



# NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

## THESIS

**MULTI-LEVEL SECURE INFORMATION SHARING  
BETWEEN SMART CLOUD SYSTEMS OF SYSTEMS**

by

Jun Ming Kelvin Lim

March 2014

Thesis Advisor:  
Co-Advisor:

Deborah E. Goshorn  
Gary Parker

**Approved for public release; distribution is unlimited**

*Reissued 1 Jul 2014 with corrections to in-text Figure and Table citations.*

THIS PAGE INTENTIONALLY LEFT BLANK

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.				
<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> March 2014	<b>3. REPORT TYPE AND DATES COVERED</b> Master's Thesis	
<b>4. TITLE AND SUBTITLE</b> MULTI-LEVEL SECURE INFORMATION SHARING BETWEEN SMART CLOUD SYSTEMS OF SYSTEMS			<b>5. FUNDING NUMBERS</b>	
<b>6. AUTHOR(S)</b> Jun Ming Kelvin Lim				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Naval Postgraduate School Monterey, CA 93943-5000			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> N/A			<b>10. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b> The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB protocol number ____N/A____.				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited			<b>12b. DISTRIBUTION CODE</b> A	
<b>13. ABSTRACT (maximum 200 words)</b>  There is a need to have secure information sharing in the industry and government sectors. For example, countries within the North Atlantic Treaty Organization (NATO) often have a common goal requiring them to communicate, but they lack a technological platform for fast information sharing, especially if the countries have different access rights to the information. Thus, the same information that an organization wants to share with multiple partners needs to be securely shared at multiple levels. In addition, the manner in which information is shared needs to be flexible enough to accommodate changes on demand, due to the nature of the information or relationship with the sharing organizations.  This thesis proposes a configurable, cloud infrastructure that enables multiple layers of secure information sharing between multiple organizations. This thesis follows a systems engineering process to propose a preliminary architecture of such a system, including an analysis of alternatives of some of the attributes of the system. Secondly, the thesis instantiates part of the proposed architecture with a proof-of-concept physical system in a laboratory environment. The proof-of-concept chooses a specific scenario of information sharing that would allow NATO members to access shared data faster, and in a secure fashion, in order to make decisions more quickly with the authorized information.				
<b>14. SUBJECT TERMS</b> Multi-Level Secure Information Sharing, System Architecture, Data Storage, Search, VPN			<b>15. NUMBER OF PAGES</b> 79	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> UU	

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release; distribution is unlimited**

**MULTI-LEVEL SECURE INFORMATION SHARING BETWEEN SMART  
CLOUD SYSTEMS OF SYSTEMS**

Jun Ming Kelvin Lim  
Civilian, Singapore Defence Science Technology Agency  
B.Eng., National University of Singapore, 2007

Submitted in partial fulfillment of the  
requirements for the degree of

**MASTER OF SCIENCE IN SYSTEMS ENGINEERING**

from the

**NAVAL POSTGRADUATE SCHOOL  
March 2014**

Author: Jun Ming Kelvin Lim

Approved by: Deborah E. Goshorn  
ThesisCo- Advisor

Gary Parker  
Thesis Co-Advisor

Clifford Whitcomb  
Chair, Department of Systems Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

## **ABSTRACT**

There is a need to have secure information sharing in the industry and government sectors. For example, countries within the North Atlantic Treaty Organization (NATO) often have a common goal requiring them to communicate, but they lack a technological platform for fast information sharing, especially if the countries have different access rights to the information. Thus, the same information that an organization wants to share with multiple partners needs to be securely shared at multiple levels. In addition, the manner in which information is shared needs to be flexible enough to accommodate changes on demand, due to the nature of the information or relationship with the sharing organizations.

This thesis proposes a configurable, cloud infrastructure that enables multiple layers of secure information sharing between multiple organizations. This thesis follows a systems engineering process to propose a preliminary architecture of such a system, including an analysis of alternatives of some of the attributes of the system. Secondly, the thesis instantiates part of the proposed architecture with a proof-of-concept physical system in a laboratory environment. The proof-of-concept chooses a specific scenario of information sharing that would allow NATO members to access shared data faster, and in a secure fashion, in order to make decisions more quickly with the authorized information.

THIS PAGE INTENTIONALLY LEFT BLANK

# TABLE OF CONTENTS

I.	INTRODUCTION.....	1
A.	OVERVIEW .....	1
B.	RESEARCH QUESTIONS .....	2
C.	DISCUSSION OF TOPIC .....	2
D.	BENEFIT OF STUDY .....	4
E.	ORIGINATION OF THESIS TOPIC FROM RADM TIGHE.....	5
F.	THESIS OUTLINE.....	6
II.	BACKGROUND OF FUNDAMENTAL TECHNOLOGIES .....	7
A.	DETECT, IDENTIFY, PREDICT, REACT .....	7
B.	SMART CLOUD SYSTEM OF SYSTEMS .....	9
C.	VIRTUALIZATION ACCESS.....	10
1.	VMWare Virtualization .....	11
2.	Virtual Private Networks .....	12
D.	“BIG DATA” STORAGE, SEARCH, AND REPLICATION TECHNOLOGY .....	13
1.	Apache Accumulo .....	13
2.	MongoDB.....	14
3.	Apache SOLR with Lucene Search Engine .....	14
III.	SYSTEM ARCHITECTURE FOR MULTI-LEVEL SECURE INFORMATION SHARING .....	15
A.	SCENARIOS .....	15
1.	Philippines Typhoon Disaster .....	16
2.	Other Scenarios.....	17
B.	FUNCTIONAL ARCHITECTURE.....	18
1.	External Systems Diagram/IDEF0 Level 1 .....	18
2.	Functional Hierarchy .....	23
3.	IDEF0 Functional Decompositions.....	27
C.	PHYSICAL ARCHITECTURE .....	32
1.	Overview.....	32
2.	Technological Platform .....	33
3.	Smart Algorithm Configurator Physical Architecture .....	38
D.	ANALYSIS OF ALTERNATIVES FOR PHYSICAL ARCHITECTURE .....	40
IV.	PROOF OF CONCEPT .....	43
A.	PROOF-OF-CONCEPT SCOPE AND TARGETED ANALYSIS OF ALTERNATIVES.....	43
1.	Scenario Review and Scope .....	43
2.	Proposed Architecture System Scope.....	43
3.	Experimental Setup .....	44
4.	Coding .....	46
a.	<i>InformationSecureSend()</i> .....	46

<b>b.</b>	<b><i>InformationSecureReceive()</i></b> .....	<b>49</b>
<b>5.</b>	<b>Results</b> .....	<b>50</b>
<b>V.</b>	<b>CONCLUSION AND FUTURE WORK</b> .....	<b>53</b>
<b>A.</b>	<b>CONCLUSION</b> .....	<b>53</b>
<b>B.</b>	<b>FUTURE WORK</b> .....	<b>53</b>
	<b>LIST OF REFERENCES</b> .....	<b>55</b>
	<b>INITIAL DISTRIBUTION LIST</b> .....	<b>57</b>

## LIST OF FIGURES

Figure 1.	A concept diagram of the proposed multi-level secure datacube sharing system. ....	3
Figure 2.	AI Systems Solution Pyramid (Goshorn 2010). ....	8
Figure 3.	DIPR as the foundation of AI and Object Behavior Modeling (Goshorn 2010). ....	9
Figure 4.	Smart cloud system of systems (Goshorn 2013). ....	10
Figure 5.	Simple Diagram of a VMWare server providing virtual PCs to multiple users who have a VMWare client. ....	11
Figure 6.	VPN setup that enables a computer to interface with a trusted Local Area Network (LAN). ....	13
Figure 7.	Information generation in humanitarian and disaster relief assistance. ....	16
Figure 8.	Information sharing scenario. ....	17
Figure 9.	External systems diagram/ IDEF0 Level 1. ....	20
Figure 10.	Provide multi-level secure information sharing system A0. ....	23
Figure 11.	Decomposition of configure system Function A1. ....	24
Figure 12.	Decomposition of provide multi-level security interface processor and storage function A2. ....	25
Figure 13.	Decomposition of enable information group to access new information function A3. ....	26
Figure 14.	Decomposition of provide secure information sharing environment function A4. ....	27
Figure 15.	Process flow of functions for A1. ....	28
Figure 16.	Process flow of functions for A2. ....	28
Figure 17.	Process flow of functions for A3. ....	30
Figure 18.	Process flow of functions for A4. ....	30
Figure 19.	Process flow of Level 3 functions for A4. ....	31
Figure 20.	Concept diagram of the physical architecture of information sharing by the home country with the information group. ....	32
Figure 21.	Concept diagram of the physical architecture of the home country. ....	33
Figure 22.	Physical of technological platform and smart algorithm configurator for a country. ....	34
Figure 23.	Physical architecture of a country with servers for sharing information with information group. ....	35
Figure 24.	Physical architecture of a country sharing data with an information group via VPN. ....	36
Figure 25.	Physical architecture of a country sharing data via data infusion platform. ....	37
Figure 26.	Full physical architecture of a country sharing data. ....	38
Figure 27.	Overview of countries sharing data using Information Sharing Authority. ....	39
Figure 28.	Data Management using smart algorithm. ....	40

Figure 29.	Full physical architecture of the system with the proof of concept portion highlighted. ....	45
Figure 30.	Schema diagram of experimental setup. ....	46
Figure 31.	InformationSecureSend() algorithm. ....	46
Figure 32.	MATLAB code for security key generator. ....	47
Figure 33.	MATLAB code for conversion of tweet data format for use with encryption code. ....	47
Figure 34.	MATLAB code for encryption. ....	48
Figure 35.	MATLAB code for preparing the tweet datas for sending. ....	48
Figure 36.	MATLAB code for sending tweet data via encrypted datalink. ....	49
Figure 37.	MATLAB code for sending tweet data via unencrypted datalink. ....	49
Figure 38.	InformationSecureReceive() algorithm. ....	49
Figure 39.	MATLAB code for receiving and decrypting tweet data. ....	50

## LIST OF TABLES

Table 1.	Comparison of methods to store, query and replicate the data. ....	41
Table 2.	Experimental table with factorial combination of the different encryptions. ....	44
Table 3.	Experimental results for the factorial combination of the different encryptions. ....	51

THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF ACRONYMS AND ABBREVIATIONS

AI	Artificial Intelligence
C2	command and control
DIPR	Detect, Identify, Predict, React
ESD	external systems diagram
HA/DR	Humanitarian Assistance/Disaster Relief
HTTP	Hypertext Transfer Protocol
ISE	Information Sharing Environment
JMX	Java Management Extensions
JSON	JavaScript Object Notation
NATO	North Atlantic Treaty Organization
NFOV	near field of view
NoSQL	Not only SQL
SoS	system of systems
SQL	Structured Query Language
SSL/TLS	Secure Sockets Layer/Transport Layer Security
VPN	virtual private network
WFOV	wide field of view
XML	Extensible Markup Language

THIS PAGE INTENTIONALLY LEFT BLANK

## EXECUTIVE SUMMARY

There is a need to have secure information sharing in the industry and government sectors (CJCSI 2013). For example, countries within the North Atlantic Treaty Organization (NATO) often have a common goal requiring them to communicate, but they lack a technological platform for fast information sharing, especially if the countries have different access rights to the information.

As the industry and government sectors are migrating toward cloud environments, this enables easier information sharing. However, there are times when an organization wants to share only partial data with one organization and a different set of partial information with another organization, so the same information that an organization wants to share with multiple partners needs to be securely shared at multiple levels.

There are several scenarios in which two countries need to share information in a fast manner. Figure 1 depicts one scenario in which a country, such as the Philippines, needs assistance from other countries, such as the United States, Japan, and Singapore. In one hypothetical scenario the Philippines may want to share certain ground knowledge and organic camera sensor data with countries that are providing aid. In addition to the food, water, and medical aid, the Philippines may want some information based on sensors that the aiding countries own, such as social media sensors. Figure 1 depicts the two types of data flows that may be desired in such a hypothetical Humanitarian Assistance/Disaster Relief (HA/DR) operation.

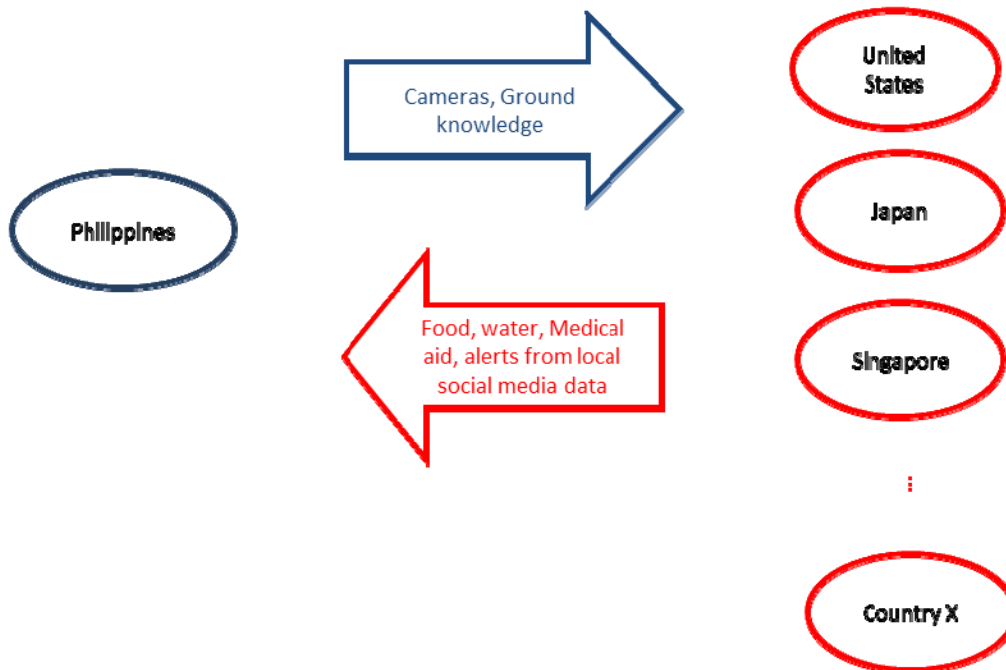


Figure 1. Hypothetical scenario in which a country, such as the Philippines, requests information from other countries' sensors, such as social media data.

It may be that countries such as the Philippines, Japan, the United States, and Singapore desire to exchange only partial information, due to different access rights associated with the country requesting information. It also may be the case that the country needing the information needs it as fast as possible. There is, therefore, a need to enable multi-level secure information sharing between countries in a fast manner.

Also, the manner in which a system shares its information with others needs to be flexible to accommodate changes on demand, due to the nature of the information or relationship with the sharing organizations. Therefore, there is a need for a configurable platform within a cloud environment that enables multiple levels of secure information sharing.

This thesis proposes a configurable, cloud infrastructure that enables multiple layers of secure information sharing between multiple organizations. Using a systems engineering process, this thesis explores the intricacies involved to enable a cloud-computing, information-sharing system of systems for

North Atlantic Treaty Organization (NATO) countries. The thesis explores the system architecting of information sharing between NATO countries. It also recommends the different information security levels required to provide a platform for different types of information to be shared. This cloud would allow NATO members to access data faster and make decisions more quickly. In particular, this thesis focuses on sharing information represented as structured files.

