Enhancing Process Visibility of the Supply Chain 'Data Pipeline'

Potchara Pruksasri[1], **Jan van den Berg**[2], **Wout Hofman**[3], and **Yao-Hua Tan**[4], Non-members

ABSTRACT

Although business process visibility is considered to be increasingly important for international trade, the visibility of existing supply chains is currently still ambiguous. Information deficiencies such as incorrectness or inconsistency of data are major determinants that decrease clarity. The Data Pipeline principle has been proposed to overcome the data quality shortcomings, enhance the visibility, and improve performance of the supply chain. In a first elaboration, the Data Pipeline model named the Distributed Trust Backbone (DTB) was designed and implemented in three different countries. In order to discover the effectiveness and feasibility of the model, the visualization of the Data Pipeline's process flow is urgently required in terms of both real-time detection of system failures and process flow representations. However, there is no process visualization feature available for the Data Pipeline. This challenge is taken up in this paper. We propose a Data Pipeline monitoring system and describe the results of performed simulation tests based on a case study of the international trade lane between Southeast Asia and Europe.

Keywords: data pipeline, process visibility, supply chain, dynamic system, monitoring

1. INTRODUCTION

Thanks to the various shipping services available today, local exporters can conveniently deliver their products to customers around the world, but global shipment is not as simple as carrying fruit home from the market. The chain of transporting goods to supply businesses' demand is hugely complex and varied. Looking at the process of the international supply chain, its characteristics are highly diverse in terms of both physical movement and information flow. This is because there are many stakeholders involved in the process from the beginning of the transportation until the delivery of the goods at their destination. In addition, during the shipment, not only goods (containers) are transported along the trade lane, but information related to the shipment also flows through many information systems that are related to the shipment [1]. This information is acknowledged to be equally essential to the physical transportation because relevant stakeholders need high quality information in order to effectively manage their tasks in the supply chain activity.

The data flows of the supply chain are complex, dynamic and constantly changing depending on business activities as well as on the number of involved systems. The supply chain information system (SCIS), therefore, requires proper mechanisms to control the exchanging of the information to drive the supply chains smoothly. However, several studies [2–4] have shown that existing SCISs such as the European SCIS are still facing information quality shortcomings and require improvements in visibility, security and performance. To overcome the problems, the new conceptual SCIS named the 'Seamless Integrated Data Pipeline' or, in short, the 'Data Pipeline' [5] was proposed by UK and Dutch Customs to the European Union (EU). The concept of the Data Pipeline aims to enhance the quality and security of data that are being exchanged between stakeholders in the current supply chains based on new data-sharing and exchanging mechanisms.

Traditionally, information on the goods is passed from one to another actor who is executing the supply chain task, but the key principle of the Data Pipeline has totally shifted from the traditional data passing (data push) to a data-requesting (data pull) scheme [6]. Information on the goods needs to be available for authorized partners who are related to the shipment at its origin. According to this concept, information provided by the source (owner) is supposed to be the most accurate and updated. As a result, the quality of information flowing in the SCIS should be higher quality because there are no intermediate tiers between the data sources and the requester. Reversing the communication scheme then creates a major challenge for designing and developing the new system.

The Data Pipeline is currently in the initial phase. The goal of creating an effective, secure and transparent system is stimulating governmental authorities

Manuscript received on February 24, 2016 ; revised on March 25, 2016.

Final manuscript received on April 5, 2016.

[1,2,4]The authors are with Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands, E-mail:p.pruksasri, j.vandenberg, y.tan @tudelft.nl

[3]The author is with TNO, Brasserplein 2, Delft, The Netherlands, E-mail: wout.hofman@tno.nl

and business bodies to cooperate in the development of the new system. Several Data Pipeline models have been introduced and aim to improve the supply chain information in different aspects, for example the Distributed Trust Backbone (DTB) [7], our proposed model, which focuses on building up a secure information sharing and exchanging mechanism for the Data Pipeline. The prototype of the DTB has been implemented in three countries spread all over the world in order to demonstrate the model by a real world case. In order to discover the effectiveness and feasibility of the Data Pipeline model, simulation tests have been performed. It turned out that the simulation is sometimes interrupted due to an error that unexpectedly occurs in some sub-systems within the process flow. The error causes insufficient information for operation processing and lowers the visibility of many supply chain activities. The process flows become unclear which decreases the performance of the SCIS. To enhance the visibility of the supply chain Data Pipeline, visualization of the process flow is recommended, particularly for the real-time detection of system failures and the information flow representations

This challenge has been taken up in this paper. The remainder of the paper is organized as follows. The background and related works are presented in the next section. After that, we show the requirement analysis and design in the third section. In the fourth section, we describe the implementation of the Data Pipeline prototype, the proposed model and then illustrate the simulation and testing. Finally, we draw conclusions in the last section.

2. BACKGROUND AND RELATED WORKS

2.1 The supply chain information system

The international supply chain is a dynamic system that consists of people, organizations, activities, resources, and information, which are involved in moving goods or services from suppliers to customers internationally. Figure 1 represents conceptual layers of the existing supply chain system.

