disguised as safe files or disguised as part of the system files, can impose a security risk if there is no final check of the entire system. Third, often the Data Center Cloud systems serve hundreds and thousands of users and have a large number and volume of data possibly at petabytes. The security system needs to check all the status of the data and check that whether the real-time security can be offered for Data services if that includes petabytes of data, and when the data is in use. In other words, we need an intelligent way to find out how to manage such a huge amount of data has been in used and in client-server requests at all times. More details will be presented in the following section.

4 The integrated solution of data security simulated by business process modeling notation (BPMN)

This section describes the integrated solution of data security which can be achieved by the simulations offered by Business Process Modeling Notation (BPMN) which can simulate the execution time of protecting and securing petabytes of data in the data center. BPM (Business Process Modeling) is a process of identifying a number of business processes that will have an impact on stakeholders and to the system. BPMN is a tool-independent graphical process definition language to study performance evaluation of the system, clarify requirements specification (such as use cases), and is executable. IBM [26] reports on saying BPM allows us to focus on our most critical business priorities first. This section of the paper is devoted to eleven habits for highly successful BPM programs with emphasis on conducting a complete BPM and the team. BPMN allows business process to be modeled visually, simulated, optimized for efficiency (time & cost), optimized for business KPIs (key performance Indicators), and quantified for KPI measurable parameters such as security improvement [27]. KPIs are the key to achieving business improvement for sustainability and performance evaluation. Most of the existing work in this area [27-29] has largely focused on performance evaluation of core business process only. This work has applied to study the performance of cloud data security process. Hence, we have developed a number of key cloud security process that is critical for cloud data. Figure 2 proposes a good principle for the cloud architectural design process which is also based on some of the key stakeholders/concepts to consider during architectural design:

- Clients who are potential cloud customers as well as cloud administrators
- Cloud Providers
- Cloud Management Team
- Data Centers/Security Pool
- Intrusion Rejection Process

The Client of Cloud Computing contains a computer software or/and a computer hardware which dependents on cloud computing architecture to support the application deliveries, or which designed specifically for cloud service deliveries.

A Client of Cloud Computing Architecture is an interface of common cloud user through the web browsers or thin terminals. Cloud provider is the one who offers the Cloud Service Delivery Models to Client through the internet. According to our proposed system the client just sends a request to the cloud then the remaining process is taking care of cloud service provider who consists of Cloud Management Teams, Data Centers/Security pools and the Intrusion Detection Mechanisms. Figure 3 and Figure 4 represent the BPMN process of Data Request flow from Client to Cloud.

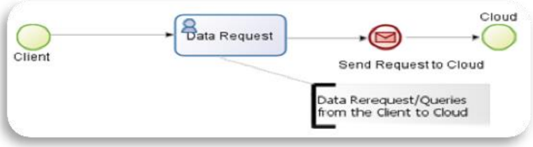

Figure 3: Data Request Business Process Model for Cloud Security

4.1 The steps involved with the simulated BPMN process

In order to understand how the BPMN process works and can offer the overall contributions to the integrated solution, this section presents the steps involved in the simulated BPMN process. Simulating a process allows us to study its behavior for external events/triggers in that process. Process simulation has been successful in several applications from low-end to high-end systems. Therefore, simulating a BPMN model helps us to study business behaviors/performance for various expected and unexpected scenarios. The BPMN simulation process consists of a number of cyclic phases (See Figure A in Appendix). BPMN starts with an actor called Client with a small circle notation which sends a message to a process (Data Request with rounded square) which task has been devoted to take action based the request and therefore send a message to the cloud (finishing circle). The second phase is to annotate each element in the process and thirdly to create tasks, assign simulation variables (different types of requests both valid and invalid) to process and tasks in that process. Finally, create messages between elements in the process and run a number of simulations. Figure 4 shows the multi-layered security demonstrated by our CCAF solution, which includes the use of firewall, identity management and encryption. This can be simulated by the BPMN following guidelines.

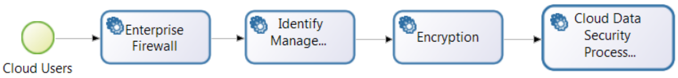

Figure 4: Our multi-layered security solution

Data centers are the essential asset of Cloud Computing corporations and private cloud deployment, these all connect to all applications, storage services and servers. The business relies on the cloud data centers supports the business values and operations and drive maximum efficiencies. The data centers are playing key roles that need to be managed and planned very carefully to meet applications and the users growing performance requirements/demands. The data centers architectures propose technologies, practices and products which help data centers engineers and management team who is responsible to answer the business goal requirements. According to our CCAF framework, any cloud data access service can follow the business process steps described in Figure 4, which include our multi-layer security protocol (Enterprise Firewall, Identity Management, Encryption/Decryption, and Cloud data security Process Controls). The Cloud data security Process Controls are further refined data security processes as shown in Figure 5, which is the BPM model for different states of models for data security. The data center's can use this model to study the performances of selected cloud data architecture. This process starts with a data status decision (diamond symbol) passes that data based on that decision to any one of the paths of the cloud storage processes (data at rest, data in use, and data in change/transition). This in turn passed on to a data security pool which is a separate lane with dedicated security processes (such as data security area and data center update) to study security controls in place before it ends.