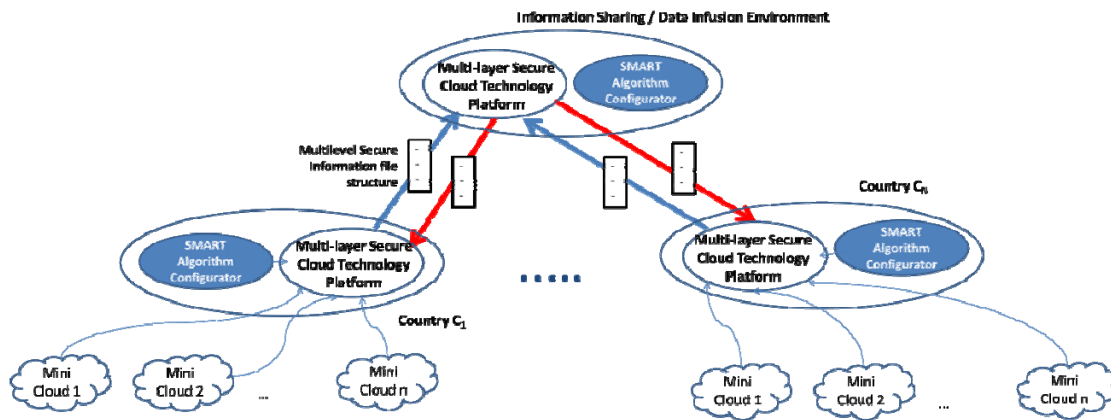


Figure 2. A concept diagram of the proposed multi-level secure datacube sharing system.

This thesis follows a systems engineering process to propose a system architecture and an analysis of alternatives for various aspects of the system. Finally, it instantiates part of the architecture with a proof-of-concept system in a laboratory environment.

## LIST OF REFERENCES

Chairman of the Joint Chiefs of Staff Instruction 6285.01C. (2013). "Multi-national and Other Mission Partners (MMNP) Information Sharing Requirements Management Process." Accessed February 28, 2014. [http://www.dtic.mil/cjcs\\_directives/cdata/unlimit/6285\\_01.pdf](http://www.dtic.mil/cjcs_directives/cdata/unlimit/6285_01.pdf).

THIS PAGE INTENTIONALLY LEFT BLANK

## **ACKNOWLEDGMENTS**

I would like to extend my sincere appreciation to my advisor, Dr. Deborah Goshorn for her valuable comments and input into this thesis. I would also like to thank Mr. Gary Parker for his constant support and important feedback for this thesis, and Mr. Daniel Zulaica for his assistance in the laboratory work. Most importantly, I would like to extend my appreciation to my lovely wife, Stella Tan, for her unwavering love and encouragement in aid of my pursuit of a master's degree at the Naval Postgraduate School.

THIS PAGE INTENTIONALLY LEFT BLANK

# I. INTRODUCTION

As an introduction to this thesis, this chapter provides a short overview of the thesis topic, the research question it seeks to answer, and the proposed thesis solution. It also highlights the benefit of this study and describes the thesis outline.

## A. OVERVIEW

There is a need to have secure information sharing in the industry and government sectors (CJCSI 2013). For example, countries within the North Atlantic Treaty Organization (NATO) often have a common goal requiring them to communicate, but they lack a technological platform for fast information sharing, especially if the two countries have different access rights to the information.

The industry and government sectors are migrating toward cloud environments which enables easier information sharing. However, there are times when an organization wants to share only partial data with one organization and a different set of partial information with another organization, so the same information that an organization wants to share with multiple partners needs to be securely shared at multiple levels.

The manner in which a system shares its information with others needs to be flexible to accommodate changes on demand, due to the nature of the information or relationship with the sharing organizations. Therefore, there is a need for a configurable platform within a cloud environment that enables multiple levels of secure information sharing.

This thesis proposes a configurable, cloud infrastructure that enables multiple layers of secure information sharing between multiple organizations. This thesis explores the intricacies involved to create a cloud computing system for NATO countries using a systems engineering process. The thesis explores the system architecting of information sharing between NATO countries. It recommends the different information security levels required to provide a

platform for different types of information to be shared. This cloud would allow NATO members to access data faster and make decisions more quickly.

## **B. RESEARCH QUESTIONS**

This thesis will answer the following primary question: “How does one apply systems engineering to create a capability architecture and proof-of-concept solution that enables countries within the North Atlantic Treaty Organization (NATO) to speed up secure information sharing so that decisions can be made promptly?”

As part of the research process, this thesis will also answer the following subsidiary questions:

1. What are the requirements for information sharing across multiple information access levels?
2. What functions should a solution for information sharing across multiple information access levels perform?
3. What is a proposed technological platform for information sharing across multiple information access levels?
4. What is an appropriate information file structure for allowing files to be accessed at multiple information access levels?
5. What are the input parameters for a smart algorithm that automatically configures the technological platform and information files to enable automated information sharing for different information access rights?

## **C. DISCUSSION OF TOPIC**

This thesis research proposes a system architecture to implement a multi-level secure methodology for how data with different information security levels can be shared via a configurable technological platform, enabling data sharing between multiple organizations (e.g., countries) that utilize a cloud framework for data storage. This includes exploring the intricacies involved in creating a cloud computing system for information sharing with multiple layers of information access rights within the NATO countries. For purposes of this thesis, information is scoped to be structured information, and is notated as a set of “datacubes.”

This thesis researches the information file structures that are appropriate for sharing between countries that have different levels of access rights to the information. The types of information file structures that can be shared include structured files, non-structured files, and semi-structured files. Files are stored in cloud environments either by Structured Query Language (SQL) database files, or NoSQL database files like “big table” and Accumulo (Accumulo). This thesis proposes a system architecture for information sharing between NATO countries, focusing on information represented in structured files.

As depicted in Figure 1, this thesis proposes a proof-of-concept solution to this problem, at the concept level, to have: (1) a technological platform for information (datacube) sharing across multiple information access levels; (2) an information file (datacube) structure for sorting different file types of multiple information access levels; and (3) a smart algorithm for automatically configuring the technological platform and information file formats to enable automated information sharing at different information access rights.

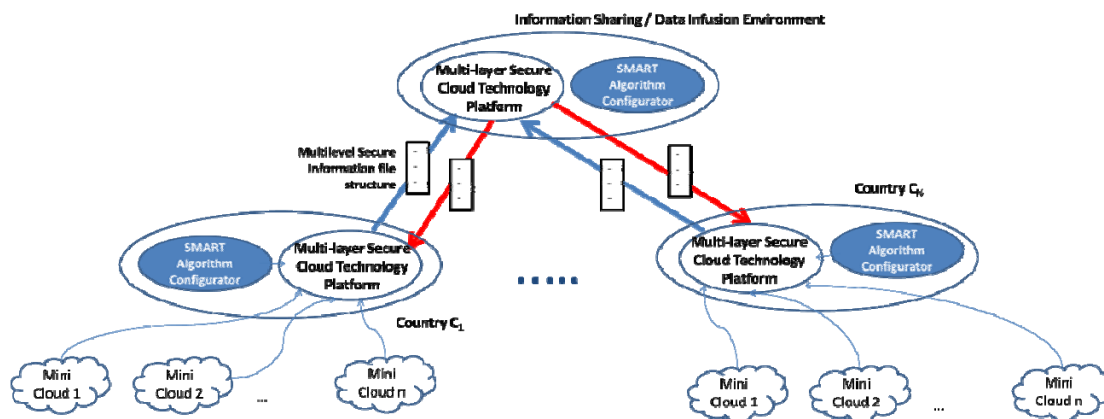


Figure 1. A concept diagram of the proposed multi-level secure datacube sharing system.

This thesis follows a systems engineering process to propose a systems architecture and an analysis of alternatives. Based on the previously mentioned concept, it instantiates part of the architecture with a proof-of-concept system in a laboratory environment. Finally, this thesis performs and documents the testing

and evaluation of such a proof-of-concept that performs multi-level secure datacube sharing. This proposed technological platform, and the automatic configuration of the technological platform and information file formats, would allow NATO members to access shared data faster in order to make decisions more quickly.

#### **D. BENEFIT OF STUDY**

Countries within the North Atlantic Treaty Organization (NATO) have a common goal requiring them to communicate, and often policies exist to implement information sharing (CJCSI 2013). However, they lack a technological platform and an automated method for fast information sharing, especially if the two countries have different access rights to the information. For example, in a peacekeeping operation, two countries may need to cooperate and work together to complete an operation. Quick information sharing is needed to execute a critical mission in a timely manner, as a lack of complete information could lead to poor decision making and loss of lives. In a hypothetical scenario, one country could have the resources, like manpower and tanks, due to its proximity in the region, and another country could be technologically advanced in its gathering of intelligence. The country with the physical resources needs to have the intelligence information from the other country in order to execute its mission well. The Indian Ocean Tsunami case study is a real case study where quick information sharing would have aided collaboration (R. Huber 2013). One type of structured information, or datacube, which is needed during natural disasters is the collection of social media information, along with structured alerts generated from such information (Social Media 2013). The proposed technological platform, and the automatic configuration of the technological platform and information file (datacube) formats in this thesis, allow NATO members to access shared data faster to make decisions more quickly.

The benefits from a systems engineering perspective are that this thesis not only implements a proof-of concept datacube sharing system, it provides a generalized system of systems architecture for any organization to implement using its own institutional physical architecture.

#### **E. ORIGINATION OF THESIS TOPIC FROM RADM TIGHE**

This thesis grew out of a serial project-based module in Systems Engineering that spanned three consecutive quarters. This serial project-based module began with system conceptualization in the first quarter when considerations were made with respect to the stakeholder capability, market, or opportunity need through early stage or conceptual design, and included systems thinking and project management. In the subsequent two quarters, the project was implemented with hands-on engineering experience applying systems engineering principles (Goshorn 2013).

The Smart Cloud System of Systems (SoS) project was chosen as the topic to complete this serial project-based module, and it subsequently spun off to a thesis with inspiration and guidance from Rear Adm. (RADM) Jan E. Tighe, president of the Naval Postgraduate School (NPS) at that time. RADM Jan E. Tighe spoke about information dominance and information warfare (IW) in a special arranged meeting (Ammon 2013). The session gave a glimpse of an insider's perspective of operations that was relevant to the course curriculum from a senior officer who has previously served as a leader in the Navy's information superiority realm.

RADM Tighe shared that in planning warfare, it is often assumed that the sensors and intelligence of the enemy are readily available and the emphasis is often on building bigger and more powerful platforms and systems instead. The author fully agrees with this view which sparked his desire to work on the smart cloud SoS. Information needs to be recognized as a form of 'lethal weapon' as well. In addition, along with information, resources like bandwidth are also often overlooked.

The project also involved a field experiment at Camp Roberts, California where there were several interactions with different stakeholders. There were discussions on policy with Mr. Al Moore and there were, discussions on Test and Evaluation with Mr. Richard Marchant. NORTHCOM was also present at Camp Roberts to discuss the importance of data alerts.

## **F. THESIS OUTLINE**

This thesis is divided into five chapters. Chapter I, Introduction, introduces the capability proposed in this thesis and the project that inspired it. Chapter II, Background and Literature Review, provides a background and literature review of the fundamental technologies used in this architecture. Chapter III, System Architecture for Multi-Level Secure Information Sharing, presents the capability's functional architecture and physical architecture. Chapter IV, Proof-of-Concept System, describes the physical proof-of-concept system of the capability, describes the experiment setup, and describes the results of the experiment. Finally, Chapter V concludes the thesis and suggests areas for future work.

## **II. BACKGROUND OF FUNDAMENTAL TECHNOLOGIES**

This chapter focuses on describing the background technologies used when determining the solution to the research questions. This includes first providing a background on a systems approach to automatically creating and categorizing data by intelligence level. Secondly, a background on smart cloud system of systems, the assumed infrastructure of organizations that desire to share data, is provided.

In addition to introducing the assumed cloud infrastructure and complementary approach to organizing data, this chapter introduces specific technical capabilities that are needed to propose a technological platform. First, for secure information hosting and sharing, a variety of virtualization solutions are needed and introduced in this chapter. Finally, background technologies behind “big data” storage, search, and replication used to handle different types of data are provided.

### **A. DETECT, IDENTIFY, PREDICT, REACT**

With the increasing amount of data to manage today, it is imperative to have an automated Artificial Intelligence (AI) systems solution (see Figure 2) to determine the data requirements, infrastructure capability, automation capability, and manpower available. Each level in the pyramid shown in Figure 2 represents a layer in the design, and they are all part of the overall AI systems solution. The foundation of the pyramid is the infrastructure capability, which is the hardware, software, and networking capabilities on which applications and data are hosted. The next layer up is the intelligence automation layer, which represents the suite of object behavior models for which the user wants an alert. The third level up represents the workflows of all of the behavior models that comprise an application workflow in order to collect a multitude of object behavior alerts. The second from the top level represents the mission areas, or operational scenarios. One mission may need several different application workflows in order to satisfy

the user. Finally, the top level represents the user who configures the entire AI systems solution, with respect to what alerts he or she wants to receive, given the threat at user requirements at hand (Goshorn 2010).

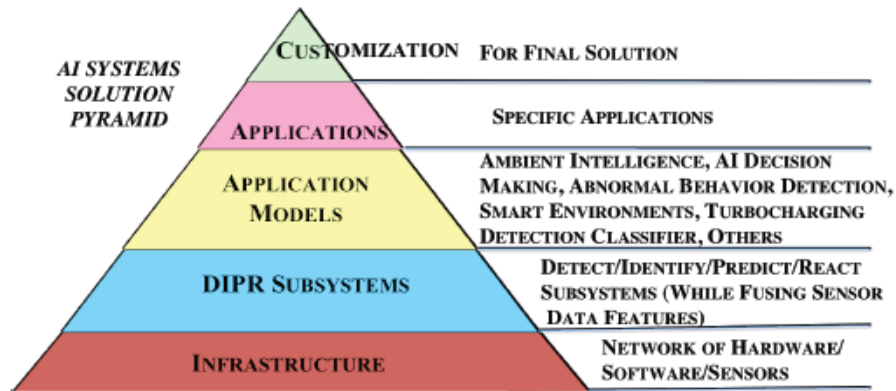


Figure 2. AI Systems Solution Pyramid (Goshorn 2010).

The rest of this section will describe the second level from the bottom of the AI Systems Solution Pyramid—the Detection, Identification, Prediction, and Reaction (DIPR) system. DIPR (see Figure 3) includes four stages of software modules, called analytics, which automatically analyze data to produce data at increasing levels of intelligence. AI and behavior modeling is implemented, together with object feature extraction and identification of intelligence states (Goshorn 2010).

First, raw data from sensors, such as cameras or social media collection tools, are input into the first subsystem, Detection. The Detection subsystem extracts useful information from the raw data, called features, and sends the features to the second subsystem, Identification. The Identification subsystem receives the processed information and outputs intelligent states (symbols). Intelligent states are made up of a logical expression of predefined features and characteristics. Within Identification it matches them with the extracted information and makes a conclusion of what the object is. Using these intelligent states, the Prediction subsystem predicts the behavioral outcome using a genetic algorithm. The Reaction subsystem then makes use of this predicted behavioral

outcome to come up with a set of responsive actions to manage the predicted behavioral outcome. Chapter III discusses how the DIPR can be used to enable multi-level secure information sharing (Goshorn 2010).

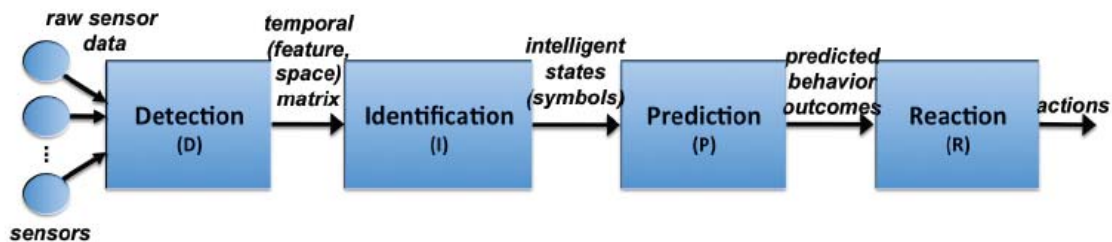


Figure 3. DIPR as the foundation of AI and Object Behavior Modeling (Goshorn 2010).

