THREE PROTOCOLS FOR SECURING THE DATA PIPELINE OF THE INTERNATIONAL SUPPLY CHAIN

Potchara Pruksasri\textsuperscript{1}, Jan van den Berg\textsuperscript{1} and Wout Hofman\textsuperscript{2}
\textsuperscript{1}Section of ICT, Faculty of Technology, Policy, and Management, Delft University of Technology, Delft, The Netherlands  
\textsuperscript{2}TNO, Brassersplein 2, Delft, The Netherlands

ABSTRACT
In order to make the international supply chain system highly efficient in worldwide transportation of goods, its business processes should be supported by high quality information systems that, “24x7”, provide the right data, at the right time, at the right place, to the right person. To do so, the idea of a (virtual) “Seamless Integrated Data Pipeline” has been proposed, a new concept of an effective information exchange system. A key requirement of the Data Pipeline is that the information to be exchanged must be safe and reliable. Starting form the idea that existing systems should be used, this paper proposes three basic protocols that enable the secure exchange of data with respect to three identified security requirements, namely, that of secure identification of supply chain partners, that of harmonization of security levels, and that of harmonization of trust levels.

KEYWORDS
Data Pipeline, supply chain security, secure identification, chain of trust.

1. INTRODUCTION
Since the “Silk Road” era, the trade of goods across international borders has grown to the high level it is today. A singular trading activity usually starts with a buyer who makes an agreement with a seller on purchasing one or more products. Based on the conditions described in the agreement, the products are next transported to the buyer. After arrival, the payment is executed. This completes the (standard) trading procedure, generally known as the “BUY-SHIP-PAY” model (UNECE, 2001). In the “SHIP” phase, which covers all activities related to the physical transfer of the goods including official controls, certain activities are more complex for international trade (compared to domestic transportations): according to the laws of each country, shipments across borders are subject to many controls by the countries’ authority agencies. These controls relate to the prevention of unwanted events like those with respect to smuggle, tax evasion, and the import of forbidden products, among others.

The main type of equipment used in international intermodal transport is the container. The physical inspection of all containers when crossing a border is however impossible due to time and cost constraints. A method to sort out only the suspicious containers can, in principle, strongly reduce the number of containers to be inspected. Within the “green lane concept” (Schaefer, 2006), a risk assessment mechanism is adopted to support this idea: goods with a low risk profile can simply bypass the import inspection. So, green lanes strongly depend on high-quality risk assessments. In their turn, high-quality risk assessments strongly depend on accurate data. These transportation data are usually collected from all kinds of data sources as available in the international supply chain. In practice however, it turned out that data with respect to the products packed into a container often contain many errors, e.g., caused by incorrect data exchanges between two or more supply chain actors (Hesketh, 2010).

The English and Dutch Customs, having become aware of the existence of incorrect data about container goods, decided to propose a new information exchange system based on the assumption that the data at the origin (i.e., the place where data enter the system) should be considered as most genuine and, therefore, most correct. They termed the new information exchange system the “Seamless Integrated Data Pipeline” (Hesketh, 2009), or simply the “Data Pipeline”. Within this virtual Data Pipeline, the original data should be (made) available, at any right moment, for all authorized stakeholders (Overbeek et al., 2011), including
business partners and governmental stakeholders with supervision tasks. The elaboration and implementation of the Digital Pipeline is however no sinecure and involves many technical and organizational challenges.

A natural implementation of the Data Pipeline involves the setup of many (types of web) information services (Hofman, 2011), which is certainly not the standard implementation approach of current information systems in the international supply chain. But this is not the only obstacle for a quick adoption. Thinking about the more detailed design of web information services as basis for the new international Data Pipeline, several questions relating to data & information security pop up immediately: these questions concern the realization of secure and efficient data access as well as the implementation of other information security characteristics like data confidentiality, data integrity, and accountability (Johnson, 2010).

Based on the above-given considerations we decided to focus some of our research on information security problems in the proposed Data Pipeline. The results obtained from these efforts are presented in this paper: we introduce two security protocols related to two information security problems, namely, (a) the registration of so-called ‘community members’ of the international supply chain, and (b) the authentication of two international community members who wish to exchange data at a given security level. We also discuss the related chains of trust that need to be built in order to ensure secure registration and communication.

The remainder of this paper is structured as follows. We will first provide some further background information on the international supply chain and the data pipeline concept, including the adopted notion of community (section 2). Next, based on a further analysis, we present two security challenges as well as the security requirements resulting from these (section 3). In section 4, we then propose our information security solutions: two protocols are described and the related chains of trust discussed. Finally, we draw our conclusions and give suggestions for future research in the last section.