Figure 5 is the BPM model for different states of models for data security. The data center's can use this model to study the performances of selected cloud data architecture. This process starts with a data status decision (diamond symbol) passes that data based on that decision to any one of the paths of the cloud storage processes (data at rest, data in use, and data in change/transition). This in turn passed on to a data security pool which is a separate lane with dedicated security processes (such as data security area and data center update) to study security controls in place before it ends.

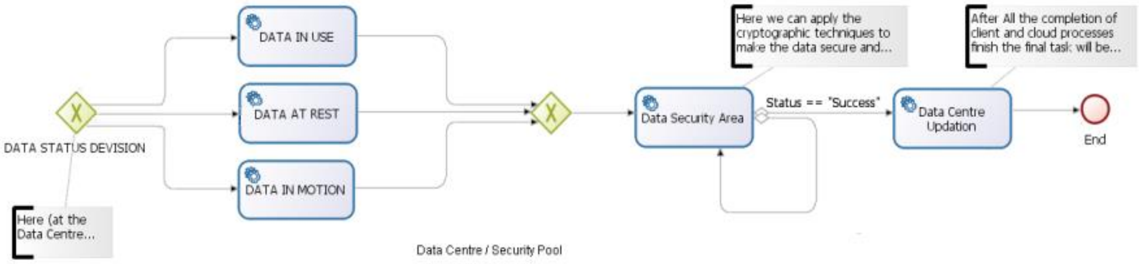
DATA IN USE
DATA AT REST
DATA IN MOTION
DATA STATUS DEVISION
Here (at the Data Centre...
Data Security Area
Status == "Success"
Here we can apply the cryptographic techniques to make the data secure and...
Data Centre Updation
After All the completion of client and cloud processes finish the final task will be...
End
Data Centre / Security Pool

Figure 5: BPMN model for three types of cloud data security

Intrusion Detection: The intrusion detection section is used to intimate the cloud management team, data centers and also its security pools about the intrusions by raising the alarms. The dangers which will happen by the intrusions are scalable (in the scale of 1-5; 1 is an ignorable danger and 5 is the most danger). In this business process and by the orders of the management team the rejection and warning messages/e-mails will be composed to send the client. The Figure 5 shows the BPMN model for simulation of Intrusion Detection/Rejection Process which can be used to study the performances of the proposed cloud architectural design.

The use of BPMN can simulate the daily operations in the data centers, which contain up to 10 petabytes (PB) of data. Figure 5 also shows the BPMN model for different states of models for data security. The data center can use this model to study the performances of selected cloud data architecture. This process starts with a data status decision (diamond symbol) passes that data based on that decision to any one of the paths of the cloud storage processes (data at rest, data in use, and data in change/transition). This in turn passed on to a data security pool which is a separate lane with dedicated security processes (such as data security area and data center update) to study security controls in place before it ends.

The process starts with a possible intrusion event (this could be an unauthorized access to a data) which triggers Raise Alarm process to compose email/message to the cloud data administrator immediately noted as the client process in this model. See Figure B in the Appendix. The following section discusses performance analysis for each of those BPM simulations.

4.2 *BPMN Simulation for petabyte data security*

This section presents Cloud big data security that is associated with the integrated solution. As explained in Section 2.3, data security for petabytes of data is a priority. In this section, we present a BPMN simulations of a Data Center, which is our data center based at the University of London Computing Center (ULCC). In our previous paper, we demonstrate that the use of Cloud bioinformatics and storage services provide added values and positive impacts, and the data center of all these services are located at ULCC [18-19]. Since Cloud security is the key to business sustainability [1-2, 6, 8], we should structure security strategy and operation to ensure all services can be delivered and optimized. This explains the importance of undertaking BPMN simulations, so that we know the execution time required if the entire data at the ULCC is at rest, or is in full use, or is involved in the transfer of data across different networks (in motion), as presented in Figure 5. In the implementation to result phase, we use BPMN for protecting the data against vulnerabilities and raising alarm in data security while all 10 PB of data in the Cloud has been intensively in used.