## B. SMART CLOUD SYSTEM OF SYSTEMS

The smart cloud system of systems contains a main cloud, several smart sensor mini clouds that could be sensors, such as cameras, an intelligence automation application/software, and translator software that standardizes the metadata outputs. Figure 4 is an example of a smart cloud system of systems. The main cloud performs the smart data fusion of the data gathered from the mini clouds. It also performs the distributed processing of data, storage of data, advanced intelligence, and search functions. Finally, the command and control mini cloud functions as the node that enables operators and analysts to both configure the smart cloud system of systems as well as visualize alerts pushed to them from the main cloud (Goshorn 2010).

Specifically, the mini clouds represented in Figure 4 consist of a command and control mini cloud, a social media mini cloud, a narrow field-of-view mini cloud, and a wide field-of-view mini cloud. The command and control mini cloud configures the smart cloud system of systems. It is alerted when new data is available, and accesses the data in the main cloud. The social media smart sensor mini cloud gathers data from social networks like Twitter and Facebook. The narrow field of view mini cloud collects detailed intelligence of an object or

target. The wide field-of-view mini cloud provides a big picture intelligence view of a situation or place. The data from the social media mini cloud, narrow field-of-view mini cloud, and wide field-of-view mini cloud, is sent to the main cloud for processing (Goshorn 2013).

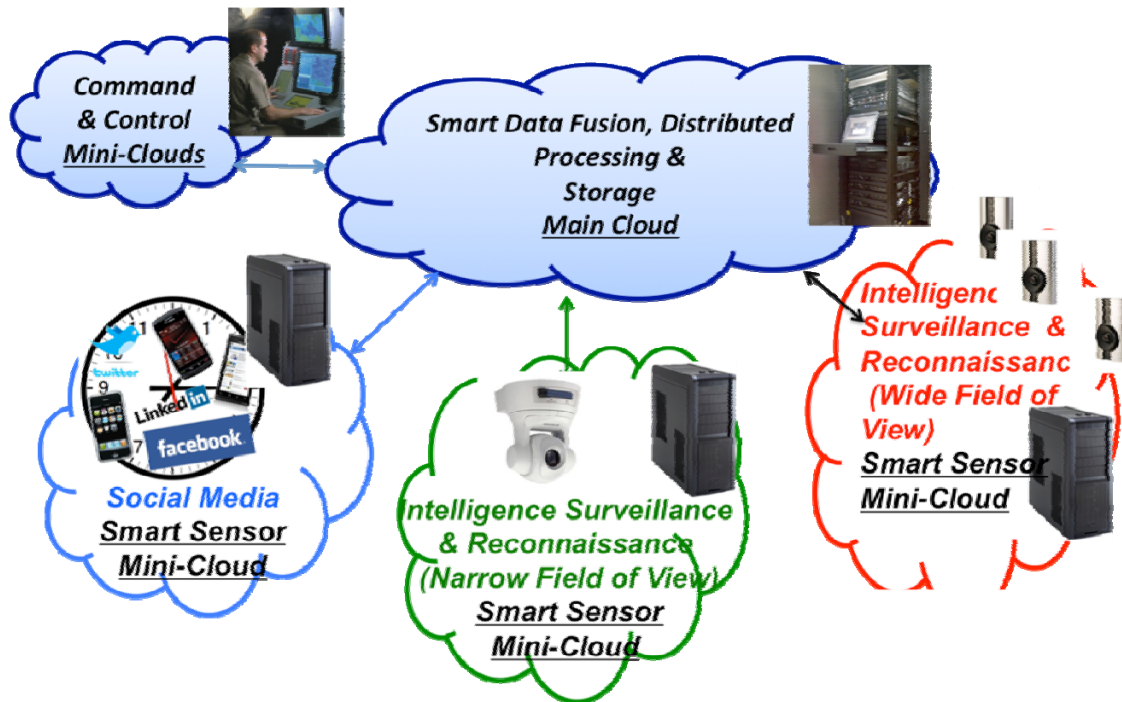


Figure 4. Smart cloud system of systems (Goshorn 2013).

This thesis focuses on how to design the main cloud to enable smart-cloud-to-smart-cloud information sharing. Chapter III discusses the new proposed design of the main cloud in order to enable multi-level secure information sharing. The rest of this chapter describes background technologies needed for secure information sharing between main clouds.

### C. VIRTUALIZATION ACCESS

One technology that aids in secure information sharing is virtualization. Virtualization is the concept of creating an imaginary version of a computer system, operating system, storage device, or network resource (Wikipedia

2014d). Virtualization of hardware and networking are both fundamental technologies in order to have a secure information-sharing environment. This section provides a brief background on a commercial implementation of virtual hardware (VMWare), along with a commercial implementation of virtual networking (VPN), such as OpenVPN.

## 1. VMWare Virtualization

VMware is a product that enables simplified infrastructure as a service (IAAS), which is a service that allows users to have infrastructure on demand. Figure 5 describes the components. It is made up of server software and client software. Multiple people can simultaneously access multiple virtual computers within the VMWare server software by having the VMWare client software on their own physical computers (VMWare 2008, 2010).

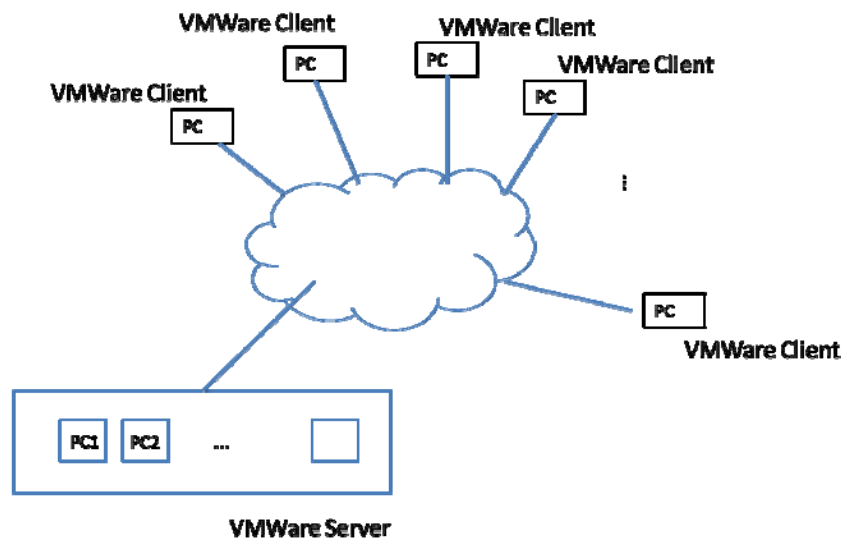


Figure 5. Simple Diagram of a VMWare server providing virtual PCs to multiple users who have a VMWare client.

Not only do hardware virtualization technologies, such as VMWare, reduce capital expenses and operating expenses, they minimize downtime of infrastructure. For example, if a virtual computer goes down functionally, a new virtual computer with the same information of the affected computer can become

functional on demand. VMWare saves operating resources, because there is less hardware involved. Finally, by providing external users to an organization's selected internal virtual computers, on which predefined data is stored, an organization can provide external users access to information without giving them the option to download the data onto their own physical computers (VMWare 2008, 2010).

## **2. Virtual Private Networks**

In order to connect a computer to an organization's private network, it is necessary to have a virtualization capability to make it seem like that computer is connected to the organization's private network, when physically it is not. This technology that provides this capability is called a virtual private network (VPN). In other words, VPNs help to put a private network across a public network. They enable computers that are in the private network to share and receive data over the public network, while maintaining the network security of the private network. This is achieved by establishing a virtual point-to-point connection using dedicated connections between the VPN client and VPN server, and by using encryption technology, as shown in Figure 6 (Wikipedia 2014b, Wikipedia 2014c).

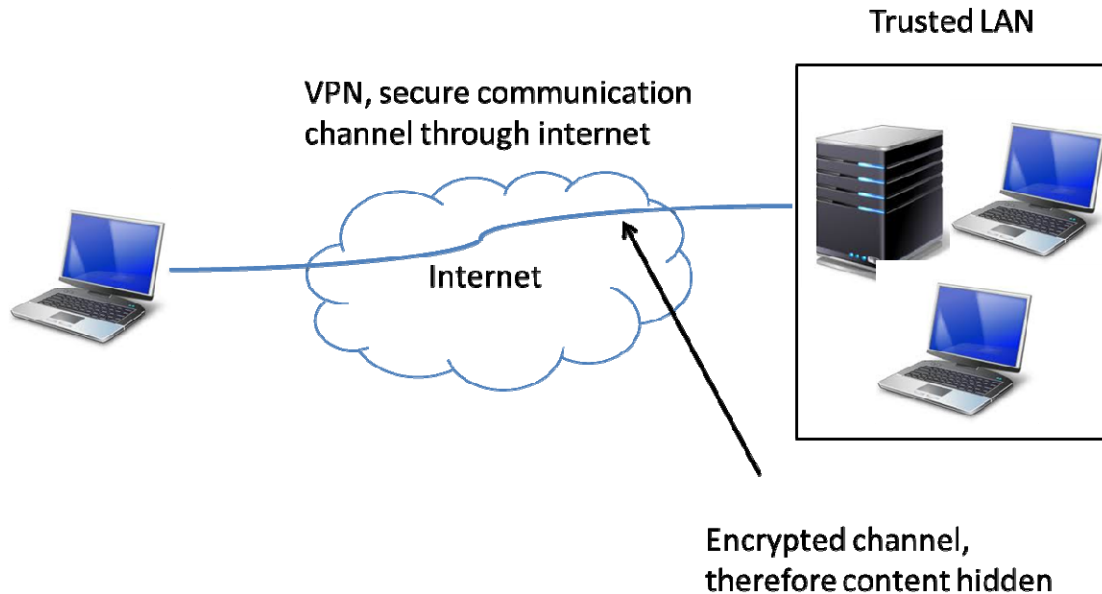


Figure 6. VPN setup that enables a computer to interface with a trusted Local Area Network (LAN).

Chapter III discusses how the VPN can be implemented to enable multi-level secure information sharing.

#### D. “BIG DATA” STORAGE, SEARCH, AND REPLICATION TECHNOLOGY

In addition to the ability to receive and transmit data, there are three main functions within the main cloud systems: storage, search, and replication. There are several technologies that enable storage, search, and replication within a cloud environment.

This section discusses the more popular open source solutions: Apache Accumulo, MongoDB, and Apache SOLR.

##### 1. Apache Accumulo

Apache Accumulo is a “big data” database system that is referred to as “NoSQL,” or “not only SQL.” It was designed based on technology that Google uses to store, replicate, and query against. In particular, Accumulo has added security functionality, as it was designed by the National Security Agency in 2008.

Accumulo is able to handle large amounts of structured, semi-structured, and unstructured data, while simultaneously enabling one to employ security protection measures on the data that it stores (Accumulo 2012, Accumulo 2014, Apache 2013).

## **2. MongoDB**

MongoDB is a “big data” document-oriented database system that is also a NoSQL solution that allows searching of the database by field, range queries, or regular expression searches. It allows for data files with varying structure formats to be stored together and queried together, even with different structured formats (Wikipedia. 2014a, MongoDB Inc. 2014).

## **3. Apache SOLR with Lucene Search Engine**

Apache SOLR is another document-oriented-database system that uses the Lucene search engine in order to search both structured and unstructured documents (Apache 2012).

Chapter III explores these technologies further and how they can be implemented to enable multi-level secure information sharing. Additionally, an analysis of alternatives is presented in Chapter III regarding which platform instantiates the main cloud’s storage, replication, and search functionalities.

### **III. SYSTEM ARCHITECTURE FOR MULTI-LEVEL SECURE INFORMATION SHARING**

This chapter describes a proposed systems architecture that could be used to implement a multi-level, secure methodology for how datacubes with different information security levels can be shared between countries on a technological cloud platform. The intricacies involved in creating a cloud computing system for information sharing with multiple layers of information access rights, within NATO countries, will be explored.

The types of information file structures that can be shared include structured files, non-structured files, SQL database files, No SQL database files like “big table,” and Accumulo. This thesis proposes a system architecture for information sharing between NATO countries, focusing on information represented in structured files.

#### **A. SCENARIOS**

Countries within NATO often have a common goal requiring them to communicate, and often policies exist to implement information sharing (CJCSI 2013). However, they lack a technological platform and an automated method for fast information sharing, especially if the countries have different access rights to the information. One common goal for information sharing between allied countries is the need to support each other in operations, such as Humanitarian Assistance / Disaster Relief operations. Figure 7 depicts an operational scenario of one country’s potential capability to sense the environment and generate alerts, which are important information with whom another allied country may want to have shared.

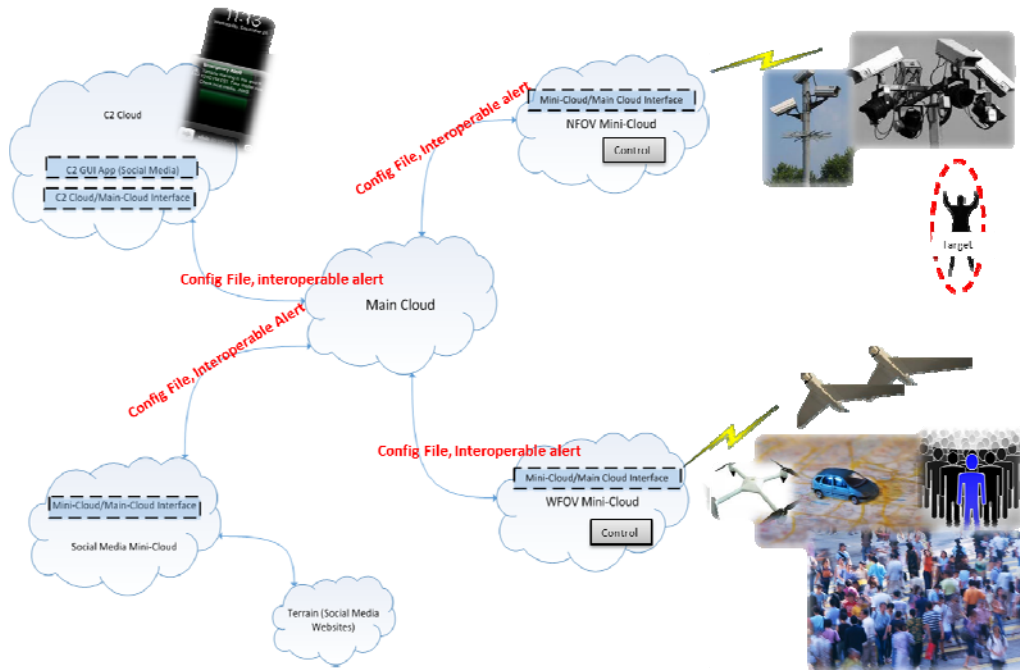


Figure 7. Information generation in humanitarian and disaster relief assistance.

## 1. Philippines Typhoon Disaster

A recent example of information sharing is the typhoon disaster in the Philippines (see Figure 8). The Philippines needed to share information with the countries coming to aid its disaster relief efforts. The government had information through various sources, such as social media, cameras, and the man on the ground, from the area that required aid. The other countries had resources like manpower, food, and intelligence gathering capabilities that could aid in the Philippine relief efforts. These countries needed to collaborate to plan and execute the relief plan as quickly as possible.