2. INTERNATIONAL SUPPLY CHAIN AND DATA PIPELINE CONCEPT

2.1 International Supply Chain

The dynamic system - consisting of organization, people, activities, information, and other resources - that aims to move products internationally from suppliers to customers can be denoted as the “International Supply Chain system”. This supply chain system can conceptually be divided into three different layers being the logistics layer, the transaction layer, and the governance layer (Oosterhout et al., 2008). These three layers can be viewed from two perspectives: the logistics (physical) perspective focuses on the moving of goods from origin to destination, while the information perspective considers the information exchange in the supply chain aimed to support the effective and efficient physical movement of the goods. In the different layers, very many stakeholders execute their different activities making the international supply chain a very complex system. To make our discussion concrete, we focus here on a few prototypical sub-processes of the transportation of goods from origin (often a factory) to destination (customer), namely those related to the import of goods by means of containers.

The description below is strongly based on the one presented in (Popal, 2008). Taking the physical perspective, the import process of the goods starts when the vessel that transports the goods arrives at the port. There, the containers mentioned on the ‘discharge list’ are unloaded and, next, stacked at the sea terminal where they should wait for possible inspection. The inspection processes concern activities related to tax declarations for customs, control of dangerous goods (e.g., related to health and environment), and import of forbidden goods (e.g., hard drugs and imitation goods). Suspicious containers may be scanned (using x-ray or nuclear scan equipment) and/or opened for physical inspection. After border control, containers may be moved to the inland terminal where they wait for further transport by inland transportation facilities. When the container (eventually) arrives at the recipient(s), it will be unloaded. Finally, it is transported to an empty container depot. This completes the import process.

According to the import process described, basic information about the goods in the container should be captured in order to facilitate smooth and fast transportation, such as product type, quantity, weight, box number, pallet number (on which boxes are stored), container number (in which the pallets are stuffed), source and destination address. In addition, during the transportation, dynamically exchanging information may be needed: inspectors for instance, might want to know the exact location of the vessel before it arrives at the port. Similarly, inland transporters want to know at which moment the container will leave the sea.
terminal. All this kind of information should be available. This means the international supply chain needs high information visibility, which is considered to be a top priority (Aberdeen, 2006). However, previous research (Hesketh, 2009, Pruksasri et al., 2011) has shown that current supply chain information systems still suffer from several information quality shortcomings related to, for example, incorrectness and inconsistency of data, as well as unclear data accountability. The above-mentioned Data Pipeline (Hesketh, 2010) has been proposed to fix all those data security related problems.

2.2 Data Pipeline Concept

The Seamless Integrated Data Pipeline or, simply, the Data Pipeline is a concept of a virtual seamless, integrated, web-based data pipeline, which is designed to provide relevant data to authorized stakeholders in the supply chain. Data flows can take place from any source to any destination. A visualization of it is provided figure 1.

According to the concept, data will become available to affiliated parties in the pipeline the moment a product has been ordered. In case it still needs to be produced, data concerning its production status may be added gradually. Next detailed data around the start of the transportation may become available. During transportation, data relating to the carriage, such as GPS position and change of transportation modality, are also made visible by linking them to the pipeline. Crucial in the concept is the idea that data are made available at their source, i.e., from the original information systems in which they are entered. It enables authorized parties to always access the original (most correct) information at the source and minimizes the need to make (possibly incorrect) copies of the data.

A consequence of the pipeline concept described so far is that all global partners would be linked up to one single pipeline, which makes it a hard-to-manage global infrastructure. In order to facilitate proper management, a type of structuring mechanism is needed. Looking in more detail at the sub-processes related to the transportation of goods, we observe that, in practice, stakeholders are cooperating together in all kinds of sub-groups, in all different phases of the transportation like export, transfer, import, etc. Sub-groups of cooperating stakeholders inside the supply chain are denoted as “Communities” (Hofman, 2011). It should be clear that a supply chain actor is usually member of more than one community. To illuminate the idea, a few example communities are shown in figure 2. To specify the characteristics of each community in more detail, all kinds of details should be specified around actors, their relationships, role, etc. This is shown in figure 3.
Figure 3 shows an example ‘actor, profile and information exchange model’ in which a company named ‘Philips’ has a different ‘business activity profile’ in two different communities: in the import community, Philips acts as an importer in which it cooperates with actors such as Dutch customs, DHL (logistics provider), and a Chinese company (supplier). Meanwhile, Philips also acts as an exporter in the export community in order to export its products to other countries. Within each community, the links between two actors represent communication lines. The profiles next to each actor outside each community describe certain general facts about the actor, while, inside a community, profiles give information about the role of the actor inside that community, including data about its expected behavior. Note that in this set-up, the same actors may have different community profiles in different countries. Data security requirements can also be added to these models, e.g., in the specification of actor profiles as well as in additional boxes next to the communication links drawn between community actors (not shown in figure 3).