At the bottom of the figure, the logistics layer demonstrates the transportation flow of the container from a supplier to an overseas customer. The following descriptions of export/import sub-processes describe how the container is shipped in practice. The export process starts when the exporter receives an order from his business partner (an importer). The exporter then packs the goods in a container and buys a transportation service from a service provider also known as a freight forwarder. The forwarder arranges inland transportation in order to move the container to the port. The forwarder typically hires a maritime carrier to ship the container to the destination country. At the port, the container is loaded onto a vessel of the carrier and starts its journey to the customer. Moving to the import process, when the

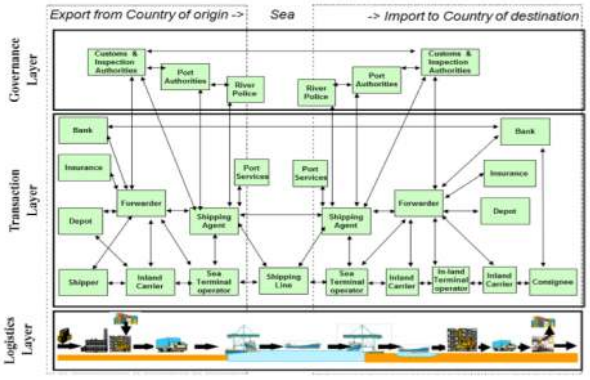


Fig.1: *Conceptual layers of the existing global supply chains [1].*

vessel arrives at the port of the destination country, the containers mentioned on the discharge list will be unloaded at the sea terminal. The process of inspection of the container will be started. Concerning tax declaration, control of dangerous goods or import of forbidden goods, suspicious containers will be physically inspected; otherwise the container will enter the country and eventually be delivered to the customer.

According to the processes, many actors perform their tasks to move the container to its destination. The transaction and governance layers in Figure 1 show a lot of relevant actors who are related to the business and governance activities of the shipment line. These actors exchange information related to the container from the beginning until the end of the journey. Thus, basic information about the goods in the container should be properly shared and exchanged between stakeholders in order to facilitate smooth and fast transportation.

Based on the characteristics of the supply chain, it can actually then be viewed from two perspectives: logistics and information. The logistics perspective considers the physical movement of the goods (containers) from the origin to the destination. Several means of transportation are linked together like a chain from suppliers to customers. On the other hand, the information perspective focuses on the information exchange and aims to support the effective movement of the container along the shipment line. Basic information such as product detail, quantity, weight, owner, and destination should be available when it is required. High quality data, e.g. accurate and timely, is essential and strongly recommended for this perspective. However, the current supply chain systems still suffer from data deficiency, for instance, incorrectness, inconsistency and unclear accountability. These poor data decrease performance and visibility, and cause security breaches in the supply chain. An improvement in the supply chain information system then becomes highly recommended [9, 10].

2.2 The seamless integrated data pipeline

Since 2001, Customs has used electronic systems such as paperless transactions and unique consignment identification in the supply chain system. Information on goods is transferred electronically through many systems along with the traveling of the container, but information is not always correct or updated when it arrives at the target system. The example below describes why data incorrectness can happen in a real situation. In the export process, for example, the exporter who packs the goods into a box creates a packing list for his exported products before providing it to the freight forwarder. The forwarder receives the package list from the exporter and then passes it to the sea carrier. The carrier lodges the list with the Customs of the destination country before the vessel departs from the port of origin. Based on the package list, Customs recognizes the goods shipped within the container. Passing information at each step, however, it needs to be re-inputted in the internal system of the handlers (exporter, forwarder, carrier and Customs) in order to be used for internal processes. Intentional or unintentional modification may occur in any system because of human error. When data has been changed at a single point, then the remaining systems all receive modified and incorrect information. Using the wrong information leads to many problems in the supply chain, for example, tax evasion or smuggling. Finally, the supply chain performance will drop and cause damages to government and business.

The UK and Dutch Customs have taken this data quality shortcoming into account and proposed a new concept of the supply chain information system named the 'Seamless Integrated Data Pipeline' or simply the 'Data Pipeline' [5] to the European Union (EU). The key concept of the Data Pipeline is that information related to the goods will become available for authorized partners at its source, i.e., from the original information system in which it is added. This enables authorized parties to get correct data because it is provided by the data owner. The Data Pipeline concept should minimize the need to make possibly incorrect copies of the data. Additionally, information in the Data Pipeline may be added to gradually when goods' status has been changed, for example, the GPS data that are related to the real-time position of the container. All the systems involved in the Data Pipeline will be virtually linked up to one single pipeline through which the supply chain stakeholders can then share and exchange data among each other, as shown in Figure 2.

The Data Pipeline is currently in the developing phase. Many organizations including both government authorities and business bodies are cooperating to build up an effective Data Pipeline system in several aspects. Our group took on the challenge of designing the secure information sharing and exchanging of the Data Pipeline in 2010. We proposed a new model of the secure Data Pipeline named the 'Distributed Trust Backbone' or DTB [7, 8] in 2012. The DTB aims to establish secure information exchange in the Data Pipeline based on the 'chain-of-trust' concept [9], underlying the basic idea that only trusted parties are allowed to exchange information within the same community. The prototype of the DTB has been constructed in three different countries: the Netherlands, Ireland, and Thailand. However, the feasibility and the effectiveness of the model need to be verified in order to ensure that the model is properly functioning.