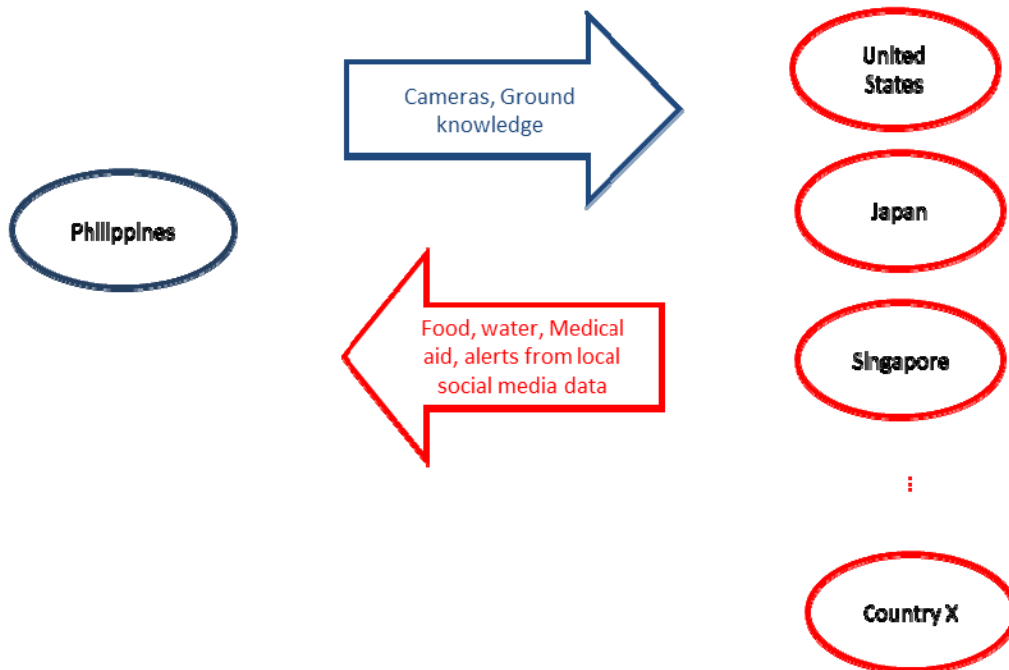


Figure 8. Information sharing scenario.

One type of structured information, or datacube, needed during natural disasters is collected social media information, along with structured alerts generated from such information (Social Media 2013). The proposed technological platform, and the automatic configuration of the technological platform and information file (datacube) formats in this thesis, allow NATO members to access shared data faster in order to make decisions more quickly. This thesis not only implements a proof-of-concept datacube sharing system, it provides a generalized system of systems architecture for anyone to implement using their own institutional physical architecture.

## 2. Other Scenarios

In a peacekeeping operation, two or more countries may need to cooperate to complete an operation. Quick information sharing is needed to execute a critical mission timely as a lack of complete information could lead to poor decision making and loss of lives. In a hypothetical scenario, one country could offer resources like manpower and tanks, due to its proximity in the region.

Another country could be technologically advanced in its gathering of intelligence. The country with the physical resources needs to have the intelligence information from the other country in order to execute its mission well. Such cooperation illustrates one positive way in which information is shared.

This technology can also be extended for use in joint service operations where the Air Force, Navy, and Army are required to work together to complete a mission.

This thesis looks at developing a solution that can allow information sharing among a multiple number of entities, and among entities with different existing information storage systems. This thesis will also look at sharing data of different security classifications (Secret, Confidential, Unclassified), and different file types (structured including JSON, non-structured, SQL, non-SQL). It will also look at the option of sharing data among entities for a specified amount of time, and allowing the shared data to be downloaded and fused with the information from different entities, to allow deeper analysis to be performed for further insights.

## **B. FUNCTIONAL ARCHITECTURE**

This section provides the functional architecture for the proposed capability, which is first the functional decomposition, followed by the functional hierarchy.

### **1. External Systems Diagram/IDEF0 Level 1**

The first functional diagram, Figure 9, acts as both the external systems diagram and the level 1 IDEF0 diagram. The External Systems Diagram is combined with the Level 1 IDEF0 Functional Decomposition for the proposed Provide Multi-Level Secure Information Sharing System SoS (see Figure 9) for clarity, as the proposed system is a complex SoS. Combining the External Systems Diagram and Level 1 IDEF0 Functional Decomposition illustrates how

each external system interfaces with different high-level internal systems of the proposed system.

There are five external systems: Country Policy Department B1, Smart Cloud SoS B2, Country Operator B3, Information Group Operator B4, and New Information Group Operator B5.

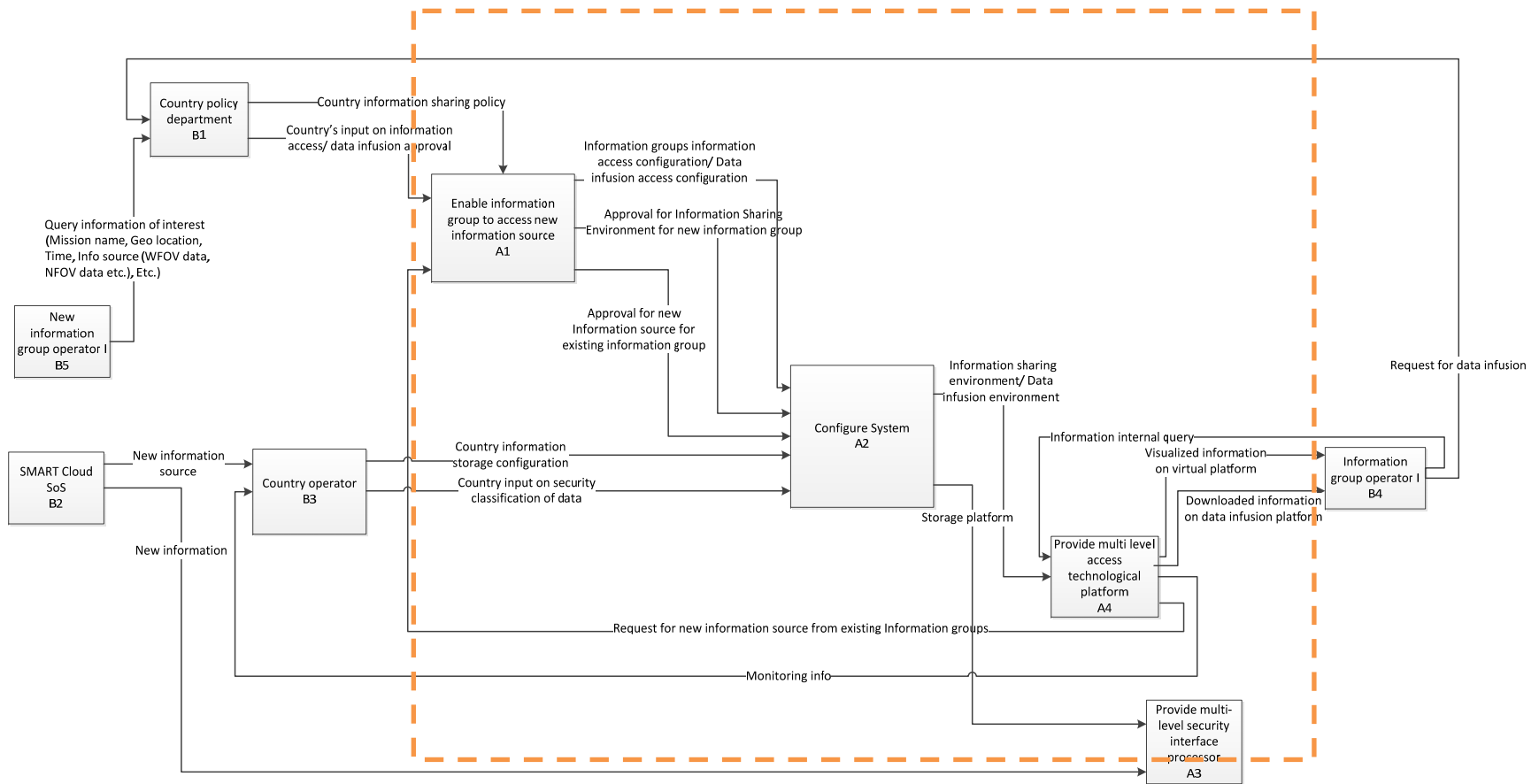


Figure 9. External systems diagram/ IDEF0 Level 1.

The Country Policy Department B1 sets the home country's information sharing policy to manage the information access privileges of the information groups. The Country Policy Department B1 also decides if its country data can be downloaded by another country for further analysis. The Smart Cloud SoS B2 is a group of sensors and cameras in which raw data is being captured and sent to the Country Operator B3 for processing. The Country Operator B3 receives new raw data from the home country's information source, and decides on the security classification of the new data. This subsystem also decides on the storage configuration in which the data should be stored.

The Information Group Operator B4 is representative of the group of countries that is requesting information sharing from the home country. This operator receives the information that is being shared by the home country via a virtualized platform. If the home country allows the data shared to be downloaded by the information group for further analysis, it would be done through a data infusion platform. A data infusion platform is a server that would contain the data from the home country and the countries in the information group. The data from these multiple sources are infused so that deeper analysis can be done for further insights. Data in this data infusion platform would not be allowed to be downloaded out of this server. A monitoring report of the data traffic in the data infusion platform would be sent to the home country daily. The Information Group Operator B4 can request data to be put on this data infusion platform, if required. The Information Group Operator B4 can also query on the information being shared and make further requests to the Country Policy Department B1 for other information to be shared if the current shared information is insufficient.

The New Information Group Operator B5 is a new group of countries that is requesting for information from the home country to be shared for the first time. New configurations need to be setup in terms of information sharing policy, computer hardware, and software to enable this sharing. The New Information Group Operator B5 makes this request to the Country Policy Department B1.

The four high-level internal systems that interact with the five external systems are: Enable Information Group To Access New Information Source A1, Configure System A2, Provide Multi-Level Security Interface Processor A3, and Provide Multi-Level Access Technological Platform A4.

Enable Information Group To Access New Information Source A1 receives input from the Country policy department B1 on the access privileges of the various Information groups. This includes the security level of the data to be shared, the type of data to be shared, and also whether the data to be shared can be downloaded to the data infusion platform for further analysis. Enable Information Group To Access New Information Source A1 then gives the configuration settings for these security inputs to Configure System A2. Enable Information Group To Access New Information Source A1 would also receive request for existing Information groups for additional information to be shared. Based on the inputs of the country information sharing policy from the Country policy department B1, Enable Information Group To Access New Information Source A1 would decide if the request can be acceded to and it gives these inputs to Configure System A2.

Configure System A2 is the brain of this architecture system. It receives input from Enable Information Group To Access New Information Source A1 to make the security access configuration for its data. It then creates an information sharing environment or a data infusion environment for the information group. Configure System A2 also receives input from the New Information Group Operator B5 on the security classification of the data that it receives and the storage configuration on how to store its data. It then creates a storage platform for all of the home country's data.

Provide Multi-Level Security Interface Processor A3 receives new data from Smart Cloud SoS B2 and the storage configuration from Configure System A2 to provide the system with a multi-level security storage solution to store its data.

Provide Multi-Level Access Technological Platform A4 uses the information sharing environment or data infusion environment from Configure System A2 to provide a technological platform for multiple security level secure information sharing. Using a virtualized platform, it provides the information group with visualized information on a virtual platform or downloaded information on a data infusion platform based on the information group's access privileges. Provide Multi-Level Access Technological Platform A4 also allows the existing information group to request more information from Enable Information Group To Access New Information Source A1. It sends the data transfer monitoring log in the information sharing, or data infusion environment platform, to the home Country Operator B3.

## 2. Functional Hierarchy

The functional hierarchy of the Provide Multi-Level Secure Information sharing system A0 (see Figure 10) has four subsystems: Configure System A1, Provide Multi-Level Security Interface Processor and Storage A2, Enable Information Group To Access New Information A3, and Provide Secure Information Sharing Environment A4. Each of these subsystems will be described further in the section on IDEF0.

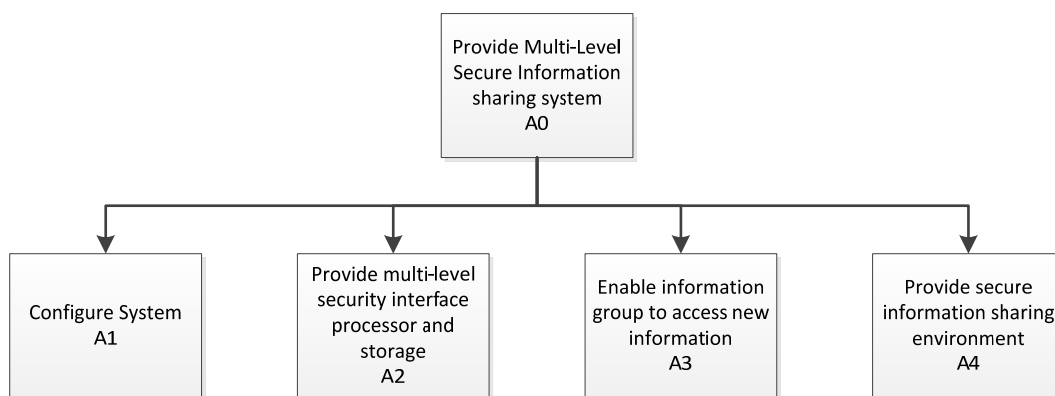


Figure 10. Provide multi-level secure information sharing system A0.

Configure System A1 is where all system configurations are performed. Configure System A1 is further decomposed to Configure Information Sharing Environment/Data Infusion Environment For Information Groups A1.1, and Configure Data Storage System A1.2, as shown in Figure 11. Each of these subsystems will be described further in the section on IDEF0.

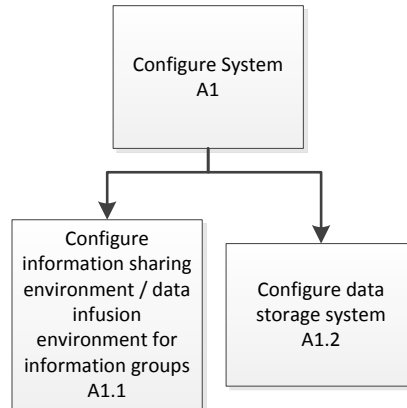


Figure 11. Decomposition of configure system Function A1.

Provide Multi-Level Security Interface Processor and Storage A2 provides a technological platform for multiple security-level secure information sharing. Provide Multi-Level Security Interface Processor and Storage A2 is further decomposed to Configure interface (buffer) A2.1, Receive raw data (buffer) A2.2, Sort raw data A2.3, Create multi-level security file versions A2.4, Route files to appropriate security classification location A2.5, and Store multi-level security data A2.6, as shown in Figure 12. Each of these subsystems will be described further in the section on IDEF0.