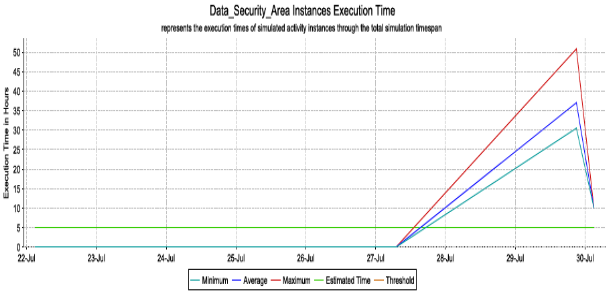

Figure 6: Data security Area Peak Access- High execution time when data in use

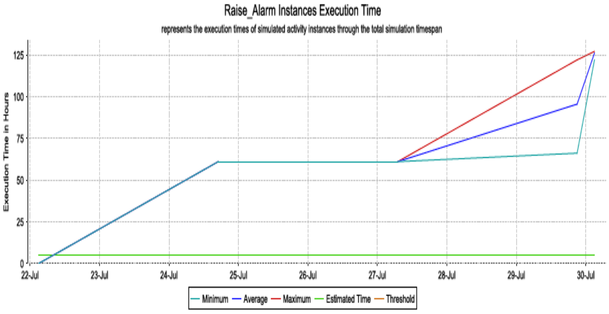

Figure 7: Raises Alarm

Figure 6 shows a graph with execution time when the BPMN process raised a security alarm. The execution time rose from 0 to three peaks (51, 36 and 30 hours) between July 27 and right before July 30, before falling to 10 hours of execution time right after July 30, 2013. The increment in execution time was necessary since BPMN alarm checked every single file and instance in 10 PB of data in the Cloud. This explained why such a long execution time was required. We plan to develop algorithms or methods that can optimize the security performance. The execution time to run each BPMN process only takes 2 seconds all the times, which has a very low execution time. This ensures that fast and efficient BPMB process can meet the requirement of business agility.

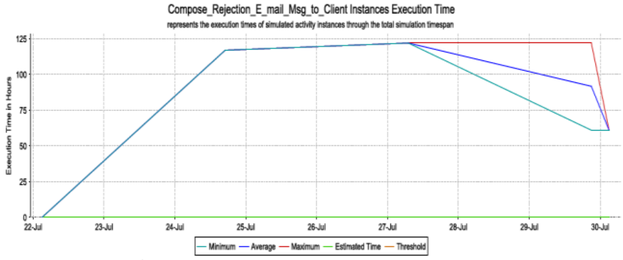

Figure 8: Rejection Message

Figure 7 shows a graph with peak execution time for entering the data security area of the business process. Results show that increased steadily from 0 to 60 hours between July 22 and the middle of July 24, 2013. The execution time stayed stable at 60 hours between the middle of July 24 and beginning of July 27. Some execution time increased due to the increasing demands in security. The implications of this result show that data security instances execution time can be high when data was constantly in use. On the other hand, the execution time was less than 2 hours if data was not in use. Figure 8 shows a performance graph for a rejection message service with peak execution times when a BPM process has sent a rejection message to allow access to data in privacy. To protect 10 PB data, it can take up to 125 hours.

For protecting cloud data, we need to distinguish different states of transitions that can occur in the cloud. This will allow us to employ appropriate data security techniques. An example model for different classes/states for cloud data is shown in Figure 5 and Figure 9. Our notion of cloud data security concept is to "Divide cloud data transactions into a few possible ways":

- Data at Rest means cloud storage servers and all types of storage on the cloud.
- Data in Change includes all types of data creation and modification processes, from file creation/deletion of folders.

4.3 Performances Evaluation and analysis of results

Section 4.2 presents BPMN simulations when all the data are in full use and capacity while they have either encountered security breach or have raised security alarm. The execution time between Figure 6 and Figure 8 represent the amount of time for the data to be fully protected or recovered after the security incident.

They do not represent the execution time of performing such BPMN simulations. This section presents the results of performing BPMN simulations in each instance. There are eight instances altogether and each time execution time was taken five times to average out, with the standard deviation of 3% of all time taken. All execution time to complete BPMN simulations for Figures 9-11 need between 1.92 and 2 seconds as shown in Figure 9. Results show that BPMN simulations support high-performance in Cloud Computing.