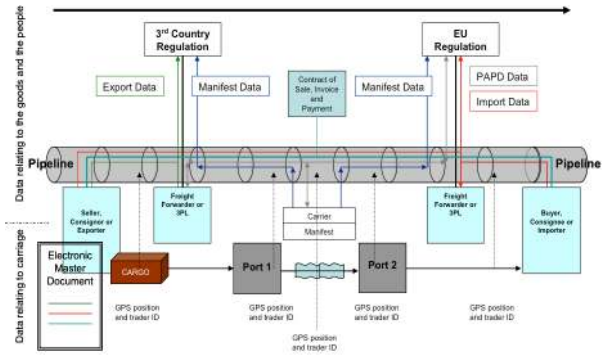


Fig.2: *Information flows of the supply chain Data Pipeline [3].*

2.3 Supply chain visibility and process visualization

Supply Chain Visibility (SCV) is everything relating to the necessary and sufficient information (identity, location, and status) required for entities, stationary or moving, that are hierarchically organized, and making their way through the supply chain. This requisite information is transmitted in messages about events, as defined within processes. The date and time of the actual event occurrence are compared to the corresponding planned date and time to render transparent the implications for decision-making [10].

Hence, transparent SCV greatly depends on high-quality information within the supply chain process. In its turn, the high-quality information heavily depends on sufficient, accurate, and updated data [11, 12]. Applying the Data Pipeline concept to supply chains could improve the SCV by providing high-quality information to the system. Sharing, requesting and exchanging data between sub-systems along the supply chain are considered fundamental operations to obtaining necessary information in the supply chain. Therefore, smooth and uninterrupted data exchange within the system is certainly vital. Design and implementation of the Data Pipeline should make sure that information flow is continuous. Any unexpected failures occurring within the process flow should be rapidly captured, discovered and corrected.

A visualization of the process flow during data exchange then becomes one of the most important tools for developing and testing the supply chain Data Pipeline model described in this paper. In contrast, some other Data Pipeline models such as the Atos Data Pipeline prototype, IBM Supply Chain Visibility Dashboard, or GS1 Visibility framework [13] have also been developed in order to demonstrate the new concept of Data Pipeline information exchanging. However, these three models mainly focus on the proper sharing and exchanging of data within the Data Pipeline since the Data Pipeline is still in the preliminary phase as mentioned earlier. Therefore, the proposed Data Pipeline models particularly aim to launch an initial version of the data exchange system based on the Data Pipeline concept, but none of them provides the process visualization feature that the model described in this paper offers.

3. REQUIREMENTS AND DESIGN

3.1 Requirements analysis

The DTB, our proposed Data Pipeline model, relies on the Confidentiality, Integrity, Availability, and Accountability (also known as CIA-A) principles [14]. All the DTB members must comply with designed security protocols in order to enable secure data exchange. In general, information flows in the DTB prototype should be safe and smooth if there is to be no accidental failure of any of the sub-systems. According to the implementation and evaluation of the DTB, some operations, for example, in the verification of the members' status that the DTB makes use of the Certificate Authority (CA) as the Trusted Third Party (TTP) [15], will be disturbed when some problems occur at the CA such as delay or a heavy load. In another case, sharing information at its source can also cause an error in the process flow if the data source system is not available or too busy to respond. These obstacles prevent the coherence of the data exchange process within the Data Pipeline. To capture and discover the mentioned problems, the visualization of the process flows is recommended for developing the Data Pipeline since there is no process visualization tool of the Data Pipeline currently available. This thus becomes the first requirement of our process visualization system.

Requirement I: *The information that is being exchanged in the system should be monitored and its status should be captured in a secure way from the time the flow of the information is started until it is finished.*

Besides that, government authorities, for example, the Customs who are mainly responsible for controlling the goods across the border, require high data quality for their processes since the sufficiency of the required data enables the control system to work efficiently. In any case of failure, it should be reported to an administrator in an understandable format in order to find out the exact point of the problem and provide a fast response to the incident. The visualization system should provide an effective reporting tool to a system administrator. We then defined the second requirement.

Requirement II: *The proposed system should represent the process flow in an understandable interface in order to facilitate discovering failure spots, improving process visibility of the supply chain data exchange and encouraging governance purposes.*

Since the prototype system of the DTB has been implemented, any alterations to the system can take time and affect the overall system. Therefore, in order to integrate the proposed system into the DTB model and prevent the system confliction, the third requirement is defined.

Requirement III: *In order to integrate the proposed system into the Data Pipeline model, it should support existing Data Pipeline operations without intensive implementation. Process visibility capability should be indicated to a satisfactory level compared to other Data Pipeline models.*

3.2 Proposed Design

To enhance the visibility and monitoring of the process flows, we propose the principle that tracking the process flow should utilize the log data of the transaction processing. The status of the transaction (message) should be recorded in the system. Meanwhile, the system should provide a communication channel to access the recorded status. By this concept, we suggest all sub-systems in the Data Pipeline must have a container where the process status will be stored and provide a secure communication interface for authorized systems to get access. The proposed monitoring model and its description are presented as follows.

3.2.1 Process Tracking Architecture

The tracking system consists of four components that are utilized to capture all the message statuses when they are exchanged between stakeholder systems in the Data Pipeline. Figure 3 shows the conceptual process tracking system embedded in the Data Pipeline model (DTB).