3. INFORMATION SECURITY REQUIREMENTS

Each community in the international supply chain consists of several actors. They are often spread all over the world. We also observe that the composition of communities can be highly dynamic due to, for example, changing transportation routes or changing import and export procedures. Therefore, the first challenge that pops up after some thinking about realizing secure information exchange inside a community is solving the problem of secure identification of all global community actors.

An important constraint when thinking about the design of a global secure identification solution is that almost every country already introduced its own identification & authentication system for supply chain actors (Eertink et al., 2008). In The Netherlands for example, a supply chain actor has to register to the local identification system known as “e-Herkenning” (e-Herkenning, 2011). This system provides several registration means such as online registration, possibly followed by the physical presentation of documents, which are applied to verify the authenticity of the organization. The number and accuracy of the various identification and authentication procedures executed depend on the security level required: high-quality procedures usually relate to high trust levels around issued IDs, vice versa. In other countries, similar but different local approaches exist to issue IDs. So, looking at the global supply chain, the second challenge identified is about harmonization of trust levels between international community members based on the local ID-systems already in use (the latter avoids the need of setting up a fully new global system).

Different communication paths between supply chain actors may need different protection levels of information. Some data may be accessible for anyone without any restriction, but a lot of data are usually restricted to certain authorized parties only. In addition, the protection levels of information are often different in each country. This indicates a third data security requirement, namely that of the harmonization of data security levels between internationally operating community members, with (again) the constraint that local arrangements do not have to change (at high costs).
4. TOWARDS SOLUTIONS: THREE SECURITY PROTOCOLS

Starting to think about solutions for meeting the security requirements mentioned in section 3, within the context described in section 2, we soon realized the need for simplicity as well as for the re-use of existing systems. So we look for simple extensions of existing systems that together implement the secure global Data Pipeline. In the example elaborated below, we assume that two actors from different countries are member of the same international community to be set up, namely, the actor ‘Customs’ in the Netherlands and the actor ‘Logistics Provider’ in England. We further assume that the Dutch customs wish to retrieve certain data from the English logistic provider to execute a certain supervision task. In line with the assumptions made in the previous section, we presume that, in both countries, a secure identification system is in place having a given security level, and that the mentioned actors are authorized users of their local e-ID system. In other words, we assume that, in each country, a chain of trust is in place for secure data exchange (Stallings, 2003). Local e-IDs can be implemented in several ways, e.g., by means of certificates signed by a Certifying Authority (CA). Furthermore, we take as a starting point that, in principle, all data communication between community members in the Data Pipeline should be secured.

Below, we introduce three protocols that meet the requirements that were put forward in the previous section. An important idea underlying these protocols is that chain of trust needed for international data exchange between community members is based on the principle that the global chain of trust is set up using the existing local chains of trust, which are linked together by chains of trust as developed in the virtual Data Pipeline. To do so, we propose to introduce, for each country, a local gateway system. A local gateway system is locally accessible for all actors that use the above-mentioned secure local e-ID system. Based on the current control tasks given to customs, it is quite natural to make customs responsible for the management of the local gateways. Local gateways play a crucial role in realizing a secure virtual global Data Pipeline.

4.1 Registration Protocol

Before supply chain members can use the Data Pipeline, they should first register. Registration enables the discovery and authentication of other registered fellow members at any later moment in time. We propose the set-up of a Registration System (RS) (in practice, for reasons of robust availability, a set of RSs) that can register community members based on requests from their side. When registering, stakeholders receive a Data Pipeline ID (d-ID). Their profile (defined in section 2), which includes their d-ID and their role in the community, is recorded at the RS during registration. The local gateway of each country plays a key supportive role during registration. To execute its role, the local gateway needs to get a secure channel with (a) the local stakeholder, (b) the local e-ID system, and (c) the global registration system: see figure 4.

![Figure 4. Protocol for secure registration of community members at a registration system.](image-url)
The registration protocol works roughly as follows (we only show its basic working and we assume no errors occur): (1) a local stakeholder (e.g., customs in the NLs, or the logistic provider in England) sends a request-to-register to the local gateway system in use, (2) the local gateway system asks the local e-ID system to verify local stakeholders e-ID, (3) the local e-ID system verifies e-ID, and returns the answer to the local gateway system, (4) if correct, the local gateway forwards the request-to-register message to the registration system RS, (5) if everything is OK, the RS returns the d-ID to the local gateway system that, in its turn, (6) forwards it to the requestor. Figure 4 provides a visualization of this protocol: note that the new elements introduced by the registration protocol are (i) the local gateway systems (in each country), (ii) the RS, and (iii), for each local gateway, the three secure channels mentioned above. Note further that, for the registration protocol, the required chain of trust in the Data Pipeline (that links to two local chains of trust together, namely that in the Netherlands and that in England) can be based on trust relations between the organizations that manage the local gateway systems and the registration system.