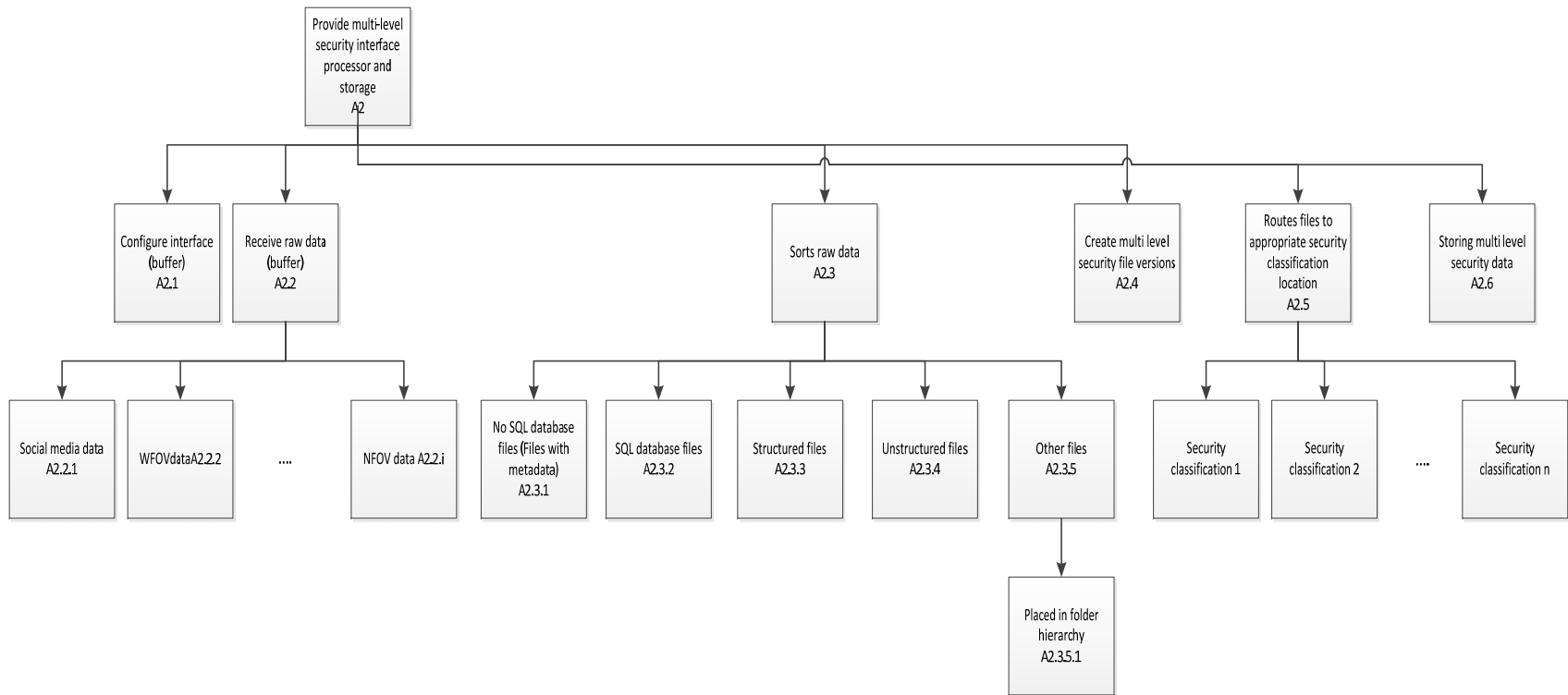


Figure 12. Decomposition of provide multi-level security interface processor and storage function A2.

Enable Information Group To Access New Information A3 handles and gives access to the information group on the data the home country can share. Enable Information Group To Access New Information A3 is further decomposed to Enable information group to describe information of interest A3.1, Search process to find related information A3.2, Manage information shared A3.3, and Share new information A3.4 subsystems, as shown in Figure 13. Each of these subsystems will be described further in the section on IDEF0.

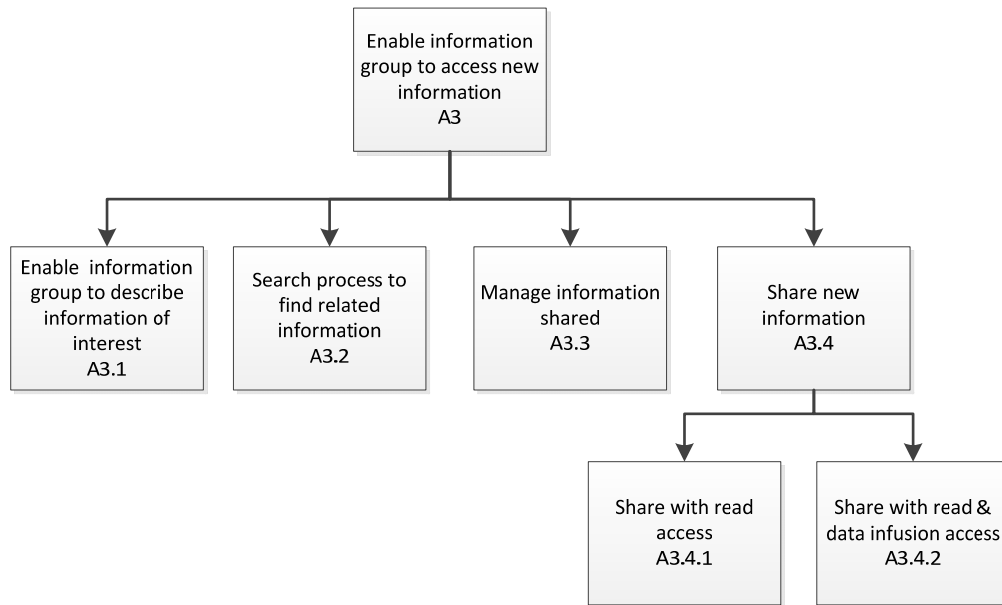


Figure 13. Decomposition of enable information group to access new information function A3.

Provide Secure Information Sharing/Data Infusion Environment A4 ensures that the sharing of information between countries is done in a secure manner. Provide Secure Information Sharing/Data Infusion Environment A4 is further decomposed to VPN connection to country's main cloud A4.1 and Provide information sharing/ data infusion environment for information group 1 A4.2, as shown in Figure 14. Each of these subsystems will be described further in the section on IDEF0.

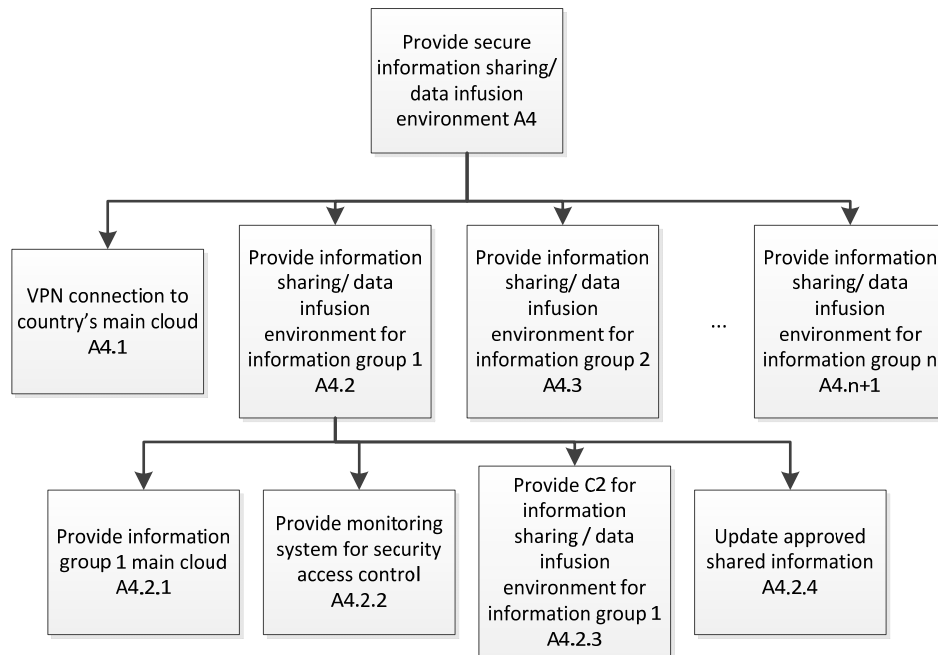


Figure 14. Decomposition of provide secure information sharing environment function A4.

### 3. IDEF0 Functional Decompositions

Configure System A1 is where the system configuration is performed. Configuration of the system includes data storage configuration, data security level configuration, and Information Group Security access level configuration. Configure Information Sharing Environment/Data Infusion Environment For Information Groups A1.1 configures the information sharing environment or the data infusion environment for the Information Group. Configure Data Storage System A1.2 configures the storage method to store the home country's data. Figure 15 shows the process flow of these subsystems.

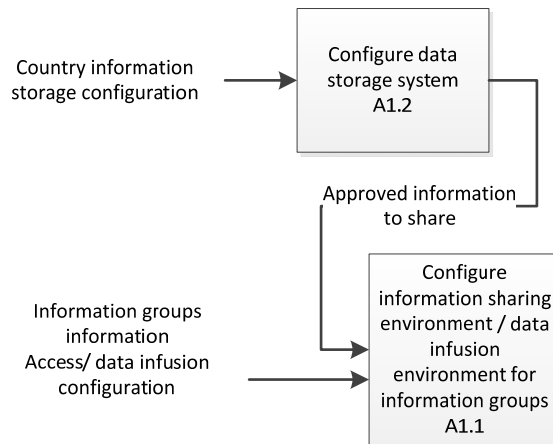


Figure 15. Process flow of functions for A1.

The home country sends the configuration of its preferred storage method to Configure Data Storage System A1.2, which then configures the storage server. Using the information group information access/data infusion configuration, Configure information sharing environment/data infusion environment for information groups A1.1 configures the access rights for the information group, and extracts the data to be shared from the storage server (see Figure 15).

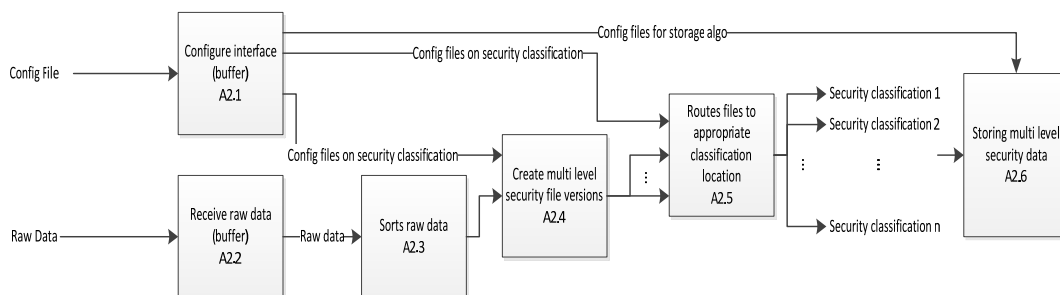


Figure 16. Process flow of functions for A2.

Provide Multi-Level Security Interface Processor and Storage A2 provides a technological platform for multiple security-level secure information sharing (see Figure 16). Configure interface (buffer) A2.1 is a buffer for the configuration settings to be integrated into the server. Receive raw data (buffer) A2.2 is another buffer for the incoming raw data as it is saved on the server. Sorts Raw

data A2.3 sorts the raw data to the server with the appropriate security classification, and into the designated location on the server. Create multi-level security file versions A2.4 creates multiple versions of the same document with each document containing only data with the appropriate security classification for storage in the different security classification levels.

When new raw data is available, classification of the security level of the data is done, and multi-level security file versions are created and routed to their appropriate classification location. The different file types are handled differently when creating the multi-level security file versions by the Smart algorithm. It is recommended that every metadata file contain the following information: Mission name, Geographical location, Time, Information source (social media data, Wide field-of-view (WFOV) data, and Near field-of-view (NFOV) data to facilitate the Smart algorithm in sorting out the raw data. The data is then stored following a predefined storage method. Route files to appropriate security classification location A2.5 sends the files to the appropriate server location. Storing multi-level security data A2.6 stores the data using the configuration for the storage algorithm.

Enable Information Group To Access New Information A3 handles and gives access to the information group on the data the home country can share (see Figure 17). Enable new information group to describe information of interest A3.1 allows the information group to describe the information of interest. This input is sent to the search engine in Search process to find related information A3.2 to find related information in the home country's data base. The search results is sent to Manage information shared A3.3. Based on the information group access privileges, Manage information shared A3.3 outputs the data that is allowed to be shared to the information group. Provide technology platform of information group n A4.n+1 provides the platform in which this data can be shared in a secure manner.

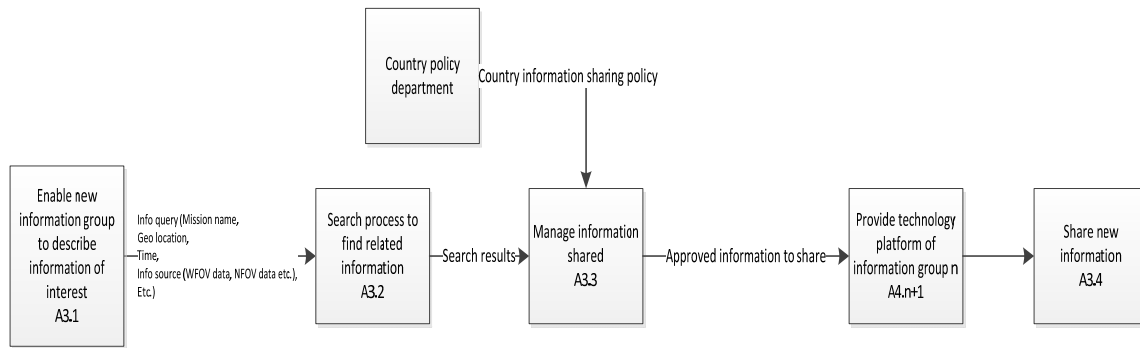


Figure 17. Process flow of functions for A3.

Provide Secure Information Sharing/Data Infusion Environment A4 ensures that the sharing of information between countries is done in a secure manner (see Figure 18). The information group logs into the home country's VPN as a virtual client, and is then able to access the information-sharing environment of his information group. If the information group has data infusion privileges, it would be able to infuse the data with the home country's data via the data infusion environment for further analysis.

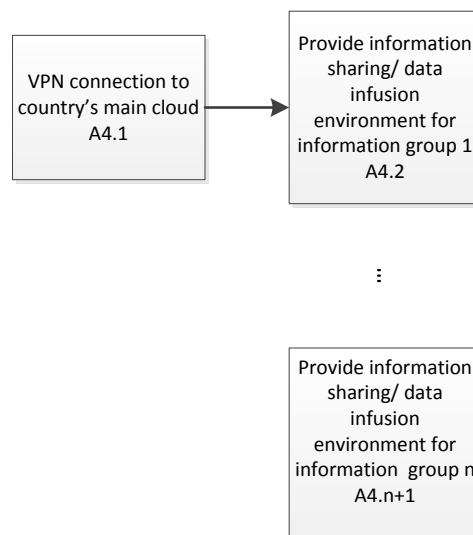


Figure 18. Process flow of functions for A4.

Provide information group 1 main cloud contains all the home country's data in which information group 1 can access (see Figure 19). Provide C2 for information sharing/data infusion environment for information group 1 A4.1.3 allows the information group to see the data on a virtual platform. The information group can send its queries on the data to Provide C2 for information sharing/data infusion environment for information group 1 A4.1.3 who would then send the query to Provide information group 1 main cloud A4.1.1. Provide information group 1 main cloud A4.1.1 would send the information request to the home country via Provide monitoring system for security access control A4.1.2. The home country would decide whether to give access to the requested query. Any updates to the information group access privileges would be sent to Provide information group 1 main cloud A4.1.1 via Update approved shared information A4.2.4. The data access log file is captured by Provide monitoring system for security access control A4.1.2 and sent to the home country to monitor for any irregular activities.

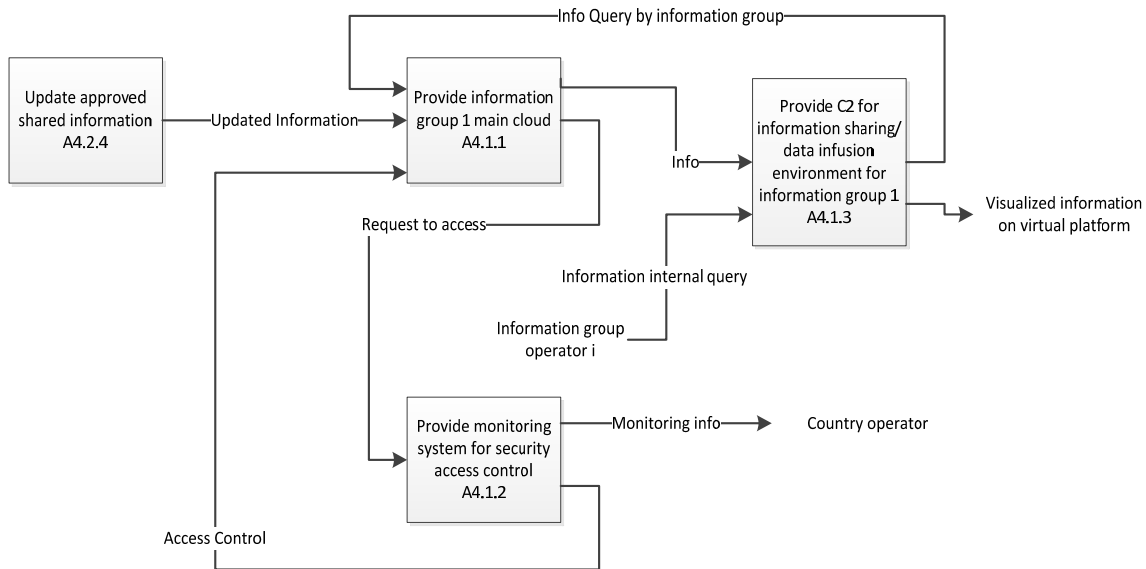


Figure 19. Process flow of Level 3 functions for A4

## C. PHYSICAL ARCHITECTURE

### 1. Overview

Figure 20 shows an overview of how the data from the home country is shared with another country. Sensor or raw data is gathered from the mini clouds to the technological platform of the home country. Using a multilayer secure technology platform, data is shared via an information sharing or a data infusion environment.