3.2.1.a Process log container: Each computer system working in the Data Pipeline must prepare a container to store the process status. The container can be in the format of a file or database system. Information about the process status of the message is called a 'Process log' and suggests its status to be

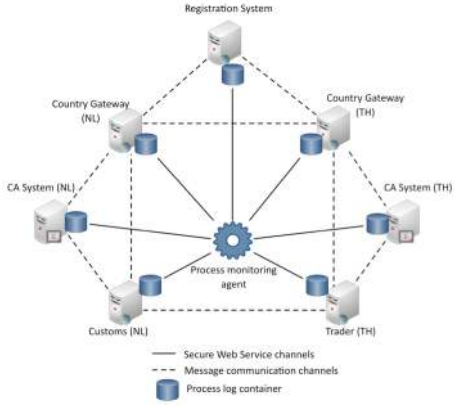


Fig.3: *Data Pipeline process tracking architecture [3].*

sent, received and processed. At each step of the message processing, the result must be recorded in this container.

3.2..1.b Secure communication interface: In order to get access to the log container, a secure communication channel is required; for example, secure web services either SOAP or REST protocols [16]. The secure web services can be applied to the Data Pipeline by using its Trust Third Party (TTP) to ensure the communication within the monitoring system is reliable. The data protection mechanisms such as transaction concealing and exposing [17] and contextual attribute-based access control [18] are potential techniques to secure the communication.

3.2..1.c Process monitoring agent: An agent software that is used for requesting, collecting and processing the log data from target systems. The Data Pipeline administrator will activate this agent to perform checking and diagnosing the system. In our case study, we assume that the Customs that have responsibility for the Data Pipeline systems want to check the message flow between them and their clients. The agent should be, for instance, implemented and embedded into the Customs' computer system.

3.2.1.d Log information: In order to make this architecture compatible with protocols of the supply chain activity, the information contained in the log is key. It should contain essential data that is usable for monitoring as well as for automatic discovering (tracking) of the log container. Based on our analysis, the content of the log should be composed of at least five basic elements, namely Process Unique ID, Timestamp, Source, Destination and Process Result. These elements should be combined and stored in the log container of the sub-systems. An example of the log template is shown below.

PUID:TIME:SOURCE:DESTINATION:RESULT

In practice, the log information will gradually be added to a system when data is processed. The mechanism of the monitoring system should start with each system generating a unique identifier of the process in the initial state, which is called the Process Unique ID (PUID). The PUID will be attached to the message and recorded in every sub-system in the process flow. The PUID can distinguish any process flow and identify the process owner in the whole system. The Universally Unique Identifier (UUID) Version 5, for example, can also be applied. Next is the Timestamp (TIME), which will be stamped by the system that is processing the transaction or message to indicate the time when processing started or finished. Here the Coordinated Universal Time (UTC) can be employed. With this value, the monitoring system can sort the order of the processing steps and present them in a sequence of time. To track the process flows, sub-systems that process the transaction are vital because information will flow from one to another system, which are called the Source and Destination systems. These could be the name of the system or a network address, which can be recognized in the whole system. In the DTB, the Data Pipeline ID and its alias can be implemented. The name of supply chain actors such as 'CustomsNL', which refers to the Dutch Customs system, or 'GatewayTH', which refers to the Gateway system of Thailand, are examples of the Source and Destination in the process flow. Finally, the result of the processing should include process information such as status and the process owner. Some other additional details could also be presented in this part to show more information about the status. In our model, sub-systems will generate the log information at every step of the message operation and store it in the log container automatically.

To sum up, the monitoring agent is embedded into the administration system and linked to log containers via secure web services channels. Log information stores the process statuses and facilitates the tracking system from the beginning until the end of the process. Consequently, process information will be captured at every state of the processing and be used to represent the process flow of the Data Pipeline in an understandable format. The process flow of any activity will be visualized from the beginning until the activity is completed. Based on the proposed design, the visibility of the supply chain process should be enhanced and facilitate delivery of better data quality to the Data Pipeline.

4. IMPLEMENTATION AND TESTING

We have implemented our proposed design on the DTB prototype in order to demonstrate how the tracking mechanism works in practice. All the DTB's core components and its security protocols are set up based on the Europe - Southeast Asia trade lane [19]. We performed testing on the protocols including discovery, identification, authentication and data

exchange protocols. In this section, we elaborate on the implementation detail and simulation results for requesting information and diagnosing failure spots between the Data Pipeline partners in Thailand and Netherlands.

4.1 Implementation

We have continued working on our proposed Data Pipeline model (the DTB). Its core components are implemented and described in this section.

The DTB, an infrastructure of secure data exchange based on the Data Pipeline principle, Public Key Infrastructure and Digital Certificate [20] technologies, consists of four main components: the Registration System (RS), the Country Gateways (GW), the Trusted Third Party (TTP), and the Data Source systems (DS). All the components are linked together by a secure communication channel in order to establish the chain of trust between the systems. Figure 4 illustrates the Distributed Trust Backbone Architecture.