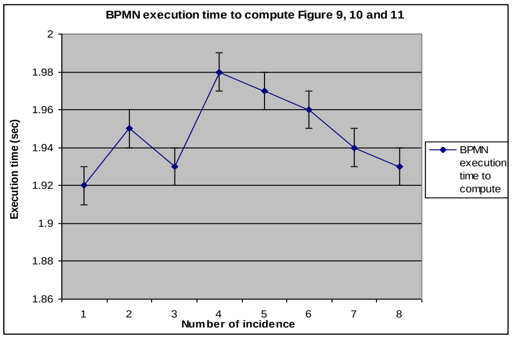

Figure 9: Execution time to perform BPMN simulations for Figures 9-11

Apart from supporting high-performance feature in Cloud Computing discussed in the earlier part of this section, this paper has introduced CCAF multi-layered security including their components and technical details. This ensures our work has made theory into practice and use multi-layered security to illustrate how to transform the conceptual framework into a real-life deployment used in Cloud security. Similarly, the use of BPMN is used to simulate the full data security of the Data Center. Results presented in Section 4.2 show that it takes between 50 and 125 hours for protecting data and raise alarm in real-time when data is in use. There is a gap time between 50 and 125 hours which leaves vulnerabilities to the data center with 10 petabytes of data in place. The use of CCAF multi-layered security provides additional protection to the data and ensures that all data can be safeguarded before the all 10 petabytes are fully optimized for services. We propose that the CCAF multi-layered security is the solution to this situation. In order to demonstrate the advantages of adopting CCAF multi-layered security, we have undertaken penetration tests to see how many viruses and trojans are trapped or cleaned, and the percentage of successful blocked rate.

5 The experiments of penetration testing for ethical hacking

To demonstrate whether the ULCC can withstand the rapid data growth due to the viruses and trojans, ethical hacking is an appropriate way to test the system performance [30-31]. Ethical hacking includes ways to penetrate into the security system in the awareness of the host. The environment for the ethical hacking was as follows. One hundred of virtual machines (VMs) were set up and each one had the CCAF multi-layered security turning on. An ethical hacking firm (which did not want its name revealed) took part in this test and provided 10,000 known viruses and trojans detected between 2010 and 2012 in the internet security breach and each of these viruses/trojans had their fix patches or repairs by the most-up-to-date security company. The objective is to test how many viruses and trojans that CCAF multi-layered security can block and quarantine. Another one hundred VMs have the Mcafee antivirus (a work partner) turning on to test the performance. This section presents the penetration testing and outcomes of the test to support that the multi-layered security can perform better for, filter out malicious attacks. To do this test, 10,000 known trojans and viruses are injected into the CCAF multi-layered security with the following numbers recorded:

- The number of viruses and trojans detected and blocked by each layer.
- The total numbers of viruses and trojans detected and blocked the system.
- The number of viruses and trojans detected but unable to be blocked and sent to quarantine.
- In the quarantine, the number of viruses and trojans that can be destroyed.
- In the quarantine, the number of viruses and trojans that cannot be destroyed.

Two types of experiments were undertaken. The first one was focused on penetration tests involved with injecting 10,000 viruses and trojans in one go. The second one was focused on continuous penetration test, such as injecting 10,000 same viruses and trojans every five hours to test that the entire data center is under the security threat as presented by BPMN simulations in Section 4.

5.1 Results of penetration tests

Figure 10 shows the results of penetration tests. 5,423 viruses and Trojans have been detected and blocked by the firewall. Another 3,742 viruses and trojans have been detected and blocked by identity management and intrusion prevention systems. 842 trojans and viruses are then detected and blocked by the encryption. All the blocked viruses and trojans can be destroyed in seconds. Amongst all these figures, there are remaining 81 viruses and trojans sent to quarantine when they cannot be destroyed directly. 79 of them can be destroyed by the quarantine. The remaining 2 viruses and trojans are unable to be destroyed but can be isolated independently. In other words, 10,000 trojans and viruses do not damage any of 10 petabytes of data in the ULCC Data Center.

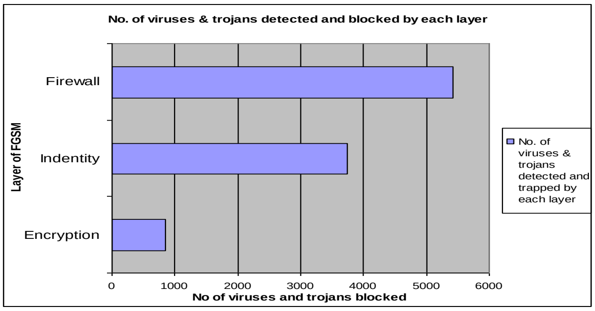

Figure 10: Results of penetration tests

Three different penetration testing metrics have been presented. First, the efficiency of penetration testing is measured based on a number of key penetration testing metrics as follows:

Penetration Test Efficiency (PTe, %) = (No. of Virus detected & blocked (V) + No. of Trojans detected & blocked (T)/Total Numbers detected & blocked (N)) x 100%

$$PTe = (\sum_1^N V + \sum_1^N T/\mathrm{N}) \text{ x } 100\% \qquad (1)$$

Second, Security Test Efficiency (STe, %) = (Number of Surface Attacks Detected, Blocked & Killed (SAs)/Total Number of Systems Surfaces Interfaces (SIs))*100%

$$STe = (\sum_0^N SAs / \sum_0^N SIs\ ) \text{ x } 100\% \qquad (2)$$

Third, Business Process Efficiency (BPMNe, %) = (PTe*Total Number of BPMN Process/Total No. of Penetration Test Hours)*100%

$$BPMNe = (PTE * BPMNt/X) \text{ x } 100\% \qquad (3)$$

Based on formula (1) to (2), the total number of viruses and trojans is 1,000 and the total number of detect and block is 9,919, total number of detect, block and kill is 9,998, hence

PTe = (9919/10000) x 100% = 99.19%
STe = (9998 / 10000) x 100% = 99.98%

5.2 Results of continuous penetration tests

Results are presented in percentages rather than the number of viruses and trojans blocked. 10,000 same viruses and trojans are injected every five hours to test how the Data Center can cope with the vulnerabilities in the most crucial 125 hours.