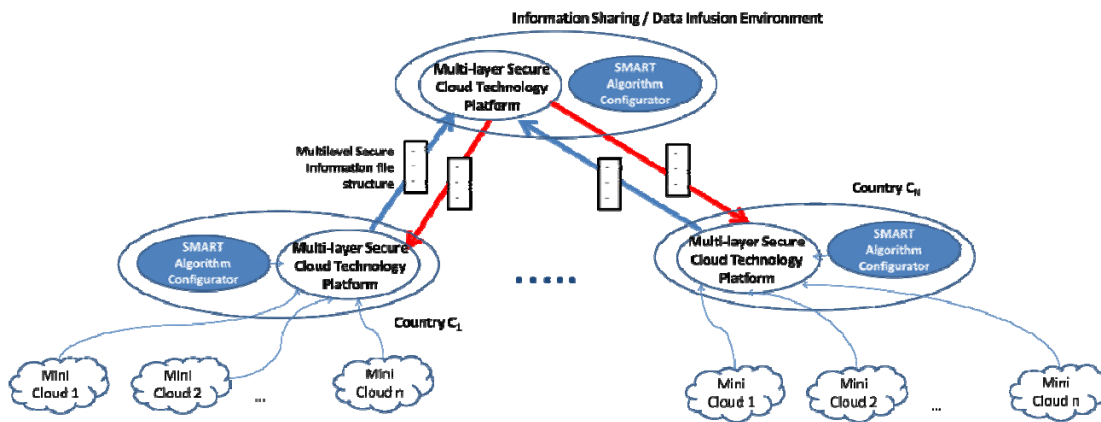


Figure 20. Concept diagram of the physical architecture of information sharing by the home country with the information group.

This automated process is configured when the home country operator obtains the requirements from the country policy department, and inputs them into an interface that auto-generates the configuration file being sent to the technological platform (see Figure 21).

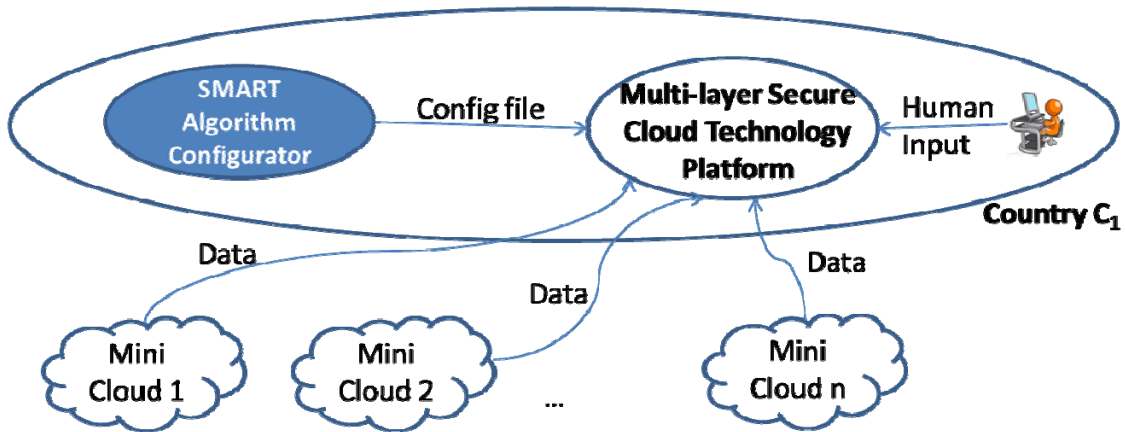


Figure 21. Concept diagram of the physical architecture of the home country.

## 2. Technological Platform

The physical architecture in Figure 22 shows the hardware required to enable the architecture to work. A country would have various sources from which to gather their information, and uses three sources of information: Social Media information, WFOV camera information, and NFOV camera information. The information from these various sources is gathered on a server where it will be further routed into separate servers with the appropriate security classification. Multiple versions of the same document could be created, with each document containing only data with the appropriate security classification for routing to the different servers.

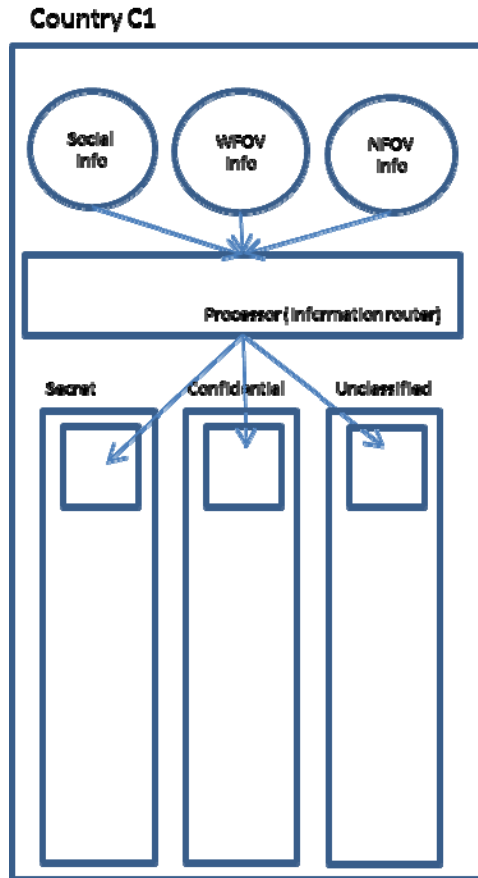


Figure 22. Physical of technological platform and smart algorithm configurator for a country.

After the data that an information group can access has been decided, this set of data would be copied into a third server (see Figure 23). This other server has a main cloud containing all the data, a monitoring cloud that monitors the data flow in and out of this server, and a C2 user interface for the information group to access the data in this server.

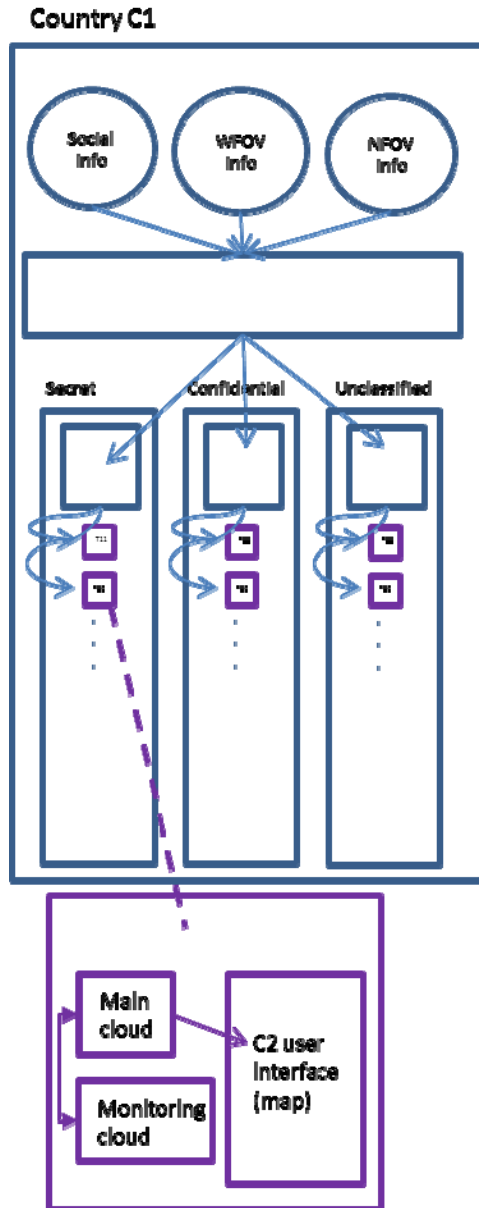


Figure 23. Physical architecture of a country with servers for sharing information with information group

The information group would access the shared information by Country C1 via VPN, as shown in Figure 24.

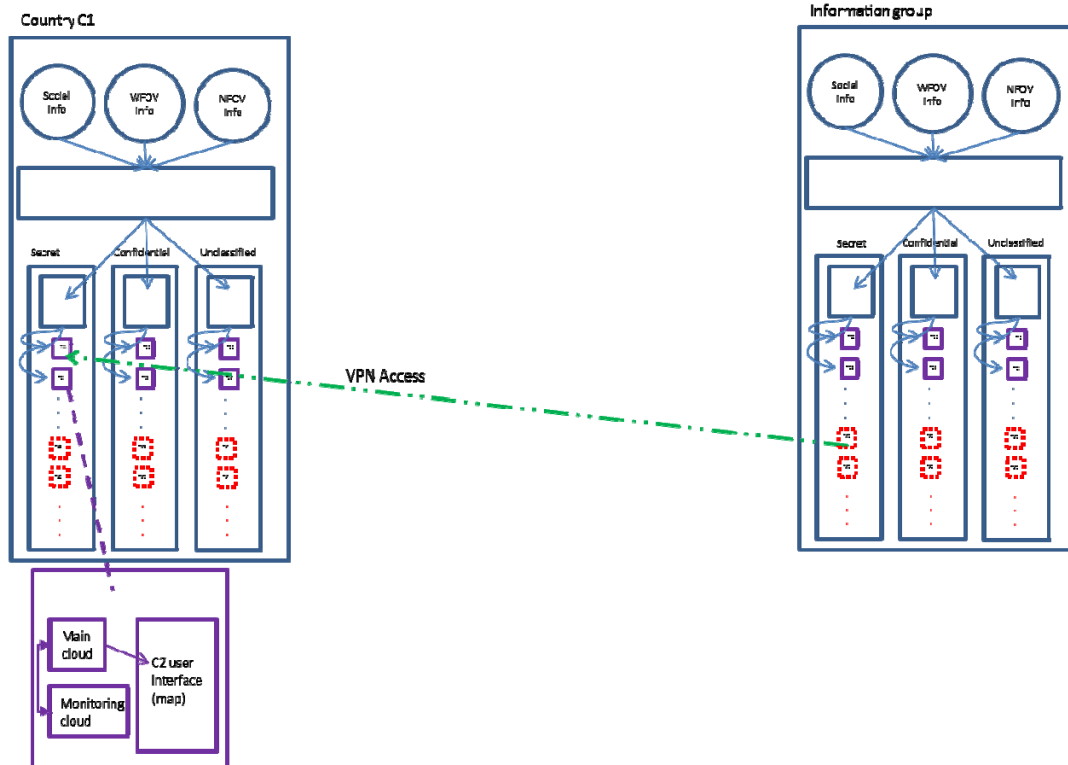


Figure 24. Physical architecture of a country sharing data with an information group via VPN.

If information from country C1 needs to be merged with information from the information group for further analysis, the data would be downloaded to a fourth server managed by the Information Sharing Authority (see Figure 25). This fourth server is also called the data infusion platform where data from various sources are infused so that deeper analysis can be done for further insights. Data in this data infusion platform would not be allowed to be downloaded out of this server.

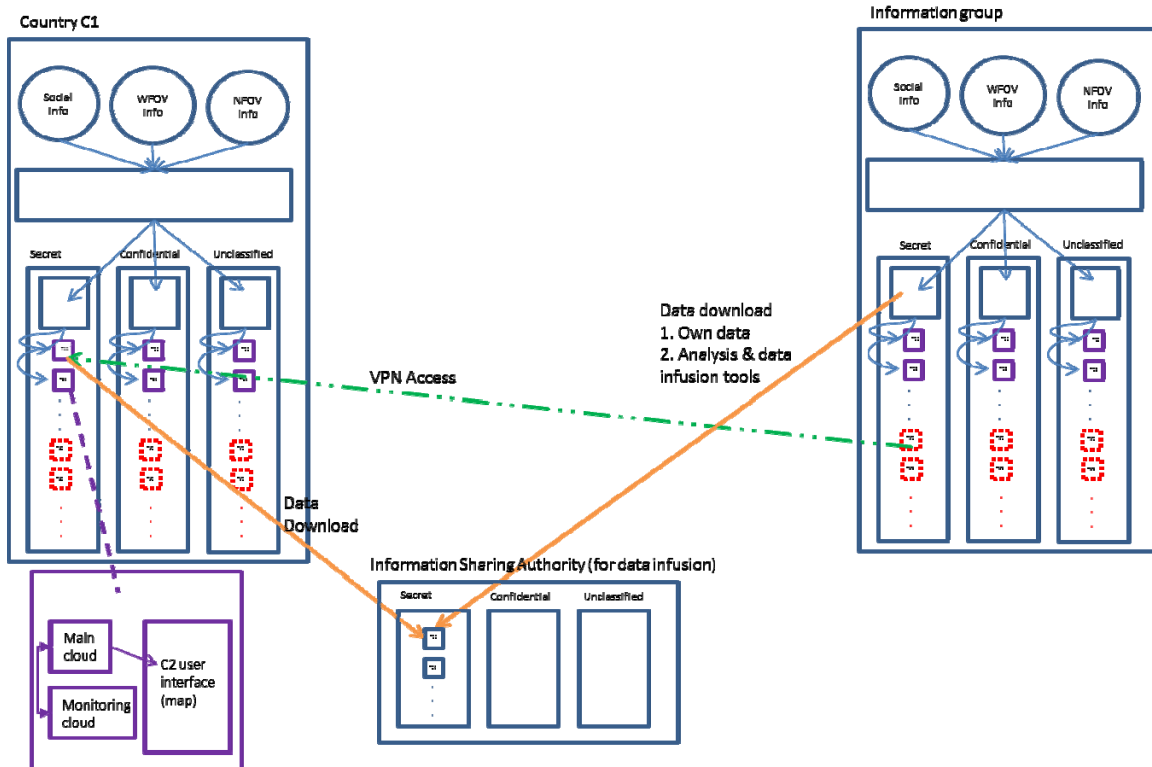


Figure 25. Physical architecture of a country sharing data via data infusion platform.

Similarly, this fourth server would have a main cloud containing all the data, a monitoring cloud that monitors the data flow in and out of this server, and a C2 user interface for the information group to access the data in this server (see Figure 26). A monitoring report of the data traffic in the data infusion platform would be sent to the Country C1 daily. The information group would access the data in this server through a VPN as well.