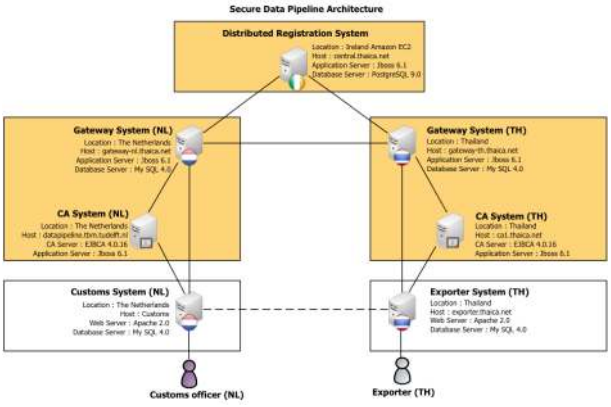


Fig.4: *Distributed Trust Backbone architecture.*

4.1.1 Registration System (RS)

The Registration System is a central system of the DTB, which provides administration tasks for all the supply chain members. Due to system availability, the RS is actually a set of RSs, which co-operate together and provide a virtual single entry channel for other systems to be communicated. The RS system is implemented on the cloud-computing environment (here at the Amazon Elastic Compute Cloud in Ireland) because the cloud-computing concept supports both the availability requirement and the scalability of the system. The implementation makes use of Java technology (J2EE) both for web application and web services. Based on the DTB model, sharing and exchanging information, all computer systems need to be registered to the Data Pipeline system as a Data Pipeline member and the RS should provide administrative actions such as member discovery and internal verification services for all the members. The verification service is crucial to the Data Pipeline according to the chain-of-trust concept. To enable the chain of trust between the RS and other components, the trustworthiness of all the systems must be examined before communicating with others. Therefore the internal verification service on connecting systems is implemented. Not only are the verification protocols tested, but other protocols such as information requesting or member discovering are also checked. The results of these processes are examples of data saved in the log container, which makes use of an open source database. The container access points are enabled via the secure web services channels. Web service applications provide communication channels for both the administration and monitoring tasks. At this point, every piece of log information is recorded, captured and made ready to be used in the monitoring system.

4.1.2 Gateway systems (GW)

In international supply chain systems, supply chain stakeholders are located in different countries over the world. Information exchanging between the countries is typically based on the international trading policy of each country. The Gateway acts as a communication entry point of the country. Before exchanging data between overseas partners, the GW should examine the request for information at the initial state. Only the countries that achieve a bilateral agreement on exchanging data can start communication. The gateway is implemented on the cloud computing systems in two countries, which are the origin and destination of the trade lane between Thailand and the Netherlands. Any result of the transaction processing particularly related to access control will be stored in the log container at the gateway. Meanwhile, the gateway also provides an implemented communication channel based on secure web services and is opened to requests from the monitoring system. We implemented both the web application and web services for the gateway by using Java EE technology.

4.1.3 Trusted Third Party (TTP)

Only trusted Data Pipeline members can exchange information with others as mentioned before. The verification process thus becomes a significant part of the model. However, the number of supply chain partners is enormous, and their systems are also highly diverse. To identify which system is a trustworthy system, the Trusted Third Party (TTP) is necessary for the Data Pipeline. According to research [21], many European countries have already set up TTP systems within their country. The DTB can benefit from these systems by making a request-for-verification from those existing TTPs. However, because of security reasons, we could not link our system to an actual TTP. We then therefore simulated

the TTP systems by using an open source CA system (EJBCA). With this system, we are ready for member identification and authentication services. However, although the TTP systems existing in the DTB are mainly based on the PKI and the digital certificate technologies, it can also be applied to other verification mechanisms depending on the policy of each country. Using the TTP, the DTB can finally bridge the chain of trust between the global (RS-GW) and the local (GW-DS) levels

4.1.4 Data Source (DS)

According to the Data Pipeline principle, information will be stored and shared with authorized parties from its source. The source system is called the Data Source (DS). The DS is a computer system of supply chain stakeholders that records all the necessary data related to an individual supply chain activity that the stakeholders are performing. We simulate the DS in Thailand and the Netherlands. In Thailand, the trader system is implemented in order to share container information with an authorized requester. In the Netherlands, we suppose that the Customs wish to get shared data from the trader. Thus the Custom's system, called the Customs dashboard, is implemented at the Netherlands' side. Data exchanging is securely exchanged within the formation of XML messages, which are called supply chain transactions. The dashboard gathers all the necessary information from the trader. The monitoring system is activated by Customs when the requesting is interrupted at certain spots. We implement the dashboard application, the monitoring agent and the trader web services to simulate the requesting of information. Both Customs and trader systems record all the processing results of the message to their log containers.

4.2 Simulation and Testing

Since we have already implemented the DTB prototype and the monitoring system, this section presents the tracking procedures on simulated situations for both normal and error cases in order to show how the prototype performs both monitoring and diagnosing a failure.

4.2.1 Simulation workflow

4.2..1.a The simulation starts with generating the requesting message and PUID at the Customs' dashboard. When the request is sent out to its destination (trader system), the Customs' system creates log information and sends records to the container including PUID, timestamp, operation result, source and destination systems.

4.2..1.b The message will flow to the Gateway as the first target. The Gateway puts the result of the message processing and process information in its log container after receiving the message.

4.2.1.c The Gateway executes an identification, authentication and verification of the message together with the CA system. In each step of security processing on the Gateway and the CA, results are stored in the container of both systems.