### 4.2 Discovery Protocol

When registration of community members has been finalized, it should be possible, for each of them, to start up data communication with any other. For example, the Dutch customs may wish to request some data from the logistic provider in England. However, due to certain changes or other reasons, customs may not yet know the global d-ID of the logistic provider. In addition, it might not know its digital certificate (DC), which would also be very helpful in order to set up a direct and secure channel with the logistic provider. (Here, we assume that the DCs signed by a local certifying authority (CA) can always be used globally, i.e., between international community partners.) The discovery protocol is meant to solve these two problems. Its working is rather straightforward: (1) customs sends a request for partner info to the local gateway system, (2) assuming it already verified customs e-ID, the local gateway forwards the request to the RS, (3) in order to receive the most updated DC, the RS forwards the request to the local gateway in England that next (4) sends a request for the DC to the local e-ID system. In the second part of the protocol, messages are sent in reverse order, where the RS adds the d-ID to the message. In this way, customs receives the required information when the message is finally forwarded by the local gateway in step (8). Note that the required chain of trust in the Data Pipeline is the same as in the registration protocol discussed above.

**Figure 5. Protocol for discovering the d-ID and DC of an international community partner.**

### 4.3 Data Exchange Protocol

Knowing the name (d-ID) and digital certificate (DC) of its community partner (here, the logistic provider), customs is able to set up a secure channel with that partner in order to request the required data. However,
before exchanging the relevant data, the security level of the secure channel should be checked since the Dutch customs is supposed to apply a certain internationally agreed data security level (such as, for example, Stork QAA levels (Eertink et al., 2008)). The protocol as visualized in figure 6 can be used to (a) check whether its community partner applies the required security level, and (b) request and receive the data.

Figure 6. Protocol for data exchange at a required security level between two community members.

The protocol shown is similar to the ones described above. The most essential elements are as follows: in steps (1) and (2) a secure direct channel is being set up between customs and the logistic provider. In step 3, the required security level is mapped onto the local level (as used in the Netherlands by the local e-ID system), in step (4), (5) and (6) it is checked by the local e-ID system has implemented that security level for Dutch Customs. If true, a request is send (7) to the local gateway in England where it is checked whether the logistic provider applies the required security level (7.1. and 7.2). If so, the logistic provider is sent a message (8) that it is allowed to send the data of the original request to customs (9). Note that the required chain of trust in the Data Pipeline is rather similar to ones needed in the first and second protocol and is here based on the trust relationship between the managers of the local gateway systems.

5. CONCLUSIONS

Based on the description of the proposed protocols, it should be clear that the data security requirements as were put forward in section 3 (based on the conceptualization in section 2) have been met, at least conceptually. The first two (registration and discovery) protocols enable secure identification of community members and, based upon that, the set-up of secure data exchange by means of protocol 3. This third protocol also takes care of the solving the problem of harmonization of data security levels. Finally we observe that in each protocol the described chains of trust implement the solution for harmonization of trust levels. The solutions provided have been designed having the important condition in mind that existing local systems for secure data exchange can be used without change.

We also paid much attention to the global chains of trust needed to make all data exchange secure at global level. From a technical perspective it is clear that the local gateways play a crucial role here since such a gateway is an essential communication unit for both the local chain of trust (e.g., that of country) and for the chains of trust implemented globally (in the international Data Pipeline) that link local chains trust together. As a consequence, the organizations that manage these local gateways should trust each other. In practice, the customs of each country may be chosen as the organizations that should manage the gateways.

The (conceptual) solutions proposed are just a first step and should be tested in practice to show proof-of-concept. During these tests all kinds of details should be elaborated, e.g., related to encryption algorithms to be used and to all kinds of organizational issues, including the organization of supervision tasks. In addition,
more specific data exchange protocols should be elaborated, implemented and tested, based on real business processes that take place in the international supply chain. These activities are planned to be executed in future research.

ACKNOWLEDGEMENT

This research has been supported by the project CASSANDRA (Common assessment and analysis of risk in global supply chains) - This project is co-funded by the 7th Framework Program of the EC.

REFERENCES

Aberdeen 2006. Global Supply Chain Benchmark report. IBM.
Schaefer, S., 2006, Secure trade lane: a sensor network solution for more predictable and more secure container shipments. Companion to the 21st ACM SIGPLAN symposium on Object-oriented programming systems, languages, and applications. Portland, Oregon, USA, 839-845