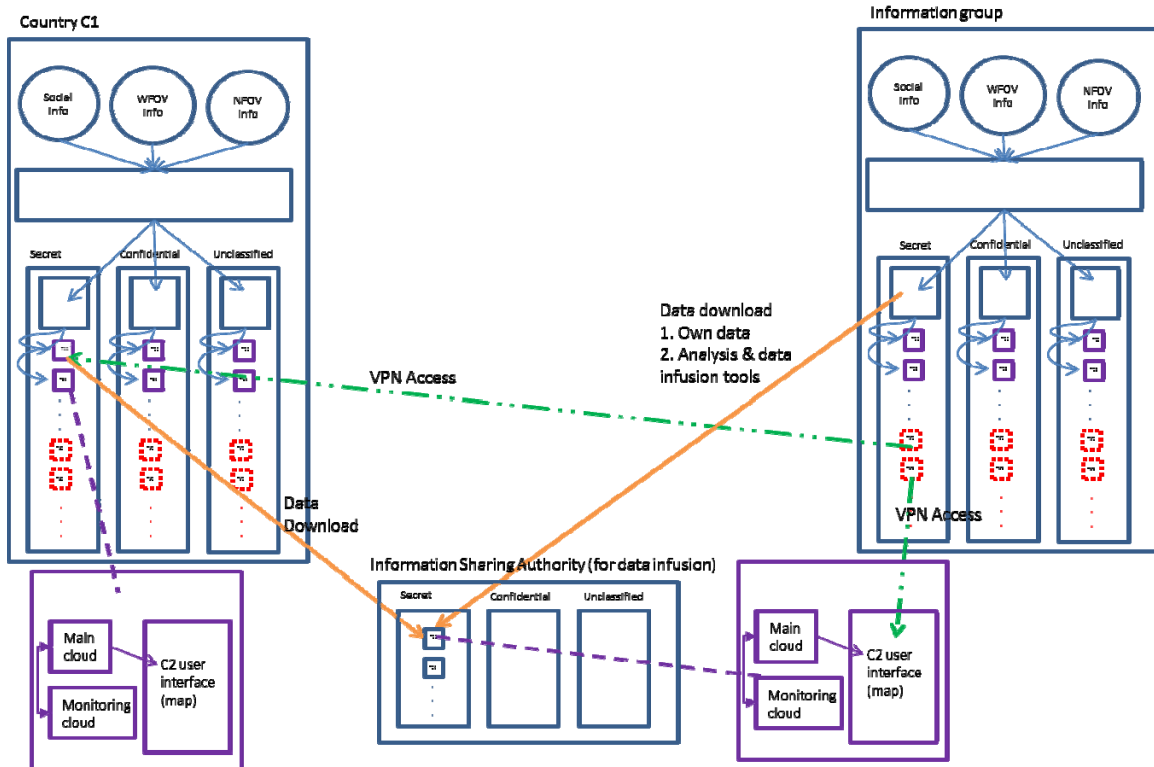


Figure 26. Full physical architecture of a country sharing data.

Initially, the architecture contemplated having the countries share data without an information sharing authority if they trusted each other. Data could be shared via VPN in this case. However, it was concluded not to have this option, as there may be political ramifications or consequences.

### 3. Smart Algorithm Configurator Physical Architecture

Figure 27 shows a broad overview of the sharing of information among countries. Every country makes use of the information sharing authority to share its data and obtain data shared by other countries.

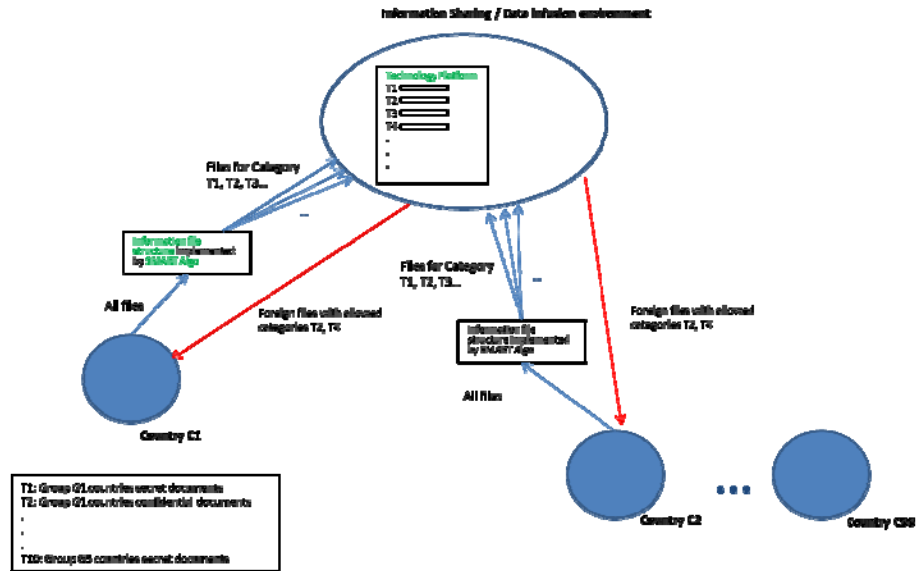


Figure 27. Overview of countries sharing data using Information Sharing Authority.

Figure 28 shows how the different data types are managed by the Smart algorithm in routing the files to their appropriate servers. Files that are in the folders with designated security classification, files with metadata, SQL database files, structured files, can be automatically sorted to their appropriate categories using the Smart algorithm. Unstructured files have to be manually sorted.

It is recommended that the metadata file should contain the following information: mission name, geographical location, time, information source, intelligence level so that the file can be uniquely differentiated. The information source can be from any sensor raw data source like the social media data, WFOV data, or NFOV data, etc. The intelligence level is classified as follows: the first level is raw or sensor data, the second level is detection alerts, fused features, predictions, or recommended reactions.

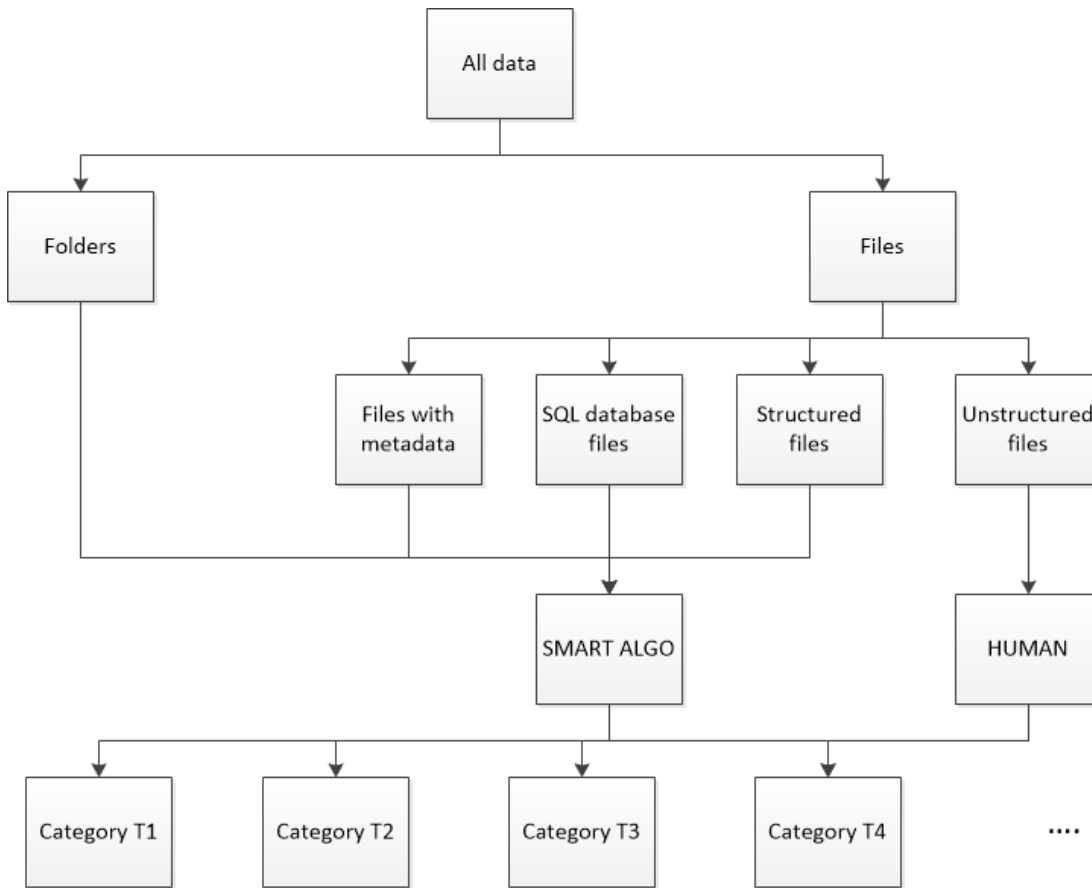


Figure 28. Data Management using smart algorithm.

#### D. ANALYSIS OF ALTERNATIVES FOR PHYSICAL ARCHITECTURE

A second important group of the technological platform would be to decide on the methods to store, query, and replicate the data. This analysis will not be provided in this thesis, but a brief comparison of the methods can be found in Table 1.

	<b>Ability to do Multi level security</b>	<b>Ability to Scale</b>	<b>Types of files</b>	<b>Remarks</b>
Accumulo	Yes	Yes	Json files No SQL database files	Data is stored in a structured manner
Mongo DB	No	Yes	Json files No SQL database files	Data is stored in a structured manner
SQL			SQL Databases files	Relational databases Data is stored in a structured manner
Apache SOLR with Lucene search engine		Yes	All files	File directory with service oriented architecture Service oriented flat file storage systems Allows fast searches no matter where the document is stored
File directory system	No	Yes	All files	Inherent to platform's operating system (windows)

Table 1. Comparison of methods to store, query and replicate the data.

THIS PAGE INTENTIONALLY LEFT BLANK

## **IV. PROOF OF CONCEPT**

This chapter describes a proof-of-concept of a portion of the architecture described in Chapter III. It will first describe the assumed scenario for this proof-of-concept. Next, it will describe the physical architecture alternatives. Finally, it will describe the experimental setup to analyze these alternatives, and discuss the pros and cons of each alternative.

### **A. PROOF-OF-CONCEPT SCOPE AND TARGETED ANALYSIS OF ALTERNATIVES**

#### **1. Scenario Review and Scope**

Assuming that the architecture described in Chapter III is used to share information in a military operation, two countries need to cooperate and work together to complete a tactical military operation. Information sharing with each other needs to be done in a timely and secure manner, as a lack of complete information could lead to poor decision making and loss of lives. One country can offer resources like manpower and tanks, due to its proximity in the region, and another country is technologically advanced in its gathering of intelligence. The country with the physical resources needs to have the intelligence information from the other country in order to execute its mission well. Such cooperation requires an effective way in which information can be shared in a timely and secure manner.

#### **2. Proposed Architecture System Scope**

Several technological components need to be considered in the building of the physical architecture. One of the most important aspects is the security of the data being shared. Often, users are protective of their data, and are most concerned with this aspect, thus they are reluctant to trust the technology with their data. Security of the data is measured by the four secure components used in the technological platform: VPN, VMWARE, Encryption of data, and Encryption of network. When there are more security components in the architecture, the

data is more secure but there would be increased latency in the data transfer. Thus, a balance needs to be struck between timeliness and data security.

It is concluded in the system architecture that the VPN component is compulsory, as every country would have their data stored on a private network. This experimental setup would find out the time it takes for a 1GB JSON file, which is approximately the size of a large tweet, to transfer between the network, when a factorial combination of two out of the remaining three security components is present in the architecture. The two security components are Encryption of data and Encryption of network. With these timing values, we can strike a balance between the objectives of timeliness and security, and the countries can better decide how many of the security components to include in the architecture. The experimental table can be found in Table 2. It is assumed that the four security components increase the security of the data by the same amount for simplicity.

VPN	VMWare	Encryption of data	Encryption of network	Total security level
Yes	No	Yes	Yes	3/4
Yes	No	Yes	No	2/4
Yes	No	No	No	1/4
Yes	No	No	Yes	2/4

Table 2. Experimental table with factorial combination of the different encryptions.

### 3. Experimental Setup

Figure 29 shows the physical architecture of the system. The proof-of-concept that is implemented is shaded in yellow. The home country C1 downloads the data that was requested to be infused to the data infusion

environment that is managed by the information sharing authority. The information sharing authority receives the data and prepares it for use by the information group for further analysis.

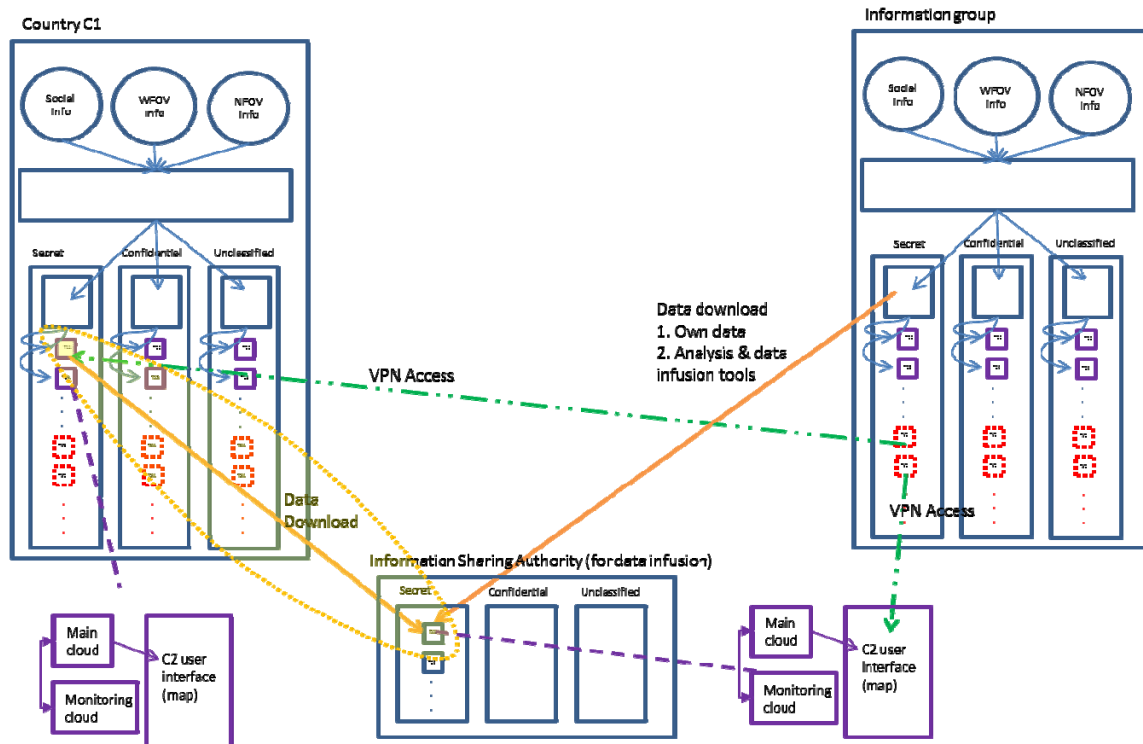


Figure 29. Full physical architecture of the system with the proof of concept portion highlighted.

The experimental setup for the proof-of-concept is shown in Figure 30. The home country that shares the information is represented by Computer A. The data to be downloaded is encrypted and sent via a CAT 5E Ethernet cable to Computer B. Computer B is the information sharing authority that manages the data infusion environment. Computer B receives the data from Computer A and decrypts the data.

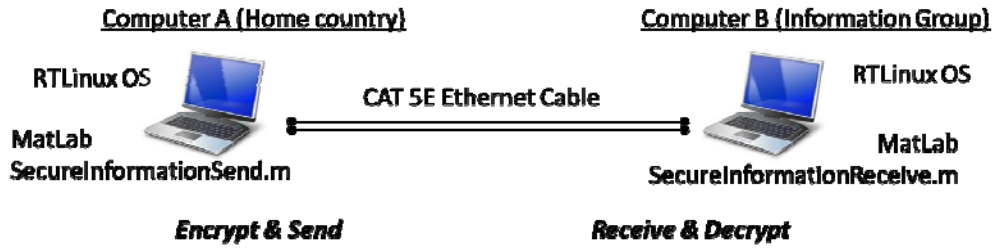


Figure 30. Schema diagram of experimental setup.

#### 4. Coding

This section explains the coding that was done in MATLAB. There are two main sets of codes that are used in this experimental setup. The first code, InformationSecureSend(), encrypts the data to be shared, and then sends the data. The second code, InformationSecureReceive(), receives the data and decrypts the data.