4.2.1.d The message continues to the trader system if there is no error in the security control. When the message arrives at the trader system, the system performs an execution of the request. All results of this state are recorded in the trader's log container.

At this point, the results of the processing are already stored in the containers, but they are spread throughout the system in different locations. We assume an administrator performs his routine task in monitoring the process flow. So, the monitoring agent captures log information and presents it to the officer.

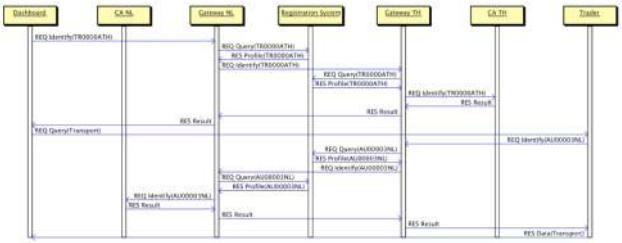
4.2.2 Monitoring

The monitoring process starts with Customs providing a PUID, which is generated by the Customs' dashboard system. The agent starts searching for log information on servers using the PUID. The destination of the message specified in the destination field is discovered. The agent then jumps to target servers via the secure web services with the PUID. At the target system, the agent makes a query on log data that contains the attached PUID. The target system responds by supplying the log information to the agent. After getting the log from the system, the agent checks the log fields. If it presents other destination systems, the agent then repeats tracking to capture further information.

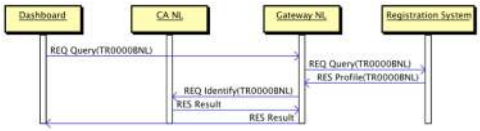
When all the logs are collected, the monitoring agent automatically sorts out the captured data by the timestamp. Then the sorted results are presented to the administrator by generating the time sequence diagram. This diagram shows the flow of the process at all states in all the systems. Figure 5a presents the message flow of the identification and authentication protocols, which are parts of the information request process within the DTB model.

The following description interprets the first diagram: the process starts with the Dutch Customs' dashboard sending a request to the trader's system via the Dutch Gateway. After the Gateway has received the message, it then performs security controls on the message together with the Registration, CAs and Thai Gateway. The process flow is working correctly in this diagram so the request has arrived at the trader's system. Then the trader's system answers the request by sending data back to the Dutch Customs in the final stage.

Similarly, Figure 5b shows the process flow of the discovery protocol generated by our proposed system when Dutch Customs have queried a member's profile from the Data Pipeline Registration system. According to the visualization of the discovery proto-

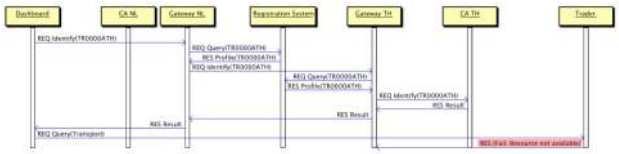


a.Information request process

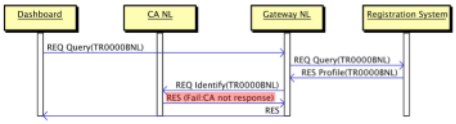


b.Member discovery process

Fig.5: *Examples of the visualization interface generated by the monitoring system.*



a.An error in the case of a trader's system being unavailable



b.An error in the case of a delay in the CA system

Fig.6: *Visualization of the failure spots generated by the monitoring system.*

col showed in Figure 5b, the Gateway acts as a broker and provides information for Customs. After the Gateway receives a request-for-query message from Customs, it then processes the request at the Registration system. With the information provided from the Registration system, the Gateway verifies the Data Pipeline member's information with the TTP (Dutch CA). Finally, the Gateway responds the member's profile, which is successfully verified by the TTP to the requester. We further perform monitoring experiments with other DTB protocols including registration, verification, data requesting and exchanging [7], access control [8], and data concealing and exposing protocols [17]. Our proposed system represents the precise visualization of all the protocols. Using this visualization, the system administrator clearly inspects the working of the Data Pipeline.

4.2.3 Diagnosing

To diagnose the system when some errors occur, the agent can simply detect the failure, for example if the dashboard cannot reach the destination system. If the target system does not receive the request from the dashboard, the log will not be available in this target system. In this case, the agent decides that there is an error in the target system. Then it generates the diagram that contains information of failure points (here it is the trader system). The detail of the error is provided and represented on the Customs system as shown in Figure 6a. In addition, to test the robustness of our system, we simulated many cases of failure in the Data Pipeline. Some of the diagnosing results, for example when a trader's system and CA are unavailable, are also presented in Figure 6b. Both diagrams in Figure 6 correctly show detecting failure points and reporting errors to the administrator. Thus, the administrator will immediately notice the point where the error has occurred, and they can execute another procedure for recovery or request information by other means.

We have implemented an agent that is able to automatically search the log container because the location of the source systems can be different in other supply chain activities. After the agent has received log information from different sources, it generates a sequence diagram dynamically without any additional implementation even if the source systems are different. Therefore, our model supports the changing of the dynamic system and its actors especially in the case of the supply chain Data Pipeline.