Results in Figure 11 show the percentage that viruses and trojans that have been blocked, which dropped from 99.19% to 76.00% in 125 hours. However, we also defend that the percentage of quarantine action is important to protect petabytes of data. If the percentage of quarantine is high, the data security can be maintained. In every 5 hours, the percentage of quarantine was measured. It started as high as 97.53% and then remains fairly constant (within 2.4% standard deviations) throughout the period of 125 hours. These results support our statements that CCAF multi-layered security can protect data security. Experiments conducted in Section 4.1 and the penetration testing took 125 hours each. The percentage of blocking has dropped to 76.00% at the end of 125 hours, in this case BPMNe=(76.00% x 125 / 125) = 76.00%.

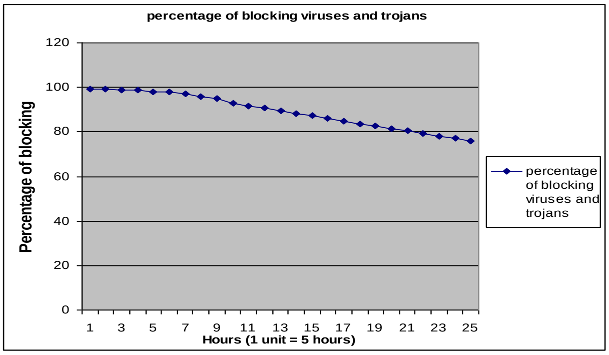

Figure 11: The percentage that viruses and trojans that have been blocked

5.3 Comparison with other approaches

This section describes comparison between our and other approaches. There are theoretical-based proposals by Goyal et al [32], Yu et al. [13], Wang et al [14] and Zissis and Lekkas [33] have addressed similar approaches with their rationale and theories in place without performing large scale experiments to check the robustness of their models. We compared CCAF multi-layered security with a single-layered approach by performing experiments. As mentioned in Section 5.1, Mcafee antivirus was used to compare performance with our CCAF multi-layered security. Mcafee service was similar to the intrusion detection system (IDS) and intrusion prevention system (IPS).

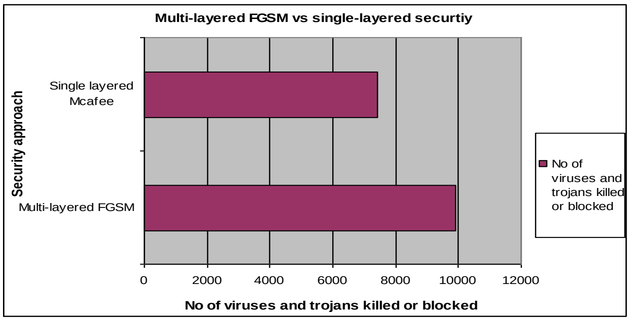

Figure 12: CCAF multi-layered security versus single-layered security (eg one Mcafee product)

10,000 viruses and trojans provided by the ethical hacking company was used and results such as the number of viruses and trojans killed or isolated were recorded. Figure 12 shows the number of viruses and trojans killed or blocked, where the CCAF multi-layered could kill/block 9,917 and the single-layered Mcafee could kill/block 7,438. We then reproduced the same experiment shown in Figure 12 to compare two approaches.

The results in Figure 13 showed that the CCAF multi-layered security has an average of 20% performance better than the adoption of a single-layered security (such as Mcafee) throughout the 125 hours of experiments. Results in our empirical studies confirm that the multi-layered approach can provide a better security service for the data center, particularly when the data security is a primary concern for the Cloud adopters and users. However, only one Mcafee product was used for comparison due to the licensing issue although multiple Mcafee products could serve like what CCAF multi-layered security could offer.