##### a. InformationSecureSend()

Figure 31 shows the process flow for which a tweet data is prepared for sending to the data infusion environment.

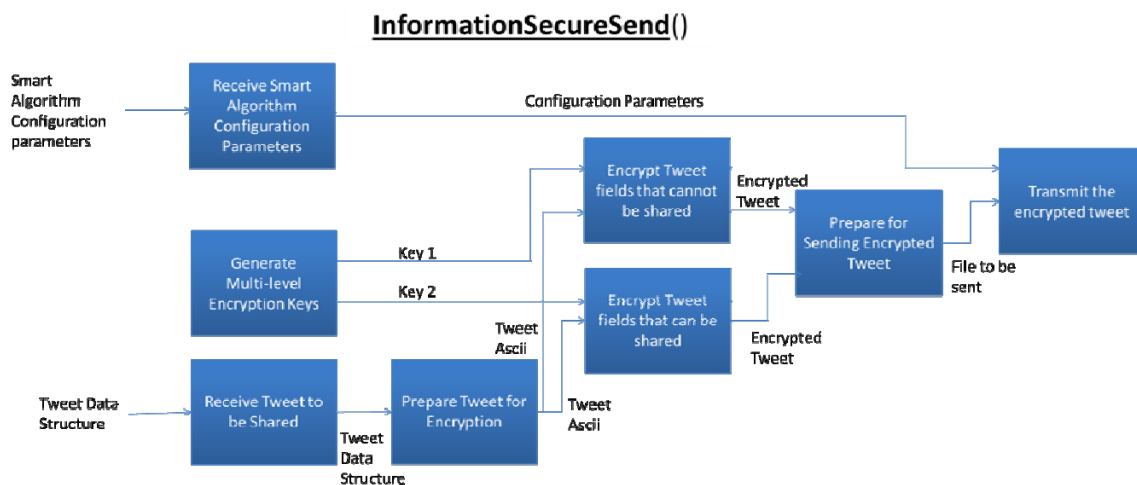


Figure 31. InformationSecureSend() algorithm.

“Generate multi-level encryption keys” would generate two security keys. One key is for encrypting the tweet fields that cannot be shared, and the second key is for encrypting the tweet fields that can be shared. Key\_generator() is a function that generates the security key. The coding for this process is shown in Figure 32.

```
[key1] = Key_generator(1,255,4); % key for encrypting non-sharing fields
[key2] = Key_generator(2,255,4); % key for encrypting sharing fields
```

Figure 32. MATLAB code for security key generator.

The tweet data to be shared is in the form of a struct data type. The “Prepare tweet for encryption” process converts the tweet struct into ASCII format for use in the encryption function. This process is shown in Figure 33. “Convert\_string\_to\_ascii\_number\_array()” is the function that does the conversion and it outputs the tweet data in ascii format.

```
%convert variables to ascii to be in the required format for the encryption function
name_ascii = convert_string_to_ascii_number_array(tweet(1).name);
user_profile_location_ascii =
convert_string_to_ascii_number_array(tweet(1).user_profile_location);
username_ascii =
convert_string_to_ascii_number_array(tweet(1).username);
tweet_geo_ascii =
convert_string_to_ascii_number_array(tweet(1).tweet_geo);
content_ascii =
convert_string_to_ascii_number_array(tweet(1).content);
time_ascii = convert_string_to_ascii_number_array(tweet(1).time);
```

Figure 33. MATLAB code for conversion of tweet data format for use with encryption code.

“Encrypt Tweet fields that cannot be shared.” and “Encrypt Tweet fields that can be shared” uses security key 1 and security key 2, respectively, to encrypt the tweet that is in ASCII format. Encryption\_Coding\_Minimize\_

Array\_Algorithm() function is the function that does this encryption, and it outputs the encrypted tweet data. This process is shown in Figure 34.

```
    %encrypt name, user_profile_location, username with key 1
    [name_encrypt,table1] =
Encryption_Coding_Minimize_Array_Algorithm(key1, name_ascii);
    [user_profile_encrypt,table2] =
Encryption_Coding_Minimize_Array_Algorithm(key1,
user_profile_location_ascii);
    [username_ascii_encrypt,table3] =
Encryption_Coding_Minimize_Array_Algorithm(key1, username_ascii);

    %encrypt tweet_geo, content, time with key 2
    [tweet_geo_encrypt,table4] =
Encryption_Coding_Minimize_Array_Algorithm(key2, tweet_geo_ascii);
    [content_encrypt,table5] =
Encryption_Coding_Minimize_Array_Algorithm(key2, content_ascii);
    [time_encrypt,table6] =
Encryption_Coding_Minimize_Array_Algorithm(key2, time_ascii);
```

Figure 34. MATLAB code for encryption.

“Prepare for Sending Encrypted Tweet” process prepares data for sending by consolidating the encrypted tweet data into a single file. This process is shown in Figure 35.

```
% prepare for sending one file

save(['tweet_fileEDEL'], 'name_encrypt', 'user_profile_encrypt', 'userna
me_ascii_encrypt', 'tweet_geo_encrypt', 'content_encrypt', 'time_encrypt
', 'table4', 'table5', 'table6');
```

Figure 35. MATLAB code for preparing the tweet datas for sending.

“Transmit the encrypted tweet” process sends the consolidated file. This encrypted datalink process is shown in Figure 36, and the unencrypted data link process is shown in Figure 37.

```

%send file via encrypted datalink
system(['scp 'tweet_fileEDEL.mat'
'user01@192.168.0.122:/export/ramdrv/computerb/receive/tweet_fileEDEL.mat
...

```

Figure 36. MATLAB code for sending tweet data via encrypted datalink.

```

% prepare & send one file via unencrypted datalink
save([pathnameUnsecure
'tweet_fileEDUL'], 'name_encrypt', 'user_profile_encrypt', 'username_ascii_e
ncrypt', 'tweet_geo_encrypt', 'content_encrypt', 'time_encrypt', 'table4',
'table5', 'table6')

```

Figure 37. MATLAB code for sending tweet data via unencrypted datalink.

**b. InformationSecureReceive()**

Figure 38 shows the process flow for which a tweet data is received by the data infusion environment and is processed to be read by the end user-information group.

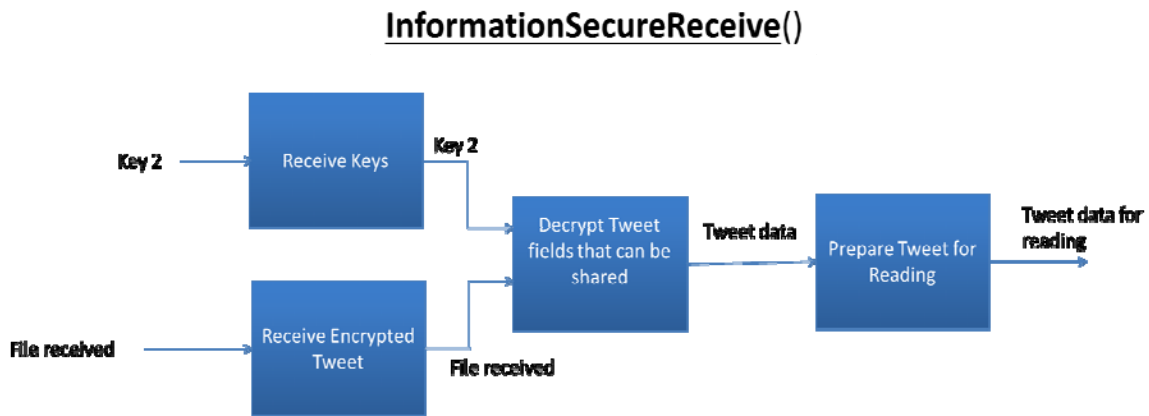


Figure 38. InformationSecureReceive() algorithm.

The “Receive keys” process receives key 2 from the home country—computer A for use of decrypting the encrypted tweet file that was received. “Receive Encrypted Tweet” process receives the encrypted tweet file from the home country. “Decrypt Tweet fields that can be shared” process uses key 2 to decrypt the encrypted tweet file and outputs the decrypted tweet data. The

“Prepare Tweet for Reading” process converts the decrypted data into character format for reading. The coding for this process is shown in Figure 39.

```
[key2] = Key_generater(2,255,4);
load([pathname 'tweet_fileEDEL'])

%decrypt      tweet_geo,      content,      time      with      key      2
tweet_geo_decrypt=Decryption_Deoding_Minimized_Array_Algorithm(key2,table4,tweet_geo_encrypt);

content_decrypt=Decryption_Deoding_Minimized_Array_Algorithm(key2,table5,content_encrypt);

time_decrypt=Decryption_Deoding_Minimized_Array_Algorithm(key2,table6,time_encrypt);

%convert variables to char
tweet_geo_decrypt_char = char(tweet_geo_decrypt);
content_decrypt_char = char(content_decrypt);
time_decrypt_char = char(time_decrypt);
```

Figure 39. MATLAB code for receiving and decrypting tweet data

## 5. Results

Informationsecurerceive() function encrypts and sends the tweet data. It also records the time it takes to encrypt the data and the amount of time it takes to send the data. Informationsecurereceive() function receives the encrypted data and decrypts it for the information group to use the data. It also records the time it takes to receive and read the data and the amount of time it takes to decrypt the data.

The experimental runs (four runs) were conducted for the factorial combination the two security components—Encryption of data and Encryption of network, and their results are summarized in the Table 3.

It is noted that encryption of the network uses a significant amount of time, as illustrated by the relatively longer time taken for the data to be received by computer B in runs 1 and 4.

RUN	VPN	VMWare	Encryption of data	Encryption of network	Total security level	Encryption time(secs)	Sending time(secs)	Read time(secs)	Decryption time(secs)	Total time (secs)
1	Yes	No	Yes	Yes	¾	0.001424	0.2782	0.000435	0.0006114	0.280665
2	Yes	No	Yes	No	2/4	0.000794	0.0051	0.000430	0.0005959	0.006935
3	Yes	No	No	No	¼	N.A	0.0034	0.000365	N.A	0.003765
4	Yes	No	No	Yes	2/4	N.A	0.3006	0.000576	N.A	0.301152

Table 3. Experimental results for the factorial combination of the different encryptions.

THIS PAGE INTENTIONALLY LEFT BLANK

## **V. CONCLUSION AND FUTURE WORK**

### **A. CONCLUSION**

The main contribution of this thesis is the proposed architecture of a configurable cloud infrastructure that enables multiple layers of secure information sharing between multiple organizations. This architecture was developed through a systems engineering process with an analysis of alternatives included. This thesis gives a broad overview of all the systems and subsystems required in order for this architecture to work. The thesis also instantiates part of the proposed architecture with a proof-of-concept system in a laboratory environment. Finally, this thesis performs and documents test and evaluation of the proof-of-concept system. The proof-of-concept chooses a specific scenario of information sharing that would allow NATO members to access shared data faster in order to make decisions more quickly.

### **B. FUTURE WORK**

Upon successful implementation of this architecture, an organization using this system would gain a technological platform and an automated method for fast information sharing among multiple organizations having different information access rights but a common goal requiring them to communicate. The proof-of-concept experiment has successfully tested information sharing between two different organizations with different information access rights. This sharing of information could be further developed to include VMWare to tighten the data security of the information shared.

Secondly, a third computer could be incorporated into the setup to access computer B's data via VPN. This would test the accessing of information by the information group (computer C) in the data infusion environment (computer B).

Thirdly, a smart algorithm could be implemented to automatically sort the new incoming raw data into its correct location based on security classification and the data information with respect to mission name, geographical location, time, and information source.

## LIST OF REFERENCES

- Ammon, Grant P. 2013. "Professional Dialogue with NPS President Helps Students Set Thesis Topics." <http://www.nps.edu/About/News/Professional-Dialogue-With-NPS-President-Helps-Students-Set-Thesis-Topics.html>.
- The Apache Software Foundation. 2012. "Apache Solr." Accessed February 28, 2014. <http://lucene.apache.org/solr/>.
- The Apache Software Foundation. 2013. "Apache ACCUMULO Notable Features, 2013." Accessed February 28, 2014. [http://accumulo.apache.org/notable\\_features.html](http://accumulo.apache.org/notable_features.html).
- Chairman of the Joint Chiefs of Staff Instruction 6285.01C. 2013. "Multi-national and Other Mission Partners (MMNP) Information Sharing Requirements Management Process." Accessed February 28, 2014. [http://www.dtic.mil/cjcs\\_directives/cdata/unlimit/6285\\_01.pdf](http://www.dtic.mil/cjcs_directives/cdata/unlimit/6285_01.pdf).
- Huber, Reiner, Tor Langsaeter, Petra Eggenhofer, Fernando Freire, António Grilo, Anne-Marie Grisogono, Jose Martins, Jens Roemer, Mink Spaans, and Klaus Titze. March 24, 2008. *The Indian Ocean Tsunami*. Accessed October 22, 2013. [http://www.dodccrp.org/files/case\\_studies/Tsunami\\_case\\_study.pdf](http://www.dodccrp.org/files/case_studies/Tsunami_case_study.pdf).
- IDA. 2013. "Social Media Technology Roadmap." Accessed October 22, 2013. <http://www.ida.gov.sg/~media/Files/Infocomm%20Landscape/Technology/TechnologyRoadmap/SocialMedia.pdf>.
- Goshorn, Deborah. "SE3201-SE3203: Smart Clouds, Using Smart Apps for Tactical Applications." Course, Department of Systems Engineering, Naval Postgraduate School, Monterey, CA, January, 2013.
- Goshorn Rachel, Deborah Goshorn, Joshua Goshorn, and Lawrence Goshorn. 2010. "Behavior Modeling for Detection, Identification, Prediction, and Reaction (DIPR) in AI Systems Solutions." In *Handbook of Ambient Intelligence and Smart Environments*, ed. Hideyuki Nakashima, Hamid Aghajan, and Juan Carlos Augusto, 669–700. New York: Springer.
- MongoDB Inc. 2014. "MongoDB." Accessed February 28. <https://www.mongodb.com/products/mongodb>.

- Natarajan, Ramesh. 2010. "VMware Virtualization Fundamentals–VMware Server and VMware." Accessed February 28, 2014. <http://www.thegeekstuff.com/2010/06/vmware-server-and-vmware-esxi-introduction/>.
- Sqrrl Data, Inc. 2014. "Accumulo." Accessed February 28. <http://sqrrl.com/product/accumulo/>.
- VMWare Inc. 2014. "VMWare." Accessed February 28. <http://www.vmware.com/virtualization/>.
- vsChart.com. 2014. "Apache Accumulo vs. MongoDB." Accessed February 28, 2014. <http://vschart.com/compare/apache-accumulo/vs/mongodb>.
- Wikibon Blog. 2012. "Accumulo: Why The World Needs Another NoSQL Database." Accessed February 28, 2014. <http://wikibon.org/blog/breaking-analysis-accumulo-why-the-world-needs-another-nosql-database/>.
- Wikipedia*. 2014a. Accessed March 22. s.v. "MongoDB," <http://en.wikipedia.org/wiki/MongoDB>.
- Wikipedia*. 2014b. Accessed February 26. s.v. "Open VPN," <http://en.wikipedia.org/wiki/OpenVPN>.
- Wikipedia*. 2014c. Accessed February 26. s.v. "Virtual Private Network." [http://en.wikipedia.org/wiki/Virtual\\_private\\_network](http://en.wikipedia.org/wiki/Virtual_private_network).
- Wikipedia*. 2014d. Accessed February 26. s.v. "Virtualization." <http://en.wikipedia.org/wiki/Virtualization>.

## INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center  
Ft. Belvoir, Virginia
2. Dudley Knox Library  
Naval Postgraduate School  
Monterey, California