Table 1: *The process visibility capabilities for Process Information Gathering*

Measurement items for Process Information Gathering	Original DP		Enhanced DP	
	CA	PRFM	CA	PRFM
The system can capture granular (detailed) events in the entire process	1	11	9	100
The system can collect process information along the entire process in a timely manner	1	11	9	100
The system can gather process information from all steps (activities) in the process	N/A	N/A	8	88.8
The system can collect process information from the external process environment in a timely manner	N/A	N/A	N/A	N/A
The system can collect granular (detailed) information about a process's current status	1	11	9	100
The system can integrate process information from a variety of data sources	N/A	N/A	9	100
Average	0.5	5.5	7.33	81.48

Table 2: *The process visibility capabilities for Process Information Analysis*

Measurement items for Process Information Analysis	Original DP		Enhanced DP	
	CA	PRFM	CA	PRFM
The system has the ability to aggregate process data	2	22	9	100
Process information, such as status, are continuously captured by the system	1	11	8	88.8
The system has the ability to analyze process data to continuously capture process detail	N/A	N/A	8	88.8
Based on preset levels (thresholds), the system can automatically detect deviations from process plans	N/A	N/A	N/A	N/A
Based on process data, the system has the ability to identify the state of multiple processes, contextualized by their relationships	N/A	N/A	9	100
The system can indicate the performance of currently executed business processes	N/A	N/A	9	100
The system can predict final results of the business process during process execution	N/A	N/A	N/A	N/A
The system offers extensive analytical capabilities to examine process information	N/A	N/A	9	100
Average	0.38	4.13	6.5	72.22

Table 3: *The process visibility capabilities for Process Information Dissemination*

Measurement items for Process Information Dissemination	Original DP		Enhanced DP	
	CA	PRFM	CA	PRFM
Process information is distributed to process participants along the entire process	3	33	9	100
The system can notify the concerned process participants regarding events that may require adjustments	N/A	N/A	N/A	N/A
Using the system, process information is widely shared among process participants	2	22	9	100
Process information is delivered to process participants through simple, understandable tools	1	11	9	100
The system can create personalized monitoring views	1	11	9	100
Process information provided by the system often reaches relevant personnel timely enough to be of use	2	22	9	100
Through the system, process information are presented to process participants	2	22	9	100
Average	1.57	17.29	7.71	85.66

As aforementioned, the Data Pipeline is currently in the initial phase; other proposed models essentially aim to initiate a proper data exchange relying on the Data Pipeline concept but not including the process visibility. Unlike the DTB, we focus on establishing secure data exchange in the Data Pipeline, in which the data availability is acknowledged to be one of the most crucial parts to be considered. Integrating the monitoring system to the DTB certainly improves the process visibility and also enhances the data availability of the Data Pipeline. Table I-III shows the process visibility capabilities (CA) [22] of nine cases of the Data Pipeline after applying the proposed system based on nine test cases of the DTB protocols.

5. CONCLUSION

In this paper, we proposed a monitoring system of the Data Pipeline in detail. This facilitates checking, monitoring, diagnosing and visualizing of the process flow, which enhances the visibility of supply chain information systems. We tested the system in different situations: both usual cases and error cases. The proposed system performs its tasks correctly based on a case study related to the international supply chain information system. This shows that the proposed model can be applied to the supply chain Data Pipeline and indicates the performance (PRFM) of its process visibility level as 81.48%, 72.22%, and 85.66% for Process Information Gathering, Analysis and Dissemination respectively, which are significantly improved and higher in comparison to the unclear visibility of initial Data Pipeline models. However, based on testing with the prototype system, there is still a lot of work to do in order to end up with a full working system. First, the Data Pipeline is currently in the developing phase. There exists no real Data Pipeline. Many parts have not been implemented nor is a final design available. So, there is a need for additional developments in many aspects. Second, the intelligence of the proposed monitoring agent to support a variety of business processes has to be further studied in order to guarantee scalability at global scale. The outcome of this work can be used as a pilot study for further development of the supply chain Data Pipeline.

6. ACKNOWLEDGEMENTS

This paper results from the CASSANDRA project and is supported by funding from the 7th Framework Program of the European Commission.

References

[1] M. van Oosterhout, P. van Baalen, R. Zuidwijk, and J. Nunen, "Port Inter-Organizational Information Systems: Capabilities to Service Global Supply Chains Appendix A: Organizations and flows in the network", *Foundations and Trends in Technology, Information and Operations Management*, Vol.2, pp.241, 2008.

[2] D. Hesketh, "Seamless electronic data and logistics pipelines shift focus from import declaration to start of commercial transaction", *World Customs Journal*, Vol.3, pp.27-32, 2009.

[3] D. Hesketh, "Weaknesses in the supply chain: Who packed the box?", *World Customs Journal*, Vol.4, pp.3-19, 2010.

[4] P. Pruksasri, J. v.d. Berg, and S. Keretho, "Accountability in Single Window systems using an Internal Certificate Authority : A case study on Thailand's National Single Window system", *Proceeding of the 5th IADIS Multi conference on computer science and information systems*, 2011, Rome, Italy, pp. 129-136.

[5] E. van Stijn, D. Hesketh, Y. Tan, B. Klievink, S. Overbeek, F. Heijmann, M. Pikart and T. Buttery, "The Data Pipeline", *Global Trade Facilitation Conference*, United Nation Economic Commission for Europe (UNECE), 2011.