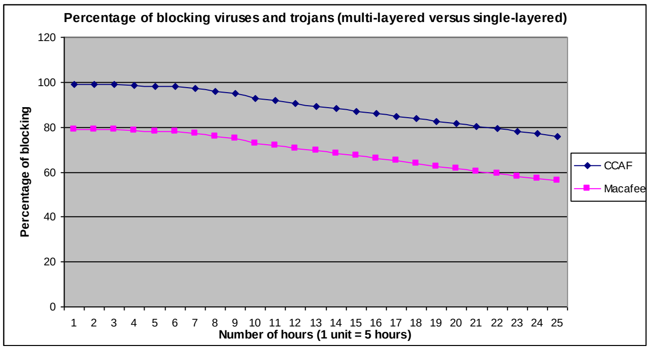

Figure 13: Percentage of blocking viruses and trojans for CCAF multi-layered versus single-layered security

5.4 Metrics, Analysis and comparison

This section presents the metrics and its analysis and comparison with other methods based on our experimental results particularly penetration testing. Antunes and Vieira [34] use four types of tools for penetration testing, explain the use of precision, recall and F-measure to justify the validity of their results. Amongst all the four tools for penetration testing, all results were very low. Although their third tool had a precision of 1, its recall and F-measure values are 0.019 and 0.037, which are extremely low. The metrics is based on our penetration testing results and transform them into precision, recall and F-measure. Their definitions are as follows. Precision is the ratio of correctly detected vulnerabilities to the number of all detected vulnerabilities:

$$precision = \frac{t_p}{t_p + f_p} \tag{4}$$

Recall is the ratio of true vulnerabilities detected to the number of known vulnerabilities:

$$recall = \frac{t_p}{t_v} \tag{5}$$

where:

- True positive (tp) refer to the number of true vulnerabilities detected;
- False positives (fp) refer to the number of vulnerabilities detected but do not exist.
- True vulnerabilities (tv) refer to the total number of vulnerabilities detected in penetration tests.

F-measure can be presented in terms of precision and recall as follows.

$$\text{F-measure} = \frac{2 \times precision \times recall}{precision + recall} \tag{6}$$

Services that can generate a high F-measure mean they are better services [34]. If a service obtains a precision of 0.8 means it can detect vulnerability with 80%. A recall of 0.9 means 90% of the known vulnerabilities is detected. While using formula (3), F-measure is equal to 0.8471. The combination of precision, recall and F-measure can determine the quality of the security services. We reproduce the experiments conducted by [34] and then compare results of CCAF multi-layered security with VS1, VS2, VS3 and VS 4 tools due to similarities with CCAF technologies except each is single layered security.

Table 3: comparison between CCAF and other single-layered services

Services	Precision	Recall	F-Measure
CCAF	**1**	**0.9919**	**0.996**
VS1	0.455	0.323	0.378
VS2	0.388	0.241	0.297
VS3	1	0.019	0.037
VS4	0.567	0.241	0.338

Results in Table 3 show that the CCAF multi-layered security can provide a much better service since all the true vulnerabilities can be detected with precision as 1. Since only 5 out of 10,000 are missed, the recall is 0.995, resulting in F-measure as 0.9975, which are above all the test results.

5.5 How to use CCAF for organizations

CCAF can be used on each VM and each server to check all the incoming data to see whether they are clean, quarantine and free of suspected malicious files. Suspected files will be alerted and moved to the quarantine section ready for further checks. Since experiments have been conducted over 125 hours with 99.19% PTe, 99.98% STe, 100% precision, 99.19% recall and 99.5% F-measure, there is a good reliability. The use of CCAF mutli-layered security can ensure the high level of protection and safeguard of data security for the organizations.

5.6 Relevance to Big Data

Our paper demonstrates data security using CCAF multi-layered security to illustrate our proofs-of-concepts. There are five characteristics with Big Data: volume, velocity, variety, veracity and value. Our work meets volume, since extensive experiments and simulations had been performed for 10 petabytes of data. Our work also meets velocity, since 10,000 viruses and trojans had been injected into our multi-layered security to test how our proposed solution can handle a large amount of infected files. The finding was that up to 125 hours were required to gain control and full data recovery. Experimental results in Section 5 also support veracity, since more than 99% of viruses and trojans can be blocked and removed under the ethical penetration test.

6 Conclusion and Future Work

Our paper has demonstrated the CCAF multi-layered security for the data security in the Data Center under the proposal and recommendation of CCAF guidelines. We explained the rationale, overview, components in the CCAF, where the design was based on the requirements and the implementation was illustrated by its multi-layered security. We explained how multi-layered security was a suitable method and recommendation, since it offered multiple protection and improvement of security for 10 PB of data in the Data Center based at the University of London Computing Center (ULCC). We explained the technical details in each layer of security and propose an integrated solution to check all the data when data is intensively used. We used the Business Process Modeling Notation (BPMN) to simulate the cases of how the data can be used, either at rest, in use, or in motion. All simulations could be completed within 2 seconds.