[6] B. Klievink, E. van Stijn, D. Hesketh, H. Aldewereld, S. Overbeek, F. Heijmann and Y. Tan, "Enhancing visibility in international supply chains: The Data Pipeline concept", *IJEGR*, Vol.8, pp.14-33, 2012.

[7] P. Pruksasri, J. v.d. Berg, W. Hofman, "Three protocols for securing the Data Pipeline of the international supply chain", *Proceeding of the 6th IADIS Multi conference on computer science and information systems*, 2012, Lisbon, Portugal, pp.27-34.

[8] P. Pruksasri, J. v.d. Berg, W. Hofman and S. Daskapan, "Multi-level access control in the data pipeline of the international supply chain system" in *Innovation in the High-Tech Economy* (Ed. P. Chuan, V. Khachidze, I.K.W. Lai, Y. Liu, S. Siddiqui, T. Wang), Springer-Verlag Berlin Heidelberg, 2013, pp.79-90.

[9] J. Francoeur and R. Peizer, "Digital chain of trust model for electronic commerce", 2002, https://www.google.com/patents/US20020065695 (Accessed: March 2015).

[10] V. Francis. "Supply chain visibility: lost in translation?", *Supply Chain Management: An International Journal*, Vol.13(3), pp.180-184, 2008.

[11] A. Zhang, M. Goh and F. Meng, "Conceptual modelling for supply chain inventory visibility", *Int. J. Production Economics*, Vol.133, pp.578-585, 2011.

[12] Paul A. Bartlett, Denyse M. Julien and Tim S. Baines, "Improving supply chain performance through improved visibility", *Int. J. Logistics Management*, Vol.18, pp.294-313, 2007.

[13] Cassandra, "Common Assessment and analysis of risk in global supply chains (CASSANDRA) project", 2011, http://www.cassandra-project.eu (Accessed: June 2015).

[14] B. C. Johnson, "Information security basics", *ISSA Journal*, Vol.8, pp.28-32, 2010.

[15] S. Daskapan, *MEDUSA: Survivable information security in critical infrastructures*, PhD Thesis, Delft University of Technology, The Netherlands, 2005.

[16] A. Nadalin, C. Kaler, R. Monzillo and P. Hallam-Baker, "Web services security", OASIS Standard Specification, 2004.

[17] P. Pruksasri, J. v.d. Berg, W. Hofman and Y. Tan, "Data Concealing of Supply Chain Transactions using the Distributed Trust Backbone", *Proceeding of the 9th International Conference for Internet Technology and Secured Transactions*, 2014, London, United Kingdom. pp. 151-156.

[18] M. J. Covington, M. R. Sastry, "A Contextual Attribute-Based Access Control Model", in *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops* (Eds. R. Meersman, Z. Tari and P. Herrero), Springer Berlin Heidelberg, 2006, pp. 1996-2006.

[19] I. Lucassen, "Cassandra WP400 Asia-NL/UK Trade lane living lab report", European Commission, 2014.

[20] C. Adam, and S. Lloyd, *Understanding PKI: concepts, standards and deployment considerations*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2002.

[21] H. Eertink, B. Hulsebosch and G. Lenzini, "STORK-eID: Framework mapping of technical/organizational issues to a quality scheme", Dutch Ministry of the Interior and Kingdom Relations, 2008.

[22] E. Graupner, E. Urbitsch, and A. Maedche, "A Conceptualization and Operationalization of Process Visibility Capabilities", *Proceeding of the 12th International Conference on Wirtschaftsinformatik*, 2015, Osnabrck, Germany.

Potchara Pruksasri has obtained both B.Sc. and M.Sc. of Computer Science at Khon Kaen University Thailand in 2000 and 2005 respectively. In 2005, he has employed as a lecturer at Mahasarakham University, Thailand. He is currently working on his PhD research at Section ICT, Faculty of Technology, Policy and Management (TPM), Delft University of Technology. His research focuses on information security of the supply chain system in order to secure data exchange of the supply chain.

Jan van den Berg studied mathematics and physics at the TUDelft while being active in the national student movement. In 1977, he received the diploma of Mathematical Engineer. From 1977-1989, he lectured courses in mathematics, physics and computer science on institutes of higher education in Breda and Eindhoven, and mathematics and physics at the secondary school of Nampula, Mozambique. From 2006, up till now he worked at TUDelft, mostly on topics related to (Big) Data Analytics and/or Cyber Security. On July 9 2013, he was appointed as full professor Cyber Security at Faculties of EEMCS and TPM Delft University of Technology.

Wout Hofman is senior research scientist at TNO, the Dutch organization for applied science, on the subject of interoperability with a specialization in government (e.g. customs) and business interoperability in logistics. He is responsible for coordinating semantic developments within the iCargo project. Wout is also as member of the Scientific Board of the EU FP7 SEC Cassandra project responsible for IT developments in that latter project.

Yao-Hua Tan is professor of Information and Communication Technology at the ICT Group of the Department of Technology, Policy and Management of the Delft University of Technology and part-time professor of Electronic Business at the Department of Economics and Business Administration of the Vrije University Amsterdam. He is coordinator of the EU-funded integrated research project ITAIDE on IT innovation to facilitate international trade. His research interests are service engineering and governance; ICT-enabled electronic negotiation and contracting; multi-agent modelling to develop automation of business procedures in international trade.