Our BPMN simulation results showed that it could take up to 50 hours to protect all the 2PB data and up to 125 hours to raise an alarm to take control of the situation in the ULCC Data Center. This means that an integrated approach was required to ensure data protection, in case that the data center is under the attack or potential threat from the rapid rise of data growth in the data center, which can be due to the external intrusion or the internal rapid consumption. We then used FGSM for the penetration testing. 10,000 viruses and trojans were injected into Data Center with two experiments performed. The first experiment showed that firewall, identity management and encryption could block 5,423, 3,742 and 842 viruses and trojans respectively. The remaining 81 could be either quarantined or isolated. The second experiment showed that continuous injection of 10,000 viruses and trojans could make the blocking rate decreased from the 99.19% to 76.00% in 125 hours. Despite of this result, the CCAF multi-layered security could quarantine and isolate 97.53% of viruses and trojans. Our work can demonstrate that the use of CCAF multi-layered security can protect the data center from the rapid data growth due to the security breach, and the use of BPMN can calculate how much time required for rescue action if the data security is compromised. In this way, we can work out the better tactics and plans for data recovery and security.

In this paper, we demonstrated that CCAF multi-layered security could provide the additional protection for all 10 PB of data in 125 hours when the Data Center was under the security threat and attack. Data security in the Cloud is an important issue for Cloud adoption. We demonstrated that our approach could provide real-time protection of all the data, block the majority of threats and quarantine the petabyte systems in the Data Center. We plan to improve our method and code in the simulation and choose the right type of algorithms to improve the overall performance in execution time of data security and blocking viruses/trojans in real-time. We will develop more services and proofs-of-concept in CCAF to improve the performance of BPMN simulation and penetration testing. Existing studies on cloud security [11, 14, 20-24; 28-29, 33] have been focused on either identify management, general issues concerning cloud security, access control or architecture layers. Our approach provides an integrated solution to cloud security based on a clear framework, business process modeling to study the impact on the performance of a user accessed service which is often learned on the fly which is costly and a CCAF three layered model.

References

[1] S., Marston, Z., Li, S., Bandyopadhyay, J., Zhang, A., Ghalsasi, "Cloud computing – The business perspective". Decision Support Systems, Elsevier, 51(1): pp 176-189, 2011.

[2] M. A., Vouk, "Cloud Computing – Issues, Research and Implementations". Journal of Com-puting and Information Technology - CIT 16, page 235–246, Volume 4, 2008.

[3] A. K., Jha, C. M., DesRoches, E. G., Campbell, K., Donelan, S. R., Rao, T. G., Ferris, & D., Blumenthal. "Use of electronic health records in US hospitals. New England Journal of Medicine", 360(16), 1628-1638, 2009.

[4] H. T., Peng, W. W., Hsu, C. H., Chen, F., Lai, J. M. Ho, "FinancialCloud: Open Cloud Framework of Derivative Pricing. In Social Computing (SocialCom), 2013 International Conference on (pp. 782-789). IEEE, 2013, September.

[5] M., Mircea, A. I., Andreescu, "Using cloud computing in higher education: A strategy to improve agility in the current financial crisis". Communications of the IBIMA, 2011, 1-15.

[6] M., Armbrust, A., Fox, R., Griffith, A. D., Joseph, R. H., Katz, A.,

Konwinski, G., Lee, D., Patterson, A., Rabkin, I., Stoica, M., Zaharia, "Above the Clouds: A Berkeley View of Cloud computing". Communications of the ACM, 53(4), 50-58, 2010.

[7] L., Liu, E., Yu, & J., Mylopoulos, "Security and privacy requirements analysis within a social setting". In Requirements Engineering Conference, 2003. Proceedings, 11th IEEE International (pp. 151-161), IEEE, 2003, September.

[8] T., Mather, S., Kumaraswamy, S. Latif, (2009), "Cloud security and privacy: an enterprise perspective on risks and compliance". ISBN: 978-0-596-80276-9, O'Reilly Media, Inc.

[9] M., Pop, S. L., Salzberg, "Bioinformatics challenges of new sequencing technology". Trends in Genetics, 24(3), 142-149, 2008.

[10] A., Greenberg, A., J., Hamilton, D. A., Maltz, P., Patel, "The cost of a cloud: research problems in data center networks". ACM SIGCOMM computer communication review, 39(1), 68-73, 2008.

[11] Q., Zhang, L., Cheng, R., Boutaba, "Cloud computing: state-of-the-art and research challenges". Journal of internet services and applications, 1(1), 7-18, 2010.

[12] J. J. Cebula, L. R. Young, "A Taxonomy of Operational Cyber Security", Technical Note: CMU/SEI-2010-TN-028, Software Engineering Institute, USA, December 2010.

[13] S., Yu, C., Wang, K., Ren, W., Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing". In INFOCOM, 2010 Proceedings IEEE, 1-9, March 2010.

[14] G., Wang, Q., Liu, J., Wu, "Hierarchical attribute-based encryption for fine-grained access control in cloud storage services". In Proceedings of the 17th ACM conference on Computer and communications security (pp. 735-737), ACM, 2010, October.

[15] X., Zhang, M., Nakae, M. J., Covington, R., Sandhu, "Toward a usage-based security framework for collaborative computing systems", ACM Transactions on Information and System Security (TISSEC), 11(1), 3, 2008.

[16] G. McGraw, " Software security: building security in, Addison Wesley, USA, 2006

[17] P., Brooks, J., Chittenden, "Metrics for Service Management: Designing for ITIL". Van Haren Publishing, ISBN: 978 90 8753 6480, 2012.

[18] V., Chang, R. J. Walters, G. Wills, "Cloud Storage and Bioinformatics in a private cloud deployment: Lessons for Data Intensive research". Springer: CLOSER 2012, CCIS 367, pp. 245–264, 2013.

[19] V., Chang, "Business Intelligence as Service in the Cloud". Future Generation Computer Systems, DOI: http://dx.doi.org/10.1016/j.future.2013.12.028, 2014.

[20] I. A., Tondel, et al., "Security requirements for rest of us: a survey". IEEE Software, Special Issue on Security and Agile requirement engineering methods, Jan/Feb, 2008.

[21] N.R., Mead, et. al., "Security Quality Requirements Engineering (SQUARE) Methodology". Technical Report, CMU/SEI-2005-TR-009, 2005.

[22] Oracle, "Data Security Challenges". Oracle9i security overview release number 2(9.2), accessed on 4th November, http://docs.oracle.com/cd/B10501_01/network.920/a96582/overview.htm , 2012.

[23] V. Kumar, Swetha M.S, Muneshwara M. S., Prof Prakash S "Cloud computing: towards case study of data security mechanism", International Journal of Advanced Technology & Engineering Research (IJATER), Volume 2, Issue 4, July 2012.

[24] F., Wen, L., Xiang, "The Study on Data Security in Cloud Computing based on Virtualization", IEEE 2011 International Symposium on IT in Medicine and Education (ITME), 2(1) Guangzhou, 2011.

[25] B., Schneier, "Beyond fear". New York: Copernicus Books, ISBN: 978-0-387-02620-6, 2003.

[26] IBM, "Eleven habits for highly successful BPMprograms". IBMThought Leadership White Paper, 2010.

[27] G. M., Cimino and G. Vaglini., An Interval-Valued Approach to Business Process Simulation Based on Genetic Algorithms and the BPMN, Information 2014, 5, 319-356; doi:10.3390/info5020319

[28] A., Behl., and K., Behl, An analysis of cloud computing security issues, 2012 World Congress on Information and Communication Technologies (WICT), November, Trivandrum, India.

[29] V., Vardharajan, and U. Tupakula, "Security as a Service Model for Cloud Environment", IEEE Transactions on Network and Service Management 11(1), 60-75, March 2014.

[30] M., Bishop, "About penetration testing". Security & Privacy, IEEE, 5(6), 84-87, 2007.

[31] M. H., Yang, N., Chandlrees, B., Lin, H. Y., Chao, "The effect of perceived ethical performance of shopping websites on consumer trust". Journal of Computer Information Systems, 50(1), 15, 2009.

[32] V., Goyal, O., Pandey, A., Sahai, B., Waters, "Attribute-based encryption for fine-grained access control of encrypted data". In Proceedings of the 13th ACM conference on Computer and communications security, ACM, pp. 89-98, October, 2006.

[33] D., Zissis, D., Lekkas, "Addressing cloud computing security issues". Future Generation Computer Systems, 28(3), pp. 583-592, 2012.

[34] N., Antunes, N., & M., Vieira, M. "Assessing and Comparing Vulnerability Detection Tools for Web Services: Benchmarking Approach and Examples". IEEE Transactions on Services Computing, 8(2), 269-283, 2015.

Biographies of the authors

Dr. Victor Chang is a Senior Lecturer at Leeds Beckett University since September 2012. Within four years, he completed PhD (CS, Southampton) and PGCert (Higher Education, Fellow) part-time. He helps organizations in achieving good Cloud design, deployment and services. He won a European Award on Cloud Migration in 2011 and a best paper in 2012, and numerous awards since 2012. He is one of the most active practitioners and researchers in Cloud Computing, Big Data and Internet of Things in the UK. He is an Editor-in-Chief of IJOCI & OJBD journals, Editor of FGCS, founding chair of two international workshops and founding Conference Chair of IoTBD 2016 www.iotbd.org and COMPLEXIS 2016 www.complexis.org.

Dr. Muthu Ramachandran is a Principal Lecturer at Leeds Beckett University. He has extensive research coupled with teaching experiences on software and systems engineering methods & lifecycle, software development, Agile software engineering, project management, process improvement, internet technology, mobile, networks, SOA, IT systems, Cloud Computing and distributed computing. He also co-edited and wrote several books.